PAUL R. HALMOS

ESPACIOS VECTORIALES FINITO-DIMENSIONALES

(FINITE-DIMENSIONAL VECTOR SPACES)

second edition

ESPACIOS VECTORIALES FINITO - DIMENSIONALES

por
PAUL R. HALMOS

Se presenta el álgebra lineal en un espíritu axiomático puro. En cada uno de las etapas el texto sólo supone el mínimo necesario para obtener el resultado deseado, ilustrado con ejemplos sobre el porqué necesariamente fallaría un número menor de suposiciones.

En esta nueva edición el profesor Halmos ha agregado literalmente cientos de problemas que se alejan de la rutina y hacen reflexionar al lector y que cubren todos los aspectos de la teoría.

Muchos de los problemas llevan el obejtivo de probar el ingenio del lector, pidiéndole que descubra si la aseveración es verdadera o falsa, demostrar que es verdadera, formular un ejemplo contrario si es falsa y estudiar las alteraciones de hipótesis y las conclusiones que haga que las verdaderas sean falsas y las falsas verdaderas.

Se presentan ejercicios tan pronto como el enunciado tiene sentido, antes de desarrollar el mecanismo para una solución rápida. El estudiante que intenta, aun sin éxito, solucionar tales ejercicios equívocos, está capacitado para resolver mejor los desarrollos subsecuentes para su intento.

(Pasa a la solapa posterior)

Vea la parte interior de este forro, encontrará una lista clasificada de algunas obras de esta Editorial

ESPACIOS VECTORIALES FINITO—DIMENSIONALES

ESPACIOS VECTORIALES FINITO–DIMENSIONALES

Por

PAUL R. HALMOS

Profesor de Matemáticas
Universidad de Michigan

COMPAÑIA EDITORIAL CONTINENTAL, S. A.
CALZADA DE TLALPAN NÚM. 4620. MÉXICO 22, D. F.

SUCURSALES, DEPOSITOS Y REPRESENTACIONES EN:
ARGENTINA — Bolivia — Brasil — Colombia — Costa Rica — CHILE
Dominicana — Ecuador — Estados Unidos — ESPAÑA — Guatemala
Honduras — Nicaragua — Panamá — Paraguay — PERU — Portugal
Puerto Rico — El Salvador — Uruguay — VENEZUELA

Título original en inglés:

FINITE—DIMENSIONAL VECTOR SPACES

Traducido por:

PROF. GABRIEL AGUIRRE CARRASCO
Catedrático de la Universidad Autónoma de Puebla

Edición autorizada por:

D. VAN NOSTRAND COMPANY INC. – PRINCETON, NEW JERSEY

Library of Congress Catalog Card No. 58-8446

Primera edición en español:
octubre de 1965

MIEMBRO DE LA CAMARA NACIONAL
DE LA INDUSTRIA EDITORIAL
Registro Núm. 43

PREFACIO

Mi propósito en este libro es tratar las transformaciones lineales sobre espacios vectoriales finito-dimensionales por los métodos de teorías más generales. La idea es hacer hincapié en las nociones geométricas simples comunes a muchas partes de las matemáticas y sus aplicaciones, y hacerlo en un lenguaje que ponga de manifiesto los secretos de la materia y diga al estudiante qué es lo que está en el fondo de la mente de gente que prueba teoremas sobre ecuaciones integrales y espacios de Hilbert. Sin embargo, el lector no tiene que participar de mi prejuiciada motivación. Con excepción de alguna referencia ocasional a matemática de subgraduados, el libro es autárquico y puede ser leído por cualquiera que trate de lograr cierta percepción de los problemas lineales usualmente discutidos en curso sobre teoría de matrices o álgebra "más elevada". Los métodos algebraicos, libres de coordenadas, no pierden poder y elegancia por especialización a un número finito de dimensiones, y son, en mi opinión, tan elementales, como el tratamiento clásico de coordenadas.

Originalmente tuve el propósito de que este libro contuviera un teorema si y sólo si existe una generalización infinito-dimensional del mismo. La tentadora facilidad de algunas nociones y resultados finito-dimensionales fue, sin embargo, irresistible y en el resultado final mis intenciones iniciales eran apenas visibles. Se ven muy claramente en el énfasis, puesto en todas partes en métodos generalizables, en vez de resultados lo más precisos posibles. El lector puede algunas veces ver algún método obvio de acortar las pruebas que doy. En esos casos las posibilidades son de que el análogo infinito-dimensional de la prueba más corta sea mucho más larga o también no existente.

Durante varios años ha estado circulando una edición preliminar del libro (Anales de Estudios Matemáticos, No. 7, publicado por primera vez por la Princeton University Press en 1942). Además de algunos cambios menores en estilo y orden, la diferencia entre la versión precedente y ésta es la que la segunda contiene el siguiente material nuevo: (1) Una breve discusión de campos y, en el tratamiento de espacios vectoriales con productos interiores, especial atención al caso real. (2) Una definición de determinantes en términos de invariantes, vía la teoría de formas multilineales. (3) Ejercicios.

Los ejercicios (más de trescientos) constituyen una importante adición; espero que resultarán útiles tanto para el estudiante, como para el maestro. Hay dos cosas sobre los mismos que el lector debería conocer. Primero, si un ejercicio no es ni imperativo ("pruébese que...") ni interrogativo ("¿es verdad que...?"), sino meramente declarativo, entonces se ofrece como un tema a discusión. Para esos ejercicios se pide al lector que descubra si la aserción es verdadera o falsa, que pruebe si es verdadera y construya un contraejemplo si es falsa y, lo que es más importante de todo, discuta las alteraciones de hipótesis y conclusión que hagan falsas las verdaderas, y verdaderas las falsas. Segundo, los ejercicios, cualquiera que sea su forma gramatical, no se colocan siempre de modo que su misma posición constituya una pista para su solución. Frecuentemente se enuncian los ejercicios tan pronto como

Ninguno de los teoremas y sólo muy pocos de los ejercicios desarrollado la maquinaria para una rápida solución. Un lector que trate (aun sin buen éxito) de resolver esos ejercicios "mal colocado", es probable que aprecie y comprenda mucho mejor los desarrollos subsecuentes para su intento. Teniendo en cuenta posibles ediciones futuras de este libro, pido al lector me dé a conocer los errores de los ejercicios y me sugiera mejoras y adiciones. (Es innecesario decir que lo mismo es bueno respecto al texto.)

Ninguno de los teoremas y sólo muy pocos de los ejercicios han sido descubrimientos míos; la mayor parte de ellos son conocidos de los matemáticos en ejercicio y han sido conocidos durante largo tiempo. Aunque no doy una detallada lista de mis fuentes, estoy, sin embargo, perfectamente enterado de mi deuda con los libros y papeles en que aprendí y con los amigos y conocidos que, antes de la publicación de la primera versión, me dieron mucha crítica y aliento valiosos. Estoy particularmente agradecido a tres personas: J. L. Dobb y Arlen Brown, que leyeron todo el manuscrito de la primera y segunda versión, respectivamente, e hicieron muchas sugestiones útiles, y a John von Neumann, quien fue uno de los creadores del método y espíritu modernos que he tratado de presentar, y cuyas enseñanzas fueron una inspiración para este libro.

P. R. H.

C O N T E N I D O

CAPITULO 1

ESPACIOS

§1. CAMPOS

En lo que sigue tendremos ocasión de usar varias clases de números (tales como la clase de todos los números reales o la clase de todos los números complejos). En virtud de que no sería posible, en esta etapa inicial, confinarnos dentro de ninguna clase en particular, adoptaremos el expediente de referirnos a los números como *escalares*. El lector no perderá nada esencial si consistentemente interpreta los escalares como números reales o como números complejos; en los ejemplos que habrán de estudiarse aparecerán ambas clases. Para ser más concretos (y también para operar en el nivel adecuado de generalidad) procederemos a listar todos los hechos generales que será necesario suponer con respecto de los escalares.

(A) A cada par, α y β, de escalares corresponde un escalar $\alpha + \beta$, llamado la suma de α y β, de manera que

(1) la adición es conmutativa, $\alpha + \beta = \beta + \alpha$,

(2) la adición es asociativa, $\alpha + (\beta + \gamma) = (\alpha + \beta) + \gamma$,

(3) existe un escalar único, 0 (llamado cero), tal que $\alpha + 0 = \alpha$ para todo escalar α, y

(4) a cada escalar no-cero α corresponde un escalar único $-\alpha$ que $\alpha + (-\alpha) = 0$.

(B) A cada par, α y β, de escalares corresponde un escalar $\alpha\beta$, llamado el *producto* de α y β, en tal forma que

(1) la multiplicación es conmutativa, $\alpha\beta = \beta\alpha$,

(2) la multiplicación es asociativa, $\alpha(\beta\gamma) = (\alpha\beta)\gamma$,

(3) existe un escalar único, distinto de cero, o no cero, 1 (llamado *uno*), tal que $\alpha 1 = \alpha$ para cada escalar α, y

(4) a cada escalar no cero α corresponde un escalar único α^{-1} $\left(\text{o } \frac{1}{\alpha}\right)$ tal que $\alpha\alpha^{-1} = 1$.

(C) La multiplicación es distributiva con respecto de la adición, $\alpha(\beta + \gamma) = \alpha\beta + \alpha\gamma$.

Si la adición y la multiplicación se definen dentro de algún conjunto de objetos (escalares), de manera que se satisfagan las condiciones (**A**), (**B**) y (**C**), entonces ese conjunto (juntamente con las operaciones dadas) se llama un campo. Así, por ejemplo, el conjunto $\mathcal{Q}$ de todos los números racionales (con las definiciones ordinarias de suma y producto) es un campo, y lo mismo es cierto respecto del conjunto $\mathcal{R}$ de todos los números ~~complejos~~ reales. Y el conjunto $\mathcal{C}$ de todos los números complejos.

EJERCICIOS

1. Casi todas las leyes de la aritmética elemental son consecuencia de los axiomas que definen un campo. Pruébese, en particular, que si $\mathcal{F}$, es un campo, y si α, β y γ pertenecen a $\mathcal{F}$, entonces se dan las siguientes relaciones:

(a) $0 + \alpha = \alpha$.

(b) Si $\alpha + \beta = \alpha + \gamma$, entonces $\beta = \gamma$.

(c) $\alpha + (\beta - \alpha) = \beta$. (Aquí $\beta - \alpha = \beta + (-\alpha)$.)

(d) $\alpha\cdot 0 = 0\cdot\alpha = 0$. (Para claridad o énfasis, usamos algunas veces el punto para indicar multiplicación.)

(e) $(-1)\alpha = -\alpha$.

(f) $(-\alpha)(-\beta) = \alpha\beta$.

(g) Si $\alpha\beta = 0$, entonces, o $\alpha = 0$, o $\beta = 0$ (o ambos).

2. (a) ¿Es un campo el conjunto de todos los números positivos? (En sistemas familiares, como el de los enteros, usaremos casi siempre las operaciones ordinarias de adición y multiplicación. En las raras ocasiones en que nos apartamos de esta convención, hacemos una amplia advertencia. En cuanto a "positivo", indicamos con esta palabra, aquí y en todo este libro, "mayor que o igual a cero". Si ha de excluirse a cero, diremos "estrictamente positivo".)

(b) ¿Qué sobre el conjunto de todos los enteros?

(c) ¿Pueden cambiarse las respuestas a estas preguntas volviendo a definir la adición o la multiplicación, o ambas?

3. Sea m un entero, $m \geqq 2$, y sea $\mathcal{Z}_m$ el conjunto de todos los enteros positivos menores que m, $\mathcal{Z}_m = \{0, 1, \cdots, m-1\}$. Si α y β están en $\mathcal{Z}_m$, sea $\alpha + \beta$ el menor resto positivo obtenido m dividiendo la suma (ordinaria) de α y β por m, y de modo semejante, sea $\alpha\beta$ el menor resto positivo obtenido dividiendo el producto (ordinario) de α y β por m. (Ejemplo: si $m = 12$, entonces $3 + 11 = 2$ y $3.11 = 9$).

(a) Pruébese que $\mathcal{Z}_m$ es un campo si y sólo si m es prima.

(b) ¿Qué es -1 en $\mathcal{Z}_5$?

(c) ¿Qué es $\frac{1}{3}$ en $\mathcal{Z}_7$?

4. El ejemplo de $\mathbb{Z}_p$ (donde p es primo) demuestra que no todas las leyes elementales de la aritmética son válidas en los campos; en $\mathbb{Z}_2$, por ejemplo, $1 + 1 = 0$. Demuéstrese que si $\mathfrak{F}$ es un campo, entonces el resultado de sumar 1 repetidamente a sí mismo es siempre diferente de 0, o la primera vez que es igual a 0 ocurre cuando el número de sumandos es primo. (La *característica* del campo $\mathfrak{F}$ se define como 0 en el primer caso y el primo crucial en el segundo.)

5. Sea $Q(\sqrt{2})$ el conjunto de todos los números reales de la forma $\alpha + \beta\sqrt{2}$, donde α y β son racionales.

(a) ¿Es $Q(\sqrt{2})$ un campo?

(b) ¿Y qué si se requiere que α y β sean enteros?

6. (a) ¿No forman un campo el conjunto de todos los polinomios con coeficientes enteros?

(b) ¿Qué si se permite que los coeficientes sean números reales?

7. Sea $\mathfrak{F}$ el conjunto de todos los pares (ordenados) (α, β) de números reales,

(a) Si la adición y la multiplicación se definen como

$$(\alpha, \beta) + (\gamma, \delta) = (\alpha + \gamma, \beta + \delta)$$

y

$$(\alpha, \beta)(\gamma, \delta) = (\alpha\gamma, \beta\delta),$$

¿se convierte $\mathfrak{F}$ en un campo?

(b) Si la adición y la multiplicación se definen como

$$(\alpha, \beta) + (\gamma, \delta) = (\alpha + \gamma, \beta + \delta)$$

y

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

¿ $\mathfrak{F}$ es, entonces, un campo?

(c) ¿Qué sucede (en los dos casos anteriores), consideramos pares ordenados de números complejos?

§ 2. ESPACIOS VECTORIALES

Llegamos ahora al concepto básico de este libro. Para la definición que sigue suponemos que se nos da un campo particular $\mathfrak{F}$; los escalares que han de usarse habrán de ser elementos de $\mathfrak{F}$.

DEFINICIÓN. Un espacio vectorial es un conjunto $\mathfrak{V}$ de elementos que satisfacen los siguientes axiomas.

(A) A cada par, x y y, de vectores en $\mathfrak{V}$ corresponde un vector $x + y$, llamado la suma de x y y, en tal forma que

(1) la adición es conmutativa, $x + y = y + x$,

(2) la adición es asociativa, $x + (y + x) = (x + y) + z$,

(3) Existe en $\mathcal{V}$ un vector único 0 (llamado el *origen*), tal que $x + 0 = x$, para cada vector x, y

(4) A cada vector x en $\mathcal{V}$ corresponde un vector único $-x$, tal que $x + (-x) = 0$.

(B) A cada par, α y x, donde α es un escalar y x es un vector en $\mathcal{V}$, corresponde un vector αx, en $\mathcal{V}$, llamado el *producto* de α y x, en tal forma que

(1) La multiplicación por escalares es asociativa, $\alpha(\beta x) = (\alpha\beta)x$, y

(2) $1x = x$ para todo vector x.

(C) (1) La mult plicación por escalares es distributiva con respecto a la adición de vectores, $(\alpha + \beta)x = \alpha x + \beta x$.

(2) La multiplicación por vectores es distributiva con respecto a la adición escalar $(\alpha + \beta)x = \alpha x + \beta x$.

No se pretende que estos axiomas son lógicamente independientes; son meramente una conveniente caracterización de los objetos que deseamos estudiar. La relación entre un espacio vectorial $\mathcal{V}$ y el campo subyacente $\mathcal{F}$ es usualmente descrita diciendo que $\mathcal{V}$ es un espacio vectorial *sobre* $\mathcal{F}$. Si $\mathcal{F}$ es el campo $\mathcal{R}$ de números reales, $\mathcal{V}$ se llama un *espacio vectorial real*; de modo semejante, si $\mathcal{F}$ es $\mathcal{Q}$ o si $\mathcal{F}$ es $\mathcal{C}$, hablamos de *espacios vectoriales racionales* o de *espacios vectoriales complejos*.

§ 3. EJEMPLOS

Antes de discutir las implicaciones de los axiomas, damos algunos ejemplos. Nos referiremos a estos ejemplos una y otra vez, y usaremos la notación aquí establecida en todo el resto de nuestro trabajo.

(1) Sea $\mathcal{C}^1(= \mathcal{C})$ el conjunto de todos los números complejos; si interpretamos $x + y$ y αx como adición numérica compleja común y multiplicación, respectivamente, $\mathcal{C}^1$ se convierte en un espacio vectorial complejo.

(2) Sea $\mathcal{P}$ el conjunto de todos los polinomios, con coeficientes complejos, en una variable t. Para convertir a $\mathcal{P}$ en un espacio vectorial complejo, interpretamos la adición vectorial y la multiplicación escalar como la adición ordinaria de dos polinomios y como la multiplicación de un polinomio por un número complejo, respectivamente; el origen de $\mathcal{P}$ es el polinomio idéntico a cero.

El ejemplo (1) es demasiado sencillo y el ejemplo (2) es de-

(como veremos después, es suficientemente general para todos nuestros propósitos.

(3) Sea $\mathcal{C}^n, n = 1, 2, \cdots$, el conjunto de todos los n tuplos de números complejos. Si $x = (\xi_1, \cdots, \xi_n)$ y $y = (\eta, \cdots, \eta_n)$ son elementos de $\mathcal{C}^n$, escribimos, por definición,

$$x + y = (\xi_1 + \eta_1, \cdots, \xi_n + \eta_n),$$

$$\alpha x = (\alpha\xi_1, \cdots, \alpha\xi_n),$$

$$0 = (0, \cdots, 0),$$

$$-x = (-\xi_1, \cdots, -\xi_n).$$

Es fácil verificar que todas las partes de nuestros axiomas **(A)**, **(B)** y **(C)**, § 2, se satisfacen, de manera que $\mathcal{C}^n$ es un espacio vectorial complejo; será llamado *espacio complejo coordenado n-dimensional.*

(4) Para cada entero positivo n, sea $\mathcal{P}_n$ el conjunto de todos los polinomios (con coeficientes complejos, como en el ejemplo (2)) de grado $\leqq n - 1$, juntamente con el polinomio idéntico a cero. (En la discusión usual de grado, el grado de este polinomio no está definido, de manera que no podemos decir que tenga un grado $\leqq n - 1$.) Con la misma interpretación de las operaciones lineales (adición y multiplicación escalar) que en (2), $\mathcal{P}_n$ es un espacio vectorial complejo.

(5) Un pariente cercano de $\mathcal{C}^n$ es el conjunto $\mathcal{R}^n$ de todos los n-tuplos de números reales. Con las mismas definiciones formales de adición y multiplicación escalar que para $\mathcal{C}^n$, excepto que ahora consideramos solamente escalares reales α, el espacio $\mathcal{R}^n$ es un espacio vectorial real; será llamado *espacio n-dimensional real coordenado.*

(6) Todos los ejemplos precedentes pueden generalizarse. Así, por ejemplo, puede describirse una obvia generalización de (1) diciendo que todo campo puede ser considerado como un espacio vectorial sobre sí mismo. Una generalización común de (3) y de (5) comienza con un campo arbitrario $\mathcal{F}$ y forma el conjunto $\mathcal{F}^n$ de n-tuples de elementos de $\mathcal{F}$; las definiciones formales de las operaciones lineales son las mismas que para el caso $\mathcal{F} = \mathcal{C}$.

(7) Un campo, por definición, tiene cuando menos dos elementos; un espacio vectorial, en cambio puede tener solamente uno. Puesto que todo espacio vectorial contiene un origen, hay esen-

cialmente (esto es, por notación), solamente un espacio vectorial que tenga solamente un vector. Este espacio vectorial sumamente trivial será denotado por $\mathcal{O}$.

(8) Si, en el conjunto $\mathcal{R}$ de todos los números reales, se define la adición como de costumbre y la multiplicación de un número real por un número racional se define como de costumbre, entonces $\mathcal{R}$ se convierte en un espacio vectorial racional.

(9) Si, en el conjunto $\mathcal{C}$ de todos los números complejos, la adición se define como de costumbre y la multiplicación de un número complejo por un número real se define como de costumbre, entonces $\mathcal{C}$ se convierte en un espacio vectorial real. (Compárese este ejemplo con (1); son enteramente diferentes.)

§ 4. COMENTARIOS

Proceden algunos comentarios sobre nuestros axiomas y notación. Hay sorprendentes similaridades (e igualmente sorprendentes diferencias) entre los axiomas de un campo y los axiomas de un espacio vectorial sobre un campo. En ambos casos, los axiomas (**A**) describen la estructura aditiva del sistema, los axiomas (**B**) describen su estructura multiplicativa, y los axiomas (**C**) describen la conexión entre las dos estructuras. Los familiarizados con la terminología algebraica habrán reconocido los axiomas (**A**) (tanto en el § 1 como en el § 2) como las condiciones que definen un grupo abeliano (conmutativo); los axiomas (**B**) y (**C**) (en el § 2), expresan el hecho de que el grupo admite escalares y operadores. Mencionamos de paso que si los escalares son elementos de un anillo (en vez de un campo), el concepto generalizado correspondiente a un espacio vectorial se llama *módulo*.

Espacios especiales vectoriales reales (como $\mathcal{R}^2$ y $\mathcal{R}^3$) son familiares en geometría. Parece no haber excusa en esta etapa por nuestra insistencia, sin aparente interés, en campos distintos de $\mathcal{R}$, y, en particular, en el campo $\mathcal{C}$ de números complejos. Esperamos que el lector esté dispuesto a confiar en que tendremos que hacer uso posteriormente de profundas propiedades de los números complejos. (conjugación, oclusión algebraica) y que los números complejos desempeñan un papel importante tanto en las aplicaciones de los espacios vectoriales a la física moderna (mecánica cuántica), como a la generalización matemática de nuestros resultados en el espacio de Hilbert. Su única gran desventaja es la dificultad de representación gráfica; la representación gráfica ordinaria (diagrama de Argand) de $\mathcal{C}^1$ no se distingue de la de $\mathcal{R}^2$, y una representación gráfica de $\mathcal{C}^2$ parece estar fuera del alcance humano. En las ocasiones en que tengamos que usar lenguaje gráfico, usaremos en consecuencia la terminología de $\mathcal{R}^n$ en $\mathcal{C}^n$, y hablar de $\mathcal{C}^2$, por ejemplo, como de un plano.

Finalmente, hacemos un comentario sobre la notación. Observamos que el símbolo 0 ha sido usado con dos significados: una vez como escalar y una vez como vector. Para hacer peor la situación, posteriormente, cuando introduzcamos las funcionales lineales y las transformaciones lineales le daremos otros significados más. Afortunadamente las relaciones entre las varias interpretaciones de 0 son tales que, después de esta palabra de advertencia, no debe suscitarse confusión por esta práctica.

EJERCICIOS

1. Pruébese que si x y y son vectores y si α es un escalar, entonces son válidas las siguientes relaciones.

(a) $0 + x = x$.
(b) $-0 = 0$.
(c) $\alpha \cdot 0 = 0$.
(d) $0 \cdot x = 0$. (Obsérvese que el mismo símbolo se usa para ambos lados de esta ecuación; a la izquierda denota un escalar, a la derecha un vector).
(e) Si $\alpha x = 0$, entonces o $\alpha = 0$, $x = 0$ (o ambos).
(f) $-x = (-1)x$.
(g) $y + (x - y) = x$. (Aquí $x - y = x + (-y)$.)

2. Si p es una prima, entonces $\mathcal{Z}_p{}^n$ es un espacio vectorial sobre $\mathcal{Z}_p$ (cf. § 1, Ex. 3); ¿cuántos vectores hay en este espacio vectorial?

3. Sea $\mathcal{V}$ el conjunto de todos los pares (ordenados) de números reales. Si $x = (\xi_1, \xi_2)$ y $y = (\eta_1, \eta_2)$ son elementos de $\mathcal{V}$, escriba

$$x + y = (\xi_1 + \eta_1, \xi_2 + \eta_2)$$
$$\alpha x = (\alpha\xi_1, 0)$$
$$0 = (0, 0)$$
$$-x = (-\xi_1, -\xi_2).$$

¿Es $\mathcal{V}$ un espacio vectorial con respecto a estas definiciones de las operaciones lineales? ¿Por qué?

4. Algunas veces un subconjunto de un espacio vectorial es él mismo un vector espacial (con respecto de las operaciones lineales ya dadas). Considérese, por ejemplo, el espacio vectorial $\mathcal{C}^3$ y los subconjuntos $\mathcal{V}$ de $\mathcal{C}^3$ consistentes en esos vectores (ξ_1, ξ_2, ξ_3), para los cuales

(a) ξ_1 es real.
(b) $\xi_1 = 0$.
(c) $\xi_1 = 0$ o $\xi_2 = 0$.

(d) $\xi_1 + \xi_2 = 0$.
(e) $\xi_1 + \xi_2 = 1$.

¿En cuál de estos casos es $\mathcal{V}$ un espacio vectorial?

5. Considérese el espacio vectorial $\mathcal{P}$ y los subconjuntos $\mathcal{V}$ de $\mathcal{P}$ que consisten en los vectores (polinomios) x, para los cuales

(a) x tiene grado 3.
(b) $2x(0) = x(1)$.
(c) $x(t) \geqq 0$ cada vez que $0 \leqq t \leqq 1$,
(d) $x(t) = x(1-t)$ para todas las t.

¿En cuál de estos casos es $\mathcal{V}$ un espacio vectorial?

§ 5. DEPENDENCIA LINEAL

Ahora que hemos descrito los espacios con que trabajaremos, debemos especificar las relaciones entre los elementos de esos espacios que serán de interés para nosotros.

Comenzaremos con unas cuantas palabras sobre la notación de la suma. Si correspondiendo a cada uno de un conjunto de índices i hay un vector dado x_i, y si no es necesario o no es conveniente especificar exactamente el conjunto de índices, simplemente hablaremos de un conjunto $\{x_i\}$ de vectores. (Admitimos la posibilidad de que el mismo vector corresponda a dos índices distintos Con toda honradez, en consecuencia, debe enunciarse que lo impor tante no es cuáles vectores aparezcan en $\{x_i\}$, sino cómo aparecen.) Si el conjunto de índices que se considera es finito, denotaremos la suma de los correspondientes vectores por $\sum_i x_i$ (o, cuando sea deseable, por un símbolo más explícito, como $\sum_{i=1}^n x_i$). A fin de evitar frecuentes y molestas distinciones de casos, es buena idea admitir dentro de la teoría general sumas como $\sum_i x_i$, aun cuando no hay índices i que hayan de sumarse o, más precisamente, aun cuando esté vacío el conjunto de índices que se considera. (En ese caso, por supuesto, no hay vectores que sumar o, más precisamente, el conjunto $\{x_i\}$ está también vacío). El valor de esa "suma vacía" se define, con toda naturalidad como el vector 0.

DEFINICIÓN. Un conjunto finito $\{x_i\}$ de vectores es *linealmente dependiente*, si existe un conjunto correspondiente $\{\alpha_i\}$ de escalares, no todos cero, tales que

$$\sum_i \alpha_i x_i = 0.$$

Si, por otra parte, $\sum_i \alpha_i x_i = 0$ implica que $\alpha_i = 0$ para cada i, el conjunto $\{x_i\}$ es *linealmente independiente*.

La redacción de esta definición está hecha en forma que cubra el caso del conjunto vacío; el resultado en ese caso, aunque posiblemente paradójico, se enlaza muy satisfactoriamente al resto de la teoría. El resultado es que el conjunto vacío de vectores es linealmente independiente. En realidad, si no hay índices i, entonces no es posible escoger algunos de ellos y asignar a los seleccionados un escalar distinto de cero, de manera que cierta suma desaparezca. La dificultad no está en evitar la asignación de cero; está en encontrar un índice al cual algo pueda ser asignado. Nótese que este argumento muestra que el conjunto vacío no es linealmente dependiente; para el lector no familiarizado con el "argumento de vacuidad", la equivalencia de la definición de independencia lineal con la negación directa de dependencia lineal, necesita un poco más de justificación intuitiva. El modo más fácil de sentirse cómodo con respecto de la aserción "$\sum_i \alpha_i x_i = 0$ implica que $\alpha_i = 0$ para cada i", en caso de que no haya índices i, es volverla a redactar en esta forma: "si $\sum_i \alpha_i x_i = 0$, entonces no hay índice i para el cual $\alpha_i \neq 0$. Esta versión es obviamente verdadera si no existe ningún índice i.

La dependencia y la independencia lineales son propiedades de conjuntos de vectores; se acostumbra, sin embargo, aplicar esos adjetivos a los vectores mismos, y así, algunas veces diremos "un conjunto de vectores linealmente independientes", en vez de "un conjunto linealmente independiente de vectores". Será conveniente también hablar de la dependencia e independencia lineales de un conjunto no necesariamente finito, $\mathfrak{X}$, de vectores. Diremos que $\mathfrak{X}$ es linealmente independiente si lo es cada subconjunto finito de $\mathfrak{X}$; de otro modo $\mathfrak{X}$ es linealmente dependiente.

Para lograr una mayor comprensión del signifcado de la dependencia lineal, estudiemos los ejemplos de espacios vectoriales que ya tenemos.

(1) Si x y y son dos vectores cualesquiera en $\mathcal{C}^1$, entonces x y y forman un conjunto linealmente dependiente. Si $x = y = 0$, esto es trivial; si no, entonces tenemos, por ejemplo, la relación $yx + (-x)y = 0$. Puesto que está claro que todo conjunto que contiene un subconjunto linealmente independiente, es él mismo linealmente dependiente, esto demuestra que en $\mathcal{C}^1$ cada conjunto que contiene más de un elemento es un conjunto linealmente dependiente.

(2) Más interesante es la situación en el espacio $\mathcal{P}$. Los vectores x, y y z, definidos por

$$x(t) = 1 - t,$$
$$y(t) = t(1 - t),$$
$$z(t) = 1 - t^2,$$

son, por ejemplo, linealmente dependientes, puesto que $x + y - z = 0$. Sin embargo, el conjunto infinito de vectores x_0, x_1 x_2, . . . , definido por

$$x_0(t) = 1, \quad x_1(t) = t, \quad x_2(t) = t^2, \cdots,$$

es un conjunto linealmente independiente, pues si tuviéramos cualquier relación de la forma

$$\alpha_0 x_0 + \alpha_1 x_1 + \cdots + \alpha_n x_n = 0,$$

tendríamos entonces una identidad polinómica

$$\alpha_0 + \alpha_1 t + \cdots + \alpha_n t^n = 0,$$

de donde

$$\alpha_0 = \alpha_1 = \cdots = \alpha_n = 0.$$

(3) Como mencionamos anteriormente, los espacios $\mathcal{C}^n$ son el prototipo de lo que deseamos estudiar; examinemos, por ejemplo el caso $n = 3$. A los que están familiarizados con la geometría de más alto número de dimensiones, la noción de dependencia lineal en este espacio (o, más propiamente hablando, en su análogo real $\mathcal{R}^3$) tiene un significado geométrico concreto, que solamente mencionaremos. En lenguaje geométrico, dos vectores son linealmente dependientes si y sólo si son coplanares con el origen. (Si uno piensa en un vector no como un punto en un espacio, sino como una flecha que apunta desde el origen a algún punto dado, el enunciado anterior podría modificarse tachando la expresión "con el origen" las dos veces que ocurre). Introduciremos inmediatamente la noción de múltiples lineales (o subespacios vectoriales) en un espacio vectorial y, en ese respecto, ocasionalmente usaremos el lenguaje sugerido por tales consideracionese geométricas.

§ 6. COMBINACIONES LINEALES

Diremos, cada vez que $x = \sum_i \alpha_i x_i$, que x es una *combinación lineal* de $\{x_i\}$; usaremos sin ninguna explicación ulterior todas las sencillas implicaciones gramaticales de esta terminología. Diremos, pues, en caso de que x sea una combinación lineal de $\{x_i\}$, que x es linealmente dependiente de $\{x_i\}$, dejaremos al lector la prueba de que si $\{x_i\}$ es liealmente independiente, entonces una condición suficiente y necesaria para que x sea una combinación lineal de $\{x_i\}$ es que el conjunto agrandado que se obtiene uniendo x a $\{x_i\}$, sea linealmente dependiente. Nótese que, de acuerdo con la definición de una suma vacía, el origen es una combinación lineal del conjunto vacío de vectores; es, además, el único vector con esta propiedad.

El siguiente teorema es el resultado fundamental concerniente a la dependencia lineal.

TEOREMA. *El conjunto de vectores no-cero $x_1, \ldots, x_n$ es linealmente dependiente si y sólo si x_k, $2 \leq k \leq n$, es una combinación lineal de los precedentes.*

PRUEBA. Supongamos que los vectores $x_1, \ldots, x_n$ son linealmente dependientes y sea k el primer entero entre 2 y n para el cual $x_1, \ldots, x_k$ son linealmente dependientes. (En el caso de lo peor, nuestro supuesto nos asegura de que bastará con hacer $k = n$). Entonces

$$\alpha_1 x_1 + \cdots + \alpha_k x_k = 0$$

para un adecuado conjunto de alfas (no todas cero); además, cualesquiera que sean las alfas, no podemos tener $\alpha_k = 0$, pues entonces tendríamos una relación de dependencia lineal entre $x_1, \ldots, x_{k-1}$. contraria a la definición de k. De aquí

$$x_k = \frac{-\alpha_1}{\alpha_k} x_1 + \cdots + \frac{-\alpha_{k-1}}{\alpha_k} x_{k-1},$$

como había de ser probado. Esto prueba la necesidad de nuestra condición; la suficiencia es clara, puesto que, como hemos observado anteriormente, cada conjunto que contiene un conjunto linealmente dependiente, lo es también.

§ 7. BASES

DEFINICIÓN. Una *base* (lineal) (o un *sistema de coordenadas*) en un espacio vectorial $\mathcal{V}$ es un conjunto $\mathcal{X}$ de vectores linealmente independientes tales que cada vector de $\mathcal{V}$ es una combinación lineal de elementos de $\mathcal{X}$. Un espacio vectorial $\mathcal{V}$ es *finito-dimensional* si tiene una base finita.

Excepto para la ocasional consideración de ejemplos restringiremos nuestra atención, en todo este libro, a espacios vectoriales finito-dimensionales.

Para ejemplos de bases, volvemos de nuevo a los espacios $\mathcal{P}$ y $\mathcal{C}^n$. En $\mathcal{P}$,el conjunto $\{x_n\}$, donde $x_n(t) = t^n$, $n = 0, 1, 2, \ldots$, es una base; cada polinomio es, por definición, una combinación lineal de un número finito de x_n. Además, $\mathcal{P}$ no tiene base finita, pues dado cualquier conjunto finito de polinomios podemos hallar un polinomio de grado más alto que cualquiera de ellos; este último polinomio, obviamente, no es una combinación lineal de los anteriores.

Un ejemplo de una base en $\mathcal{C}^n$ es el conjunto de vectores x_i, $i = 1, \ldots, n$, definido por la condición de que la coordenada de orden j de x_i sea δ_{ij}. (Aquí usamos por primera vez la popular δ de Kronecker; se define como $\delta_{ij} = 1$, si $i = j$ y $\delta_{ij} = 0$, si $i \neq j$). Afirmamos así que en $\mathcal{C}^3$ los vectores $x_1 = (1, 0, 0)$, $x_2 = (0, 1, 0)$ y $x_3 = (0, 0, 1)$ forman una base. es fácil ver que son linealmente in dependientes; la fórmula

$$x = (\xi_1, \xi_2, \xi_3) = \xi_1 x_1 + \xi_2 x_2 + \xi_3 x_3$$

prueba que cada x de $\mathcal{C}^3$ es una combinación lineal de los mismos.

En un espacio vectorial finito dimensional general $\mathcal{V}$, con base $\{x_1, \ldots, x_n\}$, sabemos que toda x puede escribrse en la forma

$$x = \sum_i \xi_i x_i;$$

afirmamos que las ξ están determinadas univócamente por x. La prueba de esta afirmación es un argumento usado a menudo en la teoría de la dependencia lineal. Si tuvieramos $x = \sum_i \eta_i x_i$, tendríamos, por sustracción,

$$\sum_i (\xi_i - \eta_i) x_i = 0.$$

Puesto que las x_i son linealmente independientes, esto implica que $\xi_i - \eta_i = 0$ para $i = 1, \ldots, n$; en otras palabras, las ξ son iguales que las η. (Obsérvese que escribir $\{x_1, \ldots, x_n\}$ para una base con n elementos no es lo que apropiadamente debe hacerse en caso de que $n = 0$. No obstante, usaremos frecuentemente esta notación. Cada vez que se hace eso es, en principio, necesario adjuntar una discusión separada, designada para cubrir el espacio vectorial $\mathcal{O}$. De hecho, sin embargo, todo lo relativo a ese espacio es tan trivial que los detalles no son dignos de asentarse y los omitiremos.)

TEOREMA. *Si $\mathcal{V}$ es un espacio vectorial finito dimensional y si $\{y_1, \ldots, y_m\}$ es un conjunto cualquiera de vectores linealmente independientes en $\mathcal{V}$, entonces, a menos de que las y formen ya una base,* *podemos encontrar vectores $y_{m+1}, \ldots, y_{m+p}$, de manera que la totalidad de las y, esto es $\{y_1, \ldots, y_m, y_{m+1}, \ldots, y_{m+p}\}$, sea una base. En otras palabras, todo conjunto linealmente independiente puede ampliarse de modo que constituya una base.*

PRUEBA. Puesto que $\mathcal{V}$ es finito-dimensional, tiene una base finita, digamos $\{x_1, \ldots, x_m\}$. Consideramos el conjunto $\mathcal{S}$ de vectores

$$y_1, \cdots, y_m, x_1, \cdots, x_n,$$

en este orden y aplicamos a este conjunto el teorema del §6 varias veces en sucesión. En primer lugar, el conjunto $\mathcal{S}$ es linealmente dependiente, puesto que las y son (como todos los vectores) combinaciones lineales de las x. De aquí que algún vector de $\mathcal{S}$ sea una combinación lineal de los precedentes; sea z el primero de esos vectores. Entonces z es diferente de cualquier y_i, $i = 1, \ldots, m$ (puesto que las y son linealmente independientes), de manera que z es igual a alguna x, digamos $z = x_i$. Consideramos el nuevo conjunto $\mathcal{S}'$ de vectores

$$y_1, \cdots, y_m, x_1, \cdots, x_{i-1}, x_{i+1}, \cdots, x_n.$$

Observamos que todo vector de $\mathcal{V}$ es una combinación lineal de vectores en $\mathcal{S}'$, puesto que por medio de $y_1, \ldots, y_m, x_1, \cdots, x_{i-1}$ podemos expresar x_i y entonces por medio de $x_1, \ldots, x_{i-1}, x_i, x_{i+1}, \ldots, x_n$ podemos expresar cualquier vector (las x forman una base). Si $\mathcal{S}'$ es linealmente independiente, ya está. Si no lo es, aplicamos el teorema § 6 una y otra vez en la misma forma, hasta que lleguemos a un conjunto linealmente independiente que contenga $y_1, \ldots, y_m$, en términos del cual podemos expresar cualquier vector de $\mathcal{V}$. Este último conjunto es una base que contiene las y.

EJERCICIOS

1. (a) Pruébese que los cuatro vectores

$$x = (1, 0, 0),$$
$$y = (0, 1, 0),$$
$$z = (0, 0, 1),$$
$$u = (1, 1, 1),$$

de $\mathcal{C}^3$ forman un conjunto linealmente independiente, pero que tres cualesquiera de ellos son linealmente independientes. (Para probar la dependencia lineal de los vectores $x = (\xi_1, \xi_2, \xi_3)$, $y = (\eta_1, \eta_2, \eta_3)$, y $z = (\zeta_1, \zeta_2, \zeta_3)$ en $\mathcal{C}^3$, procédase como sigue. Supóngase que α, β, y γ pueden determinarse de manera que $\alpha x + \beta y + \gamma z = 0$. Esto significa que

$$\alpha\xi_1 + \beta\eta_1 + \gamma\zeta_1 = 0,$$
$$\alpha\xi_2 + \beta\eta_2 + \gamma\zeta_2 = 0,$$
$$\alpha\xi_3 + \beta\eta_3 + \gamma\zeta_3 = 0.$$

Los vectores x, y y z son linealmente dependientes si y sólo si estas ecuaciones tienen una solución distinta de $\alpha = \beta = \gamma = 0$).

(b) Si los vectores x, y, z y u de $\mathcal{P}$ se definen por medio de $x(t) = 1$, $y(t) = t$, $z(t) = t$, $z(t) = t^2$, y $u(t) = 1 + t + t^2$, pruébese que x, y, z y u son linealmente dependientes, pero que tres cualesquiera de ellos son linealmente independientes.

2. Pruébese que si $\mathcal{R}$ es considerado como un espacio racional vectorial (véase § 3, (8),), entonces una condición necesaria y suficiente para que los vectores 1 y ξ de $\mathcal{R}$ sean linealmente independientes es que el número real ξ sea irracional.

3. ¿Es verdad que si x, y y z son vectores linealmente independientes, entonces también lo son $x + y$, $y + z$, y $z + x$?

4. (a) Bajo qué condiciones impuestas al escalar ξ son linealmente dependientes los vectores $(1 + \xi, 1 - \xi)$ y $(1 - \xi, 1 + \xi)$ de $\mathcal{C}^2$.

(b) ¿Bajo qué condiciones impuestas al escalar ξ son linealmente dependientes los vectores $(\xi, 1, 0)$, $(1, \xi, 1)$, y $(0, 1, \xi)$ de $\mathcal{R}^3$?

(c) ¿Cuál es la respuesta a (b) para $\mathcal{Q}^3$ (en lugar de $\mathcal{R}^3$)?

5. (a) Los vectores (ξ_1, ξ_2) y (η_1, η_2) son linealmente independientes si y sólo si $\xi_1\eta_2 = \xi_2\eta_1$.

(b) Encuentre una condición necesaria y suficiente similar para la dependencia lineal de dos vectores de $\mathcal{C}^3$. Haga lo mismo para tres vectores de $\mathcal{C}^3$.

(c) ¿Hay en $\mathcal{C}^2$ un conjunto de tres vectores linealmente independientes?

6. (a) ¿Bajo qué condiciones impuestas a los escalares ξ y η los vectores $(1, \xi)$ y $(1, \eta)$ de $\mathcal{C}^2$ son linealmente dependientes?

(b) Bajo qué condiciones impuestas a los escalares ξ, η, y ζ son linealmente independientes los vectores $(1, \xi, \xi^2)$, $(1, \eta, \eta^2)$ y $(1, \zeta, \zeta^2)$ de $\mathcal{C}^3$?

(c) Estime y pruebe una generalización de (a) y (b) a $\mathcal{C}^n$.

7. (a) Encuentre dos bases en $\mathcal{C}^4$ tales que los únicos vectores comunes a ambas sean $(0, 0, 1, 1)$ y $(1, 1, 0, 0)$.

(b) Encuentre dos bases en $\mathcal{C}^4$ que no tengan vectores en común, de manera que uno de ellos contenga los vectores $(1, 0, 0, 0)$ y $(1, 1, 0, 0)$ y el otro uno que contenga los vectores $(1, 1, 1, 0)$ y $(1, 1, 1, 1)$.

8 (a) ¿Bajo qué condiciones impuestas al escalar ξ forman una base de $\mathcal{C}^3$ los vectores $(1, 1, 1)$ y $(1, \xi, \xi^2)$?

(b) ¿Bajo qué condiciones impuestas al escalar ξ forman una base de ξ los vectores $(0, 1, \xi)$, $(\xi, 0, 1)$ y $(\xi, 1, 1, + \xi)$?

9. Considérese el conjunto de todos los vectores de $\mathcal{C}^3$ cada una de cuyas coordenadas es 0 o 1; ¿cuántas bases diferentes contiene este conjunto?

10. Si $\mathfrak{X}$ es el conjunto consistente en los seis vectores $(1, 1, 0, 0)$, $(1, 0, 1, 0)$, $(1, 0, 0, 1)$, $(0, 1, 1, 0)$, $(0, 1, 0, 1)$, $(0, 0, 1, 1)$ de $\mathcal{C}^4$, encuéntrense dos subconjuntos maximales de $\mathfrak{X}$, linealmente independientes. (Un subconjunto maximal de $\mathfrak{X}$ linealmente independiente es un subconjunto $\mathfrak{Y}$ de $\mathfrak{X}$ que se hace linealmente independiente cada vez que se adjunta a $\mathfrak{X}$ un vector de $\mathfrak{Y}$ que no está aún en $\mathfrak{Y}$.)

11. Pruebe que todo espacio vectorial tiene una base. (La prueba de este hecho está fuera del alcance de los que no estén familiarizados con algún artificio transfinito, como el lema de Zorn.)

§ 8. DIMENSION

TEOREMA 1. *El número de elementos de cualquier base de un espacio finito-dimensional* $\mathcal{V}$, *es el mismo que para cualquiera otra base.*

PRUEBA. La prueba de este teorema es un ligero refinamiento del método usado en el § 6, también, incidentalmente, prueba algo más de lo que enuncia el teorema. Sean $\mathfrak{X} = \{x_1, \ldots, x_n\}$ y $\mathcal{Y} = \{y_1, \ldots, y_m\}$ dos conjuntos finitos de vectores, cada uno con una de las dos propiedades definitorias de una base; esto es, suponemos que todo vector de $\mathcal{V}$ es una combinación lineal de las x (pero que no que las x son linealmente independientes), y suponemos que las y son linealmente independientes (pero no que cada vector es una combinación lineal de las mismas. Aplicamos el teorema § 6, justamente como anteriormente, al conjunto $\mathcal{S}$ de vectores

$$y_m, x_1, \ldots, x_n$$

Sabemos, por otra parte, que todo vector es una combinación de vectores de $\mathcal{S}$ y que $\mathcal{S}$ es linealmente independiente. Razonado justamente como antes, obtenemos un conjunto $\mathcal{S}'$ de vectores

$$y_m, x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n,$$

de nuevo con la propiedad de que todo vector es una combinación lineal de vectores de $\mathcal{S}'$. Ahora escribimos y_{m-1} en frente de los vectores de $\mathcal{S}'$ y aplicamos el mismo argumento. Continuando en esta forma, vemos que las x no se agotarán antes de las y, pues de otro modo las y que permanecieran tendrían que ser combinaciones lineales de las ya incorporadas a $\mathcal{S}$, pero sabemos que las y son linealmente independientes. En otras palabras, después de que el argumento ha sido aplicado m veces, obtenemos un conjunto con la misma propiedad que tenían las x, y este conjunto difiere del conjunto de las x en que m de ellas son reemplazadas por las y. Este enunciado, aparentemente inocente, es lo que estamos buscando; implica que $n \geqq m$. Consecuentemente, si tanto $\mathfrak{X}$ como $\mathcal{Y}$ son bases (de manera que cada una tenga ambas propiedades), entonces $n \geqq m$ y $m \geqq n$.

DEFINICIÓN. La *dimensión* de un espacio vectorial finito-dimensional es el número de una base de $\mathcal{V}$.

Obsérvese que puesto que el conjunto vacío de vectores es una base del espacio trivial $\mathcal{O}$, la definición implica que ese espacio tiene dimensión 0. Al mismo tiempo la definición (juntamente con el hecho de que ya hemos exhibido, en el § 7, una base particular de $\mathcal{C}^n$) justifica por fin nuestra terminología y nos faculta para anunciar el grato resultado: el espacio de coordenadas n-dimensionales, es n-dimensional. (Puesto que el argumento es el mismo para $\mathcal{R}^n$ y para $\mathcal{C}^n$, la aserción es verdadera tanto en el caso real como en el complejo).

Nuestro siguiente resultado es un corolario del Teorema 1 (vía el teorema del § 7).

TEOREMA 2. *Todo conjunto de $n + 1$ vectores en un espacio vectorial n-dimensional $\mathcal{V}$ es linealmente dependiente. Un conjunto de n vectores de $\mathcal{V}$ es una base si y sólo si es linealmente independiente, o, alternativamente, si y sólo si todo vector de $\mathcal{V}$ es una combinación lineal de elementos del conjunto.*

§ 9. ISOMORFISMO

Como aplicación de la noción de base lineal, o sistema de coordenadas, cumpliremos ahora una previa promesa implícita, demostrando que todo espacio vectorial finito-dimensional sobre un campo $\mathcal{F}$ es esencialmente el mismo (en lenguaje técnico, es isomórfico a) que algunos $\mathcal{F}^n$.

DEFINICIÓN. Dos espacios vectoriales $\mathcal{U}$ y $\mathcal{V}$ (sobre el mismo campo) son *isomórficos* si hay una correspondencia uno a uno entre los vectores x de $\mathcal{U}$ y los vectores y de $\mathcal{V}$, digamos $y = T(x)$, tal que

$$T(\alpha_1 x_1 + \alpha_2 x_2) = \alpha_1 T(x_1) + \alpha_2 T(x_2).$$

En otras palabras, $\mathcal{U}$ y $\mathcal{V}$ son isomórficos si hay un isomorfismo (tal como T) entre ellos, donde un *isomorfismo* es una correspondencia uno a uno que preserva todas las relaciones lineales.

Es fácil ver que los espacios vectoriales finito-dimensionales isomórficos tienen la misma dimensión; a cada base en un espacio corresponde una base en el otro espacio. Así, la dimensión es un isomorfismo invariante; demostraremos ahora que es el único isomorfismo invariante, en el sentido de que cada dos espacios vectoriales con la misma dimensión finita (sobre el mismo campo, por supuesto), son isomórficos. Puesto que el isomosfismo de $\mathcal{U}$ y $\mathcal{V}$ por una parte, y de $\mathcal{V}$ y $\mathcal{W}$ por la otra, implica que $\mathcal{U}$ y $\mathcal{W}$ son isomórficos, será suficiente comprobar el siguiente teorema.

TEOREMA. *Todo espacio vectorial n-dimensional $\mathcal{V}$ sobre un campo es isomórfico a $\mathcal{F}^n$.*

PRUEBA. Sean $\{x_1, \ldots, x_n\}$ cualquier base en $\mathcal{V}$. Cada x de $\mathcal{V}$ puede ser escrita en la forma $\xi_1 x_1 + \ldots + \xi_n x_n$, y sabemos que los escalares $\xi_1, \ldots, \xi_n$ están determinados unívocamente por x. Consideramos la correspondencia uno a uno

$$x \rightleftarrows (\xi_1, \cdots, \xi_n)$$

entre $\mathcal{V}$ y $\mathcal{F}^n$. Si $y = \eta_1 x_1 + \ldots + \eta_n x_n$, entonces

$$\alpha x + \beta y = (\alpha\xi_1 + \beta\eta_1)x_1 + \cdots + (\alpha\xi_n + \beta\eta_n)x_n;$$

esto establece el deseado isomorfismo.

Podría estar uno tentado de decir que de ahora en adelante sería inútil tratar de preservar una apariencia de generalidad hablando del espacio vectorial general n-dimensional, puesto que sabemos que, desde el punto de vista del estudio de los problemas lineales, los espacios vectoriales isomórficos son indistinguibles y consecuentemente, podríamos igualmente estudiar $\mathcal{F}^n$. Hay algo importante. Las propiedades más importantes de los vectores y los espacios vectoriales son las que son independientes de los sistemas de coordenadas o, en otras palabras, las que son invariantes bajo isomorfismos. La correspondencia entre $\mathcal{V}$ y $\mathcal{F}^n$, era, sin embargo, establecida escogiendo un sistema de coordenadas; si fuéramos siempre a estudiar $\mathcal{F}^n$, siempre estaríamos atados a ese sistema de coordenadas en particular, o de otro modo siempre estaríamos enfrentados a la tarea de demostrar que nuestras definiciones y teoremas son independientes del sistema de coordenadas en que se enuncian. (Este horrible dilema se verá claro más tarde, en las pocas ocasiones en que nos veremos forzados a usar un sistema particular de coordenadas para dar una definición). Consecuentemente, en la mayor parte de este libro pasaremos por alto el teorema que se acaba de probar, y trataremos los espacios vectoriales n-dimensionales como entidades relativas a sí mismas, independientemente de cualquier base. Además de las razones que acabamos de mencionar, hay otra razón para hacer esto: muchos ejemplos de espacios vectoriales, como por ejemplo $\mathcal{P}_n$, perderían mucho de su contenido intuitivo si fuéramos a transformarlos en $\mathcal{C}^n$ y hablar solamente de coordenadas. Al estudiar espacios vectoriales, tales como $\mathcal{P}_n$, y su relación con otros espacios vectoriales, debemos ser capaces de manejarlos con igual facilidad en diferentes sistemas de coordenadas, o, como esto es esencialmente la misma cosa, debemos ser capaces de manejarlos sin usar en absoluto ningún sistema de coordenadas.

EJERCICIOS

1. (a) ¿Cuál es la dimensión del conjunto $\mathcal{C}$ de todos los números complejos, considerados como un espacio real vectorial? (Véase el § 3 (9).)

(b) Todo espacio vectorial complejo $\mathcal{V}$ está íntimamente asociado con un espacio vectorial real $\mathcal{V}^-$; el espacio $\mathcal{V}^-$ se obtiene de $\mathcal{V}$ rehusando multiplicar vectores de $\mathcal{V}$ por otra cosa que no sean escalares reales. Si la dimensión del espacio escalar complejo $\mathcal{V}$ es n, ¿cuál es la dimensión del espacio vectorial real $\mathcal{V}^-$?

2. ¿Es el conjunto $\mathcal{R}$ de todos los números reales un espacio vectorial finito dimensional sobre el campo Q de todos los números racionales? (véase § 3 (8). La cuestión no es trivial ayuda a saber algo sobre números cardinales).

3. ¿Cuántos vectores hay en un espacio vectorial finito dimensional sobre el campo Z_p (donde p es un subíndice)?

4. Discuta la siguiente afirmación: si dos espacios vectoriales racionales tienen el mismo número cardinal (esto es, si hay entre ellos cierta correspondencia uno a uno), entonces son isomórficos (esto es, hay entre ellos una correspondencia uno a uno preservativa de la linearidad). Para una inteligente discusión se necesita un conocimiento de estos hechos básicos de la aritmética de cardinales.

§ 10. SUBESPACIOS

Los objetos de interés en geometría no son solamente los puntos del espacio que se consideran, sino también sus líneas, planos, etc. Procedemos a estudiar los análogos, en general espacios vectoriales, de estos elementos de más elevada dimensionalidad.

DEFINICIÓN. Un subconjunto no vacío $\mathfrak{M}$ de un espacio vectorial es un *subespacio* $\mathcal{V}$ o un *múltiple lineal*, si juntamente con cada par, x y y, de vectores contenidos en $\mathfrak{M}$, cada combinación lineal $\alpha x + \beta y$ está también contenido en $\mathfrak{M}$.

Una palabra de advertencia: juntamente con cada vector x, un subespacio también contiene a $x - x$. De aquí que si interpretamos los subespacios como líneas y planos generalizados, debemos tener cuidado de considerar solamente planos y líneas que pasan por el origen.

Un subespacio $\mathfrak{M}$ de un espacio vectorial $\mathfrak{V}$ es él mismo un espacio vectorial; el lector puede fácilmente verificar que, con las mismas definiciones de adición y multiplicación escalar que tuvimos en $\mathfrak{V}$, el conjunto satisface los axiomas (**A**), (**B**) y (**C**) del § 2.

Dos ejemplos especiales de subespacios son: (i) el conjunto $\mathfrak{O}$, que consiste en el origen solamente, y (ii) el espacio total $\mathfrak{V}$. Los siguientes ejemplos son menos triviales.

(1) Sean n y m dos enteros estrictamente positivos cualesquiera, $m \leq n$. Sea $\mathfrak{M}$ el conjunto de todos los vectores $x = (\xi_1, \ldots, \xi_n)$ en $\mathcal{C}^n$ para los cuales $\xi_1 = \ldots = \xi_m = 0$.

(2) Con m y n como en (1) consideremos el espacio $\mathcal{P}_n$, y cualesquiera números reales $t_1, \ldots, t_m$. Sea $\mathfrak{M}$ el conjunto de todos los vectores (polinomios) x en $\mathcal{P}_n$ para los cuales $x(t_1) = \ldots = x(t_m) = 0$.

3. Sea $\mathfrak{M}$ el conjunto de todos los vectores x en $\mathcal{P}$ para los cuales $x(t) = x(-t)$ vale idénticamente en t.

Necesitamos alguna notación y alguna terminología. Para cualquier colección $\{\mathfrak{M}_\nu\}$ de subconjuntos de un conjunto dado (digamos, por ejemplo, para una colección de subespacios en un espacio vectorial $\mathfrak{V}$), escribimos $\bigcap_\nu \mathfrak{M}_\nu$ para la *intersección* de todos los $\mathfrak{M}_\nu$, esto es, para el conjunto de puntos común a todos ellos. También, si $\mathfrak{M}$ y $\mathfrak{N}$ son subconjuntos de un conjunto, escribimos $\mathfrak{M} \subset \mathfrak{N}$ si $\mathfrak{M}$ es un subconjunto de $\mathfrak{N}$, esto es, si todo elemento de $\mathfrak{M}$ queda también en $\mathfrak{N}$. (Obsérvese que no excluimos la posibilidad de $\mathfrak{M} = \mathfrak{N}$; así, escribimos $\mathfrak{V} \subset \mathfrak{V}$ así como también $\mathfrak{O} \subset \mathfrak{V}$.) Para una colección finita $\{\mathfrak{M}_1, \ldots, \mathfrak{M}_n\}$, escribiremos $\mathfrak{M}_1 \cap \cdots \cap \mathfrak{M}_n$ en vez de $\bigcap_\nu \mathfrak{M}_\nu$; en caso de que dos subespacios $\mathfrak{M}$ y $\mathfrak{N}$ sean tales que $\mathfrak{M} \cap \mathfrak{N} = \mathfrak{O}$, diremos que $\mathfrak{M}$ y $\mathfrak{N}$ están *disjuntos*.

§ 11. CALCULO DE SUBESPACIOS

TEOREMA 1. *La intersección de cualquier colección de subespacios es un subespacio.*

PRUEBA. Si usamos un índice ν para designar aparte los miembros de la colección, de manera que los subespacios dados sean $\mathfrak{M}_\nu$, escribiremos

$$\mathfrak{M} = \bigcap_\nu \mathfrak{M}_\nu.$$

Puesto que cada $\mathfrak{M}_\nu$ contiene a 0, también $\mathfrak{M}$, y, en consecuencia $\mathfrak{M}$ no está vacío. Si x y y pertenecen a $\mathfrak{M}$ (esto es, a todas las $\mathfrak{M}_\nu$), entonces $\alpha x + \beta y$ pertenece a todas las $\mathfrak{M}_\nu$, y, en consecuencia, $\mathfrak{M}$ es un subespacio.

Para ver una aplicación de este teorema supóngase que es un conjunto arbitrario de vectores (no necesariamente un subespacio), en espacio vectorial $\mathcal{V}$. Ciertamente existen subespacios $\mathfrak{M}$ que contienen todos los elementos de $\mathcal{S}$ (esto es, tales que $\mathcal{S} \subset \mathfrak{M}$); el espacio total $\mathcal{V}$ es, por ejemplo, ese subespacio. Sea $\mathfrak{M}$ la intersección de todos los subespacios que contienen $\mathcal{S}$; está claro que $\mathfrak{M}$ mismo es un subespacio que contiene $\mathcal{S}$. Está claro, además, que $\mathfrak{M}$ es el más pequeño de estos subespacios, si $\mathcal{S}$ también está contenido en el subespacio $\mathfrak{N}$, $\mathcal{S} \subset \mathfrak{N}$, entonces $\mathfrak{M} \subset \mathfrak{N}$. El subespacio $\mathfrak{M}$ así definido se llama el espacio sobretendido por $\mathcal{S}$ o la sobretensión de $\mathcal{S}$. El siguiente resultado establece la conexión entre la noción de sobretensión y los conceptos estudiados en los §§ 5-9.

TEOREMA 2. *Si $\mathcal{S}$ es cualquier conjunto de vectores en un espacio vectorial $\mathcal{V}$ y si $\mathfrak{M}$ es el subespacio sobretendido por $\mathcal{S}$, entonces $\mathfrak{M}$ es lo mismo que todo el conjunto de combinaciones lineales de $\mathcal{S}$.*

PRUEBA. Está claro que una combinación lineal de combinaciones lineales de elementos de $\mathcal{S}$ puede de nuevo ser escrito como una combinación lineal de elementos de $\mathcal{S}$. De aquí que el conjunto de todas las combinaciones lineales de elementos de $\mathcal{S}$ es un subespacio que contiene a $\mathcal{S}$; se sigue que este subespacio debe también contener a $\mathfrak{M}$. Pongamos ahora el argumento a la recíproca: $\mathfrak{M}$ contiene a $\mathcal{S}$ y es un subespacio; de aquí que $\mathfrak{M}$ contenga todas las combinaciones lineales de elementos de $\mathcal{S}$.

Vemos, en consecuencia, que en nuestra terminología podemos definir una base lineal como un conjunto de vectores linealmente independientes que sobrepasa todo el espacio.

Nuestro siguiente resultado será una fácil consecuencia del Teorema 2; su prueba puede ser dejada con confianza al lector.

TEOREMA 3. *Si $\mathcal{H}$ y $\mathcal{K}$ son dos subespacios cualesquiera y si $\mathfrak{M}$ es el subespacio sobretendido por $\mathcal{H}$ y $\mathcal{K}$ juntamente, entonces $\mathfrak{M}$ es igual al conjunto de todos los vectores de la forma $x + y$, con x en $\mathcal{H}$ y y en $\mathcal{K}$.*

Teniendo en cuenta este teorema, usaremos la notación $\mathcal{H} + \mathcal{K}$ para el subespacio $\mathfrak{M}$ sobretendido por $\mathcal{H}$ y $\mathcal{K}$. Diremos que un subespacio $\mathcal{K}$ de un espacio vectorial $\mathcal{V}$ es un *complemento* de un subespacio $\mathcal{H}$ si $\mathcal{H} \cap \mathcal{K} = \mathcal{O}$ y

§ 12. DIMENSION DE UN SUBESPACIO

TEOREMA 1. *Un subespacio $\mathfrak{M}$ de un espacio vectorial n-dimensional $\mathcal{V}$ es un espacio vectorial de dimensión $\leqq n$.*

PRUEBA. Es posible dar una prueba de este teorema, engañosamente corta, que va como sigue. Todo conjunto de $n + 1$ vectores de $\mathcal{V}$ es linealmente dependiente; de aquí que lo mismo sea verdad de $\mathfrak{M}$; de aquí, en particular, que el número de elementos de cada base de $\mathfrak{M}$ es $\leqq n$, Q.E.D.

La dificultad con este argumento, estriba en que definimos la dimensión n requiriendo en primer lugar que exista una base finita, y a continuación demandando que esta base contenga exactamente, n elementos. La prueba anterior demuestra solamente que ninguna base puede contener más de n elementos; no demuestra que exista base alguna. Una vez que se observa la dificultad, sin embargo, es fácil llenar el hueco. Si $\mathfrak{M} = \mathcal{O}$, entonces $\mathfrak{M}$ es 0-dimensional y está listo. Si $\mathfrak{M}$ contiene un vector no cero, x_1, sea $\mathfrak{M}_1$ ($\subset \mathfrak{M}$) el subespacio sobretendido por x_1. Si $\mathfrak{M} = \mathfrak{M}_1$ entonces $\mathfrak{M}$ es 1-dimensional y ya está. Si $\mathfrak{M} \neq \mathfrak{M}_1$ sea x_2 un elemento de $\mathfrak{M}$ no contenido en $\mathfrak{M}_1$ y sea $\mathfrak{M}_2$ el subespacio sobretendido por x_1 y x_2 y así sucesivamente. Ahora podemos emplear legítimamente el argumento dado anteriormente; después de no más de n pasos de este tipo, el proceso llega a su fin, puesto que (por el § 8, Teorema 2), no podemos encontrar $n + 1$ vectores linealmente independientes

El siguiente resultado es una importante consecuencia de esa segunda y correcta prueba del Teorema 1.

TEOREMA 2. *Dado cualquier subespacio m-dimensional $\mathfrak{M}$ en un espacio vectorial n-dimensional $\mathcal{V}$, podemos encontrar una base $\{x_1, \ldots, x_m, x_{m+1}, \ldots, x_n\}$ en $\mathcal{V}$, de manera que $x_1, \ldots, x_m$ están en $\mathfrak{M}$ y forman, en consecuencia, una base de $\mathfrak{M}$.*

Denotaremos la dimensión de un espacio vectorial $\mathcal{V}$ por el símbolo dim $\mathcal{V}$. En esta notación, el teorema 1 enuncia que si $\mathfrak{M}$ es un subespacio de un espacio vectorial finito-dimensional $\mathcal{V}$, entonces dim $\mathfrak{M} \leqq$ dim $\mathcal{V}$.

EJERCICIOS

1. Si $\mathfrak{M}$ y $\mathfrak{N}$ son subespacios finito-dimensionales con la misma dirección, y si $\mathfrak{M} \subset \mathfrak{N}$, entonces $\mathfrak{M} = \mathfrak{N}$.

2. Si $\mathfrak{M}$ y $\mathfrak{N}$ son subespacios de un espacio vectorial $\mathfrak{V}$ y si todo vector de $\mathfrak{V}$ pertenece a $\mathfrak{M}$ o a $\mathfrak{N}$ (o a ambos), entonces, o $\mathfrak{M} = \mathfrak{V}$ o $\mathfrak{N} = \mathfrak{V}$ (o ambos).

3. Si x, y y z son vectores tales que $x + y + z = 0$, entonces x y y sobretienden el mismo subespacio que y y z.

4. Supóngase que x y y son vectores y $\mathfrak{M}$ es un subespacio en un espacio vectorial $\mathfrak{V}$; sea $\mathfrak{H}$ el subespacio sobretendido por $\mathfrak{M}$ y x, y sea $\mathfrak{K}$ el espacio sobretendido por $\mathfrak{M}$ y y. Pruébese que si y está en $\mathfrak{H}$ pero no en $\mathfrak{M}$, entonces x está en $\mathfrak{K}$.

5. Supóngase que $\mathfrak{L}, \mathfrak{M}$, y $\mathfrak{N}$ su subespacios de un espacio vectorial.
(a) Demuéstrese que la ecuación

$$\mathfrak{L} \cap (\mathfrak{M} + \mathfrak{N}) = (\mathfrak{L} \cap \mathfrak{M}) + (\mathfrak{L} \cap \mathfrak{N})$$

no es necesariamente verdadera.
(b) Pruébese que

$$\mathfrak{L} \cap (\mathfrak{M} + (\mathfrak{L} \cap \mathfrak{N})) = (\mathfrak{L} \cap \mathfrak{M}) + (\mathfrak{L} \cap \mathfrak{N}).$$

6. (a) ¿Puede suceder que un subespacio no trivial de un espacio vectorial $\mathfrak{V}$ (esto es, un espacio diferente de $\mathfrak{O}$ y de $\mathfrak{V}$) tenga un complemento único?
(b) Si $\mathfrak{M}$ es un subespacio vectorial m-dimensional, entonces todo complemento de $\mathfrak{M}$ tiene dimensión n-m.

7. (a) Demuestre que si tanto $\mathfrak{M}$ como $\mathfrak{N}$ son subespacios tridimensionales de un espacio vectorial de cinco dimensiones, entonces $\mathfrak{M}$ y $\mathfrak{N}$ no están disjuntos.
(b) Si $\mathfrak{M}$ y $\mathfrak{N}$ son subespacios finito-dimensionales de un espacio vectorial, entonces

$$\dim \mathfrak{M} + \dim \mathfrak{N} = \dim (\mathfrak{M} + \mathfrak{N}) + \dim (\mathfrak{M} \cap \mathfrak{N}).$$

8. Un polinomio x se llama par si $x(-t) = x(t)$ idénticamente en t (véase § 10 (3), y se llama impar si $x(-t) = -x(t)$.
(a) Tanto la clase $\mathfrak{M}$ de polinomios pares como la clase $\mathfrak{N}$ de polinomios impares son subespacios del espacio $\mathcal{P}$ de todos los polinomios (complejos).
(b) Pruébese que $\mathfrak{M}$ y $\mathfrak{N}$ son completamente uno del otro.

§ 13. ESPACIOS DUALES

DEFINICIÓN. Una función lineal en un espacio vectorial $\mathfrak{V}$ es una función y, valuada escalarmente, definida para cada vector x, con la propiedad de que (idénticamente en los vectores x_1 y x_2 y en los escalares α_1 y α_2)

$$y(\alpha_1 x_1 + \alpha_2 x_2) = \alpha_1 y(x_1) + \alpha_2 y(x_2).$$

Leamos algunos ejemplos de funcionales lineales

(1) Para $x = (\xi_1, \ldots, \xi_n)$ en $\mathfrak{C}^n$, escríbase $y(x) = \xi_1$. Más generalmente, sean $\alpha_1, \ldots, \alpha_n$ cualesquiera escalares n y escribamos

$$y(x) = \alpha_1\xi_1 + \cdots + \alpha_n\xi_n.$$

Observamos que para cualquier funcional lineal y en cualquier espacio vectorial

$$y(0) = y(0 \cdot 0) = 0 \cdot y(0) = 0;$$

por esta razón una funcional lineal, como la definimos, es algunas veces llamada homogénea. En particular $\mathfrak{C}^n$, si y es definida por

$$y(x) = \alpha_1\xi_1 + \cdots + \alpha_n\xi_n + \beta,$$

entonces y no es una funcional lineal, a menos que $\beta = 0$.

(2) Para cualquier polinomio x en $\mathcal{P}$, escríbase $y(x) = x(0)$. Más generalmente, sea $\alpha_1, \ldots, \alpha_n$, n escalares cualesquiera, sean $t_1, \ldots, t_n$, n cualesquiera números reales y escríbase

$$y(x) = \alpha_1 x(t_1) + \cdots + \alpha_n x(t_n).$$

Otro ejemplo, en un sentido, se obtiene como sigue un caso límite del que sea acaba de dar. Sea (a, b) cualquier intervalo finito sobre el eje real de las t y sea α cualquier función integrable valuada complejamente, definida sobre (a, b); defínase y por

$$y(x) = \int_a^b \alpha(t)x(t)\,dt.$$

(3) En un espacio vectorial arbitrario $\mathfrak{V}$, defínase y escribiendo

$$y(x) = 0$$

para toda x de $\mathfrak{V}$.

El último ejemplo es la primera insinuación de una situación general. Sea $\mathfrak{V}$ cualquier espacio vectorial y sea $\mathfrak{V}'$ la colección de todas las funcionales lineales sobre $\mathfrak{V}$. Denotemos por medio de 0 la funcional lineal definida en (3) (compárese el comentario al final del § 4). Si y_1 y y_2 son funcionales lineales definidas sobre $\mathfrak{V}$ y si α_1 y α_2 son escalares, escribamos y para la función definida por

$$y(x) = \alpha_1 y_1(x) + \alpha_2 y_2(x).$$

Es fácil ver que y es una funcional lineal; la denotamos por $\alpha_1 y_1 + \alpha_2 y_2$. Con estas definiciones de los conceptos lineales (cero, adición, multiplicación escalar), el conjunto $\mathfrak{V}'$ forma un espacio vectorial el *espacio dual* de $\mathfrak{V}$.

§ 14. CORCHETES

Antes de estudiar las funcionales lineales y los espacios duales con más detalle, deseamos introducir una notación que puede parecer extraña a primera vista, pero que aclarará muchas situaciones posteriores. Generalmente denotamos una función lineal por una sola letra tal como y. Algunas veces, sin embargo, es necesario usar plenamente la notación de función e indicar de algún modo que si y es una funcional lineal sobre $\mathcal{V}$ y si x es un vector de $\mathcal{V}$, entonces $y(x)$ es un escalar particular. De acuerdo con la notación que nos proponemos adoptar aquí, no escribiremos y seguida de x en paréntesis, sino que escribiremos x y y encerradas en corchetes y separadas por una coma. A causa de la naturaleza desusada de esta notación, le dedicaremos algunas palabras más.

Como hemos hecho notar $[x, y]$ es un sustituto del símbolo ordinario de función $y(x)$; ambos símbolos denotan el escalar que obtenemos si tomamos el valor de la función lineal y en el vector x. Tomemos una situación análoga (relativa a funciones que no son, sin embargo, lineales). Sea y la función real de una variable real definida para cada número real x por $y(x) = x^2$. La notación $[x\ y]$ es una forma simbólica de escribir la receta para las operaciones que de hecho se ejecutan; corresponde a la oración [tome un número y elévelo al cuadrado].

Usando esta notación, podemos sumar: a cada espacio vectorial $\mathcal{V}$ hacemos corresponder el espacio dual $\mathcal{V}'$ consistente en todas las funciones lineales sobre $\mathcal{V}$; a cada par, x y y, donde x es un vector de $\mathcal{V}$ y y es una función lineal de $\mathcal{V}'$, hacemos corresponder el escalar $[x, y]$, definido como el valor de y en x. En términos del símbolo $[x, y]$ la propiedad definitoria de una funcional lineal es

$$(1) \qquad [\alpha_1 x_1 + \alpha_2 x_2, y] = \alpha_1[x_1, y] + \alpha_2[x_2, y],$$

y la definición de las operaciones lineales para las funcionales lineales es

$$(2) \qquad [x, \alpha_1 y_1 + \alpha_2 y_2] = \alpha_1[x, y_1] + \alpha_2[x, y_2].$$

Las dos relaciones juntas se expresan diciendo que $[x, y]$ es una *funcional bilineal* de los vectores x de $\mathcal{V}$ y y de $\mathcal{V}'$.

EJERCICIOS

1. Considérese el conjunto $\mathcal{C}$ de números complejos como un espacio vectorial real (como en el § 3(9)). Supóngase que para cada $x = \xi_1 + i\xi_2$ (donde ξ_1 y ξ_2 son números reales e $i = \sqrt{-1}$ la función y se define por

(a) $y(x) = \xi_1$,
(b) $y(x) = \xi_2$,
(c) $y(x) = \xi_1^2$,
(d) $y(x) = \xi_1 - i\xi_2$,
(e) $y(x) = \sqrt{\xi_1^2 + \xi_2^2}$. (El signo de raíz cuadrada adjunto a un número positivo siempre denota la raíz cuadrada positiva de ese número.)

2. Supóngase que para cada $x = (\xi_1, \xi_2, \xi_3)$ en $\mathcal{C}^3$ la función y es definida por

(a) $y(x) = \xi_1 + \xi_2$,
(b) $y(x) = \xi_1 - \xi_2^2$,
(c) $y(x) = \xi_1 + 1$,
(d) $y(x) = \xi_1 - 2\xi_2 + 3\xi_3$.

¿En cuál de estos casos es y una funcional lineal?

3. Supóngase que para cada x de $\mathcal{P}$ la función y es definida por

(a) $y(x) = \int_{-1}^{+2} x(t)\, dt$,

(b) $y(x) = \int_0^2 (x(t))^2\, dt$,

(c) $y(x) = \int_0^1 t^2 x(t)\, dt$,

(d) $y(x) = \int_0^1 x(t^2)\, dt$,

(e) $y(x) = \frac{dx}{dt}$,

(f) $y(x) = \left.\frac{d^2x}{dt^2}\right|_{t=1}$.

¿En cuál de estos casos es y una funcional lineal?

4. Si $(\alpha_0, \alpha_1, \alpha_2, \ldots)$ es una secuencia arbitraria de números complejos y si x es un elemento de $\mathcal{P}$, $x(t) = \sum_{i=0}^{n} \xi_i t^i$, escriba $y(x) = \sum_{i=0}^{n} \xi_i \alpha_i$. Pruébese que y es un elemento de $\mathcal{P}'$ y que todo elemento de $\mathcal{P}'$ puede obtenerse en esta forma mediante una adecuada selección de las α.

5. Si y es una funcional lineal no cero sobre un espacio vectorial $\mathcal{V}$, y si α es un escalar arbitrario, ¿existe necesariamente un vector x de $\mathcal{V}$ tal que $[x, y] = \alpha$?

6. Pruébese que si y y z son funcionales lineales (sobre el mismo espacio vectorial), tales como $[x, y] = 0$, cada vez que $[x, z] = 0$ entonces existe un escalar α tal que $y = \alpha z$. (Sugestión: si $[x_0, z] \neq 0$, escríbase $\alpha = [x_0, y]/[x_0, z]$).

§ 15. BASES DUALES

Una palabra más antes de embarcarnos en las pruebas de los teoremas importantes. El concepto de espacio dual se definió sin ninguna referencia a sistemas de coordenadas, una mirada a las siguientes pruebas mostrará una superabundancia de sistemas de coordenadas. Deseamos hacer notar que este fenómeno es inevitable; estableceremos resultados concernientes a la dimensión y la dimensión es el único concepto (hasta ahora) cuya definición se da en términos de una base

TEOREMA 1. *Si* $\mathcal{V}$ *es un espacio vectorial n-dimensional, si* $\{x_1, \ldots, x_n\}$ *es una base de* $\mathcal{V}$, *y si* $\{\alpha_1, \ldots, \alpha_n\}$ *es un conjunto cualquiera de n escalares, entonces hay una, y sólo una y funcional lineal sobre* $\mathcal{V}$ *tal que* $[x_i, y] = \alpha_i$ *para* $i = 1, \ldots, n$.

PRUEBA. Toda x de $\mathcal{V}$ puede escribirse en la forma $x = \xi_1 x_1 + \ldots + \xi_n x_n$ de un modo y sólo uno; si y es cualquier funcional lineal, entonces

$$[x, y] = \xi_1[x_1, y] + \cdots + \xi_n[x_n, y].$$

De esta relación resulta clara la univocidad de y; si $[x_i, y] = \alpha_i$, entonces el valor de $[x, y]$ está determinado, para cada x, por $[x, y] = \sum_i \xi_i\alpha_i$. Puede darse también la vuelta al argumento; si definimos y por

$$[x, y] = \xi_1\alpha_1 + \cdots + \xi_n\alpha_n,$$

entonces y es verdaderamente una funcional lineal, y $[x_i, y] = \alpha_i$.

TEOREMA 2. *Si* $\mathcal{V}$ *es un espacio vectorial n-dimensional y si* $= \{x_1, \ldots, x_n\}$ *es una base de* $\mathcal{V}$, *entonces hay una base determinada en forma unívoca,* $\mathfrak{X}'$ *en* $\mathcal{V}'$, $\mathfrak{X}' = \{y_1, \ldots, y_n\}$, *con la propiedad de que* $[x_i, y_j] = \delta_{ij}$. *Consecuentemente, el espacio dual de un espacio n-dimensional es n-dimensional.*

La base $\mathfrak{X}'$ es llamada la *base dual* de $\mathfrak{X}$.

PRUEBA. Se sigue del Teorema 1 que, para cada $j = 1, \ldots, n$, una y_j única de $\mathcal{V}'$ puede determinarse de modo que $[x_i, y_j] = \delta_{ij}$; potenemos solamente que probar el conjunto $\mathfrak{X}' = \{y_1, \ldots, y_n\}$ es una base de $\mathcal{V}'$.

En primer lugar, $\mathfrak{X}'$ es un conjunto linealmente independiente, pues si tuvieramos $\alpha_1 y_1 + \ldots + \alpha_n y_n = 0$, en otras palabras, si

$$[x, \alpha_1 y_1 + \cdots + \alpha_n y_n] = \alpha_1[x, y_1] + \cdots + \alpha_n[x, y_n] = 0$$

para todas las x, entonces tendríamos, para $x = x_i$

$$0 = \sum_j \alpha_j[x_i, y_j] = \sum_j \alpha_j\delta_{ij} = \alpha_i.$$

En segundo lugar, toda y de $\mathcal{V}'$ es una combinación lineal de $y_1, \ldots, y_n$. Para probar esto, escríbase $[x_i, y] = \alpha_i$, entonces, para $x = \sum_i \xi_i x_i$, tenemos

$$[x, y] = \xi_1\alpha_1 + \cdots + \xi_n\alpha_n.$$

Por otra parte

$$[x, y_j] = \sum_i \xi_i[x_i, y_j] = \xi_j,$$

de manera que, sustituyendo en la ecuación precedente, tenemos

$$[x, y] = \alpha_1[x, y_1] + \cdots + \alpha_n[x, y_n]$$

$$= [x, \alpha_1 y_1 + \cdots + \alpha_n y_n].$$

Consecuentemente, $y = \alpha_1 y_1 + \ldots + \alpha_n y_n$, y la prueba del teorema es completa.

Necesitaremos también la siguiente fácil consecuencia del Teorema 2.

TEOREMA 3. *Si u y v son dos vectores diferentes cualesquiera del espacio vectorial n-dimensional $\mathcal{V}$, entonces existe una función lineal y sobre $\mathcal{V}$ tal que $[u, y] \neq [v, y]$; o, equivalentemente, a cualquier vector x no cero de $\mathcal{V}$ corresponde un y en $\mathcal{V}'$ tal que $[x, y] \neq 0$.*

PRUEBA. Que los dos enunciados contenidos en el teorema son en realidad equivalentes, se ve considerando $x = u - v$. En consecuencia, probaremos únicamente el último enunciado.

Sea $= \{x_1, \ldots, x_n\}$ cualquiera base de $\mathcal{V}$, y sea $\mathfrak{X}' = \{y_1, \ldots, y_n\}$ la base dual de $\mathcal{V}'$. Si $x = \sum_i \xi_i x_i$, entonces (como anteriormente. $[x, y_j] = \xi_j$. De aquí que si $[x, y] = 0$ para todas las y, y en particular, si $[x, y_j] = 0$ para $j = 1, \ldots, n$, entonces $x = 0$.

§ 16. REFLEXIBILIDAD

Es natural pensar que si el espacio dual $\mathcal{V}'$ de un espacio vectorial $\mathcal{V}$ y las relaciones entre un espacio y su dual, son de algún interés para $\mathcal{V}$ entonces, son justamente del mismo interés para $\mathcal{V}'$. En otras palabras, proponemos ahora formar el espacio dual $(\mathcal{V}')'$ de $\mathcal{V}'$; para la sencillez de la notación, lo indicaremos por $\mathcal{V}''$. La descripción verbal de un elemento de $\mathcal{V}''$ es torpe; ese elemento es una funcional lineal de funcionales lineales. Es, sin embargo, en este punto donde aparece la ventaja máxima de la notación $[x, y]$; por medio de ella es fácil discutir $\mathcal{V}$ y su relación con $\mathcal{V}''$.

Si consideramos el símbolo $[x, y]$ para alguna $y = y_0$ fija, no obtenemos nada nuevo: $[x, y_0]$ es meramente otro modo de escribir el valor $y_0(x)$ de la función y_0 en el vector x. Si, en cambio, consideramos el símbolo $[x, y]$ para alguna $x = x_0$ fija, entonces observamos que la función de los vectores de $\mathcal{V}'$ cuyo valor en y es $[x_0, y]$ es una función de valor escalar que resulta lineal [véase el § 14, (2)]; en otras palabras, $[x_0, y]$ define una función lineal sobre $\mathcal{V}'$ y, consecuentemente. un elemento de $\mathcal{V}''$.

Por este método hemos exhibido *algunas* funcionales lineales sobre $\mathcal{V}'$. ¿Las hemos exhibido todas? Para el caso finito-dimensional, el teorema siguiente suministra la respuesta afirmativa.

TEOREMA. *Si $\mathcal{V}$ es un espacio vectorial finito-dimensional entonces, correspondiendo a cada funcional lineal z_0 sobre $\mathcal{V}'$ hay un vector x_0 de $\mathcal{V}$ tal que $z_0(y) = [x_0, y\ ] = y(x_0)$ para toda y de $\mathcal{V}'$; la correspondencia $z_0 \rightleftarrows x_0$ entre $\mathcal{V}''$ y $\mathcal{V}$ es un isomorfismo.*

La correspondencia descrita en este enunciado se llama la *correspondencia natural* entre $\mathcal{V}''$ y $\mathcal{V}$.

PRUEBA. Revisemos la correspondencia desde el punto de vista de ir de $\mathcal{V}$ a $\mathcal{V}''$; en otras palabras, a cada x_0 de $\mathcal{V}$ hacemos corresponder un vector z_0 de $\mathcal{V}''$ definido por $z_0(y) = y(x_0)$ para cada y de $\mathcal{V}'$. Puesto que $[x, y]$ depende linealmente de x, la transformación $x_0 \rightarrow z_0$ es lineal.

Demostraremos que esta transformación es uno a uno, tal como va. Afirmamos, en otras palabras, que si x_1 y x_2 están en $\mathcal{V}$, y si z_1 y z_2 son los vectores correspondientes de $\mathcal{V}''$ (de manera que $z_1(y) = [x_1, y]$ y $z_2(y) = [x_2, y]$ para toda y de $\mathcal{V}'$), y si $z_1 = z_2$, entonces, $x_1 = x_2$. Decir que $z_1 = z_2$, significa que $[x_1, y] = [x_2\ y]$ para toda y de $\mathcal{V}'$; la conclusión deseada se sigue del § 15, Teorema 3.

Los dos últimos párrafos juntos demuestran que los conjuntos de funcionales lineales z sobre $\mathcal{V}'$ (esto es, elementos de $\mathcal{V}''$) que tienen la forma deseada (esto es, $z(y)$ es idénticamente igual a $[x, y]$ para una adecuada x de $\mathcal{V}$) es un subespacio de $\mathcal{V}''$ que es isomórfico a $\mathcal{V}$ y que es, en consecuencia, n-dimensional. Pero la n-dimensionalidad de $\mathcal{V}$ implica la de $\mathcal{V}'$, que a su vez implica que $\mathcal{V}''$ sea n-dimensional. Si sigue que $\mathcal{V}''$ debe coincidir con el espacio n-dimensional que se acaba de describir y la prueba del teorema es completa.

Es importante observar que el teorema demuestra no solamente que $\mathcal{V}$ y $\mathcal{V}''$ son isomórficos —hasta aquí es trival, por el hecho de que tienen la misma dimensión—, sino que la correspondencia natural es un isomorfismo. Esta propiedad de los espacios vectoriales se llama *reflexibilidad;* todo espacio vectorial finito-dimensional es reflexivo.

Frecuentemente, conviene cierta amplitud de criterio con respecto de $\mathcal{V}''$: para espacios vectoriales finito-dimensionales, identificaremos $\mathcal{V}''$ con $\mathcal{V}$ (por el natural isomorfismo) y diremos que el elemento z_0 de $\mathcal{V}''$ es *igual* al elemento z_0 de $\mathcal{V}$ cada vez que $z_0(y) = [x_0, y]$ para todas las y de $\mathcal{V}'$. En este lenguaje, es fácil expresar la relación entre una base $\mathfrak{X}$, en $\mathcal{V}$, y la base dual, de su base dual, en $\mathcal{V}''$; la simetría de la relación $[x_i, y_j] = \delta_{ij}$ demuestra que $\mathfrak{X}'' = \mathfrak{X}$.

§ 17. ANIQUILADORES

DEFINICIÓN. El *aniquilador* $\mathcal{S}^0$ de cualquier subconjunto $\mathcal{S}$ de un espacio vectorial $\mathcal{V}$ (no es necesario que $\mathcal{S}$ sea un subespacio) es el conjunto de todos los vectores y de $\mathcal{V}'$ tales que $[x, y]$ sea idénticamente igual a cero para toda x de $\mathcal{S}$.

Así $\mathcal{O}^0 = \mathcal{V}'$ y $\mathcal{V}^0 = \mathcal{O}$ $(\subset \mathcal{V}')$. Si $\mathcal{V}$ es finito-dimensional y $\mathcal{S}$ contiene un vector no cero, de manera que $\mathcal{S} \neq \mathcal{O}$, entonces el Teorema 3 del § 15 demuestra que $\mathcal{S}^0 \neq \mathcal{V}'$.

TEOREMA 1. *Si $\mathfrak{M}$ es un subespacio m-dimensional de un espacio vectorial $\mathcal{V}$, n-dimensional, entonces $\mathfrak{M}^0$ es un espacio (n-m)-dimensional de $\mathcal{V}'$.*

PRUEBA. Dejamos a cargo del lector verificar que $\mathfrak{M}^0$ (de hecho, $\mathcal{S}^0$ para una $\mathcal{S}$ arbitraria) es siempre un subespacio; probaremos solamente el enunciado concerniente a la dimensión de $\mathfrak{M}^0$.

Sea $\mathfrak{X} = \{x_1, \ldots, x_n\}$ una base en $\mathcal{V}$ cuyos primeros m elementos están en $\mathfrak{M}$ (y forman, en consecuencia, una base para $\mathfrak{M}$); sea $\mathfrak{X}' = \{y_1, \ldots, y_n\}$ la base dual en $\mathcal{V}'$. Denotamos por $\mathfrak{N}$ el subespacio (en $\mathcal{V}'$) sobretendido por $y_{m+1}, \ldots, y_n$; claramente $\mathfrak{N}$ tiene dimensión n-m. Probaremos que $\mathfrak{M}^0 = \mathfrak{N}$.

Si x es cualquier vector de $\mathfrak{M}$, entonces x es una combinacion lineal de $x_1, \ldots, x_m$,

$$x = \sum_{i=1}^{m} \xi_i x_i,$$

y, para cualquier $j = m + 1, \ldots, n$, tenemos

$$[x, y_j] = \sum_{i=1}^{m} \xi_i[x_i, y_j] = 0.$$

En otras palabras, y_j está en $\mathfrak{M}^0$ para $j = m + 1, \ldots, n$; se sigue que $\mathfrak{N}$ está contenido en $\mathfrak{M}^0$,

$$\mathfrak{N} \subset \mathfrak{M}^0.$$

Supóngase, por otra parte, que y es cualquier elemento de $\mathfrak{M}^0$. Puesto que y, estando en $\mathfrak{V}'$, es una combinación lineal de los vectores de la base $y_1, \ldots, y_n$, podemos escribir

$$y = \sum_{j=1}^{n} \eta_j y_j.$$

Puesto que, por hipótesis, y está en $\mathfrak{M}^0$, tenemos, para cada $i = 1, \ldots, m$

$$0 = [x_i, y] = \sum_{j=1}^{n} \eta_j[x_i, y_j] = \eta_i;$$

en otras palabras, y es una combinación lineal de $y_{m+1}, \ldots, y_n$. Esto prueba que y está en $\mathfrak{N}$, y, consecuentemente, que

$$\mathfrak{M}^0 \subset \mathfrak{N},$$

y el teorema se sigue.

Teorema 2. *Si $\mathfrak{M}$ es un subespacio en un espacio vectorial finito-dimensional $\mathfrak{V}$, entonces $\mathfrak{M}^{00}$ $(=(\mathfrak{M}^0)^0) = \mathfrak{M}$.*

Prueba. Obsérvese que usamos aquí la convención, establecida al fin del § 16, que identifica $\mathfrak{V}$ y $\mathfrak{V}''$. Por definición, $\mathfrak{M}^{00}$ es el conjunto de todos los vectores x tales que $[x, y] = 0$, para toda y de $\mathfrak{M}^0$. Puesto que, por definición de $\mathfrak{M}^0$, $[x, y] = 0$ para toda x de $\mathfrak{M}$ y toda y de $\mathfrak{M}^0$ se sigue que $\mathfrak{M} \subset \mathfrak{M}^{00}$. La conclusión deseada se sigue ahora de un argumento dimensional. Sea $\mathfrak{M}$ m-dimensional; entonces la dimensión de $\mathfrak{M}^0$ es $n - m$, y la de $\mathfrak{M}^{00}$ es $n - (n - m) = m$. De aquí que $\mathfrak{M} = \mathfrak{M}^{00}$ como se quería probar.

EJERCICIOS

1. Defínase una funcional lineal no cero y, sobre $\mathfrak{C}^3$ tal que si $x_1 = (1, 1, 1)$ y $x_2 = (1, 1, -1)$, entonces $[x_1, y] = [x_2, y] = 0$.

2. Los vectores $x_1 = (1, 1, 1)$, $x_2 = (1, 1, -)$ y $x_3 = (1, -1, -1)$ forman una base de $\mathfrak{C}^3$ Si $\{y_1, y_2, y_3\}$ es la base dual, y si $x = (0, 1, 0)$, encuentre $[x, y_1]$, $[x, y_2]$ y $[x, y_3]$.

3. Pruébese que si y es una funcional lineal sobre un espacio vectorial n-dimensional $\mathfrak{V}$, entonces el conjunto de todos los vectores x para los cuales

$[x, y] = 0$ es un subespacio de $\mathcal{V}$; ¿cuál es la dimensión de ese subespacio?

4. Si $y(x) = \xi_1 + \xi_2 + \xi_3$, cada vez que $x = (\xi_1, \xi_2, \xi_3)$ es un vector de $\mathcal{C}^3$ entonces y es una funcional lineal sobre $\mathcal{C}^3$ encuéntrese una base del subespacio consistente en todos los vectores x, para los cuales $[x, y] = 0$.

5. Pruébese que si $m < n$, y si $y_1, \ldots, y_m$ son funcionales lineales sobre un espacio vectorial n-dimensional $\mathcal{V}$, entonces existe allí un vector no cero x de $\mathcal{V}$ tal que $[x, y_j] = 0$ para $j = 1, \ldots, m$. ¿Qué dice este resultado sobre las soluciones de las ecuaciones lineales?

6. Supóngase que $m < n$ y que $y_1, \ldots, y_m$ son funcionales lineales sobre un espacio vectorial n-dimensional $\mathcal{V}$. ¿Bajo qué condiciones impuestas a los escalares $\alpha_1, \ldots, \alpha_m$ es verdad que existe un vector x de $\mathcal{V}$ tal que $[x, y_j] = \alpha_j$ para $j = 1, \ldots, m$? ¿Qué dice este resultado sobre las soluciones de las ecuaciones lineales?

7. Si $\mathcal{V}$ es un espacio vectorial n-dimensional sobre un campo finito, y si $0 \leq m \leq n$, entonces el número de subespacios m-dimensionales de $\mathcal{V}$ es igual al número de subespacio $(n - m)$-dimensionales.

8. (a) Pruébese que si $\mathcal{S}$ es un subconjunto de un espacio vectorial finito-dimensional, entonces $\mathcal{S}^{00}$ coincide con el subespacio sobretendido por $\mathcal{S}$.

(b) Si $\mathcal{S}$ y $\mathcal{T}$ son subconjuntos de un espacio vectorial, si y $\mathcal{S} \subset \mathcal{T}$, entonces $\mathcal{T}^0 \subset \mathcal{S}^0$.

(c) Si $\mathfrak{M}$ y $\mathfrak{N}$ son subespacios de un espacio vectorial finito-dimensional, entonces$(\mathfrak{M} \cap \mathfrak{N})^0 = \mathfrak{M}^0 + \mathfrak{N}^0$ y $(\mathfrak{M} + \mathfrak{N})^0 = \mathfrak{M}^0 \cap \mathfrak{N}^0$. (Sugestión: hágase uso repetido de (b) y del §17, Teorema 2).

(d) ¿Es válida la conclusión de (c) para espacios vectoriales no necesariamente finito-dimensionales?

9. Este ejercicio se ocupa de espacios vectoriales que no necesitan ser finito-dimensionales; la mayoría de sus partes (aunque no todas), dependen de la clase de razonamiento transfinito que se necesita para probar que todo espacio vectorial tiene una base (véase el § 7, Ej. 11).

(a) Supóngase que f y g son funcionales valuadas escalarmente, definidas sobre un conjunto $\mathfrak{X}$; si α y β son escalares, escriba $h = \alpha f + g$ para la función definida por $h(x) = \alpha f(x) + \beta g(x)$ para toda x de $\mathfrak{X}$. El conjunto de todas esas funciones es un espacio vectorial con respecto de esta definición de las operaciones lineales, y lo mismo es verdadero de todas las funciones finitamente no cero. (Una función f sobre $\mathfrak{X}$ es *finitamente no cero* si es finito el conjunto de los elementos x de $\mathfrak{X}$, para los cuales $f(x) \neq 0$.

(b) Todo espacio vectorial es isomórfico al conjunto de todas las funciones finitamente no cero sobre algún conjunto.

(c) Si $\mathcal{V}$ es un espacio vectorial con base $\mathfrak{X}$ y si f es una función valuada escalarmente, definida sobre el conjunto $\mathfrak{X}$, entonces existe una funcional lineal única y sobre $\mathcal{V}$ tal que $[x, y] = f(x)$ para toda x de $\mathfrak{X}$.

(d) Usese (a), (b) y (c) para concluír que todo espacio vectorial $\mathcal{V}$ es isomórfico a un subespacio de $\mathcal{V}'$.

(e) ¿Qué espacios vectoriales son isomórficos a sus propios duales?

(f) Si $\mathcal{Y}$ es un subconjunto linealmente independiente de un espacio vectorial $\mathcal{V}$, entonces existe una base de $\mathcal{V}$ que contiene a $\mathcal{Y}$. (Compárese este resultado con el Teorema del § 7).

(g) Si $\mathfrak{X}$ es un conjunto y si y es un elemento de $\mathfrak{X}$, escriba f_y para la función valuada escalarmente, definida sobre $\mathfrak{X}$ escribiendo $f_y(x) = 1$ o 0, según sea $x = y$ o $x \neq y$. Sea $\mathcal{Y}$ el conjunto de todas las funciones f_y, juntamente con la función g, definida por $g(x) = 1$ para toda x de $\mathfrak{X}$. Pruébese que si $\mathfrak{X}$ es infinita, entonces $\mathcal{Y}$ es un subconjunto linealmente independiente del espacio vectorial de todas las funciones valuadas escalarmente sobre $\mathfrak{X}$.

(h) La correspondencia natural de $\mathcal{V}$ a $\mathcal{V}''$ se define para todos los espacios vectoriales (no solamente para los finito-dimensionales); si x_0 está en $\mathcal{V}$, defínase el correspondiente elemento z_0 de $\mathcal{V}''$ escribiendo $z_0(y) = [x_0, y]$ para toda y de $\mathcal{V}'$. Pruébese que si $\mathcal{V}$ es reflexiva (esto es, si toda z de $\mathcal{V}''$ puede obtenerse en esta forma por una adecuada selección de x_0), entonces $\mathcal{V}$ es finito-dimensional. (Sugestión: represéntense $\mathcal{V}'$ como el conjunto de todas las funcionales valuadas escalarmente sobre algún conjunto, y úsese (g), (f) y (c) para construir un elemento de $\mathcal{V}''$ que no esté inducido por ningún elemento de $\mathcal{V}$.)

Advertencia: La aserción de que un espacio vectorial es reflexivo si y sólo si es finito-dimensional, resultaría chocante para la mayor parte de los expertos en la materia. La razón es que la acostumbrada y fructífera generalización del concepto de reflexividad a los espacios finito-dimensionales no es la sencilla dada en (h).

§ 18. SUMAS DIRECTAS

Estudiaremos varios importantes métodos generales de convertir en nuevos viejos espacios vectoriales; en esta sección comenzamos por estudiar el más fácil.

DEFINICIÓN. Si $\mathcal{U}$ y $\mathcal{V}$ son espacios vectoriales (sobre el mismo campo), su *suma directa* es el espacio vectorial $\mathcal{W}$ (denotado por $\mathcal{U} \oplus \mathcal{V}$), cuyos elementos son todos de pares ordenados $\langle x, y\rangle$ con x en $\mathcal{U}$ y y en $\mathcal{V}$, con las operaciones lineales definidas por

$$\alpha_1\langle x_1, y_1\rangle + \alpha_2\langle x_2, y_2\rangle = \langle \alpha_1 x_1 + \alpha_2 x_2, \alpha_1 y_1 + \alpha_2 y_2\rangle.$$

Observamos que la formación de la suma directa es análoga al modo en que el plano es construido desde sus dos ejes de coordenadas.

Procedemos a investigar la relación de esta noción con algunas de nuestras nociones.

El conjunto de todos los vectores (en $\mathcal{W}$) de la forma $\langle x, 0\rangle$ es un subespacio de $\mathcal{W}$; la correspondencia $\langle x, 0\rangle \rightleftarrows x$ demuestra que este subespacio es isomórfico a $\mathcal{U}$. Es conveniente, una vez más, caer en una inexactitud lógica e, identificando x y $\langle x, 0\rangle$, hablar de $\mathcal{U}$ como un subespacio de $\mathcal{W}$. De modo semejante, por supuesto, los vectores y de $\mathcal{V}$ pueden ser identificados con los vectores de la forma

$\langle 0, y\rangle$ de $\mathcal{W}$ y podemos considerar $\mathcal{V}$ como un subespacio de $\mathcal{W}$. Esta terminología es, de seguro, no del todo exacta, pero es mucho más fácil encontrar aquí una salida que en el caso del segundo espacio dual. Podríamos haber definido la suma directa de $\mathcal{U}$ y $\mathcal{V}$ (cuando menos en el caso en que $\mathcal{U}$ y $\mathcal{V}$ no tienen vectores no-cero en común), como el conjunto consistente en todas las x de $\mathcal{U}$, todas las y de $\mathcal{V}$, y todos los pares $\langle x, y\rangle$ para los cuales $x \neq 0$ y $y \neq 0$. Esta definición produce una teoría análoga en cada detalle a la que desarrollaremos, pero hace difícil probar a causa de las distinciones de casos que necesita. Está claro, sin embargo, que desde el punto de vista de esta definición $\mathcal{U}$ es de hecho un subconjunto de $\mathcal{U} \oplus \mathcal{V}$. En este sentido entonces, o en el sentido del isomorfismo de la definición que hemos adoptado, suscitamos la cuestión: ¿cuál es la relación entre $\mathcal{U}$ y $\mathcal{V}$ cuando consideramos estos espacios y subespacios del gran espacio $\mathcal{W}$?

Teorema. *Si $\mathcal{U}$ y $\mathcal{V}$ son subespacios de un espacio vectorial $\mathcal{W}$, entonces son equivalentes las tres condiciones siguientes.*

(1) $\mathcal{W} = \mathcal{U} \oplus \mathcal{V}$.

(2) $\mathcal{U} \cap \mathcal{V} = \mathcal{O}$ *y* $\mathcal{U} + \mathcal{V} = \mathcal{W}$ *(esto es, $\mathcal{U}$ y $\mathcal{V}$ son complementos uno de otro).*

(3) *Todo vector z de $\mathcal{W}$ puede ser escrito en la forma $z = x + y$, con x en $\mathcal{U}$ y y en $\mathcal{V}$, de un modo y sólo de uno.*

Prueba. Probaremos las implicaciones (1) $\Rightarrow$ (2) $\Rightarrow$ (3) $\Rightarrow$ (1).

(1) $\Rightarrow$ (2). Suponemos que $\mathcal{W} = \mathcal{U} \oplus \mathcal{V}$. Si $z = \langle x, y\rangle$ queda en ambos $\mathcal{U}$ y $\mathcal{V}$, entonces $x = y = 0$, de manera que $z = 0$; esto prueba que $\mathcal{U} \cap \mathcal{V} = \mathcal{O}$. Puesto que la representación $z = \langle x, 0\rangle + \langle 0, y\rangle$ es válida para toda z, se sigue también que $\mathcal{U} + \mathcal{V} = \mathcal{W}$.

(2) $\Rightarrow$ (3). Si suponemos (2), de manera que, en particular, $\mathcal{U} + \mathcal{V} = \mathcal{W}$, entonces está claro que toda z de $\mathcal{W}$ tiene la representación deseada, $z = x + y$. Para probar la unicidad, suponemos que $z = x_1 + y_1$ y $z = x_2 + y_2$, con x_1 y x_2 en $\mathcal{U}$ y y_1 y y_2 en $\mathcal{V}$. Puesto que $x_1 + y_1 = x_2 + y_2$, se sigue que $x_1 - x_2 = y_2 - y_1$. Puesto que el segundo miembro de esta última ecuación está en $\mathcal{U}$ y el primer miembro está en $\mathcal{V}$, disyunción de $\mathcal{U}$ y de $\mathcal{V}$ implica que $x_1 = x_2$ y $y_1 = y_2$.

(3) $\Rightarrow$ (1). Esta implicación es prácticamente indistinguible de la definición de suma directa. Si formamos la suma directa $\mathcal{U} \oplus \mathcal{V}$, y a continuación identificamos $\langle x, 0\rangle$ y $\langle 0, y\rangle$ con x y y respectivamente, estamos comprometidos a identificar la suma $\langle x, y\rangle = \langle x, 0\rangle + \langle 0, y\rangle$ con lo que estamos suponiendo que es el elemento general $z = x + y$ de $\mathcal{W}$; de la hipótesis de que la representación

de z en la forma $x + y$ es única, concluimos que la correspondencia entre $\langle x, 0\rangle$ y x (y también entre $\langle 0, y\rangle$ y y) es uno a uno.

Si dos subespacios $\mathcal{U}$ y $\mathcal{V}$ en un espacio vectorial $\mathcal{W}$ están disjuntos y sobretienden $\mathcal{W}$ (esto es, satisfacen (2), es usual decir que $\mathcal{W}$ es la *suma interna directa* de $\mathcal{U}$ y $\mathcal{V}$; simbólicamente, como antes, $\mathcal{W} = \mathcal{U} \oplus \mathcal{V}$. Si deseamos hacer resaltar la distinción entre estos conceptos y el definido antes, escribiremos el primero diciendo que $\mathcal{W}$ es *la suma externa directa* de $\mathcal{U}$ y $\mathcal{V}$. En vista de los isomorfismos naturales descritos antes y, especialmente, en vista del teorema precedente, la distinción es más pedante que conceptual. De acuerdo con nuestra convención de identificación, generalmente la pasaremos por alto.

§ 19. DIMENSION DE UNA SUMA DIRECTA

¿Qué puede decirse sobre la dimensión de una suma directa? Si $\mathcal{U}$ es n-dimensional, $\mathcal{V}$ es m-dimensional y $\mathcal{W} = \mathcal{U} \oplus \mathcal{V}$, ¿cuál es la dimensión de $\mathcal{W}$? Esta cuestión es fácil de contestar.

TEOREMA 1. *La dimensión de una suma directa es la suma de las dimensiones de sus sumandos.*

PRUEBA. Afirmamos que si $\{x_1, \ldots, x_n\}$ es una base de $\mathcal{U}$, y si $\{y_1, \ldots, y_m\}$ es una base en $\mathcal{V}$, entonces el conjunto $\{x_1, \ldots, x_n, y_1, \ldots, y_m\}$ (o, más precisamente, el conjunto $\{\langle x_1, 0\rangle, \ldots, \langle x_n, 0\rangle, \langle 0, y_1\rangle, \ldots, \langle 0, y_m\rangle\}$ es una base de $\mathcal{W}$. La prueba más fácil de esta aserción consiste en usar la implicación (1) $\Rightarrow$ (3) del teorema de la sección precedente. Puesto que toda z de $\mathcal{W}$ puede escribirse en la forma $z = x + y$, donde x es una combinación lineal de $x_1 \ldots, x_n$ y y es una combinación lineal de $y_1 \ldots, y_m$, se sigue que nuestro conjunto sobretiende a $\mathcal{W}$. Para demostrar que el conjunto es también linealmente independiente, supóngase que

$$\alpha_1 x_1 + \cdots + \alpha_n x_n + \beta_1 y_1 + \cdots + \beta_m y_m = 0.$$

La unicidad de la representación de 0 en la forma $x + y$ implica que

$$\alpha_1 x_1 + \cdots + \alpha_n x_n = \beta_1 y_1 + \cdots + \beta_m y_m = 0,$$

y de aquí que la dependencia lineal de las x y las y implique que

$$\alpha_1 = \cdots = \alpha_n = \beta_1 = \cdots = \beta_m = 0.$$

TEOREMA 2. *Si $\mathcal{W}$ es cualquier espacio vectorial $(n + m)$-dimensional, y si $\mathcal{U}$ es cualquier subeqpacio n-dimensional de $\mathcal{W}$, existe entonces un subespacio m-dimensional $\mathcal{V}$ de $\mathcal{W}$ tal que $\mathcal{W} = \mathcal{U} \oplus \mathcal{V}$.*

PRUEBA. Sea $\{x_1, \ldots, x_n\}$ cualquier base de $\mathcal{U}$; por el teorema del § 7, podemos encontrar un conjunto $\{y_1, \ldots, y_m\}$ de vectores de $\mathcal{W}$ con la propiedad de que $\{x_1, \ldots, x_n, y_1, \ldots, y_m\}$ es una base de $\mathcal{W}$. Sea $\mathcal{V}$ el subespacio sobretendido por $y_1, \ldots, y_m$; omitimos la verificación de que $\mathcal{W} = \mathcal{U} \oplus \mathcal{V}$.

El Teorema 2 dice que todo subespacio de un espacio vectorial finito-dimensional tiene un complemento.

§ 20. DUAL DE UNA SUMA DIRECTA

En la mayor parte de lo que sigue veremos la noción de suma directa, como se define para subespacios de un espacio vectorial $\mathcal{V}$; esto evita la confusión con la convención de identificación del § 18 y resulta, incidentalmente, ser el concepto más útil para nuestro trabajo posterior. Concluimos, por ahora, nuestro estudio de sumas directas, observando la relación simple que conecta los espacios duales, aniquiladores y sumas directas. Para hacer resaltar nuestro actual punto de vista de suma directa, retornamos a las letras de nuestra primera notación.

TEOREMA. *Si $\mathfrak{M}$ y $\mathfrak{N}$ son subespacios de un espacio vectorial $\mathcal{V}$, y si $\mathcal{V} = \mathfrak{M} \oplus \mathfrak{N}$, entonces $\mathfrak{M}'$ es isomórfico a $\mathfrak{N}^0$ y $\mathfrak{N}'$ a $\mathfrak{M}^0$, y $\mathcal{V}' = \mathfrak{M}^0 \oplus \mathfrak{N}^0$.*

PRUEBA. Para simplificar la notación, usaremos, en toda esta prueba, x, x' y x^0 para elementos de $\mathfrak{M}$, $\mathfrak{M}'$, y $\mathfrak{M}^0$, respectivamente y reservamos, de modo semejante, las letras y para $\mathfrak{N}$ y z para $\mathcal{V}$. (Esta notación no tiene el propósito de sugerir que exista alguna relación particular entre, digamos, los vectores x de $\mathfrak{M}$ y los vectores x' de $\mathfrak{M}'$).

Si z' pertenece tanto a $\mathfrak{M}^0$ como a $\mathfrak{N}^0$, esto es, si $z'(x) = z'(y) = 0$ para toda x y y, entonces $z'(z) = z'(x + y) = 0$ para toda z; esto implica que $\mathfrak{M}^0$ y $\mathfrak{N}^0$ están disjuntos. Si, además, z' es un vector cualquiera de $\mathcal{V}'$, y si $z = x + y$, escribimos $x^0(z) = z'(y)$ y $y^0(z) = z'(x)$. Es fácil ver que las funciones x^0 y y^0 así definidas son funciones lineales sobre $\mathcal{V}$ (esto es, elementos de $\mathcal{V}'$) que pertenecen a $\mathfrak{M}^0$ y $\mathfrak{N}^0$ respectivamente; puesto que $z' = x^0 + y^0$, se sigue que $\mathcal{V}'$ es en realidad la suma directa de $\mathfrak{M}^0$ y $\mathfrak{N}^0$.

Para establecer los isomorfismos afirmados, hacemos corresponder a cada x^0 una y' de $\mathfrak{N}'$ definida por $y'(y) = x^0(y)$. Dejamos al lector la verificación de rutina de que la correspondencia $x^0 \to y'$ es lineal y uno a uno y, en consecuencia, un isomorfismo entre $\mathfrak{M}^0$

y $\mathfrak{N}'$; el resultado correspondiente para $\mathfrak{N}^0$ y $\mathfrak{M}'$ se sigue de la simetría, intercambiando x y y. (Obsérvese que, para espacios vectoriales finito-dimensionales, la mera existencia de un isomorfismo entre, digamos, $\mathfrak{M}^0$ y $\mathfrak{N}'$ es trivial desde el punto de vista de un argumento dimensional; las dimensiones de $\mathfrak{M}^0$ y de $\mathfrak{N}'$ son iguales a la dimensión de $\mathfrak{N}$).

Observamos, con respecto de toda nuestra presentación de la teoría de sumas directas, que no hay nada mágico sobre el número dos; podríamos haber definido la suma directa de cualquier número de espacios vectoriales, y podríamos haber probado los obvios análogos de todos los teoremas de las últimas tres secciones, con sólo que la notación se volviera más complicada. Hacemos la advertencia de que usaremos posteriormente esta observación y trataremos los problemas que implica como si los hubiéramos probado.

EJERCICIOS

1. Supóngase que x, y, u y v son vectores de $\mathbb{C}^4$; sean $\mathfrak{M}$ y $\mathfrak{N}$ los subespacios de $\mathbb{C}^4$ supertendidos por $\{x, y\}$ y $\{u, v\}$ respectivamente. ¿En cuál de los siguientes casos es verdadero que $\mathbb{C}^4 = \mathfrak{M} \oplus \mathfrak{N}$?

(a) $x = (1, 1, 0, 0)$, $y = (1, 0, 1, 0)$
$u = (0, 1, 0, 1)$, $v = (0, 0, 1, 1)$.
(b) $x = (-1, 1, 1, 0)$, $y = (0, 1, -1, 1)$
$u = (1, 0, 0, 0)$, $v = (0, 0, 0, 1)$.
(c) $x = (1, 0, 0, 1)$, $y = (0, 1, 1, 0)$
$u = (1, 0, 1, 0)$, $v = (0, 1, 0, 1)$.

2. Si $\mathfrak{M}$ es el subespacio consistente en los vectores $(\xi_1, \ldots, \xi_n, \xi_{n+1}, \ldots, \xi_{2n})$ de $\mathbb{C}^{2n}$ para los cuales $\xi_1 = \ldots = \xi_n = 0$, y si $\mathfrak{N}$ es el subespacio de todos los vectores para los cuales $\xi_j = \xi_{n+j}$, $j = 1, \ldots, n$, entonces $\mathbb{C}^{2n} = \mathfrak{M} \oplus \mathfrak{N}$.

3. Constrúyanse tres subespacios $\mathfrak{M}$, $\mathfrak{N}_1$, y $\mathfrak{N}_2$ de un espacio vectorial $\mathfrak{V}$ de manera que $\mathfrak{M} \oplus \mathfrak{N}_1 = \mathfrak{M} \oplus \mathfrak{N}_2 = \mathfrak{V}$ pero $\mathfrak{N}_1 \neq \mathfrak{N}_2$. (Nótese que esto significa que no hay ley de cancelación para las sumas directas). ¿Cuál es la representación geométrica que responde a esta situación?

4. (a) Si $\mathfrak{U}$, $\mathfrak{V}$, y $\mathfrak{W}$ son espacios vectoriales, ¿cuál es la relación entre $\mathfrak{U} \oplus (\mathfrak{V} \oplus \mathfrak{W})$ y $(\mathfrak{U} \oplus \mathfrak{V}) \oplus \mathfrak{W}$ (esto es, ¿en qué sentido es la formación de sumas directas una operación asociativa?

(b) ¿En qué sentido es la formación de sumas directas conmutativa?

5. (a) Tres subespacios $\mathfrak{L}$, $\mathfrak{M}$, y $\mathfrak{N}$ de un espacio vectorial, $\mathfrak{V}$ son llamados *independientes* si cada uno es disjunto de la suma de los otros dos. Pruébese que una condición necesaria y suficiente para $\mathfrak{V} = \mathfrak{L} \oplus (\mathfrak{M} \oplus \mathfrak{N})$ (y también para $\mathfrak{V} = (\mathfrak{L} \oplus \mathfrak{M}) \oplus \mathfrak{N}$) es que $\mathfrak{L}$, $\mathfrak{M}$, y $\mathfrak{N}$ sean independientes y que $\mathfrak{V} = \mathfrak{L} + \mathfrak{M} + \mathfrak{N}$. (El subespacio $\mathfrak{L} + \mathfrak{M} + \mathfrak{N}$ es el conjunto de todos los

vectores de la forma $x + y + z$, con x en $\mathcal{L}$, y en $\mathfrak{M}$, y z en $\mathfrak{N}$).

(b) Dé un ejemplo de tres subespacios de un espacio vectorial $\mathcal{V}$ tal que la suma de los tres es $\mathcal{V}$, tal que cada dos de los tres sea disjunto, pero tal que los tres no sean independientes.

(c) Supóngase que x, y y z sean elementos de un espacio vectorial y que $\mathcal{L}$, $\mathfrak{M}$ y $\mathfrak{N}$ sean los espacios sobretendidos por x, y, y z, respectivamente. Pruébese que los vectores x, y, y z son linealmente independientes si y sólo si los subespacios $\mathcal{L}$, $\mathfrak{M}$ y $\mathfrak{N}$ son independientes.

(d) Pruébese que tres subespacios finito-dimensionales son independientes si y sólo si la suma de sus dimensiones es igual a la dimensión de su suma.

(e) Generalice los resultados (a)-(d) de los tres subespacios a cualquier número finito.

§ 21. ESPACIOS COCIENTES

Ya sabemos que si $\mathfrak{M}$ es un subespacio de un espacio vectorial $\mathcal{V}$, entonces hay, usalmente, muchos otros subespacios $\mathfrak{N}$ en $\mathcal{V}$ tales que $\mathfrak{M} \oplus \mathfrak{N} = \mathcal{V}$. No hay un modo natural de escoger uno de entre la riqueza de complementos de $\mathfrak{M}$. Hay, sin embargo, una construcción natural que asocia a $\mathfrak{M}$ y $\mathcal{V}$ un nuevo espacio vectorial que, para todos los propósitos prácticos, juega el papel de un complemento de $\mathfrak{M}$. La ventaja teórica que tiene la construcción sobre la formación de un complemento arbitrario es precisamente su carácter "natural", esto es, el hecho de que no depende de la selección de una base, o, para el caso, de la selección de cualquier otra cosa.

A fin de comprender la construcción, es una buena idea tener una imagen en la mente. Supóngase, por ejemplo, que $\mathcal{V} = \mathcal{R}^2$ (el plano de coordenadas reales) y que $\mathfrak{M}$ consiste en todos los vectores (ξ_1, ξ_2) para los cuales $\xi_2 = 0$ (el eje horizontal). Cada complemento de $\mathfrak{M}$ es una línea (distinta del eje horizontal), que pasa por el origen. Obsérvese que cada uno de esos complementos tiene la propiedad de que intersecta cada línea horizontal exactamente en un punto. La idea de la construcción que describiremos es la de hacer un espacio vectorial, utilizando el conjunto de todas las líneas horizontales.

Comenzamos por usar $\mathfrak{M}$ para singularizar ciertos subconjuntos de $\mathcal{V}$. (Estamos de nuevo en el caso general. Si x es un vector arbitrario en $\mathcal{V}$, escribimos $x + \mathfrak{M}$ para el conjunto de todas las sumas $x + y$, con y en $\mathfrak{M}$; cada conjunto de la forma $x + \mathfrak{M}$ se llama *coconjunto* de $\mathfrak{M}$. (En el caso del ejemplo anterior de líneas planas, los coconjuntos son las líneas horizontales). Nótese que uno y el mismo coconjunto se puede suscitar de dos vectores diferentes, esto es, que aun si $x \neq y$, es posible que $x + \mathcal{M} = y + \mathcal{M}$. Tiene el mismo sentido hablar de un coconjunto, digamos $\mathcal{H}$, de $\mathcal{M}$, sin especificar de qué elemento (o elementos) viene $\mathcal{H}$, decir que $\mathcal{H}$ es un coconjunto (de $\mathcal{M}$) significa simplemente que hay cuando menos una x tal que $\mathcal{H} = x + \mathcal{M}$.

Si $\mathcal{H}$ y $\mathcal{K}$ son coconjuntos (de $\mathcal{M}$), escribimos $\mathcal{H} + \mathcal{K}$ para el conjunto de todas las sumas $u + v$ con u en $\mathcal{H}$ y v en K; afirmamos que $\mathcal{H} + \mathcal{K}$ es también un coconjunto de $\mathcal{M}$. En realidad, si $\mathcal{H} = x + \mathcal{M}$ y $\mathcal{K} = y + \mathcal{M}$, entonces todo elemento de $\mathcal{H} + \mathcal{K}$ pertenece al coconjunto $(x + y) + \mathcal{M}$ (nótese que $\mathcal{M} + \mathcal{M} = \mathcal{M}$), y, recíprocamente, todo elemento de $(x + y) + \mathcal{M}$ está en $\mathcal{H} + \mathcal{K}$. (Si, por ejemplo, z está en $\mathcal{M}$, entonces $(x + y) + z = (x + z) + (y + 0)$). En otras palabras, $\mathcal{H} + \mathcal{K} = (x + y) + \mathcal{M}$, de manera que $\mathcal{H} + \mathcal{K}$ es un coconjunto, como se afirmó. Dejamos al lector la verificación de que la adición de coconjuntos es conmutativa y asociativa. El coconjunto $\mathcal{M}$ (esto es, $0 + \mathcal{M}$) es tal que $\mathcal{H} + \mathcal{M} = \mathcal{H}$ para todo coconjunto $\mathcal{H}$, y, además $\mathcal{M}$ es el único coconjunto con esa propiedad. (Si $(x + \mathcal{M}) + (y + \mathcal{M}) = x + \mathcal{M}$, entonces $x + \mathcal{M}$ contiene $x + y$, de manera que $x + y = x + u$, para alguna u de $\mathcal{M}$; esto implica que y está en $\mathcal{M}$, y de aquí que $y + \mathcal{M} = \mathcal{M}$.)

Si $\mathcal{H}$ es un conjunto, entonces el conjunto consistente en todos los vectores $-u$, con u en $\mathcal{H}$, es él mismo coconjunto, que denotaremos por $-\mathcal{H}$. El coconjunto $-\mathcal{H}$ es tal que $\mathcal{H} + (-\mathcal{H}) = \mathcal{M}$, y, además, $-\mathcal{H}$ es el único coconjunto con esta propiedad. Resumiendo; la adición de coconjuntos satisface los axiomas (**A**) del § 2.

Si $\mathcal{H}$ es un coconjunto y si α es un escalar, escribimos $\alpha\mathcal{H}$ para el conjunto consistente en todos los vectores αu, con u en $\mathcal{H}$ en caso de que $\alpha \neq 0$; el coconjunto $0 \cdot \mathcal{H}$ se definió como $\mathcal{M}$. Una simple verificación demuestra que este concepto de multiplicación satisface los axiomas (**B**) y (**C**) del § 2.

Ha sido así probado que el conjunto de todos los coconjuntos es un espacio vectorial con respecto de las operaciones lineales definidas anteriormente. Este espacio vectorial se llama el *espacio cociente* de $\mathcal{V}$ módulo $\mathcal{M}$; es denotado por $\mathcal{V}/\mathcal{M}$.

§ 22. DIMENSION DE UN ESPACIO COCIENTE

TEOREMA 1. *Si $\mathcal{M}$ y $\mathcal{N}$ son subespacios complementarios de un espacio vectorial $\mathcal{V}$, entonces, la correspondencia que asigna a cada vector y de $\mathcal{N}$ el coconjunto $y + \mathcal{M}$ es un isomorfismo entre $\mathcal{N}$ y $\mathcal{V}/\mathcal{M}$.*

PRUEBA. Si y_1 y y_2 son elementos de $\mathfrak{N}$ tales que $y_1 + \mathfrak{M} = y_2 + \mathfrak{M}$, entonces, en particular, y_1 pertenece a $y_2 + \mathfrak{M}$, de manera que $y_1 = y_2 + x$ para alguna x de $\mathfrak{M}$. Puesto que esto significa que $y_1 - y_2 = x$, y puesto que $\mathfrak{M}$ y $\mathfrak{N}$ están disjuntos, se sigue que $x = 0$, y de aquí que $y_1 = y_2$. (Recuérdese que $y_1 - y_2$ pertenece a $\mathfrak{N}$ juntamente con y_1 y y_2). Este argumento prueba que la correspondencia que estamos estudiando es uno a uno, hasta ahora. Para probar que va suficientemente lejos, considérese un coconjunto arbitrario de $\mathfrak{M}$, digamos $z + \mathfrak{M}$. Puesto que $\mathfrak{V} = \mathfrak{N} + \mathfrak{M}$, podemos escribir z en la forma $y + x$, con x en $\mathfrak{M}$ y y en $\mathfrak{N}$; se sigue (puesto que $x + \mathfrak{M} = \mathfrak{M}$) que $z + \mathfrak{M} = y + \mathfrak{M}$. Esto prueba que todo coconjunto de $\mathfrak{M}$ puede obtenerse usando un elemento de $\mathfrak{N}$ (y no solamente cualquier viejo elemento de $\mathfrak{V}$); consecuentemente $y \to y + \mathfrak{M}$ es realmente una correspondencia uno a uno entre $\mathfrak{N}$ y $\mathfrak{V}/\mathfrak{M}$. La propiedad lineal de la correspondencia se sigue inmediatamente de la definición de las operaciones lineales en $\mathfrak{V}/\mathfrak{M}$; en realidad, tenemos,

$$(\alpha_1 y_1 + \alpha_2 y_2) + \mathfrak{M} = \alpha_1(y_1 + \mathfrak{M}) + \alpha_2(y_2 + \mathfrak{M}).$$

TEOREMA 2. *Si $\mathfrak{M}$ es un subespacio m-dimensional de un espacio vectorial n-dimensional $\mathfrak{V}$, entonces $\mathfrak{V}/\mathfrak{M}$ tiene dimensión $n - m$.*

PRUEBA. Usese el § 19, Teorema 2, para encontrar un subespacio $\mathfrak{N}$ de manera que $\mathfrak{M} \oplus \mathfrak{N} = \mathfrak{V}$. El espacio $\mathfrak{N}$ tiene dimensión n-m (según el § 19, Teorema 1) y es isomórfico a $\mathfrak{V}/\mathfrak{M}$ (según el Teorema 1 anterior).

Hay más tópicos en la teoría de espacios cociente que podríamos discutir (tales como su relación a los espacios duales y a los aniquiladores). No obstante, puesto que muchos tópicos son apenas algo más que ejercicios que implican el uso de técnicas que ya están a nuestra disposición, nos ocupamos mejor de algunos modos nuevos y no obvios de manufacturar útiles espacios vectoriales.

EJERCICIOS

1. Considérense los espacios cociente obtenidos reduciendo el espacio $\mathcal{P}$ de polinomios módulo varios subespacios. Si $\mathfrak{M} = \mathcal{P}_n$, ¿es $\mathcal{P}/\mathfrak{M}$ finito-dimensional? ¿Qué si $\mathfrak{M}$ es el subespacio consistente de todos los polinomios constantes? ¿Qué si $\mathfrak{M}$ es el subespacio consistente en todos los polinomios divisibles entre x_n (donde $x_n(t) = (t^n)$?

2. Si $\mathcal{S}$ y $\mathcal{T}$ son subconjuntos arbitrarios de un espacio vectorial (no necesariamente coconjuntos de un subespacio), no hay nada que nos impida

definir $S + \mathcal{J}$ justamente como la adición fue definida para coconjuntos y, similarmente podemos definir αS (donde α es un escalar). Si la clase de todos los subconjuntos de un espacio vectorial está dotada con estas "operaciones lineales", ¿cuál de los axiomas de un espacio vectorial se satisface?

3. (a) Supóngase que $\mathfrak{M}$ es un subespacio de un espacio vectorial $\mathcal{V}$. Los vectores x y y de $\mathcal{V}$ son *congruentes* módulo $\mathfrak{M}$, en símbolos, $x \equiv y$, $(\mathfrak{M})$. si $x - y$ está en $\mathfrak{M}$. Pruébese que la congruencia módulo $\mathfrak{M}$ es una *relación de equivalencia*, esto es, que es reflexiva ($x \equiv x$), simétrica (si $x \equiv y$, entonces $y \equiv x$), y transitiva (si $x \equiv y$ y $y \equiv z$, entonces $x \equiv z$).

(b) Si α_1 y α_2 son escalares, y si x_1, x_2, y_1 y y_2 son vectores tales que $x_1 \equiv y_1$ $(\mathfrak{M})$ y $x_2 \equiv y_2 (\mathfrak{M})$, entonces $\alpha_1 x_1 + \alpha_2 x_2 \equiv \alpha_1 y_1 + \alpha_2 y_2 (\mathfrak{M})$.

(c) La congruencia módulo $(\mathfrak{M})$ parte a $\mathcal{V}$ en clases de equivalencia, esto es, en conjuntos tales que dos vectores pertenecen al mismo conjunto si y sólo si son congruentes. Pruébese que el subconjunto de $\mathcal{V}$ es una clase de equivalencia módulo $\mathfrak{M}$ si y sólo si es un subconjunto de $\mathfrak{M}$.

4. (a) Supóngase que $\mathfrak{M}$ es un subespacio de un espacio vectorial $\mathcal{V}$. Correspondiendo a cada funcional lineal y sobre $\mathcal{V}/\mathfrak{M}$ (esto es, a cada elemento y de $(\mathcal{V}/\mathfrak{M})'$), hay una funcional lineal z sobre $|\mathcal{V}$ (esto es, un elemento de $\mathcal{V}'$); la funcional lineal z se define como $z(x) = y(x + \mathfrak{M})$. Pruébese que la correspodencia $y \rightarrow z$ es un isomorfismo entre $(\mathcal{V}/\mathfrak{M})'$ y $\mathfrak{M}^0$.

(b) Supóngase que $\mathfrak{M}$ es un subespacio de un espacio vectorial $\mathcal{V}$. Correspondiendo a cada coconjunto $y + \mathfrak{M}^0$ de $\mathfrak{M}^0$ en $\mathcal{V}'$ (esto es, a cada elemento $\mathcal{K}$ de $\mathcal{V}'/\mathfrak{M}^0$) hay una funcional| lineal z sobre $\mathfrak{M}$ (esto es, un elemento z de $\mathfrak{M}'$); la funcional lineal z se define como $z(x) = y(x)$. Pruébese que z está inambiguamente determinada por el coconjunto $\mathcal{K}$ (esto es, no depende de la selección particular de y) y que la correspondencia $\mathcal{K} \rightarrow z$ es un isomorfismo entre $\mathcal{V}'/\mathfrak{M}^0$ y $\mathfrak{M}'$.

5. Dado un espacio vectorial finito-dimensional $\mathcal{V}$, fórmese la suma directa $\mathcal{W} = \mathcal{V} \oplus \mathcal{V}'$, y pruébese que la correspondencia $\langle x, y\rangle \rightarrow \langle y, x\rangle$ es un isomorfismo entre W y W'.

§ 23. FORMAS BILINEALES

Si $\mathcal{U}$ y $\mathcal{V}$ son espacios vectoriales (sobre el mismo campo), entonces su suma directa $\mathcal{W} = \mathcal{U} \oplus \mathcal{V}$ es otro espacio vectorial: nos proponemos estudiar ciertas funciones sobre $\mathcal{W}$. (Para los propósitos actuales, la definición de $\mathcal{U} \oplus \mathcal{V}$, vía pares ordenados, es la conveniente). El valor de esa función, digamos w, en un elemento $< x, y >$ de $\mathcal{W}$ será denotado por $w(x, y)$. El estudio de las funciones lineales de $\mathcal{W}$ no tiene ya interés para nosotros; los principales hechos concernientes a las mismas fueron discutidos en el § 20. Las funciones que deseamos considerar ahora son las bilineales, son, por definición, las funciones valuadas escalarmente sobre $\mathcal{W}$ con la propiedad de que para cada valor fijo de un argumento dependen linealmente del otro argumento. Más precisamente, una función valuada escalarmente w sobre $\mathcal{W}$ es una *forma bilineal* (o *funcional bilineal*), si

$$w(\alpha_1 x_1 + \alpha_2 x_2, y) = \alpha_1 w(x_1, y) + \alpha_2 w(x_2, y)$$

y

$$w(x, \alpha_1 y_1 + \alpha_2 y_2) = \alpha_1 w(x, y_1) + \alpha_2 w(x, y_2),$$

idénticamente en los vectores y escalares implicados.

En una situación especial, ya hemos encontrado funcionales bilineales. Si, a saber, $\mathcal{V}$ es el espacio dual de $|\mathcal{U}, \mathcal{V} = \mathcal{U}',|$ y si volvemos a escribir $w(x, y) = [x, y]$ (véase el § 14), entonces w es una funcional bilineal de $\mathcal{U} \oplus \mathcal{U}'$. Para un ejemplo, en una situación más general, sean $\mathcal{U}$ y $\mathcal{V}$ espacios vectoriales arbitrarios (sobre el mismo campo, como siempre), sean u y v elementos de $\mathcal{U}'$ y $\mathcal{V}'$ respectivamente, y escribamos $w(x, y) = u(x)v(y)$ para toda x de $\mathcal{U}$ y y de $\mathcal{V}$. Se obtiene un ejemplo todavía más general escogiendo un número finito de elementos de $\mathcal{U}'$, digamos $u_1, \ldots, u_k$, escogiendo el mismo número de elementos finitos de $\mathcal{V}'$, digamos $v_1, \ldots, v_k$ y escribiendo $w(x, y) = u_1(x)v_1(y) + \ldots + u_k(x)v_k(y)$. Cuál de las palabras, "funcional" o "forma" se use depende un tanto del contexto y otro tanto más en el capricho del que lo aplique. En este libro usaremos generalmente "funcional" con "lineal" y "forma" con "bilineal" (y sus generalizaciones dimensionales superiores).

Si w_1 y w_2 son formas lineales sobre $\mathcal{W}$, y si α_1 y α_2 son escalares, escribimos w para la función sobre $\mathcal{W}$ definida por

$$w(x, y) = \alpha_1 w_1(x, y) + \alpha_2 w_2(x, y).$$

Es fácil ver que w es una forma bilineal; lo denotamos por medio $\alpha_1 w_1 + \alpha_2 w_2$. Con esta definición de las operaciones lineales, el conjunto de todas las formas bilineales sobre $\mathcal{W}$ es un espacio vectorial. El principal propósito del resto de esta sección es el determinar (en el caso finito-dimensional) cómo la dimensión de este espacio depende de las dimensiones de $\mathcal{U}$ y de

Teorema 1. *Si $\mathcal{U}$ es un espacio vectorial n-dimensional con base $\{x_1, \ldots, x_n\}$, si $\mathcal{V}$ es un espacio vectorial m-dimensional con base $\{y_1, \ldots, y_m\}$ y si $\{\alpha_{ij}\}$ es cualquier conjunto de nm escalares ($i = 1, \ldots, n$; $j = 1, \ldots m$), entonces hay una y sólo una forma bilineal w sobre $\mathcal{U} \oplus \mathcal{V}$ tal que $w(x_i, y_j) = \alpha_{ij}$, para toda i y j.*

PRUEBA. Si $x = \sum_i \xi_i x_i$, $y = \sum_j \eta_j y_j$, y w es una forma bilineal sobre $\mathfrak{U} \oplus \mathfrak{V}$ tal que $w(x_i, y_j) = \alpha_{ij}$, entonces

$$w(x, y) = \sum_i \sum_j \xi_i \eta_j w(x_i, y_j) = \sum_i \sum_j \xi_i \eta_j \alpha_{ij}.$$

De esta ecuación resulta clara la unicidad de w; la existencia de una w adecuada se prueba leyendo la misma ecuación de derecha a izquierda, esto es, definiendo por medio de ello a w. (Compárese este resultado con el § 15, Teorema 1).

TEOREMA 2. *Si $\mathfrak{U}$ es un espacio vectorial n-dimensional con base $\{x_1, \ldots, x_n\}$, y si $\mathfrak{V}$ es un espacio vectorial m-dimensional con base $\{y_1, \ldots, y_m\}$, entonces hay una base determinada en forma única $\{w_{pq}\}$ $(p = 1, \ldots, n;\ q = 1, \ldots, m)$ en el espacio vectorial de todas los formas bilineales sobre $\mathfrak{U} \oplus \mathfrak{V}$ con la propiedad de que $w_{pq}(x_i, x_j) = \delta_{ip}\delta_{jq}$. Consecuentemente, la dimensión del espacio de formas bilineales sobre $\mathfrak{U} \oplus \mathfrak{V}$ es el producto de las dimensiones de $\mathfrak{U}$ y $\mathfrak{V}$.*

PRUEBA. Aplicando el Teorema 1, determinamos w_{pq} (para cada p y q fijas), mediante la condición dada $w_{pq}(x_i, y_j) = \delta_{ip}\delta_{jq}$. Las formas bilineales así determinadas son linealmente independientes, puesto que

$$\sum_p \sum_q \alpha_{pq} w_{pq} = 0$$

implica que

$$0 = \sum_p \sum_q \alpha_{pq} \delta_{ip} \delta_{jq} = \alpha_{ij}.$$

Si, además, w es un elemento arbitrario de $\mathfrak{W}$, y si $w(x_i, y_j) = \alpha_{ij}$, entonces $w = \sum_p \sum_q \alpha_{pq} w_{pq}$. En realidad, si $x = \sum_i \xi_i x_i$ y $y = \sum_j \eta_j y_j$, entonces

$$w_{pq}(x, y) = \sum_i \sum_j \xi_i \eta_j \delta_{ip} \delta_{jq} = \xi_p \eta_q,$$

y, consecuentemente,

$$w(x, y) = \sum_i \sum_j \xi_i \eta_j \alpha_{ij} = \sum_p \sum_q \alpha_{pq} w_{pq}(x, y).$$

Se sigue que la w_{pq} forma una base en el espacio de formas bilneales; esto completa la prueba del teorema. (Compárese este resultado con el § 15, Teorema 2).

EJERCICIOS

1. (a) Si w es una forma bilineal sobre $\mathfrak{R}^n \oplus \mathfrak{R}^n$, entonces existen escalares α_{ij}, $i, j = 1, \ldots, n$, tales que si $x = (\xi_1, \ldots, \xi_n)$ y $y = (\eta_1, \ldots, \eta_n)$, en-

tonces $w(x, y) = \sum_i \sum_j \alpha_{ij}\xi_i\eta_j$. Los escalares α_{ij} están determinados en forma única por w.

(b) Si z es una función lineal sobre el espacio de todas las formas bilineales sobre $\mathcal{R}^n \oplus \mathcal{R}^n$, entonces existen escalares β_{ij} tales que [en la notación de (a)] $z(w) = \sum_i \sum_j \alpha_{ij}\beta_{ij}$ para toda w. Los escalares β_{ij} son determinados en forma única por z.

2. Una forma bilineal w sobre $\mathfrak{U} \oplus \mathfrak{V}$ es *degenerada* si, como función de uno de sus dos argumentos, desaparece idénticamente para algún valor no cero de su otro argumento; de otro modo es *no degenerada*.

(a) Dé un ejemplo de una forma bilineal degenerada (no idéntica a cero) sobre $\mathcal{C}^2 \oplus \mathcal{C}^2$.

(b) Dé un ejemplo de una forma bilineal no degenerada sobre $\mathcal{C}^2 \oplus \mathcal{C}^2$.

3. Si w es una forma bilineal sobre $\mathfrak{U} \oplus \mathfrak{V}$, si y_0 está en $\mathfrak{V}$, y si una función y se define sobre $\mathfrak{U}$ por medio de $y(x) = w(x, y_0)$, entonces y es una función lineal sobre $\mathfrak{U}$. ¿Es verdad que si w es no degenerada, entonces toda función lineal sobre $\mathfrak{U}$ puede obtenerse de este modo (por una adecuada selección de y_0)?

4. Supóngase que para cada x y y de $\mathcal{P}_n$ la función w se define

(a) $w(x, y) = \int_0^1 x(t)y(t)\,dt$,

(b) $w(x, y) = x(1) + y(1)$,

(c) $w(x, y) = x(1)\cdot y(1)$,

(d) $w(x, y) = x(1)\left(\frac{dy}{dt}\right)_{t=1}$.

¿En cuál de estos casos es w una forma bilineal sobre $\mathcal{P}_n \oplus \mathcal{P}_n$? ¿En cuáles casos es no degenerada?

5. ¿ Existe un espacio vectorial $\mathfrak{V}$ y una forma bilineal w sobre $\mathfrak{V} \oplus \mathfrak{V}$ tal que w no es idénticamente cero, sino $w(x, x) = 0$ para toda x de $\mathfrak{V}$?

6. (a) Una forma bilineal w sobre $\mathfrak{V} \oplus \mathfrak{V}$ es *simétrica* si $w(x, y) = w(y, x)$ para toda x y y. Una *forma cuadrática* sobre $\mathfrak{V}$ es una función q sobre $\mathfrak{V}$ obtenida de una forma bilineal w escribiendo $q(x) = w(x, x)$. Pruébese que si la característica del campo escalar subyacente es diferente de 2, entonces toda forma simétrica bilineal está determinada en forma única por la correspondiente forma cuadrática. ¿Qué sucede si la característica es 2?

(b) ¿Puede una forma bilineal no simétrica definir la misma forma cuadrática como simétrica?

§ 24. PRODUCTOS TENSORIALES

En esta sección describiremos un nuevo método de poner juntos dos espacios vectoriales para hacer un tercero, a saber, la formación de su producto tensorial. Aunque tendremos relativamente poca ocasión de hacer uso de productos tensoriales en este libro, su teoría está estrechamente aliada a algunos de los temas que trataremos, y es útil en otras partes relacionadas de las matemáticas, tales como la teoría de representaciones de grupos y el cálculo tensorial. La noción es esencialmente más complicada que la de suma directa; en consecuencia, comenzaremos por dar algunos ejemplos de lo que sería un producto tensorial y el estudio de estos ejemplos nos guiará para establecer una definición.

Sea $\mathcal{U}$ el conjunto de todos los polinomios en una variable s, con, digamos, coeficientes complejos; sea $\mathcal{V}$ el conjunto de todos los polinomios en otra variable t; y, finalmente, sea $\mathcal{W}$ el conjunto de todos los polinomios en dos variables s y t. Con respecto de las definiciones obvias de las operaciones lineales $\mathcal{U}$, $\mathcal{V}$, y $\mathcal{W}$ todas son espacios vectoriales complejos; en este caso, desearíamos llamar a $\mathcal{W}$, o algo similar, el producto tensorial de $\mathcal{U}$ y $\mathcal{V}$. Una razón para esta terminología es que si tomamos cualquier x de $\mathcal{V}$, y cualquier y de , podemos formar su producto, esto es, el elemento z de $\mathcal{W}$ definido por $z(s, t) = x(s)y(t)$. (Este es el producto ordinario de dos polinomios. Aquí como antes, estamos tercamente pasando por alto el hecho intrascedente de que podemos hasta multiplicar dos elementos de $\mathcal{U}$, esto es, que el producto de dos polinomios en la misma variable es otro polinomio en esa variable. Los espacios vectoriales en que se define un concepto conveniente de multiplicación se llaman *álgebras*, y su estudio, como tal, queda fuera del alcance de este libro).

En el ejemplo precedente consideramos espacios vectoriales cuyos elementos son funciones. Podemos, si deseamos, considerar el simple espacio vectorial $\mathcal{C}^n$ como una colección de funciones también: El dominio de definición de la función es, en este caso, un conjunto consistente en exactamente n puntos, digamos, los primeros n (estrictamente) enteros positivos. En otras palabras, un vector $(\xi_1, \ldots, \xi_n)$ puede considerarse como una función de ξ cuyo valor $\xi(i)$ está definido por $i = 1, \ldots, n$; la definición de las operaciones vectoriales en $\mathcal{C}^n$ es tal, que corresponden, en la nueva notación, a las operaciones ordinarias ejecutadas sobre las funciones ξ. Si, simultáneamente, consideramos $\mathcal{C}^m$ como la colección de funciones η cuyo valor $\eta(j)$ se define por $j = 1, \ldots, m$, entonces querríamos que el producto tensorial de $\mathcal{C}^n$ y $\mathcal{C}^m$ sea el conjunto de todas las funciones ζ cuyo valor $\zeta(i, j)$ se define por $i = 1, \ldots, m$ y $j = 1, \ldots, m$. El producto tensorial, en otras palabras, es la colección de todas las funciones definidas sobre un conjunto, consiste en exactamente nm objetos y, en consecuencia, naturalmente isomórfico a $\mathcal{C}^{nm}$. Este ejemplo pone de manifiesto una propiedad de los productos tensoriales, a saber, la multiplicatividad de dimensión, que desearíamos retener en el caso general.

Tratemos ahora de abstraer las más importantes propiedades de estos ejemplos. La definición de suma directa fue una posible rigorización de la cruda idea intuitiva de escribir, formalmente, la suma de dos vectores que pertenecen a espacios vectoriales diferentes. De modo semejante, nuestros ejemplos sugieren que el producto tensorial $\mathcal{U} \otimes \mathcal{V}$ de los dos espacios vectoriales $\mathcal{U}$ y $\mathcal{V}$ deberían ser tales que cada x de $\mathcal{U}$ y y de $\mathcal{V}$ corresponda a un "producto" $z = x \otimes y$ en $\mathcal{U} \otimes \mathcal{V}$, en tal forma que la correspondencia entre x y z, para cada y fija, así como la correspondencia entre y y z, para cada x fija, es lineal. (Esto significa, por supuesto, que $(\alpha_1 x_1 + \alpha_2 x_2) \otimes y$ debería ser igual a $\alpha_1(x_1 \otimes y) + \alpha_2(x_2 \otimes y)$ y que una ecuación similar sería válida para $x \otimes (\alpha_1 y_1 + \alpha_2 y_2)$.) Para decirlo más sencillamente, $x \otimes y$ deberían definir una función bilineal (valuada vectorialmente) de x y y.

La noción de multiplicación formal requiere también que si u y v son funcionales lineales sobre $\mathcal{U}$ y $\mathcal{V}$ respectivamente, entonces es su producto w, definido por $w(x, y) = u(x)v(y)$, el que debería ser, en cierto sentido, el elemento general del espacio dual $(\mathcal{U} \otimes \mathcal{V})'$. Obsérvese que este producto es una función bilineal (valuada escalarmente), de x y y.

§ 25. BASES DE PRODUCTOS

Después de una palabra más de explicación preliminar estaremos listos para discutir la definición formal de productos tensoriales. Resulta ser técnicamente preferible llegar a $\mathcal{U} \otimes \mathcal{V}$ indirectamente, definiéndolo como el dual de otro espacio; haremos un uso tácito de la reflexividad para obtener $\mathcal{U} \otimes \mathcal{V}$ mismo. Puesto que hemos probado la reflexividad solamente para espacios finito-dimensionales, restringiremos la definición a esos espacios.

DEFINICIÓN. El *producto tensorial* $\mathcal{U} \otimes \mathcal{V}$ de dos espacios vectoriales finito-dimensionales $\mathcal{U}$ y $\mathcal{V}$ (sobre el mismo campo) es el dual del espacio vectorial de todas las formas bilineales sobre $\mathcal{U} \oplus \mathcal{V}$. Para cada par de vectores x y y, con x en $\mathcal{U}$ y y en $\mathcal{V}$, el producto tensorial $z = x \otimes y$ de x y y es el elemento de $\mathcal{U} \otimes \mathcal{V}$ definido por $z(w) = w(x, y)$ para cada forma bilineal w.

Esta definición es uno de los más rápidos abordajes rigurosos de la teoría, pero posteriormente lleva a algunas desagradables complicaciones técnicas. Sin embargo, cualesquiera que sean sus desventajas, observamos que obviamente tienen las dos propiedades deseadas; está claro, a saber, que la dimensión es multiplicativa (véase el § 23, Teorema 2, y el § 15, Teorema 2) y está claro que $x \otimes y$ depende linealmente de cada uno de sus factores.

Otra posible (y merecidamente popular) definición de productos tensoriales es por productos formales. De acuerdo con esa definición, $\mathfrak{U} \otimes \mathfrak{V}$ se obtiene considerando todos los símbolos de la forma $\sum_i \alpha_i(x_i \otimes y_i)$, y, dentro del conjunto de esos símbolos, haciendo las identificaciones demandadas por la linearidad de las operaciones vectoriales y la bilinearidad de la multiplicación tensorial. (Para el purista: en esta definición $x \otimes y$ representan meramente el par ordenado de x y y; el signo de multiplicación es solamente una insinuación de lo que se espera). Ninguna definición es sencilla; adoptamos la que dimos porque parecía más de acuerdo con el espíritu del resto del libro. La principal desventaja nuestra es que no se extiende fácilmente a las generalizaciones más útiles de espaciones vectoriales finito-dimensionales, esto es, a módulos y a espacios finito-dimensionales.

Por el momento probaremos solamente un teorema sobre productos tensoriales. El teorema es una ulterior justificación de la teoría de productos e, incidentalmente, es un afinamiento de la aserción de que la dimensión es multiplicativa.

Teorema. *Si $\mathfrak{X} = \{x_1, \ldots, x_n\}$ y $\mathfrak{Y} = \{y_1, \ldots, y_m\}$ son bases de $\mathfrak{U}$ y de $\mathfrak{V}$ respectivamente, entonces el conjunto $\mathfrak{Z}$ de vectores $z_{ij} = x_i \otimes y_j$ $(i = 1, \ldots, n; j = 1, \ldots, m)$ es una base de $\mathfrak{U} \otimes \mathfrak{V}$.*

Prueba. Sea w_{pq} la forma bilineal sobre $\mathfrak{U} \oplus \mathfrak{V}$ tal que $w_{qp}(x_i y_j) = \delta_{ip}\delta_{jq}$ $(i, p = 1, \ldots, n; j, q = 1, \ldots, m)$; la existencia de esas formas bilineales y el hecho de que constituyen una base para todas las formas bilineales, se sigue de § 23, Teorema 2. Sea $\{w'_{pq}\}$ la base dual de $\mathfrak{U} \otimes \mathfrak{V}$, de manera que $[w_{ij}, w'_{pq}] = \delta_{ip}\delta_{jq}$. Si $w = \sum_p \sum_q \alpha_{pq} w_{pq}$ es una forma bilineal arbitraria sobre $\mathfrak{U} \oplus \mathfrak{V}$, entonces

$$w'_{ij}(w) = [w, w'_{ij}] = \sum_p \sum_q \alpha_{pq}[w_{pq}, w'_{ij}]$$

$$= \alpha_{ij} = w(x_i, y_j) = z_{ij}(w).$$

La conclusión se sigue del hecho de que los vectores w'_{ij} constituyen una base de $\mathfrak{U} \otimes \mathfrak{V}$.

EJERCICIOS

1. Si $x = (1, 1)$ y $y = (1, 1, 1)$ son vectores de $\mathbb{R}^2$ y $\mathbb{R}^3$ respectivamente, encuéntrense las coordenadas de $x \otimes y$ en $\mathbb{R}^2 \otimes \mathbb{R}^3$ con respecto a la base del producto $\{x_i \otimes y_j\}$ donde $x_i = (\delta_{i1}, \delta_{i2})$ y $y_j = (\delta_{1j}, \delta_{2j}, \delta_{3j})$.

2. Sea $\mathcal{P}_{n,m}$ el espacio de todos los polinomios z con coeficientes complejos, en dos variables s y t, tales que cada $z = 0$ o el grado de $z(s, t)$ sea $\leq m - 1$ para cada s fija y $\leq n - 1$ para cada t fija. Pruébese que existe un isomorfismo entre $\mathcal{P}_n \otimes \mathcal{P}_m$ y $\mathcal{P}_{n,m}$ tal que el elemento z de $\mathcal{P}_{n,m}$ que corresponde a $x \otimes y$ (x en $\mathcal{P}_n$, y en $\mathcal{P}_m$) es dado por $z(s, t) = x(s)y(t)$.

3. ¿Hasta qué grado es la formación de productos tensoriales conmutativa y asociativa? ¿Qué hay sobre la ley distributiva $\mathcal{U} \otimes (\mathcal{V} \oplus \mathcal{W}) = (\mathcal{U} \otimes \mathcal{V}) \oplus (\mathcal{U} \otimes \mathcal{W})$?

4. Si $\mathcal{V}$ es un espacio vectorial finito-dimensional, y si x y y están en $\mathcal{V}$, ¿es verdad que $x \otimes y = y \otimes x$?

5. (a) Supóngase que $\mathcal{V}$ es un espacio vectorial finito-dimensional real y que $\mathcal{U}$ es el conjunto $\mathcal{C}$ de todos los números complejos considerados como un espacio vectorial real (bidimensional). Fórmense el producto tensorial $\mathcal{V}^+ = \mathcal{U} \otimes \mathcal{V}$. Pruebese que hay un modo de definir productos de números complejos con elementos de $\mathcal{V}^+$ de manera que $\alpha(x \otimes y) = \alpha x \otimes y$ cada vez que α y x están en $\mathcal{C}$ y y está en $\mathcal{V}$.

(b) Pruébese que con respecto de la adición vectorial y con respecto de la multiplicación escalar compleja, como se define en (a), el espacio $\mathcal{V}^+$ es un espacio vectorial complejo.

(c) Encuéntrese la dimensión del espacio vectorial complejo $\mathcal{V}^+$ en términos de las dimensiones del espacio vectorial real $\mathcal{V}$.

(d) Pruébese que el espacio vectorial $\mathcal{V}$ es isomórfico a un subespacio en $\mathcal{V}^+$ (cuando este último es considerado como un espacio vectorial real).

La moral de este ejercicio es que no sólo cada espacio vectorial complejo puede ser considerado como un espacio vectorial real, sino que, en cierto sentido, la recíproca es verdadera. El espacio vectorial $\mathcal{V}^+$ se llama la *complejificación* de $\mathcal{V}$.

6. Si $\mathcal{U}$ y $\mathcal{V}$ son espacios vectoriales finito-dimensionales, ¿cuál es el espacio dual de $\mathcal{U}' \otimes \mathcal{V}'$?

§ 26. PERMUTACIONES

El principal tema de este libro es generalmente conocido como álgebra lineal. En las últimas tres secciones, sin embargo, se puso énfasis en algo llamado álgebra multilineal. Es difícil decir exactamente dónde queda la línea divisoria entre los dos temas. Puesto que, en cualquier caso, ambos son completamente extensos, no sería práctico tratar de acumular un tratamiento extenso de ambos en el mismo volumen. Ni es deseable discutir el álgebra lineal en su estado absolutamente puro; la adición de aun una pequeña parte de la teoría multilineal (tal como está implicada en el moderno punto de vista de productos tensoriales y determinantes) extiende gratamente el dominio de la teoría multilineal fuera de proporción al esfuerzo involucrado. Nos proponemos, en cosecuencia, continuar el estudio del álgebra multilineal; nuestra intención es la de trazar una línea más o menos recta entre lo que ya sabemos y los hechos básicos de los determinantes. Teniendo eso presente, dedicaremos tres secciones a la discusión de algunos sencillos hechos sobre combinatoria; la conexión entre esos hechos y el álgebra multilineal aparecerá inmediatamente después de esa discusión.

Por una *permutación* de los enteros entre 1 y k (inclusive), significaremos una transformación uno a uno que asigna a cada uno de esos enteros otro (o posiblemente el mismo). Decir que la transformación π es uno a uno significa, por supuesto, que si $\pi(1), \ldots, \pi(k)$ son los enteros que π asigna a $1, \ldots, k$, respectivamente, entonces $\pi(i) = \pi(j)$ puede acontecer sólo en caso de que $i = j$. Puesto que esto implica que ambos conjuntos $\{1, \ldots, k\}$ y $\{\pi(1), \ldots, \pi(k)\}$ consisten en exactamente k elementos, se sigue que constan exactamente de los mismos elementos. De esto, a su vez, inferimos que una permutación π del conjunto $\{1, \ldots, k\}$ traza un mapa de ese conjunto *sobre* sí mismo, esto es, que si $1 \leqq j \leqq k$, entonces existe cuando menos una i (y, de hecho, exactamente una) tal que $\pi(i) = j$. El número total de enteros que se consideran, a saber, k, será mantenido fijo en toda la discusión siguiente.

La teoría de las permutaciones, como todo lo demás, es mejor entendida observando fijamente algunos ejemplos no triviales. Sin embargo, antes de presentar cualesquiera ejemplos, mencionaremos primero algunas de las cosas generales que pueden hacerse con permutaciones; por este medio los ejemplos ilustrarán no sólo los conceptos básicos, sino también sus propiedades básicas.

Si σ y τ son permutaciones arbitrarias, una permutación (que será denotada por $\sigma\tau$) puede definirse escribiendo

$$(\sigma\tau)(i) = \sigma(\tau i)$$

para cada i. Para probar que $\sigma\tau$ es en realidad una permutación, obsérvese que si $(\sigma\tau)(i) = (\sigma\tau)(j)$, entonces $\tau(i) = \tau(j)$ (puesto que σ es uno a uno) y, en consecuencia, $i = j$, (puesto que τ es uno a uno). La permutación $\sigma\tau$ se llama el *producto* de las permutaciones σ y τ. Advertencia: el orden es importante. En general, $\sigma\tau \neq \tau\sigma$, o, en otras palabras, la multiplicación de permutaciones no es conmutativa.

La multiplicación de permutaciones es asociativa; esto es, si π, σ y τ son permutaciones, entonces

(1) $$(\pi\sigma)\tau = \pi(\sigma\tau).$$

Para probar esto, debemos demostrar que

$$((\pi\sigma)\tau)(i) = (\pi(\sigma\tau))(i)$$

para toda i. La prueba consiste en diversas aplicaciones de la definición de producto, como sigue

$$((\pi\sigma)\tau)(i) = (\pi\sigma)(\tau i) = \pi(\sigma(\tau(i))),$$

y

$$(\pi(\sigma\tau))(i) = \pi((\sigma\tau)(i)) = \pi(\sigma(\tau(i))).$$

En vista de este resultado podemos omitir y omitiremos paréntesis al escribir el producto de tres o más permutaciones. El resultado nos faculta también para probar las obvias leyes de los exponentes. Las potencias de una permutación π se definen inductivamente escribiendo $\pi^1 = \pi$ y $\pi^{p+1} = \pi \cdot \pi^p$ para toda $p = 1, 2, 3, \ldots$; la ley asociativa implica que $\pi^p\pi^q = \pi^{p+q}$ y $(\pi^p)^q = \pi^{pq}$ para toda p y q. Obsérvese que cualesquiera dos potencias de una permutación conmutan una con otra, esto es, que $\pi^p\pi^q = \pi^q\pi^p$.

La permutación más sencilla es la *identidad* (que se denota con ϵ); se define como $\epsilon(i) = i$ para cada i. Si π es una permutación arbitraria, entonces

(2) $$\epsilon\pi = \pi\epsilon = \pi,$$

o, en otras palabras, la multiplicación por ϵ deja inafectada a cada permutación. La prueba es directa; para cada i tenemos

$$(\epsilon\pi)(i) = \epsilon(\pi(i)) = \pi(i)$$

y

$$(\pi\epsilon)(i) = \pi(\epsilon(i)) = \pi(i).$$

La permutación ϵ se comporta, desde el punto de vista de la multiplicación, como el número 1. En analogía con la convención numérica usual, la potencia de orden cero de cada permutación π se definió escribiendo $\pi^0 = \epsilon$.

Si π es una permutación arbitraria, entonces existe una permutación (denotada por π^{-1}), tal que

(3) $$\pi^{-1}\pi = \pi\pi^{-1} = \epsilon.$$

Para definir $\pi^{-1}(j)$, donde, por supuesto, $1 \leq j \leq k$, encuéntrese la i única, tal que $\pi(i) = j$ y escríbase $\pi^{-1}(j) = i$; la validez de (3) es una consecuencia inmediata de las definiciones. La permutación π^{-1} es llamada la *inversa* de π

Sea $\mathcal{S}_k$ el conjunto de todas las permutaciones de los enteros entre 1 y k. Lo que hemos probado hasta ahora es que una operación de multiplicación puede ser definida por los elementos de $\mathcal{S}_k$ de manera que (1) la multiplicación es asociativa, (2) existe un elemento de identidad, esto es, un elemento tal que la multiplicación por el mismo deja fijo cada elemento de S_k, y (3) cada elemento tiene un inverso, esto es, un elemento cuyo producto con el elemento dado es la identidad. Un conjunto que satisface a (1)-(3) se llama un *grupo* con respecto del concepto de producto a que se refieren esas condiciones; el conjunto S_k, en particular se llama el *grupo simétrico* de *grado* k. Obsérvese que los enteros $1, \ldots, k$ podrían ser reemplazados por k objetos distintos sin afectar ninguno de los conceptos definidos anteriormente; el cambio sería meramente cuestión de notación.

§ 27. CICLOS

Un ejemplo simple de permutación se obtiene como sigue: escójanse dos enteros distintos cualesquiera entre 1 y k, digamos p y q, y escríbase

$$\begin{aligned} \tau(p) &= q, \\ \tau(q) &= p, \\ \tau(i) &= i \text{ cada vez que } i \neq p \text{ e } i \neq q. \end{aligned}$$

La permutación τ así definida se denota por (p, q); cada permutación de esta forma se llama una *transposición*. Si τ es una transposición, entonces $\tau^2 = \varepsilon$.

Otro modo usual de construir ejemplos, es elegir p enteros distintos entre 1 y k, digamos $i_1, \ldots, i_p$, y escribir

$$\begin{aligned} \sigma(i_j) &= i_{j+1} \text{ siempre que } 1 \leqq j < p, \\ \sigma(i_p) &= i_1, \\ \sigma(i) &= i \text{ siempre que } i \neq i_1, \cdots, i \neq i_p. \end{aligned}$$

La permutación σ así definida es denotada por $(i_1, \ldots i_p)$. Si $p = 1$, entonces $\sigma = \varepsilon$; si $p = 2$, entonces σ es una transposición. Para cualquier p con $1 < p \leqq k$, toda permutación de la forma $(i_1, \ldots, i_p)$ es llamada un *ciclo-p* o simplemente un *ciclo;* los ciclos-2 son exactamente las transposiciones. Advertencia: no se supone que $i_1 < \ldots < i_p$. Si, por ejemplo, $k = 5$ y $p = 3$, entonces hay veinte ciclos distintos. Obsérvese también que la notación para ciclos no es única; los símbolos (1, 2,3), (2, 3, 1) y (3, 1, 2) denotan todos la misma permutación. Dos ciclos $(i_1, \ldots, i_p)$ y $(j_1, \ldots, j_q)$, son *disjuntos* si ninguna de las i es igual a ninguna de las j. Si σ y τ son ciclos disjuntos entonces $\sigma\tau = \tau\sigma$ o, en otras palabras, σ y τ conmutan.

Teorema 1. Toda permutación es el producto de ciclos disjuntos en pares.

Prueba. Si π es una permutación y si i es tal que $\pi\ (i) \neq i$ (supóngase, por el momento, que $\pi \neq \varepsilon$), forma la secuencia $(i, \pi(i), \pi^2(i), \ldots)$. Puesto que hay sólo un número finito de enteros distintos entre 1 y k, deben existir exponentes p y $q (0 \leqq p < q)$ tales que $\pi^p(i) = \pi^q(i)$. El carácter uno a uno de π implica que $\pi^{q-p}(i) = i$, o, con un obvio cambio de notación, lo que hemos probado es que debe existir un exponente p estrictamente positivo, tal que $\pi^p(i) = i$. Si se escoge a p, como el exponente más pequeño con esta propiedad, entonces los enteros $i, \ldots, \pi^{p-1}(i)$ son distintos uno de otro. (En realidad, si $0 \leqq q < r < p$ y $\pi^q(i) = \pi^r(i)$, entonces $\pi^{r-q}(i) = i$, contradiciendo la minimalidad de p). Se sigue que $(i, \ldots, \pi^{p-1}(i))$ es un ciclo-p. Si hay una j entre 1 y k diferente de cada una de las $i, \ldots, \pi^{p-1}(i)$ y diferente de $\pi(j)$, repetimos el procedimiento que nos condujo a este ciclo, con j en lugar de i. Continuamos formando ciclos de esta manera, mientras después de cada paso podamos todavía hallar un nuevo entero que π no envíe sobre sí misma; el producto de los ciclos disjuntos así construidos es π. El caso $\pi = \varepsilon$ es cubierto por la convención un tanto natural de que un producto sin factores, un "producto vacío", ha de interpretarse como la permutación de identidad.

Teorema 2. *Todo ciclo es un producto de trasposiciones.*

Prueba. Supóngase que σ es un ciclo-p, en obsequio de la sencillez de notación, daremos la prueba, que es perfectamente general, en el caso especial $p = 5$. La prueba misma consiste en una línea:

$$(i_1, i_2, i_3, i_4, i_5) = (i_1, i_5)(i_1, i_4)(i_1, i_3)(i_1, i_2).$$

Unas cuantas palabras más de explicación podrían ser útiles. En vista de la definición del producto de permutaciones, el segundo miembro de la última ecuación opera sobre cada entero entre 1 y k de adentro hacia afuera o, quizás más sugestivamente, de derecha a izquierda. Así, por ejemplo, el resultado de aplicar $(i_1, i_5)\ (i_1, i_4)\ (i_1, i_3)\ (i_1, i_2)$ a i_3 se calcula como sigue: $(i_1, i_2)(i_3) = i_3$, $(i_1, i_3)(i_3) = i_1$, $(i_1, i_4)(i_1) = i_4$, $(i_1, i_5)(i_4) = i_4$, de manera que $(i_1, i_5)(i_1, i_4)(i_1, i_3)(i_1, i_2)(i_3) = i_4$.

Por motivos de referencia registramos el siguiente corolario inmediato de los teoremas precedentes.

TEOREMA 3. *Toda permutación es un producto de trasposiciones.*

Obsérvese que de las trasposiciones de los teoremas 2 y 3 no se dijo que fueran disjuntas; en general no lo son.

EJERCICIOS

1. (a) ¿Cuántas permutaciones hay en $\mathcal{S}_k$?

(b) ¿Cuántos ciclos-p distintos hay en $\mathcal{S}_k$ ($1 \leqq p \leqq k$?

2. Si σ y τ son permutaciones (en $\mathcal{S}_k$), entonces $(\sigma\tau)^{-1} = \tau^{-1}\sigma^{-1}$.

3. (a) Si σ y τ son permutaciones (en $\mathcal{S}_k$), entonces existe una permutación única en π tal que $\sigma\pi = \tau$.

(b) Si π, σ y τ son permutaciones tales que, $\pi\sigma = \pi\tau$, entonces $\sigma = \tau$.

4. Dese un ejemplo de una permutación que no sea el producto de trasposiciones disjuntas.

5. Pruébese que toda permutación en $\mathcal{S}_k$ es el producto de trasposiciones de la forma $(j, j+1)$, donde $1 \leqq j < k$. ¿Es esta factorización única?

6. ¿Es el inverso de un ciclo también un ciclo?

7. Pruébese que la representación de una permutación como el producto de ciclos disjuntos es única, excepto posiblemente por el orden de los factores.

8. El *orden* de una permutación π es el entero mínimo $p (< 0)$ tal que $\pi^p = \varepsilon$.

(a) Toda permutación tiene un orden.

(b) ¿Cuál es el orden de un ciclo-p?

(c) Si σ es un ciclo-p, τ es un ciclo-q y σ y τ son disyuntas, ¿cuál es el orden de $\sigma\tau$?

(d) Dese un ejemplo para mostrar que el supuesto de disyunción es esencial en (c).

(e) Si π es una permutación de orden p y si $\pi^p = \varepsilon$, entonces q es divisible por p.

9. Cada permutación de $\mathcal{S}_k$ $(k > 1)$ puede escribirse como un producto, cada factor del cual es una de las trasposiciones $(1, 2)$, $(1, 3)$, $(1, 4)$, $\ldots$, $(1, k)$.

10. Dos permutaciones σ y τ son llamadas *conjugadas* si existe una permutación π tal que $\sigma\pi = \pi\tau$. Pruébese que σ y τ son conjugadas si y sólo si tienen la misma estructura del ciclo. (Esto significa que en la representación de σ como un producto de ciclos disyuntos, el número de ciclos-p es, para cada p, el mismo que el número correspondiente para τ).

§ 28. PARIDAD

Puesto que $(1, 3)(1, 2) = (1, 2)(2, 3)(= (1, 2, 3))$, vemos que la representación de una permutación (hasta un ciclo) como

$(t_3 - t_4)$, o el par $(t_1 - t_2)$ y $t_1 - t_4$). Cada factor en esos pares entra al otro factor, excepto posiblemente que su signo pueda cambiar; si cambia para un factor, cambiará para su compañero. (3) Si σ y τ *son permutaciones*, entonces $(\sigma\tau)f = \sigma(\tau f)$; consecuentemente $\sigma\tau$ es par si y sólo si σ y τ tienen la misma paridad. Obsérvese que $\operatorname{sgn}(\sigma\tau) = (\operatorname{sgn}\sigma)(\operatorname{sgn}\tau)$.

Se sigue de (2) y (3) que un producto de un manojo de trasposiciones es par si y sólo si hay un número par de ellas y es impar en el otro caso. (Nótese, en particular, observando la prueba del § 27, Teorema 2, que un ciclo-p es par si y sólo si p es impar; en otras palabras, si σ es un ciclo-p, entonces $\operatorname{sgn}\sigma = (-1)^{p+1}$. Conclusión: Sin importar cómo una permutación π es factorizada en trasposiciones, el número de factores es siempre par (este es el caso si π es par) o también es siempre impar (este es el caso si π es impar).

El producto de dos permutaciones pares es par; la inversa de una permutación par es par; la permutación de identidad es par. Estos hechos se resumen diciendo que el conjunto de todas las permutaciones pares es un *subgrupo* de $\mathcal{S}_k$; este subgrupo (denotado por $\mathcal{A}_k$) se llama el *grupo alterno de grado k*.

EJERCICIOS

1. ¿Cuántas permutaciones hay en $\mathcal{A}_k$?
2. Dense ejemplos de permutaciones pares con orden par, y de permutaciones pares con orden impar; hágase lo mismo para permutaciones impares.
3. Toda permutación en $\mathcal{A}_k (k > 2)$ puede escribirse como un producto, cada factor del cual es uno de los ciclos-3 $(1, 2, 3)$ $(1, 2, 4)$, $\ldots$, $(1, 2, k)$.

§ 29. FORMAS MULTILINEALES

Estamos ahora listos para proceder con el álgebra multilineal. El concepto básico es el de forma multilineal (o funcional), una fácil generalización del concepto de forma bilineal. Supóngase que $\mathcal{V}_1, \ldots, \mathcal{V}_k$ son espacios vectoriales (sobre el mismo campo); una forma *k-lineal* ($k = 1, 2, 3, \ldots$) es una función valuada escalarmente sobre la suma directa $\mathcal{V}_1 \oplus \cdots \oplus \mathcal{V}_k$, con la propiedad de que para cada valor fijo de cualesquiera argumentos $k - 1$, depende linealmente del argumento restante. Las formas 1-lineales son simplemente las funcionales lineales (sobre $\mathcal{V}_1$) y las formas 2-lineales son las formas bilineales (sobre $\mathcal{V}_1 \oplus \mathcal{V}_2$). Las formas 3-lineales (o trilineales) son las funciones w (sobre $\mathcal{V}_1 \oplus \mathcal{V}_2 \oplus \mathcal{V}_3$), valuadas escalarmente, tales que

$$w(\alpha_1 x_1 + \alpha_2 x_2, y, z) = \alpha_1 w(x_1, y, z) + \alpha_2 w(x_2, y, z),$$

y tales que identidades similares son válidas para $w(x, \alpha_1 y_1 + \alpha_2 y_2, z)$, y $w(x, y, \alpha_1 z_1 + \alpha_2 z_2)$. Una función que es k-lineal para algunas k, se llama una *forma multilineal.*

Gran parte de la teoría de las formas bilineales se extiende fácilmente al caso multilineal. Así, por ejemplo, si w_1 y w_2 son formas k-lineales, si α_1 y α_2 son escalares, y si w se define por

$$w(x_1, \cdots, x_k) = \alpha_1 w_1(x_1, \cdots, x_k) + \alpha_2 w_2(x_1, \cdots, x_k)$$

siempre que x_i esté en $\mathcal{V}_i$, $i = 1, \ldots, k$, entonces w es una forma k-lineal, denotada por $\alpha_1 w_1 + \alpha_2 w_2$. El conjunto de todas las formas k-lineales es un espacio vectorial con respecto de esta definición de las operaciones lineales; la dimensión de ese espacio vectorial es el producto $n_1, \ldots, n_k$, donde, por supuesto, n_i es la dimensión de $\mathcal{V}_i$

Las pruebas de todos estos enunciados son justamente como las pruebas (en el § 23) de los enunciados correspondientes para el caso bilineal. Podríamos seguir imitando la teoría bilineal y, en particular, estudiando productos tensoriales múltiples. A fin de mantener dentro de un mínimo nuestra digresión multilineal, marcharemos, en cambio, en una dirección diferente, más especial y, para nuestros propósitos, más útil.

En lo que sigue restringiremos nuestra atención al caso en que los k espacios $\mathcal{V}_i$ son todos iguales a uno y el mismo espacio vectorial, digamos, $\mathcal{V}$; supondremos que $\mathcal{V}$ es finito dimensional. En este caso llamaremos a una "forma k-lineal sobre $\mathcal{V}_1 \oplus \cdots \oplus \mathcal{V}_k$" simplemente "una forma k-lineal sobre $\mathcal{V}$", o, aún más sencillamente, una "forma k-lineal"; el lenguaje es ligeramente inadecuado, pero, en el contexto, completamente no ambiguo. Si la dimensión de $\mathcal{V}$, es n, entonces la dimensión del espacio vectorial de todas las formas k-lineales es n^k. El espacio $\mathcal{V}$ y, por supuesto, la dimensión n, se mantendrá fija en toda la discusión siguiente.

El carácter especial del caso que estamos estudiando nos faculta para aplicar una técnica que no es universalmente disponible; la técnica consiste en operar sobre las formas k-lineales por permutaciones en $\mathcal{S}_k$. Si w es una forma k-lineal, y si π está en $\mathcal{S}_k$, escribimos

$$\pi w(x_1, \cdots, x_k) = w(x_{\pi(1)}, \cdots, x_{\pi(k)})$$

siempre que $x_1, \ldots, x_k$ estén en $\mathcal{V}$. La función πw así definida es de nuevo una forma k-lineal. (El valor de πw en $(x_1, \ldots, x_k)$ es más honradamente denotado por $(\pi w)(x_1, \ldots, x_k)$; puesto que, no obstante, la notación más sencilla no parece conducir a confusión alguna, continuaremos usándola).

Usando el modo como las permutaciones actúan sobre formas k-lineales, podemos definir algunos interesantes conjuntos de esas formas. Así, por ejemplo, una forma k-lineal w se llama *simétrica* si $\pi w = w$ para cada permutación π de $\mathcal{S}_k$. (Nótese que si $k = 1$, entonces esta condición se satisface trivialmente). El conjunto de todas las formas simétricas k-lineales es un subespacio del espacio de todas las formas k-lineales. De aquí que, en particular, el origen de ese espacio, la forma k-lineal 0, es simétrico. Para un ejemplo no trivial, supóngase que $k = 2$, sean y_1 y y_2 funcionales lineales sobre $\mathcal{V}$, y escríbase

$$w(x_1, x_2) = y_1(x_1)y_2(x_2) + y_1(x_2)y_2(x_1).$$

Este procedimiento para construir formas k-lineales tiene generalizaciones útiles. Así, por ejemplo, si $1 \leqq h < k \leqq n$ y si u es una forma h-lineal y v es una forma $(k - h)$-lineal, entonces la ecuación

$$w(x_1, \cdots, x_k) = u(x_1, \cdots, x_h) \cdot v(x_{h+1}, \cdots, x_k)$$

define una forma k-lineal w, que, en general, es no simétrica. Una forma k-lineal simétrica puede obtenerse de w (o, para el caso, de cualquier forma k-lineal dada), formando $\sum \pi w$, donde la suma se extiende sobre todas las permutaciones π en $\mathcal{S}_k$.

No estudiaremos más las formas simétricas k-lineales. Las introducimos aquí porque constituyen una clase muy natural de funciones definibles en términos de permutaciones. Las abandonamos ahora en favor de otra clase de funciones, que juegan un papel mucho más importante en la teoría.

§ 30. FORMAS ALTERNAS

Una forma k-lineal w es *simétrico-oblicua* si $\pi w = -w$ para cada permutación impar π en $\mathcal{S}_k$. Equivalente, w es simétrico-oblicua si $\pi w = (\operatorname{sgn} \pi)w$ para toda permutación π en $\mathcal{S}_k$. (Si $\pi w = (\operatorname{sgn} \pi)w$ para toda π, entonces, en particular, $\pi w = -w$ cada vez que π es impar. Si, recíprocamente, $\pi w = -w$ para toda π impar, entonces, dada una π arbitraria, factorícese en transposiciones, digamos $\pi = \tau_1 \ldots \tau_q$, obsérvese que $\operatorname{sgn} \pi = (-1)^q$ y, puesto que $\pi w =$

$(-1)^q w$, conclúyase que $\pi w = (\operatorname{sgn} \pi) w$, como se enunció. Esta prueba hace uso de un hecho no probado, pero fácilmente verificable de que si σ y τ son permutaciones en $\mathcal{S}_k$, entonces $\sigma(\tau w) = (\sigma\tau) w$). El conjunto de todas las formas k-lineales simétrico-oblicuas es un subespacio del espacio de todas las formas k-lineales. Para obtener un ejemplo no trivial de una forma bilineal simétrico-oblicua w, sean y_1 y y_2 funcionales lineales y escribamos

$$w(x_1, x_2) = y_1(x_1)y_2(x_2) - y_1(x_2)y_2(x_1).$$

Más generalmente, si w es una forma k-lineal arbitraria, una forma k-lineal simétrico-oblicua puede obtenerse de w, formando $\sum (\operatorname{sgn} \pi)\, \pi w$, donde la suma se extiende sobre todas las permutaciones π en $\mathcal{S}_k$.

Una forma k-lineales w se llama alterna si $w(x_1, \ldots, x_k) = 0$, siempre que dos de las x son iguales. (Nótese que si $k = 1$, entonces esta condición se satisface vacuamente). El conjunto de todas las formas alternas k-lineales es un subespacio del espacio de todas las formas k-lineales. Hay una importante relación entre formas alternas y simétrico-oblicuas.

TEOREMA 1. *Toda forma alterna multilineal es simétrico-oblicua.*

PRUEBA. Supóngase que w es una forma alterna k-lineal, y que i y j son enteros, $1 \leq i < j \leq k$. Si $x_1, \ldots, x_k$ son vectores, escribimos

$$w_0(x_i, x_j) = w(x_1, \cdots, x_k);$$

si las x distintas de x_i y x_j se mantienen fijas (temporalmente), entonces w_0 es una forma alterna bilineal de sus dos argumentos. Puesto que, por bilinearidad

$$w_0(x_i + x_j, x_i + x_j) = w_0(x_i, x_i) + w_0(x_i, x_j) + w_0(x_j, x_i) + w_0(x_j, x_j),$$

y puesto que, por el carácter alterno de w_0, el primer miembro y los términos extremos del segundo miembro de esta ecuación desaparecen, vemos que $w_0(x_j, x_i) = - w_0(x_i, x_j)$. Esto, sin embargo, dice que

$$(i, j)w(x_1, \cdots, x_k) = -w(x_1, \cdots, x_k),$$

o, puesto que las x son arbitrarias, que $(i, j)w = - w$. Puesto que toda permutación impar es el producto de un número impar de trasposiciones, tales como (i, j), se sigue que $\pi w = - w$ para toda π impar, y la prueba de este teorema es completa. ∎

La conexión entre formas alternas y formas simétrico-oblicuas implica un punto sutil. Considérese la siguiente "prueba" del recíproco del Teorema 1: si w es una forma k-lineal simétrico-oblicua, si $1 \leq i < j \leq k$ y si $x_1, \ldots, x_k$ son vectores tales que $x_i = x_j$ entonces $(i, j)w(x_1, \ldots, x_k) = w(x_1, \ldots, x_k)$ puesto que $x_i = x_j$ y al mismo tiempo $(i, j)w(x_1, \ldots, x_k) = -w(x_1, \ldots, x_k)$, puesto que w es simétrico-oblicua; consecuentemente $w(x_1, \ldots, x_k) = -w(x_1, \ldots, x_k)$ de manera que w es alterna. Este argumento es erróneo; la dificultad está en la inferencia "si $w = -w$, entonces $w = 0$". Si examinamos esa inferencia con más detenimiento, encontramos que está basada en el razonamiento siguiente: si $w = -w$, entonces $w + w = 0$, de manera que $(1 + 1)w = 0$. Esto es correcto. La dificultad está en que, en ciertos campos $1 + 1 = 0$, y en consecuencia, la inferencia de $(1 + w1 = 0$ a $w = 0$, no está justificada; la recíproca del Teorema 1 es, de hecho, falsa para espacios vectoriales sobre esos campos.

TEOREMA 2. *Si $x_1, \ldots, x_k$ son vectores linealmente dependientes y si w es una forma k-lineal alterna, entonces $w(x_1, \ldots, x_k) = 0$.*

PRUEBA. Si $x_i = 0$ para alguna i, la conclusión es trivial. Si todas las x_i son diferentes de 0, aplicamos el teorema del § 6, para determinar una x_h, $2 \leq h \leq k$, que es una combinación lineal de las precedentes. Si, digamos, $x_h = \sum_{i=0}^{h-1} \alpha_i x_i$, reemplácese x_h en $w(x_1, \ldots, x_k)$ en este argumento h-gésimo y trácese la conclusión deseada mediante una $(h - 1)$-tuple aplicación en supuesto de que w es alterna.

En un caso extremo (a saber, cuando $k = n$), una especie de recíproco del Teorema 2, es verdadero.

TEOREMA 3. *Si w es una forma n-lineal alterna, no cero, y si $x_1, \ldots, x_n$ son vectores linealmente independientes, entonces $w(x_1, \ldots, x_n) \neq 0$.*

PRUEBA. Puesto que (§ 8, Teorema 2), los vectores $x_1, \ldots, x_n$ forman una base, podemos, dado un conjunto arbitrario de n vectores $y_1, \ldots, y_n$ escribir cada y como una combinación lineal de las x. Si reemplazamos cada y de $w(y_1, \ldots, y_n)$ por la correspondiente combinación lineal de x y expandimos el resultado por multilinealidad, obtenemos una larga combinación lineal de términos tales como $w(z_1, \ldots, z_n)$, donde cada z es una de las x. Si, en ese término, coinciden dos de las z, entonces, puesto que w es alterna, ese término debe desaparecer. Si, por otra parte, todas las z son distintas, entonces $w(z_1, \ldots, z_n) = \pi w(x_1, \ldots, x_n)$ para alguna permutación π. Puesto que (Teorema 1) w es simétrico-oblicua, se sigue que $w(z_1, \ldots, z_n) = (\operatorname{sgn} \pi) w(x_1, \ldots, x_n)$. Si $w(x_1 \ldots, x_n)$ fuera 0 se seguiría que $w(z_1, \ldots, z_n) = 0$, y de aquí que $w\ (y_1, \ldots, y_n) = 0$ para toda $y_1, \ldots, y_n$, contradiciendo el supuesto de que $w \neq 0$. La prueba (no el enunciado) de este resultado da un valioso corolario.

TEOREMA 4. *Dos formas n-lineales alternas cualesquiera son linealmente dependientes.*

PRUEBA. Supóngase que w_1 y w_2 son formas n-lineales alternas y que $\{x_1, \ldots, n_n\}$ es una base. Dados n vectores cualesquiera $y_1, \ldots, y_n$, escriba cada uno de ellos como una combinación lineal de las x, y, tal como anteriormente, reemplace cada uno de ellos, tanto en $w_1(y_1, \ldots, y_n)$ como en $w_2(y_1, \ldots, y_n)$, por la correspondiente combinación lineal. Se sigue que cada $w_1(y_1, \ldots, y_n)$ y $w_2(y_1, \ldots, y_n)$ es una combinación lineal (la misma combinación lineal) de términos tales como $w_1(z_1, \ldots, z_n)$ y $w_2(z_1, \ldots, z_n)$, donde cada z es una de las x. Puesto que $w_1(x_1, \ldots, x_n)$ y $w_2(x_1, \ldots, x_n)$ son escalares, son linealmente dependientes, de manera que existen escalares α_1 y α_2 no ambos ceros, tales que $\alpha_1 w_1(x_1 \ \ldots, x_n) + \alpha_2 w_2(x_1, \ldots, x_n) = 0$; de estos hechos podemos inferir que $\alpha_1 w_1 + \alpha_2 w_2 = 0$, como se afirmó.

§ 31. FORMAS ALTERNAS DE GRADO MAXIMO

Revisando la última sección, el lector observará que no dimos ningún ejemplo no trivial de formas k-lineales alternas y ni siquiera indirectamente sugerimos teorema alguno de existencia relativo a las mismas. De hecho, no siempre existen; el § 30, Teorema 2, implica, por ejemplo, que $k > n$, entonces 0 es la única forma k-lineal alterna. (Véase el § 8, Teorema 2). Para las aplicaciones que tenemos a la vista, necesitamos solamente un teorema de existencia; procedemos a probar una forma del mismo un tanto concisa.

TEOREMA. *El espacio vectorial de formas n-lineales alternas sobre un espacio vectorial n-dimensional es unidimensional.*

PRUEBA. Demostramos primero que si $1 \leqq k \leqq n$, entonces existe cuando menos una forma k-lineal alterna, no cero; la prueba va por inducción sobre k. Si $k = 1$, el resultado deseado se sigue de la existencia de funcionales lineales no triviales (véase el § 15, Teorema 3). Si $1 \leq k < n$, suponemos que v es una forma k-lineal alterna, no cero; usando v construiremos una forma lineal w alterna $(k + 1)$ no cero. Puesto que $v \neq 0$, podemos encontrar vectores $x_1^0 \ldots x_k^0$, tales que $v(x_1^0 \ldots x_k^0) \neq 0$ (los supraíndices son aquí justamente índices). Puesto que $k < n$, podemos encontrar un vector x_{k+1}^0 que no pertenece al subespacio sobretendido por $x_1^0, \ldots, x_k^0$ y (véase el § 17, Teorema 1), podemos entonces encontrar una funcional lineal u, tal que $u(x_1^0) = \ldots = u(x_k^0) = 0$ y $u(x_{k+1}^0) \neq 0$.

La prometida forma $(k + 1)$-lineal w se obtiene de la funcional lineal u y la forma k-lineal v, escribiendo

$$(1) \quad w(x_1, \cdots, x_k, x_{k+1}) = \sum_{i=1}^{k} (i, k+1)v(x_1, \cdots, x_k)u(x_{k+1}) - v(x_1, \cdots, x_k)u(x_{k+1}).$$

Así, por ejemplo, si $k = 3$, entonces

$$w(x_1, x_2, x_3, x_4) = v(x_4, x_2, x_3)u(x_1) + v(x_1, x_4, x_3)u(x_2) + v(x_1, x_2, x_4)u(x_3) - v(x_1, x_2, x_3)u(x_4).$$

Se sigue de la discusión elemental del § 29, que w es en realidad una forma $(k + 1)$-lineal; tenemos que probar que es no cero y alterna.

El hecho de que w no sea idénticamente cero es fácil de probar. En realidad, puesto que $u(x_i^0) = 0$ para $i = 1, \ldots, k$, se sigue que si reemplazamos x_i por x_i^0 en (1), $i = 1, \ldots, k + 1$, entonces los primeros k términos de la suma a la derecha desaparecen y, consecuentemente,

$$(2) \quad w(x_1^0, \cdots, x_k^0, x_{k+1}^0) = -v(x_1^0, \cdots, x_k^0)u(x_{k+1}^0) \neq 0.$$

Supóngase ahora que $x_1, \ldots, x_{k+1}$ son vectores e i y j son enteros tales que $1 \leq i < j \leq k + 1$ y $x_i = x_j$. Tenemos que probar que, bajo estas circunstancias, $w(x_1, \ldots, x_{k+1}) = 0$. Notamos que tanto x_i como x_j ocurren en el argumento de v en todos, excepto en dos de los $k + 1$ términos del primer miembro de (1). Puesto que v es alterna, los términos en que tanto x_i y x_j ocurren así, desaparecen todos.

El resto de la prueba se divide naturalmente en dos casos. Si $j = k + 1$, entonces todo lo que queda a la izquierda es

$$(i, k+1)v(x_1, \cdots, x_k)u(x_{k+1}) - v(x_1, \cdots, x_k)u(x_{k+1}),$$

y, puesto que $x_i = x_{k+1}$, esto es claramente igual a 0. Si $j \leqq k$, entonces cada uno de los dos términos que posiblemente no desaparezcan y que quedan todavía, pueden obtenerse del otro por una aplicación de la trasposición (i, j). Se sigue que esos términos difieren en signo solamente y de aquí que su suma sea cero. Esto prueba que w es alterna y prueba, en consecuencia, que la dimensión del espacio de formas n-lineales alternas es no menor de 1.

El hecho de que la dimensión del espacio de formas n-lineales alternas no es más de 1, es una consecuencia inmediata del § 30, Teorema 4.

Esto concluye nuestra discusión del álgebra multilineal. El lector podría bien hacer el cargo de que la discusión no estuvo fuertemente motivada. Este libro no puede contener una motivación completa; la justificación del estudio del álgebra multilineal es la amplia aplicabilidad de esa materia. La única aplicación que haremos es a la teoría de los determinantes (que, de seguro, podría ser tratada por métodos más directos, pero menos elegantes, que involucran mucho mayor dependencia de selecciones arbitrarias de bases); esa aplicación pertenece al capítulo siguiente.

EJERCICIOS

1. Si w es una forma k-lineal y si la característica del campo subyacente de escalares es diferente de 2 (esto es, si $1 + 1 \neq 0$), entonces w es la suma de una forma k-lineal simétrica y de una oblicuosimétrico. ¿Qué si la característica es 2?
2. Dé un ejemplo de una forma multilineal oblicuosimétrica que no sea alterna. (Recuérdese que en vista de la discusión del § 30, el campo de escalares debe tener característica 2).
3. Dese un ejemplo de una forma k-lineal alterna w sobre un espacio n-dimensional ($k < n$), tal que $w(x_1, \ldots, x_k) = 0$, para algún conjunto de vectores linealmente independientes $x_1, \ldots, x_k$.
4. ¿Cuál es la dimensión del espacio de todas las formas simétricas k-lineales? ¿Qué sobre los oblicuosimétrico? ¿Qué sobre los alternos?

CAPITULO 2

TRANSFORMACIONES LINEALES

§32. TRANSFORMACIONES LINEALES

Llegamos ahora a los objetos que realmente hacen interesantes a los espacios vectoriales.

DEFINICIÓN. *Una transformación lineal (u operador)* A *sobre un espacio vectorial* $\mathcal{V}$ es una correspondencia que asigna a cada vector x de $\mathcal{V}$ un vector Ax de $\mathcal{V}$, de tal modo que

$$A(\alpha x + \beta y) = \alpha Ax + \beta Ay$$

idénticamente en los vectores x y y y en los escalares α y β.

Hacemos de nuevo la observación que hicimos en relación con la definición de funcionales lineales, a saber, que para una transformación lineal A, como la definimos, $A0 = 0$. Por esta razón, esas transformaciones son algunas veces llamadas transformaciones lineales *homogéneas*.

Antes de discutir cualesquiera propiedades de las transformacio nes lineales damos varios ejemplos. No nos molestaremos en probar que las transformaciones que mencionamos son, en realidad, lineales; en todos los casos la verificación de la ecuación que define la linearidad, es un simple ejercicio.

(1) Dos transformaciones especiales de considerable importancia para el estudio que sigue, y para el cual consistentemente reservamos los símbolos 0 y 1, respectivamente, se definen (para todas las x) por $0x = 0$ y $1x = x$.

(2) Sea x_0 cualquier vector fijo en $\mathcal{V}$, y sea y_0 cualquier funcional lineal sobre $\mathcal{V}$; escríbase $Ax = y_0(x)\cdot x_0$. Más generalmente: sea $\{x_1, \ldots, x_n\}$ un conjunto finito arbitrario de vectores en $\mathcal{V}$; y sea $\{y_1, \ldots, y_n\}$ un conjunto correspondiente de funcionales linea-

les sobre $\mathcal{V}$ escríbase $Ax = y_1(x)x_1 + \ldots + y_n(x)x_n$. No es difícil probar que si, en particular, $\mathcal{V}$ es n-dimensional, y los vectores $x_1, \ldots, x_n$ forman una base para $\mathcal{V}$, entonces toda transformación lineal A tiene la forma que se acaba de describir.

(3) Sea π una permutación de los enteros $\{1, \ldots, n\}$; si $x = (\xi_1, \ldots, \xi_n)$ es un vector en $\mathcal{C}^n$, escríbase $Ax = (\xi_{\pi(1)}, \ldots, \xi_{\pi(n)})$. De modo semejante, sea π un polinomio con coeficientes complejos; si x es un vector (polinomio) en $\mathcal{P}$, escríbase $Ax = y$ para el polinomio definido por $y(t) = x(\pi(t))$.

(4) Para cualquier x de $\mathcal{P}_n$, $x(t) = \sum_{j=0}^{n-1} \xi_j t^j$, escríbase $(Dx)(t) = \sum_{j=0}^{n-1} j\xi_j t^{j-1}$.

(Usamos aquí la letra D para recordar que Dx es la derivada del polinomio x. Hacemos la observación de que podríamos haber definido a D sobre $\mathcal{P}$ así como sobre $\mathcal{P}_n$; posteriormente haremos uso de este hecho. Obsérvese que para los polinomios la definición de diferenciación puede darse en forma puramente algebraica y no necesita la teoría usual de los procesos límite.)

(5) Para cada x *de* $\mathcal{P}$, $x(t) = \sum_{j=0}^{n-1} \xi_j t^j$, debe escribirse $Sx = \sum_{j=0}^{n-1} \frac{\xi_j}{j+1} t^{j+1}$.

(Una vez más estamos disfrazando con notación algebraica un concepto analítico bien conocido.) Justamente como en (4) $(Dx)(t)$ estaba en lugar de $\frac{dx}{dt}$, de manera que aquí $(Sx)(t)$ es igual que $\int_0^t x(s)\,ds$.)

(6) Sea m un polinomio con coeficientes complejos en una variable t. (Podemos, aunque no es particularmente provechoso hacerlo, considerar a m como un elemento de $\mathcal{P}$.) Para cada x de $\mathcal{P}$, escribimos Mx para el polinomio definido por $(Mx)(t) = m(t)x(t)$. Para fines ulteriores, introduciremos un símbolo especial; en caso de que $m(t) = t$, escribiremos T para la transformación M, de manera que $(Tx)(t) = tx(t)$.

§33. TRANSFORMACIONES COMO VECTORES

Procedemos ahora a deducir ciertas propiedades elementales de, y relaciones entre, transformaciones lineales sobre un espacio vectorial. Más particularmente, indicaremos varios modos de hacer nuevas transformaciones, partiendo de las antiguas; generalmente nos satisfaremos con dar la definición de las nuevas transformaciones y omitiremos la prueba de linearidad.

Si A y B son transformaciones lineales, definimos su *suma*, $S = A + B$, por la ecuación $Sx = Ax + Bx$ (para toda x). Observamos que la conmutividad y la asociatividad de la adición en $\mathcal{V}$ implica inmediatamente que la adición de transformaciones lineales es conmutativa y asociativa. Muchos más que esto es verdad. Si consideramos la suma de cualquier transformación lineal A y la transformación lineal 0 (definida en la sección precedente, vemos que $A + 0 = A$. Si, para cada A, denotamos por $-A$ la transformación definida por $(-A)x = -(Ax)$, vemos que $A + (-A) = 0$, y que la transformación $-A$, así definida, es la única transformación lineal B con la propiedad de que $A + B = 0$. Resumiendo: Las propiedades de un espacio vectorial, descrito en los axiomas(**A**) del § 2, aparecen de nuevo en el conjunto de todas las transformaciones lineales sobre el espacio; el conjunto de todas las transformaciones lineales es un grupo abeliano con respecto a la operación de adición.

Continuamos en el mismo espíritu. Por ahora no sorprenderá a nadie si los axiomas (**B**) y (**C**) de espacios vectoriales se satisfacen también para el conjunto de todas las transformaciones lineales. Se satisfacen. Para cualquier A y cualquier escalar α, definimos el producto αA por medio de la ecuación $(\alpha A)x = \alpha(Ax)$. Los axiomas (**B**) y (**C**) son inmediatamente verificados; resumimos como sigue:

TEOREMA. *El conjunto de todas las transformaciones lineales sobre un espacio vectorial es él mismo un espacio vectorial.*

Usualmente pasaremos por alto este teorema; la razón es que podemos decir mucho más sobre transformaciones lineales y el mero hecho de que formen un espacio vectorial se usa muy rara vez. El "mucho más" que podemos decir, es que existe para las transformaciones lineales una más o menos conveniente definición de multiplicación, que discutiremos en la sección siguiente.

EJERCICIOS

1. Pruébese que cada una de las correspondencias descritas abajo es una transformación lineal.

(a) $\mathcal{V}$ es el conjunto $\mathcal{C}$ de números complejos, considerado como un espacio vectorial real; Ax es el conjunto complejo de x.

(b) $\mathcal{V}$ es $\mathcal{P}$; si x es un polinomio, entonces $(Ax)(t) = x(t + 1) - x(t)$.

(c) $\mathcal{V}$ es el producto tensorial k-duplo de un espacio vectorial consigo mismo; A es tal que $A(x_1 \otimes \cdots \otimes x_k) = x_{\pi(1)} \otimes \cdots \otimes x_{\pi(k)}$, donde π es una permutación de $\{1, \ldots, k\}$.

(d) $\mathcal{V}$ es el conjunto de todas las formas k-lineales sobre un espacio vectorial; $(Aw)(x_1, \ldots, x_k) = w(x_{\pi(1)}, \ldots, x_{\pi(k)})$, donde π es una permutación de $\{1, \ldots, k\}$.

(e) $\mathcal{V}$ es el conjunto de todas las formas k-lineales sobre un espacio vectorial; si w está en $\mathcal{V}$, entonces $Aw = \sum \pi w$, donde la suma se extiende sobre todas las permutaciones π en $\mathcal{S}_k$.

(f) Igual que (e), con la excepción de que $Aw = \sum (\operatorname{sgn} \pi)\, \pi w$.

2. Demuéstrese que si $\mathcal{V}$ es un espacio vectorial finito-dimensional, entonces el espacio de todas las transformaciones lineales sobre $\mathcal{V}$ es finito-dimensional y encuéntrese su dimensión.

3. El concepto de "transformación lineal", como se define en el texto, es demasiado especial para algunos fines. De acuerdo con una definición más general, una transformación lineal de un espacio vectorial $\mathcal{U}$ a un espacio vectorial $\mathcal{V}$ sobre el mismo campo es una correspondencia A que asigna a cada vector x de $\mathcal{U}$ un vector Ax de $\mathcal{V}$ de manera que

$$A(\alpha x + \beta y) = \alpha Ax + \beta Ay.$$

Pruébese que cada una de las correspondencias descritas abajo es una transformación lineal en este sentido generalizado.

(a) $\mathcal{V}$ es el campo de escalares de $\mathcal{U}$; A es una función lineal sobre $\mathcal{U}$.

(b) $\mathcal{U}$ es la suma directa de $\mathcal{V}$ con algún otro espacio; A traza el mapa de cada par de $\mathcal{U}$ sobre su primera coordenada.

(c) $\mathcal{V}$ es el coeficiente de $\mathcal{U}$ módulo de un subespacio; A traza el mapa de cada vector de $\mathcal{U}$ sobre el conjunto que determina.

(d) Sea w una funcional bilineal sobre una suma directa $\mathcal{U} \oplus \mathcal{V}_0$. Sea $\mathcal{V}$ el dual de $\mathcal{V}_0$, y defínase A como la correspondencia que asigna a cada x_0 de $\mathcal{U}$ la funcional lineal sobre $\mathcal{V}_0$ obtenida de w haciendo su primer argumento igual a x_0.

4. (a) Supóngase que $\mathcal{U}$ y $\mathcal{V}$ son espacios vectoriales sobre el mismo campo. Si A y B son transformaciones lineales de $\mathcal{U}$ a $\mathcal{V}$, si α y β son escalares y si

$$Cx = \alpha Ax + \beta Bx$$

para cada x de $\mathcal{U}$, entonces C es una transformación lineal de $\mathcal{U}$ a $\mathcal{V}$.

(b) Si escribimos por definición, $C = \alpha A + \beta B$, entonces el conjunto de todas las transformaciones lineales de $\mathcal{U}$ a $\mathcal{V}$ se convierte en un espacio vectorial con respecto a esta definición de las operaciones lineales.

(c) Pruébese que si $\mathcal{U}$ y $\mathcal{V}$ son finito-dimensionales, entonces lo es el espacio de todas las transformaciones lineales de $\mathcal{U}$ a $\mathcal{V}$, y encuéntrese su dimensión.

5. Supóngase que $\mathfrak{M}$ es un subespacio m-dimensional de un espacio vectorial n-dimensional $\mathcal{V}$. Pruébese que el conjunto de las transformaciones lineales A sobre $\mathcal{V}$ para las cuales $Ax = 0$, siempre que x esté en $\mathfrak{M}$ es un subespacio del conjunto de todas las transformaciones lineales sobre $\mathcal{V}$, y encuéntrese la dimensión de ese subespacio.

§34. PRODUCTOS

El *producto* P de dos transformaciones lineales A y B, $P = AB$, se define por la ecuación $Px = A(Bx)$.

La noción de multiplicación es fundamental para todo lo que sigue. Antes de dar ejemplos para ilustrar el significado de productos de tranformación, observemos las implicaciones del simbolismo $P = AB$. Decir que P es una transformación significa, por supuesto, que dado un vector x, P le hace algo a ese vector. Lo que le hace se encuentra operando sobre x con B, esto es, encontrando Bx, y a continuación operando sobre el resultado con A. En otras palabras, si consideramos el símbolo para una transformación como una receta para ejecutar un acto determinado, entonces el símbolo para el producto de dos transformaciones ha de leerse de derecha a izquierda. El orden para transformar por AB significa transformar primero por B y a continuación por A. Esto puede parecer como crear demasiada confusión por un motivo pequeño; sin embargo, como pronto veremos, la multiplicación de transformaciones es, en general, no conmutativa y el orden en que tranformamos tiene mucha importancia.

Los ejemplos más notorios de no conmutatividad se encuentran sobre el espacio $\mathcal{P}$. Consideremos las transformaciones de diferenciación y multiplicación D y T definidas por $(Dx)(t) = \dfrac{dx}{dt}$ y $(Tx)(t) = tx(t)$; tenemos

$$(DTx)(t) = \frac{d}{dt}(tx(t)) = x(t) + t\frac{dx}{dt}$$

y

$$(TDx)(t) = t\frac{dx}{dt}.$$

En otras palabras, no sólo es falso que $DT = TD$ (de manera que $DT - TD = 0$), sino que, de hecho, $(DT - TD)x = x$ para toda x, de manera que $DT - TD = 1$.

Sobre la base de los ejemplos del § 32, el lector podría estar en condiciones de construir muchos ejemplos de pares de transformaciones no conmutativas. Los que están acostumbrados a pensar geométricamente en las transformaciones lineales, pueden, por ejemplo, convencerse fácilmente de que el producto de dos rotaciones de $\mathcal{R}^3$ (alrededor del origen) depende, en general, del orden en que se ejecuten.

La mayor parte de las propiedades algebraicas formales de la multiplicación numérica (con la ya mencionada notable excepción de la conmutatividad) son válidas en el álgebra de transformaciones. Así, tenemos

(1) $$A0 = 0A = 0,$$

(2) $$A1 = 1A = A,$$

(3) $$A(B + C) = AB + AC,$$

(4) $$(A + B)C = AC + BC,$$

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

Las pruebas de todas estas identidades son consecuencias inmediatas de las definiciones de adición y multiplicación; para ilustrar los principios probamos (3), una de las leyes distributivas. La prueba consiste en el siguiente cómputo:

$$\begin{aligned}(A(B + C))x &= A((B + C)x) = A(Bx + Cx)\\ &= A(Bx) + A(Cx) = (AB)x + (AC)x\\ &= (AB + AC)x.\end{aligned}$$

§35. POLINOMIOS

La ley asociada de la multiplicación nos faculta para escribir el producto de tres (o más) factores sin ningún paréntesis; en particular podemos considerar el producto de cualquier número finito, digamos m, de factores todos iguales a A. Este producto depende solamente de A sobre m (y no, como acabamos de observar, de cualquier puesta en corchetes de los factores); lo denotaremos por A^m. La justificación de esta notación es que, aunque, en general, la multiplicación de transformaciones no es conmutativa, para las potencias de una transformación tenemos las leyes usuales de los exponentes, $A^nA^m = A^{n+m}$ y $(A^n)^m = A^{nm}$. Observamos que $A^1 = A$; es costumbre escribir también, por definición, $A^0 = 1$. Con estas definiciones el cálculo de potencias de una sola transformación es casi exactamente el mismo que el de la aritmética ordinaria. Podemos, en particular, definir los polinomios en una transformación lineal, así, si p es un polinomio con coeficientes escalares en una variable t, digamos $p(t) = \alpha_0 + \alpha_1 t + \ldots + \alpha_n t^n$, po-demos formar la transformación lineal

$$p(A) = \alpha_0 1 + \alpha_1 A + \cdots + \alpha_n A^n.$$

Las reglas para la manipulación algebraica de esos polinomios son fáciles. Así, $p(t)q(t) = r(t)$ implica $p(A)q(A) = r(A)$ (de manera que, en particular, cualquier $p(A)$ y $q(A)$ son conmutativas); si $p(t) = \alpha$ (idénticamente), usualmente escribiremos $p(A) = \alpha$ en vez de $p(A) = \alpha \cdot 1$); esto está en armonía con el uso de los símbolos 0 y 1 para las transformaciones lineales.

Si p es un polinomio en dos variables y si A y B son transformaciones lineales, no es usualmente posible dar una interpretación razonable a $p(A, B)$. La dificultad, por supuesto, es que A y B pueden no conmutar y aun un simple monomio, como s^2t, causará confusión. Si $p(s, t) = s^2t$, ¿qué entenderíamos por $p(A, B)$? ¿Sería A^2B, o ABA, o BA? Es importante reconocer que hay una dificultad aquí; afortunadamente para nosotros no es necesario tratar de vencerla. Trabajaremos con polinomios en varias variables sólo en conexión con transformaciones conmutativas, y entonces todo es sencillo. Observamos que si $AB = BA$, entonces $A^nB^m = B^mA^n$ y, en consecuencia, $p(A, B)$ tiene un significado no ambiguo para todo polinomio p. Las propiedades formales de la correspondencia entre transformaciones (conmutativas) y polinomios son igualmente válidas para varias variables que para una; omitimos los detalles.

Para un ejemplo del posible comportamiento de las potencias de una transformación vemos la transformación de diferenciación D sobre $\mathcal{P}$ (o, justamente igual, sobre $\mathcal{P}_n$, para alguna n). Es fácil ver que para cada entero positivo k, y para todo polinomio x en $\mathcal{P}$, tenemos $(D^kx)(t) = \dfrac{d^kx}{dt^k}$. Observamos que cualquier cosa que D haga, baja el grado del polinomio sobre el cual actúa en exactamente una unidad (suponiendo, por supuesto, que el grado es $\geqq 1$). Sea x un polinomio de grado $n - 1$, digamos; ¿qué es D^nx? O dicho de otro modo: ¿cuál es el producto de las dos transformaciones (conmutativas) D^k y D^{n-k} (donde k es un entero entre 0 y n), considerado sobre el espacio $\mathcal{P}_n$? Mencionamos este ejemplo para poner de manifiesto el hecho desconcertante implicado por la respuesta a la última cuestión; el producto de dos transformaciones puede desaparecer aun cuando ninguna de ellas sea cero. Una transformación no cero, cuyo producto con alguna transformación no cero sea cero, se llama un *divisor de cero*.

EJERCICIOS

1. Calcúlese la transformación lineal D^nS^n y S^nD^n, $n = 1, 2, 3, \ldots$; en otras palabras, compútese el efecto de cada una de esas transformaciones sobre un elemento arbitrario de $\mathcal{P}$. (Aquí D y S denotan las transformaciones de diferenciación e integración definidas en el § 32.)

2. Si A y B son transformaciones lineales tales que $AB - BA$ conmuta con A, entonces $A^kB - BA^k = kA^{k-1}(AB - BA)$ para todo entero positivo k.

3. Supóngase que $Ax(t) = x(t+1)$ para toda x de $\mathcal{P}_n$; pruébese que si D es el operador de diferenciación, entonces

$$1 + \frac{D}{1!} + \frac{D^2}{2!} + \cdots + \frac{D^{n-1}}{(n-1)!} = A.$$

4. (a) Si A es una transformación lineal sobre un espacio vectorial n-dimensional, entonces existe un polinomio p no cero, de grado $\leqq n^2$, tal que $p(A) = 0$.

(b) Si $Ax = y_0(x)x_0$ (véase § 32, (2)), encuéntrese un polinomio no cero p, tal que $p(A) = 0$. ¿Cuál es el grado p más pequeño posible que puede tener?

5. El producto de transformaciones lineales entre diferentes espacios vectoriales se define solamente si "concuerdan" en el sentido siguiente. Supóngase que $\mathcal{U}$, $\mathcal{V}$, y $\mathcal{W}$ son espacios vectoriales sobre el mismo campo y supóngase que A y B son transformaciones lineales de $\mathcal{U}$ a $\mathcal{V}$ y de $\mathcal{V}$ a $\mathcal{W}$, respectivamente. El producto $C = BA$ (el orden es importante) se define como la transformación lineal de $\mathcal{U}$ a $\mathcal{W}$ dada por $Cx = B(Ax)$. Interprétense y pruébense tantas como sean posibles entre las ecuaciones § 34, (1)–(5) para este concepto de multiplicación.

6. Sea A una transformación lineal sobre un espacio vectorial n-dimensional $\mathcal{V}$. (a) Pruébese que el conjunto de todas las transferencias lineales B sobre $\mathcal{V}$ para las cuales $AB = 0$ es un subespacio del espacio de todas las transformaciones lineales sobre $\mathcal{V}$.

(b) Demuéstrese que mediante una adecuada selección de A la dimensión del subespacio descrita en (a) puede hacerse igual a 0, o n, o n^2. ¿Qué valores puede alcanzar esta dimensión?

(c) ¿Puede cada subespacio del espacio de todas las transformaciones lineales obtenerse en la manera descrita en (a) (por selección de una A adecuada)?

7. Sea A una transformación lineal sobre un espacio vectorial $\mathcal{V}$, y considérese la correspondencia que asigna a cada transformación lineal X sobre $\mathcal{V}$ la transformación lineal AX. Pruebe que esta correspondencia es una transformación lineal (sobre el espacio de todas las transformaciones li-

neales). ¿Puede obtenerse de esta manera (por selección de una A adecuada) cada transformación lineal sobre ese espacio?

§36. INVERSOS

En cada una de las dos secciones precedentes dimos un ejemplo; estos dos ejemplos ponen de manifiesto las dos desagradables propiedades que tiene la multiplicación de transformaciones lineales, a saber, la no conmutativa y la existencia de divisores de cero. Pasamos ahora a propiedades más agradables que tienen las transformaciones lineales.

Puede suceder que la transformación lineal A tenga una o ambas de las dos siguientes propiedades muy especiales.

(i) Si $x_1 \neq x_2$, entonces $Ax_1 \neq Ax_2$.

(ii) A cada vector y corresponde (cuando menos) un vector x tal que $Ax = y$.

Si cada A tiene estas dos propiedades, diremos que A es *invertible*. Si A es invertible, definimos una transformación lineal, llamada el *inverso* de A y denotada por A^{-1}, como sigue. Si y_0 es un vector cualquiera, podemos (por (ii)) encontrar una x_0 para la cual $Ax_0 = y_0$. Esta x_0 está, además, determinada en forma única, puesto que $x_0 \neq x_1$ implica (por (i)) que $y_0 = Ax_0 \neq Ax_1$. Definimos $A^{-1}y_0$ como x_0. Para probar que A^{-1} es lineal, evaluamos $A^{-1}(\alpha_1 y_1 + \alpha_2 y_2)$. Si $Ax_1 = y_1$ y $Ax_2 = y_2$, entonces la linearidad de A nos dice que $A(\alpha_1 x_1 + \alpha_2 x_2) = \alpha_1 y_1 + \alpha_2 y_2$, de manera que $A^{-1}(\alpha_1 y_1 + \alpha_2 y_2) = \alpha_1 x_1 + \alpha_2 x_2 = \alpha_1 A^{-1} y_1 + \alpha_2 A^{-1} y_2$.

Como un ejemplo trivial de transformación invertible, mencionamos la transformación de identidad 1; claramente $1^{-1} = 1$. La transformación 0 no es invertible; viola las dos condiciones (i) y (ii) casi tan fuertemente como pueden ser violadas.

Se sigue inmediatamente de la definición que para cualquier A invertible tenemos

$$AA^{-1} = A^{-1}A = 1;$$

demostraremos ahora que estas ecuaciones sirven para caracterizar a A^{-1}.

Teorema 1. *Si A, B y C son transformaciones lineales tales que*

$$AB = CA = 1,$$

entonces A es invertible y $A^{-1} = B = C$.

PRUEBA. Si $Ax_1 = Ax_2$, entonces $CAx_1 = CAx_2$, de manera que (puesto que $CA = 1$) $x_1 = x_2$; en otras palabras, la primera condición de la definición de invertibilidad queda satisfecha. La segunda condición también se satisface, pues si y es cualquier vector y $x = By$, entonces $y = ABy = Ax$. Multiplicando $AB = 1$ sobre la izquierda y $CA = 1$ sobre la derecha, por A^{-1}, vemos que $A^{-1} = B = C$.

Para demostrar que ni $AB = 1$, ni $CA = 1$, es, por sí mismo, suficiente para asegurar la invertibilidad de A, llamamos la atención a las transformaciones de diferenciación e integración D y S, definidas en el § 32, (4) y (5). Aunque $DS = 1$, ni D ni S son invertibles; D viola (i) y S viola (ii).

En espacios finito-dimensionales la situación es mucho más sencilla.

TEOREMA 2. *Una transformación lineal A sobre un espacio vectorial $\mathcal{V}$ es invertible si y sólo si $Ax = 0$ implica que $x = 0$, o, alternativamente, si y sólo si cada y en $\mathcal{V}$ puede ser escrita en la forma $y = Ax$.*

PRUEBA. Si A es invertible, ambas condiciones se satisfacen; esto es trivial. Supóngase ahora que $Ax = 0$ implica que $x = 0$. Entonces, $u \neq v$, esto es, $u - v \neq 0$, implica que $A(u - v) \neq 0$, esto es, que $Au \neq Av$; esto prueba (i). Para probar (ii), sea $\{x_1, \ldots, x_n\}$ una base de $\mathcal{V}$; afirmamos que $\{Ax_1, \ldots, Ax_n\}$ es también una base. De acuerdo con el § 8, Teorema 2, necesitamos solamente probar la dependencia lineal. Pero $\sum_i \alpha_i A x_i = 0$ significa $(\sum_i \alpha_i x_i) = 0$, y, por hipótesis, esto implica que $\sum_i \alpha_i x_i = 0$; la independencia lineal de x_i nos dice ahora que $\alpha_1 = \ldots, \alpha_n = 0$. Se sigue, por supuesto, que todo vector y puede ser escrito en la forma $y = \sum_i \alpha_i A x_i = A(\sum_i \alpha_i x_i)$.

Supongamos, a continuación, que toda y es una Ax y sea $\{y_1, \ldots, y_n\}$ cualquier base de $\mathcal{V}$. Correspondiendo a cada y_i podemos encontrar una x_i (no necesariamente única) para la cual $y_i = Ax_i$; afirmamos que $\{x_1, \ldots, x_n\}$ es también una base. Para $\sum_i \alpha_i x_i = 0$ implica $\sum_i \alpha_i A x_i = \sum_i \alpha_i y_i = 0$, de manera que $\alpha_1 = \ldots = \alpha_n = 0$. Consecuentemente cada x puede ser escrita en la forma $x = \sum_i \alpha_i x_i$, y $Ax = 0$ implica, como en el argumento que se acaba de dar, que $x = 0$.

TEOREMA 3. *Si A y B son invertibles, entonces AB es invertible y $(AB)^{-1} = B^{-1}A^{-1}$. Si A es invertible y $\alpha \neq 0$, entonces αA es invertible y $(\alpha A)^{-1} = \frac{1}{\alpha} A^{-1}$. Si A es invertible, entonces A^{-1} es invertible y $(A^{-1})^{-1} = A$.*

Prueba. De acuerdo con el Teorema 1, es suficiente probar (para el primer enunciado) que el producto de AB con $B^{-1}A^{1-}$, en ambos órdenes, es la identidad; esta verificación la dejamos al lector. Las pruebas de los dos enunciados restantes son idénticas en principio con esta prueba del primer enunciado; el último enunciado, por ejemplo, se sigue del hecho de que las ecuaciones $AA^{-1} = A^{-1}A = 1$ son completamente simétricas en A y A^{-1}.

Concluimos nuestra discusión de inversos con el siguiente comentario, definir funciones racionales de A, cada vez que sea posible, usando A^{-1}. No encontraremos útil hacer esto, excepto en un caso: si A es invertible, entonces sabemos que A^n es también invertible, $n = 1, 2, \ldots$; escribiremos A^{-n} para $(A^n)^{-1}$, de manera que $A^{-n} = (A^{-1})^n$.

EJERCICIOS

1. ¿Cuál de las transformaciones lineales descritas en el § 33, Ej. 1, son invertible?
2. Una transformación lineal A se define sobre $\mathbb{C}^2$ por

$$A(\xi_1, \xi_2) = (\alpha\xi_1 + \beta\xi_2, \gamma\xi_1 + \delta\xi_2),$$

donde α, β, γ y δ son escalares fijos. Pruébese que A es invertible si y sólo si $\alpha\delta - \beta\gamma \neq 0$.

3. Si A y B son transformaciones lineales (sobre el mismo espacio vectorial), entonces una condición necesaria y suficiente para que A y B sean invertibles, es que AB y BA sean invertibles.
4. Si A y B son transformaciones lineales sobre un espacio vectorial finito-dimensional y si $AB = 1$ entonces, tanto A como B sòn invertibles.
5. (a) Si A, B, C y D son transformaciones lineales (todas sobre el mismo espacio vectorial) y si tanto $A + B$ como $A - B$ son invertibles, existen entonces transformaciones lineales X y Y tales que

$$AX + BY = C$$

y

$$BX + AY = D.$$

(b) ¿Hasta qué grado son necesarios en (a) los supuestos de invertibilidad?

6. (a) Una transformación lineal sobre un espacio vectorial finito-dimensional es invertible si y sólo si conserva independencia lineal. Decir que A conserva independencia lineal significa que siempre que $\mathfrak{X}$ es un conjunto linealmente independiente en el espacio $\mathfrak{V}$ con el cual actúa A, entonces $A\mathfrak{X}$ es también un conjunto linealmente independiente de $\mathfrak{V}$. (El símbolo $A\mathfrak{X}$ denota, por supuesto, el conjunto de todos los vectores de la forma Ax, con x en $\mathfrak{X}$.)

(b) ¿Es necesario el supuesto de dimensionalidad finita para la validez de (a)?

7. Demuéstrese que si A es una transformación lineal tal como $A^2 - A + 1 = 0$, entonces A es invertible.

8. Si A y B son transformaciones lineales (sobre el mismo espacio vectorial) y si $AB = 1$, entonces A es llamada *inverso izquierdo* de B, y B es llamado un *inverso derecho* de A. Pruébese que si A tiene exactamente un inverso derecho, digamos B, entonces A es invertible. (Sugestión: considérese $BA + B - 1$).

9. Si A es una transformación lineal invertible sobre un espacio vectorial finito-dimensional $\mathcal{V}$, existe entonces un polinomio p, tal que $A^{-1} = p(A)$. (Sugestión: encuéntrese un polinomio q, no cero, de grado mínimo, tal que $q(A) = 0$ y pruébese que su término constante no puede ser 0).

10. Diséñese una definición razonable de invertibilidad para transformaciones lineales de un espacio vectorial a otro. Usando esta definición, decídase cuáles (si es que alguna) de las transformaciones lineales descritas en el § 33, Ej. 3, son invertibles.

§37. MATRICES

Recojamos ahora algunos cabos sueltos; habiendo introducido el nuevo concepto de transformación lineal, debemos encontrar lo que tiene que ver con los viejos conceptos de bases, funcionales lineales, etc.

Una de las más importantes herramientas en el estudio de las transformaciónes lineales sobre espacios vectoriales finito-dimensionales es el concepto de una matriz. Puesto que este concepto no tiene usualmente un análogo conveniente en espacios finito-dimensionales, y puesto que es posible en la mayor parte de las consideraciones prescindir del mismo, trataremos de no usarlo en la prueba de los teoremas. Es, sin embargo, importante, saber qué es una matriz; entramos ahora en la discusión detallada.

Definición. Sea $\mathcal{V}$ un espacio vectorial finito-dimensional, sea $\mathfrak{X} = \{x_1, \ldots, x_n\}$ una base cualquiera de $\mathcal{V}$, y sea A una transformación lineal sobre $\mathcal{V}$. Puesto que todo vector es una combinación lineal de las x_i, tenemos en particular

$$Ax_j = \sum_i \alpha_{ij} x_i$$

para $j = 1, \ldots, n$. El conjunto (α_{ij}) de n^2 escalares, que llevan el doble subíndice i, j, es la *matriz* de A en el sistema de coordenadas $\mathfrak{X}$; generalmente lo denotaremos por $[A]$, o, si se hace necesario, indicar la base particular $\mathfrak{X}$ que se considera, por $[A; \mathfrak{X}]$. Una matriz (α_{ji}) se escribe usualmente mediante una ordenación cuadrada:

$$[A] = \begin{bmatrix} \alpha_{11} & \alpha_{12} & \cdots & \alpha_{1n} \\ \alpha_{21} & \alpha_{22} & \cdots & \alpha_{2n} \\ \vdots & \vdots & & \vdots \\ \alpha_{n1} & \alpha_{n2} & \cdots & \alpha_{nn} \end{bmatrix};$$

los escalares $(\alpha_{i1}, \ldots, \alpha_{in})$ una *fila* es $(\alpha x_{1j}, \ldots, \alpha_{nj})$ una *columna* de $[A]$.

Esta definición no define el concepto de "matriz"; define "la matriz asociada en ciertas condiciones con una transformación lineal". Es útil, a menudo, considerar una matriz como algo existente por su propio derecho, como una ordenación cuadrada de escalares; en general, una matriz en este libro estará ligada a una transformación lineal y a una base.

Comentamos sobre la notación. Es costumbre usar el mismo símbolo, digamos, A, para la matriz y para la transformación. La justificación para esto se encuentra en la discusión que sigue (de propiedades de matrices). No seguimos aquí esta costumbre, porque una de nuestras principales metas, en conexión con las matrices, es poner de relieve que dependen de un sistema de coordenadas (mientras que no es así por lo que hace a la noción de transformación lineal) y estudiar cómo cambia la relación entre matrices y transformaciones lineales, al pasar de un sistema de coordenadas a otro.

Llamamos también la atención a una peculiaridad de los índices de los elementos α_{ij} de una matriz $[A]$. Una base es una base y hasta ahora, aunque generalmente dábamos a sus elementos como índices los primeros n enteros positivos, el orden de los elementos en la misma carecía enteramente de importancia. Es costumbre, sin embargo, cuando se habla de matrices, referirse a, digamos, la primera fila o la primera columna. Este lenguaje se justifica solamente si pensamos en los elementos de la base $\mathfrak{X}$ como ordenados en un orden definido. Puesto que en la mayoría de nuestras consideraciones el orden de las filas y de las columnas de una matriz es tan irrelevante como el orden de los elementos de una base, no incluimos este aspecto de las matrices en nuestra definición. Es importante, sin embargo, advertir que la apariencia de la ordenación cuadrada asociada con $[A]$ varía con la ordenación de $\mathfrak{X}$. Todo lo que diremos sobre matrices puede, en consecuencia, ser interpretado desde dos puntos de vista diferentes; o en estricto acuerdo con la letra de nuestra definición o, de otro modo, sgiuiendo una definición modificada que hace corresponder una matriz (con filas y columnas ordenadas) no meramente a una transformación lineal y a una base, sino también a una ordenación de la base.

Un palabra más a los que conocen. Es una perversidad, no del autor, sino de la naturaleza, lo que nos hace escribir

$$Ax_j = \sum_i \alpha_{ij}x_i,$$

en vez de la ecuación más usual

$$Ax_i = \sum_j \alpha_{ij}x_j.$$

La razón es que deseamos que las fórmulas de multiplicación matricial y de aplicación de matrices a vectores numéricos (esto es, vectores $(\xi_1, \ldots, \xi_n)$ en $\mathcal{C}^n$) aparezcan normales y que en algún punto del proceso de pasar de vectores a sus coordenadas se inviertan los índices. Para enunciar explícitamente nuestra regla: escríbase Ax_j como una combinación lineal de $x_1, \ldots, x_n$, y escríbanse los coeficientes así obtenidos como la *columna* j-ésima de la matriz $[A]$. (El primer índice de α_{ij} es siempre el índice de la fila; el segundo, el de la columna.

Para un ejemplo, consideremos la transformación de diferenciación D sobre el espacio $\mathcal{P}_n$, y la base $\{x_1, \ldots, x_n\}$ definida por $x_i(t) = t^{i-1}$, $i = 1, \ldots, n$. ¿Cuál es la matriz de D en esta base? Tenemos

$$\begin{aligned}
Dx_1 &= 0x_1 + 0x_2 + \cdots + \quad\quad 0x_{n-1} + 0x_n \\
Dx_2 &= 1x_1 + 0x_2 + \cdots + \quad\quad 0x_{n-1} + 0x_n \\
Dx_3 &= 0x_1 + 2x_2 + \cdots + \quad\quad 0x_{n-1} + 0x_n \\
&\quad\cdot \qquad\qquad \cdots \qquad\qquad \cdot \\
&\quad\cdot \\
Dx_n &= 0x_1 + 0x_2 + \cdots + (n-1)x_{n-1} + 0x_n,
\end{aligned} \tag{1}$$

de manera que

$$[D] = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 2 & \cdots & 0 & 0 \\ \cdot & & & & & \cdot \\ \vdots & & & & & \vdots \\ 0 & 0 & 0 & \cdots & 0 & n-1 \\ 0 & 0 & 0 & \cdots & 0 & 0 \end{bmatrix}. \tag{2}$$

El desagradable fenómeno de los índices que se invierten se ve comparando (1) con (2).

§38. MATRICES DE TRANSFORMACIONES

Hay ahora cierta cantidad de trabajo rutinario que hacer, la mayor parte del cual se dejará a la imaginación. El problema es éste: en un sistema fijo de coordenadas $\mathfrak{X} = \{x_1, \ldots, x_n\}$, conociendo las matrices de A y B, ¿cómo podemos encontrar las matrices de $\alpha A + \beta B$, de AB, de 0, 1, etc.?

Escríbase $[A] = (\alpha_{ij})$, $[B] = (\beta_{ij})$, $C = \alpha A + \beta B$, $[C] = (\gamma_{ij})$; afirmamos que

$$\gamma_{ij} = \alpha\alpha_{ij} + \beta\beta_{ij};$$

también si $[0] = (o_{ij})$ y $[1] = (e_{ij})$, entonces

$$o_{ij} = 0$$

y

$$e_{ij} = \delta_{ij} \text{ (= delta de Kronecker).}$$

Una regla más complicada es la siguiente: si $C = AB$, $[C] = (\gamma_{ij})$, entonces

$$\gamma_{ij} = \sum_k \alpha_{ik}\beta_{kj}.$$

Para probar esto, usamos la definición de matriz asociada con una transformación y la manipulamos así:

$$\begin{aligned} Cx_j = A(Bx_j) = A(\textstyle\sum_k \beta_{kj}x_k) &= \textstyle\sum_k \beta_{kj}Ax_k \\ = \textstyle\sum_k \beta_{kj}(\sum_i \alpha_{ik}x_i) &= \textstyle\sum_i (\sum_k \alpha_{ik}\beta_{kj})x_i. \end{aligned}$$

La relación entre transformaciones y matrices es exactamente la misma que la relación entre vectores y sus coordenadas, y el análogo del teorema de isomorfismo del § 9 es verdadero en el mejor sentido posible. Daremos precisión a estos enunciados.

Con ayuda de una base fija $\mathfrak{X}$, hemos hecho corresponder una matriz $[A]$ a cada transformación lineal A; la correspondencia es descrita por las relaciones $Ax_j = \sum_i \alpha_{ij}x_i$. Afirmamos ahora que esta correspondencia es uno a uno (esto es, que las matrices de dos transformaciones diferentes son diferente) y que toda ordenación (α_{ij}) de n^2 es la matriz de alguna transformación. Para probar esto, observamos, en primer lugar, que el conocimiento de la matriz de A determina completamente a A (esto es, que Ax está, por lo mismo, definido en forma única para cada x), como sigue: si $x = \sum_j \xi_j x_j$, entonces $Ax = \sum_j \xi_j A x_j = \sum_j \xi_j (\sum_i \alpha_{ij} x_i) = \sum_i (\sum_j \alpha_{ij} \xi_j) x_i$. (En otras palabras, si $y = Ax = \sum_i \eta_i x_i$, entonces

$$\eta_i = \sum_j \alpha_{ij} \xi_j.$$

Compárese esto con los comentarios en el § 37 sobre la perversidad de índices). En segundo lugar, no hay ley contra la lectura hacia atrás de la región $Ax_j = \sum_i \alpha_{ij} x_i$ Si, en otras palabras, (α_{ij}) es una ordenación cualquiera, podemos usar esta relación para definir una transformación lineal A; está claro que la matriz será exactamente (α_{ij}). (Una vez más, sin embargo, ponemos de relieve el hecho fundamental de que esta correspondencia uno a uno entre transformaciones y matrices fue establecida por medio de un sistema de coordenadas a otro, la misma transformación lineal puede corresponder a varias matrices y una matriz puede ser la correspondiente a muchas transformaciones lineales). El siguiente enunciado resume la parte esencial de la discusión precedente.

TEOREMA. *Entre el conjunto de todas las matrices* (α_{ij}), (β_{ij}), *etc.*, $i, j = 1, \ldots, n$ *(no consideradas en relación con transformaciones lineales), definimos la suma, la multiplicación, el producto,* (o_{ij}), *y* (e_{ij}), *por*

$$(\alpha_{ij}) + (\beta_{ij}) = (\alpha_{ij} + \beta_{ij}),$$
$$\alpha(\alpha_{ij}) = (\alpha\alpha_{ij}),$$
$$(\alpha_{ij})(\beta_{ij}) = (\textstyle\sum_k \alpha_{ik}\beta_{kj}),$$
$$o_{ij} = 0, \quad e_{ij} = \delta_{ij}.$$

Entonces la correspondencia (establecida por medio de un sistema arbitrario de coordenadas $\mathfrak{X} = \{x_1, \ldots, x_n\}$ *del espacio vectorial n-dimensional* $\mathfrak{V}$*), entre todas las transformaciones lineales* A *sobre* $\mathfrak{V}$ *y todas las matrices* (α_{ij}), *descritas por* $Ax_j = \sum_i \alpha_{ij} x_i$, *es un isomorfismo; en otras palabras, es una correspondencia uno a uno que preserva la suma, la multiplicación escalar, el producto, 0 y 1.*

Hemos evitado cuidadosamente discutir la matriz de A^{-1}. Es posible dar una expresión para $[A^{-1}]$ en términos de los elementos α_{ij}

de $[A]$, pero la expresión no es tan sencilla y, afortunadamente, no es útil para nosotros.

EJERCICIOS

1. Sea A la transformación lineal sobre $\mathcal{P}_n$ definida por $(Ax)(t) = x(t+1)$ y sea $\{x_0, \ldots, x_{n-1}\}$ la base de $\mathcal{P}_n$ definida por $x_j(t) = t^j$, $j = 0, \ldots, n-1$. Encuéntrese la matriz de A con respecto de esta base.

2. Encuéntrese la matriz de la operación de conjugación sobre $\mathcal{C}$, considerando como un espacio vectorial real, con respecto de la base $\{1, i\}$ (donde $i = \sqrt{-1}$).

3. (a) Sea π una permutación de los enteros $1, \ldots, n$; si $x = (\xi_1, \ldots, \xi_n)$ es un vector en $\mathcal{C}^n$, escríbase $Ax = (\xi_{\pi(1)}, \ldots, \xi_{\pi(n)})$. Si $x_i = (\delta_{i1}, \ldots, \delta_{in})$, encuéntrese la matriz de A con respecto de $\{x_1, \ldots, x_n\}$.

(b) Encuéntrense todas las matrices que conmutan con la matriz de A.

4. Considérese el espacio vectorial de todas las matrices reales dos por dos y sea A la transformación lineal sobre este espacio que envía cada matriz X hacia PX, donde $P = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$. Encuéntrese la matriz de A $\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$, con respecto de la base consistente en $\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$.

5. Considérese el espacio vectorial consistente en todas las transformaciones lineales sobre el espacio vectorial $\mathcal{V}$, y sea A la transformación de multiplicación (izquierda) que envía cada transformación sobre X hacia PX, donde P es alguna transformación prescrita sobre $\mathcal{V}$. ¿Bajo qué condiciones sobre P es A invertible?

6. Pruébese que I, J y K son las matrices complejas

$$\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix}, \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}$$

respectivamente (donde $i = \sqrt{-1}$), entonces $I^2 = J^2 = K^2 = -1$, $IJ = -JI = K$, $JK = -KJ = I$ y $KI = -IK = J$.

7. (a) Pruébese que si A, B y C son transformaciones lineales sobre un espacio vectorial bidimensional, entonces $(AB - BA)^2$ conmuta con C.

(b) ¿Es la conclusión de (a) verdadera para espacios de más dimensiones?

8. Sea A la transformación lineal sobre $\mathcal{C}^2$ definida por $A(\xi_1, \xi_2) = (\xi_1 + \xi_2, \xi_2)$. Pruébese que si una transformación lineal B conmuta con A, entonces existe un polinomio p tal que $B = p(A)$.

9. ¿Para cuál de los siguientes polinomios p y matrices A es verdadero que $p(A) = 0$?

(a) $p(t) = t^3 - 3t^2 + 3t - 1$ $A = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$.

(b) $p(t) = t^2 - 3t$, $A = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}$.

(c) $p(t) = t^3 + t^2 + t + 1$, $A = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 & 1 \\ 0 & 1 & 1 \end{pmatrix}$.

(d) $p(t) = t^3 - 2t$, $A = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$.

10. Pruébese que si A y B son las matrices complejas

$$\begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{bmatrix} \quad \text{y} \quad \begin{bmatrix} i & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -i & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

respectivamente, (donde $i = \sqrt{-1}$), y si $C = AB - iBA$, entonces $C^3 + C^2 + C = 0$.

11. Si A y B son transformaciones lineales sobre un espacio vectorial, y si $AB = 0$, se sigue que $BA = 0$?

12. ¿Qué sucede a la matriz de una transformación lineal sobre un espacio finito dimensional cuando los elementos de la base con respecto a la cual la matriz se computa se permutan entre sí mismo?

13. (a) Supóngase que $\mathcal{V}$ es un espacio vectorial finito-dimensional con base $\{x_1, \ldots, x_n\}$. Supóngase que $\alpha_1, \ldots, \alpha_n$ son escalares distintos por pares. Si A es una transformación lineal tal que $Ax_j = \alpha_j x_j$, $j = 1, \ldots, n$, y si B es una transformación lineal que conmuta con A, entonces existen escalares $\beta_1, \ldots, \beta_n$ tales que $Bx_j = \beta_j x_j$.

(b) Pruébese que si B es una transformación lineal sobre un espacio vectorial finito-dimensional $\mathcal{V}$, y si B conmuta con dada transformación lineal sobre $\mathcal{V}$ entonces B es un escalar (esto es (existe un escalar β tal que $Bx = \beta x$ para toda x de $\mathcal{V}$).

14. Si $\{x_1, \ldots, x_k\}$ y $\{y_1, \ldots, y_k\}$ son con juntos linealmente independientes de vectores en un espacio finito-dimensional $\mathcal{V}$, entonces existe una transformación lineal invertible A sobre $\mathcal{V}$ tal que $Ax_j = y_j$, $j = 1, \ldots, k$.

15. Si una matriz $[A] = (\alpha_{ij})$ es tal que $\alpha_{ii} = 0$, $i = 1, \ldots, n$, entonces existen matrices $[B] = (\beta_{ij})$ y $[C] = (\gamma_{ij})$ tales que $[A] = [B][C] - [C][B]$. (Sugestión: pruébese $(\beta_{ij}) = \beta_i \delta_{ij}$).

16. Decídase cuál de las siguientes matrices son invertibles y encuéntrense las inversas de las que lo son.

(a) $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$.

(b) $\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$.

(c) $\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$.

(e) $\begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}$.

(f) $\begin{pmatrix} 1 & 0 & 1 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \end{pmatrix}$.

(d) $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$.

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

17. ¿Para cuáles valores de α son invertibles las siguientes matrices? Encuéntrense las inversas siempre que sea posible.

(a) $\begin{pmatrix} \alpha & 1 \\ 1 & 0 \end{pmatrix}$.

(c) $\begin{pmatrix} 1 & \alpha \\ 1 & \alpha \end{pmatrix}$.

(b) $\begin{pmatrix} 1 & \alpha \\ 1 & 0 \end{pmatrix}$.

(d) $\begin{pmatrix} 1 & 1 \\ 1 & \alpha \end{pmatrix}$.

18. ¿Para cuáles valores de α son invertibles las siguientes matrices? Encuéntrense las inversas siempre que sea posible.

(a) $\begin{pmatrix} 1 & \alpha & 0 \\ \alpha & 1 & \alpha \\ 0 & \alpha & 1 \end{pmatrix}$.

(c) $\begin{pmatrix} 0 & 1 & \alpha \\ 1 & \alpha & 0 \\ \alpha & 0 & 1 \end{pmatrix}$.

(b) $\begin{pmatrix} \alpha & 1 & 0 \\ 1 & \alpha & 1 \\ 0 & 1 & \alpha \end{pmatrix}$.

(d) $\begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & \alpha \\ 1 & \alpha & 1 \end{pmatrix}$.

19. (a) Es fácil extender la teoría matricial a las transformaciones lineales entre diferentes vectores espaciales. Supóngase que $\mathcal{U}$ y $\mathcal{V}$ son espacios vectoriales sobre el mismo campo, sean $\{x_1, \ldots, x_n\}$ y $\{y_1, \ldots, y_m\}$ bases de $\mathcal{U}$ y $\mathcal{V}$ respectivamente y sea A una transformación lineal de $\mathcal{U}$ a $\mathcal{V}$. La matriz de A es, por definición, la ordenación rectangular, m por n, de escalares, definida por

$$Ax_j = \sum_i \alpha_{ij} y_i.$$

Defínase la adición y la multiplicación de matrices rectangulares de manera que se generalicen tantos como sean posibles de los resultados del §38. (Nótese que el producto de una matriz m_1 por n_1 y una matriz m_2 por n_2, en ese orden, se define como $n_1 = m_2$).

(b) Supóngase que A y B son matrices multiplicables. Divídase A en cuatro bloques rectangulares (superior izquierdo, superior derecho, inferior izquierdo, inferior derecho) y entonces divídase B de modo semejante, de manera que el número de columnas en el extremo superior izquierdo de A sea el mismo que el número de filas en el extremo superior izquierdo de B. Si, en una obvia notación taquigráfica, estas divisiones se indican por

$$A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}, \quad B = \begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix},$$

entonces

$$AB = \begin{pmatrix} A_{11}B_{11} + A_{12}B_{21} & A_{11}B_{12} + A_{12}B_{22} \\ A_{21}B_{11} + A_{22}B_{21} & A_{21}B_{12} + A_{22}B_{22} \end{pmatrix}.$$

(c) Usense subespacios y complementos para expresar el resultado de (b) en términos de transformaciones lineales (en vez de matrices).

(d) Generalícense (b) y (c) a un mayor número de partes (en vez de cuatro).

§39. INVARIANCIA

Una posible relación entre subespacios $\mathfrak{M}$ de un espacio vectorial y transformaciones lineales A sobre ese espacio es la invariancia. Decimos que $\mathfrak{M}$ es *invariante* bajo A si x en $\mathfrak{M}$ implica que Ax esté en $\mathfrak{M}$ (Obsérvese que se requiere la relación de implicación solamente en una dirección; no suponemos que toda y de $\mathfrak{M}$ pueda escribirse en la forma $y = Ax$ con x en $\mathfrak{M}$; ni siquiera suponemos que Ax en $\mathfrak{M}$ implica x en $\mathfrak{M}$. Inmediatamente veremos ejemplos en que definitivamente no se dan las condiciones que no supusimos). Sabemos que un subespacio de un espacio vectorial es él mismo un espacio vectorial; si sabemos que $\mathfrak{M}$ es invariante bajo A, podemos pasar por alto el hecho de que A está definida fuera de $\mathfrak{M}$ y podemos considerar A como una transformación lineal definida sobre el espacio vectorial. La invariancia es a menudo considerada para conjuntos de transformaciones lineales, así como una sola; $\mathfrak{M}$. es invariante bajo un conjunto, si es invariante bajo cada miembro del conjunto.

¿Qué puede decirse sobre la matriz de una transformación lineal A sobre un espacio vectorial n-dimensional $\mathcal{V}$ si sabemos que algún $\mathfrak{M}$ es invariante bajo A? En otras palabras: ¿hay un modo hábil de escoger una base $\mathfrak{X} = \{x_1, \ldots, x_n\}$ en $\mathcal{V}$ de manera que $[A] = [A; \mathfrak{X}\ ]$ tenga alguna forma particularmente sencilla? La respuesta está en el §12, Teorema 2; podemos escoger $\mathfrak{X}$ de manera que $x_1, \ldots, x_m$ estén en $\mathfrak{M}$ y $x_{m+1}, \ldots, x_n$ no estén. Expresemos Axj en términos de $x_1, \ldots, x_n$. Para $m + 1 \leqq j \leqq n$, no hay mucho que podamos decir: $Ax_j = \Sigma_i \alpha_{ij} x_i$. Para $1 \leqq j \leqq m$, sin embargo, x_j está en $\mathfrak{M}$, y, en consecuencia, (puesto que $\mathfrak{M}$. es invariante bajo A)Ax_j está en $\mathfrak{M}$. Consecuentemente, en este caso Ax_j es una combinación lineal de $x_1, \ldots, x_m$; las α_{ij} con $m + 1 \leqq i \leqq n$ son cero. De aquí la matriz $[A]$ de A, en este sistema de coordenadas tenga la forma

$$[A] = \begin{pmatrix} [A_1] & [B_0] \\ [0] & [A_2] \end{pmatrix},$$

donde $[A_1]$ la matriz (de m filas) de A, considerada como una transformación lineal sobre el espacio $\mathfrak{M}$ con respecto del sistema de coordenadas $\{x_1, \ldots, x_m\}$, $[A_2]$ y $[B_0]$ son algunas ordenaciones de escalares (en tamaño $(n-b)$ por $(n-m)$ y m por $(n-m)$ respectivamente) y 0 denota la ordenación rectangular $((n-m)$ por $m)$, consiste únicamente en ceros. (Es importante observar el hecho desagradable de que $[B_0]$ no necesita ser cero).

§40. REDUCIBILIDAD

Un subcaso particularmente importante de la noción de invariancia es el de reducibilidad. Si $\mathfrak{M}$ y $\mathfrak{N}$ son dos subespacios tales que ambos son invariantes bajo A y tales que $\mathfrak{V}$ es su suma directa, entonces A es *reducido* (descompuesto) por el par $(\mathfrak{M}, \mathfrak{N})$. La diferencia entre invariancia y reducibilidad consiste en que, en el primer caso, entre la colección de todos los subespacios invariantes bajo A podemos no estar en la posibilidad de escoger dos cualesquiera, distintos de $\mathfrak{O}$ y de $\mathfrak{V}$, con la propiedad de que $\mathfrak{V}$ sea su suma directa. O, diciéndolo de otro modo, si $\mathfrak{M}$ es invariante bajo A, hay, de seguro, muchos modos de determinar una $\mathfrak{N}$ tal que $\mathfrak{V} = \mathfrak{M} \oplus \mathfrak{N}$, pero puede suceder que ninguna de esas $\mathfrak{N}$ sea invariante bajo A.

El proceso descrito puede ser también invertido. Sean $\mathfrak{M}$ y $\mathfrak{N}$ dos vectores espaciales cualesquiera, y sean A y B dos transformaciones lineales cualesquiera (sobre $\mathfrak{M}$ y $\mathfrak{N}$ respectivamente). Sea $\mathfrak{V}$ la suma directa $\mathfrak{M} \oplus \mathfrak{N}$; podemos definir sobre $\mathfrak{V}$ una transformación lineal C llamada la suma directa de A y de B, escribiendo

$$Cz = C(x, y) = (Ax, By).$$

Omitiremos la discusión detallada de las sumas directas de transformaciones; mencionaremos solamente los resultados. Sus pruebas son fáciles. Si $(\mathfrak{M}, \mathfrak{N})$ reduce a C, y si denotamos por A la transformación lineal C considerada solamente sobre $\mathfrak{N}$ entonces C es la suma directa de A y de B. Por una selección de base (a saber, escogiendo $x_1, \ldots, x_m$ en $\mathfrak{M}$ y $x_{m+1}, \ldots, x_n$ en $\mathfrak{N}$) podemos poner la matriz de la suma directa de A y B en la forma mostrada en la sección precedente, con $[A_1] = [A]$, $[B_0] = 0$, y $[A_2] = [B]$. Si p es un polinomio cualquiera, y si escribimos $A' = p(A)$, $B' = p(B)$, entonces la suma directa C' de A' y B' será $p(C)$.

EJERCICIOS

1. Supóngase que la matriz de una transformación lineal (sobre un espacio vectorial de dos dimensiones), con respecto de algún sistema de coordenadas es $\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$. ¿Cuántos subespacios hay invariantes bajo la transformación?

2. Dé un ejemplo de una transformación lineal A sobre un espacio vectorial finito-dimensional $\mathcal{V}$ tal que $\mathcal{O}$ y $\mathcal{V}$ son los únicos subespacios invariantes bajo A.

3. Sea D el operador de diferenciación sobre $\mathcal{P}_n$. Si $m \leq n$, entonces el subespacio $\mathcal{P}_m$ ¿es invariante bajo D? ¿Es D sobre $\mathcal{P}_m$ invariable, ¿Hay un complemento de $\mathcal{P}_m$ en $\mathcal{P}_n$ tal que junto con $\mathcal{P}_m$ reduzca a D?

4. Pruébese que el subespacio sobretendido por dos subespacios, cada uno de los cuales es invariante bajo alguna transformación lineal A, es él mismo invariante bajo A.

§41. PROYECCIONES

Especialmente importante para nuestro propósito es otra conexión entre sumas directas y transformaciones lineales.

DEFINICIÓN. Si $\mathcal{V}$ es la suma directa de $\mathcal{M}$ y $\mathcal{N}$, de manera que toda z de $\mathcal{V}$ pueda ser descrita, únicamente en la forma $z = x + y$, con x en $\mathcal{M}$ y y en $\mathcal{N}$, la *proyección* sobre $\mathcal{M}$ a lo largo de $\mathcal{N}$ es la transformación E, definida por $Ez = x$.

Si las sumas direct: ; son importantes, entonces las ,proyecciones lo son también, puesto que, como veremos, son una herramienta algebraica muy poderosa para el estudio del concepto geométrico de suma directa. El lector se satisfará sobre la razón para usar la palabra "proyección", trazando un par de ejes (múltiples lineales) en el plano (suma directa). Para hacer que la imagen aparezca suficientemente general, no tiene ejes perpendiculares.

Pasamos por alto un punto cuya prueba es lo bastante fácil para ser omitida, pero cuya existencia debe reconocerse; debe mostrarse que E es una transformación *lineal*. Dejamos esta demostración al lector, y pasamos a buscar propiedades especiales de las proyecciones.

TEOREMA 1. *A Una transformación lineal E es una proyección sobre algún subespacio si y sólo si es idempotente, esto es,* $E^2 = E$.

PRUEBA. Si E es la proyección sobre $\mathcal{M}$ a lo largo de $\mathcal{N}$, y si $z = x + y$, con x en $\mathcal{M}$ y en $\mathcal{N}$, entonces la descomposición de x es $x + 0$, de manera que

$$E^2z = EEz = Ex = x = Ez.$$

Inversamente, supóngase que $E^2 = E$. Sea $\mathfrak{N}$ el conjunto de todos los vectores z de $\mathfrak{V}$ para los cuales $E = 0$; sea $\mathfrak{M}$ el conjunto de todos los vectores z para los cuales $Ez = z$. Está claro que tanto $\mathfrak{M}$ como $\mathfrak{N}$ son subespacios. Probaremos que $\mathfrak{V} = \mathfrak{M} \oplus \mathfrak{N}$. En vista del teorema del § 18, necesitamos probar que $\mathfrak{M}$ y $\mathfrak{N}$ son disjuntos y que juntos sobretienden a $\mathfrak{V}$.

S. z está en $\mathfrak{M}$, entonces $Ez = z$, si z está en $\mathfrak{N}$, entonces $Ez = 0$; por lo que, si z está tanto en como en $\mathfrak{N}$, entonces $z = 0$. Para una z arbitraria tenemos

$$z = Ez + (1 - E)z.$$

Si escribimos $Ez = x$ y $(1 - E)z = y$, entonces $Ex = E^2z = Ez = x$, y $Ey = E(1 - E)z = Ez - E^2z = 0$, de manera que x está en $\mathfrak{M}$ y y está en $\mathfrak{N}$. Esto prueba que $\mathfrak{V} = \mathfrak{M} \oplus \mathfrak{N}$, y que la proyección sobre $\mathfrak{M}$ a lo largo de $\mathfrak{N}$ es precisamente E.

Como consecuencia inmediata de la prueba anterior, obtenemos también el siguiente resultado.

Teorema 2. *Si E es la proyección $\mathfrak{M}$ a lo largo de $\mathfrak{N}$, entonces $\mathfrak{M}$ y $\mathfrak{N}$ son, respectivamente, los conjuntos de todas las soluciones de las ecuaciones $Ez = z$ y $Ez = 0$.*

Por medio de estos dos teoremas podemos eliminar la aparente asimetría, en la definición de proyecciones, entre los papeles desempeñados por $\mathfrak{M}$ y $\mathfrak{N}$. Si a toda $z = x + y$ hacemos corresponder no x, sino y, obtenemos también una transformación lineal idempotente. Esta transformación (a saber, $1 - E$) es la proyección de $\mathfrak{N}$ a lo largo de $\mathfrak{M}$. Resumimos los hechos como sigue:

Teorema 3. *Una transformación lineal E es una proyección si y sólo si $1 - E$ es una proyección; si E es la proyección de $\mathfrak{M}$ a lo largo de $\mathfrak{N}$, entonces $1 - E$ es la proyección de $\mathfrak{N}$ a lo largo de $\mathfrak{M}$.*

§42. COMBINACIONES DE PROYECCIONES

Continuando en el espíritu del Teorema 3 de la sección precedente, investigamos las condiciones bajo las cuales varias combinaciones de proyecciones algebraicas son esas mismas proyecciones.

Teorema. *Suponemos que E_1 y E_2 son proyecciones sobre $\mathfrak{M}_1$ y $\mathfrak{M}_2$ a lo largo de $\mathfrak{N}_1$ y $\mathfrak{N}_2$ respectivamente, y que el campo subyacente es tal que $1 + 1 \neq 0$. Hacemos tres aserciones.*

(*i*) *$E_1 + E_2$ es una proyección si y sólo si $E_1E_2 = E_2E_1 = 0$; si esta condición se satisface, entonces $E = E_1 + E_2$ es la prolongación sobre $\mathfrak{M}$ a lo largo de $\mathfrak{N}$, donde* $\mathfrak{M} = \mathfrak{M}_1 \oplus \mathfrak{M}_2$ y $\mathfrak{N} = \mathfrak{N}_1 \cap \mathfrak{N}_2$.

(*ii*) *$E_1 - E_2$ es una proyección si y sólo si $E_1E_2 = E_2E_1 = E_2$; si esta condición se satisface, entonces $E = E_1 - E_2$ es la proyección de $\mathfrak{M}$ a lo largo de $\mathfrak{N}$, donde $\mathfrak{M} = \mathfrak{M}_1 \cap \mathfrak{N}_2$ y $\mathfrak{N} = \mathfrak{N}_1 \oplus \mathfrak{M}_2$.*

(*iii*) *Si $E_1E_2 = E_2E_1 = E$, entonces E es la proyección sobre $\mathfrak{M}$ a lo largo de $\mathfrak{N}$, donde $\mathfrak{M} = \mathfrak{M}_1 \cap \mathfrak{M}_2$ y $\mathfrak{N} = \mathfrak{N}_1 + \mathfrak{N}_2$.*

PRUEBA. Recordaremos la notación. Si $\mathfrak{H}$ y $\mathfrak{K}$ son subespacios entonces $\mathfrak{H} + \mathfrak{K}$ es el subespacio sobretendido por $\mathfrak{H}$ y $\mathfrak{K}$; escribir $\mathfrak{H} \oplus \mathfrak{K}$ implica que $\mathfrak{H}$ y $\mathfrak{K}$ están disjuntos, y entonces $\mathfrak{H} \oplus \mathfrak{K} = \mathfrak{H} + \mathfrak{K}$; y $\mathfrak{H} \cap \mathfrak{K}$ es la intersección de $\mathfrak{H}$ y $\mathfrak{K}$.

(i) Si $E_1 + E_2 = E$ es una proyección, entonces $(E_1 + E_2)^2 = E^2 = E = E_1 + E_2$, de manera que el producto cruz de los términos debe desaparecer:

$$E_1E_2 + E_2E_1 = 0. \tag{1}$$

Si multiplicamos (1) por E_1 tanto a la izquierda como a la derecha, obtenemos,

$$E_1E_2 + E_1E_2E_1 = 0,$$

$$E_1E_2E_1 + E_2E_1 = 0;$$

sustrayendo, obtenemos $E_1E^2 - E_2E_1 = 0$. De aquí que E_1 y E_2 son conmutativos, y (1) implica que su producto es cero. (Aquí es donde necesitamos el supuesto de que $1 + 1 \neq 0$.) Puesto que, inversamente $E_1E_2 - E_2E_1 = 0$ claramente implica (1), vemos que la condición es también suficiente para asegurar que E es una proyección.

Supongamos, de ahora en adelante, que E es una proyección; por el § 41, Teorema 2, $\mathfrak{M}$ y $\mathfrak{N}$ son, respectivamente, los conjuntos de todas las soluciones de las ecuaciones $Ez = z$ y $Ez = 0$. Escribamos $z = x_1 + y_1 = x_2 + y_2$, donde $x_1 = E_1z$ y $x_2 = E_2z$ están en $\mathfrak{M}_1$ y $\mathfrak{M}_2$, respectivamente. Si z está en $\mathfrak{N}_1$ $E_2z + E_2z = z$, entonces

$$z = E_1(x_2 + y_2) + E_2(x_1 + y_1) = E_1y_2 + E_2y_1.$$

Puesto que $E_1(E_1y_2) = E_1y_2$ y $E_2(E_2y_1) = E_2y_1$, hemos exhibido z como una suma de un vector de $\mathfrak{M}_1$ y un vector de $\mathfrak{M}_2$, de manera que $\mathfrak{M} \subset \mathfrak{M}_1 + \mathfrak{M}_2$. Recíprocamente, si z es una suma de un vector de $\mathfrak{M}_1$ y un vector de $\mathfrak{M}_2$ entonces $(E_1 + E_2)z = z$, de manera que z está en $\mathfrak{M}$, y consecuentemente $\mathfrak{M} = \mathfrak{M}_1 + \mathfrak{M}_2$. Finalmente, si z pertenece tanto a $\mathfrak{M}_1$ como a $\mathfrak{M}_2$, de manera que $E_1z = E_2z = z$, entonces $z = E_1z = E_1(E_2z) = 0$, de manera que $\mathfrak{M}_1$ y $\mathfrak{M}_2$ están disjuntos. Hemos probado que $\mathfrak{M} = \mathfrak{M}_1 \oplus \mathfrak{M}_2$.

Falta por determinar $\mathfrak{N}$, esto es, encontrar todas las soluciones $E_1z + E_2z = 0$. Si z está en $\mathfrak{N}_1 \cap \mathfrak{N}_2$, esta ecuación se satisface claramente; recíprocamente $E_1z + E_2z = 0$ implica (por multiplicación a la izquierda de E_1 y E_2, respectivamente,) que $E_1z + E_2E_1z = 0$ y $E_2E_1z + E_2z = 0$. Puesto que $E_1E_2z = E_2E_1z = 0$, para toda z, obtenemos, finalmente, $E_1z = E_2z = 0$, de manera que z pertenece tanto a $\mathfrak{N}_1$ como a $\mathfrak{N}_2$.

Con la técnica y los resultados obtenidos en esta prueba, las pruebas de las partes restantes del teorema son fáciles.

(ii) De acuerdo con el § 41, Teorema 3, $E_1 - E_2$ es una proyección si y sólo si $1 - (E_1 - E_2) = (1 - E_1) + E_2$ es una proyección. De acuerdo con (i) esto sucede (puesto que, por supuesto, $1 - E_1$ es la proyección $\mathfrak{N}_1$ a lo largo de $\mathfrak{M}_1$) si y sólo si

$$(1 - E_1)E_2 = E_2(1 - E_1) = 0,$$

y en este caso $(1 - E_1) + E_2$ es la proyección sobre $\mathfrak{N}_1 \oplus \mathfrak{M}_2$ a lo largo de $\mathfrak{M}_1 \cap \mathfrak{N}_2$. Puesto que (2) es equivalente a $E_1E_2 = E_1E_2 = E_2$, la prueba de (ii) es completa.

(iii) Que $E = E_1E_2 = E_2E_1$ implica que E es una proyección está claro, puesto que E es idempotente. Suponemos, en consecuencia, que E_1 y E_2 conmutante y determinamos $\mathfrak{M}$ y $\mathfrak{N}$. Si $Ez = z$, entonces $E_1z = E_1Ez = E_1E_1E_2z = E_1E_2z = z$, y de modo semejante, $E_2z = z$, de manera que z está contenida tanto en $\mathfrak{M}_1$ como en $\mathfrak{M}_2$. El recíproco está claro; si $E_1z = z = E_2z$, entonces $Ez = z$. Supóngase, a continuación, que $E_1E_2z = 0$; se sigue que E_2z pertenece a $\mathfrak{N}_1$, y de la conmutatividad de E_1 y E_2, que E_1z pertenece a $\mathfrak{N}_2$. Esta es más simetría de la que necesitamos; puesto que $z = E_2z + (1 - E_2)z$, y puesto que $(1 - E_2)z$ está en $\mathfrak{N}_2$, tenemos exhibida z como una suma de un vector de $\mathfrak{N}_1$ y un vector de $\mathfrak{N}_2$. Recíprocamente, si z es una suma tal, entonces $E_1E_2z = 0$; esto concluye la prueba de que $\mathfrak{N} = \mathfrak{N}_1 + \mathfrak{N}_2$.

Volveremos, posteriormente, a teoremas de este tipo y obtendremos, en ciertos casos, resultados más precisos. Antes de dejar el tema, sin embargo, llamamos la atención a unas cuantas peculiaridades de menor cuantía del teorema de esta sección. Observamos, primero, que tanto en (i) como en (ii) una de $\mathfrak{M}$ y $\mathfrak{N}$ fue una suma *directa* de los subespacios dados, en (iii) enunciamos solamente que $\mathfrak{N} = \mathfrak{N}_1 + \mathfrak{N}_2$. La consideración de la posibilidad de $E_1 = E_2 = E$ muestra que esto es inevitable. También se afirmó que la condición de (iii) era sólo suficiente; es posible construir las proyecciones E_1 y E_2, cuyo producto E_1E_2 y es una proyección, pero para la cual E_1E_2 y E_2E_1 son diferentes. Finalmente, puede conjeturarse que es posible extender el resultado de (i) por inducción a más de dos sumandos, Aunque esto es verdad, es sorprendentemente no trivial; lo probaremos posteriormente en un caso de especial interés.

§43. PROYECCIONES E INVARIANCIA

Hemos visto ya que el estudio de las proyecciones es equivalente al estudio de descomposiciones directas de sumas. Por medio de proyecciones podemos también estudiar las nociones de variancia y reducibilidad.

Teorema 1. *Si un subespacio $\mathfrak{M}$ es invariante bajo la transformación lineal A, entonces $EAE = AE$ para toda proyección E sobre $\mathfrak{M}$. Recíprocamente, si $EAE = AE$ para alguna proyección E sobre $\mathfrak{M}$, $\mathfrak{M}$ entonces es invariablemente bajo A.*

Prueba. Supóngase que $\mathfrak{M}$ es invariante bajo A y que $\mathcal{V} = \mathfrak{M} \oplus \mathfrak{N}$ para alguna $\mathfrak{N}$; sea E la proyección sobre $\mathfrak{M}$ a lo largo de $\mathfrak{N}$. Para cualquiera $z = x + y$ (con x en $\mathfrak{M}$ y y en $\mathfrak{N}$) tenemos $AEz = Ax$ y $EAEz = EAx$; puesto que la presencia de x en $\mathfrak{M}$ garantiza la presencia de Ax en $\mathfrak{M}$, se sigue que EAx es también igual a Ax, como se desea.

Recíprocamente, supóngase que $\mathcal{V} = \mathfrak{M} \oplus \mathfrak{N}$, y que $EAE = AE$ para la proyección E sobre $\mathfrak{M}$ a lo largo de $\mathfrak{N}$. Si x está en $\mathfrak{M}$, entonces $Ex = x$, de manera que

$$EAx = EAEx = AEx = Ax,$$

y consecuentemente Ax está también en $\mathfrak{M}$.

Teorema 2. *Si $\mathfrak{M}$ y $\mathfrak{N}$ son subespacios con $\mathcal{V} = \mathfrak{M} \oplus \mathfrak{N}$, entonces una condición necesaria y suficiente para que la transformación lineal A sea reducida por el par $(\mathfrak{M}, \mathfrak{N})$ es que $AE = AE$, donde E es la proyección sobre $\mathfrak{M}$ a lo largo de $\mathfrak{N}$.*

Prueba. Primero suponemos que $EA = .AE$, y probamos que A es reducida por $(\mathfrak{M}$, Si x está en $\mathfrak{N})$. si x está en $\mathfrak{M}$, entonces $Ex = 0$ y $EAx = AEx = A0 = 0$, de manera que Ax está también en $\mathfrak{M}$;

A continuación suponemos que A es reducida por $(\mathfrak{M}, \mathfrak{N})$, y probamos que $EA = AE$. Puesto que $\mathfrak{M}$ es invariante bajo A, el Teorema 1 nos asegura que $EAE = AE$; puesto que $\mathfrak{N}$ es también invariante bajo A, y puesto que $1 - E$ es una proyección sobre $\mathfrak{N}$, tenemos, de modo semejante $(1-E)A(1-E) = A(1-E)$. De la segunda edición, después de efectuar las multiplicaciones indicadas y simplificar, obtenemos $EAE = EA$; esto concluye la prueba del teorema.

EJERCICIOS

1. (a) supóngase que E es una proyección sobre un espacio vectorial $\mathfrak{V}$, y supóngase que se redefine la multiplicación escalar, de manera que el nuevo producto de un escalar α y un vector x es el antiguo producto de α y Ex. Demuéstrese que la (antigua) edición vectcrial y la (nueva) multiplicación escalar satisfacen todos los axiomas sobre un espacio vcctorial, excepto $1 \cdot x = x$.

(b) ¿A qué extremo es verdadero que el método descrito en (a) es el único modo de construir sistemas que satisfagan todos los axiones sobre un espacio vectorial, excepto $1 \cdot x = x$?

2. (a) Supóngase que $\mathfrak{V}$ es un espacio vectorial, x_0 es un vector $\mathfrak{V}$, y y_0 es una funcional lineal sobre $\mathfrak{V}$; escríbase $Ax = [x, y_0]x_0$ para toda x de $\mathfrak{V}$. ¿Bajo qué condiciones impuestas a x_0 y y_0 es A una proyección?

(b) Si A es la proyección sobre, digamos $\mathfrak{M}$ a lo largo de $\mathfrak{N}$, caracterice $\mathfrak{M}$ y $\mathfrak{N}$ en términos de x_0 y y_0.

3. Si A es multiplicación izquierda por P sobre un espacio de transformaciones lineales (véase § 8, Ej. 5), ¿bajo qué condiciones impuestas a P es A una proyección?

4. Si A es una transformación lineal, si E es una proyección y si $F = 1 - E$, entonces

$$A = EAE + EAF + FAE + FAF.$$

Usese este resultado para probar la regla de multiplicación para matrices (cuadradas) subdivididas (como en el § 38, Ej. 19).

5. (a) Si E_1 y E_2 son proyecciones sobre $\mathfrak{M}_1$ y $\mathfrak{M}_2$ a lo largo de $\mathfrak{N}_1$ y $\mathfrak{N}_2$ respectivamente, y si E_1 y E_2 conmutan, entonces $E_1 + E_2 - E_1E_2$ es una proyección.

(b) Si $E_1 + E_2 - E_1E_2$ es la proyección sobre $\mathfrak{M}$ a lo largo de $\mathfrak{N}$ descríbase $\mathfrak{M}$ y $\mathfrak{N}$, en términos de $\mathfrak{M}_1$, $\mathfrak{M}_2$, $\mathfrak{N}_1$, y $\mathfrak{N}_2$.

6. (a) Encuéntrese una transformación lineal A tal que $A^2(1 - A = 0$, pero que A no sea idempotente.

(b) Encuéntrese una transformación lineal A tal que $A(1 - A)^2 = 0$, pero A no es idempotente.

(c) Pruébese que si A es una transformación lineal tal que $A^2(1 - A) = A(1 - A)^2 = 0$, entonces A es idempotente.

7. (a) Pruébese que si E es una proyección sobre un espacio vectorial finito-dimensional, entonces existe una base $\mathfrak{X}$ tal que la matriz (e_{ij}) de E

con respecto de $\mathcal{X}$ tiene la siguiente forma especial: $e_{ij} = 0$ o 1 para toda i y j, y $e_{ij} = 0$, si $i \neq j$.

(b) Una *involución* es una transformación lineal U tal que $U^2 = 1$. Demuéstrese que si $1 + 1 \neq 0$, entonces la ecuación $U = 2E - 1$ establece una correspondencia uno a uno entre todas las proyecciones E y todas las involuciones U.

(c) ¿Qué implican (a) y (b) respecto de la matriz de una involución sobre un espacio vectorial finito-dimensional?

8. (a) En el espacio $\mathcal{C}^2$ de todos los vectores (ξ_1, ξ_2), sean $\mathfrak{M}^+$, $\mathfrak{N}_1$, y $\mathfrak{N}_2$ los subespacios caracterizados por $\xi_1 = \xi_2$, $\xi_1 = 0$ y $\xi_2 = 0$, respectivamente. Si E_1 y E_2 son las proyecciones sobre $\mathfrak{M}^+$ a lo largo de $\mathfrak{N}_1$ y $\mathfrak{N}_2$ respectivamente, demuéstre que $E_1E_2 = E_2$ y $E_2E_1 = E_1$.

(b) Sea $\mathfrak{M}^-$ el subespacio caracterizado por $\xi_1 = -\xi_2$. Si E_0 es la proyección sobre $\mathfrak{N}_2$ a lo largo de $\mathfrak{M}^-$, entonces E_2E_0 es una proyección, pero E_0E_2 no lo es.

9. Demuéstrese que si E, F y G son proyecciones sobre un espacio vectorial sobre un campo cuya característica no es igual a 2, y si $E + F + G = 1$, entonces $EF = FE = EG = GE = FG = GF = 0$.

¿Es válida la prueba para cuatro proyecciones en vez de tres?

§44. ADJUNTOS

Estudiemos a continuación la relación entre las nociones de transformación lineal y espacio dual. Sea $\mathcal{V}$ cualquier espacio vectorial y sea y cualquier elemento de $\mathcal{V}'$; para cualquier transformación lineal A sobre $\mathcal{V}$ consideramos la expresión $[Ax, y]$. Para cada y fija y, la función y' definida por $y'(x) = [Ax, y]$ es una función lineal sobre $\mathcal{V}$; usando la notación de corchetes para y' así como para y, tenemos $[Ax, y] = [x, y']$. Si permitimos que y varíe sobre $\mathcal{V}'$, entonces este procedimiento hace corresponder a cada y una y', que depende, por supuesto, de y; escribimos $y' = A'\, y$. La propiedad definitoria de A' es

$$(1) \qquad [Ax, y] = [x, A'y].$$

Afirmamos que A' es una transformación lineal sobre $\mathcal{V}'$. En realidad, si $y = \alpha_1 y_1 + \alpha_2 y_2$, entonces

$$\begin{aligned}[x, A'y] &= [Ax, y] = \alpha_1[Ax, y_1] + \alpha_2[Ax, y_2] \\ &= \alpha_1[x, A'y_1] + \alpha_2[x, A'y_2] = [x, \alpha_1 A'y_1 + \alpha_2 A'y_2].\end{aligned}$$

La transformación lineal A' es llamada el *adjunto* (o dual) de A; dedicamos esta sección y la siguiente a estudiar las propiedades de A'. Veamos primero las reglas algebraicas formales; van como sigue:

$$(2) \qquad 0' = 0,$$

$$(3) \qquad 1' = 1,$$

$$(4) \qquad (A + B)' = A' + B',$$

$$(5) \qquad (\alpha A)' = \alpha A',$$

$$(6) \qquad (AB)' = B'A',$$

$$(7) \qquad (A^{-1})' = (A')^{-1}.$$

De aquí que (7) deba ser interpretada en el sentido siguiente: si A es invertible, entonces también lo es A', y la ecuación es válida. Las pruebas de estas relaciones son elementales; para indicar el procedimiento, efectuamos los cómputos para (6) y (7). Para probar (6) simplemente observamos que

$$[ABx, y] = [Bx, A'y] = [x, B'A'y].$$

Para probar (7), supóngase que A es invertible, de manera que $AA^{-1} = A^{-1}A = 1$. Aplicando (3) y (6) a estas ecuaciones, obtenemos

$$(A^{-1})'A' = A'(A^{-1})' = 1;$$

El Teorema 1 del §36 implica que A' es invertible y que (7) es válida.

En espacios finito-dimensionales es válida otra importante relación:

$$(8) \qquad A'' = A.$$

Esta relación tiene que leerse con un grano de sal. Tal como está, A'' es una transformación, no sobre $\mathcal{V}$ sino sobre el espacio dual $\mathcal{V}''$ de $\mathcal{V}'$. Si, no obstante, identificamos $\mathcal{V}''$ y $\mathcal{V}$ de acuerdo con el isomorfismo natural, entonces A'' actúa sobre $\mathcal{V}$ y (8) tiene sentido. En esta interpretación la prueba de (8) es trivial. Puesto que $\mathcal{V}$ es reflexiva, obtenemos cada funcional lineal sobre $\mathcal{V}'$ considerando $[x, y]$ como una función de y, con x fija en $\mathcal{V}$. Puesto que $[x, A'y]$ define una función (una funcional lineal) de y, puede ser escrita en la forma $[x', y]$. El vector x' es aquí, por definición, $A''x$, De aquí tenemos, para toda y de $\mathcal{V}'$ y para toda x de $\mathcal{V}$,

$$[Ax, y] = [x, A'y] = [A''x, y];$$

la igualdad del primero y último términos de esta cadena prueba (8).

Bajo la hipótesis de (8) (esto es, finito-dimensionalidad), puede ser eliminada la asimetría en la interpretación de (7); afirmamos que en este caso la invertibilidad de A' implica la de A y, en consecuencia, la validez de (7). Prueba: aplíquese la antigua interpretación de (7) a A' y A'' en lugar de A y A'.

Nuestra discusión se resume, en el caso reflexivo finito-dimensional, por la aserción de que el trazo del mapa de $A \to A'$ es uno a uno, y, de hecho, un antisomorfismo algebraico, desde el conjunto de todas las transformaciones lineales sobre $\mathcal{V}$ hacia y sobre el conjunto de todas las transformaciones lineales sobre $\mathcal{V}'$. (Se puso el prefijo "anti" a causa de la regla de conmutación (6)).

§45. ADJUNTOS DE PROYECCIONES

Hay un caso importante en que la multiplicación no se invierte, esto es, cuando $(AB)' = A'B'$; a saber, el caso cuando A y B conmutan. Tenemos, en particular, $(A^n)' = (A')^n$, y, más generalmente, $(p(A))' = p(A')$ para todo polinomio p. Se sigue de esto que si E es una proyección, entonces también lo es E'. La cuestión se suscita: ¿con cuál descomposición directa de suma ésta asociada E'?

TEOREMA 1. *Si E es la proyección sobre $\mathfrak{M}$ a lo largo de $\mathfrak{N}$, entonces E' es la proyección sobre $\mathfrak{N}^0$ a lo largo de $\mathfrak{M}^0$.*

PRUEBA. Sabemos ya que $(E')^2 = E'$ y $\mathcal{V}' = \mathfrak{N}^0 \oplus \mathfrak{M}^0$ (véase el §20). Es necesario solamente determinar los subespacios consistentes en las soluciones de $E'y = 0$ y $E'y = y$. Esto lo hacemos en cuatro pasos.

(i) Si y está en $\mathfrak{M}^0$, entonces, para toda x

$$[x, E'y] = [Ex, y] = 0,$$

de manera que $E'y = 0$.

(ii) Si $E'y = 0$, entonces, para toda x de $\mathfrak{M}$,

$$[x, y] = [Ex, y] = [x, E'y] = 0,$$

de manera que y está en $\mathfrak{M}^0$.

(iii) Si y está en $\mathfrak{N}^0$, entonces, para toda x

$$[x, y] = [Ex, y] + [(1 - E)x, y] = [Ex, y] = [x, E'y],$$

de manera que $E'y = y$.

(iv) Si $E'y = y$, entonces para toda x en $\mathfrak{N}$,

$$[x, y] = [x, E'y] = [Ex, y] = 0,$$

de manera que ya está en $\mathfrak{N}^0$.

Los pasos (i) y (ii) juntos muestran que el conjunto de soluciones de $E'y = 0$ es precisamente $\mathfrak{M}^0$; los pasos (iii) y (iv) juntos muestran que el conjunto de soluciones de $E'y = y$ es precisamente $\mathfrak{N}^0$. Esto concluye la prueba del teorema.

TEOREMA 2. *Si $\mathfrak{M}$ es invariante bajo A, entonces $\mathfrak{M}^0$ es invariante bajo A'; si A es reducida por $(\mathfrak{M}, \mathfrak{N})$, entonces A' es reducida por $(\mathfrak{M}^0, \mathfrak{N}^0)$.*

PRUEBA. Probaremos solamente el primer enunciado; el segundo se sigue claramente del primero. Observamos primero la siguiente identidad, válida para cualesquiera tres transformaciones lineales E, F y A, sujetas a la relación $F = 1 - E$;

(1) $$FAF - FA = EAE - AE.$$

(Compárese esto con la prueba del §43, Teorema 2). Sea E cualquier proyección sobre $\mathfrak{M}$; por el §43, Teorema 1, desaparece el primer miembro de (1) y, en consecuencia, el segundo miembro. Tomando adjuntos, obtenemos $F'A'F' = A'F'$; puesto que, por el Teorema 1 de la presente sección, $F' = 1 - E'$ es una proyección sobre $\mathfrak{M}^0$, la prueba del Teorema 2 es completa. (Aquí está una prueba alternativa del primer enunciado del Teorema 2, una prueba que no hace uso del hecho de que $\mathfrak{V}$ es la suma directa de $\mathfrak{M}$ y de algún otro subespación. Si y está en $\mathfrak{M}^0$, entonces $[x, A'y] = [Ax, y] = 0$ para toda x de $\mathfrak{M}$, y, en consecuencia, $A'y$ está en $\mathfrak{M}^0$. La única ventaja de la prueba algebraica dada anteriormente sobre esta simple prueba geométrica es que la primera prepara el terreno para trabajo futuro con proyecciones).

Concluimos nuestro tratamiento de los adjuntos discutiendo sus matrices; esta discusión tiene el propósito de iluminar toda la teoría y facultar al lector para construir muchos ejemplos.

Necesitaremos el hecho siguiente: si $\mathfrak{X} = \{x_1, \ldots, x_n\}$ es una base cualquiera en el espacio vectorial n-dimensional $\mathfrak{V}$, si $\mathfrak{X}' = \{y_1, \ldots, y_n\}$ es la base dual de $\mathfrak{V}'$, y si la matriz de la transformación lineal A en el sistema de coordenadas $\mathfrak{X}$ es (α_{ij}), entonces

(2) $$\alpha_{ij} = [Ax_j, y_i].$$

Esto se sigue de la definición de la matriz de una transformación lineal; puesto que $Ax_j = \sum_k \alpha_{kj}x_k$, tenemos

$$[Ax_j, y_i] = \sum_k \alpha_{kj}[x_k, y_i] = \alpha_{ij}.$$

Para llevar las cosas derechas en las aplicaciones, volvemos a redactar la Fórmula (2), así: para encontrar el elemento (i, j) de $[A]$ en la base $\mathfrak{X}$, aplicamos A al j-ésimo elemento de $\mathfrak{X}$ y a continuación tomamos el valor de la función lineal i-ésima (en $\mathfrak{X}'$) del vector así obtenido.

Es ahora muy fácil encontrar la matriz $(\alpha_{ij}) = [A']$ en el sistema de coordenadas $\mathfrak{X}'$; simplemente seguimos la receta que se acaba de dar. En otras palabras, consideramos $A'y_j$ y tomamos el valor de la funcional lineal i-ésima en $\mathfrak{X}''$ (esto es, considerada como funcional lineal sobre $\mathfrak{X}'$) en este vector; el resultado es que

$$\alpha'_{ij} = [x_i, A'y_j].$$

Puesto que $[x_i, A'y_j] = [Ax_i, y_j] = \alpha_{ji}$ de manera que $\alpha_{ij} = \alpha_{ji}$, esta matriz $[A']$ se llama la *transpuesta* de $[A]$.

Obsérvese que nuestros resultados sobre la relación entre E y E' (donde E es una proyección) podrían también haber sido deducidos usando los hechos sobre la representación matricial de una proyección, justamente con el presente resultado sobre las matrices de transformaciones de adjuntos.

§46. CAMBIO DE BASE

Aunque lo que hemos estado haciendo hasta ahora con las transformaciones lineales puede haber sido complicada, fue en gran parte automático. Habiendo introducido el nuevo concepto de transformación lineal, solamente dejamos que algunos de los conceptos precedentes sugieran modos en que están conectados a las transformaciones lineales. Comenzamos ahora el estudio apropiadamente dicho de las transformaciones lineales. Como primera aplicación de la teoría resolveremos los problemas que se suscitan del cambio de base. Estos problemas pueden formularse sin mencionar transformaciones lineales, pero su solución es dada en forma sumamente efectiva en términos de transformaciones lineales.

Sea $\mathcal{V}$ un espacio vectorial n-dimensional y sean $\mathfrak{X} = \{x_1, \ldots, x_n\}$y $\mathcal{Y} = \{y_1, \ldots, y_n\}$ dos bases de $\mathcal{V}$. Podemos hacer las dos preguntas siguientes

CUESTION I. *Si x está en $\mathcal{V}$, $x = \sum_i \xi_i x_i = \sum_i \eta_i y_i$, ¿cuál es la relación entre sus coordenadas $(\xi_1, \ldots, \xi_n)$ con res-*

pecto de $\mathfrak{X}$ *y sus coordenadas* $(\eta_1, \ldots, \eta_n)$ *con respecto de* $\mathfrak{Y}$?

CUESTIÓN II. *Si* $(\xi_1, \ldots, \xi_n)$ *en un conjunto ordenado de* n *escalares, ¿cuál es la relación entre los vectores* $x = \sum_i \xi_i x_i$ y $y = \sum_i \xi_i y_i$?

Ambas cuestiones pueden ser fácilmente contestadas en el lenguaje de las transformaciones lineales. Consideramos, a saber, que la transformación lineal A, definida por $Ax_i = y_i$, $i = 1, \ldots, n$. Más explícitamente

$$A(\textstyle\sum_i \xi_i x_i) = \sum_i \xi_i y_i.$$

Sea (α_{ij}) la matriz de A en la base $\mathfrak{X}$, esto es, $y_j = Ax_j = \sum_i \alpha_{ij} x_i$. Observamos que A es invertible, puesto que $\sum_i \xi_i y_i = 0$ implica que $\xi_1 = \xi_2 = \ldots = \xi_n = 0$

RESPUESTA A LA CUESTIÓN I. Puesto que

$$\begin{aligned}\textstyle\sum_j \eta_j y_j &= \textstyle\sum_j \eta_j A x_j = \sum_j \eta_j \sum_i \alpha_{ij} x_i \\ &= \textstyle\sum_i (\sum_j \alpha_{ij} \eta_j) x_i,\end{aligned}$$

tenemos

$$\xi_i = \textstyle\sum_j \alpha_{ij} \eta_j. \tag{1}$$

RESPUESTA A LA CUESTIÓN II.

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

Hablando en términos generales, la transformación lineal invertible A(o, más propiamente, la matriz (α_{ij})) puede ser considerada como una transformación de coordenadas (como en 1)), o puede ser consideradas (como usualmente la consideramos en (2)) como una transformación de vectores.

En tratados clásicos sobre espacios vectoriales es costumbre tratar los vectores como n-tuples numéricos, más bien que como entidades abstractas. Esto necesita la introducción de cierta molesta terminología. Damos aquí un breve glosario de algunos de los términos y notaciones más desconcertantes que se su suscitan en conexión con espacios duales y transformaciones de adjuntos.

Si $\mathfrak{V}$ es un espacio vectorial n-dimensional, un vector x es dado por sus coordenadas con respecto de algún sistema absoluto de coordenadas preferido; estas coordenadas forman un conjunto ordenado de n escalares. Es costumbre escribir este conjunto de escalares en una columna,

$$x = \begin{bmatrix} \xi_1 \\ \cdot \\ \cdot \\ \xi_n \end{bmatrix}.$$

Elementos del espacio dual $\mathcal{V}'$ se escriben como filas, $x' = (\xi'_1, \ldots, \xi'_n)$. Si consideramos a x como una matriz rectangular n por uno y a x' como una matriz uno por n, entonces el producto matricial $x'x$ es una matriz uno por uno, esto es, un escalar. En nuestra notación, este escalar es $[x, x'] = \xi_1\xi'_1 + \ldots + \xi_n\xi'_n$. El artificio de considerar los vectores como matrices delgadas es efectivo aun cuando consideramos las matrices de magnitud normal de las transformaciones lineales. Así, el producto matricial de (α_{ij}) con la columna (ξ_j) es la columna cuyo i-ésimo elemento es $\eta_i = \sum_j \alpha_{ij}\xi_j$. En vez de preocuparse por bases duales y transformaciones de adjuntos, podemos formar similarmente el producto de la fila (ξ'_j) con la matriz (α_{ij}) en el orden (ξ_j) (α_{ij}); el resultado es la fila que al principio denotamos por $y' = A'x'$. La expresión $[Ax, x']$ se abrevia ahora como $x' \cdot A \cdot x$; ambos puntos significan multiplicación matricial ordinaria. Los vectores x de $\mathcal{V}$ se llaman *covariantes* y los vectores x' de $\mathcal{V}'$ se llaman *contravariantes*. Puesto que la noción del producto $x' \cdot x$ (esto es, $[x, x']$)depende, desde este punto de vista, de las coordenadas de x y x', resulta de importancia hacer la siguiente pregunta: si cambiamos la base en $\mathcal{V}$, de acuerdo con la transformación lineal invertible A, ¿qué debemos hacer en $\mathcal{V}'$ para preservar el producto $x' \cdot x$? En nuestra notación: si $[x, x'] = [y, y']$ donde $y = Ax$, entonces, ¿cómo está y' relacionada con x'? Respuesta: $y' = (A')^{-1}x'$. Para expresar toda esta maraña de ideas, la terminología clásica dice que los vectores x varían *cogradientemente* mientras que los x' varían *contragradientemente*.

§47. SIMILARIDAD

Las dos cuestiones siguientes están estrechamente relacionadas con las de la sección precedente.

CUESTIÓN III. *Si B es una transformación lineal sobre $\mathcal{V}$, ¿cuál es la relación entre su matriz (β_{ij}) con respecto de $\mathfrak{X}$ y de su matriz (γ_{ij}) con respecto de $\mathfrak{Y}$?*

CUESTIÓN IV. *Si (β_{ij}) es una matriz, ¿cuál es la relación* *entre las transformaciones lineales B y C definidas, respectivamente, por* $Bx_j = \sum_i \beta_{ij}x_i$ *y* $Cy_j = \sum_i \beta_{ij}y_i$?

Las cuestiones III y IV son formulaciones explícitas de un problema que hemos suscitado antes: a una transformación corresponde (en diferentes sistemas de coordenadas) muchas matrices (cuestión III) y a una matriz corresponden muchas transformaciones (cuestión IV).

RESPUESTA A LA CUESTIÓN III. Tenemos

$$(1) \qquad Bx_j = \sum_i \beta_{ij}x_i$$

y

$$(2) \qquad By_j = \sum_i \gamma_{ij}y_i.$$

Usando la transformación lineal A definida en la sección precedente, podemos escribir

$$(3) \qquad \begin{aligned} By_j &= BAx_j = B(\textstyle\sum_k \alpha_{kj}x_k) \\ &= \textstyle\sum_k \alpha_{kj}Bx_k = \sum_k \alpha_{kj}\sum_i \beta_{ik}x_i = \sum_i (\sum_k \beta_{ik}\alpha_{kj})x_i, \end{aligned}$$

y

$$(4) \qquad \begin{aligned} \textstyle\sum_k \gamma_{kj}y_k &= \textstyle\sum_k \gamma_{kj}Ax_k = \sum_k \gamma_{kj}\sum_i \alpha_{ik}x_i \\ &= \textstyle\sum_i (\sum_k \alpha_{ik}\gamma_{kj})x_i. \end{aligned}$$

comparando (2), (3) y (4), vemos que

$$\sum_k \alpha_{ik}\gamma_{kj} = \sum_k \beta_{ik}\alpha_{kj}.$$

Usando multiplicación matricial, escribimos esto en la forma peligrosamente simple

$$(5) \qquad [A][C] = [B][A].$$

El peligro radica en el hecho de que tres de las cuatro matrices de (5) corresponden a sus transformaciones lineales en la base $\mathfrak{X}$; la cuarta —a saber, la denotada por $[C]$ corresponde a B en la base $\mathcal{Y}$. En esta inteligencia, (5) es correcto. Una forma más usual de (5), adaptada, en principio, a computar $[C]$ cuando $[A]$ y $[B]$ son conocidos, es

(6) $$[C] = [A]^{-1}[B][A].$$

Respuesta a la cuestión IV. Para poner de relieve el carácter esencialmente geométrico de esta cuestión y su respuesta, observamos que

$$Cy_j = CAx_j$$

y

$$\sum_i \beta_{ij} y_i = \sum_i \beta_{ij} A x_i = A(\sum_i \beta_{ij} x_i) = ABx_j.$$

De aquí que C es tal que

$$CAx_j = ABx_j,$$

o, finalmente

(7) $$C = ABA^{-1}.$$

No hay dificultad con (7) similar a la que nos obligó a hacer una reserva sobre la interpretación de (6);encontrar la transformación lineal (no matricial) C, multiplicamos las transformaciones A, B y A^{-1}, y no es necesario decir nada sobre el sistema de coordenadas. Compárense, sin embargo, las fórmulas (6) y (7) y obsérvese una vez más la innata perversidad de los símbolos matemáticos. Esto es solamente otro aspecto de los hechos ya mencionados en los §§37 y 38.

Dos matrices $[B]$ y $[C]$ se llaman *similares* si existe una matriz invertible $[A]$ que satisface a (6); dos transformaciones lineales B y C se llaman similares si existe una transformación invertible A que satisface a (7). En este lenguaje, las respuestas a las cuestiones III y IV pueden expresarse muy brevemente, en ambos casos la respuesta es que las matrices o transformaciones dadas deben ser similares.

Habiendo obtenido la respuesta a la cuestión IV, podemos ver ahora que hay demasiados subíndices en su formulación. La validez de (7) es un hecho geométrico completamente independiente de la linearidad, finito-dimensionalidad o de cualquiera otra propiedad accidental que puedan poseer A, B y C; la respuesta a la cuestión IV es también la respuesta a una cuestión mucho más general. Esta cuestión geométrica, una paráfrasis de la formulación analítica de la cuestión IV, es ésta: Si B transforma a $\mathcal{V}$, y si C transforma a $A\,\mathcal{V}$ del mismo modo, ¿cuál es la relación entre B y C? La expresión "del mismo modo" no es tan vaga como parece; si significa que si B toma x dentro de, digamos, u, entonces C toma Ax dentro de Au.

La respuesta, es, por supuesto, la misma de antes: puesto que $Bx = u$ y $Cy = v$ (donde $y = Ax$ y $v = Au$), tenemos

$$ABx = Au = v = Cy = CAx.$$

La situación se resume convenientemente en el siguiente diagrama nemónico

$$\begin{tikzcd} x \arrow[r, "B"] \arrow[d, "A"'] & u \arrow[d, "A"] \\ y \arrow[r, "C"'] & v \end{tikzcd}$$

Podemos ir de y a v usando el camino corto C, o yendo alrededor del cuadrilátero; en otras palabras, $C = ABA^{-1}$. Recuérdese que ABA^{-1} ha de aplicarse a y de derecha a izquierda: primera A^{-1}, a continuación B y a continuación A.

Hemos visto que la teoría del cambio de bases es coextensiva con la teoría de las transformaciones lineales invertibles. Una transformación lineal invertible es un *automorfismo* donde, por automorfismo significamos un isomorfismo de un espacio vectorial consigo mismo. (Véase el §9). Observamos que, inversamente, todo automorfismo es una transformación lineal invertible.

Esperamos que la relación entre transformaciones lineales y matrices sea por ahora suficientemente clara para que el lector no objete si en lo que sigue, cuando deseamos dar ejemplos de transformaciones lineales con varias propiedades, nos contentemos con escribir una matriz. La interpretación que ha de darse siempre a este procedimiento es que tenemos presente el espacio vectorial concreto $\mathfrak{C}^n$ (o una de sus versiones generalizadas $\mathfrak{F}^n$) y la base concreta $\mathfrak{X} = \{x_1, \ldots, x_n\}$ definida por $x_i = (\delta_{i1}, \ldots, \delta_{in})$. En esta inteligencia, una matriz (α_{ij}) define, por supuesto, una transformación lineal única A, dada por la fórmula usual $A(\sum_i \xi_i x_i) = \sum_i (\sum_j \alpha_{ij}\xi_j)x_i$.

EJERCICIOS

1. Si A es una transformación lineal de un espacio vectorial $\mathcal{U}$ a un espacio vectorial $\mathcal{V}$, entonces, correspondiendo a cada y fija de $\mathcal{V}'$ existe un vector, que podría bien ser denotado por $A'y$, en $\mathcal{U}'$ de manera que

$$[Ax, y] = [x, A'y]$$

para toda x de $\mathfrak{U}$. Pruébese que A' es una transformación lineal de $\mathfrak{V}'$ a $\mathfrak{U}'$. (La transformación A' se llama el *adjunto* de A). Interprétense y pruébense tantas como sea posible entre las ecuaciones §44, (2)–(8) para este concepto de adjunto).

(2). (a) Pruébese que la similaridad de las transformaciones lineales sobre un espacio vectorial es una relación de equivalencia (esto es, reflexiva, simétrica y transitiva).

(b) Si A es similar a un escalar α, entonces $A = \alpha$.

(c) Si A y B son similares, entonces también lo son A^2 y B^2, A' y B' y, en caso de que A y B sean invertibles, A^{-1} y B^{-1}.

(d) Generalícese el concepto de similaridad a dos transformaciones definidas sobre espacios vectoriales diferentes. ¿Cuál de los resultados precedentes permanece válido para el concepto generalizado?

3. (a) Si A y B son transformaciones lineales sobre el mismo espacio vectorial y si cuando menos una de ellas es invertible, entonces AB y BA son similares.

(b) ¿Permanece válida la conclusión de (a) si ni A ni B son invertibles?

4. Si la matriz de una transformación lineal A sobre $\mathbb{C}^2$, con respecto de la base $\{(1, 0), (0, 1)\}$ es $\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$, ¿cuál es la matriz de A con respecto de la base $\{(1, 1), (1, -1)\}$? ¿Qué sobre la base $\{(1, 0), (1, 1)\}$?

5. Si la matriz de una transformación lineal A sobre $\mathbb{C}^3$, con respecto de la base $\{(1, 0, 0), (0, 1, 0), (0, 0, 1)\}$ es $\begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & -1 \\ -1 & -1 & 0 \end{pmatrix}$, ¿cuál es la matriz de A con respecto de esta base $\{(0, 1, -1), (1, -1, 1), (-1, 1, 0)\}$?

6. (a) La construcción de una matriz asociada con una transformación lineal depende de dos bases, no de una. En realidad, si $\mathfrak{X} = \{x_1, \cdots, x_n\}$ y $\bar{\mathfrak{X}} = \{\bar{x}_1, \cdots, \bar{x}_n\}$ son bases de $\mathfrak{V}$, y si A es una transformación lineal sobre V, entonces la matriz $[A; \mathfrak{X}, \bar{\mathfrak{X}}]$ de A con respecto a $\mathfrak{X}$ y $\bar{\mathfrak{X}}$ debería definirse como

$$Ax_j = \sum_i \alpha_{ij}\bar{x}_i.$$

La definición adoptada en el texto corresponde al caso especial en que $\bar{\mathfrak{X}} = \mathfrak{X}$. El caso especial lleva a la definición de similaridad (B y C son similares si existen bases $\mathfrak{X}$ y $\mathfrak{Y}$ tales que $[B; \mathfrak{X}] = [C; \mathfrak{Y}]$). La relación análoga sugerida por el caso general se llama equivalencia; B y C son *equivalentes* si existen bases pares $(\mathfrak{X}, \bar{\mathfrak{X}})$ y $(\mathfrak{Y}, \bar{\mathfrak{Y}})$ tales que $[B; \mathfrak{X}, \bar{\mathfrak{X}}] = [C; \mathfrak{Y}, \mathfrak{Y}]$. Pruébese que esta noción es, en realidad, una relación de equivalencia.

(b) Dos transformaciones lineales B y C son equivalentes si y sólo si existen transformaciones lineales invertibles P y Q tales que $PB = CQ$.

(c) Si A y B son equivalentes, entonces lo son también A' y B'.

(d) ¿Existe una transformación lineal A tal que A sea equivalente a un escalar α, pero $A \neq \alpha$?

(e) ¿Existen transformaciones lineales A y B tales que A y B sean equivalentes, pero A^2 y B^2 no lo sean?

(f) Generalícese el concepto de equivalencia a dos transformaciones definidas sobre diferentes vectores espaciales. ¿Cuáles de los resultados precedentes permanece válido para el concepto generalizado?

§48. TRANSFORMACIONES POR COCIENTE

Supóngase que A es una transformación lineal sobre un espacio vectorial $\mathcal{V}$ y que $\mathfrak{M}$ es un subespacio de $\mathcal{V}$ invariante bajo A. Bajo estas circunstancias hay un modo natural de definir una transformación lineal (que ha de ser denotada por $A/\mathfrak{M}$) sobre el espacio $\mathcal{V}/\mathfrak{M}$; esta "transformación por cociente" está relacionada con A justamente en la misma forma en que el espacio cociente está relacionado con $\mathcal{V}$. Será conveniente (en esta sección) denotar $\mathcal{V}/\mathfrak{M}$ por el símbolo más compacto $\mathcal{V}^-$, y usar símbolos relacionados para los factores y las transformaciones que ocurren. Así, por ejemplo, si x es un vector cualquiera de $\mathcal{V}$, denotaremos el coconjunto $x + \mathfrak{M}$ por x^-; objetos tales como x^- son los elementos típicos de $\mathcal{V}^-$.

Para definir la transformación por cociente $A/\mathfrak{M}$ (que será denotada, alternativamente, por A^-), escribimos

$$A^-x^- = (Ax)^-$$

para todo vector x de $\mathcal{V}$. En otras palabras, para encontrar la transforma por $A/\mathfrak{M}$ del coconjunto $x + \mathfrak{M}$, primero encuéntrese la transforma por A del vector x, y a continuación fórmese el coconjunto de $\mathfrak{M}$ determinado por el vector transformado. Esta definición debe ser sostenida por un argumento de inambigüedad; debemos estar seguros de que si dos vectores determinan el mismo coconjunto, entonces lo mismo es verdadero de sus transformas por A. El hecho clave es la invariancia de $\mathfrak{M}$. En realidad, si $x + \mathfrak{M} = y + \mathfrak{M}$, entonces $x - y$ está en $\mathfrak{M}$, de manera que $Ax - Ay$ (invariancia) está en $\mathfrak{M}$, y en consecuencia, $Ax + \mathfrak{M} = Ay + \mathfrak{M}$.

¿Qué sucede si $\mathfrak{M}$ no es meramente invariante bajo A, sino, juntamente con un subespacio adecuado $\mathfrak{N}$, reduce a A? Si esto sucede, entonces A es la suma directa, digamos $A = B \oplus C$, de dos transformaciones lineales definidas sobre los subespacios $\mathfrak{M}$ y $\mathfrak{N}$ de $\mathcal{V}$, respectivamente; la cuestión es, ¿cuál es la relación entre A^- y C? Ambas transformaciones pueden considerarse como complementarias de A; la transformación B describe lo que A hace sobre

$\mathfrak{M}$, y tanto A^- como C describe, de diferentes modos, lo que A hace en otra parte.

Sea T la correspondencia que asigna a cada vector x de $\mathfrak{N}$ el coconjunto x^- $(= x + \mathfrak{M})$. Sabemos ya que T es un isomorfismo entre $\mathfrak{N}$ y $\mathfrak{V}/\mathfrak{M}$ (véase el § 22, Teorema 1); demostraremos ahora que el isomorfismo lleva la transformación C sobre la transformación A^-. Si $Cx = y$ (donde, por supuesto, x está en $\mathfrak{N}$), entonces $A^-x^- = (Ax)^- = (Cx)^- = y^-$; se sigue que $TCx = Ty = A^-Tx$. Esto implica que $TC = A^-T$, como se prometió. Hablando en términos generales (véase el §47), podemos decir que A^- transforma a $\mathfrak{V}^-$ del mismo modo que C transforma $\mathfrak{N}$. En otras palabras, las transformaciones lineales A^- y C son abstractamente idénticas (isomórficas). Este hecho es de gran importancia en las aplicaciones del concepto de espacio cociente.

§49. ALCANCE Y ESPACIO NULO

DEFINICIÓN. Si A es una transformación lineal sobre un espacio vectorial $\mathfrak{V}$ y si $\mathfrak{M}$ es un subespacio de $\mathfrak{V}$, la *imagen* de $\mathfrak{M}$ bajo A, en símbolos $A\mathfrak{M}$, es el conjunto de todos los vectores de la forma Ax, con x en $\mathfrak{M}$. El *alcance* de A es el conjunto $\mathfrak{R}(A) = A\mathfrak{V}$; el *espacio nulo* de A es el conjunto $\mathfrak{N}(A)$ de todos los vectores x para los cuales $Ax = 0$.

Se verifica inmediatamente que $\mathfrak{M}$ y $\mathfrak{N}(A)$ son subespacios. Si, como es usual, denotamos por $\mathfrak{O}$ el subespacio que contiene solamente el vector 0, es fácil describir algunos conceptos familiares en la terminología que se acaba de introducir; listamos algunos de los resultados.

(i) La transformación A es invertible si y sólo si $\mathfrak{R}(A) = \mathfrak{V}$ y $\mathfrak{N}(A) = \mathfrak{O}$.

(ii) En caso de que $\mathfrak{V}$ sea finito-dimensional, A es invertible si y sólo $\mathfrak{R}(A) = \mathfrak{V}$ o $\mathfrak{N}(A) = \mathfrak{O}$.

(iii) El subespacio $\mathfrak{M}$ es invariante bajo A si y sólo si $A\mathfrak{M} \subset \mathfrak{M}$.

(iv) Un par de subespacios complementarios $\mathfrak{M}$ y $\mathfrak{N}$ reduce a A si y sólo si $A\mathfrak{M} \subset \mathfrak{M}$ y $A\mathfrak{N} \subset \mathfrak{N}$.

(v) Si E es la proyección sobre $\mathfrak{M}$ a lo largo de $\mathfrak{N}$, entonces $\mathfrak{R}(E) = \mathfrak{M}$ y $\mathfrak{N}(E) = \mathfrak{N}$.

Todos estos enunciados son fáciles de probar; indicamos la prueba de (v). Según el § 41, Teorema 2, sabemos que $\mathfrak{N}$ es el conjunto de todas las soluciones de la ecuación $Ex = 0$; esto coincide con nuestra definición de $\mathfrak{N}(E)$. Sabemos también que $\mathfrak{M}$ es el conjunto de todas las soluciones de la ecuación $Ex = x$. Si x está en $\mathfrak{N}$, entonces x está también en $\mathfrak{R}(E)$, (puesto que x es la imagen bajo E de algo (a saber de sí misma). Recíprocamente, si un vector x es la imagen bajo E de algo, digamos $x = Ey$ (de manera que x está en $\mathfrak{R}(E)$), entonces $Ex = E^2x = Ey = x$, de manera que x está en $\mathfrak{N}$.

Advertencia: es accidental que para proyecciones $\mathfrak{R} \oplus \mathfrak{N} = \mathfrak{V}$. En general, no es necesario que sea cierto que $\mathfrak{R} = \mathfrak{R}(A)$ y $\mathfrak{N} = \mathfrak{N}(A)$ estén disjuntos. Puede suceder, por ejemplo, que para cierto vector x tengamos $x \neq 0$, $Ax \neq 0$ y $A^2x = 0$; para ese vector, Ax claramente pertenece al alcance y al espacio nulo de A.

Teorema. *Si A es una transformación lineal sobre un espacio vectorial $\mathfrak{V}$, entonces*

$$(\mathfrak{R}(A))^0 = \mathfrak{N}(A'); \tag{1}$$

si $\mathfrak{V}$ es finito-dimensional, entonces

$$(\mathfrak{N}(A))^0 = \mathfrak{R}(A'). \tag{2}$$

Prueba. Si y está en $(\mathfrak{R}(A))^0$, entonces, para toda x de $\mathfrak{V}$,

$$0 = [Ax, y] = [x, A'y],$$

de manera que $A'y = 0$ y y está en $\mathfrak{N}(A')$. Si, por otra, y está en $\mathfrak{N}(A')$, entonces para toda x de $\mathfrak{V}$,

$$0 = [x, A'y] = [Ax, y],$$

de manera que y está en $(\mathfrak{R}(A))^0$.

Si aplicamos (1) a A' en lugar de A, obtenemos

$$(\mathfrak{R}(A'))^0 = \mathfrak{N}(A''). \tag{3}$$

Si $\mathfrak{V}$ es finito-dimensional (y de aquí reflexivo), podemos reemplazar A'' por A en (3) y entonces podemos formar el aniquilador de ambos lados; la conclusión deseada (2) se sigue del § 17, Teorema 2.

EJERCICIOS

1. Usese el operador de diferenciación sobre $\mathcal{P}_n$ para demostrar que el rango y el espacio nulo de una transformación lineal no necesita ser disjunta.

2. (a) Dese un ejemplo de una transformación lineal sobre un espacio tridimensional con un rango bidimensional.

(b) Dese un ejemplo de una transformación lineal sobre un espacio tridimensional con un espacio nulo bidimensional.

3. Encuéntrese una matriz cuatro por cuatro cuyo rango sea sobreentendido por (1, 0, 1, 0) y (0, 1, 0, 1).

4. (a) Dos proyecciones E y F tienen el mismo rango si y sólo si $EF = F$ y $FE = F$.

(b) Dos proyecciones E y F tienen el mismo espacio nulo si y sólo si $EF = E$ y $FE = F$.

5. Si $E_1, \ldots, E_k$ son proyecciones con el mismo rango y si $\alpha_1, \ldots, \alpha_k$ son escalares tales que $\sum_i \alpha_i = 1$, entonces $\sum_i \alpha_i E_i$ es una proyección.

§50. RANGO Y NULIDAD

Restringiremos ahora nuestra atención al caso finito-dimensional y sacaremos ciertas fáciles conclusiones del teorema de la sección precedente.

Definición. El *rango* $\rho(A)$, de una transformación lineal A sobre un espacio vectorial finito-dimensional es la dimensión de $\mathfrak{R}(A)$; la *nulidad*, $\nu(A)$, es la dimensión de $\mathfrak{N}(A)$.

Teorema 1. *Si A es una transformación lineal sobre un espacio vectorial, entonces* $\rho(A) = \rho(A')$ y $\nu(A) = n - \rho(A)$.

Prueba. El teorema de la sección precedente y el § 17, Teorema 1, juntamente implican que

(1) $$\nu(A') = n - \rho(A).$$

Sea $\mathfrak{X} = \{x_1, \ldots, x_n\}$ cualquier base para la cual $x_1, \ldots, x_\nu$ están en $\mathfrak{N}(A)$; entonces, para cualquier $x = \sum_i \xi_i x_i$, tenemos

$$Ax = \sum_i \xi_i A x_i = \sum_{i=\nu+1}^{n} \xi_i A x_i.$$

En otras palabras, Ax es una combinación lineal de los $n - \nu$ vectores $Ax_{\nu+1}, \ldots, Ax_n$; se sigue que $\rho(A) \leqq n - \nu(A)$. Aplicando este resultado a A' y usando (1), obtenemos

(2) $$\rho(A') \leqq n - \nu(A') = \rho(A)$$

En (2) podemos reemplazar A por A', obteniendo

(3) $$\rho(A) = \rho(A'') \leqq \rho(A');$$

(2) y (3) juntos demuestran que

(4) $$\rho(A) = \rho(A'),$$

y (1) y (4) juntos demuestran que

(5) $$\nu(A') = n - \rho(A').$$

Reemplazando A por A' en (5) da, finalmente

(6) $$\nu(A) = n - \rho(A),$$

y concluye la prueba del teorema.

Estos resultados se discuten generalmente desde un punto de vista un poco diferente. Sea A una transformación lineal sobre un espacio vectorial n-dimensional, y sea $\mathfrak{X} = \{x_1, \ldots, x_n\}$ una base de ese espacio; sea $[A] = (\alpha_{ij})$ la matriz de A en el sistema de coordenadas $\mathfrak{X}$, de manera que

$$Ax_j = \sum_i \alpha_{ij} x_i.$$

Puesto que $x = \sum_j \xi_j x_j$, entonces $Ax = \sum_j \xi_j Ax_j$, se sigue que todo vector de $\mathfrak{R}(A)$ es una combinación lineal de las Ax_j, en consecuencia, de cualquier subconjunto máximo linealmente independiente de las Ax_j. Se sigue que el número máximo de Ax_j, linealmente independientes es precisamente $\rho(A)$. En términos de las coordenadas $(\alpha_{1j}, \ldots, \alpha_{nj})$ de Ax_j podemos expresar esto diciendo que $\rho(A)$ es el número máximo de columnas linealmente independientes de la matriz $[A]$. Puesto que (§ 45) las columnas de $[A']$ (expresada la matriz en términos de la base dual de $\mathfrak{X}$) son las filas de $[A]$, se sigue del Teorema 1 que $\rho(A)$ es también el número máximo de filas linealmente independientes de $[A]$. De aquí que "el rango de hilera de $[A]$ = el rango de columna de $[A]$ = el rango de $[A]$".

Teorema 2. *Si A es una transformación lineal del espacio vectorial n-dimensional $\mathcal{V}$, y si $\mathcal{H}$ es cualquier espacio h-dimensional de $\mathcal{V}$, entonces la dimensión de $A\mathcal{H}$ es $\geqq h - \nu(A)$.*

Prueba. Sea $\mathcal{K}$ cualquier subespacio para el cual $\mathcal{V} = \mathcal{H} \oplus \mathcal{K}$, de manera que si k es la dimensión de $\mathcal{K}$, entonces $k = n - h$. Al operar con A, obtenemos

$$A\mathcal{V} = A\mathcal{H} + A\mathcal{K}.$$

(La suma no es necesariamente una suma directa; véase el § 11). Puesto que $A\mathcal{V} = \mathfrak{R}(A)$ tiene dimensión $n - \nu(A)$, puesto que la dimensión de $A\mathcal{K}$ es claramente $\leqq k = n - h$, y puesto que la dimensión de la suma es $\leqq$ la suma de las dimensiones, tenemos el resultado deseado.

Teorema 3. *Si A y B son transformaciones lineales sobre un espacio vectorial finito-dimensional, entonces*

(7) $$\rho(A + B) \leqq \rho(A) + \rho(B),$$

(8) $$\rho(AB) \leqq \text{mín } \{\rho(A), \rho(B)\},$$

y

(9) $$\nu(AB) \leqq \nu(A) + \nu(B).$$

Si B es invertible, entonces

(10) $$\rho(AB) = \rho(BA) = \rho(A).$$

PRUEBA. Puesto que $(AB)x = A(Bx)$, se sigue que $\mathfrak{R}(AB)$ está contenido en $\mathfrak{R}(A)$, de manera que $\rho(AB) \leqq \rho(A)$, o, en otras palabras, el rango de un producto no es mayor que el rango del primer factor. Apliquemos este resultado auxiliar al BA'; esto, juntamente con lo que ya sabemos, da (8). Si B es invertible, entonces

$$\rho(A) = \rho(AB \cdot B^{-1}) \leqq \rho(AB)$$

y

$$\rho(A) = \rho(B^{-1} \cdot BA) \leqq \rho(BA);$$

juntamente con (8) esto da (10). La ecuación (7) es una consecuencia inmediata de un argumento que ya hemos usado en la prueba del Teorema 2. La prueba del (9) la dejamos como ejercicio al lector. (Sugestión: aplíquese el Teorema 2 con $\mathfrak{X} = B\mathfrak{V} = \mathfrak{R}(B)$.) Juntas las dos fórmulas (8) y (9) se conocen como *ley de nulidad* de Silvester.

§51. TRANSFORMACION DE RANGO UNO

Concluimos nuestra discusión del rango por una descripción de las matrices de transformaciones lineales de rango $\leqq 1$.

TEOREMA 1. *Si una transformación lineal A sobre un espacio vectorial finito-dimensional* $\mathfrak{V}$ *es tal que* $\rho(A) \leqq 1$ *[esto es,* $\rho(A) = 0$ *o* $\rho(A) = 1$*], entonces los elementos de la matriz* $[A] = (\alpha_{ij})$ *de A tienen la forma* $\alpha_{ij} = \beta_i\gamma_j$ *en todo sistema de coordenadas; recíprocamente, si la matriz de A tiene esta forma en algún otro sistema de coordenadas, entonces* $\rho(A) \leqq 1$.

PRUEBA. Si $\rho(A) = 0$, entonces $A = 0$ y el enunciado es trivial. Si $\rho(A) = 1$, esto es, $\mathfrak{R}(A)$ es unidimensional, entonces existe en

$\mathfrak{R}(A)$ un vector no cero x_0 [una base de $\mathfrak{R}(A)$] tal que todo vector de $\mathfrak{R}(A)$ es un múltiplo de x_0. De aquí que, para toda x,

$$Ax = y_0 x_0$$

donde el coeficiente escalar $y_0[= y_0(x)]$ depende, por supuesto, de x. La linearidad de A implica que y_0 es una funcional lineal sobre $\mathfrak{V}$. Sea $\mathfrak{X} = \{x_1, \ldots, x_n\}$ una base de $\mathfrak{V}$, y sea (α_{ij}) la correspondiente matriz de A, de manera que

$$Ax_j = \sum_i \alpha_{ij} x_i.$$

Si $\mathfrak{X}' = \{y_1, \ldots, y_n\}$ es la base dual de $\mathfrak{V}'$, entonces [véase el § 45, (2)]

$$\alpha_{ij} = [A_{xj}, y_i].$$

En el presente caso

$$\alpha_{ij} = [y_0(x_j)x_0, y_i] = y_0(x_j)[x_0, y_i] = [x_0, y_i][x_j, y_0];$$

en otras palabras, tomamos $\beta_i = [x_0, y_i]$ y $\gamma_j = [x_j, y_0]$.

Recíprocamente, supóngase que en un sistema fijo de coordenadas $\mathfrak{X} = \{x_1, \ldots, x_n\}$ la matriz (α_{ij}) de A es tal que $\alpha_{ij} = \beta_i\gamma_j$. Podemos encontrar una función lineal y_0 tal que $\gamma_j = [x_j, y_0]$ y podemos definir un vector x_0 por $x_0 = \sum_k \beta_k x_k$. La transformación lineal $\tilde{A}$ definida por $\tilde{A}x = y_0(x)x_0$ es claramente de rango uno (a menos, por supuesto, que $\alpha_{ij} = 0$ para toda i y j) y su matriz $(\tilde{\alpha}_{ij})$ en el sistema de coordenadas $\mathfrak{X}$ es dada por

$$\tilde{\alpha}_{ij} = [\tilde{A}x_j, y_i]$$

(donde $\mathfrak{X}' = \{y_1, \ldots, y_n\}$ es la base dual de $\mathfrak{X}$). De aquí que

$$\tilde{\alpha}_{ij} = [y_0(x_j)x_0, y_i] = [x_0, y_i][x_j, y_0] = \beta_i\gamma_j,$$

y, puesto que A y $\tilde{A}$ tienen la misma matriz en un sistema de coordenadas, se sigue que $\tilde{A} = A$. Esto concluye la prueba del teorema.

El siguiente teorema hace algunas veces posible aplicar el Teorema 1 para obtener resultados con una arbitraria transformación lineal.

TEOREMA 2. *Si A es una transformación lineal de rango ρ sobre un espacio vectorial finito-dimensional $\mathfrak{V}$, entonces A puede ser escrito como la suma de ρ transformaciones de rango uno.*

PRUEBA. Puesto que $A\mathcal{V} = \mathcal{R}(A)$ tiene una dimensión ρ, podemos encontrar ρ vectores $x_1, \ldots, x_\rho$ que formen una base para $\mathcal{R}(A)$. Se sigue que, para todo vector x de $\mathcal{V}$, tenemos

$$Ax = \sum_{i=1}^{\rho} \xi_i x_i,$$

donde cada ξ_i depende, por supuesto, de x; escribimos $\xi_i = y_i(x)$. Es fácil ver que y_i es una funcional lineal. En términos de estas y_i definimos, para cada $i = 1, \ldots, \rho$, una transformación lineal A_i por $A_i x = y_i(x)x_i$. Se sigue que cada A_i tiene rango uno y $A = \sum_{i=1}^{\rho} A_i$. [Compárese este resultado con el § 32, Ej. (2)].

Un ligero refinamiento de la prueba que se acaba de dar da el siguiente resultado.

TEOREMA 3. *Correspondiendo a cualquier transformación lineal A sobre un espacio vectorial finito-dimensional* $\mathcal{V}$ *hay una transformación lineal P para la cual PA es una proyección.*

PRUEBA. Sean $\mathcal{R}$ y $\mathfrak{N}$, respectivamente, el alcance y el espacio nulo de A y sea $\{x_1, \ldots, x_\rho\}$ una base para $\mathcal{R}$. Sean $x_{\rho+1}, \ldots, x_n$ vectores tales que $\{x_1, \ldots, x_n\}$ es una base de $\mathcal{V}$. Puesto que x_i está en $\mathcal{R}$ para $i = 1, \ldots, \rho$, podemos encontrar vectores y_i tales que $Ay_i = x_i$; finalmente, escogemos una base para $\mathfrak{N}$, que podemos denotar por $\{y_{\rho+1}, \ldots, y_n\}$. Afirmamos que $\{y_1, \ldots, y_n\}$ es una base para $\mathcal{V}$. Necesitamos, por supuesto, probar solamente que las y son linealmente independientes. Para este propósito suponemos que $\sum_{i=1}^{n} \alpha_i y_i = 0$; .Entonces tenemos (recordando que para $i = \rho + 1, \ldots, n$ el vector y_i pertenece a $\mathfrak{N}$)

$$A(\sum_{i=1}^{n} \alpha_i y_i) = \sum_{i=1}^{\rho} \alpha_i x_i = 0,$$

de donde $\alpha_1 = \ldots = \alpha_\rho = 0$. Consecuentemente, $\sum_{i=\rho+1}^{n} \alpha_i y_i = 0$; la independencia lineal de $y_{\rho+1}, \ldots, y_n$ demuestra que las α restantes deben también desaparecer.

Una transformación lineal P de esta clase, cuya existencia fue afirmada, se determina ahora por las condiciones $Px_i = y_i$, $i = 1, \ldots, n$. En realidad, si $i = 1, \ldots, \rho$, entonces $PAy_i = Px_i = y_i$ y si $i = \rho + 1, \ldots, n$, entonces $PAy_i = P0 = 0$.

La consideración del adjunto de A, juntamente con la reflexividad de $\mathcal{V}$, demuestra que podemos también encontrar un invertible Q para el cual AQ es una proyección. En caso de que A misma sea invertible, debemos tener $P = Q = A^{-1}$.

EJERCICIOS

1. ¿Cuál es el rango del operador de diferenciación sobre $\mathcal{P}_n$? ¿Cuál es su nulidad?

2. Encuéntrense los rangos de las siguientes matrices.

(a) $\begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}$, (c) $\begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}$,

(b) $\begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}$, (d) $\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$.

3. Si A es una multiplicación izquierda por P sobre un espacio de transformaciones lineales (véase el § 38, Ej. 5) y si P tiene rango m, ¿cuál es el rango de A?

4. El rango de la suma directa de dos transformaciones lineales (sobre espacios vectoriales finito-dimensionales) es la suma de sus rangos.

5. (a) Si A y B son transformaciones lineales sobre un espacio vectorial n-dimensional y si $AB = 0$, entonces $\rho(A) + \rho(B) \leqq n$.

(b) Para cada transformación lineal A sobre un espacio vectorial n-dimensional existe una transformación lineal B tal que $AB = 0$ y tal que $\rho(A) + \rho(B) = n$.

6. Si A, B y C son transformaciones lineales sobre un espacio vectorial finito-dimensional, entonces

$$\rho(AB) + \rho(BC) \leqq \rho(B) + \rho(ABC).$$

7. Pruébese que dos transformaciones lineales (sobre el mismo espacio vectorial finito-dimensional) son equivalentes si y sólo si tienen el mismo rango.

8. (a) Supóngase que A y B son transformaciones lineales (sobre el mismo espacio vectorial finito-dimensional) tales que $A^2 = A$ y $B^2 = B$. ¿Es verdad que A y B son similares si y sólo si $\rho(A) = \rho(B)$?

(b) Supóngase que A y B son transformaciones lineales (sobre el mismo espacio vectorial finito-dimensional) tales como $A \neq 0$, $B \neq 0$ y $A^2 = B^2 = 0$. ¿Es verdad que A y B son similares si y sólo si $\rho(A) = \rho(B)$?

9. (a) Si A es una transformación lineal de rango uno, existe entonces un escalar único α tal que $A^2 = \alpha A$.

(b) Si $\alpha \neq 1$, entonces $1 - A$ es invertible.

§52. PRODUCTOS TENSORIALES DE TRANSFORMACIONES

Liguemos ahora las transacciones lineales con la teoría de los productos tensoriales. Sean $\mathcal{U}$ y $\mathcal{V}$ espacios vectoriales finito-dimensionales (sobre el mismo campo) y sean A y B dos transformaciones lineales cualesquiera sobre $\mathcal{U}$ y $\mathcal{V}$ respectivamente. Definimos una transformación lineal $\bar{C}$ sobre el espacio $\mathcal{W}$ de todas las formas bilineales sobre $\mathcal{U} \oplus \mathcal{V}$ escribiendo

$$(\bar{C}w)(x, y) = w(Ax, By).$$

El *producto tensorial* $C = A \otimes B$ de las transformaciones A y B es, por definición, el dual de la transformación $\bar{C}$, de manera que

$$(Cz)(w) = z(\bar{C}w)$$

siempre que z esté en $\mathcal{U} \otimes \mathcal{V}$ y w en $\mathcal{W}$. Si aplicamos C a un elemento $z_0 = x_0 \otimes y_0$ [recuérdese que esto significa que $z_0(w) = w(x_0, y_0)$ para toda w de $\mathcal{W}$], obtenemos

$$(Cz_0)(w) = z_0(\bar{C}w) = (x_0 \otimes y_0)(\bar{C}w)$$
$$= (\bar{C}w)(x_0, y_0) = w(Ax_0, By_0) = (Ax_0 \otimes By_0)(w).$$

Inferimos que

(1) $$Cz_0 = Ax_0 \otimes By_0.$$

Puesto que hay pocos elementos en $\mathcal{U} \otimes \mathcal{V}$ de la forma $x \otimes y$, suficientes en cualquier caso para formar una base (véase el § 25), esta relación caracteriza a C.

Las reglas formales de productos tensoriales son como sigue.

(2) $$A \otimes 0 = 0 \otimes B = 0,$$

(3) $$1 \otimes 1 = 1,$$

(4) $$(A_1 + A_2) \otimes B = (A_1 \otimes B) + (A_2 \otimes B),$$

(5) $$A \otimes (B_1 + B_2) = (A \otimes B_1) + (A \otimes B_2),$$

(6) $$\alpha A \otimes \beta B = \alpha\beta(A \otimes B),$$

(7) $$(A \otimes B)^{-1} = A^{-1} \otimes B^{-1},$$

(8) $$(A_1A_2) \otimes (B_1B_2) = (A_1 \otimes B_1)(A_2 \otimes B_2).$$

Las pruebas de estas relaciones, excepto quizás las dos últimas, son directas.

La fórmula (7), como toda fórmula que implica inversos, tiene que ser leída con precaución. Tiene el propósito de significar que si A y B son invertibles, entonces lo son $A \otimes B$ y la ecuación es válida, y, recíprocamente, que si $A \otimes B$ es invertible, también lo son A y B. Probaremos (7) y (8) en orden inverso.

La fórmula (8) se sigue de la caracterización (1) de productos tensoriales y del siguiente cómputo:

$$(A_1A_2 \otimes B_1B_2)(x \otimes y) = A_1A_2x \otimes B_1B_2y$$
$$= (A_1 \otimes B_1)(A_2x \otimes B_2y) = (A_1 \otimes B_1)(A_2 \otimes B_2)(x \otimes y).$$

Como consecuencia inmediata de (8) obtenemos

$$(9) \qquad A \otimes B = (A \otimes 1)(1 \otimes B) = (1 \otimes B)(A \otimes 1).$$

Para probar (7), supóngase que A y B son invertibles, y fórmese $A \otimes B$ y $A^{-1} \otimes B^{-1}$. Puesto que, según (8), el producto de estas dos transformaciones, en cualquier orden, es 1, se sigue que $A \otimes B$ es invertible y (7) es válida. Recíprocamente, supóngase que $A \otimes B$ es invertible. Recordando que definimos los productos tensoriales solamente para espacios finito-dimensionales, podemos invocar el § 36, Teorema 2; basta con probar que $Ax = 0$ implica que $x = 0$ y $By = 0$ implica que $y = 0$. Usamos (1):

$$Ax \otimes By = (A \otimes B)(x \otimes y).$$

Si cualquiera de los factores sobre la izquierda es cero, entonces $(A \otimes B)(x \otimes y) = 0$, de donde $x \otimes y = 0$, de manera que o $x = 0$ o $y = 0$. Puesto que [por (2)] $B = 0$ es imposible, podemos encontrar un vector y tal que $By \neq 0$. Aplicando el argumento anterior a esta y, con cualquiera x para la cual $Ax = 0$, concluimos que $x = 0$. El mismo argumento con los papeles de A y B intercambiados prueba que B es invertible.

Un interesante (y complicado) aspecto de la teoría de los productos tensoriales de transformaciones es la teoría de los productos de matrices de Kronecker. Sean $\mathfrak{X} = \{x_1, \ldots, x_n\}$ y $\mathcal{Y} = \{y_1, \ldots, y_m\}$ bases en $\mathcal{U}$ y $\mathcal{V}$, y sean $[A] = [A; \mathfrak{X}] = (\alpha_{ij})$ y $[B] = [B; \mathcal{Y}] = (\beta_{pq})$ las matrices de A y B. ¿Cuál es la matriz de $A \otimes B$ en el sistema de coordenadas $\{x_i \otimes y_p\}$?

Para resolver la cuestión, debemos recordar la discusión del § 37, relativa al arreglo de una base en orden lineal. Puesto que, infortunadamente, es imposible escribir una matriz sin comprometerse en cuanto al orden de las filas y las columnas seremos francos al respecto y arreglaremos los n veces m vectores $x_i \otimes y_p$ en el llamado orden lexicográfico, como sigue:

$$x_1 \otimes y_1, x_1 \otimes y_2, \cdots, x_1 \otimes y_m, x_2 \otimes y_1, \cdots,$$
$$x_2 \otimes y_m, \cdots, x_n \otimes y_1, \cdots, x_n \otimes y_m.$$

Procedemos también a efectuar el siguiente cómputo:

$$(A \otimes B)(x_j \otimes y_q) = Ax_j \otimes By_q = (\sum_i \alpha_{ij}x_i) \otimes (\sum_p \beta_{pq}y_p)$$
$$= \sum_i \sum_p \alpha_{ij}\beta_{pq}(x_i \otimes y_p).$$

Este proceso indica exactamente hasta dónde podemos llegar sin ordenar los elementos básicos; si, por ejemplo, convenimos en poner índices a los elementos de una matriz, no con un par de enteros, sino con un par de pares, digamos (i, p) y (j, q), entonces sabemos ahora que el elemento en la fila (i, p) y la columna (j, q), es $\alpha_{ij}\beta_{pq}$. Si usamos el ordenamiento lexicográfico, la matriz $A \otimes B$ tiene la forma

$$\begin{bmatrix} \alpha_{11}\beta_{11} \cdots \alpha_{11}\beta_{1m} & \cdots & \alpha_{1n}\beta_{11} \cdots \alpha_{1n}\beta_{1m} \\ \vdots & & \vdots \\ \alpha_{11}\beta_{m1} \cdots \alpha_{11}\beta_{mm} & \cdots & \alpha_{1n}\beta_{m1} \cdots \alpha_{1n}\beta_{mm} \\ \vdots & & \vdots \\ \alpha_{n1}\beta_{11} \cdots \alpha_{n1}\beta_{1m} & \cdots & \alpha_{nn}\beta_{11} \cdots \alpha_{nn}\beta_{1m} \\ \vdots & & \vdots \\ \alpha_{n1}\beta_{m1} \cdots \alpha_{n1}\beta_{mm} & \cdots & \alpha_{nn}\beta_{m1} \cdots \alpha_{nn}\beta_{mm} \end{bmatrix}.$$

En una notación condensada cuyo significado es claro, podemos escribir esta matriz con

$$\begin{bmatrix} \alpha_{11}[B] & \cdots & \alpha_{1n}[B] \\ \vdots & & \vdots \\ \alpha_{n1}[B] & \cdots & \alpha_{nn}[B] \end{bmatrix}.$$

Esta matriz es conocida como el *producto de Kronecker* de $[A]$ y $[B]$ en ese orden. La regla para formarlo es fácil de describir en palabras: reemplácese cada elemento α_{ij} de la matriz n por n $[A]$ por la matriz m por $m\alpha_{ij}[B]$. Si en esta regla intercambiamos los papeles de A y B (y consecuentemente intercambiamos n y m) obtenemos la definición de producto de Kronecker de $[B]$ y $[A]$.

EJERCICIOS

1. Sabemos que el producto tensorial de $\mathcal{P}_n$ y $\mathcal{P}_m$ puede identificarse con el espacio $\mathcal{P}_{n,m}$ de polinomios en dos variables (véase el § 25, Ej. 2). Pruébese que si A y B son diferenciación sobre $\mathcal{P}_n$ y $\mathcal{P}_m$ respectivamente, y si $C = A \otimes B$, entonces C es diferenciación parcial mixta, esto es, si z $\mathcal{P}_{n,m}$, entonces $Cz = \dfrac{\partial^2 z}{\partial s \partial t}$.

2. Con la ordenación lexicográfica de la base de producto $\{x_i \otimes y_p\}$ resultó que la matriz de $A \otimes B$ es el producto de Kronecker de las matrices de A y B. ¿Hay una ordenación de los vectores de base, tal que la matriz de $A \otimes \cdot B$, referida al sistema de coordenadas así ordenado, es el producto de Kronecker de las matrices de B y A (en ese orden)?

3. Si A y B son transformaciones lineales, entonces

$$\rho(A \otimes B) = \rho(A)\rho(B).$$

§53. DETERMINANTES

Es, por supuesto, posible generalizar las consideraciones de la sección procedente a formas multilineales y productos tensoriales múltiples. En vez de entrar en esa parte del álgebra multilineal, marchamos en una dirección distinta; vamos directamente tras de los determinantes.

Supóngase que A es una transformación lineal sobre un espacio vectorial n-dimensional $\mathcal{V}$ y sea w una forma alterna n-lineal sobre $\mathcal{V}$. Si escribimos $\bar{A}w$ para la función definida por

$$(\bar{A}w)(x_1, \cdots, x_n) = w(Ax_1, \cdots, Ax_n),$$

entonces $\bar{A}w$ es una forma alterna n-lineal sobre $\mathcal{V}$, y, de hecho, $\bar{A}$ es una transformación n-lineal sobre el espacio de esas formas. Puesto que (véase el § 31) ese espacio es unidimensional, se sigue que $\bar{A}$ es igual a la multiplicación por un escalar apropiado. En otras palabras, existe un escalar δ tal que $\bar{A}w = \delta w$ para toda forma alterna n-lineal w. Por este procedimiento un tanto indirecto (de A a $\bar{A}$ a δ) hemos asociado un escalar δ determinado en forma única con cada transformación lineal A sobre $\mathcal{V}$; llamamos δ al *determinante* de A, y escribimos $\delta = \det A$. Obsérvese que *det* no es ni un escalar ni una transformación, sino una función que asocia un escalar con cada transformación lineal.

Nuestro propósito inmediato es estudiar la función *det*. Comenzamos por encontrar los determinantes de las transformaciones lineales más sencillas, esto es, las multiplicaciones por escalares. Si $Ax = \alpha x$ para toda x de $\mathcal{V}$, entonces

$$(\bar{A}w)(x_1, \cdots, x_n) = w(\alpha x_1, \cdots, \alpha x_n) = \alpha^n w(x_1, \cdots, x_n)$$

para toda forma alterna n-lineal w; se sigue que $\det A = \alpha^n$. Notamos, en particular, que $\det 0 = 0$ y $\det 1 = 1$.

En seguida preguntamos sobre las propiedades multiplicativas de *det*. Supóngase que A y B son transformaciones lineales sobre $\mathcal{V}$, y escríbase $C = AB$. Si w es una forma alterna n-lineal, entonces

$$\begin{aligned}(\bar{C}w)(x_1, \cdots, x_n) &= w(ABx_1, \cdots, ABx_n)\\ &= (\bar{A}w)(Bx_1, \cdots, Bx_n) = (\bar{B}\bar{A}w)(x_1, \cdots, x_n),\end{aligned}$$

de manera que $\bar{C} = \bar{B}\bar{A}$. Puesto que

$$\bar{C}w = (\det C)w$$

y

$$\bar{B}\bar{A}w = (\det B)\bar{A}w = (\det B)(\det A)w,$$

se sigue que

$$\det (AB) = (\det A)(\det B).$$

(Los valores de *det* son escalares y, en consecuencia, conmutan entre sí).

Una transformación lineal A es llamada *singular* si $\det A = 0$ y *no-singular* en el otro caso. Nuesto siguiente resultado es que A es invertible si y solamente si es no-singular. En realidad, si A es invertible, entonces

$$1 = \det 1 = \det (AA^{-1}) = (\det A)(\det A^{-1}),$$

y, en consecuencia, $\det A \neq 0$. Supóngase, por otra parte, que $\det A \neq 0$. Si $\{x_1, \ldots, x_n\}$ es una base de $\mathcal{V}$, y si w es una forma n-lineal alterna no cero sobre $\mathcal{V}$, entonces $(\det A)w(x_1, \ldots, x_n) \neq 0$, según el § 30, Teorema 3. Esto implica, según el § 30, Teorema 2, que el conjunto $\{Ax_1, \ldots, Ax_n\}$ es linealmente independiente (y, en consecuencia, una base); de esto, a su vez, inferimos que A es invertible.

En la terminología clásica el determinante se define como una función de matrices (no de transformaciones lineales); estamos ahora en condiciones de hacer contacto con ese punto de vista. Deduciremos una expresión para $\det A$ en términos de los elementos

α_{ij} de la matriz correspondiente a A en algún sistema de coordenadas $\{x_1, \ldots, x_n\}$. Sea w una forma n-lineal alterna no cero; sabemos que

$$(1) \qquad (\det A)w(x_1, \cdots, x_n) = w(Ax_1, \cdots, Ax_n).$$

Si reemplazamos cada Ax_j en el segundo miembro de (1) por $\sum_i \alpha_{ij}x_i$ y expandimos el resultado por multilinealidad, obtenemos una larga combinación lineal de términos tales como $w(z_1, \ldots, z_n)$ donde cada z es una de las x. (Compárese esta parte del argumento con la prueba del §30, Teorema 3). Si, en un término tal, dos de las z coinciden, entonces, puesto que w es alterna, ese término debe desaparecer. Si, por otra parte, todas las z son distintas, entonces $w(z_1, \ldots, z_n) = \pi w(x_1, \ldots, x_n)$ para alguna permutación π, y, además, para alguna permutación π puede ocurrir en esta forma. El coeficiente del término $\pi w(x_1, \ldots, x_n)$ es el producto $\alpha_{\pi(1),1} \ldots \alpha_{\pi(n),n}$. Puesto que (§ 30, Teorema 1) w es simétrico-oblicua, se sigue que

$$(2) \qquad \det A = \sum_\pi (\operatorname{sgn} \pi)\alpha_{\pi(1),1} \cdots \alpha_{\pi(n),n}$$

donde la suma se extiende sobre todas las permutaciones π de $\mathcal{S}_n$. [Recuérdese que $w(x_1, \ldots, x_n) \neq 0$, según el § 30, Teorema 3, de manera que la división por $w(x_1, \ldots, x_n)$ es legítima.]

De esta ecuación clásica (2) podemos deducir muchas propiedades especiales de determinantes por cómputo directo. Aquí está un ejemplo. Si σ y π son permutaciones (en $\mathcal{S}_n$) entonces (puesto que $\pi\sigma$ es también una permutación), se sigue que los productos $\alpha_{\pi(1),1} \ldots \alpha_{\pi(n),n}$ y $\alpha_{\pi\sigma(1),\sigma(1)} \ldots \alpha_{\pi\sigma(n),\sigma(n)}$ difieren solamente en el orden de sus factores. Si, para cada π, tomamos $\sigma = \pi^{-1}$ y, a continuación, alteramos cada sumando de (2) consecuentemente, obtenemos

$$\det A = \sum_\pi (\operatorname{sgn} \pi)\alpha_{1,\pi(1)} \cdots \alpha_{n,\pi(n)}.$$

(Nótese que $\operatorname{sgn} \pi = \operatorname{sgn} \pi^{-1}$ y que la suma sobre todas las π es la misma que la suma sobre todas las π^{-1}.) Puesto que esta última suma es justamente como la suma en (2), excepto que $\alpha_{i,\pi(i)}$ aparece en lugar de $\alpha_{\pi(i),i}$, se sigue de una aplicación de (2) a A' en lugar de A que

$$\det A = \det A'.$$

Aquí está otro hecho útil sobre determinantes. Si $\mathfrak{M}$ es un subespacio invariante bajo A, si B es la transformación A considerada sobre $\mathfrak{M}$ solamente, y si C es la transformación por cociente $A/\mathfrak{M}$, entonces

$$\det A = \det B \cdot \det C.$$

Esta relación multiplicativa es válida si, en particular, A es la suma directa de dos transformaciones B y C. La prueba puede basarse directamente sobre la definición de determinantes o, alternativamente, sobre la expansión obtenida en el párrafo precedente.

Si, para una transformación lineal fija A, escribimos $p(\lambda) = \det\,(A - \lambda)$ entonces p es una función del escalar λ; afirmamos que es, de hecho, un polinomio de grado n en λ y que el coeficiente de λ^n es $(-1)^n$. Para la prueba podemos usar la notación de (1). Es fácil ver que $w[(A - \lambda)x_1, \ldots, (A - \lambda)x_n]$ es una suma de términos tales como $\lambda^k w(y_1, \ldots, y_n)$, donde $y_i = x_i$ para exactamente k valores de i y $y_i = Ax_i$ para los restantes $n - k$ valores de $i (k = 0, 1, \ldots, n)$. El polinomio p se llama el *polinomio característico* de A; la ecuación $p = 0$, esto es, $\det(A - \lambda) = 0$ es la *ecuación característica de* A. Las raíces de la ecuación característica de A [esto es, los escalares α tales que $\det\,(A - \alpha) = 0$] se llaman las *raíces características* de A.

EJERCICIOS

1. Usense determinantes para obtener una nueva prueba del hecho de que si A y B son transformaciones lineales sobre un espacio vectorial finito-dimensional, y si $AB = 1$, entonces tanto A como B son invertibles.

2. Si A y B son transformaciones lineales tales que $AB = 0$, $A \neq 0$, $B \neq 0$, entonces $\det A = \det B = 0$.

3. Supóngase que (α_{ij}) es una matriz n por n, no singular, y supóngase que $A_1, \ldots, A_n$ son transformaciones lineales (sobre el mismo espacio vectorial). Pruébese que si las transformaciones lineales $\sum_j \alpha_{ij} A_j$, $i = 1, \cdots, n$, conmutan entre sí, entonces lo mismo es cierto para $A_1, \ldots, A_n$

4. Si $\{x_1, \ldots, x_n\}$ y $\{y_1, \ldots, y_n\}$ son bases en el mismo espacio vectorial, y si A es la transformación lineal tal que $Ax_i = y_i$, $i =,1 \ldots, n$, entonces $\det A \neq 0$.

5. Supóngase que $\{x_1, \ldots, x_n\}$ es una base en un espacio vectorial finito-dimensional $\mathfrak{V}$. Si $y_1, \ldots, y_n$ son vectores de $\mathfrak{V}$, escríbase $w(y_1, \ldots, y_n)$ para el determinante de la transformación lineal A, tal que $Ax_j = y_j$, $j = 1, \ldots, n$. Pruébese que w no es una forma alterna n-lineal.

6. Si, de acuerdo con el § 53, (2), el determinante de una matriz (α_{ij}) (no una transformación lineal) se define como $\sum_\pi (\operatorname{sgn} \pi)\alpha_{\pi(1),1} \cdots \alpha_{\pi(n),n}$, en-

tonces, para cada transformación lineal A, los determinantes de todas las matrices $[A; \mathfrak{X}]$ son todos iguales entre sí. (Aquí $\mathfrak{X}$ es una base arbitraria.)

7. Si (α_{ij}) es una matriz n por n, tal que $\alpha_{ij} = 0$ para más de $n^2 - n$ pares de valores de i y j, entonces $\det(\alpha_{ij}) = 0$.

8. Si A y B son trasformaciones lineales sobre espacios vectoriales de dimensiones n y m, respectivamente, entonces

$$\det(A \otimes B) = (\det A)^m \cdot (\det B)^n.$$

9. Si A, B, C y D son matrices tales que C y D conmuta y D es invertible, entonces (véase § 38, Ej. 19)

$$\det \begin{pmatrix} A & B \\ C & D \end{pmatrix} = \det(AD - BC).$$

(Sugestión: multiplíquese sobre la derecha por $\begin{pmatrix} 1 & 0 \\ X & 1 \end{pmatrix}$ ·) ¿Qué si D no es invertible? ¿Qué si C y D no conmutan?

10. ¿Tienen siempre A y A' el mismo polinomio característico?

11. (a) Si A y B son similares, entonces $\det A = \det B$.

(b) Si A y B son similares, entonces A y B tienen el mismo polinomio característico.

(c) Si A y B tienen el mismo polinomio característico, entonces $\det A = \det B$.

(d) ¿Es verdad la recíproca de cualquiera de estas aserciones?

12. Determínese el polinomio característico de la matriz (o, más bien, de la transformación lineal definida por la matriz)

$$\begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ \alpha_{n-1} & \alpha_{n-2} & \alpha_{n-3} & \cdots & \alpha_0 \end{bmatrix},$$

y conclúyase que todo polinomio es el polinomio característico de alguna transformación lineal.

13. Supóngase que A y B son transformaciones lineales sobre el mismo espacio vectorial finito-dimensional.

(a) Pruébese que si A es una proyección, entonces AB y BA tienen el mismo polinomio característico. (Sugestión: escójase una base que haga a la matriz de A tan simple como sea posible y, a continuación, calcúlese directamente con matrices.)

(b) Pruébese que, en todos los casos, AB y BA tienen el mismo polinomio característico. (Sugestión: encuéntrese un invertible P tal que PA sea una proyección y aplíquese (a) a PA y BP^{-1}.)

§54. VALORES PROPIOS

Un escalar λ es un *valor propio* y un vector x no cero es un *vector propio* de una transformación lineal A si $Ax = \lambda x$. Casi toda combinación de los adjetivos propio, latente, característico, eigen, y secular, con los nombres raíz, número y valor, han sido usados en la terminología para lo que llamamos un valor propio. Es importante estar sobre aviso del orden de selección en la definición; es un valor propio de A si existe un vector no cero x para el cual $Ax = \lambda x$ y un vector no cero x es un vector propio de A si existe un escalar λ para el cual $Ax = \lambda x$.

Supóngase que λ es un valor propio de A; sea $\mathfrak{M}$ la colección de todos los vectores x que son vectores propios de A que pertenecen a este valor propio, esto es, para el cual $Ax = \lambda x$. Puesto que, por nuestra definición, 0 no es un vector propio, $\mathfrak{M}$ no contiene a 0; si, en cambio, alargamos $\mathfrak{M}$ adjuntándole el origen, entonces se convierte en un subespacio. Definimos la multiplicidad del valor propio λ como la dimensión del subespacio $\mathfrak{M}$; un valor propio simple es aquel cuya multiplicidad es igual a 1. Por una obvia extensión de esta terminología, podemos expresar el hecho de que un escalar λ no es un valor propio de A, diciendo que λ es un valor de multiplicidad cero. El conjunto de valor propios de A se llama, algunas veces, el *espectro de* A. Nótese que el espectro de A es igual al conjunto de todos los escalares λ para los cuales $A - \lambda$ no es invertible.

Si el espacio vectorial con que estamos trabajando tiene la dimensión n, entonces el escalar 0 es un valor propio de multiplicidad n de la transformación lineal 0, y, de modo semejante, el escalar 1 es un valor propio de multiplicidad n de la transformación lineal 1. Puesto que $Ax = \lambda x$ si y sólo si $(A - \lambda)x = 0$, esto es, si y sólo si x está en el espacio nulo de $A - \lambda$, se sigue que la multiplicidad de λ como valor propio de A es igual a la nulidad de la transformación lineal $A - \lambda$. De esto, a su vez, se infiere (véase el § 50, Teorema 1), que los valores propios de A, juntamente con sus multiplicidades asociadas, son exactamente iguales a los de A'.

Observamos que B es cualquier transformación invertible, entonces

$$BAB^{-1} - \lambda = B(A - \lambda)B^{-1},$$

de manera que $(A - \lambda)x = 0$, si y sólo si $(BAB^{-1} - \lambda)Bx = 0$. Esto implica que todos los conceptos espectrales (por ejemplo, el espectro y las multiplicidades de los valores propios) son invariantes bajo el reemplazo de A por BAB^{-1}. Notamos también que si $Ax = \lambda x$, entonces

$$A^2x = A(Ax) = A(\lambda x) = \lambda(Ax) = \lambda(\lambda x) = \lambda^2 x.$$

Más generalmente, si p es un polinomio cualquiera, entonces $p(A)x = p(\lambda)x$, de manera que todo vector propio de A, que pertenece al valor propio λ, es también un vector propio de $p(A)$, que pertenece al valor propio $p(\lambda)$. De aquí que si A satisface cualquier ecuación de la forma $p(A) = 0$, entonces $p(\lambda) = 0$, para todo valor propio λ de A.

Puesto que una condición necesaria y suficiente para que $A - \lambda$ tenga un espacio nulo no trivial es que sea singular, esto es, que det $(A - \lambda) = 0$, se sigue que λ es un valor propio de A si y sólo si es una raíz característica de A. Este hecho es la razón de la importancia de los determinantes en el álgebra lineal. El concepto geométrico útil es el de un valor propio. Por la geometría de la situación, sin embargo, es imposible probar que existen valores propios. Por medio de determinantes reducimos el problema a un problema algebraico; resulta que los valores propios son lo mismo que raíces de cierta ecuación polinómica. No es extraño ahora que sea difícil probar que los valores propios existen siempre; las ecuaciones polinómicas no siempre tienen raíces y, correspondientemente, hay ejemplos fáciles de transformaciones lineales sin valores propios.

§55. MULTIPLICIDAD

La discusión de la sección precedente indica una de nuestras razones para desear el estudio de los espacios vectoriales complejos. Por el llamado teorema fundamental del álgebra, una ecuación polinomia sobre el campo de números complejos tiene siempre, cuando menos, una raíz; se sigue que una transformación lineal sobre un espacio vectorial complejo tiene siempre, cuando menos, un valor propio. Hay otros campos, además del campo de los números complejos, sobre los cuales es resoluble toda ecuación polinomia; se llaman campos *algebraicamente cerrados*. El resultado más general de la clase que buscamos en este momento es que toda transformación lineal de un espacio vectorial finito-dimensional tenga, cuando menos, un valor propio. En todo el resto de este capítulo (en las siguientes cuatro secciones) supondremos que nuestro campo de escalares es algebraicamente cerrado. El uso que haremos de este supuesto es el que se acaba de mencionar, a saber, que del mismo podemos concluir que existen siempre valores propios.

El punto de vista algebraico sobre valores propios sugiere otra posible definición de multiplicidad. Supóngase que A es una transformación lineal sobre un espacio vectorial finito-dimensional, y supóngase que λ es un valor propio de A. Podríamos desear considerar la multiplicidad de λ como una raíz de la ecuación característica de A. Este es un concepto útil, que llamaremos la multiplicidad *algebraica* de λ, para distinguirlo de otra noción *geométrica* anterior de multiplicidad.

Los dos conceptos de multiplicidad no coinciden, como lo muestra el ejemplo siguiente. Si D es diferenciación sobre el espacio $\mathcal{P}_n$ de todos los polinomios de grado $\leqq n-1$, entonces una condición necesaria y suficiente para que un vector x de $\mathcal{P}_n$ sea un vector propio de D es que $\frac{dx}{dt} = \lambda x(t)$ para algún número complejo λ. Tomamos en préstamo de la teoría elemental de las ecuaciones diferenciales el hecho de que toda solución de esta ecuación es una constante múltiple de $e^{\lambda t}$. Puesto que, a menos que $\lambda = 0$, solamente el múltiplo cero de $e^{\lambda t}$ es un polinomio (lo que debe ser, si pertenece a $\mathcal{P}_n$),debemos tener $\lambda = 0$ y $x(t) = 1$. En otras palabras, esta transformación particular tiene sólo un valor propio (que debe, en consecuencia, ocurrir con la multiplicidad algebraica n), a saber, $\lambda = 0$; pero, y esto es más perturbador, la dimensión del múltiple linear de dimensiones es exactamente uno. De aquí que si $n > 1$, las dos definiciones de multiplicidad dan valores diferentes. (En este argumento usamos el simple hecho de que una ecuación polinomia de grado n sobre un campo algebraicamente cerrado tiene exactamente n raíces, si las multiplicidades son contadas adecuadamente. Se sigue que una transformación lineal sobre un espacio vectorial n-dimensional tiene exactamente n valores propios, contando las multiplicidades algebraicas.)

Es completamente fácil ver que la multiplicidad geométrica de λ no es nunca mayor que su multiplicidad algebraica. En realidad, si A es una transformación lineal, si λ_0 es cualquiera de sus valores propios y si $\mathfrak{M}$ es el subespacio de soluciones de $Ax = \lambda_0 x$, entonces está claro que $\mathfrak{M}$ es invariante bajo A. Si A_0 es la transformación lineal A considerada sobre $\mathfrak{M}$ solamente, entonces está claro que $\det(A_0 - \lambda)$ es un factor de $\det(A - \lambda)$. Si la dimensión de $\mathfrak{M}$ ($=$ la multiplicidad geométrica de λ_0) es m, entonces $\det(A_0 - \lambda) = (\lambda_0 - \lambda)^m$; el resultado deseado se sigue de la definición de multiplicidad algebraica. Se sigue también que si $\lambda_1, \ldots, \lambda_p$ son valores propios distintos de A, con multiplicidades geométricas respectivas, $m_1, \ldots, m_p$, y si sucede que $\sum_{i=1}^{p} m_i = n$, entonces m_i es igual a la multiplicidad algebraica de λ_i para cada $i = 1, \ldots, p$.

Por medio de valores propios y sus multiplicidades algebraicas podemos caracterizar dos interesantes funciones de transformaciones lineales; una de ellas es el determinante y la otra es algo nuevo. (Advertencia: estas caracterizaciones son válidas sólo bajo el supuesto usual de que el campo escalar está algebraicamente cerrado.)

Sea A cualquier transformación lineal sobre un espacio vectorial n-dimensional, y sean $\lambda_1, \ldots, \lambda_p$ sus diferentes valores propios. Denotemos por m_j la multiplicidad algebraica de λ_j, $j = 1, \ldots, p$, de manera que $m_1 + \ldots + m_p = n$. Para cualquier ecuación polinomia

$$\alpha_0 + \alpha_1\lambda + \cdots + \alpha_n\lambda^n = 0,$$

el producto de las raíces es $(-1)^n\alpha_0/\alpha_n$ y la suma de las raíces es $-\alpha_{n-1}/\alpha_n$. Puesto que el coeficiente principal ($= \alpha_n$) del polinomio característico $\det(A - \lambda)$ es $(-1)^n$ y puesto que el término constante es $\det(A - 0) = \det A$, tenemos

$$\det A = \prod_{j=1}^{p} \lambda_j^{m_j}.$$

Esta caracterización del determinante motiva la definición

$$\operatorname{tr} A = \sum_{j=1}^{p} m_j\lambda_j;$$

la función así definida se llama la *traza* de A. No tendremos ocasión de usar la traza en lo que sigue; dejaremos la deducción de las propiedades básicas de la traza al lector que se interese.

EJERCICIOS

1. Encuéntrense todos los valores propios (complejos) y vectores propios de las siguientes matrices

(a) $\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$.

(b) $\begin{pmatrix} 1 & 0 \\ 0 & i \end{pmatrix}$.

(c) $\begin{pmatrix} 1 & 1 \\ 0 & i \end{pmatrix}$.

(d) $\begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}$.

(e) $\begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$.

2. Sea π una permutación de los enteros $\{1, \ldots, n\}$; si $x = (\xi_1, \ldots, \xi_n)$ es un vector de $\mathcal{C}^n$, escríbase $Ax = (\xi_{\pi(1)}, \ldots, \xi_{\pi(n)})$. Encuéntrese el espectro de A.
3. Pruébese que todos los valores propios de una proyección son 0 o 1 y que todos los valores propios de una involución son $+1$ o -1. (Este resultado no depende de la finito-dimensionalidad del espacio vectorial.)
4. Supóngase que A es una transformación lineal y que p es un polinomio. Sabemos que si λ es un valor propio de A, entonces $p(\lambda)$ es un valor propio de $p(A)$; ¿qué puede decirse sobre el recíproco?
5. Pruébese que el operador D de diferenciación sobre el espacio $\mathcal{P}_n$ ($n > 1$) no es reducible (esto es, no es reducido por ningún par no trivial de subespacios complementarios $\mathfrak{M}$ y $\mathfrak{N}$).
6. Si A es una transformación lineal sobre un espacio vectorial finito-dimensional, y si λ es un valor propio de A, entonces la multiplicidad algebraica de λ para A es igual a la multiplicidad algebraica de λ para BAB^{-1}. (Aquí B es una transformación arbitraria invertible.)
7. ¿Tienen AB y BA siempre el mismo espectro?
8. Supóngase que A y B son transformaciones lineales sobre espacios vectoriales finito-dimensionales.
(a) $\operatorname{tr}(A \oplus B) = \operatorname{tr} A + \operatorname{tr} B$.
(b) $\operatorname{tr}(A \otimes B) = (\operatorname{tr} A)(\operatorname{tr} B)$.
(c) El espectro de $A \oplus B$ es la unión de los espectros de A y B.
(d) El espectro de $A \otimes B$ consiste en todos los escalares de la forma $\alpha\beta$, con α y β en el espectro de A y de B, respectivamente.

§56. FORMA TRIANGULAR

Es ahora completamente fácil probar el más fácil de los llamados teoremas canónicos de forma. Nuestro supuesto sobre el campo escalar (a saber, que es algebraicamente cerrado) todavía está en vigor.

TEOREMA 1. *Si A es una transformación lineal sobre un espacio vectorial finito-dimensional $\mathcal{V}$, existen entonces $n+1$ subespacios $\mathfrak{M}_0, \mathfrak{M}_1, \cdots, \mathfrak{M}_{n-1}, \mathfrak{M}_n$ con las propiedades siguientes.*

(i) *cada $\mathfrak{M}_j$ $(j = 0, 1, \ldots, n-1, n)$ es invariante bajo A,*
(ii) *la dimensión de $\mathfrak{M}_j$ es j,*
(iii) $(\mathcal{O} =)\ \mathfrak{M}_0 \subset \mathfrak{M}_1 \subset \cdots \subset \mathfrak{M}_{n-1} \subset \mathfrak{M}_n\ (= \mathcal{V})$.

PRUEBA. Si $n = 0$ o $n = 1$, el resultado es trivial; procedemos por inducción, suponiendo que el enunciado es correcto para $n-1$. Considérese la transformación dual A' sobre $\mathcal{V}'$; puesto que tiene, cuando menos, un vector propio, digamos x', existe un subespacio unidimensional $\mathfrak{M}$ invariante bajo el mismo, a saber, el conjunto de todos los múltiplos de x'. Denotemos por $\bar{\mathfrak{M}}_{n-1}$ el aniquilador (en

$\mathcal{V}'' = \mathcal{V}$) de $\mathfrak{M}$, $\mathfrak{M}_{n-1} = \mathfrak{M}^0$; entonces $\mathfrak{M}_{n-1}$ es un subespacio $(n-1)$-dimensional de $\mathcal{V}$, y $\mathfrak{M}_{n-1}$ es invariante bajo A. Consecuentemente, podemos considerar A como una transformación lineal sobre $\mathfrak{M}_{n-1}$ solo y podemos encontrar $\mathfrak{M}_0, \mathfrak{M}_1, \cdots, \mathfrak{M}_{n-2}, \mathfrak{M}_{n-1}$, que satisfagan las condiciones (i), (ii) (iii). Escribimos $\mathfrak{M}_n = \mathcal{V}$, y está hecho.

El principal interés de este teorema viene de su interpretación matricial. Puesto que $\mathfrak{M}_1$ es unidimensional, podemos encontrar en el mismo un vector $x_1 \neq 0$. Puesto que $\mathfrak{M}_1 \subset \mathfrak{M}_2$, se sigue que x_1 está también en $\mathfrak{M}_2$, y puesto que $\mathfrak{M}_2$ es bidimensional, podemos encontrar en el mismo un vector x_2 tal que x_1 y x_2 sobretiendan a $\mathfrak{M}_2$. Procedemos, en esta forma, por inducción, escogiendo vectores x_j, de manera que $x_1, \ldots, x_j$ quedan en $\mathfrak{M}_j$ y sobretienden $\mathfrak{M}_j$ para $j = 1, \ldots, n$. Obtenemos, finalmente, una base $\mathfrak{X} = \{x_1, \ldots, x_n\}$ de $\mathcal{V}$; calculemos la matriz de A en este sistema de coordenadas. Puesto que x_j está en $\mathfrak{M}_j$ y puesto que $\mathfrak{M}_j$ es invariante bajo A, se sigue que Ax_j debe ser una combinación lineal de $x_1, \ldots, x_j$. De aquí que, en la expresión

$$Ax_j = \sum_i \alpha_{ij} x_i$$

el coeficiente de x_i debe desaparecer siempre que $i > j$; en otras palabras, $i > j$ implica $\alpha_{ij} = 0$. De aquí que la matriz de A tenga la *forma triangular*

$$[A] = \begin{bmatrix} \alpha_{11} & \alpha_{12} & \alpha_{13} & \cdots & \alpha_{1n} \\ 0 & \alpha_{22} & \alpha_{23} & \cdots & \alpha_{2n} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & \cdots & \alpha_{n-1,n} \\ 0 & 0 & 0 & \cdots & \alpha_{nn} \end{bmatrix}.$$

Está claro de esta representación que det $(A - \alpha_{ii}) = 0$. para $i = 1, \ldots, n$, de manera que α_{ii} son los valores propios de A, que aparecen en la diagonal principal de $[A]$ con las multiplicidades propias. Resumimos como sigue:

Teorema 2. *Si A es una transformación lineal sobre un espacio vectorial n-dimensional $\mathcal{V}$, existe entonces una base $\mathfrak{X}$ de $\mathcal{V}$ tal que la matriz $[A; \mathfrak{X}]$ es triangular; o, equivalentemente, si $[A]$ es una matriz cualquiera, existe una matriz no singular $[B]$ tal que $[B]^{-1}[A][B]$ es triangular.*

La forma triangular es útil para probar muchos resultados sobre transformaciones lineales. Se sigue de la misma, por ejemplo, que para cualquier polinomio p, los valores propios de $p(A)$, incluyendo sus multiplicidades algebraicas, son precisamente los números $p(\lambda)$, donde λ corre a través de los valores propios de A.

Una gran parte de la teoría de las transformaciones lineales está dedicada a mejorar el resultado de triangularización que se acaba de obtener. Lo mejor que puede ser una matriz, no es ser triangular, sino *diagonal* (esto es, $\alpha_{ij} = 0$, a menos que $i = j$); si una transformación lineal es tal que su matriz es diagonal con respecto de un sistema adecuado de coordenadas, llamaremos a la transformación *diagonable*.

EJERCICIOS

1. Interprete las siguientes matrices como transformaciones lineales sobre $\mathcal{C}^2$ y, en cada caso, encuentre una base de $\mathcal{C}^2$ tal que la matriz de la transformación con respecto de esa base sea triangular.

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

(b) $\begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}$.

(c) $\begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$.

(d) $\begin{pmatrix} \cdot & 1 \\ 1 & 1 \end{pmatrix}$.

(e) $\begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}$.

(f) $\begin{pmatrix} 0 & 1 & 1 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}$.

2. Dos transformaciones lineales conmutativas sobre un espacio vectorial finito-dimensional $\mathcal{V}$ sobre un campo algebraicamente cerrado, pueden ser triangularizadas simultáneamente. En otras palabras, si $AB = BA$, entonces existe una base $\mathfrak{X}$ tal que tanto $[A, \mathfrak{X}]$ como $[B, \mathfrak{X}]$ son triangulares. [Sugestión: para imitar la prueba del §56 es deseable encontrar un subespacio $\mathfrak{M}$ de $\mathcal{V}$ invariante tanto bajo A como bajo B. Con esto presente, considérese cualquier valor propio de λ de A y examínese el conjunto de todas las soluciones de $Ax = \lambda x$ para el papel de $\mathfrak{M}$.]

3. Formúlense y pruébense los análogos de los resultados de §56 para matrices triangulares bajo la diagonal (en vez de arriba de ella).

4. Supóngase que A es una transformación lineal sobre un espacio vectorial n-dimensional. Para cada forma alterna n-lineal w, escríbase $\bar{A}w$ para la función definida por

$$(\bar{A}w)(x_1, \cdots, x_n) = w(Ax_1, x_2, \cdots, x_n) + w(x_1, Ax_2, \cdots, x_n) + \cdots + w(x_1, x_2, \cdots, Ax_n).$$

Puesto que $\bar{A}w$ es una forma alterna n-lineal y, de hecho, $\bar{A}$ es una transformación lineal sobre el espacio (unidimensional) de esas formas, se sigue que $\bar{A}w = \tau(A)\cdot w$, donde $\tau(A)$ es un escalar.

(a) $\tau(0) = 0$.

(b) $\tau(1) = n$.

(c) $\tau(A + B) = \tau(A) + \tau(B)$.

(d) $\tau(\alpha A) = \alpha\tau(A)$.

(e) Si el campo escalar tiene característica cero y si A es una proyección, entonces $\tau(A) = \rho(A)$.

(f) Si (α_{ij}) es la matriz de A en un sistema de coordenadas, entonces $\tau(A) = \sum_i \alpha_{ii}$.

(g) $\tau(A') = \tau(A)$.

(h) $\tau(AB) = \tau(BA)$.

(i) ¿Para cuáles permutaciones π de los enteros $1, \ldots, k$ es verdad que $\tau(A_1 \ldots A_k) = \tau(A_{\pi(1)} \ldots A_{\pi(k)})$ para todos los k-tuples $(A_1, \ldots, A_k)$ de transformaciones lineales?

(j) Si el campo de escalares está algebraicamente cerrado, entonces $\tau(A) = \operatorname{tr} A$. (Por esta razón la traza se define usualmente como τ; el procedimiento más popular es usar (f) como la definición.)

5. (a) Supóngase que el campo escalar tiene característica cero. Pruébese que si $E_1, \ldots, E_k$ y $E_1 + \ldots + E_k$ son proyecciones, entonces $E_iE_j = 0$ siempre que $i \neq j$. (Sugestión: del hecho de que $\operatorname{tr}(E_1 + \ldots + E_k) = \operatorname{tr}(E_1) + \ldots + \operatorname{tr}(E_k)$ conclúyase que el alcance de $E_1 + \ldots + E_k$ es la suma directa de los alcances de $E_1 \ldots E_k$).

(b) Si $A_1, \ldots, A_k$ son transformaciones lineales sobre un espacio vectorial n-dimensional, y si $A_1 + \ldots + A_k = 1$ y $\rho(A_1) + \ldots + \rho(A_k) \leqq n$; entonces cada A_i es una proyección y $A_iA_j = 0$, siempre que $i \neq j$. (Comiéncese con $k = 2$ y procédase por inducción; úsese un argumento de suma directa, como en (a).)

6. (a) Si A es una transformación lineal sobre un espacio vectorial finito-dimensional de característica cero, y si $\operatorname{tr} A = 0$, entonces existe una base $\mathfrak{X}$ tal que si $[A; \mathfrak{X}] = (\alpha_{ij})$, entonces $\alpha_{ii} = 0$ para toda i. (Sugestión: usando el hecho de que A no es escalar, pruébese primero que existe un vector x tal que x y Ax son linealmente independientes. Esto prueba que puede hacerse que α_{11} desaparezca; procédase por inducción.)

(b) Demuéstrese que si la característica no es cero, la conclusión de (a) es falsa. (Sugestión: si la característica es 2, calcúlese $BC - CB$, donde $B = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ y $C = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$.)

§57. NILPOTENCIA

Como ayuda para obtener un teorema de representación más informativo que el triangular, procedemos a introducir y a estudiar una clase muy especial, pero muy útil de transformaciones. Una transformación lineal A es llamada *nilpotente* si existe un entero estrictamente positivo q, tal que $A^q = 0$; el menor entero q es el *índice* de nilpotencia.

TEOREMA 1. *Si A es una transformación lineal nilpotente de índice q sobre un espacio vectorial finito-dimensional $\mathcal{V}$, y si x_0 es un vector para el cual $A^{q-1}x_0 \neq 0$,* entonces los vectores *$x_0, Ax_0, \ldots, A^{q-1}x_0$ son linealmente independientes. Si $\mathcal{H}$ es el espacio sobretendido por estos vectores, entonces existe un subespacio $\mathcal{K}$ tal que $\mathcal{V} = \mathcal{H} \oplus \mathcal{K}$* y tal que el par $(\mathcal{H}, \mathcal{K})$ *reduce a A.*

PRUEBA. Para probar la independencia lineal, supóngase que $\sum_{i=0}^{q-1} \alpha_i A^i x_0 = 0$, y sea j el índice mínimo tal que $\alpha_j = 0$. (No excluimos la posibilidad de que $j = 0$.) Dividiendo por $-\alpha_j$ y cambiando la notación en una forma obvia, obtenemos una relación de la forma

$$A^j x_0 = \sum_{i=j+1}^{q-1} \alpha_i A^i x_0 = A^{j+1}(\sum_{i=j+1}^{q-1} \alpha_i A^{i-j-1} x_0) = A^{j+1} y.$$

Se sigue de la definición de q que

$$A^{q-1} x_0 = A^{q-j-1} A^j x_0 = A^{q-j-1} A^{j+1} y = A^q y = 0;$$

puesto que esto contradice la selección de x_0, debemos tener $\alpha_j = 0$ para cada j.

Está claro que $\mathcal{H}$ es invariante bajo A; para construir $\mathcal{K}$ procedemos por inducción sobre el índice q de nilpotencia. Si $q = 1$, el resultado es trivial; suponemos ahora el teorema para $q - 1$. El alcance $\mathcal{R}$ de A es un subespacio que es invariante bajo A; restringida a $\mathcal{R}$ la transformación lineal A es nilpotente de índice $q - 1$. Escribimos $\mathcal{H}_0 = \mathcal{H} \cap \mathcal{R}$ y $y_0 = Ax_0$; entonces $\mathcal{H}_0$ es sobretendido por los vectores linealmente independientes $y_0, Ay_0, \ldots, A^{q-2}y_0$. La hipótesis de inducción puede aplicarse y podemos concluir que $\mathcal{R}$ es la suma directa de $\mathcal{H}_0$ y algún otro subespacio invariante $\mathcal{K}_0$.

Escribimos $\mathcal{K}_1$ para el conjunto de todos los vectores x, tales que Ax está en $\mathcal{K}_0$; está claro que $\mathcal{K}_1$ es un subespacio. Es grande la tentación de hacer $\mathcal{K} = \mathcal{K}_1$ e intentar probar que $\mathcal{K}$ tiene las propiedades deseadas. Infortunadamente, no es necesario que esto sea verdadero; $\mathcal{H}$ y $\mathcal{K}_1$ no necesitan ser disjuntos. (Es verdad, pero no usaremos ese hecho, que la intersección de $\mathcal{H}$ y $\mathcal{K}_1$ está contenida en el espacio nulo de A.) Que, a pesar de esto, $\mathcal{K}_1$ sea útil es motivado por el hecho de que $\mathcal{H} + \mathcal{K}_1 = \mathcal{V}$. Para probar esto, obsérvese que Ax está en $\mathcal{R}$ para toda x y, consecuentemente, $Ax =$

$y + z$ con y en $\mathcal{K}_0$ y z en $\mathcal{K}_0$. El elemento general de $\mathcal{K}_0$ es una combinación lineal de $Ax_0, \ldots, A^{q-1}x_0$; de aquí tenemos

$$y = \sum_{i=1}^{q-1} \alpha_i A^i x_0 = A(\sum_{i=0}^{q-2} \alpha_{i+1} A^i x_0) = Ay_1,$$

donde y_1 está en $\mathcal{K}$. Se sigue que $Ax = Ay_1 + z$, o $A(x - y_1) = z$, de manera que $A(x - y_1)$ está en $\mathcal{K}_0$. Esto significa que $x - y_1$ está en $\mathcal{K}_1$, de manera que x es la suma de un elemento (a saber y_1) de $\mathcal{K}$ y un elemento (a saber, $x - y_1$) de $\mathcal{K}_1$.

Por lo que concierne a la disyunción, podemos decir cuando menos que $\mathcal{K} \cap \mathcal{K}_0 = \mathcal{O}$. Para probar esto, supóngase que x está en $\mathcal{K} \cap \mathcal{K}_0$, y obsérvese primero que Ax está en $\mathcal{K}_0$ (puesto que x está en $\mathcal{K}$). Puesto que $\mathcal{K}_0$ es también invariante bajo A, el vector Ax pertenece a $\mathcal{K}_0$ juntamente con x, de manera que $Ax = 0$. De esto inferimos que x está en $\mathcal{K}_0$. (Puesto que x está en $\mathcal{K}$, tenemos $x = \sum_{i=0}^{q-1} \alpha_i A^i x_0$; y, en consecuencia, $0 = Ax = \sum_{i=1}^{q-1} \alpha_{i-1} A^i x_0$; de la independencia lineal de las $A^j x_0$ se sigue que $\alpha_0 = \ldots \alpha_{q-2} = 0$, de manera que $x = \alpha_{q-1} A^{q-1} x_0$). Hemos probado que si x pertenece a $\mathcal{K} \cap \mathcal{K}_0$, entonces pertenece también a $\mathcal{K}_0 \cap \mathcal{K}_0$, y de aquí que $x = 0$.

La situación es ahora ésta: $\mathcal{K}$ y $\mathcal{K}_1$ juntamente sobretienden a $\mathcal{V}$, y $\mathcal{K}_1$ contiene los dos subespacios disjuntos $\mathcal{K}_0$ y $\mathcal{K} \cap \mathcal{K}_1$. Si suponemos que $\mathcal{K}'_0$ es un complemento cualquiera de $\mathcal{K}_0 \oplus (\mathcal{K} \cap \mathcal{K}_1)$ en $\mathcal{K}_1$, esto es, si

$$\mathcal{K}'_0 \oplus \mathcal{K}_0 \oplus (\mathcal{K} \cap \mathcal{K}_1) = \mathcal{K}_1,$$

entonces podemos escribir $\mathcal{K} = \mathcal{K}'_0 \oplus \mathcal{K}_0$; afirmamos que esta $\mathcal{K}$ tiene las propiedades deseadas. En primer lugar $\mathcal{K} \subset \mathcal{K}_1$ y $\mathcal{K}$ es disjunta de $\mathcal{K} \cap \mathcal{K}_1$; se sigue que $\mathcal{K} \cap \mathcal{K} = \mathcal{O}$. En segundo lugar $\mathcal{K} \oplus \mathcal{K}$ contiene tanto a $\mathcal{K}$ como a $\mathcal{K}_1$, de manera que $\mathcal{K} \oplus \mathcal{K} = \mathcal{V}$. Finalmente, $\mathcal{K}$ es invariante bajo A, puesto que el hecho de que $\mathcal{K} \subset \mathcal{K}_1$ implica que $A\mathcal{K} \subset \mathcal{K}_0 \subset \mathcal{K}$. La prueba del teorema es completa.

Posteriormente, necesitaremos la siguiente observación. Si $\tilde{x}_0$ es cualquier otro vector para el cual $A^{q-1}x_0 \neq 0$, si $\tilde{\mathcal{K}}$ es el subespacio sobretendido por los vectores $\tilde{x}_0, A\tilde{x}_0, \ldots, A^{q-1}\tilde{x}_0$ y si, finalmente, $\tilde{\mathcal{K}}$ es cualquier subespacio que juntamente con $\tilde{\mathcal{K}}$ reduce a A, entonces el comportamiento de A sobre $\tilde{\mathcal{K}}$ y $\tilde{\mathcal{K}}$ es el mismo que su comportamiento sobre $\mathcal{K}$ y $\mathcal{K}$ respectivamente. (En otras palabras, a pesar de la aparente univocidad en el enunciado del Teorema 1, todo está, de hecho, determinado en forma unívoca, hasta los isomorfismos.) La verdad de esta observación se sigue del hecho de que el índice de nilpotencia de A sobre $\mathcal{X}$ (r, digamos) es el mismo que el índice de nilpotencia de A sobre $\bar{\mathcal{X}}$ ($\bar{r}$, digamos). Este hecho, a su vez, se prueba como sigue. Puesto que $A^r\mathcal{V} = A^r\mathcal{X} + A^r\mathcal{X}$ y también $A^r\mathcal{V} = A^r\bar{\mathcal{X}} + A^r\bar{\mathcal{X}}$ (estos resultados dependen de la invariancia de todos los subespacios involucrados), se sigue que pueden ser igualadas las dimensiones de los segundos miembros de estas ecuaciones, y de aquí que $(q - r) + 0 = (q - r) + (\bar{r} - r)$.

Usando el Teorema 1 podemos encontrar una completa caracterización geométrica de transformaciones nilpotentes.

Teorema. 2. *Si A es una transformación lineal nilpotente de índice q sobre un espacio vectorial finito-dimensional $\mathcal{V}$, entonces existen enteros positivos r, $q_1, \ldots, q_r$ y vectores $x_1, \ldots, x_r$ tales que (i) $q_1 \geqq \ldots \geqq q_r$, (ii) los vectores*

$$\begin{array}{l} x_1, Ax_1, \cdots, A^{q_1-1}x_1, \\ x_2, Ax_2, \cdots, A^{q_2-1}x_2, \\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot \\ x_r, Ax_r, \cdots, A^{q_r-1}x_r \end{array}$$

forman una base para $\mathcal{V}$, y (iii) $A^{q_1}x_1 = A^{q_2}x_2 = \ldots = A^{q_r}x_r = 0$. Los enteros r, $q_1, \ldots, q_r$ que forman un conjunto completo de invariantes de isomorfismo de A. Si, en otras palabras, B es cualquier otra transformación lineal nilpotente sobre un espacio vectorial finito-dimensional $\mathcal{W}$, entonces una condición necesaria y suficiente para que exista un isomorfismo T entre $\mathcal{V}$ y $\mathcal{W}$ tal como $TAT^{-1} = B$ es que los enteros r, $q_1, \ldots, q_r$ aplicados a B sean iguales que los aplicados a A.

Prueba. Escribimos $q_1 = q$ y escogemos x_1 como un vector cualquiera para el cual $A^{q_1-1}x_1 \neq 0$. El subespacio sobretendido por $x_1, Ax_1, \ldots, A^{q_1-1}x_1$ es invariante bajo A, y, por el Teorema 1, posee un complemento invariante, que, naturalmente, tiene una dimensión estrictamente más baja que $\mathcal{V}$. Sobre este subespacio complementario, A es nilpotente de índice q_2, digamos; aplicamos el mismo proceso de reducción a este subespacio (comenzando con un vector x_2 para el cual $A^{q_2-1}x_2 \neq 0$). Continuamos así por inducción hasta agotar el espacio. Esto prueba la parte existencial del teorema; la parte restante se sigue de la univocidad (hasta llegar a isomorfismos) de la descomposición dada por el Teorema 1.

Con respecto a la base $\{A^i x_j\}$ descrita en el Teorema 2, la matriz de A toma una forma sencilla. Todo elemento de la matriz que no está en la diagonal situada exactamente debajo de la diagonal principal, desaparece (esto es, $a_{ij} \neq 0$ implica $j = i - 1$), y los elementos situados debajo de la diagonal principal comienzan, (en la parte superior) con una serie de unos seguidos por un solo 0, continúan con otra serie de unos seguidos por un 0 y continúan así hasta el fin, con las longitudes de las series de unos disminuyendo monótonamente (o, en todo caso, no decreciendo).

Obsérvese que nuestro supuesto sobre la oclusión algebraica del campo de escalares no se usó en esta sección.

EJERCICIOS

1. ¿Existe una transformación nilpotente de índice 3 sobre un espacio bidimensional?
2. (a) Pruébese que una transformación lineal nilpotente sobre un espacio vectorial finito-dimensional tiene traza cero.
(b) Pruébese que si A y B son transformaciones lineales (sobre el mismo espacio vectorial finito-dimensional) y si $C = AB - BA$, entonces $1 - C$ no es nilpotente.
3. Pruébese que si A es una transformación lineal nilpotente de índice q, sobre espacio vectorial tridimensional, entonces

$$\nu(A^{k+1}) + \nu(A^{k-1}) \leq 2\nu(A^k)$$

para $k = 1, \ldots, q - 1$.
4. Si A es una transformación lineal (sobre un espacio vectorial finito-dimensional, sobre un espacio algebraicamente cerrado), entonces existen transformaciones lineales B y C tales que $A = B + C$, B es diagonable, C es nilpotente y $BC = CB$; las transformaciones B y C son determinadas en forma única por estas condiciones.

§58. FORMA DE JORDAN

Es una sana intuición geométrica lo que nos hace conjeturar que, para transformaciones lineales, el ser invertibles y el ser cero en algún sentido son nociones exactamente opuestas. Nuestro desconcierto al encontrar que el alcance y el espacio nulo no es necesario que estén disjuntos, está conectado con esta conjetura. La situación puede ser arreglada relajando el sentido en que interpretamos "ser cero"; para la mayor parte de los propósitos prácticos una transformación lineal, alguna potencia de la cual es cero (esto es, una transformación nilpotente) es tan de naturaleza cero, como puede esperarse que lo sea. Aunque no podemos decir que una transformación lineal es invertible o "cero", aún en el sentido ampliado de "ceridad" o naturaleza cero, podemos decir cómo cualquier transformación está compuesta de estas dos clases extremas.

Teorema 1. *Toda transformación A sobre un espacio vectorial finito-dimensional $\mathcal{V}$ es la suma directa de una transformación nilpotente y de una transformación invertible.*

Prueba. Consideramos el espacio nulo de la k-ésima potencia de A; es éste un subespacio $\mathfrak{N}_k = \mathfrak{N}(A^k)$. Claramente $\mathfrak{N}_1 \subset \mathfrak{N}_2 \subset \ldots$. Afirmamos primero que si siempre $\mathfrak{N}_k = \mathfrak{N}_{k+1}$, entonces $\mathfrak{N}_k = \mathfrak{N}_{k+j}$ para todos los enteros positivos j. En efecto, si $A^{k+j}x = 0$, entonces $A^{k+1}A^{j-1}x = 0$, de donde (por el hecho de que $\mathfrak{N}_k = \mathfrak{N}_{k+1}$) se sigue que $A^kA^{j-1}x = 0$, y, en consecuencia, que $A^{k+j-1}x = 0$. En otras palabras $\mathfrak{N}_{k+j}$ está contenida en (y, en consecuencia, es igual a) $\mathfrak{N}_{k+j-1}$; la inducción sobre j establece nuestra aserción.

Puesto que $\mathcal{V}$ es finito-dimensional, los subespacios $\mathfrak{N}_k$ no pueden continuar aumentando indefinidamente; sea q el entero positivo más pequeño para el cual $\mathfrak{N}_q = \mathfrak{N}_{q+1}$. Está claro que $\mathfrak{N}_q$ es invariante bajo A (de hecho cada $\mathfrak{N}_k$ es así). Escribimos $\mathfrak{R}_k = \mathfrak{R}(A^k)$ para el alcance de A^k (de manera que, de nuevo, está claro que $\mathfrak{R}_q$ es invariante bajo A); probaremos que $\mathcal{V} = \mathfrak{N}_q \oplus \mathfrak{R}_q$ y que A sobre $\mathfrak{N}_q$ es nilpotente, mientras que sobre $\mathfrak{R}_q$ es invertible.

Si x es un vector común a $\mathfrak{N}_q$ y a $\mathfrak{R}_q$, entonces $A^qx = 0$ y $x = A^qy$, para alguna y. Se sigue que $A^{2q}y = 0$ y de aquí que, según la definición de q, que $x = A^qy = 0$. Hemos demostrado así que el alcance y el espacio nulo de A^q son disjuntos; un argumento de dimensionalidad (véase el §50, Teorema 1) demuestra que sobretienden a $\mathcal{V}$, de manera que $\mathcal{V}$ es su suma directa. Se sigue de la definición de q y $\mathfrak{N}_q$ que A sobre $\mathfrak{N}_q$ es nilpotente de índice q. Finalmente, si x está en $\mathfrak{R}_q$ (de manera que $x = A^qy$ para alguna y) y si $Ax = 0$, entonces $A^{q+1}y = 0$, de donde $x = A^qy = 0$; esto demuestra que A es invertible sobre $\mathfrak{R}_q$. La prueba del Teorema 1 es completa.

La descomposición de A en sus partes nilpotentes e invertible es única. Supóngase, en realidad, que $\mathcal{V} = \mathcal{H} \oplus \mathcal{K}$ de manera que A sobre $\mathcal{H}$ es nilpotente y A sobre $\mathcal{K}$ es invertible. Puesto que $\mathcal{H} \subset \mathfrak{N}(A^k)$ para alguna k, se sigue que $\mathcal{H} \subset \mathfrak{N}_q$, y. puesto que $\mathcal{K} \subset \mathfrak{R}(A^k)$ para toda k, se sigue que $\mathcal{K} \subset \mathfrak{R}_q$; estos hechos juntos implican que $\mathcal{H} = \mathfrak{N}_q$ y $\mathcal{K} = \mathfrak{R}_q$.

Podemos ahora usar nuestros resultados sobre transformaciones nilpotentes para estudiar la estructura de transformaciones arbitrarias. El método para obtener una transformación nilpotente de una transformación arbitraria puede parecer como una treta mágica, pero es una treta útil, que se emplea a menudo. Lo que es esencial es la existencia garantizada de valores propios; por esa razón continuamos suponiendo que el campo escalar está algebraicamente cerrado (véase el §55).

TEOREMA 2. *Si A es una transformación lineal sobre un espacio vectorial finito-dimensional $\mathcal{V}$, y si $\lambda_1, \ldots, \lambda_p$ son los distintos valores propios de A con multiplicidades algebraicas respectivas $m_1, \ldots, m_p$, entonces $\mathcal{V}$ es la suma directa de p subespacios $\mathfrak{M}_1, \cdots, \mathfrak{M}_p$ de dimensiones respectivas $m_1, \ldots, m_p$, tales que cada $\mathfrak{M}_j$ es invariante bajo A y tales que $A - \lambda_j$ es nilpotente sobre $\mathfrak{M}_j$.*

PRUEBA. Tómese cualquier $j = 1, \ldots, p$, fija y considérese la transformación $A_j = A - \lambda_j$. A A_i podemos aplicar la descomposición del Teorema 1 para obtener los subespacios $\mathfrak{M}_j$ y $\mathfrak{N}_j$ tales que A_j es nilpotente sobre $\mathfrak{M}_j$ e invertible sobre $\mathfrak{N}_j$. Puesto que $\mathfrak{M}_j$ es invariante bajo A_j, es también invariante bajo $A_j + \lambda_j = A$. De aquí que, para toda λ el determinante $A - \lambda$ es el producto de los dos determinantes correspondientes para las dos transformaciones lineales en que A se convierte, cuando la consideramos sobre $\mathfrak{M}_j$ y $\mathfrak{N}_j$ separadamente. Puesto que el único valor propio de A sobre $\mathfrak{M}_j$ es λ_j, y puesto que A sobre $\mathfrak{N}_j$ no tiene el valor propio λ_j (esto es $A - \lambda_j$ es invertible sobre $\mathfrak{N}_j$), se sigue que la dimensión de $\mathfrak{M}_j$ es exactamente $\mathfrak{M}_j$ y que cada uno de los subespacios $\mathfrak{M}_j$ está disjunto del ámbito de todos los otros. Un argumento de dimensión prueba que $\mathfrak{M}_1 \oplus \cdots \oplus \mathfrak{M}_p = \mathcal{V}$ y, en consecuencia, concluye la prueba del teorema.

Procedemos a describir los principales resultados de esta sección y la precedente en lenguaje matricial. Si A es una transformación lineal sobre un espacio vectorial finito-dimensional $\mathcal{V}$, entonces con respecto de una base adecuada de $\mathcal{V}$, la matriz de A tiene la forma siguiente. Todo elemento que no está inmediatamente sobre o debajo de la diagonal principal desaparece. Sobre la diagonal principal aparecen los distintos valores propios de A, cada uno un número de veces igual a su multiplicidad algebraica. Debajo de cualquier valor particular propio aparecen sólo unos y ceros y éstos en la forma siguiente: hay cadenas de unos, seguidas de un solo 0, con las longitudes de las cadenas decreciendo al leerse del extremo superior al inferior. Esta matriz es la *forma de Jordan* o la *forma canónica* *clásica* de A; tenemos $B = TAT^{-1}$ si y sólo si las formas canónicas de A y B son iguales, excepto por lo que hace al orden de los valores propios. (Así, en particular, una transformación lineal A es diagonable si y sólo si su forma canónica clásica es ya diagonal, esto es, si cada cadena de unos tiene longitud cero.)

Introduzcamos alguna notación. Supóngase que A tiene p valores propios distintos $\lambda_1, \ldots, \lambda_p$, con multiplicidades algebraicas $m_1, \ldots, m_p$, como antes; sea r_j el número de cadenas de unos bajo λ_j y sean $q_{j,1} - 1, q_{j,2} - 1, \ldots, q_{j,r_j} - 1$ las longitudes de estas cadenas. El polinomio e_{ji} definido por $e_{ji}(\lambda) = (\lambda - \lambda_j)^{q_{j,i}}$ es llamado un *divisor elemental* de A de *multiplicidad* $q_{j,i}$, perteneciente al valor propio λ_j. Un divisor elemental es llamado *simple* si su multiplicidad es 1 (de manera que la correspondiente longitud de la cadena es 0); vemos que una transformación lineal es diagonal si y sólo si sus divisores elementales son simples.

Para ilustrar la potencia del Teorema 2 hacemos una aplicación. Podemos expresar el hecho de que la transformación $A - \lambda_j$ sobre $\mathfrak{M}_j$ es nilpotente de índice q_{j1} diciendo que la transformación A sobre $\mathfrak{M}_j$ es anulada por el polinomio e_{j1}. Se sigue que A sobre $\mathcal{V}$ es anulada (por el producto de los divisores elementales de las multiplicidades más altas); este producto se llama el *polinomio mínimo de A*. Es sumamente fácil ver (puesto que el índice de nilpotencia de $A - \lambda_j$ sobre $\mathfrak{M}_j$ es exactamente $q_{j,1}$) que este polinomio está determinado en forma única (hasta un factor multiplicativo) como el polinomio de menor grado que anula a A. Puesto que el polinomio característico de A es el producto de todos los divisores elementales y, en consecuencia, un múltiplo del polinomio mínimo, obtenemos la *ecuación de Hamilton-Cayley*: toda transformación lineal es anulada por su polinomio característico.

EJERCICIOS

1. Encuéntrese la forma de Jordan de $\begin{pmatrix} 1 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix}$.

2. ¿Cuál es el número máximo de transformaciones lineales no similares por pares, sobre un espacio vectorial tridimensional, cada una de las cuales tiene un polinomio característico $(\lambda - 1)^3$?

3. ¿Tiene raíz cuadrada toda transformación lineal invertible? (Decir que A es una raíz cuadrada de B significa, por supuesto, que $A^2 = B$).

4. (a) Pruébese que si ω es una raíz cúbica de 1 ($\omega \neq 1$), entonces las matrices

$$\begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix} \quad \text{y} \quad \begin{pmatrix} 1 & 0 & 0 \\ 0 & \omega & 0 \\ 0 & 0 & \omega^2 \end{pmatrix}$$

son similares.

(b) Descúbrase y pruébese una generalización de (a) a dimensiones más elevadas.

5. (a) Pruébese que las matrices $\begin{pmatrix} 0 & 1 & \alpha \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$ y $\begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$ son similares.

(b) Descríbase y pruébese una generalización de (a) a dimensiones más elevadas.

6. (a) Demuéstrese que las matrices

$$\begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix} \quad \text{y} \quad \begin{pmatrix} 3 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

son similares (sobre, digamos, el campo de números complejos).

(b) Descríbase y pruébese una generalización de (a) a dimensiones más elevadas.

7. Si dos matrices reales son similares sobre $\mathcal{C}$, entonces son similares sobre $\mathcal{R}$.

8. Pruébese que toda matriz es similar a su transpuesta.

9. Si A y B son matrices n por n tales que las matrices $2n$ por $2n$ $\begin{pmatrix} A & 0 \\ 0 & A \end{pmatrix}$ y $\begin{pmatrix} B & 0 \\ 0 & B \end{pmatrix}$ son similares, entonces A y B son similares.

10. ¿Cuáles de las siguientes matrices son diagonales (sobre el campo de números complejos)?

(a) $\begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$.

(b) $\begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$.

c) $\begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix}$.

(d) $\begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}$.

(e) $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$.

¿Qué sobre el campo de números reales?

11. Demuéstrese que la matriz

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

es diagonable sobre el campo de números complejos, pero no sobre el campo de números reales.

12. Sea π una permutación de los enteros $\{1, \ldots, n\}$; si $x = (\xi_1, \ldots, \xi_n)$ es un vector de $\mathbb{C}^n$, escríbase $Ax = (\xi_{\pi(1)}, \ldots, \xi_{\pi(n)})$. Pruébese que A es diagonable y encuéntrese una base con respecto de la cual es diagonal la matriz de A.

13. Supóngase que A es una transformación lineal y que $\mathfrak{M}$ es un subespacio invariante bajo A. Pruébese que si A es diagonable, también lo es la restricción de A a $\mathfrak{M}$.

14. ¿Bajo qué condiciones impuestas a los números complejos $\alpha_1, \ldots, \alpha_n$ es diagonable (sobre el campo de números complejos) la matriz

$$\begin{bmatrix} 0 & \cdots & 0 & \alpha_1 \\ 0 & \cdots & \alpha_2 & 0 \\ \vdots & \cdots & \vdots & \vdots \\ \alpha_n & \cdots & 0 & 0 \end{bmatrix}$$

15. ¿Son verdaderas o falsas las siguientes aserciones?

(a) Una matriz real dos por dos, con un determinante negativo es similar a una matriz diagonal.

(b) Si A es una transformación lineal sobre un espacio vectorial complejo, y si $A^k = 1$ para algún entero positivo k, entonces A es diagonable.

(c) Si A es una transformación lineal nilpotente sobre un espacio vectorial finito-dimensional, entonces A es diagonable.

16. Si A es una transformación lineal sobre un espacio vectorial finito-dimensional sobre un campo algebraicamente cerrado, y si todo valor propio de A tiene multiplicidad algebraica 1, entonces A es diagonable.

17. Si el polinomio mínimo de una transformación lineal A sobre un espacio vectorial n-dimensional tiene grado, entonces A es diagonable.

18. Encuéntrense los polinomios mínimos de todas las proyecciones e involuciones.

19. ¿Cuál es el polinomio mínimo de la matriz?

$$\begin{bmatrix} \lambda_1 & 0 & 0 & \cdots & 0 \\ 0 & \lambda_2 & 0 & \cdots & 0 \\ 0 & 0 & \lambda_3 & \cdots & 0 \\ \cdot & \cdot & \cdot & \cdots & \cdot \\ 0 & 0 & 0 & \cdots & \lambda_n \end{bmatrix}?$$

20. (a) ¿Cuál es el polinomio mínimo del operador de diferenciación sobre $\mathcal{P}_n$?

(b) ¿Cuál es el polinomio mínimo de la transformación A sobre $\mathcal{P}_n$ definido por $(Ax)(t) = x(t+1)$?

21. Si A es una transformación lineal con polinomio mínimo p, y si q es un polinomio tal que $q(A) = 0$, entonces q es divisible entre p.

22. (a) Si A y B son transformaciones lineales, si p es un polinomio tal que $p(AB) = 0$, y si $q(t) = tp(t)$, entonces $q(BA) = 0$.

(b) ¿Qué puede ser inferido de (a) sobre la relación entre los polinomios mínimos de AB y de BA?

23. Una transformación lineal es invertible si y sólo si el término constante del polinomio mínimo es diferente de cero.

CAPITULO 3

ORTOGONALIDAD

§59. PRODUCTOS INTERIORES

Pongamos ahora nuestros pies sobre la tierra. Comenzamos en el Cap. 1 señalando que deseábamos generalizar ciertas propiedades elementales de ciertos espacios elementales como $\mathcal{R}^2$. Hasta ahora hemos hecho esto en nuestro estudio, pero hemos dejado por completo de considerar un aspecto de $\mathcal{R}^2$. Hemos estudiado el concepto de linearidad; lo que hemos pasado completamente por alto es el concepto usual cuantitativo de ángulo y longitud. En el presente capítulo colmaremos ese vacío, sobrepondremos a los espacios vectoriales que han de estudiarse ciertas funciones numéricas, correspondientes a las nociones ordinarias de ángulo y longitud y estudiaremos la nueva estructura así obtenida (espacio vectorial, más función numérica dada). Por la mayor profundidad de visión geométrica que ganamos de este modo, debemos sacrificar algo de generalidad; en todo el resto de este libro tendremos que suponer que el campo subyacente de escalares es, o el campo $\mathcal{R}$ de números reales, o el campo $\mathcal{C}$ de números complejos.

Para una sugestión sobre cómo proceder, primero inspeccionamos $\mathcal{R}^2$. Si $x = (\xi_1, \xi_2)$ y $y = (\eta_1, \eta_2)$ son dos puntos cualesquiera de $\mathcal{R}^2$, la fórmula usual para la distancia entre x y y, o la longitud del segmento que une x y y, es $\sqrt{(\xi_1 - \eta_1)^2 + (\xi_2 - \eta_2)^2}$. Es conveniente introducir la notación

$$\| x \| = \sqrt{\xi_1^2 + \xi_2^2}$$

para la distancia desde x al origen $0 = (0, 0)$; en esta notación la distancia entre x y y se vuelve $\| x - y \|$.

Esto, por ahora, sobre longitudes y distancia. ¿Qué hay sobre ángulos? Resulta que es mucho más conveniente estudiar, en el caso general, no cualquiera de las medidas usuales de ángulos, sino sus cosenos. (Hablando en términos generales, la razón para esto es que el ángulo, en la representación usual del círculo de radio uno, es la longitud de cierto arco circular, mientras que el coseno del ángulo es la longitud de un segmento de línea; este último es mucho más fácil de relacionar con nuestro estudio precedente de las funciones lineales). Supóngase entonces que α es el ángulo entre el segmento de 0 a x y el eje positivo ξ_1, y sea β el ángulo entre el segmento de 0 a y y el mismo eje; el ángulo entre los dos vectores x y y es $\alpha - \beta$, de manera que su coseno es

$$\cos(\alpha - \beta) = \cos\alpha \cos\beta + \operatorname{sen}\alpha \operatorname{sen}\beta = \frac{\xi_1\eta_1 + \xi_2\eta_2}{\|x\| \cdot \|y\|}.$$

Considérese la expresión $\xi_1\eta_1 + \xi_2\eta_2$; por medio de la misma podemos expresar tanto el ángulo como la longitud, por medio de fórmulas muy sencillas. Hemos visto ya que si conocemos la distancia entre 0 y x para toda x, entonces podemos calcular la distancia entre x y y; afirmamos ahora que si para cada par de vectores x y y se nos da el valor de $\xi_1\eta_1 + \xi_2\eta_2$, entonces, en términos de este valor, podemos calcular todas las distancias y todos los ángulos. En realidad, si tomamos $x = y$, entonces $\xi_1\eta_1 + \xi_2\eta_2$ se vuelve $\xi_1^2 + \xi_2^2 = \|x\|^2$, y esto resuelve la cuestión de las longitudes; la fórmula anterior del coseno nos da el ángulo en términos de $\xi_1\eta_1 + \xi_2\eta_2$ y de las dos longitudes $\|x\|$ y $\|y\|$. Para tener una notación concisa, escribimos, para $x = (\xi_1, \xi_2)$ y $y = (\eta_1, \eta_2)$,

$$\xi_1\eta_1 + \xi_2\eta_2 = (x, y);$$

lo que dijimos anteriormente se resume por las relaciones

$$\text{distancia de } 0 \text{ a } x = \|x\| = \sqrt{(x, x)},$$

$$\text{distancia de } x \text{ a } y = \|x - y\|,$$

$$\text{coseno del ángulo entre } x \text{ y } y = \frac{(x, y)}{\|x\| \cdot \|y\|}.$$

Las importantes propiedades de (x, y), considerado como una función numérica del par de vectores x y y, son las siguientes: es simétrico en x y y, depende linealmente de cada una de sus dos variables y (a menos que $x = 0$) el valor de (x, x) es siempre estrictamente positivo. (El conflicto notacional entre el uso de paréntesis en (x, y) y en $(\xi_1\ \xi_2)$ es sólo aparente. Podría suscitarse solamente en espacios bidimensionales y aun allí se evita fácilmente la confusión).

Obsérvese por un momento la muy trivial imagen de $\mathcal{R}^1$. Para $x = (\xi_1)$ y $y = (\eta_1)$ deberíamos tener, en este caso, $(x, y) = \xi_1\eta_1$ (y esto es por la razón de que (x, y) es conocido como el *producto interior* o *producto escalar* de x y y). El ángulo entre dos vectores cualesquiera es 0 o π, de manera que su coseno es $+1$ o -1. Esto demuestra la mayor sensibilidad de la función dada por (x, y), que toma todos los posibles valores numéricos.

§60. PRODUCTOS INTERIORES COMPLEJOS

¿Qué sucede si deseamos considerar $\mathcal{C}^2$ en vez de $\mathcal{R}^2$? La generalización parece quedar a la mano; para $x = (\xi_1, \xi_2)$ y $y = (\eta_1, \eta_2)$ (donde ahora las ξ y las η pueden ser números complejos) escribimos $(x, y) = \xi_1\eta_1 + \xi_2\eta_2$ y esperamos que las expresiones $\|x\| = (x, x)$ y $\|x - y\|$ puedan usarse como medidas sensibles de distancia. Sin embargo, obsérvese el siguiente extraño fenómeno (cuando $i = \sqrt{-1}$):

$$\| ix \|^2 = (ix, ix) = i(x, ix) = i^2(x, x) = -\| x \|^2.$$

Esto significa que si $\|x\|$ es positivo, esto es, si x está a una distancia positiva desde el origen, entonces ix no está; de hecho, la distancia de 0 a ix es imaginaria. Esto es muy desagradable; seguramente es razonable exigir que cualquier cosa que sea lo que va a jugarse el papel de (x, y) en este caso, debería tener la propiedad de que para $x = y$ nunca se hiciera negativa. Está a mano un remedio formal; podríamos tratar de escribir

$$(x, y) = \xi_1\bar{\eta}_1 + \xi_2\bar{\eta}_2$$

(donde la barra denota conjugación compleja. En esta definición la expresión (x, y) pierde mucho de su anterior belleza; no es ya simétrica en x y y no es más completamente lineal en cada una de sus variables. Pero, y esto es lo que nos impulsó a dar nuestra nueva definición

$$(x, x) = \xi_1\bar{\xi}_1 + \xi_2\bar{\xi}_2 = |\xi_1|^2 + |\xi_2|^2$$

seguramente nunca es negativa. Es dudoso a priori si una útil y elegante teoría puede construirse sobre la base de una función que no llega a poseer tantas de las propiedades que la recomendaron en primer lugar a nuestra atención; la aparente inelegancia se justificará en lo que sigue por su buen éxito. Es éste un alegre portento. Considérese el espacio $\mathcal{C}^1$ (esto es, el conjunto de todos los números complejos). Es imposible trazar una imagen de alguna configuración en este espacio y a continuación estar en condiciones es distinguirlo de una configuración en $\mathcal{R}^2$, pero conceptualmente es claramente un espacio diferente. El análogo de (x, y) en este espacio, para $x = (\xi_1)$ y $y = (\eta_1)$ es dado por $(x, y) = \xi_1\eta_1$ y esta expresión tiene una sencilla interpretación geométrica. Si unimos x y y al origen por segmentos de rectas, (x, y), no será, de seguro, el coseno del ángulo entre los dos segmentos; resulta que, para $||x|| = ||y|| = 1$, su parte real es exactamente ese coseno.

Los conjugados complejos que nos vimos forzados a introducir aquí vendrán a abrumarnos después; por ahora dejamos esta introducción heurística y volvemos al trabajo formal, después de sólo un comentario más sobre la notación. La similaridad de los símbolos (,) y [,], uno usado aquí para el producto interior y el otro usado anteriormente para las funciones lineales no es accidental. Demostraremos más tarde que es, de hecho, sólo la presencia de la conjugación compleja en (,) lo que hace necesario usar para la misma un símbolo diferente de [,]. Por ahora, sin embargo, no nos podemos dar el lujo de confundir los dos.

§61. ESPACIOS DE PRODUCTOS INTERIORES

DEFINICIÓN. Un *producto interior*, en un espacio vectorial (real o complejo) es una función valuada numéricamente (respectivamente, real o compleja), del par ordenado de vectores x y y, tal que

(1) $$(x, y) = \overline{(y, x)},$$

(2) $$(\alpha_1 x_1 + \alpha_2 x_2, y) = \alpha_1(x_1, y) + \alpha_2(x_2, y),$$

(3) $$(x, x) \geqq 0; \quad (x, x) = 0 \text{ si y sólo si } x = 0.$$

Un *espacio de producto interior* es un espacio vectorial con un producto interior.

Observamos que en el caso de un espacio vectorial real, la conjugacion en (1) puede ser pasada por alto. Sin embargo, en cualquier caso, real o complejo (1) implica que (x, x) es siempre real, de manera que tiene sentido la desigualdad de (3). En un espacio de producto interior usaremos la notación

$$\sqrt{(x, x)} = \| x \|;$$

el número $\| x \|$ se llama la *norma* o *longitud* del vector x. Un espacio real de producto interior es algunas veces llamado un *espacio euclidiano;* su análogo complejo es llamado un *espacio unitario.*

Como ejemplos de espacios unitarios podemos considerar $\mathbb{C}^n$ y $\mathcal{P}$; en el primer caso escribimos, para $x = (\xi_1, \ldots, \xi_n)$ y $y = (\eta_1, \ldots, \eta_n)$,

$$(x, y) = \sum_{i=1}^{n} \xi_i \bar{\eta}_i,$$

y, en $\mathcal{P}$, escribimos

$$(x, y) = \int_0^1 x(t)\overline{y(t)}\, dt.$$

Las modificaciones que convierten estos ejemplos en espacios euclidianos (esto es, en espacios reales de productos interiores), son obvias.

En un espacio unitario tenemos

$$(2') \qquad (x, \alpha_1 y_1 + \alpha_2 y_2) = \bar{\alpha}_1 (x, y_1) + \bar{\alpha}_2 (x, y_2).$$

(Para transformar el primer miembro de (2′) en el segundo miembro, usamos (1), desarrollamos (2) y usamos (1) de nuevo. Este hecho, juntamente con la definición de un producto interior explica la terminología usada algunas veces para describir las propiedades (1), (2) y (3) [y su consecuencia (2′)]. De acuerdo con esa terminología (x, y) es una forma *simétrica hermitiana* (1), *bilineal conjugada* [(2) y (2′)], y *definida positiva* (3). En un espacio euclidiano la conjugación en (2′) puede ser pasada por alto, juntamente con la conjugación en (1); en el caso (x, y), se llama una forma simétrica, bilineal y definida positiva. Observamos que en ambos casos, las condiciones impuestas a (x, y) implican para $\| x \|$ la propiedad de homogeneidad

$$\| \alpha x \| = | \alpha | \cdot \| x \|.$$

(Prueba: $\| \alpha x \|^2 = (\alpha x, \alpha x) = \alpha \bar{\alpha} (x, x)$.)

§62. ORTOGONALIDAD

La más importante relación entre los vectores de un espacio de producto interior es la ortogonalidad. Por definición, los vectores x y y se llaman *ortogonales* si $(x, y) = 0$. Observamos que esta relación es simétrica; puesto que $(x, y) = (y, x)$, se sigue que (x, y) y (y, x) desaparecen juntamente. Si recordamos la motivación para la introducción de (x, y) la terminología se explica por sí misma, dos vectores son ortogonales (o perpendiculares) si el ángulo que hacen entre sí es de 90°, esto es, si el coseno de este ángulo es 0. Dos subespacios se llaman ortogonales si cada vector de uno es ortogonal a cada vector del otro.

Un conjunto $\mathfrak{X}$ de vectores es *ortonormal* si siempre que x y y estén en $\mathfrak{X}$ se sigue que $(x, y) = 0$, o $(x, y) = 1$, según que $x \neq y$ o $x = y$. (Si $\mathfrak{X}$ es finito, digamos, $\mathfrak{X} = \{x_1, \ldots, x_n\}$, tenemos $(x_i, x_j) = \delta_{ij}$). Llamamos *completo* a un conjunto ortonormal si no está contenido en un conjunto ortonormal mayor.

Para dar nuestra última definición en este respecto, observamos primero que un conjunto ortonormal es linealmente independiente. En realidad, si $\{x_1, \ldots, x_k\}$ es un cualquier subconjunto finito de un conjunto ortonormal $\mathfrak{X}$, entonces $\sum_i \alpha_i x_i = 0$ implica que

$$0 = (\sum_i \alpha_i x_i, x_j) = \sum_i \alpha_i (x_i, x_j) = \sum_i \alpha_i \delta_{ij} = \alpha_j;$$

en otras palabras, una combinación lineal de las x puede desaparecer sólo si desaparecen todos los coeficientes. De aquí concluimos que en un espacio finito-dimensional de producto interior el número de vectores de un conjunto ortonormal es siempre finito y, de hecho, no mayor que la dimensión lineal del espacio. Definimos, en este caso, la *dimensión ortogonal* del espacio como el mayor número de vectores que puede contener un conjunto ortonormal.

Advertencia: por lo que sabemos en esta etapa, los conceptos de ortogonalidad y de conjunto ortonormales son vacuos. Pueden usarse ejemplos triviales para demostrar que las cosas no están así de mal; el vector 0, por ejemplo, es siempre ortogonal a todo vector y, si el espacio contiene un vector x no cero, entonces el conjunto que consiste en sólo $\dfrac{x}{\|x\|}$ es un conjunto ortonormal. Garantizamos que estos ejemplos no son muy inspiradores. Por ahora, sin embargo, nos contentamos con ellos; pronto veremos que hay siempre "suficientes" vectores ortogonales para operar cómodamente.

Obsérvese también que no tenemos derecho a suponer que el número de elementos de un conjunto completo ortonormal es igual a la dimensión ortonormal. El punto es éste: si tuviéramos un conjunto ortonormal con ese número de elementos, está claro que sería completo; es concebible, igualmente, que algún otro conjunto contenga menos elementos, pero que también sea completo porque su ingrata estructura excluya la posibilidad de expandirlo. Estas dificultades son puramente verbales y se desvanecerán en el momento en que comencemos a probar cosas; ocurren solamente porque de entre las diversas posibilidades de la definición de completividad tuvimos que escoger una definida y tenemos que probar su equivalencia con las otras.

Necesitamos alguna notación. Si $\mathcal{E}$ es cualquier conjunto de vectores de un espacio de producto interior $\mathcal{V}$, denotamos por $\mathcal{E}^{\perp}$ el conjunto de todos los vectores de $\mathcal{V}$ que son ortogonales a todo vector de $\mathcal{E}$. Está claro que $\mathcal{E}^{\perp}$ es un subespacio de $\mathcal{V}$ (sea o no $\mathcal{E}$ uno) y que $\mathcal{E}$ está contenido en $\mathcal{E}^{\perp\perp} = (\mathcal{E}^{\perp})^{\perp}$. Se sigue que el subespacio sobretendido por $\mathcal{E}$ está contenido en $\mathcal{E}^{\perp\perp}$. En caso de que $\mathcal{E}$ sea un subespacio, llamaremos a $\mathcal{E}^{\perp}$ el *complemento ortogonal* de $\mathcal{E}$. Usamos el signo a fin de recordar la ortogonalidad (o perpendicularidad). En discusiones informales, $\mathcal{E}^{\perp}$ puede ser llamado "E perp".

EJERCICIOS

1. Dados cuatro números complejos α, β, γ y δ, tratamos de definir un producto interior en $\mathcal{C}^2$ escribiendo

$$(x, y) = \alpha\xi_1\bar{\eta}_1 + \beta\xi_2\bar{\eta}_1 + \gamma\xi_1\bar{\eta}_2 + \delta\xi_2\bar{\eta}_2$$

siempre que $x = (\xi_1, \xi_2)$ y $y = (\eta_1, \eta_2)$. ¿Bajo qué condiciones impuestas a α, β, γ y δ define esta ecuación un producto interior?

2. Pruébese que si x y y son vectores de un espacio unitario, entonces

$$4(x, y) = \| x + y \|^2 - \| x - y \|^2 + i \| x + iy \|^2 - i \| x - iy \|^2.$$

3. Si un producto interior en $\mathcal{P}_{n+1}$ se define como $(x, y) = \int_0^1 x(t)\overline{y(t)}\, dt$, y si $x_j(t) = t^j$, $j = 0, \ldots, n - 1$, encuéntrese un polinomio de grado n, ortogonal a $x_0, x_1, \ldots, x_{n-1}$.

4. (a) Dos vectores x y y en un espacio real de producto interior son ortogonales si y sólo si $\| x + y \|^2 = \| x \|^2 + \| y \|^2$.

(b) Demuéstrese que (a) es falsa si "real" es cambiado por "complejo"

(c) Dos vectores x y y de un espacio complejo de producto interior son ortogonales si y sólo si $\| \alpha x + \beta y \|^2 = \| \alpha x \|^2 + \beta y \|^2$ para todos los pares de escalares α y β.

(d) Si x y y son vectores en un espacio real de producto interior, y si $\| x \| = \| y \|$, entonces $x - y$ y $x + y$ son ortogonales. (¿Representación?) Discútase el enunciado correspondiente para espacios complejos.

(e) Si x y y son vectores, de un espacio de producto interior, entonces

$$\|x+y\|^2 + \|x-y\|^2 = 2\|x\|^2 + 2\|y\|^2.$$

¿Representación?

§63. COMPLETIVIDAD

TEOREMA 1. *Si $\mathfrak{X} = \{x_1, \cdots, x_n\}$ es un conjunto finito ortogonal cualquiera en un espacio de producto interior, si x es un vector cualquiera, y si $\alpha_i = (x, x_i)$ entonces (desigualdad de Bessel)*

$$\sum_i |\alpha_i|^2 \leq \|x\|^2.$$

El vector $x' = x - \sum_i \alpha_i x_i$ es ortogonal a cada x_j y, consecuentemente, al subespacio sobretendido por $\mathfrak{X}$.

PRUEBA. Para la primera aserción:

$$\begin{aligned} 0 \leq \|x'\|^2 &= (x', x') = (x - \textstyle\sum_i \alpha_i x_i,\, x - \sum_j \alpha_j x_j) \\ &= (x, x) - \textstyle\sum_i \alpha_i (x_i, x) - \sum_j \bar{\alpha}_j (x, x_j) + \sum_i \sum_j \alpha_i \bar{\alpha}_j (x_i, x_j) \\ &= \|x\|^2 - \textstyle\sum_i |\alpha_i|^2 - \sum_i |\alpha_i|^2 + \sum_i |\alpha_i|^2 \\ &= \|x\|^2 - \textstyle\sum_i |\alpha_i|^2; \end{aligned}$$

para la segunda aserción:

$$(x', x_j) = (x, x_j) - \sum_i \alpha_i (x_i, x_j) = \alpha_j - \alpha_j = 0.$$

TEOREMA 2. *Si $\mathfrak{X}$ es un conjunto finito ortogonal cualquiera en un espacio de producto interior $\mathfrak{V}$, las siguientes seis condiciones impuestas a $\mathfrak{X}$ son equivalentes entre sí.*

(1) *El conjunto ortonormal $\mathfrak{X}$ es completo.*
(2) *Si $(x, x_i) = 0$, para $i = 1, \ldots, n$, entonces $x = 0$.*
(3) *El espacio sobretendido por $\mathfrak{X}$ es todo el espacio $\mathfrak{V}$.*
(4) *Si x está en $\mathfrak{V}$, entonces $x = \sum_i (x, x_i) x_i$.*
(5) *Si x y y están en $\mathfrak{V}$, entonces (identidad de Parseval)*

$$(x, y) = \sum_i (x, x_i)(x_i, y).$$

(6) *Si x está en $\mathfrak{V}$, entonces*

$$\|x\|^2 = \sum_i |(x, x_i)|^2.$$

PRUEBA. Estableceremos las implicaciones (1) $\Rightarrow$ (2) $\Rightarrow$ (3) $\Rightarrow$ (4) $\Rightarrow$ (5) $\Rightarrow$ (6) $\Rightarrow$ (1). Así, primero suponemos (1) y probamos (2), a continuación suponemos (2) y probamos (3), y así, sucesivamente, hasta que finalmente probamos (1) suponiendo (6).

(1) $\Rightarrow$ (2). Si $(x, x_i) = 0$ para toda i y $x \neq 0$, entonces podemos adjuntar $x/\| x \|$ a $\mathfrak{X}$ y así obtener un conjunto ortonormal mayor que $\mathfrak{X}$.

(2) $\Rightarrow$ (3). Si hay una x que no sea una combinación lineal de las x_i, entonces, por la segunda parte del Teorema 1, $x' = x - \sum_i (x, x_i)x_i$ es diferente de 0 y es ortogonal a cada x_i.

(3) $\Rightarrow$ (4). Si toda x tiene la forma $x = \sum_j \alpha_j x_j$, entonces

$$(x, x_i) = \sum_j \alpha_j(x_j, x_i) = \alpha_i.$$

(4) $\Rightarrow$ (5). Si $x = \sum_i \alpha_i x_i$ y $y = \sum_j \beta_j x_j$, con $\alpha_i = (x, x_i)$ y $\beta_j = (y, x_j)$, entonces

$$(x, y) = (\textstyle\sum_i \alpha_i x_i, \sum_j \beta_j x_j) = \sum_i \alpha_i \bar{\beta}_j (x_i, x_j) = \sum_i \alpha_i \bar{\beta}_i.$$

(5)$\Rightarrow$(6). Conjunto $x = y$.

(6) $\Rightarrow$ (1). Si $\mathfrak{X}$ estuviera contenida en un conjunto ortogonal más grande, digamos si x_0 es ortogonal a cada x_i, entonces

$$\| x_0 \|^2 = \textstyle\sum_i |(x_0, x_i)|^2 = 0,$$

de manera que $x_0 = 0$.

§64. DESIGUALDAD DE SCHWARZ

TEOREMA. *Si x y y son vectores de un espacio de producto interior, entonces (desigualdad de Schwarz)*

$$|(x, y)| \leqq \| x \| \cdot \| y \|.$$

PRUEBA. Si $y = 0$, ambos miembros desaparecen. Si $y \neq 0$, entonces el conjunto consistente en el vector $y/\| y \|$ es ortonormal y, consecuentemente, por la desigualdad de Bessel

$$|(x, y/\| y \|)|^2 \leqq \| x \|^2.$$

La desigualdad de Schwarz tiene importantes consecuencias aritméticas, geométricas y analíticas.

(1) En cualquier espacio de producto interior definimos la distancia $\delta(x, y)$ entre dos vectores x y y como

$$\delta(x, y) = \| x - y \| = \sqrt{(x - y, x - y)}.$$

A fin de que δ merezca ser llamada una distancia, debería tener las siguientes tres propiedades

(i) $\delta(x, y) = \delta(y, x)$,
(ii) $\delta(x, y) \geqq 0$; $\delta(x, y) = 0$ si y sólo si $x = y$,
(iii) $\delta(x, y) \leqq \delta(x, z) + \delta(z, y)$.

(En un espacio vectorial es también grato estar seguro de que la distancia es invariante bajo las traslaciones:

(iv) $\delta(x, y) = \delta(x + z, y + z)$.)

Las propiedades (i), (ii) y (iv) son obviamente poseídas por la δ particular que definimos; la única cuestión es la validez de la "desigualdad triangular" (iii). Para probar (iii), observamos que

$$\begin{aligned}\| x + y \|^2 &= (x + y, x + y) = \| x \|^2 + (x, y) + (y, x) + \| y \|^2 \\ &= \| x \|^2 + (x, y) + \overline{(x, y)} + \| y \|^2 \\ &= \| x \|^2 + 2 \operatorname{Re} (x, y) + \| y \|^2 \\ &\leqq \| x \|^2 + 2 |(x, y)| + \| y \|^2 \\ &\leqq \| x \|^2 + 2\| x \| \cdot \| y \| + \| y \|^2 \\ &= (\| x \| + \| y \|)^2;\end{aligned}$$

reemplazando x por $x - z$ y y por $z - y$, obtenemos

$$\| x - y \| \leqq \| x - z \| + \| z - y \|,$$

y esto es equivalente a (iii). (Usamos $\operatorname{Re} \zeta$ para denotar la parte real del número completo ζ; si $\zeta = \xi + i\eta$ con ξ y η reales, entonces $\operatorname{Re} \zeta = \xi$. La parte imaginaria de ζ, esto es, el número real η, es denotado por $\operatorname{Im} \zeta$).

(2) En el espacio euclidiano $\mathfrak{R}^n$, la expresión

$$\frac{(x, y)}{\| x \| \cdot \| y \|}$$

da el coseno del ángulo entre x y y. La desigualdad de Schwarz en este caso simplemente equivale al enunciado de que el coseno de un ángulo real es $\leqq 1$.

(3) En el espacio unitario $\mathbb{C}^n$, la desigualdad de Schwarz se vuelve la llamada desigualdad de Cauchy; afirma que para dos secuencias cualesquiera $(\xi_1, \ldots, \xi_n)$ y $(\eta_1, \ldots, \eta_n)$ de números complejos, tenemos

$$|\sum_{i=1}^{n} \xi_i \bar{\eta}_i|^2 \leq \sum_{i=1}^{n} |\xi_i|^2 \cdot \sum_{i=1}^{n} |\eta_i|^2.$$

(4) En el espacio $\mathcal{P}$, la desigual de Schwarz se vuelve

$$|\int_0^1 x(t)\overline{y(t)}\, dt|^2 \leq \int_0^1 |x(t)|^2\, dt \cdot \int_0^1 |y(t)|^2\, dt.$$

Es útil observar que las relaciones mencionadas en (1)-(4) anteriormente, son no sólo análogas a la desigualdad general de Schwarz, sino de hecho consecuencias o casos especiales de la misma.

(5) Mencionamos de paso que hay lugar entre las dos nociones (espacios generales vectoriales y espacios de productos interiores) para un concepto intermedio de cierto interés. Este concepto es el de un espacio *normado*, un espacio vectorial en que hay una definición aceptable de longitud, pero no se dice nada de ángulos. Una norma en un espacio vectorial (real o complejo) es una función valuada numéricamente $||x||$ de todos los vectores x, tal que $||x|| \geq 0$, a menos que $x = 0$, $||\alpha x|| = |\alpha| . ||x||$, y $||x + y|| \leq ||x|| + ||y||$. Nuestra discusión hasta ahora muestra que un espacio de producto interior es un espacio vectorial normado; la recíproca no es en general verdadera. En otras palabras, si todo lo que se nos da es una norma que satisface las tres condiciones que se acaban de dar, puede no ser posible encontrar un producto interior para el cual (x, x) sea idénticamente igual a $||x||^2$. En términos un tanto vagos, aunque quizás sugestivos, podemos decir que la norma de un espacio de producto interior tiene un carácter esencialmente "cuadrático" que las normas no poseen por lo general.

§65. CONJUNTOS ORTONORMALES COMPLETOS

TEOREMA. *Si $\mathcal{V}$ es un espacio n-dimensional de producto interior, existen entonces conjuntos completos ortonormales en $\mathcal{V}$ y todo conjunto completo ortonormal de $\mathcal{V}$, contiene exactamente n elementos. La dimensión ortonormal de $\mathcal{V}$ es la misma que su dimensión lineal.*

PRUEBA. Para gentes que no tienen inconveniente en dedicarse a la búsqueda de un elemento en un posible conjunto incontable, es obvia la existencia de conjuntos completos ortonormales. En realidad, ya hemos visto que existen conjuntos ortonormales, de manera que escogemos uno; si no es completo, lo agrandamos y si el conjunto ortonormal resultante todavía no es completo, lo agrandamos de nuevo, y procedemos en esta forma por inducción. Puesto que un conjunto ortonormal puede contener cuando más n elementos, a lo sumo en n pasos llegaremos a un conjunto ortonormal completo. Este conjunto sobretiende todo el espacio [véase el § 63, Teorema 2, (1) $\Rightarrow$ (3)], y, puesto que es también linealmente independiente, es una base y contiene en consecuencia precisamente n elementos. Esto prueba la primera aserción del teorema; la segunda aserción resulta ahora obvia de las definiciones.

Hay un modo constructivo de evitar esta cruda inducción y puesto que arroja aún más luz sobre las nociones implicadas, lo reproducimos aquí, como una prueba alternativa del teorema.

Sea $\mathfrak{X} = \{x_1, \ldots, x_n\}$ una base cualquiera de $\mathfrak{V}$. Construiremos un conjunto completo ortonormal $\mathfrak{Y} = \{y_1, \ldots, y_n\}$ con la propiedad de que cada y_j es una combinación lineal de $x_1, \ldots, x_j$. Para comenzar la construcción, observamos que $x_1 \neq 0$ (puesto que $\mathfrak{X}$ es linealmente independiente) y escribimos $y_1 = x_1/\|x_1\|$. Supóngase ahora que $y_1, \ldots, y_r$ han sido determinadas de modo que forman un conjunto ortonormal y así, que cada y_j $(j = 1, \ldots, r)$ es una combinación lineal de $x_1, \ldots, x_j$. Escribimos

$$z = x_{r+1} - (\alpha_1 y_1 + \cdots + \alpha_r y_r),$$

donde los valores de los escalares $\alpha_1, \ldots, \alpha_r$ están todavía por determinarse. Puesto que

$$(z, y_j) = (x_{r+1} - \textstyle\sum_i \alpha_i y_i, y_j) = (x_{r+1}, y_j) - \alpha_j$$

para $j = 1, \ldots, r$, se sigue que si escogemos $\alpha_j = (x_{r+1}, y_j)$, entonces $(z, y_j) = 0$ para $j = 1, \ldots, r$. Puesto que, además z es una combinación lineal de x_{r+1} y $y_1, \ldots, y_r$, es también una combinación lineal de x_{r+1} y $x_1, \ldots, x_r$. Finalmente, z es diferente de cero, puesto que $x_1, \ldots, x_r, x_{r+1}$ son linealmente independientes y el coeficiente de x_{r+1} en la expresión de z no es cero. Escribimos $y_{r+1} = z/\|z\|$; claramente $\{y_1, \ldots, y_r, y_{r+1}\}$ es de nuevo un conjunto ortogonal con todas las propiedades deseadas y el paso de inducción queda terminado. Haremos uso del hecho de que no solamente es cada y_j una combinación lineal de las x, con índices entre 1 y j, sino, viceversa, cada x_j es una combinación lineal de las y, con índices entre 1 y j. El método para convertir una base lineal en un conjunto ortonormal completo que acabamos de describir, es conocido como el *proceso de ortogonalización de Gram-Schmidt.*

Encontraremos conveniente y natural, en espacios de productos interiores, trabajar exclusivamente con bases que son también conjuntos ortonormales completos. Llamaremos a esa base una *base ortonormal* o un *sistema ortonormal de coordenadas;* en el futuro, cada vez que discutamos las bases que no son necesariamente ortonormales, pondremos este hecho de manifiesto, llamándolas bases lineales.

EJERCICIOS

1. Conviértase $\mathcal{P}_2$ en un espacio de producto interior, escribiendo $(x, y) = \int_0^1 x(t)\overline{y(t)}\,dt$ cada vez que x y y estén en $\mathcal{P}_2$, y encontrando un conjunto ortonormal completo en ese espacio.

2. Si x y y son vectores unitarios ortogonales (eso es, $\{x, y\}$ es un conjunto ortogonal), ¿cuál es la distancia entre x y y?

3. Pruébese que si $|(x, y)| = ||x|| \cdot ||y||$ (esto es, si la desigualdad de Schwarz se reduce a una igualdad) entonces x y y son linealmente dependientes.

4. (a) Pruébese que la desigualdad de Schwarz sigue siendo verdadera si, en la definición de un producto interior "estrictamente positivo" es reemplazado por "no negativo".

(b) Pruébese que para un producto interior "no negativo" del tipo mencionado en (a), el conjunto de todos los vectores x para los cuales $(x, x) = 0$ es un subespacio.

(c) Fórmese el espacio cociente, módulo el subespacio mencionado en (b) y demuéstrese que el "producto interior" dado induce sobre ese espacio cociente, de una manera natural, un producto interior honesto (estrictamente positivo).

(d) ¿Se extienden las consideraciones de (a), (b) y (c) a espacios normados (posiblemente sin producto interior)?

5. (a) Dado un número estrictamente positivo, α, trátese de definir una norma que $\mathcal{R}^2$ escribiendo

$$\|x\| = (|\xi_1|^\alpha + |\xi_2|^\alpha)^{1/\alpha}$$

siempre que $x = (\xi_1, \xi_2)$. ¿Bajo qué condiciones impuestas a α define esta ecuación una norma?

(b) Pruébese que la ecuación

$$\|x\| = \text{máx}\,\{|\xi_1|, |\xi_2|\}$$

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

(c) ¿A cuál de las normas definidas en (a) y (b) corresponde un producto interior en $\mathcal{R}^2$ tal que $||x||^2 = (x, x)$ para toda x de $\mathcal{R}^2$?

6. (a) Pruébese que una condición necesaria y suficiente impuesta a un espacio real normado para que exista un producto interior que satisfaga la ecuación $||x||^2 = (x, x)$ para toda x es que

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

para toda x y y.

(b Discútase la correspondiente aserción para espacios complejos.

(c) Pruébese que una condición necesaria y suficiente impuesta a una norma en $\mathcal{R}^2$ para que exista un producto interior que satisfaga la ecuación $\|x\|^2 = (x, x)$ para toda x de $\mathcal{R}^2$ es que el lugar geométrico de la ecuación $\|x\| = 1$ sea una elipse.

7. Si $\{x_1, \ldots, x_n\}$ es un conjunto completo ortonormal, en un espacio de producto interior, y si $y_j = \sum_{i=1}^{j} x_i, j = 1, \cdots, n$, expresa en términos de las x los vectores obtenidos aplicando el proceso de ortogonalización de Gram-Schmidt a las y.

§66. TEOREMA DE LAS PROYECCIONES

Puesto que un subespacio de producto interior puede él mismo considerarse como espacio de producto interior, puede aplicarse el teorema de la sección precedente. El siguiente resultado, llamado el *teorema de las proyecciones*, es el de más importante aplicación.

TEOREMA. *Si $\mathfrak{M}$ es un subespacio cualquiera de un espacio finito dimensional $\mathfrak{V}$, de producto interior, entonces $\mathfrak{V}$ es la suma directa de $\mathfrak{M}$ y $\mathfrak{M}^\perp$, y $\mathfrak{M}^{\perp\perp} = \mathfrak{M}$.*

PRUEBA. Sea $\mathfrak{X} = \{x_1, \ldots, x_m\}$ un conjunto ortogonal que es completo en $\mathfrak{M}$, y sea z un vector cualquiera de $\mathfrak{V}$. Escribimos $x = \sum_i \alpha_i x_i$, donde $\alpha_i = (z, x_i)$; se sigue del §63, Teorema 1, que $y = z - x$ está en $\mathfrak{M}^\perp$, de manera que z es la suma de dos vectores, $z = x + y$, con x en $\mathfrak{M}$ y y en $\mathfrak{M}^\perp$. Que $\mathfrak{M}$ y $\mathfrak{M}^\perp$ están disjuntos está claro; si x perteneciera a ambos, entonces tendríamos $\|x\|^2 = (x, x) = 0$. Se sigue del teorema del §18 que $\mathfrak{V} = \mathfrak{M} \oplus \mathfrak{M}^\perp$.

Observamos que en la descomposición $z = x + y$, tenemos

$$(z, x) = (x + y, x) = \|x\|^2 + (y, x) = \|x\|^2,$$

y, de modo semejante,

$$(z, y) = \|y\|^2.$$

De aquí que, si z está en $\mathfrak{M}^{\perp\perp}$, de manera que $(z, y) = 0$, entonces $\|y\|^2 = 0$, de manera que $z\ (= x)$ está en $\mathfrak{M}$; en otras palabras, $\mathfrak{M}^{\perp\perp}$ está contenida en $\mathfrak{M}$. Puesto que ya sabemos que $\mathfrak{M}$ está contenida en $\mathfrak{M}^{\perp\perp}$, la prueba del teorema es completa.

Esta clase de descomposición de suma directa de un espacio de producto interior (vía un subespacio y su complemento ortogonal), es de considerable interés geométrico. Estudiaremos un poco después las proyecciones asociadas; resultan ser una interesante e importante subclase de la clase de todas las proyecciones. Por el momento sólo hacemos una observación sobre la conexión con el teorema de Pitágoras; puesto que $(z, x) = \| x \|^2$ y $(z, y) = \| y \|^2$, tenemos

$$\| z \|^2 = (z, z) = (z, x) + (z, y) = \| x \|^2 + \| y \|^2.$$

En otras palabras, el cuadrado de la hipotenusa es igual a la suma de los cuadrados de los lados. Más generalmente, si $\mathfrak{M}_1, \cdots, \mathfrak{M}_k$ son subespacios ortogonales por pares, en un espacio de producto interior $\mathfrak{V}$, y si $x = x_1 + \ldots + x_k$ con x_j en $\mathfrak{M}_j$ para $j = 1, \ldots, k$, entonces

$$\| x \|^2 = \| x_1 \|^2 + \cdots + \| x_k \|^2.$$

§67. FUNCIONALES LINEALES

Estamos ahora en condiciones de estudiar funcionales lineales sobre espacios de productos. Para un espacio vectorial n-dimensional el espacio dual es también n-dimensional y es, en consecuencia, isomórfico al espacio original. Sin embargo, no hay un isomorfismo natural obvio que podamos establecer; tenemos que esperar que el segundo espacio dual regrese allí de donde partimos. El punto principal del teorema que probaremos ahora es que en espacios de producto interior hay una correspondencia "natural" entre $\mathfrak{V}$ y $\mathfrak{V}'$; la única nube en el horizonte es que, en general, no es completamente un isomorfismo.

TEOREMA. *A cualquier funcional lineal y' sobre un espacio finito dimensional de producto interior $\mathfrak{V}$ corresponde un vector y único en $\mathfrak{V}$ tal que $y'(x) = (x, y)$ para toda x.*

PRUEBA. Si $y' = 0$, podemos escoger $y = 0$; supongamos desde ahora que $y'(x)$ no es idénticamente cero. Sea $\mathfrak{M}$ el subespacio consistente en todos los vectores x para los cuales $y'(x) = 0$, y sea $\mathfrak{N} = \mathfrak{M}^\perp$ el completamente ortogonal de $\mathfrak{M}$. El subespacio $\mathfrak{N}$ contiene un vector no cero, y_0; multiplicando por una constante adecuada, podemos suponer que $\| y_0 \| = 1$. Escribimos $y = \overline{y'(y_0)} \cdot y_0$. (La barra denota conjugación compleja, como de costumbre; en caso de que $\mathfrak{V}$ sea un espacio real de producto interior y no un espacio unitario, la barra puede ser omitida). Tenemos entonces la relación deseada,

$$y'(x) = (x, y) \tag{1}$$

cuando menos para $x = y_0$ y para toda x de $\mathfrak{M}$. Para una x arbitraria de $\mathcal{V}$, escribimos $x_0 = x - \lambda y_0$, donde

$$\lambda = \frac{y'(x)}{y'(y_0)};$$

entonces $y'(x_0) = 0$ y $x = x_0 + \lambda y_0$ es una combinación lineal de dos vectores, para cada uno de los cuales (1) es válida. De la linearidad de ambos miembros de (1) se sigue que (1) es válida para x, como se tenía que probar.

Para probar la univocidad, supóngase que $(x, y_1) = (x, y_2)$ para toda x. Se sigue que $(x, y_1 - y_2) = 0$ para toda x y, en consecuencia, en particular para $x = y_1 - y_2$, de manera que $\| y_1 - y_2 \|^2 = 0$ y $y_1 = y_2$.

La correspondencia $y' \rightleftarrows y$ es una correspondencia uno a uno entre $\mathcal{V}$ y $\mathcal{V}'$, con la propiedad de que $y'_1 + y'_2$ corresponde $y_1 + y_2$, y a $\alpha y'$ corresponde $\bar{\alpha} y$; por esta razón nos referimos a ella como un *isomorfismo conjugado*. A pesar del hecho de que este isomorfismo conjugado hace a $\mathcal{V}'$ prácticamente indistinguible de $\mathcal{V}$, es prudente conservar a ambos conceptualmente separados. Una razón para esto es que quisiéramos que $\mathcal{V}'$ fuera un espacio de producto interior juntamente con $\mathcal{V}$; si, en cambio, seguimos la pista dada por el isomorfismo conjugado entre $\mathcal{V}$ y $\mathcal{V}'$, la conjugación causa de nuevo dificultades. Sean y'_1 y y'_2 dos elementos cualesquiera de $\mathcal{V}'$; si $y'_1(x) = (x, y_1)$ y $y'_2(x) = (x, y_2)$, es grande la tentación de escribir

$$(y'_1, y'_2) = (y_1, y_2).$$

Un momento de reflexión mostrará que esta expresión puede no satisfacer el §61, (2) y no es consecuencia un adecuado producto interior. La dificultad se origina solamente en espacios complejos (esto es, unitarios); tenemos, por ejemplo,

$$(\alpha y'_1, y'_2) = (\bar{\alpha} y_1, y_2) = \bar{\alpha}(y_1, y_2) = \bar{\alpha}(y'_1, y'_2).$$

El remedio es claro; escribimos

$$(2) \qquad (y'_1, y'_2) = \overline{(y_1, y_2)} = (y_2, y_1);$$

dejamos al lector la comprobación de que con esta definición $\mathcal{V}$ se convierte en todos los casos en un espacio de producto interior. Denotaremos este espacio de producto interior con $\mathcal{V}^*$.

Hacemos la observación de que nuestras dificultades (si así pueden llamarse) con la conjugación compleja han sido ahora más notacionales que conceptuales; sigue siendo verdadero que la única diferencia entre la teoría de los espacios euclidianos y la teoría de los espacios unitarios es que en los últimos aparece ocasionalmente una barra. Se suscitarán diferencias más profundas entre las dos teorías cuando pasemos a estudiar las transformaciones lineales.

§68. PARENTESIS CONTRA CORCHETES

Se hace necesario ahora aclarar la relación entre espacios vectoriales generales y espacios de productos interiores. El teorema de la sección precedente muestra que, mientras procedamos con cuidado con respecto de la conjugación compleja, (x, y) puede tomar por completo el lugar de $[x, y]$. Podría parecer q. habría sido deseable desarrollar todo el tema de espacios vectoriales generales en forma tal que el concepto de ortogonalidad en un espacio se vuelva no solamente un análogo, sino un caso especial de alguna relación general previamente estudiada entre vectores y funcionales. Una forma, por ejemplo, de evitar el hecho desagradable de la conjugación (o, más bien, de cambiarla a una posición menos conspicua), habría sido definir el espacio dual de un espacio vectorial complejo como el conjunto de funcionales lineales conjugadas, esto es, el conjunto de funciones y numéricamente valuadas, para las cuales

$$y(\alpha_1 x_1 + \alpha_2 x_2) = \bar{\alpha}_1 y(x_1) + \bar{\alpha}_2 y(x_2).$$

A causa de que parecía sin objeto (y contrario al uso común) introducir esta complicación en la teoría general, escogimos el camino indirecto que hemos seguido. Puesto que de ahora en adelante trataremos sólo con espacios de productos interiores, pedimos al lector que revise mentalmente todo el trabajo anterior, reemplazando, en todas partes, el corchete $[x, y]$ por el paréntesis (x, y). Examinemos el efecto de este cambio sobre los teoremas y definiciones de los primeros dos capítulos.

El reemplazo de $\mathcal{V}'$ por $\mathcal{V}^*$ es meramente un cambio de notación; el nuevo símbolo se supone que nos recuerda que algo nuevo (a saber, un producto interior) ha sido añadido a $\mathcal{V}'$. De un poco más interés es el isomorfismo (conjugado) entre $\mathcal{V}$ y $\mathcal{V}^*$; por medio de él, los teoremas del §15, que afirman la existencia de funcionales lineales con varias propiedades, pueden ser ahora interpretados como afirmando la existencia de ciertos vectores en $\mathcal{V}$ mismo. Así, por ejemplo, la existencia de una base dual a cualquier base dada $\mathfrak{X} = \{x_1, \ldots, x_n\}$ implica ahora la existencia de una base $\mathcal{Y} = \{y_1, \ldots, y_n\}$ (de $\mathcal{V}$), con la propiedad de que $(x_i, y_j) = \delta_{ij}$.

Más excitante aún es el reemplazo implicado del aniquidador $\mathfrak{M}^0$ de un subespacio $\mathfrak{M}$ ($\mathfrak{M}^0$ queda en $\mathcal{V}'$ o $\mathcal{V}^*$) por el complemento ortogonal $\mathfrak{M}^\perp$ (que queda, juntamente con $\mathfrak{M}$, en $\mathcal{V}$). El nuevo desarrollo más radical, sin embargo, se refiere al adjunto de una transformación lineal. Así, podemos escribir el análogo del § 44 (1) y correspondiendo a cada transformación lineal A sobre $\mathcal{V}$ podemos definir una transformación lineal A^* escribiendo

$$(Ax, y) = (x, A^*y)$$

para cada x. Se sigue de esta definición que A^* es de nuevo una transformación lineal definida sobre el mismo espacio vectorial $\mathcal{V}$, pero, a causa de la simetría hermitiana de (x, y), la relación entre A y A^* no es completamente la misma que la relación entre A y A'. La diferencia más notable es que (en un espacio unitario) $(\alpha A)^* = \bar{\alpha}A^*$ (y *no* $(\alpha A)^* = \alpha A^*$). Asociado con este fenómeno está el hecho de que si la matriz de A, con respecto de alguna base fija, es (α_{ij}) entonces la matriz de A^*, con respecto de la base dual, no es (α_{ji}), sino $(\bar{\alpha}_{ji})$. Para determinantes, no tenemos det A^* = det A, sino det $A^* = \overline{\det A}$, y, consecuentemente, los valores propios de A^* no son los mismos que los de A, sino más bien sus conjugados. Aquí, sin embargo, se detiene la diferencia. Todos los demás resultados del §44 sobre la naturaleza antisomórfica de la correspondencia $A \leftrightarrows A^*$ son válidos; la identidad $A = A^{**}$ es estrictamente verdadera y no necesita la ayuda de un isomorfismo para interpretarla.

De inmediato discutiremos transformaciones lineales sobre espacios de productos interiores y veremos que la principal característica nueva que diferencia su estudio de la discusión del Cap. II es la posibilidad de comparar A y A^* como transformaciones lineales sobre el mismo espacio y de investigar las clases de transformaciones lineales que tienen una relación particularmente sencilla con sus adjuntos.

§69. ISOMORFISMOS NATURALES

Hay ahora sólo una duda posible más que el lector podría (o, en todo caso, debería) tener. Muchos de nuestros resultados precedentes fueron consecuencia de relaciones de reflexividad como $A^{**} = A$; ¿permanecen éstas válidas después de la revolución de corchetes a paréntesis? Más al grano es el siguiente modo de presentar la cuestión. Todo lo que decimos de un espacio unitario $\mathcal{V}$ debe ser también verdad de un espacio unitario $\mathcal{V}^*$; en particular está también en una relación isomórfica conjugada natural con su espacio dual $\mathcal{V}^{**}$. Si ahora a cada vector de $\mathcal{V}$ hacemos corresponder un vector de $\mathcal{V}^{**}$, aplicando primero el isomorfismo conjugado general de $\mathcal{V}$ a $\mathcal{V}^*$ y a continuación procediendo en la misma forma de $\mathcal{V}^*$ a $\mathcal{V}^{**}$; entonces este mapa es un rival para el título de mapa natural $\mathcal{V}$ a $\mathcal{V}^{**}$, título ya concedido en el Cap. 1 a una correspondencia aparentemente diferente. ¿Cuál es la relación entre las dos correspondencias naturales? Nuestros enunciados sobre las coincidencias, excepto en cuanto a modificaciones triviales, de las teorías de los paréntesis y los corchetes, se justifican realmente por el hecho, que probaremos ahora, de que dos mapas son iguales. (No debe sorprender puesto que $\bar{\bar{\alpha}} = \alpha$. que después de dos aplicaciones desaparezca la molesta conjugación). La prueba es más corta que la introducción a la misma.

Sea y_0 cualquier elemento de $\mathcal{V}$; al mismo corresponde la funcional lineal y_0^* de $\mathcal{V}^*$, definida por $y_0^*(x) = (x, y_0)$ y a y_0^*, a su vez corresponde la funcional lineal y_0^{**} de $\mathcal{V}^{**}$, definida por $y_0^{**}(y^*) = (y^*, y_0^*)$. Ambas correspondencias se dan por el mapa introducido en este capítulo. Anteriormente (véase §16), el correspondiente y_0^{**} en $\mathcal{V}^{**}$ de y_0 en $\mathcal{V}$ fue definido como $y_0^{**}(y^*) = y^*(y_0)$ para toda y^* en $\mathcal{V}^*$; debemos demostrar que y_0^{**}, como lo hemos definido aquí, satisface esta identidad. Sea y^* cualquier funcional lineal sobre $\mathcal{V}$ (esto es, cualquier elemento de $\mathcal{V}^*$); tenemos

$$y_0^{**}(y^*) = (y^*, y_0^*) = (y_0, y) = y^*(y_0).$$

(la igualdad de en medio viene de la definición de producto interior en $\mathcal{V}^*$). Esto resuelve todos nuestros problemas.

EJERCICIOS

1. Si $\mathfrak{M}$ y $\mathfrak{N}$ son subespacios de un espacio finito-dimensional de producto interior, entonces

$$(\mathfrak{M} + \mathfrak{N})^\perp = \mathfrak{M}^\perp \cap \mathfrak{N}^\perp$$

y

$$(\mathfrak{M} \cap \mathfrak{N})^\perp = \mathfrak{M}^\perp + \mathfrak{N}^\perp.$$

2. Si $y'(x) = \frac{1}{3}(\xi_1 + \xi_2 + \xi_3)$ para cada $x = (\xi_1, \xi_2, \xi_3)$ de $\mathbb{C}^3$, encontrar un vector y de $\mathbb{C}^3$ tal que $y'(x) = (x, y)$.

3. Si y es un vector de un espacio de producto interior, si A es una transformación lineal sobre ese espacio, y si $f(x) = \overline{(y, Ax)}$ para todo vector x,

entonces f es una funcional lineal; encuéntrese un vector y^* tal que $f(x) = (x, y^*)$ para toda x.

4. (a) Si A es una transformación lineal sobre un espacio finito dimensional de producto interior, entonces $tr(A^*A) \geq 0$; una condición necesaria y suficiente para que tr $(A^*A) = 0$ es que $A = 0$. (Sugestión: Véanse las matrices). Esta propiedad de las trazas puede usarse a menudo para obtener hechos algebraicos que de otro modo serían elusivos, sobre productos de transformaciones y sus adjuntos.

(b) Pruébese por un argumento de traza, y también directamente, que si $A_1, \ldots, A_k$ son transformaciones lineales sobre un espacio finito-dimensional de producto interior, y si $\sum_{j=1}^{k} A_j^* A_j = 0$, entonces $A_1 = \ldots = A_k = 0$.

(c) Si $A^*A = B^*B - BB^*$, entonces $A = 0$.

(d) Si A^* conmuta con A y si A conmuta con B, entonces A^* conmuta con B. (Sugestión: si $C = A^*B - BA^*$ y $D = AB - BA$, entonces tr $(C^*C) =$ tr $(D^*D) +$ tr$[(A^*A - AA^*)(B^*B - BB^*)]$.

5. (a) Supóngase que $\mathcal{H}$ es un espacio unitario, y forma el conjunto de todos los pares ordenados $\langle x, y\rangle$ con x y y en $\mathcal{H}$ (esto es, la suma directa de $\mathcal{H}$ consigo mismo). Pruébese que la ecuación

$$(\langle x_1, y_1\rangle, \langle x_2, y_2\rangle) = (x_1, x_2) + (y_1, y_2)$$

define un producto interior en la suma directa $\mathcal{H} \oplus \mathcal{H}$.

(b) Si U es definida por $U\langle x, y\rangle = \langle y, -x\rangle$, entonces $U^*U = 1$.

(c) La *gráfica* de una transformación lineal A sobre $\mathcal{H}$ es el conjunto de todos los elementos $\langle x, y\rangle$ de $\mathcal{H} \oplus \mathcal{H}$ para los cuales $y = Ax$. Pruébese que la gráfica de toda transformación lineal sobre $\mathcal{H}$ es un subespacio de $\mathcal{H} \oplus \mathcal{H}$

(d) Si A es una transformación lineal sobre $\mathcal{H}$ con gráfica $\mathcal{G}$, entonces la gráfica de A^* es el complemento ortogonal (en $\mathcal{H} \oplus \mathcal{H}$) de la imagen bajo U [véase (b)] de la gráfica de A.

6. (a) Si para cada transformación lineal A sobre un espacio finito-dimensional de producto interior $N(A) = \sqrt{\operatorname{tr}(A^*A)}$, entonces N es una norma (sobre el espacio de todas las transformaciones lineales).

(b) ¿Es la norma N inducida por un producto interior?

7. (a) Dos transformaciones lineales A y B sobre un espacio de producto interior se llaman *congruentes* si existe una transformación lineal invertible P tal que $B = P^*AP$. (El concepto se define frecuentemente para las "formas cuadráticas" asociadas con transformaciones lineales y no para las transformaciones lineales mismas; esto es en gran parte cuestión de gusto. Nótese que si $\alpha(x) = (Ax, x)$ y $\beta(x) = (Bx, x)$, entonces $B = P^*AP$ implica que $\beta(x) = \alpha(Px)$. Pruébese que la congruencia es una relación de equivalencia.

(b) Si A y B son congruentes, entonces lo son también A^* y B^*.

(c) ¿Existe una transformación lineal A tal que A sea congruente a un escalar α, pero $A \neq \alpha$?

(d) ¿Existen transformaciones lineales A y B tales que A y B sean congruentes, pero A^2 y B^2 no lo sean?

(e) Si dos transformaciones invertibles son congruentes, entonces son inversas.

§70. TRANSFORMACIONES AUTO-ADJUNTAS

Estudiemos ahora la estructura algebraica de la clase de todas las transformaciones lineales sobre un espacio de producto interior. En muchos respectos fundamentales esta clase se parece a la clase de todos los números complejos. En ambos sistemas las nociones de adición, multiplicación, 0 y 1 se definen y tienen propiedades similares y en ambos sistemas hay un involuntario anti-automorfismos del sistema sobre sí mismo (a saber $A \to A^*$ y $\zeta \to \bar{\zeta}$) Usaremos esta analogía como principio heurístico e intentaremos llevar a transformaciones lineales algunos bien conocidos conceptos del dominio complejo. Estaremos estorbados en este trabajo por dos dificultades en la teoría de las transformaciones lineales, de las cuales, lo que es posiblemente sorprendente, la segunda es mucho más seria; son la imposibilidad de división irrestricta y la no conmutatividad de transformaciones lineales generales.

Los tres más importantes subconjuntos del plano de números complejos son el conjunto de los números reales, el conjunto de los números reales positivos y el conjunto de números de valor absoluto uno. Procederemos ahora sistemáticamente a usar nuestra analogía heurística de transformaciones con números complejos y a tratar de descubrir los análogos entre transformaciones de estos bien conocidos conceptos numéricos.

¿Cuándo es real un número complejo? Claramente, una condición necesaria y suficiente para la realidad de ζ es la validez de la ecuación $\zeta = \bar{\zeta}$. Podríamos en consecuencia (recordando que el análogo del conjugado complejo para transformaciones lineales es el adjunto), definir una transformación lineal A como real si $A = A^*$. Más comúnmente, las transformaciones lineales A, para las cuales $A = A^*$ son llamadas *auto-adjuntas;* en espacios reales de productos interiores la palabra usual es *simétrico* y en espacios complejos de productos interiores, *hermitiano.* Veremos que las transformaciones auto-adjuntas juegan en realidad el mismo papel que los números reales.

Es completamente fácil caracterizar la matriz de una transformación auto-adjunta con respecto de una base ortonormal $\mathfrak{X} = \{x_1, \ldots, x_n\}$. Si la matriz de A es (α_{ij}) entonces sabemos que la matriz de A^* con respecto de la base dual de $\mathfrak{X}$ es $(\alpha_{ij}{}^*)$, donde $\alpha_{ij}{}^* = \overline{\alpha_{ji}}$ puesto que una base ortonormal es autodual y puesto que $A = A^*$, tenemos

$$\alpha_{ij} = \overline{\alpha_{ji}}.$$

Dejamos al lector la verificación de la recíproca; si definimos una transformación lineal A por medio de una matriz (α_{ij}) y un sistema arbitrario de coordenadas ortonormales $\mathfrak{X} = \{x_1, \ldots, x_n\}$ vía las ecuaciones usuales

$$A(\sum_j \xi_j x_j) = \sum_i \eta_i x_i,$$
$$\eta_i = \sum_j \alpha_{ij}\xi_j,$$

y si la matriz (α_{ij}) es tal que $\alpha_{ij} = \overline{\alpha_{ji}}$, entonces A es auto-adjunta.

Las reglas algebraicas para la manipulación de transformaciones auto-adjuntas son fáciles de recordar si pensamos en esas transformaciones como análogos de números reales. Así, si A y B son auto-adjuntas, también lo son $A + B$; si A es auto-adjunta y diferentes de 0, y si α es un escalar no cero, entonces una condición necesaria y suficiente para que αA sea auto-adjunta es que α sea real; y si A es invertible, entonces ambas o ninguna de A y A^{-1} son auto-adjuntas. El lugar donde hay siempre algo mal es en la multiplicación; el producto de dos transformaciones auto-adjuntas no necesita ser auto-adjunto. Los hechos positivos sobre productos son dados por los siguientes dos teoremas.

TEOREMA 1. *Si A y B son auto-adjuntos, entonces una condición necesaria y suficiente para que AB (o BA) sean auto-adjuntos es que $AB = BA$ (esto es, que A y B conmuten).*

PRUEBA. Si $AB = BA$, entonces $(AB)^* = B^*A^* = BA = AB$. Si $(AB)^* = AB$, entonces $AB = (AB)^* = B^*A^* = BA$.

TEOREMA 2. *Si A es auto-adjunto entonces B^*AB es auto-adjunto para toda B; si B es invertible y B^*AB es auto-adjunto, entonces A es auto-adjunto.*

PRUEBA. Si $A = A^*$, entonces $(B^*AB)^* = B^*A^*B^{**} = B^*AB$. Si B es invertible y $B^*AB = (B^*AB)^* = B^*A^*B$, entonces (multiplíquese por B^{*-1} sobre la izquierda y B^{-1} sobre la derecha) $A = A^*$.

Un número complejo ζ es puramente imaginario si y sólo si $\bar{\zeta} = -\zeta$. El concepto correspondiente para la transformación lineal se identifica por la palabra *oblicuo*, si una transformación lineal A sobre un producto interior es tal que $A^* = -A$, entonces A se llama *simétrico oblicuo* o *hermitiano oblicuo* según el espacio sea real o complejo. Aquí hay alguna evidencia de la naturaleza total de nuestra analogía entre números complejos y transformaciones lineales: una transformación lineal arbitraria puede ser expresada, de un modo y sólo de uno, en la forma $A = B + C$, donde B es un auto-adjunto y C es oblicuo. (La representación de A en esta forma es, algunas veces, llamada la *descomposición cartesiana* de A.) En realidad, si escribimos

$$B = \frac{A + A^*}{2}, \tag{1}$$

$$C = \frac{A - A^*}{2}, \tag{2}$$

entonces tenemos $B^* = \dfrac{A^* + A}{2} = B$ y $C^* = \dfrac{A^* - A}{2} = -C$ y, por supuesto, $A = B + C$. De esta prueba de la existencia de la descomposición cartesiana, es clara su univocidad; si tenemos $A = B + C$, entonces $A^* = B - C$ y, consecuentemente, A, B y C están de nuevo conectadas por (1) y (2).

En el caso complejo hay una forma simple de obtener transformaciones hermitianas oblicuas de hermitianas, y viceversa: multiplíquese sólo por $i(= \sqrt{-1})$. Se sigue que, en el caso complejo, toda transformación lineal A tiene una representación única en la forma $A = B + iC$, donde B y C son hermitianas. Nos referiremos a B y C como las partes real e imaginaria de A.

EJERCICIOS

Dé un ejemplo de dos transformaciones auto-adjuntas cuyo producto no sea auto-adjunto.

2. Considérese el espacio $\mathcal{P}_n$. con el producto interior dado por $(x, y) = \int_0^1 x(t)\overline{y(t)}\,dt$.

(a) ¿Es auto-adjunto el operador de multiplicación T (definido por $(Tx)(t) = tx(t)$?

(b) ¿Es auto-adjunto el operador de diferenciación D?

3. (a) Pruébese que la ecuación $(x, y) = \sum_{j=0}^{n} x\left(\frac{j}{n}\right) \overline{y\left(\frac{j}{n}\right)}$ define un producto interior en el espacio $\mathcal{P}_n$.

(b) ¿Es el operador de multiplicación T [definido por $(Tx)(t) = tx(t)$] auto-adjunto [con respecto del producto interior definido en (a)]?

(c) ¿Es auto-adjunto el operador de diferenciación D?

4. Si A y B son transformaciones lineales tales que A y AB sean auto-adjuntos y tales que $\mathfrak{N}(A) \subset \mathfrak{N}(B)$ entonces existe una transformación auto-adjunta $CA = B$.

5. Si A y B son congruentes y A es oblicua, ¿se sigue que B es oblicua?

6. Si A es oblicua, ¿se sigue que también lo es A^2? ¿Qué hay sobre A^3?

7. Si tanto A como B son auto-adjuntos, o si ambos son oblicuos, entonces $AB + BA$ es auto-adjunto y $AB - BA$ es oblicuo. ¿Qué sucede si uno de A y B es autoadjunto y el otro es oblicuo?

8. Si A es una transformación simétrico oblicua sobre un espacio euclidiano, entonces $(Ax, x) = 0$ para cada vector x. ¿Recíproca?

9. Si A es autoadjunta, u oblicua, y si $A^2x = 0$, entonces $Ax = 0$.

10. (a) Si es una transformación simétrico oblicua sobre un espacio euclidiano de dimensión impar, entonces $\det A = 0$.

(b) Si es A una transformación simétrico oblicua sobre un espacio euclidiano finito-dimensional, entonces $\rho(A)$ es par.

§71. POLARIZACION

Antes de continuar con el programa de estudiar las analogías entre números complejos y transformaciones lineales, nos damos tiempo para escoger algunos importantes resultados auxiliares sobre espacios de productos interiores.

TEOREMA 1. *Una condición necesaria y suficiente para que una transformación lineal A sobre un espacio de producto interior A sea 0, es que $(Ax, y) = 0$ para toda x y y.*

PRUEBA. La necesidad de la condición es obvia; la suficiencia se sigue de hacer y igual a Ax.

TEOREMA 2. *Una condición necesaria y suficiente para que una transformación lineal auto-adjunta A sobre un espacio de producto interior A sea 0 es que $(Ax, x) = 0$ para toda x.*

PRUEBA. La necesidad es obvia. La prueba de suficiencia comienza por verificar la identidad

(1) $(Ax, y) + (Ay, x) = (A(x + y), (x + y)) - (Ax, x) - (Ay, y)$.

(Desarróllese el primer término del segundo miembro). Puesto que A es auto-adjunto, el primer miembro de la ecuación es igual a 2 Re (Ax, y). La condición supuesta implica que desaparece el segundo miembro y de aquí que Re $(Ax, y) = 0$. En este punto es necesario dividir la prueba en dos casos. Si el espacio de producto interior es real (esto es, A es simétrica), entonces (Ax, y) es real y, en consecuencia, $(Ax, y) = 0$. Si el espacio de producto interior es complejo (esto es, A es hermitiana), entonces encontramos un número complejo θ tal que $|\theta| = 1$ y $\theta(Ax, y) = |(Ax, y)|$. (Aquí x y y son temporalmente fijos). El resultado que ya tenemos, aplicado a θx en lugar de x, da $0 = \text{Re}\,(A(\theta x), y) = \text{Re}\,(\theta(Ax, y) = \text{Re}\,|(Ax, y)| = |(Ax, y)|$. En todo caso, en consecuencia, $(Ax, y) = 0$ para toda x y y, y el resultado deseado se sigue del Teorema 1.

Es útil preguntar qué tan importante es la auto-adjunción de A en el Teorema 2; la respuesta es que en el caso complejo no es importante en lo más mínimo.

TEOREMA 3. *Una condición necesaria y suficiente para una transformación lineal A sobre un espacio unitario sea 0, es que $(Ax, x) = 0$ para toda x.*

PRUEBA. Como antes, la necesidad es obvia. Para la prueba de suficiencia, usamos la llamada identidad de *polarización*:

(2) $\alpha\bar{\beta}(Ax, y) + \bar{\alpha}\beta(Ay, x)$

$$= (A(\alpha x + \beta y), (\alpha x + \beta y)) - |\alpha|^2(Ax, x) - |\beta|^2(Ay, y).$$

(Justamente como para (1), la prueba consiste en desarrollar el primer término del segundo miembro). Si (Ax, x) es idéntico a cero, entonces obtenemos, escogiendo primero $\alpha = \beta = 1$, y a continuación $\alpha = i\ (= \sqrt{-1})$, $\beta = 1$.

$$(Ax, y) + (Ay, x) = 0$$

$$i(Ax, y) - i(Ay, x) = 0.$$

Dividiendo la segunda de estas ecuaciones por i y formando a continuación su medio aritmético, vemos que $(Ax, y) = 0$ para toda x y y, de manera que, por el Teorema 1, $A = 0$.

Este proceso de polarización es a menudo usado para obtener información sobre la "forma bilineal" (Ax, y) cuando solamente se supone el conocimiento de la "forma cuadrática" (Ax, x).

Es importante observar que, a pesar de esta aparente inocencia, el Teorema 3 hace un uso muy esencial del sistema de números complejos, éste y muchas de sus consecuencias dejan de ser verdaderas para espacios reales de productos interiores. La prueba, por supuesto, falla en nuestra selección de $\alpha = \sqrt{-1}$. Como ejemplo considérese una rotación de 90° en el plano; claramente tiene la propiedad de que convierte a cada vector x en un vector ortogonal a x.

Hemos visto que las transformaciones hermitianas juegan el mismo papel que los números reales; el siguiente teorema indica que están ligadas con el concepto de realidad en formas más profundas que a través de la analogía formal que sugirió su definición.

TEOREMA 4. *Una condición necesaria y suficiente para que una transformación lineal A sobre un espacio unitario sea hermitiana es que (Ax, x) sea real para todas las x.*

PRUEBA. Si $A = A^*$, entonces

$$(Ax, x) = (x, A^*x) = (x, Ax) = \overline{(Ax, x)},$$

de manera que (Ax, x) es igual a su propio conjugado y, en consecuencia, real. Si, recíprocamente, (Ax, x) es siempre real, entonces

$$(Ax, x) = \overline{(Ax, x)} = (x, A^*x) = (A^*x, x),$$

de manera que $([A - A^*]x, x) = 0$, para toda x y, por el Teorema 3, $A = A^*$.

El Teorema 4 es falso para productos reales de espacios interiores. Es o era de esperarse, porque, en primer lugar, su prueba depende de un teorema que es verdadero solamente para espacios unitarios y, en segundo lugar, en un espacio real la realidad de (Ax, x) es automática, mientras que la identidad $(Ax, y) = (x, Ay)$ no se satisface necesariamente.

§72. TRANSFORMACIONES POSITIVAS

¿Cuándo es un número complejo ζ positivo (esto es, $\geqq 0$)? Dos condiciones igualmente necesarias y suficientes son que ζ pueda ser escrito en la forma $\zeta = \xi^2$ con algún ξ real, o que ζ pueda ser escrito en la forma $\xi = \sigma\bar{\sigma}$ con alguna σ (en general compleja). Recordando también el hecho de que (cuando menos para espacios unitarios) el carácter hermitiano de una transformación A puede ser descrito en términos de productos interiores (Ax, x) podemos considerar cualquiera de las tres condiciones siguientes e intentar usarla como definición de positividad para transformaciones

(1) $A = B^3$ para alguna B auto-adjunta,

(2) $A = C^*C$ para alguna C,

(3) A es auto-adjunta y $(Ax, x) \geqq 0$ para toda x.

Antes de decidir cuál de estas tres condiciones usar como definición, observamos que (1) $\Rightarrow$ (2) $\Rightarrow$ (3). En realidad: Si $A = B^2$, y $B = B^*$, entonces $A = BB = B^*B$ y Si $A = C^*C$, entonces $A^* = C^*C = A$ y $(Ax, x) = (C^*Cx, X) = Cx, Cx) = \| Cx \|^2 \geqq 0$. Es de hecho verdadero que (3) implica a (1), de manera que las tres condiciones son equivalentes, pero no podremos probar esto sino hasta más tarde. Adoptamos como nuestra definición la tercera condición.

DEFINICIÓN. Una transformación lineal A sobre un espacio de producto interior es *positiva*, en símbolos $A \geqq 0$, si es autoadjunta y si $(Ax, x) \geqq 0$ para toda x.

Más generalmente, escribiremos $A \geqq B$ (o $B \leqq A$) siempre que $A - B \geqq 0$. Aunque, por supuesto, es completamente posible que la diferencia de dos transformaciones que no son auto-adjuntas pares resulte positiva, generalmente escribiremos desigualdades solamente para transformaciones autoadjuntas. Obsérvese que para un espacio complejo de producto interior, una parte de la definición de positividad es superflua; si $(Ax, x) \geqq 0$ para toda x, entonces, en particular, (Ax, x) es real para toda x, y por el Teorema 4 de la sección precedente, A debe ser positiva.

Las transformaciones positivas son generalmente llamadas *semidefinidas no negativas*. Si $A \geqq 0$ y $(Ax, x) = 0$ implica que $x = 0$, diremos que A es *estrictamente* positiva; el término usual es *definida positiva*. Puesto que la desigualdad de Schwarz implica que

$$|(Ax, x)| \leqq \| Ax \| \cdot \| x \|,$$

vemos que si A es una transformación estrictamente positiva y si $Ax = 0$, entonces $x = 0$, de manera que, sobre un espacio finito-dimensional de producto interior, una transformación estrictamente positiva es invertible. Veremos posteriormente que la recíproca es verdadera. Si $A \leqq 0$ y A es invertible, entonces A es estrictamente positiva. Es algunas veces conveniente indicar el hecho de que una transformación A es estrictamente positiva escribiendo $A > 0$, si $A - B > 0$, podemos también escribir $A > B$ (o $B < A$).

Es posible dar una caracterización matricial de las transformaciones positivas; pospondremos esta discusión para después. Mientras tanto tendremos ocasión de referirnos a matrices positivas, significando con eso matrices simétricas hermitianas (α_{ij}) (esto es, $\alpha_{ij} = \overline{\alpha_{ji}}$) con la propiedad de que para cada secuencia $(\xi_1, \ldots, \xi_n)$ de n escalares tenemos $\sum_i \sum_j \alpha_{ij}\xi_i\bar{\xi}_j \geqq 0$. (En el caso real pueden omitirse las barras; en el caso complejo, la simetría hermitiana se sigue de la otra condición). Estas condiciones son claramente equivalentes a la condición de que (α_{ij}) es la matriz, con respecto de algún sistema de coordenadas ortonormales, de una transformación positiva.

Las reglas algebraicas para combinar transformaciones positivas son similares a las de las transformaciones auto-adjuntas, por lo que concierne a sumas, múltiples escalares e inversos; aun el §70, Teorema 2, sigue siendo válido si reemplazamos totalmente "auto-adjunto" por "positivo". Es también verdad que si A y B son positivos, es que $AB = BA$ (esto es, que A y B conmuten), pero tendremos que posponer algún tiempo la prueba de este enunciado.

EJERCICIOS

1. ¿Bajo qué condiciones impuestas a una transformación lineal A la función de dos variables cuyo valor en x y y es (Ax, y), satisface las condiciones impuestas a un producto interior?

2. ¿Cuáles de las siguientes matrices son positivas?

(a) $\begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}$.

(b) $\begin{pmatrix} 0 & i \\ -i & 0 \end{pmatrix}$.

(c) $\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$.

(d) $\begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}$.

(e) $\begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$.

3. ¿Para cuáles valores de α es positiva la matriz

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

4. (a) Si A es auto-adjunto, entonces tr A es real.
(b) Si $A \geqq 0$, entonces tr $A \geqq 0$.

5. (a) Dése un ejemplo de una matriz positiva, algunas de cuyas notaciones sean negativas.

(b) Dése un ejemplo de una matriz no positiva, todas cuyas notaciones sean positivas.

6. Una condición necesaria y suficiente para que una matriz dos por dos $\begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}$ (considerada como una transformación lineal sobre $\mathcal{C}^2$) sea positiva

es que sea simétrica hermitiana (esto es, que α y δ sean reales y $\gamma = \bar{\beta}$) y que $\alpha \geqq 0$, $\delta \geqq 0$ y $\alpha\delta - \beta\gamma \geqq 0$.

7. Asociada con cada secuencia $(x_1, \ldots, x_k)$ de k vectores en un espacio de producto interior, hay una matriz k por k (no una transformación lineal) llamada la *gramiana* de $(x_1, \ldots, x_k)$ y denotada por $G(x_1, \ldots, x_k)$; el elemento de la i-ésima fila y j-ésima columna de $G(x_1, \ldots, x_k)$ es el product interior (x_i, x_j). Pruébese que toda gramiana es una matriz positiva.

8. Si x y y son vectores no-cero (en un espacio finito dimensional de producto interior), entonces una condición necesaria y suficiente para que exista una transformación positiva A tal que $Ax = y$, es que $(x, y) > 0$.

9. (a) si las matrices $A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$ y $B = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$ se consideran como transformaciones lineales sobre $\mathbb{C}^2$, y si C es una matriz hermitiana (transformación lineal sobre $\mathbb{C}^2$) tal que $A \leqq C$ y $B \leqq C$, entonces

$$C = \begin{pmatrix} 1+\epsilon & \theta \\ \bar{\theta} & 1+\delta \end{pmatrix},$$

donde ϵ y δ son números reales positivos y $|\theta|^2 \leqq \min\{\epsilon(1+\delta), \delta(1+\epsilon)\}$.

(b) Si, además, $C \geqq 1$, entonces $\epsilon = \delta = \theta = 0$. En terminología moderna estos hechos juntos demuestran que las matrices hermitianas con el concepto inducido por la noción de positividad no forman una *celosía*. En el caso real, si la matriz $\begin{pmatrix} \alpha & \beta \\ \beta & \gamma \end{pmatrix}$ es interpretada como el punto (α, β, γ) del espacio tridimensional, la ordenación y su carácter de no celosía toma un divertido aspecto geométricos.

§73. ISOMETRIAS

Continuamos con nuestro programa de investigar la analogía entre números y transformaciones. ¿Cuándo un número complejo ζ tiene valor absoluto uno? Claramente una condición necesaria y suficiente es que $\bar{\zeta} = 1/\zeta$; guiados por nuestro principio heurístico nos vemos conducidos a considerar transformaciones lineales U para las cuales $U^* = U^{-1}$, o, equivalentemente, para las cuales $UU^* = U^*U = 1$. (Observamos que sobre un espacio vectorial finito-dimensional cualquiera de las dos condiciones $UU^* = 1$ y $U^*U = 1$ implica la otra; véase el § 36, Teoremas 1 y 2.) Esas transformaciones se llaman *ortogonales o unitarias*, según el espacio subyacente de producto interior sea real o complejo. Procedemos a deducir un par de útiles caracterizaciones alternativas de los mismos.

Teorema. *Las siguientes tres condiciones impuestas a una transformación lineal U sobre un espacio de producto interior, son equivalentes entre sí*

(1) $$U^*U = 1,$$

(2) $$(Ux, Uy) = (x, y) \text{ para toda } x \text{ y } y,$$

(3) $$\| Ux \| = \| x \| \text{ para toda } x.$$

Prueba. Si (1) es válido, entonces

$$(Ux, Uy) = (U^*Ux, y) = (x, y)$$

para toda x y y y, en particular

$$\| Ux \|^2 = \| x \|^2$$

para toda x; esto prueba ambas implicaciones (1) $\Rightarrow$ (2) y (2) $\Rightarrow$ (3). La prueba puede completarse demostrando que (3) implica a (1). Si (3) es válido, esto es, si $(U^*Ux, x) = (x, x)$ para toda x, entonces el § 71, Teorema 2, es aplicable a la transformación (autoadjunta) $U^*U - 1$; la conclusión es que $U^*U = 1$ (como se desea).

Puesto que (3) implica

(4) $$\| Ux - Uy \| = \| x - y \|$$

para toda x y y (la implicación recíproca (4) $\Rightarrow$ (3) es también verdadera y trivial), vemos que las transformaciones del tipo que trata el teorema se caracterizan por el hecho de que preservan distancia. Por esta razón llamaremos a esa transformación una *isometría.* Puesto que, como ya hemos observado, una isometría sobre un espacio finito dimensional es necesariamente ortogonal o unitaria (según el espacio sea real o complejo), el uso de la terminología nos facultará para tratar simultáneamente los casos real y complejo. Observamos que (sobre un espacio finito-dimensional) una isometría es siempre invertible y que $U^{-1}(= U^*)$ es una isometría, juntamente con U.

En cualquier sistema algebraico, y en particular en espacios vectoriales generales y espacios de productos interiores, es de interés considerar los automorfismos del sistema, esto es, considerar esos mapas uno a uno del sistema sobre sí mismo que preservan todas las relaciones estructurales entre sus elementos. Ya hemos visto que los automorfismos de un espacio vectorial son las transformaciones lineales invertibles. En un espacio de producto interior requerimos más de un isomorfismo, a saber, que también preserve los productos interiores (y, consecuentemente, longitudes y distancias). El teorema precedente demuestra que este requerimiento es equivalente a la condición de que la transformación sea una isometría. (Estamos suponiendo aquí la finito-dimensionalidad; sobre espacios finito-dimensionales, el alcance de una isometría no necesita ser todo el espacio. Este sacrificio sin importancia de la generalidad es en obsequio de la comodidad terminológica; para espacios finito-dimensionales no hay palabras de uso común que describan simultáneamente las transformaciones ortogonales y unitarias). Así, las dos preguntas, "¿qué transformaciones lineales son los análogos de números complejos de valor absoluto uno?" y "Cuáles son los automorfismos más generales de un espacio finito-dimensional de producto interior?" tienen la misma respuesta: isometrías. En la siguiente sección demostraremos que las isometrías suministran también la respuesta a una tercera importante cuestión.

§74. CAMBIO DE BASES ORTONORMALES

Hemos visto que la teoría del pase de una base lineal de un espacio vectorial a otro se estudia mejor por medio de una transformación lineal asociada A (§ § 46, 47); se suscita la cuestión de qué propiedades especiales tiene A cuando pasamos de una base *ortonormal* de un espacio de producto interior a otro. La respuesta es fácil.

Teorema 1. *Si $\mathfrak{X} = \{x_1, \ldots, x_n\}$ es una base ortonormal de un espacio n-dimensional $\mathfrak{V}$, de producto interior, y si U es una isometría sobre $\mathfrak{V}$, entonces $U\mathfrak{X} = \{Ux_1, \ldots, Ux_n\}$ es también una base ortonormal de $\mathfrak{V}$. Recíprocamente, si U es una transformación lineal y $\mathfrak{X}$ es una base ortonormal con la propiedad de que $U\mathfrak{X}$ es también una base ortonormal, entonces U es una isometría.*

Prueba. Puesto que $(Ux_i, Ux_j) = x_i, x_j) = \delta_{ij}$, se sigue que $U\mathfrak{X}$ es un conjunto ortonormal juntamente con $\mathfrak{X}$; es completo si $\mathfrak{X}$ lo es, puesto que $(x, Ux_i) = 0$ para $i = 1, \ldots, n$ implica que $(U^*x, x_i) = 0$ y de aquí que $U^*x = x = 0$. Si, recíprocamente $U\mathfrak{X}$ es un conjunto completo ortonormal juntamente con x, entonces tenemos $(Ux, Uy) = (x, y)$ cada vez que x y y están en $\mathfrak{X}$, y está claro que por linearidad obtenemos $(Ux, Uy) = (x, y)$ para toda x y y.

Observamos que la matriz (u_{ij}) de una transformación isométrica, con respecto de una base arbitraria ortonormal, satisface las condiciones

$$\sum_k \bar{u}_{ki} u_{kj} = \delta_{ij},$$

y que, recíprocamente, cualquier matriz como esa, juntamente con una base ortonormal, define una isometría. (Prueba: $U^*U = 1$. En el caso real las barras pueden ser omitidas). Por razones de brevedad diremos que una matriz que satisface estas condiciones es una *matriz isométrica.*

Una consecuencia interesante y fácil de nuestras consideraciones concernientes a isometrías es el siguiente corolario del § 56, Teorema 1.

TEOREMA 2. *Si A es una transformación lineal sobre un espacio complejo n-dimensional de producto interior $\mathcal{V}$, existe entonces una base ortonormal $\mathfrak{X}$ en $\mathcal{V}$ tal que la matriz $[A;\mathfrak{X}]$ es triangular, o equivalentemente, si $[A]$ es una matriz, entonces existe una matriz isométrica $[U]$ tal que $[U]^{-1}[A][U]$ es triangular.*

PRUEBA. En el § 56, en la deducción del Teorema 2 del Teorema 1, construimos una base (lineal) $\mathfrak{X} = \{x_1, \ldots, x_n\}$ con la propiedad de que $x_1, \ldots, x_j$ queda en $\mathfrak{M}_j$ y sobrentiende a $\mathfrak{M}_j$ para $j = 1, \ldots, n$ y demostramos que con respecto de esta base la matriz de A es triangular. Si supiéramos que esta base es también una base ortonormal, podríamos aplicar el Teorema 1 de la presente sección, para obtener el resultado deseado. Si $\mathfrak{X}$ no es una base ortonormal, es fácil convertirla en una; esto es precisamente lo que puede hacer el proceso de ortogonalización de Gram-Schmidt (§65). Aquí usamos una propiedad especial del proceso de Gram-Schmidt, a saber, que el elemento j-ésimo de la base ortonormal que construye es una combinación lineal de $x_1, \ldots, x_j$ y queda, en consecuencia, en $\mathfrak{M}_j$.

EJERCICIOS

1. Si $(Ax)(t) = x(-t)$ sobre $\mathcal{P}$ (con el producto interior dado por $(x, y) = \int_0^1 x(t)\overline{y(t)}\, dt$) es la transformación lineal de A isométrica. ¿Es auto-adjunto?

2. ¿Para cuáles valores de α son isométricas las matrices sigiuentes?

(a) $\begin{pmatrix} \alpha & 0 \\ 1 & 1 \end{pmatrix}$. (b) $\begin{pmatrix} \alpha & \frac{1}{2} \\ -\frac{1}{2} & \alpha \end{pmatrix}$.

3. Encuéntrese una matriz isométrica 3 por 3 cuya primera fila sea un múltiplo de $(1, 1, 1)$.

4. Si una transformación lineal tiene dos cualesquiera de las propiedades de ser auto-adjunta, isométrica o involutoria, entonces tiene la tercera. (Recuérdese que una involución es una transformación lineal A tal que $A^2 = 1$).

5. Si una matriz isométrica es triangular, entonces es diagonal.

6. Si $(x_1, \ldots, x_k)$ y $(y_1, \ldots, y_k)$ son dos secuencias de vectores en el mismo espacio de producto interior, entonces una condición necesaria y suficiente para que exista una isometría U tal que $Ux_i = y_i$, $i = 1, \ldots, k$, es que $(x_1, \ldots, x_k)$ y $(y_1, \ldots, y_k)$ tengan la misma gramiana.

7. El trazado $\xi \rightarrow \dfrac{\xi + 1}{\xi - 1}$ es el mapa del eje imaginario en el plano complejo. una vez alrededor del círculo unitario, sin tocar el punto 1; el trazado inverso (del círculo menos un punto al eje imaginario) es dado por la misma fórmula. Los análogos de transformación de estos hechos geométricos son como sigue.

(a) Si A es oblicua, entonces $A - 1$ es invertible.

(b) Si $U = (A + 1)(A - 1)^{-1}$, entonces U es isométrica. (Sugestión: $\|(A + 1)y\|^2 = \|(A - 1)y\|^2$ para toda y).

(c) $U - 1$ es invertible.

(d) Si U es isométrica y $U - 1$ es invertible, y si $A = (U + 1)(U - 1)^{-1}$, entonces A es oblicua.

Tanto A como U son conocidas como la *transforma de Cayley*, una de otra.

8. Supóngase que U es una transformación (que no se supone que sea lineal) que traza el mapa de un espacio de producto interior $\mathcal{V}$ sobre sí misma (esto es, si x está en $\mathcal{V}$, entonces Ux está en $\mathcal{V}$, y si y está en $\mathcal{V}$, entonces $y = Ux$ para alguna x en $\mathcal{V}$), en tal forma que $(Ux, Uy) = (x, y)$ para toda x y y.

(a) Pruébese que U es uno a uno y que si la transformación inversa es denotada por U^{-1}, entonces $(U^{-1}x, U^{-1}y) = (x, y)$ y $(Ux, y) = (x, U^{-1}y)$ para toda x y y.

(b) Pruébese que U es lineal. (Sugestión: $(x, U^{-1}y)$ depende linealmente de x).

9. Una *conjugación* es una transformación J (que no se supone lineal) que traza el mapa de un espacio unitario sobre sí misma y es tal que $J^2 = 1$ y $(Jx, Jy) = (y, x)$ para toda x y y.

(a) Dése un ejemplo de conjugación.

(b) Pruébese que $(Jx, y) = (Jy, x)$.

(c) Pruébese que $J(x + y) = Jx + Jy$.

(d) Pruébese que $J(\bar{\alpha} x) = \alpha . Jx$.

10. Se dice que una transformación lineal A es *real* con respecto de una conjugación J si $AJ = JA$.

(a) Dése un ejemplo de transformación hermitiana que no sea real y dése un ejemplo de una transformación real que no sea hermitiana.

(b) Si A es real, entonces el espectro de A es simétrico con respecto del eje real.

(c) Si A es real, entonces lo es A^*.

11. § 74. El Teorema 2 demuestra que la forma triangular puede ser lograda por una base ortonormal; ¿es cierto lo mismo para la forma de Jordan?

12. Si tr $A = 0$, existe entonces una matriz isométrica U tal que todas las notaciones diagonales de $[U]^{-1}[A][U]$ son cero. (Sugestión: véase el § 56, Ej. 6).

§75. PROYECCIONES PERPENDICULARES

Estamos ahora en condiciones de cumplir nuestra promesa anterior de investigar las proyecciones asociadas con las descomposiciones particulares de suma directa $\mathcal{V} = \mathfrak{M} \oplus \mathfrak{M}^\perp$. Llamaremos a esa proyección una *proyección perpendicular*. Puesto que $\mathfrak{M}^\perp$ está determinada en forma única por el subespacio $\mathfrak{M}$, no necesitamos especificar ambos sumandos directos asociados con una proyección si ya sabemos que es perpendicular. Llamaremos a la proyección (perpendicular) E sobre $\mathfrak{M}$ a lo largo de $\mathfrak{M}^\perp$ simplemente la proyección sobre $\mathfrak{M}$ y escribiremos $E = P_{\mathfrak{M}}$.

TEOREMA 1. *Una transformación lineal E es una proyección perpendicular si y sólo si* $E = E^2 = E^*$. *Las proyecciones perpendiculares son transformaciones lineales positivas y tienen la propiedad de que* $\| Ex \| \leqq \| x \|$ para toda x.

PRUEBA. Si E es una proyección perpendicular, entonces el §45, Teorema 1 y el teorema del §20 demuestran (después, por supuesto, de las sustituciones naturales como $\mathfrak{M}^\perp$ por $\mathfrak{M}^0$ y A^* por A'), que $E = E^*$. Recíprocamente, si $E = E^2 = E^*$, entonces la idempotencia de E nos garantiza que E es la proyección sobre $\mathfrak{R}$ a lo largo de $\mathfrak{N}$, donde, por supuesto $\mathfrak{R} = \mathfrak{R}(E)$ y $\mathfrak{N} = \mathfrak{N}(E)$ son el alcance y el espacio nulo de E, respectivamente. De aquí que necesitemos solamente demostrar que $\mathfrak{R}$ y $\mathfrak{N}$ son ortogonales. Para este propósito, sea x un elemento de $\mathfrak{R}$ y y cualquier elemento de $\mathfrak{N}$; el resultado deseado se sigue de la relación

$$(x, y) = (Ex, y) = (x, E^*y) = (x, Ey) = 0.$$

El carácter positivo de una E que satisface $E = E^2 = E^*$ se sigue de

$$(Ex, x) = (E^2x, x) = (Ex, E^*x) = (Ex, Ex) = \| Ex \|^2 \geqq 0.$$

Aplicando este resultado a la proyección perpendicular $1 - E$, vemos que

$$\| x \|^2 - \| Ex \|^2 = (x, x) - (Ex, x) = ([1 - E]x, x) \geqq 0;$$

esto concluye la prueba del teorema.

Para algunas de las generalizaciones de nuestra teoría es útil saber que la idempotencia, juntamente con la última propiedad mencionada en el Teorema 1 es también característica de las proyecciones perpendiculares.

TEOREMA 2. *Si una transformación lineal E es tal que* $E = E^2$ *y* $\| Ex \| \leqq \| x \|$ para toda x, entonces $E = E^*$.

PRUEBA. Vamos a demostrar que el alcance $\mathcal{R}$ y el espacio nulo $\mathfrak{N}$ de E son ortogonales. Si x está en $\mathfrak{N}^\perp$, entonces $y = Ex - x$ está en $\mathfrak{N}$, puesto que $Ey = E^2x - Ex = Ex - Ex = 0$. De aquí que $Ex = x + y$ con $(x, y) = 0$, de manera

$$\| x \|^2 \geqq \| Ex \|^2 = \| x \|^2 + \| y \|^2 \geqq \| x \|^2,$$

y, en consecuencia, $y = 0$. Consecuentemente, $Ex = x$, de manera que x está en $\mathcal{R}$; esto prueba que $\mathfrak{N}^\perp \subset \mathcal{R}$. Recíprocamente, si z está en $\mathcal{R}$, de manera que $Ez = z$, escribimos $z = x + y$ con x en $\mathfrak{N}^\perp$ y y en $\mathfrak{N}$. Entonces $z = Ez = Ex + Ey = Ex = x$. (La razón para esta igualdad es que x está en $\mathfrak{N}^\perp$ y, en consecuencia, en $\mathcal{R}$.) De aquí que z esté en $\mathfrak{N}^\perp$, de manera que $\mathcal{R} \subset \mathfrak{N}^\perp$, y, en consecuencia $\mathcal{R} = \mathfrak{N}^\perp$.

Necesitaremos también el hecho de que el teorema del § 42 sigue siendo verdadero si a la palabra "proyección" se le agrega la calificación de "perpendicular". Esta es una consecuencia inmediata de la caracterización precedente de proyecciones perpendiculares y del hecho de que las sumas y diferencias de transformaciones auto-adjuntas son auto-adjuntas si y sólo si conmutan. Por nuestros actuales métodos geométricos es también completamente fácil generalizar la parte del teorema que trata de sumas desde dos sumandos hasta un número finito cualquiera. La generalización es enunciada muy convenientemente en términos del concepto de ortogonalidad para proyecciones; diremos que dos proyecciones (perpendiculares) E y F son *ortogonales* si $EF = 0$. (La consideración de los adjuntos demuestra que esto es equivalente a $FE = 0$). El siguiente teorema demuestra que el lenguaje geométrico está justificado.

TEOREMA 3. *Dos proyecciones perpendiculares* $E = P_{\mathfrak{M}}$ *y* $F = P_{\mathfrak{N}}$ *son ortogonales si y sólo si los subespacios* $\mathfrak{M}$ *y* $\mathfrak{N}$ *(esto es, los alcances o árbitos de E y F) son ortogonales.*

PRUEBA. Si $EF = 0$, y si x y y están en los alcances (ámbitos) de E y F respectivamente, entonces

$$(x, y) = (Ex, Fy) = (x, E^*Fy) = (x, EFy) = 0.$$

Si, recíprocamente, $\mathfrak{M}$ y $\mathfrak{N}$ son ortogonales (de manera que $\mathfrak{N} \subset \mathfrak{M}^\perp$), entonces el hecho de que $Ex = 0$ para x en $\mathfrak{M}^\perp$ implica que $EFx = 0$ para toda x (puesto que Fx está en $\mathfrak{M}$ y, consecuentemente, en $\mathfrak{M}^{\perp}$).

§76. COMBINACIONES DE PROYECCIONES PERPENDICULARES

El teorema de la suma para proyecciones perpendiculares es ahora fácil.

Teorema 1. *Si $E_1, \ldots, E_n$ son proyecciones (perpendiculares), entonces una condición necesaria y suficiente para que $E = E_1 + \ldots + E_n$ sea una proyección (perpendicular) es que $E_iE_j = 0$, siempre que $i \neq j$ (esto es, que las E_i sean ortogonales dos a dos).*

Prueba. La prueba de la suficiencia de la condición es trivial; probamos explícitamente su necesidad solamente, de manera que ahora suponemos que E es una proyección perpendicular. Si x pertenece al alcance de alguna E_i, entonces

$$\begin{aligned} \| x \|^2 &\geqq \| Ex \|^2 = (Ex, x) = (\textstyle\sum_j E_jx, x) \\ &= \textstyle\sum_j (E_jx, x) = \sum_j \| E_jx \|^2 \geqq \| E_ix \|^2 = \| x \|^2, \end{aligned}$$

de manera que debemos tener igualdad a todo lo largo. Puesto que, en particular, debemos tener

$$\textstyle\sum_j \| E_jx \|^2 = \| E_ix \|^2,$$

se sigue que $E_jx = 0$, siempre que $j \neq i$. En otras palabras, toda x en el alcance de E_i está en el espacio nulo (y, consecuentemente, es ortogonal al alcance) de cada E_j con $j \neq i$; usando el § 75, Teorema 3, obtenemos la conclusión deseada.

Terminamos nuestra discusión de las proyecciones con un breve estudio de relaciones de orden. Es tentador escribir $E \leqq F$, para dos proyecciones perpendiculares $E = P$ y $F = P$, siempre que
Anteriormente, sin embargo, interpretamos el signo $\leqq$ cuando se usó en una expresión que implicaba las transformaciones lineales E y F (como en $E \leqq F$), como significado que $F - E$ es una transformación positiva. Hay también otras razones posibles para considerar que E es más pequeño que F; podríamos tener $\| Ex \| \leqq \| Fx \|$ para toda x, o $FE = EF = E$ [véase el § 42, (ii)]. La situación se subsana por el siguiente teorema, que juega aquí un papel similar al del § 75,

Teorema 3, esto es, establece la coincidencia de varios conceptos aparentemente diferentes, concernientes a proyecciones, algunos de los cuales se definen algebraicamente, mientras que otros se refieren a los objetos geométricos subyacentes.

TEOREMA 2. *Para proyecciones perpendiculares $E = P\mathfrak{M}$ y $F = P\mathfrak{N}$ las condiciones siguientes son mutuamente equivalentes*

(i) $E \leqq F$.

(ii) $\| Ex \| \leqq \| Fx \|$ para toda x

(iii) $\mathfrak{M} \subset \mathfrak{N}$.

(iva) $FE = E$,

(ivb) $EF = E$.

PRUEBA. Probaremos las relaciones de implicación (i) $\Rightarrow$ (ii) $\Rightarrow$ (iii) $\Rightarrow$ (iva $\Rightarrow$ (ivb) $\Rightarrow$ (i).

(i) $\Rightarrow$ (ii). Si $E \leqq F$, entonces, para toda x,

$$0 \leqq ([F - E]x, x) = (Fx, x) - (Ex, x) = \| Fx \|^2 - \| Ex \|^2$$

(puesto que E y F son proyecciones perpendiculares).

(ii) $\Rightarrow$ (iii). Suponemos que $|| Ex || \leqq || Fx ||$ para toda x. Tomemos anora cualquier x de $\mathfrak{M}$; entonces tenemos

$$\| x \| \geqq \| Fx \| \geqq \| Ex \| = \| x \|,$$

de manera que $|| Fx || x ||$ o $(x, x) - (Fx, x) = 0$, de donde

$$([1 - F]x, x) = \| (1 - F)x \|^2 = 0,$$

y, consecuentemente $x = Fx$. En otras palabras, x en $\mathfrak{M}$ implica que x está en $\mathfrak{N}$, pero había de probarse.

(iii) $\Rightarrow$ (iva). Si $\mathfrak{M} \subset \mathfrak{N}$, entonces Ex está en $\mathfrak{N}$ para toda x, de manera que $FEx = Ex$ para toda x, como había de probarse.

Que (iva) implica a (ivb) y que es de hecho equivalente al mismo, se sigue tomando adjuntos.

(iv) $\Rightarrow$ (i). Si $EF = FE = E$, entonces, para toda x,

$$(Fx, x) - (Ex, x) = (Fx, x) - (FEx, x) = (F[1 - E]x, x).$$

Puesto que E y F son proyecciones conmutativas, también lo son $(1 - E)$ y F, y consecuentemente, $G = F(1 - E)$ es una proyección. De aquí que

$$(Fx, x) - (Ex, x) = (Gx, x) = \| G\bar{x} \|^2 \geqq 0.$$

Esto completa la prueba del Teorema 2.

En términos de los conceptos introducidos hasta ahora, es posible dar una formulación que suena completamente intuitiva del teorema del § 42 (en cuanto se aplica a las proyecciones perpendiculares), como sigue. Para dos proyecciones perpendiculares E y F su suma, producto o diferencia, es también una proyección perpendicular si y sólo si F es respectivamente ortogonal a, conmutativo con o mayor que E.

EJERCICIOS

1. (a) Dé un ejemplo de una proyección que no sea una proyección perpendicular.

(b) Dése un ejemplo de dos proyecciones E y F, no pueden ser ambas perpendiculares) tales que $EF = 0$ y $FE \neq 0$.

2. Encuéntrese la proyección (perpendicular) de $(1, 1, 1)$ sobre el subespacio (unidimensional) de $\mathcal{C}^3$ sobretendido por $(1, -1, 1)$. (En otras palabras: encüéntrese la imagen del vector dado bajo la proyección sobre el subespacio dado).

3. Encuéntrense las matrices de todas las proyecciones perpendiculares sobre $\mathcal{C}^2$.

4. Si $U = 2E - 1$, entonces una condición necesaria y suficiente para que U sea una isometría involutoria es que E sea una proyección perpendicular.

5. Una transformación lineal U se llama una *isometría parcial* si existe un subespacio $\mathfrak{M}$ tal que $\| Ux \| = \| x \|$ siempre que x esté en $\mathfrak{M}$ y $Ux = 0$, siempre que x esté en $\mathfrak{M}^{\perp}$.

(a) El adjunto de una isometría parcial es una isometría parcial.

(b) Si U es una isometría parcial y si $\mathfrak{M}$ es un subespacio tal que $\| Ux \| = \| x \|$ o 0 según que x esté en $\mathfrak{M}$ o en $\mathfrak{M}^{\perp}$, entonces U^*U es la proyección perpendicular sobre $\mathfrak{M}$.

(c) Cada una de las siguientes cuatro condiciones es necesaria y suficiente para que una transformación lineal U sea una isometría parcial. (i) $UU^*U = U$, (ii) U^*U es una proyección, (iii) $U^*UU^* = U^*$, (iv) UU^* es una proyección.

(d) Si λ es un valor propio de una isometría parcial, entonces $|\lambda| \leqq 1$.

(e) Dése un ejemplo de una isometría parcial que tiene $\frac{1}{2}$ como valor propio.

6. Supóngase que A es una transformación lineal sobre, y $\mathfrak{M}$ es un sub espacio de un espacio vectorial finito-dimensional $\mathcal{V}$. Pruébese que si dim $\mathfrak{M} \leq$ dim $\mathfrak{M}^{\perp}$, entonces existen transformaciones lineales B y C sobre $\mathcal{V}$, tales que $Ax = (BC - CB)x$ para toda x de $\mathfrak{M}$. (Sugestión: sea B una isometría parcial tal que $\| Bx \| = \| x \|$ o 0, según que x esté en $\mathfrak{M}$ o en $\mathfrak{M}^{\perp}$ y tal que $\mathfrak{R}(B) = \mathfrak{M}^{\perp}$).

§77. COMPLEJIFICACION

En las últimas secciones hemos estado tratando simultáneamente espacios vectoriales reales y complejos. Algunas veces esto no es posible; el sistema de números complejos es más rico que el real. Hay teoremas que son verdaderos tanto para espacios reales como para complejos, pero para los cuales la prueba es mucho más fácil en el caso complejo y hay teoremas que son verdaderos para espacios complejos, pero no para espacios reales. (Un ejemplo de esta última clase es la aserción de que si el espacio es finito-dimensional, entonces toda transformación lineal tiene un valor propio). Por estas razones, es frecuentemente cómodo estar en condiciones de "complejificar" un espacio vectorial real, esto es, asociarlo con un espacio vectorial complejo de las mismas propiedades esenciales. El propósito de esta sección es describir ese proceso de complejificación.

Supóngase que $\mathcal{V}$ es un espacio vectorial real, y sea $\mathcal{V}^+$ el conjunto de todos los pares ordenados $\langle x, y\rangle$ con x y y en $\mathcal{V}$. Defínase la suma de los dos elementos de $\mathcal{V}^+$ por

$$\langle x_1, y_1\rangle + \langle x_2, y_2\rangle = \langle x_1 + x_2, y_1 + y_2\rangle,$$

y defínase el producto de un elemento de $\mathcal{V}^+$ por un número complejo $\alpha + i\beta$ (α y β real, $i = \sqrt{-1}$) por

$$(\alpha + i\beta)\langle x, y\rangle = \langle \alpha x - \beta y, \beta x + \alpha y\rangle.$$

(Para recordar estas fórmulas, pretendase que $\langle x, y\rangle$ significan $x + iy$). Un cálculo directo y sólo ligeramente laborioso muestra que el conjunto $\mathcal{V}^+$ se convierte en un espacio vectorial complejo con respecto de estas definiciones de las operaciones lineales.

El conjunto de los elementos $\langle x, y\rangle$ de $\mathcal{V}^+$ para los cuales $y = 0$ está en una correspondencia natural uno a uno con el espacio $\mathcal{V}$. Siendo un espacio vectorial complejo, el espacio $\mathcal{V}$ puede ser también considerado como un espacio vectorial real, si identificamos cada elemento x de $\mathcal{V}^+$ con su réplica $\langle x, 0\rangle$ de $\mathcal{V}^+$ (es excesivamente conveniente hacer esto) podemos decir que $\mathcal{V}^+$ (como espacio vectorial real) incluye a $\mathcal{V}$. Puesto que $\langle 0, y\rangle = i\langle y, 0\rangle$, de manera que $\langle x, y\rangle = \langle x, 0\rangle + i\langle y, 0\rangle$, nuestra convención de identificación nos faculta para decir que todo vector de $\mathcal{V}^+$ tiene la forma $x + iy$, con x y y en $\mathcal{V}$. Puesto que $\mathcal{V}$ e i $\mathcal{V}$ (donde i $\mathcal{V}$ denota el conjunto de todos los elementos $\langle x, y\rangle$ de $\mathcal{V}^+$ con $x = 0$) son subconjuntos de $\mathcal{V}^+$ con solamente 0 (esto es, $\langle 0, 0\rangle$) en común, se sigue que la representación de un vector de $\mathcal{V}^+$ en la forma $x + iy$ (con x y y en $\mathcal{V}$) es única. Hemos así construido un espacio vectorial complejo $\mathcal{V}^+$

con la propiedad de que $\mathcal{V}^+$ considerado como un espacio real incluye a $\mathcal{V}$ como un subespacio, y tal que $\mathcal{V}^+$ es la suma directa de $\mathcal{V}$ y de $i\ \mathcal{V}$. Aquí $i\mathcal{V}$ denota el conjunto de todos los elementos de $\mathcal{V}^+$ que tienen la forma iy para alguna y de $\mathcal{V}$). Llamaremos a $\mathcal{V}^+$ la *complejificación* de $\mathcal{V}$.

Si $\{x_1, \ldots, x_n\}$ es un conjunto linealmente independiente en $\mathcal{V}$ (coeficientes reales), entonces es también un conjunto linealmente independiente en $\mathcal{V}^+$ (coeficientes complejos). En realidad, si $\alpha_1, \ldots, \alpha_n, \beta_1, \ldots, \beta_n$ son números reales tales que $\sum_j (\alpha_j + i\beta_j)x_j = 0$, entonces $(\sum_j \alpha_j x_j) + i(\sum_j \beta_j x_j) = 0$,
y, consecuentemente, por la univocidad de la representación de vectores de $\mathcal{V}^+$ por medio de vectores de $\mathcal{V}$ se sigue que $(\sum_j \alpha_j x_j = (\sum_j \beta_j x_j) = 0$; el resultado deseado está) ahora implicado por la supuesta independencia lineal (real) de $\{x_1, \ldots x_n\}$ en $\mathcal{V}$. Si, además, $\{x_1, \ldots, x_n\}$ es una base de $\mathcal{V}$ (coeficientes reales), entonces es también una base de $\mathcal{V}^+$ (coeficientes complejos). En realidad si x y y están en $\mathcal{V}$ entonces existen números reales $\alpha_1, \ldots, \alpha_n, \beta_1, \ldots, \beta_n$ tales que $x = \sum_j \alpha_j x_j$ y $y = \sum_j \beta_j x_j$; se sigue que $x + iy = \sum_j (\alpha_j + i\beta_j)x_j$, y de aquí que $\{x_1, \ldots, x_n\}$ sobretiende a $\mathcal{V}^+$. Estos resultados implican que el espacio vectorial complejo $\mathcal{V}^+$ tiene la misma dimensión que el espacio vectorial real $\mathcal{V}$.

Hay un modo natural de extender cada transformación lineal A sobre $\mathcal{V}$ a una transformación lineal A^+ sobre $\mathcal{V}^+$; escribimos

$$A^+(x + iy) = Ax + iAy$$

cada vez que x y y están en $\mathcal{V}$. (La verificación de que A^+ es en realidad una transformación lineal sobre $\mathcal{V}^+$ es rutinaria). Una extensión similar es válida para funcionales lineales y hasta multilineales. Si, por ejemplo, w es una función bilineal (real) sobre $\mathcal{V}$, su extensión a $\mathcal{V}^+$ es la funcional bilineal (compleja) definida por

$$w^+(x_1 + iy_1, x_2 + iy_2) = w(x_1, x_2) - w(y_1, y_2) + i(w(x_1, y_2) + w(y_1, x_2)).$$

Si, por otra parte, w es alterna, entonces lo mismo es cierto de w^+. En realidad, las partes real e imaginaria de $w^+(x + iy, x + iy)$ son $w(x, x) - w(y, y)$ y $w(x, y) + w(y, x)$ respectivamente; si w es alterna, entonces w es simétrico oblicua (§ 30, Teorema 1) y, en consecuencia, w^+ es alterna. La misma prueba establece el resultado correspondiente también para funcionales k-lineales, para todos los valores de k. De esto y de la definición de determinantes se sigue que $\det A = \det A^+$ para toda transformación lineal A sobre $\mathcal{V}$.

El método de desarrollar funcionales bilineales es válido también para funcionales bilineales conjugadas. Si, esto es, $\mathcal{V}$ es un espacio (real) de producto interior, entonces hay un modo natural de introducir un producto interior (complejo) en $\mathcal{V}^+$; escribimos, por definición,

$$(x_1 + iy_1, x_2 + iy_2) = (x_1, x_2) + (y_1, y_2) - i((x_1, y_2) - (y_1, x_2)).$$

Obsérvese que si x y y son vectores ortogonales de $\mathcal{V}$, entonces

$$\| x + iy \|^2 = \| x \|^2 + \| y \|^2.$$

La correspondencia de A a A^+ preserva todas las propiedades algebraicas de las transformaciones. Así, si $B = \alpha A$ (con α real), entonces $B^+ = \alpha A^+$; si $C^+ = A^+ + B^+$, entonces $C^+ = A^+ + B^+$; y si $C = AB$, entonces $C^+ = A^+B^+$. Si, además $\mathcal{V}$ es un espacio de producto interior y si $B = A^*$, entonces $B^+ = (A^+)^*$. (Prueba: evalúese $[A^+(x_1 + iy_1), (x_2 + iy_2)]$ y $[x_1 + iy_1, B^+(x_2 + iy_2)]$.

Si A es una transformación lineal sobre $\mathcal{V}$ y si A^+ tiene un vector propio $x + iy$, con valor propio $\alpha + i\beta$ (donde x y y serán en $\mathcal{V}$ y α y β son reales), de manera que

$$Ax = \alpha x - \beta y,$$

$$Ay = \beta x + \alpha y,$$

entonces el subespacio de $\mathcal{V}$ sobretendido por x y y es invariante bajo A. (Puesto que toda transformación lineal sobre un espacio vectorial complejo tiene un vector propio, concluimos que toda transformación lineal sobre un espacio vectorial real deja invariante un subespacio de dimensión igual a 1 o 2). Si, en particular, A^+ resulta tener un valor real propio (esto es, si $\beta = 0$, entonces A tiene el mismo valor propio (puesto que $Ax = \alpha x$, $Ay = \alpha y$, y no desaparecen tanto x como y).

Ya hemos visto que toda base (real) de $\mathcal{V}$ tiene al mismo tiempo una base (compleja) en $\mathcal{V}^+$. Se sigue que la matriz de una transformación lineal A sobre $\mathcal{V}$, con respecto de alguna base $\mathfrak{X}$ de $\mathcal{V}$, es la misma que la matriz de A^+ sobre $\mathcal{V}^+$, con respecto de la base $\mathfrak{X}$ de $\mathcal{V}^+$. Este comentario está en la raíz de toda la teoría de las complejificación, el punto de vista ingenuo de la materia es que las matrices reales constituyen un caso especial de matrices complejas

EJERCICIOS

1. ¿Qué sucede si el proceso de complejificación descrito en el § 77 se aplica a un espacio vectorial que ya es complejo?

2. Pruébese que existe un isomorfismo único entre la complejificación descrita en el § 77 y la descrita en el§ 25, Ej. 5, con la propiedad de que cada vector "real" (esto es, cada vector en el espacio vectorial real originalmente dado) corresponde a sí mismo.

3. (a) ¿Cuál es la complejificación de $\mathcal{R}^1$?

(b) Si $\mathcal{V}$ es un espacio vectorial real n-dimensional, ¿cuál es la dimensión de su complejificación $\mathcal{V}^+$, considerada como espacio vectorial real?

4. Supóngase que $\mathcal{V}^+$ es el espacio complejo de producto interior obtenido complejificando un espacio real de producto interior $\mathcal{V}$.

(a) Pruébese que si $\mathcal{V}^+$ es considerado como un espacio vectorial real y si $A(x + iy) = x - iy$ siempre que x y y están en $\mathcal{V}$, entonces A es una transformación lineal sobre $\mathcal{V}^+$.

(b) ¿Es A auto-adjunto? ¿Isométrico? ¿Idempotente? ¿Involuntario?

(c) ¿Qué es $\mathcal{V}^+$ si se considera como espacio complejo?

5. Discútase la relación entre dualidad y complejificación y, en particular, la relación entre el adjunto de una transformación lineal sobre un espacio vectorial real y el adjunto de su complejificación.

6. Si A es una transformación sobre un espacio vectorial real $\mathcal{V}$ y si un subespacio $\mathfrak{M}$ de la complejificación $\mathcal{V}^+$ es invariante bajo A^+ entonces $\mathfrak{M}^\perp \cap \mathcal{V}$ es invariante bajo A.

§78. CARACTERIZACION DE ESPECTROS

Los siguientes resultados apoyan más que cualquier otra cosa hasta ahora, la analogía entre los números y transformaciones; aseguran que las propiedades que hacen que definamos las clases especiales de transformaciones que hemos estado considerando están reflejadas por su espectro.

Teorema 1. *Si A es una transformación autoadjunta de un espacio de producto interior, entonces el valor propio de A es real, si A es positivo, o estrictamente positivo, entonces todo valor propio de A es positivo, o estrictamente positivo, respectivamente.*

Prueba. Podemos pasar por alto el hecho de que la primera aserción es trivial en el caso real; la misma prueba sirva para establecer ambas aserciones tanto en el caso real como en el complejo. En realidad, si $Ax = \lambda x$, con $x \neq 0$, entonces

$$\frac{(Ax, x)}{\|x\|^2} = \frac{\lambda(x, x)}{\|x\|^2} = \lambda;$$

se sigue que si (Ax, x) es real, (véase el § 71, Teorema 4), entonces lo es y si (Ax, x) es positivo (o estrictamente positivo), entonces lo es λ.

Teorema 2. *Toda raíz de la ecuación característica de una transformación auto-adjunta sobre un espacio finito-dimensional de producto interior, es real.*

Prueba. En el caso complejo, las raíces de la ecuación característica son la misma cosa que valores propios, y el resultado se sigue del Teorema 1. Si A es una transformación simétrica sobre un espacio Euclidiano, entonces su complejificación A^+ es Hermitiana y el resultado se sigue del hecho de que A y A^+ tienen la misma ecuación característica.

Observamos que es una consecuencia inmediata del Teorema 2 que una transformación autoadjunta sobre un espacio finito dimensional de producto interior tenga siempre un valor propio.

Teorema 3. *Todo valor propio de una isometría tiene valor absoluto uno.*

Prueba. Si U es una isometría, y si $Ux = \lambda x$, con $x \neq 0$, entonces $\| x \| = \| Ux \| = \| \lambda \| \cdot \| x \|$.

Teorema 4. *Si A es autoadjunta o isométrica, entonces los vectores propios de A que pertenecen a valores propios distintos son ortogonales.*

Prueba. Supóngase $Ax_1 = \lambda_1 x_1$, $Ax_2 = \lambda_2 x_2$, $\lambda_1 \neq \lambda_2$. Si A es autoadjunta, entonces

$$(1) \qquad \lambda_1(x_1, x_2) = (Ax_1, x_2) = (x_1, Ax_2) = \lambda_2(x_1, x_2).$$

(El paso medio hace uso del carácter autoadjunto de A, y el último paso de la realidad de λ_2). En caso de que A sea una isometría, (1) es reemplazado por

$$(2) \qquad (x_1, x_2) = (Ax_1, Ax_2) = (\lambda_1/\lambda_2)(x_1, x_2);$$

recuérdese que $\bar{\lambda}_2 = 1/\lambda_2$. En cualquier caso, $(x_1, x_2) \neq 0$ implicaría que $\lambda_1 = \lambda_2$, de manera que debemos tener $(x_1, x_2) = 0$.

Teorema 5. *Si un subespacio $\mathfrak{M}$ es invariante bajo una isometría U sobre un espacio finito-dimensional de producto interior, entonces lo es $\mathfrak{M}^\perp$.*

Prueba. Considerada sobre el subespacio finito-dimensional $\mathfrak{M}$, la transformación U es aún una isometría y, consecuentemente, es invertible. Se sigue que toda x de $\mathfrak{M}$ puede escribirse en la forma $x = Uy$, con y en $\mathfrak{M}$; en otras palabras, si x está en $\mathfrak{M}$ y si $y = U^{-1}x$, entonces y está en $\mathfrak{M}$. De aquí que $\mathfrak{M}$ sea invariante bajo $U^{-1} = U^*$. Se sigue del § 45, Teorema 2, que $\mathfrak{M}^\perp$ es invariante bajo $(U^*)^* = U$.

Observamos que el mismo resultado para transformaciones autoadjuntas (aun en espacios no necesariamente finito-dimensionales) es trivial, puesto que si $\mathfrak{M}$ es invariante bajo A, entonces $\mathfrak{M}^\perp$ es invariante bajo $A^* = A$.

Teorema 6. *Si A es una transformación autoadjunta sobre un espacio finito-dimensional de producto interior, entonces la multiplicidad algebraica de cada valor propio λ_0 de A es igual a su multiplicidad geométrica, esto es, a la dimensión del subespacio $\mathfrak{M}$ de todas las soluciones de $Ax = \lambda_0 x$.*

Prueba. Está claro que $\mathfrak{M}$ es invariante bajo A y, en consecuencia, lo es $\mathfrak{M}^\perp$; denotamos por B y C la transformación lineal A, considerada solamente sobre $\mathfrak{M}$ y $\mathfrak{M}^\perp$ respectivamente. Tenemos

$$\det (A - \lambda) = \det (B - \lambda) \cdot \det (C - \lambda)$$

para toda λ. Puesto que B es una transformación autoadjunta sobre un espacio finito-dimensional, con un solo valor propio, a saber, λ_0, se sigue que λ_0 debe ocurrir como un valor propio de B con multiplicidad algebraica igual a la dimensión de $\mathfrak{M}$. Si esa dimensión es m, entonces $\det (B - \lambda) = (\lambda_0 - \lambda)^m$. Puesto que, por otra parte, λ_0 no es un valor propio de C en lo más mínimo y, puesto que consecuentemente, $\det (C - \lambda_0) \neq 0$, vemos que $\det (A - \lambda)$ contiene $(\lambda_0 - \lambda)$ como factor, exactamente m veces, como tenía que probarse.

Lo que hizo válida esta prueba fue la invariancia de $\mathfrak{M}^\perp$ y el hecho de que toda raíz de la ecuación característica de A es un valor propio de A. La última aserción es verdadera para toda transformación lineal sobre un espacio unitario; el siguiente resultado es una consecuencia de estas observaciones y del Teorema 5.

Teorema 7. *Si U es una transformación unitaria sobre un espacio unitario finito-dimensional, entonces la multiplicidad algebraica de cada valor propio de U es igual a su multiplicidad geómetrica.*

EJERCICIOS

1. Dése un ejemplo de una transformación lineal con dos vectores propios no ortogonales, pertenecientes a valores propios distintos.

2. Dése un ejemplo de una transformación lineal no positiva (sobre un espacio finito dimensional unitario) cuyos valores propios sean positivos.
3. (a) Si A es autoadjunto, entonces det A es real.
(b) Si A es unitaria, entonces $|\det A| = 1$.
(c) ¿Qué puede decirse sobre el determinante de una isometría parcial?

§79. TEOREMA ESPECTRAL

Estamos ahora listos para probar el principal teorema de este libro, el teorema del cual son corolarios todos los otros resultados de este capítulo. Hasta cierto punto, lo que hemos estado haciendo hasta ahora fue una cuestión de deporte (útil, sin embargo, para generealizaciones); deseábamos demostrar cuanto puede hacerse cómodamente con la teoría espectral, antes de probar el teorema espectral. En el caso complejo, incidentalmente, puede hacerse que el teorema espectral se siga del proceso de triangularización que ya hemos descrito; a causa de la importancia del teorema preferimos dar a continuación una prueba directa (completamente fácil). El lector puede encontrar provechoso adaptar el método de prueba (no el resultado) del § 56, Teorema 2, para probar tanto como pueda del teorema espectral y sus consecuencias.

TEOREMA 1. *A cada transformación lineal autoadjunta A sobre un espacio finito-dimensional de producto interior, corresponden números reales* $\alpha_1, \ldots, \alpha_r$ *y proyecciones perpendiculares* $E_1 \ldots, E_r$ *(donde r es un entero estrictamente positivo, no mayor que la dimensión del espacio), de manera que*

(1) *las* α_j *son distintas dos a dos (por pares),*
(2) *las* E_j *son ortogonales dos a dos (por pares) y diferentes de* 0,
(3) $\sum_j E_j = 1$,
(4) $\sum_j \alpha_j E_j = A$.

PRUEBA. Sean $\alpha_1, \ldots, \alpha_r$ los distintos valores propios de A, y sea E_j la proyección perpendicular sobre el subespacio consistente en todas las soluciones de $Ax = \alpha_j x$ $(j = 1, \ldots, r)$. La condición (1) se satisface entonces por definición: el hecho de que las α sean reales se sigue del § 78, Teorema 1. La condición (2) se sigue del § 78, Teorema 4. De la ortogonalidad de E_j inferimos que si $E = \sum_j E_j$, entonces E es una proyección perpendicular. La dimensión del alcance o ámbito de E es la suma de las dimensiones de los alcances de la E_j y, consecuentemente, según el § 78, Teorema 6, la dimensión del alcance de E es igual a la dimensión del espacio entero; esto implica a (3). (Alternativamente, si $E \neq 1$, entonces A considerada sobre el alcance de $1 - E$ sería una transformación autoadjunta sin valores propios). Para probar (4) tómese cualquier vector x y escríbase $x_j = E_j x$; se sigue que $Ax_j = \alpha_j x_j$ y de aquí que

$$Ax = A(\sum_j E_j x) = \sum_j Ax_j = \sum_j \alpha_j x_j = \sum_j \alpha_j E_j x.$$

Esto completa la prueba del teorema espectral.

La representación $A \sum_j \alpha_j E_j$ (donde las α y las E satisfacen las condiciones (1)-(3) del Teorema 1), es llamada una *forma espectral* de A; el principal efecto del resultado siguiente es el de probar la univocidad de la forma espectral.

TEOREMA 2. *Si $\sum_{j=1}^{r} \alpha_j E_j$ es la forma espectral de una transformación autoadjunta A, sobre un espacio finito-dimensional de producto interior, entonces las α son todas los distintos valores propios de A. Si, además, $1 \leq k \leq r$, entonces existen polinomios p_k, con coeficientes reales, tales que $p_k(\alpha_j) = 0$, siempre que $j \neq k$ y tales que $p_k(\alpha_k) = 1$; para cada uno de esos polinomios $p_k(A) = E_k$.*

PRUEBA. Puesto que $E_j \neq 0$, existe un vector x en el alcance de E_j. Puesto que $E_j x = x$ y $E_i x = 0$, siempre que $i \neq j$, se sigue que

$$Ax = \sum_i \alpha_i E_i x = \alpha_j E_j x = \alpha_j x,$$

de manera que cada α_j es un valor propio de A. Si, recíprocamente, λ es un valor propio de A, digamos $Ax = \lambda x$, con $x \neq 0$, entonces escribimos $x_j = E_j x$ y vemos que

$$Ax = \lambda x = \lambda \sum_j x_j$$

y

$$Ax = A \sum_j x_j = \sum_j \alpha_j x_j,$$

de manera que $\sum_j (\lambda - \alpha_j) x_j = 0$. Puesto que las x_j son ortogonales dos a dos, las que no son cero de entre ellas forman un conjunto linealmente independiente. Se sigue, para cada j, o $xj = 0$, o $\lambda = \alpha_j$. Puesto que $x \neq 0$, debemos tener $x_j \neq 0$ para alguna j y, consecuentemente, λ es en realidad igual a una de las α.

Puesto que $E_i E_j = 0$, si $i \neq j$, y $E_j^2 = E_j$, se sigue que

$$A^2 = (\sum_i \alpha_i E_i)(\sum_j \alpha_j E_j) = \sum_i \sum_j \alpha_i \alpha_j E_i E_j$$
$$= \sum_j \alpha_j^2 E_j.$$

De modo semejante

$$A^n = \sum_j \alpha_j{}^n E_j$$

para todo entero positivo n [en caso de que $n = 0$, úsese (3)], y de aquí que

$$p(A) = \sum_j p(\alpha_j) E_j$$

para todo polinomio p. Para concluir la prueba del teorema, todo lo que necesitamos hacer es exhibir un polinomio (real) p_k tal que $p_k(\alpha_j) = 0$ cada vez que $j \neq k$ y tal que $p_k(\alpha_k) = 1$. Si escribimos

$$p_k(t) = \prod_{j \neq k} \frac{t - \alpha_j}{\alpha_k - \alpha_j},$$

entonces p_k es un polinomio con todas las propiedades requeridas.

TEOREMA 3. *Si $\sum_{j=1}^{r} \alpha_j E_j$ es la forma espectral de una transformación autoadjunta A sobre un espacio finito-dimensional de producto interior, entonces una condición necesaria y suficiente para que una transformación lineal B conmute con A, es que conmute con cada E_j.*

PRUEBA. La suficiencia de la condición es trivial. Si $\sum_j \alpha_j E_j$ y $E_j B = B E_j$ para toda j, entonces $AB = BA$. La necesidad se sigue del Teorema 2; si B conmuta con A, entonces B conmuta con todo polinomio en A y, en consecuencia, B conmuta con cada E_j.

Antes de continuar explotando el teorema espectral, hacemos notar su interpretación matricial. Si escogemos una base ortonormal en el alcance de cada E_j, entonces la totalidad de los vectores de estas pequeñas bases es una base para el espacio entero; expresada en esta base, la matriz de A será diagonal. El hecho de que por una adecuada selección de una base ortonormal puede hacerse diagonal la matriz de una transformación autoadjunta o, equivalentemente, que cualquier matriz autoadjunta puede ser transformada isométricamente (esto es, reemplazada por $[U]^{-1}[A][U]$, donde U es una isometría) en una matriz diagonal, ya se sigue (en el caso complejo) de la teoría de la forma triangular. Dimos la versión algebraica por dos razones: primero, es la versión que se generaliza fácilmente hasta el caso infinito-dimensional y, segundo, aun en el caso finito-dimensional, escribiendo $\sum_j \alpha_j E_j$ tiene a menudo grandes ventajas notacionales y tipográficas sobre la notación matricial.

Haremos uso del hecho de que una transformación A no necesariamente autoadjunta es isométricamente diagonable (es decir, que su matriz es diagonal con respecto de una base adecuadamente ortonormal) si y sólo si son válidas para la misma las condiciones (1)-(4) del Teorema 1. En realidad, si tenemos (1)-(4), entonces es aplicable la prueba de la diagonabilidad, dada para transformaciones autoadjuntas; el recíproco lo dejamos como ejerecicio para el lector.

EJERCICIOS

1. Supóngase que A es una transformación lineal sobre un espacio complejo de producto interior. Pruébese que si A es hermitiana, entonces los factores lineales del polinomio mínimo de A son distintos. ¿Es verdadero el recíproco?

2. (a) Dos transformaciones lineales A y B sobre un espacio unitario son *unitariamente equivalentes* si existe una transformación unitaria U tal que $A = U^{-1}BU$. (El concepto correspondiente al caso real es llamado *equivalencia ortogonal.*) Pruébese que la equivalencia unitaria es una relación de equivalencia.

b) ¿Son A^*A y AA^* siempre unitariamente equivalentes?

(c) ¿Son A y A^* siempre unitariamente equivalentes?

3. ¿Cuál de los siguientes pares de matrices son unitariamente equivalentes?

(a) $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ y $\begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$.

(b) $\begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}$ y $\begin{pmatrix} \frac{1}{2} & \frac{1}{2} & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 0 & -1 \end{pmatrix}$.

(c) $\begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix}$ y $\begin{pmatrix} -1 & 0 & 0 \\ 0 & i & 0 \\ 0 & 0 & i \end{pmatrix}$.

(d) $\begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix}$ y $\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$.

4. Si dos transformaciones lineales son unitariamente equivalentes, entonces son similares y congruentes; si dos transformaciones lineales son o similares o congruentes, entonces son equivalentes. Demuestre por medio de ejemplos que estas relaciones de implicación son las únicas válidas entre estos conceptos.

§80. TRANSFORMACIONES NORMALES

Las más fáciles (y al mismo tiempo las más útiles) generalizaciones del teorema espectral se aplican a espacios complejos de producto interior (esto es, espacios unitarios). A fin de evitar complicaciones irrelevantes, excluimos en esta sección el caso real y concentramos nuestra atención solamente sobre espacios unitarios.

Hemos visto que toda transformación hermitiana es diagonable, y que una transformación arbitraria A puede ser escrita en la forma $B + iC$, con B y C hermitianos, ¿por qué no es cierto que simplemente diagonalizando B y C separadamente podemos diagonalizar A? La respuesta es, por supuesto, que la ortogonalización implica la selección de una adecuada base ortonormal, y no hay razón para esperar que una base que diagonalice B tendrá el mismo efecto sobre C. Es de considerable importancia saber la clase precisa de transformaciones para las cuales es válido el teorema espectral y afortunadamente esta clase es fácil de describir.

Llamaremos *normal* a una transformación lineal A si conmuta con su adjunta $A^*A = AA^*$. (Esta definición tiene sentido y se usa tanto en espacios reales como complejos de producto interior) continuaremos, sin embargo, usando técnicas que están inextricablemente ligadas con el caso complejo.) Hacemos notar primero que A es normal si y sólo si conmutan sus partes real e imaginaria. Supóngase, en realidad, que A es normal y que $A = B + iC$ con B y C hermitianas; puesto que $B = ½(A + A^*)$ y $C = \frac{1}{2i}(A - A^*)$, está claro que $BC = CB$. Si, recíprocamente, $BC = CB$, entonces las dos relaciones $A = B + iC$ y $A^* = B - iC$ implican que A es normal. Observamos que las transformaciones hermitiana y unitaria son normales.

La clase de transformaciones que poseen una forma espectral en el sentido del § 79, es precisamente la clase de transformaciones normales. La mitad de este enunciado es fácil de probar: si $A = \sum_j \alpha_j E_j$, entonces $A^* = \sum_j \bar{\alpha}_j E_j$, y se necesita un sencillo cálculo para demostrar que $A^*A = AA^* = \sum_j |\alpha_j|^2 E_j$. Para probar la recíproca, esto es, para probar que la normalidad implica la existencia de una forma espectral, tenemos dos alternativas. Podríamos deducir este resultado del teorema espectral para transformaciones hermitianas, usando las partes reales e imaginarias, o podríamos probar que los lemas esenciales del § 78, sobre los cuales descansa la prueba del caso hermitiano, son igualmente válidos para una transformación normal arbitraria. A causa de que estos métodos son de cierto interés, adoptamos el segundo procedimiento. Observamos que en el § 78 dispusimos de la maquinaria que se necesita para probar los lemas que siguen, de manera que podríamos haber enunciado inmediatamente el teorema de las transformaciones normales; la principal razón de que hayamos seguido el camino actual fue motivar la definición de normalidad.

TEOREMA 1. *Si A es normal, entonces una condición necesaria y suficiente para que x sea un vector propio de A es que sea un vector propio de A^*; si $Ax = \lambda x$, entonces $A^*x = \bar{\lambda}x$.*

PRUEBA. Observamos que la normalidad de A implica que

$$\| Ax \|^2 = (Ax, Ax) = (A^*Ax, x) = (AA^*x, x) = (A^*x, A^*x) = \| A^*x \|^2.$$

Puesto que $A - \lambda$ es normal juntamente con A y puesto que $(A - \lambda)^* = A^* - \bar{\lambda}$, obtenemos la relación

$$\| Ax - \lambda x \| = \| A^*x - \bar{\lambda}x \|,$$

de la cual se siguen inmediatamente las aserciones del teorema.

TEOREMA 2. *Si A es normal, entonces los vectores propios que pertenecen a distintos valores propios son ortogonales.*

PRUEBA. Si $Ax_1 = \lambda_1 x_1$ y $Ax_2 = \lambda_2 x_2$, entonces

$$\lambda_1(x_1, x_2) = (Ax_1, x_2) = (x_1, A^*x_2) = \lambda_2(x_1, x_2).$$

Este teorema generaliza el § 78, Teorema 4; en la prueba del teorema espectral para transformaciones hermitianas necesitamos también el § 78, Teoremas 5 y 6. El siguiente resultado toma el lugar del primero de éstos.

TEOREMAS 3. *Si A es normal, λ es un valor propio de A, y $\mathfrak{M}$ es el conjunto de todas las soluciones de $Ax = \lambda x$,* entonces *tanto $\mathfrak{M}$ como $\mathfrak{M}^\perp$ son invariantes bajo A.*

PRUEBA. El hecho de que $\mathfrak{M}$ sea invariante bajo A lo hemos visto antes; esto nada tiene que ver con la normalidad. Para probar que $\mathfrak{M}^\perp$ es también invariante bajo A, es suficiente con probar que $\mathfrak{M}$ es invariante bajo A^*. Esto es fácil; si x está en $\mathfrak{M}$, entonces

$$A(A^*x) = A^*(Ax) = \lambda(A^*x),$$

de manera que A^*x está también en $\mathfrak{M}$.

Este teorema es mucho más débil que su correspondiente del § 78. La cosa importante que observar, sin embargo, es que la prueba del § 78, Teorema 6, dependía sólo de la correspondiente versión debilitada del Teorema 5, los únicos subespacios que necesitan ser considerados son los del tipo mencionado en el teorema precedente.

Esto concluye el trabajo de preparación; el teorema espectral para operadores normales se sigue justamente como en el caso hermitiano. Si en los Teoremas del § 79 reemplazamos la palabras "autoadjunto" por "normal", destruimos toda referencia a la realidad, e insistimos en que el espacio subyacente de producto interior sea complejo, las partes restantes de los enunciados y todas las pruebas quedan sin cambio.

Es la teoría de las transformaciones normales la que es de principal interés en el estudio de los espacios unitarios. Uno de los hechos más útiles sobre transformaciones normales es que las condiciones espectrales del tipo dado en el § 78, Teoremas 1 y 3, que allí se demostró eran necesarias para el carácter autoadjunto, positivo e isométrico de una transformación, son suficientes en el caso normal

TEOREMA 4. *Una transformación normal sobre un espacio finito-dimensional unitario es* (1) *hermitiana,* (2) *positiva,* (3) *estrictamente positiva,* (4) *unitaria,* (5) *invertible,* (6) *idempotente si y sólo si todos sus valores propios son* (1′) *reales,* (2′) *positivos,* (3′) *estrictamente positivos,* (4′) *de valor absoluto uno* (5′) *diferentes de cero,* (6′) *iguales a cero o uno.*

PRUEBA. El hecho de que (1), (2), (3) y (4) impliquen (1′), (2′), (3′) y (4′), respectivamente, se sigue del § 78. Si A es invertible, y $Ax = \lambda x$, con $x \neq 0$, entonces $x = A^{-1}Ax = \lambda A^{-1}x$ y, en consecuencia, $\lambda \neq 0$; esto prueba que (5) implica (5′). Si A es idempotente y $Ax = \lambda x$, con $x \neq 0$, entonces $\lambda x = Ax = A^2x = \lambda^2 x$, de manera que $(\lambda - \lambda^2)x = 0$ y, en consecuencia, $\lambda = \lambda^2$; esto prueba que (6) implica a (6′). Obsérvese que estas pruebas son válidas para un espacio arbitrario de producto interior (ni siquiera necesariamente finito-dimensional) y que el supuesto auxiliar de que A es normal, también es superfluo.

Supóngase ahora que la forma espectral de A es $\sum_j \alpha_j E_j$. . Puesto que $A^* = \sum_j \bar{\alpha}_j E_j$, , vemos que (1′) implica a (1). Puesto que

$$(Ax, x) = \sum_j \alpha_j(E_j x, x) = \sum_j \alpha_j \| E_j x \|^2,$$

se sigue que (2') implica a (2). Si $\alpha_j > 0$ para toda j y si $(Ax, x) = 0$, entonces debemos tener $E_jx = 0$ para toda j y, en consecuencia $x = \sum_j E_jx = 0$; esto prueba que (3') implica a (3). La implicación de (4') a (4) se sigue de la relación

$$A^*A = \sum_j |\alpha_j|^2 E_j.$$

Si $\alpha_j \neq 0$ para toda j, podemos formar la transformación lineal $B = \sum_j \frac{1}{\alpha_j} E_j$; puesto que $AB = BA = 1$, se sigue que (5') implica a (5). Finalmente $A^2 = \sum_j \alpha_j^2 E_j$; de esto podemos inferir que (6') implica a (6).

Observamos que las relaciones de implicación (5) $\Rightarrow$ (5'), (2) $\Rightarrow$ (2'), y (3') $\Rightarrow$ (3) juntas cumplen una promesa que hicimos en el § 72. Si A es positiva e invertible, entonces A es estrictamente positiva.

EJERCICIOS

1. Dése un ejemplo de una transformación normal que no es hermitiana ni unitaria.
2. (a) Si A es una transformación lineal arbitraria (sobre un espacio finito-dimensional unitario) y si α y β son números complejos tales que $|\alpha| = |\beta| = 1$, entonces $\alpha A + \beta A^*$ es normal.
(b) Si $\|Ax\| = \|A^*x\|$ para toda x, entonces A es normal.
(c) ¿Es siempre normal la suma de, dos transformaciones normales?
3. Si A es una transformación normal sobre un espacio finito-dimensional unitaria y si $\mathfrak{M}$ es un subespacio invariante bajo A, entonces la restricción de A a $\mathfrak{M}$ es también normal.
4. Una transformación lineal A sobre un espacio finito-dimensional unitario $\mathcal{V}$ es normal si y sólo si $A\mathfrak{M} \subset \mathfrak{M}$ implica $A\mathfrak{M}^\perp \subset \mathfrak{M}^\perp$ para cada subespacio $\mathfrak{M}$ de $\mathcal{V}$.
5. (a) Si A es normal e idempotente, entonces es autoadjunto.
(b) Si A es normal y nilpotente, entonces es cero.
(c) Si A es normal y $A^3 = A^2$, entonces A es idempotente. ¿Permanece verdadera la conclusión si se omite el supuesto de normalidad?
(d) Si A es autoadjunta y si $A^k = 1$ para algún entero estrictamente positivo k, entonces $A^2 = 1$.
6. Si A y B son normales y si $AB = 0$, ¿se sigue que $BA = 0$?
7. Supóngase que A es una transformación lineal sobre un espacio n-dimensional unitario; sean $\lambda_1, \ldots, \lambda_n$ los valores propios de A (ocurriendo cada uno un número de veces igual a su multiplicidad algebraica). Pruébese que

$$\sum_i |\lambda_i|^2 \leq \operatorname{tr}(A^*A),$$

y que A es normal si y sólo si la igualdad es válida.
8. El *alcance numérico* de una transformación lineal A sobre un espacio

finito-dimensional unitario es el conjunto $W(A)$ de todos los números complejos de la forma (Ax, x), con $\|x\| = 1$.

(a) Si A es normal, entonces $W(A)$ es convexa. [Esto significa que si ξ y η están en $W(a)$ y si $0 \leqq \alpha \leqq 1$ entonces $\alpha\xi + (1-\alpha)\eta$ está también en $W(A)$].

(b) Si A es normal, entonces todo punto extremo de $W(A)$ es un valor propio de A. (Un punto extremo es uno que no tiene la forma $\alpha\xi + (1-\alpha)\eta$ para alguna ξ y η de $W(A)$ y para cualquier α propiamente entre 0 y 1)

(c) Es sabido que la conclusión de (a) permanece verdadera aún si no se supone la normalidad. Este hecho puede ser expresado como sigue: si A_1 y A_2 son transformaciones hermitianas, entonces el conjunto de todos los puntos de la forma $[(A_1x, x), (A_2x, x)]$ en el plano de coordenadas reales (con $\|x\| = 1$) es convexo. Demuéstrese que la generalización de esta aserción a más de dos transformaciones hermitianas es falsa.

(d) Pruébese que la conclusión de (b) puede ser alta para transformaciones no normales.

§81. TRANSFORMACIONES ORTOGONALES

Puesto que una transformación unitaria sobre un espacio unitario es normal, los resultados de la sección precedente incluyen la teoría de las transformaciones unitarias como un caso especial. Sin embargo, puesto que, una transformación ortogonal sobre un espacio real de producto interior no tienen valores propios, el teorema espectral, como lo conocemos hasta ahora, no nos da información sobre transformaciones ortogonales. No es difícil llegar a los hechos; la teoría de la complejificación se hizo a la medida para este propósito.

Supóngase que U es una transformación ortogonal sobre un espacio $\mathcal{V}$; finito-dimensional real de producto interior; sea U^+ la extensión de U a la complejificación $\mathcal{V}^+$. Puesto que $U^*U = 1$ (sobre $\mathcal{V}$), se sigue que $(U^+)^*U^+ = 1$ (sobre $\mathcal{V}^+$), esto es, que U^+ es unitario.

Sea $\lambda = \alpha + i\beta$ un número complejo (α y β son reales), y sea $\mathfrak{M}$ el subespacio consistente en todas las soluciones de $U^+z = \lambda z$ en $\mathcal{V}^+$. (Si λ no es un valor propio de U^+, entonces $\mathfrak{M} = \mathcal{O}$.) Si z está en $\mathfrak{M}$, escríbase $z = x + iy$, con x y y en $\mathcal{V}$. La ecuación

$$Ux + iUy = (\alpha + i\beta)(x + iy)$$

implica (véase el § 77) que

$$Ux = \alpha x - \beta y$$

y

$$Uy = \beta x + \alpha y.$$

Si multiplicamos el segundo del último par de ecuaciones por i y a continuación lo sustraemos del primero, obtenemos

$$Ux - iUy = (\alpha - i\beta)(x - iy).$$

Esto significa que $U^+\bar{z} = \bar{\lambda}\bar{z}$, donde el sugestivo y conveniente símbolo $\bar{z}$ denota, por supuesto, el vector $x - iy$. Puesto que el argumento (esto es, el paso de $U^+z = \lambda z$ a $U^+\bar{z} = \bar{\lambda}\bar{z}$) es reversible, hemos probado que el mapa $z \to \bar{z}$ es una correspondencia uno a uno entre $\mathfrak{M}$ y el subespacio $\bar{\mathfrak{M}}$ consistente en todas las soluciones $\bar{z}$ de $U^+\bar{z} = \overline{\lambda z}$. El resultado implica, entre otras cosas, que los valores propios complejos de U^+ vienen en pares; si λ es uno de ellos, entonces también lo es $\bar{\lambda}$. (Esta observación sola podríamos haberla obtenido más rápidamente del hecho de que los coeficientes de la característica polinomial de U^+ son reales).

No hemos hecho uso todavía del carácter unitario de U^+. Una forma en que podemos hacer uso del mismo es ésta. Si λ es un valor propio complejo (definitivamente no real) de U^+, entonces $\lambda \neq \bar{\lambda}$ se sigue que si $U^+z = \lambda z$, de manera que $U^+\bar{z} = \overline{\lambda z}$, entonces z y $\bar{z}$ son ortogonales. Esto significa que

$$0 = (x + iy, x - iy) = \| x \|^2 - \| y \|^2 + i((x, y) + (y, x)),$$

y de aquí que $|| x ||^2 = || y ||^2$ y $(x, y) = -(y, x)$. Puesto que un producto interior real es simétrico $[(x, y) = (y, x)]$ se sigue que $(x, y) = 0$. Esto, a su vez, implica que $|| z ||^2 = || x ||^2 + || y ||^2$ y de aquí que $|| x || = || y || = \dfrac{1}{\sqrt{2}} || z ||$.

Si λ_1 y λ_2 son valores propios de U^+ con $\lambda_1 \neq \lambda_2$ y $\lambda_1 \neq \bar{\lambda}_2$ y si $z_1 = x_1 + iy_1$ y $z_2 = x_2 + iy_2$ son vectores propios correspondientes (x_1, x_2 y_1, y_2 en $\mathcal{V}$), entonces z_1 y z_2 son ortogonales y (puesto que $\bar{z}_2$ es un vector propio que pertenece al valor propio $\bar{\lambda}_2$) z_1 y $\bar{z}_2$ son también ortogonales. Usando de nuevo la expresión para el producto interior complejo sobre $\mathcal{V}^+$ en términos del producto interior real sobre $\mathcal{V}$, vemos que

$$(x_1, x_2) + (y_1, y_2) = (x_1, y_2) - (y_1, x_2) = 0$$

y

$$(x_1, x_2) - (y_1, y_2) = (x_1, y_2) + (y_1, x_2) = 0.$$

Se sigue que los cuatro vectores x_1, x_2, y_1 y y_2 son ortogonales dos a dos (por pares).

La transformación unitaria U^+ podría también tener valores reales propios. Sin embargo, puesto que sabemos que los únicos valores reales propios posibles de U^+ tienen valor absoluto uno, se sigue que los únicos valores reales propios posibles de U^+ son 1 y -1. Si $U^+(x + iy) = \pm (x + iy)$, entonces $Ux = \pm x$ y $Uy = \pm y$, de manera que los vectores propios de U^+ con valores reales propios se obtienen colocando juntos los vectores propios de U de una manera obvia

Estamos ahora listos para dar el paso final. Dada U escogemos una base ortonormal, digamos $\mathfrak{X}_1$, en el múltiple lineal de soluciones de $Ux = x$ (en $\mathcal{V}$), y, de modo semejante, escogemos una base otornormal, digamos $\mathfrak{X}_{-1}$, en el múltiple lineal de soluciones de $Ux = -x$ (en $\mathcal{V}$). (Los conjuntos $\mathfrak{X}_1$ y $\mathfrak{X}_{-1}$ pueden estar vacíos). A continuación, para cada par conjugado de valores propios complejos λ y $\bar{\lambda}$ de U^+ escogemos una base ortonormal $\{z_1, \ldots, z_r\}$ en el múltiple lineal de soluciones de $U^+z = \lambda z$ (en $\mathcal{V}^+$). Si $z_j = x_j + iy_j$ (con x_j y y_j en $\mathcal{V}$), sea $\mathfrak{X}_\lambda$ el conjunto $\{\sqrt{2}\, x_1, \sqrt{2}\, y_1, \ldots, \sqrt{2}\, x_r, \sqrt{2}\, y_r\}$ de vectores en $\mathcal{V}$. Los resultados que hemos obtenido implican que si formamos la unión de todos los conjuntos $\mathfrak{X}_1$, $\mathfrak{X}_{-1}$, y $\mathfrak{X}_\lambda$, para todos los valores propios λ de U^+, obtenemos una base ortonormal de $\mathcal{V}$. En caso de que $\mathfrak{X}_1$ tenga tres elementos, $\mathfrak{X}_{-1}$ tiene cuatro elementos y hay dos pares conjugados $\{\lambda_1, \bar{\lambda}_1\}$ y $\{\lambda_2, \bar{\lambda}_2\}$, entonces la matriz de U con respecto de la base así construida se ve así:

$$\begin{bmatrix}
1 &&&&&&&&&& \\
& 1 &&&&&&&&& \\
&& 1 &&&&&&&& \\
&&& -1 &&&&&&& \\
&&&& -1 &&&&&& \\
&&&&& -1 &&&&& \\
&&&&&& -1 &&&& \\
&&&&&&& \alpha_1 & -\beta_1 && \\
&&&&&&& \beta_1 & \alpha_1 && \\
&&&&&&&&& \alpha_2 & -\beta_2 \\
&&&&&&&&& \beta_2 & \alpha_2
\end{bmatrix}$$

(Todos los términos no explícitamente indicados son iguales a cero). En general, hay una hilera de $+1$ sobre el diagonal principal, seguida por una hilera de -1 y a continuación hay una hilera de cajas dos por dos que corren diagonal abajo; cada caja tiene la forma $\begin{pmatrix} \alpha & -\beta \\ \beta & \alpha \end{pmatrix}$ con $\alpha^2 + \beta^2 = 1$. El hecho de que $\alpha^2 + \beta^2 = 1$ implica que podemos determinar un número real θ tal que $\alpha = \cos\theta$ y $\beta = \operatorname{sen}\theta$; es costumbre usar esta representación trigonométrica al escribir la forma canónica de la matriz de una transformación ortogonal.

EJERCICIOS

1. Todo valor propio de una transformación ortogonal tiene valor absoluto 1.
2. Si $A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$, ¿cuántas matrices ortogonales (reales) P hay con la propiedad $P^{-1}AP$ es diagonal?
3. Enúnciese y pruébese un análogo razonable del teorema espectral para transformaciones normales sobre un espacio real de producto interior.

§82. FUNCIONES DE TRANSFORMACIONES

Uno de los más útiles conceptos en la teoría de las transformaciones normales sobre espacios unitarios es el de una función de una transformación. Si A es una transformación normal con forma espectral $\sum_j \alpha_j E_j$ (para esta discusión suponemos temporalmente que el espacio vectorial subyacente es un espacio unitario) y si f es una función arbitraria valuada complejamente, definida cuando menos en los puntos α_j, entonces definimos una transformación lineal $f(A)$ como

$$f(A) = \sum_j f(\alpha_j)E_j.$$

Puesto que para polinomios p (y aún para funciones racionales) hemos ya visto que nuestra definición anterior de $p(A)$ da, si A es normal, $p(A) = \sum_j p(\alpha_j)E_j$, vemos que el nuevo concepto es una generalización del anterior. La ventaja de considerar $f(A)$ para funciones arbitrarias f es para nosotros en gran parte notacional; no introduce nada nuevo conceptualmente. en realidad, para una f arbitraria, podemos escribir $f(\alpha_j) = \beta_j$, y en seguida podemos encontrar un polinomio p que, en el conjunto finito de números complejos distintos α_j, tome, respectivamente, los valores β_j. Con este polinomio p tenemos $f(A) = p(A)$, de manera que la clase de transformaciones definidas por la formación de funciones arbitrarias no es nada esencialmente nuevo; sólo ahorra la dificultad de construir un polinomio adecuado a cada caso especial. Así, por ejemplo, si para cada número complejo λ escribimos

$$f_\lambda(\zeta) = 0 \quad \text{siempre que } \zeta \neq \lambda$$

y

$$f_\lambda(\lambda) = 1,$$

entonces $f_\lambda(A)$ es la proyección perpendicular sobre el subespacio de soluciones de $Ax = \lambda x$.

Observamos que si $f(\zeta) = \frac{1}{\zeta}$, entonces (suponiendo por supuesto que f está definida para toda α_j, esto es $\alpha_j \neq 0$) $f(A) = A^{-1}$ y si $f(\zeta) = \overline{\zeta}$, entonces $f(A) = A^*$. Estos enunciados implican que si f es una función arbitraria racional de ζ y $\overline{\zeta}$, obtenemos $f(A)$ por los reemplazos $\zeta \to A$, $\overline{\zeta} \to A^*$ y $\frac{1}{\zeta} = A^{-1}$. El símbolo $f(A)$ está, sin embargo, definido para funciones mucho más generales y en lo que sigue nos sentiremos libres para hacer uso de expresiones tales como e^A y $\sqrt{A}$.

Una función particularmente importante es la raíz cuadrada de transformaciones positivas. Consideramos $f(\zeta) = \sqrt{\zeta}$ definida para toda $\zeta \geqq 0$ real como la raíz cuadrada positiva de ζ, y para toda $A = \sum_j \alpha_j E_j$ positiva escribimos

$$\sqrt{A} = \sum_j \sqrt{\alpha_j}\, E_j.$$

(Recuérdese que $\alpha_j \geqq 0$ para toda J. La discusión que sigue se aplica tanto a espacios reales como complejos de prodccto anterior). Está claro que $\sqrt{A} \geqq 0$ y que $(\sqrt{A})^2 = A$; desearíamos investigar hasta qué punto estas propiedades caracterizan a $\sqrt{A}$. A primera vista podría parecer sin esperanza buscar alguna univocidad, puesto que si consideramos $B = \sum_j \pm \sqrt{\alpha_j}\, E_j$, con una selección arbitraria de signo en cada lugar, tenemos todavía $A = B^2$. La transformación $\sqrt{A}$ que construimos fue, sin embargo, positiva y podemos demostrar que esta propiedad adicional garantiza la univocidad. En otras palabras, si $A = B^2$ y $B \geqq 0$, entonces $B = \sqrt{A}$. Para probar esto, sea $B = \sum_k \beta_k F_k$ la forma espectral de B. entonces

$$\sum_k \beta_k{}^2 F_k = B^2 = A = \sum_j \alpha_j E_j.$$

Puesto que las β_k son distintas y positivas, también lo son las $\beta_k{}^2$; la univocidad de la forma espectral de A implica que cada $\beta_k{}^2$ es igual a alguna α_j (y viceversa) y que las correspondientes E y F son iguales. Por una permutación de los índices podemos, en consecuencia, lograr $\beta_j{}^2 = \alpha_j$ para toda j, de manera que $\beta_j = \sqrt{\alpha_j}$, como tuvo que demostrarse.

Hay varias importantes aplicaciones de la existencia de raíces cuadradas para operadores positivos; daremos ahora dos de ellas. Primero; recordamos que en el § 72 mencionamos tres posibles definiciones de una transformación positiva A y adoptamos la más débil, a saber, que A es autoadjunto y $(Ax, x) \geqq 0$ para toda x. La más fuerte de las tres definiciones posibles era que podíamos escribir A en la forma $A = B^2$ para alguna B autoadjunta. Señalamos que el resultado de esta sección concerniente a raíces cuadradas implica que la (aparentemente) más débil de nuestras condiciones es, en consecuencia, equivalente a la más fuerte. (De hecho, podemos hasta obtener una raíz cuadrada única positiva.)

Segundo: en el § 72 enunciamos también que si A y B son positivas y conmutativas, entonces AB es también positiva; podemos ahora dar una fácil prueba de esta aserción. Puesto que $\sqrt{A}$ y $\sqrt{B}$ son funciones (polinomiales en) A y B, respectivamente, la conmutividad de A y B implica que $\sqrt{A}$ y $\sqrt{B}$ conmutan entre sí; consecuentemente

$$AB = \sqrt{A}\sqrt{A}\sqrt{B}\sqrt{B} = \sqrt{A}\sqrt{B}\sqrt{A}\sqrt{B} = (\sqrt{A}\sqrt{B})^2.$$

Puesto que $\sqrt{A}$ y $\sqrt{B}$ son autoadjuntas y conmutativas, su producto es autoadjunto y, en consecuencia, su cuadrado es positivo.

La teoría espectral hace también completamente fácil caracterizar la matriz (con respecto de un sistema arbitrario de coordenadas ortonormales) de una transformación positiva A. Puesto que det A es el producto de los valores propios de A, está claro que $A \geqq 0$ implica det $A \geqq 0$. (La discusión del § 55 se aplica directamente sólo a espacios complejos de producto interior; la modificación apropiada necesitada para la discusión de transformación autoadjuntas sobre espacios reales posibles es, sin embargo, completamente fácil de suministrar.) Si consideramos la propiedad definitoria de positividad expresada en términos de la matriz (α_{ij}) de A, que es $\sum_i \sum_j \alpha_{ij}\xi_i\bar{\xi}_j \geqq 0$, observamos que la última expresión permanece positiva si restringimos las coordenadas $(\xi_1, \quad, \xi_n)$ requiriendo que algunas de ellas desaparezcan. En términos de la matriz esto significa que si tachamos las columnas numeradas $j_1, \ldots, j_k$ y tachamos también las filas que llevan los *mismos* números, la pequeña matriz remanente es todavía positiva y, consecuentemente, lo es su determinante. Este hecho se expresa usualmente diciendo que los *menores principales* del determinante de una matriz positiva son positivos. El coeficiente de la potencia j-ésima de λ en el polinomio característico det $(A - \lambda)$ de A es (excepto por el signo) la suma de todos los menores principales de n-j filas y columnas. El signo es alternativamente más y menos; esto implica que si A tiene menores positivos principales y es autoadjunta (de manera que se sepa que los ceros de det $[A - \lambda]$ son reales), entonces los valores propios de A son positivos. Puesto que el carácter autoadjunto de una matriz es discernible observando si es o no simétrica hermitiana ($\alpha_{ij} = \overline{\alpha_{ij}}$), nuestros comentarios reducen el problema de determinar si es o no positiva una matriz, a un número finito de cálculos elementales.

EJERCICIOS

1. Correspondiendo a cada transformación unitaria U hay una transformación hermitiana A tal que $U = e^{iA}$.
2. Discútase la teoría de funciones de una transformación normal sobre un espacio real de producto interior.
3. Si $A \leqq B$ y si C es una transformación positiva que conmuta con A y con B, entonces $AC \leqq BC$.
4. Una transformación autoadjunta tiene una raíz cúbica autoadjunta única.
5. Encuentre todas las raíces cúbicas hermitianas de la matriz

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 8 \end{pmatrix}.$$

6. (a) Dése un ejemplo de una transformación lineal A sobre un espacio finito dimensional unitario tal que A no tenga raíz cuadrada.
(b) Pruébese que toda transformación hermitiana sobre un espacio unitario finito-dimensional tiene raíz cuadrada.
(c) ¿Tiene raíz cuadrada toda transformación autoadjunta sobre un espacio finito dimensional euclidiano?
7. (a) Pruébese que si A es una transformación lineal positiva sobre espacio finito dimensional de producto interior, entonces $\rho(\sqrt{A}) = \rho(A)$.
(b) Si A es una transformación lineal sobre un espacio finito-dimensional de producto interior, ¿es cierto qué $\rho(A^*A) = \rho(A)$?

8. Si $A \geqq 0$ y si $(Ax, x) = 0$ para alguna x, entonces $Ax = 0$.

9. Si $A \geqq 0$, entonces $|(Ax, y)|^2 \leqq (Ax, x)(Ay, y)$ para toda x y y.

10. Si los vectores $x_1, \ldots, x_k$ son linealmente independientes, entonces su gramiana es no singular.

11. Toda matriz positiva es una gramiana.

12. Si A y B son transformaciones lineales sobre un espacio finito-dimensional de producto interior, y si $0 \leqq A \leqq B$, entonces $\det A \leqq \det B$. (Sugestión: la conclusión es trivial si $\det B = 0$; si $\det B \neq 0$, entonces $\sqrt{B}$ es invertible.)

13. Si una transformación lineal A sobre un espacio finito-dimensional de producto interior es estrictamente positiva, y si $A \leqq B$, entonces $B^{-1} \leqq A^{-1}$. (Sugestión: ensáyese primero $A = 1$.)

14. (a) Si B es una transformación hermitiana sobre un espacio finito-dimensional unitario, entonces $1 + iB$ es invertible.

(b) Si A es positiva e invertible y B es hermitiana, entonces $A + iB$ es invertible.

15. Si $0 \leqq A \leqq B$, entonces $\sqrt{A} \leqq \sqrt{B}$. (Sugestión: calcúlese

$$(\sqrt{B} + \sqrt{A} + \epsilon)(\sqrt{B} - \sqrt{A} + \epsilon),$$

y pruébese en consecuencia que el segundo factor es invertible, siempre que $\varepsilon > 0$).

16. Supóngase que A es una transformación autoadjunta sobre un espacio finito-dimensional de producto interior; escríbase $|A| = \sqrt{A^2}$, $A_+ = \frac{1}{2}(|A| + A)$, y $A_- = \frac{1}{2}(|A| - A)$.

(a) Pruébese que $|A|$ es la más pequeña transformación hermitiana que conmuta con A y para la cual tanto $A \leqq |A|$ como $-A \leqq |A|$. ("La más pequeña" se refiere, por supuesto, a la ordenación de las transformaciones hermitianas).

(b) Pruébese que A es la más pequeña transformación positiva que conmuta con A y para la cual $-A \leqq .A+$.

(c) Pruébese que A es la más pequeña transformación positiva que conmuta con A y para la cual $-A \leqq A-$.

(d) Pruébese que si A y B son autoadjuntas y conmutativas, entonces existe una transformación mínima autoadjunta C que conmuta tanto con A y B como con $A \leqq C$ y $B \leqq C$.

17. (a) Si A y B son transformaciones lineales positivas sobre un espacio unitario finito-dimensional, y si A^2 y B^2 son unitariamente equivalentes, entonces A y B son unitariamente equivalentes.

(b) ¿Es verdadero el análogo real de (a)?

§83. DESCOMPOSICION POLAR

Hay otra consecuencia útil de la teoría de las raíces cuadradas, a saber, el análogo de la representación polar $\zeta = \rho e^{i\theta}$ de un número complejo.

TEOREMA 1. *Si A es una transformación lineal arbitraria sobre un espacio finito-dimensional de producto interior, entonces hay una transformación positiva P (determinada unicamente), y hay una isometría U, tal que $A = UP$. Si A es invertible, entonces U está también unicamente determinada por A.*

PRUEBA. Aunque no es lógicamente necesario hacerlo así, daremos primero la prueba en el caso en que A es invertible; la prueba general es una modificación obvia de esta especial, y la prueba especial da mayor comprensión de la estructura geométrica de la transformación A.

Puesto que la transformación A^*A es positiva, podemos encontrar su raíz cuadrada positiva (única) $P = \sqrt{A^*A}$. Escribimos $V = PA^{-1}$; puesto que $VA = P$, el teorema quedará probado si podemos probar que V es una isometría, puestos entonces podemos escribir $U = V^{-1}$. Puesto que

$$V^* = (A^{-1})^*P^* = (A^*)^{-1}P,$$

Vemos que

$$V^*V = (A^*)^{-1}PPA^{-1} = (A^*)^{-1}A^*AA^{-1} = 1,$$

de manera que V es una isometría, y estamos listos.

Para probar la univocidad, observamos que $UP = U_0P_0$ implica $PU^* = P_0U_0^*$ y, en consecuencia

$$P^2 = PU^*UP = P_0U_0^*U_0P_0 = P_0^2.$$

Puesto que la transformación positiva $P^2 = P_0^2$ tiene sólo una raíz cuadrada positiva, se sigue que $P = P_0$. (En esta parte de la prueba no usamos la invertibilidad de A). Si A es invertible, entonces lo es P (puesto que $P = U^{-1}A$), y de esto obtenemos (multiplicando la relación $UP = U_0P_0$ sobre la derecha por $P^{-1} = P_0^{-1}$) que $U = U_0$.

Pasamos ahora al caso general, donde no suponemos que A es invertible. Formamos P exactamente en la misma forma que anteriormente, de manera que $P^2 = A^*A$, y entonces observamos que para todo vector x tenemos

$$\| Px \|^2 = (Px, Px) = (P^2x, x) = (A^*Ax, x) = \| Ax \|^2.$$

Si para cada vector $y = Px$ en el alcance $\mathcal{R}(P)$ de P escribimos $Uy = Ax$, entonces la transformación U es conservadora de la longitud, dondequiera que se defina. Debemos demostrar que U está determinada inambiguamente, esto es, que $Px_1 = Px_2$ implica $Ax_1 = Ax_2$. Esto es verdad puesto que $P(x_1 - x_2) = 0$ es equivalente a $\| P(x_1 - x_2)\| = 0$ y esta última condición implica $\| A(x_1 - x_2)\| = 0$. El alcance de la transformación U, definido hasta ahora sobre el subespacio $\mathcal{R}(P)$ solamente, es $\mathcal{R}(A)$. Puesto que U es lineal, $\mathcal{R}(A)$ y $\mathcal{R}(P)$ tienen la misma dimensión, y, por lo tanto, $(\mathcal{R}(A))^{\perp}$ y $(\mathcal{R}(P))^{\perp}$ tienen la misma dimensión. Si definimos a U sobre $(\mathcal{R}(P))^{\perp}$ como cualquier transformación lineal e isométrica de $(\mathcal{R}(P))^{\perp}$ sobre $(\mathcal{R}(A))^{\perp}$, entonces U, determinada por lo mismo sobre toda $\mathcal{V}$, es una isometría con la propiedad $Upx = Axp$ para toda x. Esto completa la prueba.

Aplicando a A^* en vez de A el teorema que se acaba de probar y tomando a continuación adjuntos, obtenemos también el hecho dual de que toda A puede ser escrita en la forma $A = PU$ con una U isométrica U y una P positiva. En contraste con la descomposición cartesiana (§ 70), llamamos la representación $A = UP$ una *descomposición polar de A*.

En términos de descomposiciones polares, obtenemos una nueva caracterización de normalidad.

TEOREMA 2. *Si $A = UP$ es una descomposición polar de la transformación A, entonces una condición necesaria y suficiente para que A sea normal es que $PU = UP$.*

PRUEBA. Puesto que U no está necesariamente determinada únicamente por A, el enunciado debe interpretarse como sigue: si A es normal, entonces P conmuta con cada U, y si P conmuta con *alguna* U, entonces A es normal. Puesto que $AA^* = UP^2U^* = UP^2U^{-1}$ y $A^*A = P^2$, está claro que A es normal si y sólo si U conmuta con P^2. Puesto que, sin embargo, P^2 es una función de P y viceversa P es una función de $P^2(P = \sqrt{P^2})$, se sigue que conmutar con P^2 es equivalente a conmutar con P.

EJERCICIOS

1. Si una transformación lineal sobre un espacio finito-dimensional de producto interior tiene sólo una descomposición polar, entonces es invertible.
2. Usese el cálculo funcional para deducir la descomposición polar de un operador normal.
3. (a) Si A es una transformación lineal arbitraria sobre un espacio finito-dimensional de producto interior, entonces hay una isometría parcial U, y hay una transfomación positiva P, tal que $\mathcal{N}(U) = \mathcal{N}(P)$ y tal que $A = UP$. Las transformaciones U y P están determinadas únicamente por estas condiciones.

(b) La transformación A es normal si y sólo si las transformaciones U y P descritas en (a) conmutan entre sí.

§84. CONMUTATIVIDAD

El teorema espectral para operadores autoadjuntas y normales y el cálculo funcional pueden también usarse para resolver ciertos problemas concernientes a la conmutatividad. Es éste un tema profundo y extenso; más que ilustrar ciertos métodos que para obtener resultados, discutimos dos teoremas del mismo.

Teorema 1. *Dos transformaciones autoadjuntas A y B sobre un espacio finito-dimensional de producto interior son conmutativas si y sólo si existe una transformación autoadjunta C y existen dos funciones real-valuadas f y g de una variable real, de manera que $A = f(C)$ y $B = g(C)$. Si existe una C tal, entonces podemos hasta escoger C en la forma $C = h(A, B)$, donde h es una adecuada función real-valuada de dos variables reales.*

Prueba. La suficiencia de la condición es clara; probamos solamente la necesidad.

Sean $A = \sum_i \alpha_i E_i$ y $B = \sum_j \beta_j F_j$ las formas espectrales de A y B; puesto que A y B conmutan, se sigue del § 79, Teorema 3, que E_i y E_j conmutan. Sea h cualquier función de dos variables reales tales que los números $h(\alpha_i, \beta_j) = \gamma_{ij}$ son todos distintos, y escribamos

$$C = h(A, B) = \sum_i \sum_j h(\alpha_i, \beta_j) E_i F_j.$$

(Está claro que h puede hasta ser escogido como polinomio, y lo mismo es cierto de las funciones f y g que vamos a describir). Sean f y g tales que $f(\gamma_{ij}) = \alpha_i$ y $g(\gamma_{ij}) = \beta_j$ para toda i y j. Se sigue que $f(C) = A$ y $g(C) = B$, y todo está probado.

Teorema 2. *Si A es una transformación normal sobre un espacio finito-dimensional unitario y si B es una transformación arbitraria, que conmuta con A, entonces B conmuta con A^*.*

Prueba. Sea $A = \sum_i \alpha_i E_i$ la forma espectral de A; entonces $A^* = \sum_i \bar{\alpha}_i E_i$. Sea f una función tal (polinomio) de una variable compleja que $f(\alpha_i) = \bar{\alpha}_i$, para toda i. Puesto que $A^* = f(A)$, se sigue la conclusión.

EJERCICIOS

1. (a) Pruébese la siguiente generalización del Teorema 2; si A_1 y A_2 son transformaciones normales (sobre un espacio finito dimensional unitario) y si $A_1B = BA_2$, entonces $A_1^*B = BA_2^*$.
(b) El Teorema 2 enuncia que la relación de conmutatividad es algunas veces transiitva: si A^* conmuta con A y si A conmuta con B, entonces A^* conmuta con B. ¿Permanece cierta esta formulación si A^* es reemplazada por una transformación arbitraria C?
2. (a) Si A conmuta con A^*A, ¿se sigue que A es normal?
(b) Si A^*A conmuta con AA^*. ¿Se sigue que A es normal?
3. (a) Una transformación lineal A es normal si y sólo si existe, un polinomio p tal que $A^* = p(A)$.
(b) Si A es normal y conmuta con B, entonces A conmuta con B^*.
(c) Si A y B son normales y conmutativos, entonces AB es normal.
4. Si A y B son normales y similares, entonces son únicamente equivalentes.
5. (a) Si A es hermitiana, si cada valor propio de A tiene multiplicidad 1, y si $AB = BA$, entonces existe un polinomio p tal que $B = p(A)$.
(b) Si A es hermitiana, entonces una condición necesaria y suficiente para que exista un polinomio p tal que $B = p(A)$ es que B conmute con cada transformación lineal que conmute con A.
6. Demuéstrese que un conjunto conmutativo de transformaciones normales sobre un espacio finito-dimensional unitario puede ser diagonalizado simultáneamente.

§85. TRANSFORMACION AUTOADJUNTA DE RANGO UNO

Hemos visto ya (§ 51, Teorema 2) que toda transformación lineal A de rango ρ es la suma de ρ transformaciones lineales de rango uno. Es fácil ver (usando el teorema espectral) que si A es autoadjunta, o positiva, entonces los sumandos pueden también tomarse autoadjuntos y positivos, respectivamente. Sabemos (§ 51, Teorema 1), lo que tiene que ser la matriz de transformación de rango uno; ¿qué más podemos decir si la transformación es autoadjunta o positiva?

Teorema 1. *Si A tiene rango uno y es autoadjunto (o positiva), entonces en cada sistema de coordenadas ortonormales la matriz (α_{ij}) de A es dada por $\alpha_{ij} = \kappa\beta_i\bar{\beta_j}$ con una real κ (o por $\alpha_{ij} = \gamma_i\bar{\gamma}_j$). Si, recíprocamente, $[A]$ tiene esta forma en algún sistema de coordenadas ortonormales, entonces A tiene rango uno y es autoadjunta (o positiva).*

Prueba. Sabemos que la matriz (α_{ij}) de una transformación A de rango uno, en cualquier sistema de coordenadas ortonormales

$\mathfrak{X} = \{x_1, \ldots, x_n\}$ es dada por $\alpha_{ij} = \beta_j\gamma_i$. Si A es autoadjunta, debemos también tener $\alpha_{ij} = \bar{\alpha}_{ji}$ de donde $\beta_i\gamma_j = \overline{\beta_j\gamma_i}$. Si $\beta_i = 0$ y $\gamma_i \neq 0$ para alguna i, entonces $\bar{\beta}_j = \beta_i\gamma_j/\bar{\gamma}_i = 0$ para toda j, de donde $A = 0$. Puesto que supusimos que el rango de A es uno (y no cero) esto es imposible. De modo semejante $\beta_i \neq 0$ y $\gamma_i = 0$ es imposible; esto es, podemos encontrar una i para la cual $\beta_i\gamma_i \neq 0$. Usando esta i, tenemos

$$\bar{\beta}_j = (\beta_{i/}\ {}_{\iota,/}\gamma_j = \kappa\gamma_j$$

con alguna constante k no cero, independiente de j. Puesto que los elementos de la diagonal $\alpha_{jj} = (Ax_j, x_j) = \beta_j\gamma_j$ de una matriz autoadjunta son reales, podemos hasta concluir que $\alpha_{ij} = \kappa\beta_i\bar{\beta}_j$ con una κ real.

Si, además, A es positiva, entonces hasta sabemos que $\kappa\beta_j\bar{\beta}_j = \alpha_{jj} = (Ax_j, x_j)$ es positiva y, en consecuencia, lo es κ. En este caso escribimos $\lambda = \sqrt{\kappa}$; la relación $\kappa\beta_i\bar{\beta}_j = (\lambda\beta_i)(\lambda\bar{\beta}_j)$ demuestra que α_{ij} es dada por $\alpha_{ij} = \gamma_i\bar{\gamma}_j$.

Es fácil ver que estas condiciones necesarias son también suficientes. Si $\alpha_{ij} = \kappa\beta_i\bar{\beta}_j$ con una κ real, entonces A es autoadjunta. Si $\alpha_{ij} = \gamma_i\bar{\gamma}_j$, y $x = \sum_i \xi_i x_i$, entonces

$$\begin{aligned}(Ax, x) &= \textstyle\sum_i \sum_j \alpha_{ij}\bar{\xi}_i\xi_j = \sum_i \sum_j \gamma_i\bar{\gamma}_j\bar{\xi}_i\xi_j \\ &= \textstyle(\sum_i \gamma_i\bar{\xi}_i)\overline{(\sum_j \gamma_j\bar{\xi}_j)} = |\sum_i \gamma_i\bar{\xi}_i|^2 \geqq 0,\end{aligned}$$

de manera que A es positiva.

Como consecuencia del Teorema 1 es muy fácil probar un notable teorema sobre matrices positivas.

TEOREMA 2. *Si A y B son transformaciones lineales positivas cuyas matrices en algún sistema ortogonal de coordenadas son (α_{ij}) y (β_{ij}), respectivamente, entonces la transformación lineal C, cuya matriz (γ_{ij}) en el mismo sistema de coordenadas es dada por $\gamma_{ij} = \alpha_{ij}\beta_{ij}$, para toda i y j, es también positiva.*

PRUEBA. Puesto que podemos escribir tanto A como B, como sumas de transformaciones positivas de rango uno, de manera que

$$\alpha_{ij} = \textstyle\sum_p \alpha_i^p\bar{\alpha}_j^p$$

y

$$\beta_{ij} = \textstyle\sum_q \beta_i^q\bar{\beta}_j^q,$$

se sigue que

$$\gamma_{ij} = \textstyle\sum_p \sum_q \alpha_i^p\beta_i^q\overline{(\alpha_j^p\beta_j^q)}.$$

(Los supraíndices no son aquí exponentes). Puesto que una suma de matrices positivas es positiva, será suficiente con probar que, para p y q fijas, la matriz $((\alpha_i^p \beta_i^q)\overline{(\alpha_j^p \beta_j^q)})$ es positiva y esto se sigue del Teorema 1.

Incidentalmente, la prueba demuestra que el Teorema 2 permanece válido si se reemplaza "positivo" por "autoadjunto" en ambas hipótesis y conclusión; en la mayoría de las aplicaciones, sin embargo, es sólo la versión enunciada la que es útil. La matriz (γ_{ij}) descrita en el Teorema 2 se llama el *producto de Hadamard* de (α_{ij}) y (β_{ij}).

EJERCICIOS

1. Supóngase que $\mathfrak{U}$ y $\mathfrak{V}$ son espacios finito-dimensionales de productos interiores (ambos reales o ambos complejos).

(a) Hay un producto interior único sobre el espacio vectorial de todas las formas bilineales sobre $\mathfrak{U} \oplus \mathfrak{V}$ tales que $w_1(x, y) = (x, x_1)(y, y_1)$ y $w_2(x, y) = (x, x_2)(y, y_2)$, entonces $(w_1, w_2) = (x_2, x_1)(y_2, y_1)$.

(b) Hay un producto interior único sobre el producto tensorial $\mathfrak{U} \otimes \mathfrak{V}$, tal que si $z_1 = x_1 \otimes y_1$ y $z_2 = x_2 \otimes y_2$, entonces $(z_1, z_2) = (x_1, x_2)(y_1, y_2)$.

(c) Si $\{x_i\}$ y $\{y_p\}$ son bases ortonormales en $\mathfrak{U}$ y $\mathfrak{V}$, respectivamente, entonces los vectores $x_i \otimes y_p$ forman una base ortonormal en $\mathfrak{U} \otimes \mathfrak{V}$.

2. ¿Es necesariamente hermitiano el producto tensorial de dos transformaciones hermitianas? ¿Qué sobre transformaciones unitarias? ¿Qué sobre transformaciones normales?

CAPITULO 4

ANALISIS

§86. CONVERGENCIA DE VECTORES

Esencialmente la única forma en que explotamos, hasta ahora, la existencia de un producto interior en un espacio de producto interior, fue introduciendo la noción de transformación normal, juntamente con ciertos importantes casos especiales de la misma. Un círculo de ideas mucho más obvio es el estudio de los problemas de convergencia que se suscitan en un espacio de producto interior.

Veámos qué podríamos dar a entender con la aserción de que una secuencia (x_n) de vectores en $\mathcal{V}$ converge a un vector x en $\mathcal{V}$. Hay dos posibilidades que se sugieren por sí mismas:

(i) $\qquad ||x_n - x|| \to 0$ a medida que $n \to \infty$;

(ii) $(x_n - x, y) \to 0$, a medida que $n \to \infty$, para cada y fija de $\mathcal{V}$.

Si (i) es verdadera, entonces tenemos, para cada y,

$$|(x_n - x, y)| \leq \| x_n - x \| \cdot \| y \| \to 0,$$

de manera que (ii) es verdadera. En un espacio finito-dimensional la implicación recíproca es válida: (ii)$\Rightarrow$(i). Para probar esto, sea $\{z_1, \ldots, z_N\}$ una base ortonormal en $\mathcal{V}$. (A menudo en este capítulo escribiremos N para la dimensión de un espacio vectorial finito-dimensional a fin de reservar n para la variable ficticia en procesos limitadores). Si suponemos (ii), entonces $(x_n - x, z_i) \to 0$ para cada $i = 1, \ldots, N$. Puesto que (§ 63, Teorema 2)

$$\| x_n - x \|^2 = \textstyle\sum_i |(x_n - x, z_i)|^2,$$

se sigue que $|| x_n - x || \to 0$, como tenía que probarse.

Con respecto de la convergencia de vectores (en cualquiera de dos sentidos equivalentes). usaremos sin prueba los siguientes hechos. (Todos estos hechos son fáciles consecuencias de nuestras definiciones y de las propiedades de la convergencia en el dominio usual de los números complejos; suponemos que el lector, tiene alguna familiaridad con estos conceptos). La expresión $\alpha x + \beta y$ define una función continua de todos sus argumentos simultáneamente; esto es, si (α_n) y (β_n) son secuencias de números y (x_n) y (y_n) son secuencias de vectores, entonces $\alpha_n \to \alpha$, $\beta_n \to \beta$, $x_n \to x$, y $y_n \to y$ implican que a $\alpha_n x_n + \beta_n y_n \to \alpha x + \beta y$. Si $\{z_i\}$ es una base ortonormal en $\mathcal{V}$, y si $x_n = \sum_i \alpha_{in} z_i$ y $x = \sum_i \alpha_i z_i$, entonces una condición necesaria y suficiente para que $x_n \to x$ es que $\alpha_{in} \to \alpha_i$ (a medida que $n \to \infty$) para cada $i = 1, \ldots, N$. (Así, el concepto de convergencia aquí definido coincide con el usual en el espacio N-dimensional de coordenadas reales o complejas). Finalmente, supondremos conocido el hecho de que un espacio finito-dimensional de producto interior con la métrica definida por la norma es completo; esto es, si (x_n) es una secuencia de vectores para los cuales $\| x_n - x_m \| \to 0$, a medida que n, $m \to \infty$, entonces hay un vector (único) x, tal que $x_n \to x$, a medida que $n \to \infty$.

§ 87. NORMA

Las propiedades métricas de los vectores tienen ciertas implicaciones importantes para las propiedades métricas de transformaciones lineales, que ahora comenzamos a estudiar.

Definición. Una transformación lineal A sobre un espacio $\mathcal{V}$ de producto interior es *ligada* si existe una constante K tal que $\| Ax \| \leqq K \| x \|$ para todo vector x de $\mathcal{V}$. La ligadura máxima inferior de todas las constantes K con esta propiedad es llamada la *norma* (o *ligadura*) de A y es denotada por $\| A \|$.

Claramente si A es ligada, entonces $\| Ax \| \leqq \| A \| \cdot \| x \|$ para toda x. Como ejemplos podemos considerar los casos en que A es una proyección perpendicular (no cero) o una isometría; el Teorema 1 del § 75 y el teorema del § 73, respectivamente, implican que en ambos casos $\| A \| = 1$. Consideraciones de los vectores definidos por $x_n(t) = t^n$ en $\mathcal{P}$ muestra que la transformación de diferenciación no está ligada.

A causa de que en lo que sigue tendremos ocasión de considerar bastantes ligaduras superiores e inferiores similares a $\| A \|$, introduciremos una notación conveniente. Si P es cualquier propiedad posible de números reales t, denotaremos el conjunto de todos los números reales t que poseen la propiedad P por el símbolo $\{t\colon P\}$, y denotaremos la máxima ligadura inferior y la mínima ligadura superior con *inf* (por infimum) y *sup* (por supremum), respectivamente. En esta notación tenemos, por ejemplo,

$$\| A \| = \inf \{K : \| Ax \| \leq K \| x \| \text{ for all } x\}.$$

La noción de ligadura esta íntimamente relacionada con la noción de continuidad. Si A está limitada y si ϵ es cualquier número positivo, escribiendo $\delta = \dfrac{\epsilon}{\| A \|}$ podemos asegurarnos de que $\| x - y \| < \delta$ implica que

$$\| Ax - Ay \| = \| A(x - y) \| \leq \| A \| \cdot \| x - y \| < \epsilon;$$

en otras palabras, la ligadura significa continuidad (uniforme). (En esta prueba supusimos tácitamente que $\| A \| \neq 0$; el otro caso es trivial). En vista de este hecho, el siguiente resultado es bienvenido.

Teorema. *Toda transformación lineal sobre un espacio finito-dimensional de producto interior es ligada.*

Prueba. Supóngase que A es una transformación lineal en $\mathfrak{V}$; sea $\{x_1, \ldots x_N\}$ una base ortonormal en $\mathfrak{V}$ y escribamos

$$K_0 = \text{máx}\,\{\| Ax_1 \|, \ldots, \| Ax_N \|\}.$$

Puesto que un vector arbitrario x puede ser escrito en la forma $x = \sum_i (x, x_i)x_i$, obtenemos, aplicando la desigualdad de Schwarz y recordando que $\| x_i \| = 1$,

$$\begin{aligned} \| Ax \| &= \| A(\textstyle\sum_i (x, x_i)x_i) \| \\ &= \| \textstyle\sum_i (x, x_i)Ax_i \| \leq \textstyle\sum_i |(x, x_i)| \cdot \| Ax_i \| \\ &\leq \textstyle\sum_i \| x \| \cdot \| x_i \| \cdot \| Ax_i \| \leq K_0 \textstyle\sum_i \| x \| \\ &= NK_0 \| x \|. \end{aligned}$$

En otras palabras, $K = NK_0$ es una ligadura de A, y la prueba es completa.

No es un accidente que la dimensión N de $\mathfrak{V}$ entre en nuestra evaluación; ya hemos visto que el teorema no es verdadero en espacios infinito-dimensionales.

EJERCICIOS

1. (a) Pruébese que el producto interior es una función continua (y en consecuencia lo es también la norma); esto es, si $x_n \to x$ y $y_n \to y$, entonces $(x_n, y_n) \to (x, y)$.

(b) ¿Es continua toda función lineal? ¿Qué sobre las formas multilineales?

2. Se dice que está ligada desde abajo una transformación lineal A sobre un espacio de producto interior, si existe una constante (estrictamente) positiva K tal que $\| Ax \| \geqq K \| x \|$ para toda x. Pruébese que (sobre un espacio finito-dimensional) A está ligada desde abajo si y sólo si es invertible.

3. Si una transformación lineal sobre un espacio de producto interior (no necesariamente finito-dimensional) es continua en un punto, en consecuencia es ligada (y consecuentemente continua sobre todo el espacio).

4. Para cada entero positivo n, constrúyase una proyección E_n (no una proyección perpendicular) tal que $\| E_n \| \geqq n$.

5. (a) Si U es una isometría parcial distinta de 0, entonces $\| U \| = 1$.

(b) Si U es una isometría, entonces $\| UA \| = \| AU \| = \| A \|$ para toda transformación lineal A.

6. Si E y F son proyecciones perpendiculares, con alcances $\mathfrak{M}$ y $\mathfrak{N}$ respectivamente, y si $\| E - F \| < 1$, entonces $\dim \mathfrak{M} = \dim \mathfrak{N}$.

7. (a) Si A es normal, entonces $\| A^n \| = \| A \|^n$ para cada entero positivo n.

(b) Si A es una transformación lineal sobre un espacio 2-dimensional unitario, y si $\| A^2 \| = \| A \|^2$, entonces A es normal.

(c). Es verdadera la conclusión de (b) para transformaciones sobre un espacio 3-dimensional.

§ 88. EXPRESIONES PARA LA NORMA

Para facilitar el trabajo con la norma de una transformación, consideramos las siguientes cuatro expresiones:

$$p = \sup \{\| Ax \| / \| x \| : x \neq 0\},$$

$$q = \sup \{\| Ax \| : \| x \| = 1\},$$

$$r = \sup \{|(Ax, y)| / \| x \| \cdot \| y \| : x \neq 0, y \neq 0\},$$

$$s = \sup \{|(Ax, y)| : \| x \| = \| y \| = 1\}.$$

De acuerdo con nuestra definición de la notación de llave, la expresión $\{\| Ax \| : \| x \| = 1$, por ejemplo, significa el conjunto de todos los números reales de la forma $\| Ax \|$, considerada para todas las x para las cuales $\| x \| = 1$.

Puesto que $\| Ax \| \leqq K \| x \|$ es trivialmente verdadera con cualquier K si $x = 0$, la definición de supremum implica que $p = \| A \|$; probaremos que, de hecho $p = q = r = s = \| A \|$. Puesto que el supremum en la expresión de q se extiende sobre un subconjunto del correspondiente conjunto para p (esto es, si $\| x \| = 1$, entonces $\| Ax \| / \| x \| = \| Ax \|$), vemos que $q \leqq p$; un argumento semejante demuestra que $s \leqq r$.

Para cualquier $x \neq 0$, consideramos $y = \dfrac{x}{\|x\|}$ (de manera que $\|y\| = 1$); tenemos $\|Ax\|/\|x\| = \|Ay\|$. En otras palabras, todo número del conjunto cuyo supremum es p ocurre también en el correspondiente conjunto para q; se sigue que $p \leqq q$, y consecuentemente que $p = q = \|A\|$.

De modo semejante, si $x \neq 0$ y $y \neq 0$, consideramos $x' = x/\|x\|$ y $y' = y/\|y\|$; tenemos

$$|(Ax, y)|/\|x\| \cdot \|y\| = |(Ax', y')|,$$

y de aquí que, por el argumento que se acaba de usar, $r \leqq s$, de manera que $r = s$.

Para consolidar nuestra posición, notamos que hasta ahora hemos probado que

$$p = q = \|A\| \quad \text{and} \quad r = s.$$

Puesto que

$$\frac{|(Ax, y)|}{\|x\| \cdot \|y\|} \leqq \frac{\|Ax\| \cdot \|y\|}{\|x\| \cdot \|y\|} = \frac{\|Ax\|}{\|x\|},$$

se sigue que $r \leqq p$; completaremos la prueba demostrando que $p \leqq r$.

Para este propósito consideramos cualquier vector x para el cual $Ax \neq 0$ (de manera que $x \neq 0$); para una tal x escribimos $y = Ax$ y tenemos

$$\|Ax\|/\|x\| = |(Ax, y)|/\|x\| \cdot \|y\|.$$

En otras palabras, probamos que todo número que ocurre en el conjunto que define a p, y es diferente de cero, ocurre también en el conjunto del cual r es supremum; esto claramente implica el resultado deseado.

La función numérica de una transformación A dada por $\|A\|$ satisface las siguientes cuatro condiciones:

(1) $$\|A + B\| \leqq \|A\| + \|B\|,$$

(2) $$\|AB\| \leqq \|A\| \cdot \|B\|,$$

(3) $$\|\alpha A\| = |\alpha| \cdot \|A\|,$$

(4) $$\|A^*\| = \|A\|.$$

La prueba de las primeras tres de éstas es inmediata de la definición de la norma de una transformación; para la prueba de (4) usamos la ecuación $\|A\| = r$, como sigue. Puesto que

$$|(Ax, y)| = |(x, A^*y)| \leqq \| x \| \cdot \| A^*y \|$$
$$\leqq \| A^* \| \cdot \| x \| \cdot \| y \|,$$

vemos que $\| A \| \leqq \| A^* \|$; reemplazando A por A^* y A^* por $A^{**} = A$, obtenemos la desigualdad recíproca.

EJERCICIOS

1. Si B es invertible, entonces $\|AB\| \geqq \| A \| / \| B^{-1} \|$ para toda A.
2. ¿Es verdad para cada transformación lineal A que $\| A^*A \| = AA^* \|$?
3. (a) Si A es hermitiana y si $\alpha \geqq 0$, entonces una condición necesaria y suficiente para que $\| A \| \leqq \alpha$ es que $-\alpha \leqq A \leqq \alpha$.
   (b) Si A es hermitiana, si $\alpha \leqq A \leqq \beta$, y si p es un polinomio tal que $p(t) \geqq 0$, cada vez que $\alpha \leqq t \leqq \beta$, entonces $p(A) \geqq 0$.
   (c) Si A es hermitiana, si $\alpha \leqq A \leqq \beta$, y si p es un polinomio tal que $p(t) \neq 0$, cada vez que $\alpha \leqq t \leqq \beta$, entonces $p(A)$ es invertible.

§89. LIMITES DE UNA TRANSFORMACION AUTOADJUNTA

Como de costumbre podemos decir un poco más del caso especial de una transformación autoadjunta que en el caso general. Consideramos, para cualquier transformación autoadjunta A, los conjuntos de números reales

$$\Phi = \{(Ax, x)/\| x \|^2 : x \neq 0\}$$

y

$$\Psi = \{(Ax, x) : \| x \| = 1\}.$$

Esta claro que $\Psi \subset \Phi$. Si para toda $x \neq 0$, escribimos $y = x/\| x\|$, entonces $\| y \| = 1$ y $(Ax, x)/\| x \|^2 = (Ay, y)$, de manera que todo número de Φ ocurre también en Ψ y, consecuentemente $\Phi = \Psi$. Escribimos

$$\alpha = \inf \Phi = \inf \Psi,$$
$$\beta = \sup \Phi = \sup \Psi,$$

y decimos que α es el *límite inferior* y β es el *límite superior* de la transformación autoadjunta A. Si recordamos la definición de transformación positiva, vemos que α es el número real máximo para cual $A - \alpha \geqq 0$ y β es el número real mínimo para el cual $\beta - A \geqq 0$. Con relación a estos números aseguramos que

$$\gamma = \text{máx}\,\{|\alpha|, |\beta|\} = \|A\|.$$

La mitad de la prueba es fácil. Puesto que

$$|(Ax, x)| \leqq \|Ax\| \cdot \|x\| \leqq \|A\| \cdot \|x\|^2,$$

está claro que tanto $|\alpha|$ como $|\beta|$ están dominados por $\|A\|$. Para probar la desigualdad recíproca, observamos que el carácter positivo de las dos transformaciones lineales $\gamma - A$ y $\gamma + A$ implica que ambas

$$(\gamma + A)^*(\gamma - A)(\gamma + A) = (\gamma + A)(\gamma - A)(\gamma + A)$$

y

$$(\gamma - A)^*(\gamma + A)(\gamma - A) = (\gamma - A)(\gamma + A)(\gamma - A)$$

son positivas y, en consecuencia, lo es también su suma $2\gamma(\gamma^2 - A^2)$. Puesto que $\gamma = 0$ implica $\|A\| = 0$, la aserción es trivial en este caso; en cualquier otro caso podemos dividir entre 2 y obtenemos el resultado de que $\gamma^2 - A^2 \geqq 0$. En otras palabras

$$\gamma^2\|x\|^2 = \gamma^2(x, x) \geqq (A^2x, x) = \|Ax\|^2,$$

de donde $\gamma \geqq \|A\|$, y la prueba es completa.

Llamamos la atención del lector al hecho de que el cómputo en el cuerpo principal de esta prueba podría haber sido evitado enteramente. Puesto que tanto $\gamma - A$ como $\gamma + A$ son positivas, y puesto que conmutan, podemos concluir inmediatamente (§82) que su producto $\gamma^2 - A^2$ es positivo. Presentamos el método indirecto consecuentemente con el principio de que, con vista a las generalizaciones de la teoría, debería uno evitar el uso del teorema espectral simpre que sea posible. Nuestra prueba del hecho de que la positividad y la conmutatividad de A y B implican la positividad de AB estaba basada sobre la existencia de raíces cuadradas para transformaciones positivas. Este hecho, de seguro, puede obtenerse por métodos llamados "elementales", esto es, método que no usan el teorema espectral, pero hasta la prueba simple más elemental implica complicaciones que son puramente técnicas, y, para nuestros propósitos no particularmente útiles.

§90. PRINCIPIO MINIMAX

Un hecho muy elegante y útil, relativo a las transformaciones autoadjuntas es el siguiente *principio minimax.*

TEOREMA. *Sea A una transformación autoadjunta sobre un espacio n-dimensional de producto interior $\mathcal{V}$, y sean $\lambda_1, \ldots,$*

λ_n *los valores propios (no necesariamente distintos) de A, con la notación escogida de tal modo que* $\lambda_1 \geqq \lambda_2 \geqq \ldots \geqq \lambda_n$. *Si, para cada subespacio,* $\mathfrak{M}$ *de* $\mathcal{V}$,

$$\mu(\mathfrak{M}) = \sup \{(Ax, x): x \text{ in } \mathfrak{M}, \| x \| = 1\},$$

y si, para $k = 1, \ldots, n$,

$$\mu_k = \inf \{\mu(\mathfrak{M}): \dim \mathfrak{M} = n - k + 1\},$$

entonces $\mu_k = \lambda_k$ *para* $k = 1, \ldots, n$.

Prueba. Sean $\{x_1, \ldots, x_n\}$ una base ortonormal en $\mathcal{V}$ para la cual $Ax_i = \lambda_i x_i$, $i = 1, \ldots, n$, (§79); sea $\mathfrak{M}_k$ el subespacio sobretendido por $x_1, \ldots, x_k$, para $k = 1, \ldots, n$. Puesto que la dimensión de $\mathfrak{M}_k$, el subespacio $\mathfrak{M}_k$ no puede ser disjunto de cualquier subespacio $(n - k + 1)$-dimensional en $\mathfrak{M}$; si $\mathcal{V}$ es uno de esos espacios, podemos encontrar un vector x que pertenece tanto a $\mathfrak{M}_k$ como a $\mathfrak{M}$ y tal que $\| x \| = 1$. Para esta $x = \sum_{i=1}^{k} \xi_i x_i$ tenemos

$$(Ax, x) = \sum_{i=1}^{k} \lambda_i |\xi_i|^2 \geqq \lambda_k \sum_{i=1}^{k} |\xi_i|^2$$
$$= \lambda_k \| x \|^2 = \lambda_k,$$

de manera que $\mu(\mathfrak{M}) \geqq \lambda_k$.

Si, por otra parte, consideramos el subespacio $\mathfrak{M}_0$ particular $(n - k + 1)$-dimensional, sobretendido por $x_k, x_{k+1}, \ldots, x_n$, entonces para cada $x = \sum_{i=k}^{n} \xi_i x_i$ de este subespacio, tenemos, (suponiendo $\| x \| = 1$)

$$(Ax, x) = \sum_{i=k}^{n} \lambda_i |\xi_i|^2 \leqq \lambda_k \sum_{i=k}^{n} |\xi_i|^2$$
$$= \lambda_k \| x \|^2 = \lambda_k,$$

de manera que $\mu(\mathfrak{M}_0) \leqq \lambda_k$.

En otras palabras, como $\mathfrak{M}$ corre sobre todos los subespacios $(n - k + 1)$-dimensionales, $\mu(\mathfrak{M})$ es siempre $\geqq \lambda_k$, y es cuando menos una vez $\leqq \lambda_k$; esto demuestra que $\mu_k = \lambda_k$, como había de probarse.

En particular, para $k = 1$, vemos (usando el §89) que si A es autoadjunta, entonces $\| A \|$ es igual al máximo de los valores absolutos de los valores propios de A.

EJERCICIOS

1. Si λ es un valor propio de una transformación lineal A sobre un espacio finito-dimensional de producto interior, entonces $|\lambda| \leqq \| A \|$.

2. Si A y B son transformaciones lineales sobre un espacio finito-dimensional unitario, y si $C = AB - BA$, entonces $\|1 - C\| \geq 1$. (Sugestión: considérense los valores propios de C).

3. Si A y B son transformaciones lineales sobre un espacio finito-dimensionales unitario, si $C = AB - BA$, y si C conmuta con A, entonces C no es invertible. (Sugestión: si C es invertible, entonces $2\ \|B\| \cdot \|A\| \cdot \|A^{k-1}\| \geq k\,\|A^{k-1}\|/\|C^{-1}\|$).

4. (a) Si A es una transformación lineal normal sobre un espacio finito-dimensional unitario, entonces $\|A\|$ es igual al máximo del valor absoluto de los valores propios de A.

(b) ¿Permanece verdadera la conclusión de (a) si se omite la hipótesis de normalidad?

5. El *radio espectral* de una transformación lineal A sobre un espacio finito-dimensional unitario, denotado por $r(A)$, es el máximo de los valores absolutos de los valores propios de A.

(a) Si $f(\lambda) = ((1 - \lambda A)^{-1}x, y)$, entonces f es una iunción analítica de λ en la región determinada por $|\lambda| < \dfrac{1}{r(A)}$ (para cada x y y fijas).

(b) Existe una constante K fija tal que $|\lambda|^n\|A^n\| \leq K$ siempre que $|\lambda| < \dfrac{1}{r(A)}$ y $n = 0, 1, 2, \ldots$. (Sugestión: para cada x y y existe una constante K tal que $|\lambda^n(A^n x, y)| \leq K$ para toda n).

(c) $\lim \sup_n \|A^n\|^{1/n} \leq r(A)$.

(d) $(r(A))_n \leq r(A^n)$, $n = 0, 1, 2, \ldots$.

(e) $r(A) = \lim_n \|A^n\|^{1/n}$.

6. Si A es una transformación lineal sobre un espacio finito-dimensional unitario, entonces una condición necesaria y suficiente para que $r(A) = \|A\|$, es que $\|A^n\| = \|A\|^n$ para $n = 0, 1, 2, \cdots$.

7. (a) Si A es una transformación lineal positiva sobre un espacio finito-dimensional de producto inferior, y si AB es autoadjunto, entonces

$$|(ABx, x)| \leq \|B\| \cdot (Ax, x)$$

para cada vector x.

(b) ¿Permanece verdadera la conclusión de (a) si $\|B\|$ es reemplazada por $r(B)$?

§91. CONVERGENCIA DE TRANSFORMADORES LINEALES

Retornamos a hora a la consideración de problemas de convergencia. Hay tres sentidos obvios en que podemos tratar de definir la convergencia de la secuencia (A_n) de transformaciones lineales a una transformación lineal fija A.

(i) $\|A_n - A\| \to 0$ a medida que $n \to \infty$.

(ii) $\|A_n x - Ax\| \to 0$, a medida que $n \to \infty$ para cada x fija.

(iii) $|(A_n x, y) - (Ax, y)| \to 0$, a medida que $n \to \infty$ para cada x y y fijas.

Si (i) es verdadera, entonces para cada x,

$$\| A_n x - Ax \| = \| (A_n - A)x \| \leqq \| A_n - A \| \cdot \| x \| \to 0,$$

de manera que (i) $\Rightarrow$ (ii). Hemos visto ya (§86) que (ii) $\Rightarrow$ (iii) y que en espacios finito-dimensionales (iii) $\Rightarrow$ (ii). Es aun verdad que en espacios finito-dimensional (ii) $\Rightarrow$ (i), de manera que las tres condiciones son equivalentes. Para probar esto, sea $\{x_1, \ldots, x_N\}$ una base ortonormal en $\mathcal{V}$. Si suponemos que (ii) es válido, entonces, para cada $\varepsilon > 0$, podemos encontrar $n_0 = n_0(\varepsilon)$ tal que $\| A_n x_i - Ax_i \| < \varepsilon$ para $n \geqq n_0$ y para $i = 1, \ldots, N$. Se sigue que para una $x = \sum_i (x, x_i)x_i$, arbitraria, tenemos

$$\begin{aligned} \| (A_n - A)x \| &= \| \textstyle\sum_i (x, x_i)(A_n - A)x_i \| \\ &\leqq \textstyle\sum_i \| x \| \cdot \| (A_n - A)x_i \| \leqq \varepsilon N \| x \|, \end{aligned}$$

y esto implica (i).

Es fácil también probar que si la norma se usa para definir una distancia para transformaciones, entonces es completo el espacio métrico resultante, esto es, si $\| A_n - A_m \| \to 0$, a medida que n, $m \to \infty$, entonces hay una A tal que $\| A_n - A \| \to 0$. La prueba de este hecho se reduce al hecho correspondiente para vectores. Si $\| A_n - A_m \| \to 0$, entonces $\| A_n x - A_m x \| \to 0$ para cada x, de manera que podemos encontrar un vector correspondiente a x, que podemos denotar por Ax, por ejemplo, tal que $\| A_n x - Ax \| \to 0$. Está claro que la correspondencia de x a Ax es dada por una transformación lineal A; la relación de implicación (ii) $\Rightarrow$ (i) anteriormente probada, completa la prueba.

Ahora que sabemos lo que significa la convergencia para transformaciones lineales, nos corresponde examinar alguna funciones simples de estas transformaciones a fin de verificar su continuidad. Afirmamos que $\| A \|$, $\| Ax \|$, (Ax, y), Ax, $A + B$, αA, AB, y A^* todas definen funciones continuas de todos sus argumentos simultáneamente. (Obsérvese que las tres primeras son funciones numérico-valuadas, la siguiente es vector-valuada y las cuatro últimas son transformaciones valuadas). Las pruebas de estos enunciados son todas completamente fáciles y similares entre sí; para ilustrar las ideas, discutimos $\| A \|$, Ax, y A^*.

(1) Si $A_n \to A$ esto es, $\| A_n - A \| \to 0$, entonces, puesto que las relaciones

$$\| A_n \| \leqq \| A_n - A \| + \| A \|,$$

y

$$\| A \| \leqq \| A - A_n \| + \| A_n \|$$

implican que

$$| \| A_n \| - \| A \| | \leqq \| A_n - A \|,$$

vemos que $|| A_n || \rightarrow || A ||$.

(2) Si $A_n \rightarrow A$ y $x_n \rightarrow x$, entonces,

$$\| A_n x_n - Ax \| \leqq \| A_n x_n - A x_n \| + \| A x_n - Ax \| \rightarrow 0,$$

de manera que $A_n x_n \rightarrow Ax$.

(3) Si $A_n \rightarrow A$, entonces, para cada x y y,

$$(A_n^* x, y) = (x, A_n y) = \overline{(A_n y, x)} \rightarrow \overline{(Ay, x)}$$

$$= \overline{(y, A^* x)} = (A^* x, y),$$

de donde $A_n^* \rightarrow A^*$.

EJERCICIOS

1. Una secuencia (A_n) de transformaciones lineales converge a una transformación lineal A si, y sólo para todo sistema de coordenadas, converge cada entrada en la matriz de A_n, a medida que $n \rightarrow \infty$, hacia la correspondiente entrada en la matriz de A.
2. Para cada transformación lineal A existe una secuencia (A_n) de transformaciones lineales invertibles, tal que $A_n \rightarrow A$.
3. Si E y F son proyecciones perpendiculares, entonces (EFE^n) converge, a medida que $n \rightarrow \infty$, a la proyección cuyo alcance es la intersección de los alcances de E y F.
4. Si A es una transformación lineal sobre un espacio finito-dimensional unitario, entonces una condición necesaria y suficiente para que $A^n \rightarrow 0$ es que todos los valores propios de A sean (estrictamente) menores que 1 en valor absoluto.
5. Pruébese que si A es la matriz n por n entonces A^k converge, a medida

$$\begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \cdot & \cdot & \cdot & \cdots & \cdot \\ 0 & 0 & 0 & \cdots & 1 \\ \frac{1}{n} & \frac{1}{n} & \frac{1}{n} & \cdots & \frac{1}{n} \end{bmatrix}$$

que $k \to \infty$, hacia una proyección cuyo rango es unidimensional; encuéntrese el rango.

6. Pruébese que *det* y *tr* son continuos.

§92. TEOREMA ERGODICO

Ha terminado el trabajo de rutina: pasamos a ilustrar la teoría general, considerando algunos problemas de convergencia muy especiales pero completamente importantes.

TEOREMA. *Si U es una isometría sobre un espacio finito-dimensional de producto interior, y si $\mathfrak{M}$ es el subespacio de todas las soluciones de $Ux = x$, entonces la secuencia definida por*

$$V_n = \frac{1}{n}(1 + U + \cdots + U^{n-1})$$

converge a medida que $n \to \infty$, hacia la proyección perpendicular $E = P_{\mathfrak{M}}$.

PRUEBA. Sea $\mathfrak{N}$ el alcance de la transformación lineal $1 - U$. Si $x = y - Uy$ está en $\mathfrak{N}$, entonces

$$V_n x = \frac{1}{n}(y - Uy + Uy - U^2y + \cdots + U^{n-1}y - U^n y)$$

$$= \frac{1}{n}(y - U^n y),$$

de manera que

$$\| V_n x \| = \frac{1}{n}\| y - U^n y \| \leq \frac{1}{n}(\| y \| + \| U^n y \|)$$

$$= \frac{2}{n}\| y \|.$$

Esto implica que $V_n x$ converge hacia cero cuando x está en $\mathfrak{N}$. Por otra parte, si x está en $\mathfrak{M}$, esto es, $Ux = x$, entonces $V_n x = x$, de manera en este caso $V_n x$ ciertamente converge hacia x.

Completaremos la prueba demostrando que $\mathfrak{N}^{\perp} = \mathfrak{M}$. (Esto implicará que todo vector es una suma de dos vectores, para los cuales (V_n) converge, de manera que (V_n) converge en todas partes). Lo que hemos ya probado sobre el límite de (V_n) en $\mathfrak{M}$ y en $\mathfrak{N}$ demuestra que $(V_n x)$ converge siempre hacia la proyección de x en $\mathfrak{M}$.).

Para demostrar que $\mathfrak{N}^{\perp} = \mathfrak{M}$, observamos que x está en el complemento ortogonal de $\mathfrak{N}$ si y sólo si $(x, y - Uy) = 0$ para toda y. Esto, a su vez, implica que

$$0 = (x, y - Uy) = (x, y) - (x, Uy) = (x, y) - (U^{*}x, y)$$
$$= (x - U^{*}x, y),$$

esto es, que $x - U^{*}x = x - U^{-1}x$ es ortogonal a cada vector y, de manera que $x - U^{-1}x = 0$, $x = U^{-1}x$, o $Ux = x$. La lectura del último cálculo de derecha a izquierda, demuestra que esta condición necesaria es también suficiente; necesitamos solamente recordar la definición de $\mathfrak{M}$ para ver que $\mathfrak{M} = \mathfrak{N}^{\perp}$.

Esta prueba tan ingeniosa, que es válida sólo con muy ligeras modificaciones en la mayor parte de los casos infinito-dimensionales, es debida a F. Riesz.

§93. SERIES DE POTENCIAS

Consideramos a continuación la llamada serie de Neuman $\sum_{n=0}^{\infty} A^{n}$, donde A es una transformación lineal con norma < 1 sobre un espacio vectorial finito-dimensional.
Si escribimos

$$S_p = \sum_{n=0}^{p} A^{n},$$

entonces

$$(1) \qquad (1 - A)S_p = S_p - AS_p = 1 - A^{p+1}.$$

Para probar que S_p tiene un límite, a medida que $p \to \infty$, consideramos (para dos índices cualesquiera p y q, con $p > q$)

$$\| S_p - S_q \| \leq \sum_{n=q+1}^{p} \| A^{n} \| \leq \sum_{n=q+1}^{p} \| A \|^{n}.$$

Puesto que $\| A \| < 1$, la última cantidad escrita se aproxima a cero a medida que p, $q \to \infty$; se sigue que S_p tiene un límite S, a media que $p \to \infty$. Para evaluar el límite, observamos que $1 - A$ es invertible. (Prueba: $(1 - A)x = 0$ implica que $Ax = x$, y, si $x \neq 0$, esto implica que $\| Ax \| = \| x \| > \| A \| \cdot \|x\|$, una contradicción). De aquí que podamos escribir (1) en la forma

$$(2) \qquad S_p = (1 - A^{p+1})(1 - A)^{-1} = (1 - A)^{-1}(1 - A^{p+1});$$

puesto que $A^{p+1} \to 0$, a medida que $p \to \infty$, se sigue que $S = (1 - A)^{-1}$.

Como otro ejemplo de una serie infinita de transformaciones, consideremos la serie exponencial. Para una transformación lineal arbitraria A (no necesariamente con $\|A\| < 1$) escribimos

$$S_p = \sum_{n=0}^{p} \frac{1}{n!} A^n.$$

Puesto que tenemos

$$\| S_p - S_q \| \leq \sum_{n=q+1}^{p} \frac{1}{n!} \| A \|^n,$$

y puesto que el lado derecho de esta desigualdad, siendo parte de la serie de potencias para $\exp \| A \| = e^{\|A\|}$, converge hacia 0, a medida que $p, q \to \infty$, vemos que hay una transformación lineal S tal que $S_p \to S$. Escribimos $S = \exp A$; solamente mencionaremos algunas de las propiedades elementales de esta función de A.

La consideración de las formas triangulares de A y de S_p demuestra que los valores propios de $\exp A$, justamente con sus multiplicidades algebraicas, son iguales a las exponenciales de los valores propios de A. (Este argumento, así como algunos de los que siguen, se aplica directamente sólo al caso complejo; el caso real tiene que ser deducido vía complejificación). De la consideración de la forma triangular se sigue también que el determinante de $\exp A$, esto es $\prod_{i=1}^{N} \exp \lambda_i$, donde $\lambda_i, \ldots, \lambda_N$ son los valores propios de A (no necesariamente distintos), es igual que $\exp (\lambda_1 + \ldots + \lambda_N) = \exp (\operatorname{tr} A)$. Puesto que $\exp \zeta \neq 0$, esto demuestra, incidentalmente, que $\exp A$ es siempre invertible.

Considerada como función de una transformación lineal, exponencial retiene muchas de las simples propiedades de la función ordinaria numérica exponencial. Tomemos, por ejemplo, cualesquiera dos transformaciones lineales *conmutativas* A y B. Puesto que $\exp (A + B) - \exp A \exp B$ es el límite (a medida que $p \to \infty$) de la expresión

$$\begin{aligned} &\sum_{n=0}^{p} \frac{1}{n!} (A + B)^n - \sum_{m=0}^{p} \frac{1}{m!} A^m \cdot \sum_{k=0}^{p} \frac{1}{k!} B^k \\ &\qquad = \sum_{n=0}^{p} \frac{1}{n!} \sum_{j=0}^{n} \binom{n}{j} A^j B^{n-j} - \sum_{m=0}^{p} \sum_{k=0}^{p} \frac{1}{m!k!} A^m B^k, \end{aligned}$$

habremos probado la regla de multiplicación para exponenciales, cuando hayamos probado que esta expresión converge hacia cero. $\left(\text{Aquí } \binom{n}{j} \text{ representa el coeficiente combinatorio } \frac{n!}{j!(n-j)!}.\right)$ Una fácil verificación resulta en el hecho de que $k + m \leqq p$ el producto $A^m B^k$ ocurre en ambos términos de la última expresión escrita, con coeficientes que difieren en signo únicamente. Los términos que no se cancelan están todos en el substraendo y juntos son iguales a

$$\sum_m \sum_k \frac{1}{m!k!} A^m B^k,$$

extendiéndose la sumo sobre los valores de m y k que son $\leqq p$ y para los cuales $m + k > p$. Puesto que $m + k > p$ implica que cuando menos uno de los dos enteros m y k es mayor que el entero parte de $\frac{p}{2}$ (en símbolos $[\frac{p}{2}]$), la norma de este resto es dominada por

$$\begin{aligned} &\sum_{m=0}^{\infty} \sum_{k=[\frac{p}{2}]}^{\infty} \frac{1}{m!k!} \| A \|^m \| B \|^k \\ &\qquad + \sum_{k=0}^{\infty} \sum_{m=[\frac{p}{2}]}^{\infty} \frac{1}{m!k!} \| A \|^m \| B \|^k \\ &= \left(\sum_{m=0}^{\infty} \frac{1}{m!} \| A \|^m\right)\left(\sum_{k=[\frac{p}{2}]}^{\infty} \frac{1}{k!} \| B \|^k\right) \\ &\qquad + \left(\sum_{k=0}^{\infty} \frac{1}{k!} \| B \|^k\right)\left(\sum_{m=[\frac{p}{2}]}^{\infty} \frac{1}{m!} \| A \|^m\right) \\ &= (\exp \| A \|)\alpha_p + (\exp \| B \|)\beta_p, \end{aligned}$$

donde $\alpha_p \to 0$ y $\beta_p \to 0$ a medida que $p \to \infty$.

Métodos similares sirven para tratar $f(A)$, donde f es una función representable por una serie de potencias

$$f(\zeta) = \sum_{n=0}^{\infty} \alpha_n \zeta^n,$$

y donde $\| A \|$ es (estrictamente) más pequeña que el radio de convergencia de la serie. Dejamos al lector la verificación de que el cálculo funcional que estamos sugiriendo es consistente con el cálculo funcional para transformaciones normales. Así, por ejemplo, $\exp A$ como se define anteriormente, es la misma transformación lineal, como se define por nuestra noción previa de $\exp A$, en caso de que A sea normal.

EJERCICIOS

1. Dése una prueba alternativa del teorema ergódico, basada sobre el teorema espectral para transformaciones unitarias.
2. Pruébese que si $\| 1 - A \| < 1$, entonces A es invertible, considerando la expansión formal de la serie de potencias de $(1 - (1 - A))^{-1}$.

APENDICE

ESPACIO DE HILBERT

Probablemente la más útil y, ciertamente, la mejor desarrollada generalización de la teoría de espacios finito-dimensionales de producto interior es la teoría de espacio de Hilbert. Sin entrar en detalles, y enteramente sin pruebas, intentaremos indicar ahora cómo procede esta generalización y cuáles son las principales dificultades que tenemos que vencer.

La definición de espacio de Hilbert es fácil: es un espacio de producto interior que satisface una condición extra. Que esta condición, (a saber, completividad), queda automáticamente satisfecha en el caso finito-dimensional, se prueba en el análisis elemental. En el caso infinito-dimensional puede ser posible que para una secuencia (x_n) de vectores $\| x_n - x_m \| \to 0$, a medida que $n, m \to \infty$, pero, no obstante, no hay un vector x para el cual $\| x_n - x \| \to 0$, la única manera efectiva de descartar esta posibilidad es suponer lo contrario. En otras palabras: un espacio de Hilbert es un espacio completo de producto interior. (Algunas veces el concepto de espacio de Hilbert se restringe por condiciones adicionales, cuyo propósito es el de limitar el tamaño del espacio, de arriba y de abajo al mismo tiempo. Las condiciones más usuales requieren que el espacio sea infinito-dimensional y separable. En años recientes, desde que se advirtió que esas restricciones adicionales no dan por sí mismas buenos resultados, se ha hecho constumbre usar el "espacio de Hilbert" para el concepto que definimos).

Es fácil ver que el espacio $\mathcal{P}$ de polinomios con el producto interior definido por $(x, y) = \int_0^1 x(t)\overline{y(t)}dt$ no es completo. En conexión con la plenitud o completividad de ciertos espacios particulares de Hilbert, hay un conocimiento matemático completamente extenso. Así, por ejemplo, la principal aserción del celebrado teorema de Riesz-Fischer es que un espacio de Hilbert, construido con el conjunto de todas las funciones x para las cuales $\int_0^1 |x(t)|^2 dt < \infty$ (en el sentido de la integración de Lebesgue) es un espacio de Hilbert (con formalmente la misma definición de producto interior que para los polinomios). Otro popular espacio de Hilbert, que recuerda por su apariencia el espacio de coordenadas finito-dimensionales, es el espacio de todas las secuencias (ξ_n) de números (reales o complejos, como sea el caso) para las cuales $\sum_n |\xi_n|^2$ converge.

Usando la plenitud a fin de discutir inteligentemente la convergencia de algunas sumas infinitas, se puede proceder durante bastante tiempo en la construcción de la teoría de los espacios de Hilbert sin encontrar dificultades debidas a la infinito-dimensionalidad. Así, por ejemplo, las nociones de ortogonalidad y de conjuntos completos ortonormales pueden definirse en el caso general exactamente como las definimos. Una prueba de la desigualdad de Bessel y de la equivalencia de las varias formulaciones posibles de plenitud para conjuntos ortonormales, tienen que sufrir solamente ligeros cambios verbales. (La convergencia de las diversas sumas infinitas que entran, es una consecuencia automática de la desigualdad de Bessel). Nuestra prueba de la desigualdad de Schwarz es válida, tal como está, en el caso más general. Finalmente, la prueba de la existencia de conjuntos ortonormales completos se parece bastante a la prueba en el caso finito. En la prueba inconstructiva, el lema de Zorn (o inducción transfinita) reemplaza la inducción ordinaria, y aun los pasos constructivos del proceso Gram-Schmidt son efectuados fácilmente.

En la discusión de múltiples, funcionales, y transformaciones la situación se vuelve incómoda si no hacemos una concesión a la topología del espacio de Hilbert. Pueden probarse buenas generalizaciones de todos nuestros enunciados para el caso finito-dimensional, si consideramos múltiples lineales *cerrados*, funcionales lineales *continuas*, y transformaciones lineales *limitadas*. (En un espacio finito-dimensional, todo múltiple lineal es cerrado, toda funcional lineal es continua y toda transformación lineal es limitada). Si, en cambio, convenimos en hacer estas concesiones, entonces una vez más podemos deslizarnos por inercia sobre nuestras pruebas finito-dimensionales, la mayor parte del tiempo sin cambio alguno, y con sólo la inserción de una ϵ ocasional el resto del tiempo. Así, una vez más obtenemos que $\mathcal{V} = \mathfrak{M} \oplus \mathfrak{M}^\perp$, que $\mathfrak{M} = \mathfrak{M}^{\perp\perp}$, y que toda funcional lineal de x tiene la forma (x, y); nuestras definiciones de transformación autoadjunta y transformación positiva todavía tienen sentido, y todos nuestros teoremas sobre proyecciones perpendiculares (así como sus pruebas) pasan sin cambio.

La primera insinuación de cómo pueden descomponerse las cosas viene del estudio de las transformaciones ortogonales y unitarias. Llamamos todavía a una transformación U ortogonal o unitaria (según el espacio es real o complejo) si $UU^* = U^*U = 1$, y es aún verdadero que esa transformación es isométrica, esto es, que $\| Ux \| = \| x \|$ para toda x, o, equivalentemente, $(Ux, Uy) = (x, y)$ para toda x y y. Es, sin embargo, fácil de construir una transformación isométrica que no sea unitaria; a causa de su importancia en la construcción de contraejemplos, describiremos esa transformación. Consideremos un espacio de Hilbert en que hay un conjunto ortonormal completo contable, digamos $\{x_0, x_1, x_2, \ldots\}$. Una transformación lineal única limitada U se define por las condiciones $Ux_n = x_n + 1$ para $n = 0, 1, 2, \ldots$. Esta U es isométrica ($U^*U = 1$), pero, puesto que $UU^*x_0 = 0$, no es verdad que $UU^* = 1$.

Es cuando llegamos a la teoría espectral cuando todo el sabor del desarrollo cambia radicalmente. La definición de valor propio con un número λ para el cual $Ax = \lambda x$ tiene una solución no cero, tiene aún sentido, y todavía es verdadero nuestro teorema sobre la realidad de los valores propios de una transformación autoadjunta. La noción de valor propio pierde, sin embargo, mucho de su importancia. Los valores propios son tan útiles en el caso finito-dimensional, porque son un modo de cómo describir el hecho de que hay algo que no anda bien con el recíproco de $A - \lambda$, y la única cosa que puede resultar errónea es que el recíproco se niegue a existir. Cosas esencialmente diferentes pueden suceder en el caso infinito-dimensional; justamente para ilustrar las posibilidades podemos mencionar, por ejemplo, que puede existir la recíproca de $A - \lambda$, pero ser ilimitada. Que no hay una generalización útil de determinante y, en consecuencia, de la ecuación característica, es la menor de nuestras preocupaciones. Toda la teoría ha, de hecho, alcanzado su plena belleza y madurez sólo después de que se renunció a la profusa imitación de los métodos finito-dimensionales.

Después de cierta valuación del hecho de que el caso infinito-dimensional tiene que vencer grandes dificultades, viene como una agradable sorpresa que el teorema espectral para transformaciones autoadjuntas (y, en el caso complejo, hasta para normales) tiene una generalización muy bella y poderosa. (Aunque describimos el teorema sólo para transformaciones limitadas, hay una gran clase de ilimitadas para las cuales es válido). A fin de estar en condiciones de establecer la analogía, reexaminemos el caso finito- dimensional.

Sea A una transformación lineal autoadjusta sobre un espacio finito-dimensional de producto interior, y sea $A = \sum_j \lambda_j F_j$ su forma espectral. Si M es intervalo en el eje real, escribimos $E(M)$ para la suma de todas las F_j para las cuales λ_j pertenece a M. está claro que $E(M)$ es una proyección perpendicular para cada M. Las siguientes propiedades de la función-intervalo, proyección-valuada E son las cruciales: Si M es una unión de una colección contable $\{M_n\}$ de intervalos disjuntos, entonces

(1) $$E(M) = \sum_n E(M_n),$$

y si M es el intervalo impropio consistente en todos los números reales, entonces $E(M) = 1$. La relación entre A y E es descrita por la ecuación

$$A = \sum_j \lambda_j E(\{\lambda_j\}),$$

donde, por supuesto, $\{\lambda_j\}$ es el intervalo degenerado consistente del solo número λ_j. Los que están familiarizados con la integración Lebesgue-Stieltjes reconocerán la última suma escrita como una suma típica aproximante a una integral de la forma $\int \lambda \, dE(\lambda)$ y, en consecuencia, verán cómo puede uno esperar que vaya la generalización. El concepto algebraico de suma es reemplazado por el concepto analítico de integración; la relación generalizada entre A y E es descrita por la ecuación

(2) $$A = \int \lambda \, dE(\lambda).$$

Con la excepción de esta alteración formal, el teorema espectral para transformaciones autoajustas es verdadero en el espacio de Hilbert. Tenemos, por supuesto, que interpretar correctamente el significado de las operaciones limitantes implicadas en (1) y (2). Una vez más estamos confrontados con las tres posibilidades mencionadas en el § 91. Se llaman, respectivamente, convergencia uniforme, fuerte y débil, y resulta que tanto a (1) como a (2) puede darse la interpretación fuerte. (El lector deduce, por supuesto, de nuestro lenguaje que en un espacio finito-dimensional de Hilbert las tres posibilidades son en realidad distintas).

Hemos visto que las proyecciones F_j que entran en la forma espectral de A en el caso finito-dimensional son funciones muy sencillas de A (§ 82). Puesto que las $E(M)$ se obtienen por suma de las

F_j, son también funciones de A, y es completamente fácil describir qué funciones. Escribimos $g_M(\zeta) = 1$, si ζ está en M y $g_M(\zeta) = 0$ de otro modo; entonces $E(M) = g_M(A)$. Este hecho da la pista principal hacia una posible prueba del teorema espectral general. El procedimiento usual consiste en discutir el cálculo funcional para polinomios y, por procesos limitadores, extenderlo a una clase de funciones que incluye todas las funciones g_M. Una vez hecho esto, podemos definir la función-intervalo E escribiendo $E(M) = g_M(A)$; no hay dificultad particular en establecer que E y A satisfacen a (1) y (2).

Después de que se ha probado el teorema espectral, es fácil deducir del mismo las versiones generalizadas de nuestros teoremas concernientes a raíces cuadradas, el cálculo funcional, la descomposición polar, y las propiedades de conmutatividad, y, de hecho, contestar prácticamente toda pregunta susceptible de hacerse sobre transformaciones normales limitadas.

Las dificultades que quedan son las consideraciones de transformaciones no normales y no limitadas. Con respecto de las transformaciones generales no normales, es completamente fácil describir el estado de nuestro conocimiento; es no existente. No existe una generalización aun insatisfactoria para la forma triangular o la forma canónica de Jordan, y la teoría de los divisores elementales. Muy diferente es la situación concerniente a las transformaciones normales (y particularmente autoadjuntas), no limitadas. (El lector simpatizará con el deseo de tratar esas transformaciones si recuerda que la primera y más importante operación funcional que la mayor parte de nosotros aprendemos es la diferenciación). En este particular, insinuaremos apenas el principal obstáculo que confronta la teoría. No es muy difícil demostrar que si una transformación lineal autoadjunta está definida por todos los vectores del espacio de Hilbert, entonces está limitada. En otras palabras, el primer requisito concerniente a transformaciones que nos vemos obligados a renunciar, es que pueden estar definidas en todas partes. La discusión del dominio preciso sobre el cual puede definirse una transformación autoadjunta hasta la cual puede ser agregado este dominio, es la nueva dificultad principal que se encuentra en el estudio de transformaciones no limitadas.

LECTURAS QUE SE RECOMIENDAN

La siguiente lista, muy breve, no tiene la pretensión de ser completa: solamente contiene un par de obras representativas de las diversas direcciones en que puede intentar marchar el lector.

Para álgebra lineal y multilineal generalizada (pero, generalmente, finitodimensional):

1. N. Bourbaki, Algèbre, Chap. II (*Algèbre linéaire*), Paris, 1947 y Chap. III (*Algèbre multilinéaire*), Paris, 1948.
2. B. L. van der Waerden, *Modern Algebra*, Nueva York, 1953.

Para conexiones con el análisis clásico y moderno:

1. S. Banach, *Théorie des opérations linéaries*, Warszawa, 1932.
2. F. Riesz y B. Sz-Nagy, *Functional Analysis*, Nueva York, 1955.

Para la geometría del espacio de Hilbert y transformaciones sobre el mismo:

1. P. R. Halmos, *Introduction to Hilbert space, Nueva York*, 1951.
2. M. H. Stone, *Linear transformation in Hilbert space*, Nueva York, 1932.

Para el contacto con la física clásica y moderna:

1. R. Courant y D. Hilbert, *Methods of mathematical physics*, Nueva York, 1953.
2. J. von Neumann, *Mathematical foundations of quantum mechanics*, Princeton, 1955.

INDICE DE SIMBOLOS

INDICE

—F—

—G—

—H—

—I—

—L—

—M—

—N—

—O—

—P—

—R—

—S—

—T—

—V—

ESTA EDICION DE 2 000 EJEMPLARES SE TERMINO DE IMPRIMIR EL DIA 14 DE OCTUBRE DE 1965, EN LOS TALLERES DE LA CIA. EDITORIAL CONTINENTAL, S. A., CALZADA DE TLALPAN No. 4620, MEXICO 22, D. F.

ALGEBRA Y TRIGONOMETRIA PLANA

Por ABRAHAM SPITZBART y ROSS H. BARDELL

Tela, 468 Págs. Ilustrado

Una modalidad del presente libro es la extensa discusión de las fórmulas trigonométricas de reducción. Las demostraciones de estas fórmulas no son fácilmente asequibles para el lector interesado, y esperamos que los estudiantes apreciarán la demostración completa dada. Puede omitirse el estudio de la demostración sin perturbar lo restante del tratado.

□

METODOS NUMERICOS EN INGENIERIA

Por MARIO G. SALVADORI y MELVIN L. BARON

Tela, 314 Págs. Ilustrado

Este curso es el último de una serie de cinco, impartido por el autor con el fin de ampliar los conocimientos matemáticos, tanto de los alumnos graduados como de los no graduados y para llenar el vacío existente entre un conocimiento teórico de las matemáticas y la resolución de problemas físicos por métodos matemáticos.

□

CALCULO AVANZADO

Por LOUIS BRAND

Tela, 700 Págs. Ilustrado

CONTENIDO: 1. El sistema de los números; 2. Secuencias y series; 3. Funciones de una variable; 4. Funciones de variables múltiples; 5. Vectores; 6. La integral definida; 7. Integrales impropias; 8. Integrales lineales; 9. Integrales múltiples; 10. Convergencia uniforme; 11. Funciones de variable compleja; 12. Series de Fourier. Apéndices: 1. Puntos de acumulación; 2. Ecuaciones de diferencia; 3. Cálculo de diferencias; 4. Comprobaciones dimensionales. Prueba general; Respuestas a los problemas.