NOTAS DE ÁLGEBRA LINEAL

A. Ibort y M.A. Rodríguez
Departamento de Matemáticas, Universidad Carlos III de Madrid
Departamento de Física Teórica II, Universidad Complutense de Madrid

30 de marzo de 2001

ÍNDICE

Prólogo

Estas notas de Álgebra Lineal responden a cursos desarrollados en la Facultad de Ciencias Físicas de la Universidad Complutense durante varios periodos. A través de ellos, el plan de estudios ha cambiado y, como consecuencia, la duración de los cursos y su contenido. Hemos procurado que este texto comprenda el actual temario de la asignatura, aunque ello dependerá en última instancia del enfoque particular que cada profesor dé al tema. Pero además aparecen en él otras cuestiones que complementan o aclaran algunos aspectos de la asignatura y que, aunque pueden ser evitados en una primera lectura, contribuyen, desde nuestro punto de vista, a una mejor comprensión del tema tratado.

El contenido se divide en cuatro grandes temas dedicados al estudio de los espacios vectoriales, las aplicaciones lineales y la teoría de matrices, los espacios con producto escalar y los operadores en estos últimos espacios. Estos temas van precedidos por una introducción con conceptos elementales sobre estructuras algebraicas. El curso se completa con un capítulo sobre tensores y algunos conceptos de geometría afín con una aplicación a la clasificación de cónicas También se incluye una colección de problemas planteados en clase o problemas de tipo clásico similares a los que pueden encontrarse en otros textos, así como problemas propuestos en exámenes de la asignatura.

En cuanto al tema primero de estructuras algebraicas, no todas las nociones expuestas son necesarias para el curso. En particular, la teoría de grupos podría parecer excesiva y no es nuestro objetivo insistir en ella. Sin embargo, el conocimiento de algunas nociones del grupo de permutaciones es esencial para una mejor comprensión de la teoría de determinantes y por eso se ha desarrollado con un cierto detalle. Asimismo, aunque desde luego no se entra en la construcción de los números reales (un tema que no es propio del álgebra) sí se desarrolla la de los números complejos, pues resulta fundamental en el curso. Algunas ideas sobre anillos y cuerpos completan este capítulo que se cierra con la primeras nociones sobre polinomios.

Los capítulos dedicados a la teoría de espacios vectoriales, aplicaciones lineales y matrices son fundamentales y forman parte de cualquiera curso básico de álgebra lineal. En el enfoque que aquí se ha dado, los sistemas lineales aparecen como asociados de forma natural a la teoría de aplicaciones lineales. Pensamos que, como el planteamiento de forma independiente ya se ha estudiado en cursos anteriores, merece la pena enfocarlo desde la perspectiva mencionada. Asimismo, los espacios de aplicaciones lineales, en especial el espacio dual, permiten la introducción de nuevos ejemplos de espacios lineales que resultarán de gran utilidad en el resto del curso.

El capítulo dedicado al estudio de las formas canónicas de endomorfismos puede simplificarse en gran manera. En particular la falta de tiempo aconseja muchas veces prescindir del estudio de formas canónicas de Jordan para endomorfismos que no son diagonalizables y limitarse a un estudio de los diagonalizables. Pero no cabe duda de que, aun sin entrar en el detalle de la construcción de estas formas de Jordan, sí merece la pena hacer algún comentario sobre su existencia y estructura.

los temas dedicado a espacios con productos escalares son básicos en las aplicaciones, en particular en Física que es el objeto final de los estudiantes a los que van dirigidas estas notas. Algunas cuestiones como la definición de formas sesquilineales pueden ser evitados pasando directamente a las cuestiones relacionadas con el producto escalar. Pero las formas canónicas de operadores simétricos, unitarios y ortogonales deberían figurar en el contenido del curso.

El capítulo sobre tensores es un tanto singular en la estructura del curso. Parte de su contenido sobrepasa ciertamente el nivel general de estas notas. Pero no es difícil extraer de él las ideas más elementales de lo que es un tensor bajo transformaciones. Además dado su carácter aislado su supresión no afectaría a un curso basado en este texto.

Finalmente se presenta un capítulo con cuestiones más geométricas como los movimientos en los espacios euclídeos y como aplicación la clasificación de cónicas, sobre el que, sin embargo, no se ha incluido ningún problema en esta versión.

Muchas cuestiones que aparecen en esta notas no pueden ser expuesta en el desarrollo del curso, pero esperamos que su lectura ayude, como hemos dicho al principio, a la comprensión de todo el temario.

Muchas han sido los fuentes de las que han surgido estas notas. Aparte de la experiencia de los autores (y de los numerosos libros consultados) citaremos la contribución de los profesores del Departamento de Física Teórica II de la Facultad de Ciencias Físicas de la Universidad Complutense que han enseñado esta materia durante diversos cursos. Aunque, como es habitual, los errores y faltas nos correspondan exclusivamente, los aciertos y utilidad de estas notas deben adjudicarse a todos ellos (en particular también a nosotros).

Alberto Ibort
Miguel A. Rodríguez

Bibliografía

La lista siguiente incluye algunos libros que, bien por su accesibilidad o por su interés general nos ha parecido oportuno incluir, entre ellos, una colección de problemas. Por supuesto, el nivel que presentan es muy variado y las preferencias de los autores de estas notas solo aparecen reflejadas en algunos de ellos.

Burgos, J. de, *Algebra Lineal*, McGraw Hill, Madrid, 1993.
Gantmacher, F.R., *Théorie des matrices*, Dunod, Paris, 1966.
Gel'fand, I.M., *Lectures on Linear Algebra*, Dover, N.Y. 1989.
Hungerford, T.W., *Algebra*, Holt, Rinehart and Winston, Inc. 1973.
Kostrikhin, A.I., *Introducción al Álgebra*, McGraw Hill, 2a. edición, Madrid, 1993.
Nomizu, K., *Fundamentals of Linear Algebra*, Academic Press, New York, 1974.
Rojo, J., Martín, I., *Ejercicios y problemas de álgebra lineal*, McGraw Hill, Madrid, 1994.
Souriau, J.M., *Calcul Linéaire*, Editions Jacques Gabay, 2 édition 1992.

Tema 1

Estructuras algebraicas

Grupos. Anillos. Números enteros. Cuerpos. Números racionales. Números reales. Números complejos. Polinomios.

1.1 Notación y teoría de conjuntos

Se supone que los alumnos se hallan familiarizados con la teoría elemental de conjuntos. A lo largo de este texto los conjuntos serán denotados habitualmente por letras latinas mayúsculas $A, B, C, \ldots, X, Y, Z$. Los elementos de un conjunto A se denotarán por letras latinas minúsculas $a, b, c, \ldots, x, y, z$. El símbolo $a \in A$ significa que el elemento a pertenece al conjunto A, así $A = \{a \in A\}$. Existe un conjunto que no posee ningún elemento, tal conjunto se llama vacío y se denota por $\emptyset$.

> Nota. Aunque no será necesario en este curso, nos gustaría hacer notar que no todas las construcciones que pueden hacerse en álgebra (incluso a este nivel elemental) conducen a conjuntos. Por ejemplo la familia formada por todos los conjuntos no es un conjunto (¿Por qué?). En este sentido es conveniente tener cuidado al definir conjuntos y utilizarlos. Por ejemplo, si "definimos" el conjunto de los números racionales cuya primera cifra decimal es cero nos encontramos que no sabemos si el número $1/10$ pertenece o no, ya que su expresión decimal es $0.1 = 0.0\bar{9}$, por tanto no hemos definido un conjunto. Un ejemplo mucho menos evidente es el siguiente: consideremos el conjunto de los números naturales "interesantes". Podemos probar inmediatamente que todo número natural es "interesante". En efecto, tomemos el complementario C de este subconjunto. Ciertamente el número 1 es interesante luego no pertenece a C. Probemos que $C = \emptyset$. Si $C \neq \emptyset$ existe un elemento m mínimo en dicho conjunto, luego m es el número natural más pequeño que no es interesante, pero ésta es desde luego una propiedad interesante, por tanto m es interesante y C debe ser vacío. QED

El símbolo $f\colon A \to B$ (o también $A \xrightarrow{f} B$) denotará a lo largo del texto una aplicación f del conjunto A, llamado dominio de f, en B, llamado rango de f. Si $f\colon A \to B$, $g\colon B \to C$ son dos aplicaciones $g \circ f$ denotará su composición. La imagen de $a \in A$ por f se denotará $f(a)$. Con esta notación definimos la composición de aplicaciones como $(g \circ f)(a) = g(f(a))$.

El producto cartesiano de dos conjuntos A, B se denotará por $A \times B$ y se define como $A \times B = \{(a, b) \mid a \in A, b \in B\}$. La unión de dos conjuntos se denotará por $A \cup B = \{x \mid x \in A \vee x \in B\}$, y la intersección por $A \cap B = \{x \mid x \in A \wedge x \in B\}$. Denotaremos por $A \setminus B = \{a \in A \mid a \notin B\}$. Así, $A \setminus A = \emptyset$.

El cuantificador lógico $\forall$ significa "para todo" y $\exists$ significa "existe". También utilizaremos $\exists!$ que significa "existe un único".

1.2 Grupos

1.2.1 Operaciones binarias internas

Una operación binaria interna en un conjunto X es una aplicación $\star\colon X \times X \to X$. Habitualmente la imagen por la aplicación $\star$ de dos elementos $x, y \in X$ se denotará por $\star(x, y) = x \star y$, y se leerá "x multiplicado por y" o "x por y". Escribimos $(X, \star)$ para denotar el conjunto X junto con la ley $\star$. Si X es un conjunto finito, una operación binaria $\star$ se puede describir dando su tabla de multiplicar: se colocará sobre el eje OX los elementos de X y sobre el eje OY de nuevo los elementos de X. En los nodos o puntos de intersección en el retículo definido por estos puntos, colocaremos los resultados de multiplicar los correspondientes elementos. Esto es, si $X = \{x_1, x_2, \ldots, x_n\}$, tendremos,

$\star$	x_1	x_2	$\cdots$	x_{n-1}	x_n
x_1	$x_1 \star x_1$	$x_1 \star x_2$	$\cdots$	$x_1 \star x_{n-1}$	$x_1 \star x_n$
x_2	$x_2 \star x_1$	$x_2 \star x_2$	$\cdots$	$x_2 \star x_{n-1}$	$x_2 \star x_n$
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$
x_n	$x_n \star x_1$	$x_n \star x_2$	$\cdots$	$x_n \star x_{n-1}$	$x_n \star x_n$

Nótese que $\star$ es una aplicación si y sólo si la tabla queda completamente llena y en cada nodo hay un único elemento.

Ejemplo 1.2.1 Sea $X = \{a, b\}$ y $\star$ la operación binaria interna con tabla de multiplicar

$\star$	a	b
a	a	b
b	b	a

La tabla anterior es equivalente a la definición de la aplicación $\star$, $a \star a = a$, $a \star b = b$, $b \star a = b$, $b \star b = a$.

Ejemplo 1.2.2 $X = \{a, b\}$. Definiremos la operación binaria interna $\perp$ a través de su tabla de multiplicar,

$\perp$	a	b
a	a	a
b	b	b

o equivalentemente $a \perp a = a$, $a \perp b = a$, $b \perp a = b$, $b \perp b = b$.

Un elemento $e \in X$ se dirá que es neutro por la derecha respecto a $\star$ si $x \star e = x$, $\forall x \in X$. Análogamente se dirá que es neutro por la izquierda si $e \star x = x$, $\forall x \in X$. Diremos que e es simplemente neutro si es neutro por la derecha y por la izquierda. En otros términos un elemento es neutro si su columna y fila en la tabla de multiplicar es simplemente una copia de X. En el ejemplo 1.2.1 a es neutro. En el ejemplo 1.2.2 no hay elemento neutro.

Ejercicio 1.2.1 Probar que si $(X, \star)$ tiene elemento neutro e, éste es único.

Sea $(X, \star)$ un conjunto con producto $\star$ y elemento neutro e. Diremos que y es un inverso a derecha (izquierda) de x si $x \star y = e$ $(y \star x = e)$. Diremos que y es un inverso de x si es inverso a derecha e izquierda.

Ejemplo 1.2.3 Sea $X = \{a, b, c\}$ con la operación binaria interna,

$\star$	a	b	c
a	a	b	c
b	b	a	a
c	c	a	b

Observamos que a es el elemento neutro. $b \star b = a$ implica que b es un elemento inverso de b. $b \star c = a = c \star b$ implica que c es un elemento inverso de b.

Diremos que una operación interna $\star$ es asociativa si $(x \star y) \star z = x \star (y \star z)$, $\forall x, y, z \in X$. En tal caso $\star$ se llamará usualmente "producto" en X.

Ejercicio 1.2.2 Probar que si $\star$ es asociativa y $x \in X$ tiene inverso, éste es único. Tal elemento se denotará habitualmente por x^{-1}.

Las operaciones de los ejemplos 1.2.1 y 1.2.2 son asociativas, no así la del 1.2.3.

Un conjunto X con una operación asociativa $\star$ se denomina semigrupo.

Ejemplo 1.2.4 $\mathbb{N} = \{1, 2, 3, \ldots\}$ denota el conjunto de los números naturales. Denotaremos por $+$ la operación binaria interna definida por la adición ordinaria de números naturales. $(\mathbb{N}, +)$ es un semigrupo que no posee elemento neutro. Denotaremos por $\cdot$ la multiplicación ordinaria de números naturales. $(\mathbb{N}, \cdot)$ es un semigrupo con elemento neutro 1.

Así como la tabla de sumar no se obliga a "memorizar" a los niños, la tabla de multiplicar de los números naturales se hace memorizar a todos los niños del mundo. Es la primera operación interna no trivial que pertenece al acervo cultural de la humanidad.

Una operación binaria $\star$ se dirá conmutativa si $x \star y = y \star x$, $\forall x, y \in X$. Las operaciones $+$, $\cdot$ en el ejemplo 1.2.4 son conmutativas. Las operaciones de los ejemplos 1.2.1 y 1.2.3 son conmutativas pero la del ejemplo 1.2.2 no lo es. Si $\star$ es conmutativa su tabla de multiplicar es simétrica respecto a la diagonal.

Si X posee dos operaciones internas $\star$, $\perp$, diremos que $\perp$ es distributiva respecto de $\star$ si $x \star (y \perp z) = (x \star z) \perp (x \star z)$, $\forall x, y, z \in X$. En $(\mathbb{N}, +, \cdot)$, la suma $+$ es distributiva respecto de $\cdot$.

1.2.2 Permutaciones y grupos

Por muy variadas razones la familia de las permutaciones de una colección finita de elementos forman un conjunto muy importante. Lo vamos a discutir detalladamente. Consideremos por ejemplo el conjunto $X = \{1, 2, \ldots, n\}$ de los n primeros números naturales. Una permutación de $1, 2, \ldots, n$ es una biyección $\alpha \colon X \to X$. Nótese que $\alpha(1)$ será por tanto un número natural entre 1 y n que podemos denotar por α_1, $\alpha(2)$ será otro denotado por α_2, etc., hasta $\alpha(n) = \alpha_n$. Diremos que la lista de números $\alpha_1 \alpha_2 \cdots \alpha_n$ se obtiene de la $123 \cdots n$ por una "permutación", la permutación α. Es convencional escribir la permutación α como una matriz

$$\alpha = \begin{pmatrix} 1 & 2 & \cdots & n \\ \alpha_1 & \alpha_2 & \cdots & \alpha_n \end{pmatrix}$$

que es autoexplicativa, esto es, $1 \mapsto \alpha_1$, $2 \mapsto \alpha_2$, ... , $n \mapsto \alpha_n$.

El conjunto de todas las permutaciones del conjunto $\{1, 2, \ldots, n\}$ se denotará por S_n. En S_n definimos una operación binaria interna $\cdot$ como la composición de aplicaciones, esto es:

$$\alpha \cdot \beta = \alpha \circ \beta, \forall \alpha, \beta \in S_n,$$

esto es, $(\alpha \cdot \beta)(i) = \alpha(\beta(i))$, $i = 1, 2, \ldots, n$.

La operación $\cdot$ es asociativa ya que la composición de aplicaciones lo es (¡probadlo!).

Denotaremos por e la permutación correspondiente a la aplicación identidad, esto es, $e(i) = i$, $i = 1, 2, \ldots, n$, o en la notación anterior

$$e = \begin{pmatrix} 1 & 2 & \cdots & n \\ 1 & 2 & \cdots & n \end{pmatrix}.$$

Claramente $\alpha \cdot e = e \cdot \alpha = \alpha$, $\forall \alpha \in S_n$, por tanto e es el elemento neutro de $(S_n, \cdot)$. Toda permutación α tiene elemento inverso (¡único!). En efecto, si $\alpha \in S_n$ como α es biyectiva existe la aplicación inversa $\alpha^{-1} \colon X \to X$, tal que $\alpha^{-1}(i) = j$ si $\alpha(j) = i$. Es evidente que $\alpha \cdot \alpha^{-1} = \alpha^{-1} \cdot \alpha = e$.

Ejemplo 1.2.5 Sea $\alpha = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 4 & 1 & 3 \end{pmatrix}$, con $\alpha^{-1} = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 1 & 4 & 2 \end{pmatrix}$. Sea $\beta = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 3 & 4 & 2 \end{pmatrix}$, entonces,

$$\alpha \cdot \beta = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 4 & 1 & 3 \end{pmatrix} \cdot \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 3 & 4 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 1 & 3 & 4 \end{pmatrix},$$

y

$$\beta \cdot \alpha = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 3 & 4 & 2 \end{pmatrix} \cdot \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 4 & 1 & 3 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 2 & 1 & 4 \end{pmatrix},$$

luego $\alpha \cdot \beta \neq \beta \cdot \alpha$ y la operación $\cdot$ no es conmutativa.

Ejemplo 1.2.6 Tablas de multiplicar de $(S_2, \cdot)$ y $(S_3, \cdot)$. Vemos que S_2 es conmutativo y S_3 no lo es. $S_2 = \left\{ e, \tau = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix} \right\}$.

$\cdot$	e	τ
e	e	τ
τ	τ	e

(Nótese que esta tabla de multiplicar coincide, salvo notación, con la tabla del ejemplo 1.2.1).

Consideremos a continuación el conjunto S_3

$$\left\{ e, \tau_1 = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}, \tau_2 = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix}, \tau_3 = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix}, \right.$$

$$\left. \sigma_1 = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}, \sigma_2 = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix} \right\}.$$

$\cdot$	e	τ_1	τ_2	τ_3	σ_1	σ_2
e	e	τ_1	τ_2	τ_3	σ_1	σ_2
τ_1	τ_1	e	σ_2	σ_1	τ_3	τ_2
τ_2	τ_2	σ_1	e	σ_2	τ_1	τ_3
τ_3	τ_3	σ_2	σ_1	e	τ_2	τ_1
σ_1	σ_1	τ_2	τ_3	τ_1	σ_2	e
σ_2	σ_2	τ_3	τ_1	τ_2	e	σ_1

Se observa fácilmente que $\tau_i^{-1} = \tau_i$, $i = 1, 2, 3$ y $\sigma_1^{-1} = \sigma_2$, $\sigma_2^{-1} = \sigma_1$.

Ejercicio 1.2.3 Escribid la tabla de multiplicar de S_4 y S_5.

Ejercicio 1.2.4 Probad que el producto en $(S_n, \cdot)$, $n \geq 3$ no es conmutativo.

Definición 1.2.1 *Un conjunto G con una operación binaria interna $\star$ se dirá que es un grupo si,*

i. Posee elemento neutro $e \in G$.

ii. La operación $\star$ es asociativa, y

iii. Todo elemento posee inverso, esto es, $\forall x \in G$, $\exists x^{-1} \in G$.

De todo lo anterior se desprende que $(S_n, \cdot)$ es un grupo. Dicho grupo se llama el grupo de permutaciones de n elementos. Cuando nos refiramos a un grupo $(G, \star)$ habitualmente omitiremos la operación $\star$ y si no hay riesgo de confusión también omitiremos el símbolo $\star$ al escribir el producto, esto es, escribiremos xy en lugar de $x \star y$.

Si el grupo G es finito, llamaremos orden del grupo G al número de sus elementos y se denotará por $\mid G \mid$. Si G no es finito diremos que $\mid G \mid = \infty$. Por ejemplo $\mid S_2 \mid = 2$, $\mid S_3 \mid = 6$.

Ejercicio 1.2.5 $\mid S_n \mid = n!$.

Definición 1.2.2 *Un subconjunto $H \subset G$ del grupo $(G, \cdot)$ se dirá que es un subgrupo si*
i. $\forall x, y \in H,\ x \cdot y \in H$,
ii. $\forall x \in H,\ x^{-1} \in H$.

Un subgrupo H de G es a su vez un grupo con la operación inducida de la del grupo G. Sea $H = \{e\}$, H es un subgrupo llamado el subgrupo trivial. $\emptyset \subset G$ no es un subgrupo. Si $H = G$, H es un subgrupo. G y $\{e\}$ se llaman subgrupos impropios (o triviales). Un subgrupo H diferente de $\{e\}$ y G se dirá propio.

Ejemplo 1.2.7 $A_3 = \{e, \sigma_1, \sigma_2\} \subset S_3$. A_3 es un subgrupo de S_3. En efecto del ejemplo 1.2.6 obtenemos que

A_3	e	σ_1	σ_2
e	e	σ_1	σ_2
σ_1	σ_1	σ_2	e
σ_2	σ_2	e	σ_1

El subconjunto $\{e, \tau_1\} \subset S_3$ es un subgrupo. Lo mismo ocurre con $\{e, \tau_2\}$, $\{e, \tau_3\}$. Ningún otro subconjunto de S_3 es un subgrupo.

1.2.3 Más sobre el grupo de permutaciones

Un ciclo es una permutación α en S_n de la forma

$$\alpha(k_1) = k_2, \alpha(k_2) = k_3, \ldots, \alpha(k_{r-1}) = k_r, \alpha(k_r) = k_1,$$

donde $\{k_1, k_2, \ldots, k_r\} \subset \{1, 2, \ldots, n\}$ y los demás elementos no cambian. Tal permutación se denotará por $\alpha = (k_1 k_2 \cdots k_r)$ y habitualmente se indica el número de elementos que se permutan cíclicamente, esto es, se dice que $(k_1 k_2 \cdots k_r)$ es un r–ciclo.

Ejemplo 1.2.8 En S_3 todo elemento es un ciclo. En S_4 no todo elemento es un ciclo. Por ejemplo, la permutación $\alpha = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 1 & 4 & 3 \end{pmatrix}$ es el producto de dos 2–ciclos, $\alpha = (12)(34)$.

Llamaremos transposiciones a los 2–ciclos, esto es a los elementos de S_n de la forma $(k_1 k_2)$. Por ejemplo en S_3 los elementos τ_1, τ_2 y τ_3 son transposiciones.

Los resultados más importantes sobre la aritmética de ciclos son:

Proposición 1.2.1 *Toda permutación admite una descomposición única salvo orden en producto de ciclos disjuntos que conmutan entre si.*

Ejemplo 1.2.9 $\sigma \in S_6$. $\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 1 & 4 & 2 & 3 & 6 & 5 \end{pmatrix} = (1)(243)(56)$.

Proposición 1.2.2 *Toda permutación admite una descomposición en producto de transposiciones (que no conmutan en general y que no es única).*

Ejemplo 1.2.10 $\sigma \in S_3$. $\sigma = (123) = (12)(23) = (23)(13)$.

Proposición 1.2.3 *La paridad del número de transposiciones en las que se puede descomponer toda permutación no depende de la descomposición sino sólo de la permutación.*

Se llama paridad o signatura de una permutación al número $\epsilon(\sigma) = (-1)^k$, donde k es el número de transposiciones de una descomposición de $\sigma \in S_n$.

Proposición 1.2.4 *La paridad de un producto es el producto de las paridades.*

$$\epsilon(\alpha\beta) = \epsilon(\alpha)\epsilon(\beta).$$

Ejemplo 1.2.11 Todas las transposiciones tienen paridad impar. Un k-ciclo tiene paridad $(-1)^{k-1}$.

El conjunto de las permutaciones de paridad par forman un subgrupo de S_n llamado el grupo de las alternaciones o grupo alternado. Se denota habitualmente por A_n y su orden es $n!/2$.

1.2.4 Homomorfismos de grupos

Una aplicación $f\colon G \to G'$ entre dos grupos $(G, \cdot)$, $(G', \star)$, se dirá que es un homomorfismo de grupos si $f(g \cdot h) = f(g) \star f(h)$, $\forall g, h \in G$. Si el homomorfismo f es inyectivo se dirá que es un monomorfismo. Si es suprayectivo se dirá que es un epimorfismo y si f es biyectivo se dirá que es un isomorfismo.

Ejercicio 1.2.6 Denotemos por D_3 el grupo de simetrías de un triángulo equilátero. Probar que D_3 es isomorfo a S_3.

Ejemplo 1.2.12 Si denotamos por T el grupo de simetrías de un tetraedro regular, entonces S_4 es isomorfo a T.

Si $f\colon G \to G'$ es un homomorfismo de grupos, llamaremos núcleo de f el subconjunto de G que se aplica en el elemento neutro de G' y se denota por $\ker f$,

$$\ker f = \{g \in G \mid f(g) = e'\}.$$

El conjunto imagen de f se denotará habitualmente por $\operatorname{im} f = \{f(g) \in G' \mid g \in G\}$.

Proposición 1.2.5 *$\ker f$, $\operatorname{im} f$ son subgrupos.*

Ejemplo 1.2.13 Considérese la aplicación $i\colon S_3 \to S_4$ definida por $i(\alpha) = \begin{pmatrix} 1 & 2 & 3 & 4 \\ \alpha_1 & \alpha_2 & \alpha_3 & 4 \end{pmatrix}$. Entonces i es un monomorfismo.

La inclusión natural $j\colon A_n \to S_n$ es un monomorfismo.

La asignación a cada permutación de su paridad, $\epsilon\colon S_n \to \mathbb{Z}_2$ es un epimorfismo debido a la proposición 1.2.4.

Un subgrupo H de G se dice normal si $gHg^{-1} \subset H$, $\forall g \in G$, esto es, si para todo $h \in H$, $ghg^{-1} \in H$ para todo $g \in G$. $\ker f$ es un subgrupo normal de G.

Ejemplo 1.2.14 A_n es un subgrupo normal de S_n.

1.3 Anillos

1.3.1 Los números enteros

En esta sección revisaremos escuetamente los números enteros y la noción de anillo.

Hemos visto que el conjunto de los números naturales $\mathbb{N}$ tiene una estructura de semigrupo respecto a la suma (también con respecto al producto). Podemos plantearnos como extender este conjunto para convertirlo en un grupo. Más concretamente, la ecuación $x + n = m$, $n, m \in \mathbb{N}$ no siempre tiene solución en los números naturales. ¿Podemos extender $\mathbb{N}$ para que la ecuación anterior siempre se pueda resolver? Hay un procedimiento natural para hacer esto y consiste en añadir las raíces de esta ecuación a $\mathbb{N}$. Si denotamos la raíz de $x + n = m$ por $m - n$ vemos inmediatamente que $m - n = (m + r) - (n + r)$ para todo $r \in \mathbb{N}$, lo que nos permite introducir una relación de equivalencia en el conjunto de todas las raíces de todas las ecuaciones $x + n = m$. Denotaremos por $(n - m)$ una de estas clases. Podemos definir la suma de raíces como sigue:

$$(m - n) + (m' - n') = ((m + m') - (n + n')).$$

El elemento neutro de la suma es la raíz de la ecuación $x + n = n$, esto es $(n - n)$ que denotaremos por 0. Si $m > n$ existe un número natural r tal que $m = n + r$ y la clase $(m - n)$ la denotaremos simplemente por r. Si $m < n$ de manera análoga existe un número natural s tal que $n = m + s$ y la clase $(m - n)$ se denotará por $-s$.

El conjunto de raíces se denotará $\mathbb{Z}$ y sus elementos se llamarán números enteros.

$$\mathbb{Z} = \{\ldots, -2, -1, 0, 1, 2, \ldots\}.$$

Alternativamente los números enteros pueden construirse considerando una relación de equivalencia en el producto cartesiano $\mathbb{N} \times \mathbb{N}$ como sigue: $(n, m) \sim (n', m')$ si y sólo si $n + m' = m + n'$. La clase de equivalencia que contiene a (m, n) se denotará como $[m, n]$. Definimos la suma en el conjunto de clases como $[m, n] + [r, s] = [m + r, n + s]$.

Ejercicio 1.3.1 Probad que la operación $+$ está bien definida y proporciona una estructura de grupo en $\mathbb{N} \times \mathbb{N}/\sim$. La operación es conmutativa con elemento neutro la clase $[m, m]$.

Es fácil comprobar que hay tres tipos de clases: la clase $[m, m]$ que denotaremos por 0; las de tipo $[m + r, m]$, que denotaremos por r; y, finalmente, las de tipo $[m, m + r]$ que denotaremos por $-r$. Esto muestra de nuevo que el conjunto de raíces de la ecuación lineal de primer orden con coeficientes naturales está formado por los elementos del conjunto $\mathbb{N} \times \mathbb{N}/\sim$. Identificaremos a partir de este momento ambos conjuntos y los llamaremos indistintamente números enteros. El subconjunto de enteros $0, 1, 2, \ldots,$ se denominarán enteros positivos y el subconjunto $0, -1, -2, \ldots,$ se denominarán enteros negativos. El cero es el único entero positivo y negativo. Diremos que p es menor o igual que q si $p - q$ es positivo y lo denotaremos $p \leq q$. La relación $\leq$ es una relación de orden total en $\mathbb{Z}$.

Tenemos la siguiente propiedad fundamental de los números enteros (y de los naturales):

Teorema 1.3.1 *Cualquier subconjunto no vacío de enteros positivos posee un elemento menor o igual que todos los demás que se denomina mínimo del conjunto.*

Demostración. Un tal subconjunto contiene un entero positivo n ya que es no vacío. Entonces el primer elemento en la lista $0, 1, 2, \ldots, n - 1, n$ contenido en el conjunto tiene la propiedad en cuestión. QED

Una propiedad equivalente al teorema anterior es el "principio de inducción completa".

Teorema 1.3.2 *Principio de inducción completa. Si una proposición sobre un número entero positivo n es cierta para $n = 0$, y su veracidad para todo $0 \leq k < n$ implica su veracidad para n, entonces es cierta para todo n.*

Demostración. Llamemos F el subconjunto de números enteros positivos para los cuales la proposición es falsa. Si F es no vacío, tomemos el mínimo de este conjunto, llamémosle n_0. Pero la proposición es cierta para todo $k < n_0$ y por hipótesis la proposición es cierta para n_0. QED

Producto de números enteros. Definimos una operación $\cdot$ en $\mathbb{Z}$ como sigue: $[m, n] \cdot [r, s] = [mr + ns, ms + nr]$, o utilizando la notación de raíces,

$$n \cdot m = nm; n \cdot (-m) = -(nm); (-n) \cdot (-m) = nm; n \cdot 0 = (-n) \cdot 0 = 0.$$

Omitiremos en lo sucesivo el punto "$\cdot$" en el producto de números enteros excepto por motivos de claridad en la notación.

Existe elemento neutro para el producto de enteros, el 1.

Proposición 1.3.1 ± 1 *son los únicos enteros que poseen inverso respecto al producto.*

Es inmediato verificar que el producto es asociativo,

$$p(qr) = (pq)r, \quad \forall p, q, r \in \mathbb{Z},$$

distributivo,

$$p(q + r) = pq + pr,$$

conmutativo,

$$pq = qp,$$

y además

$$0 \cdot p = p \cdot 0 = 0.$$

Definición 1.3.1 *Un anillo es un conjunto A dotado de dos operaciones binarias internas denotadas respectivamente por $+$ y $\cdot$, tales que $(A,+)$ es un grupo Abeliano y $(A,\cdot)$ es un semigrupo, satisfaciéndose además la propiedad distributiva:*

$$x\cdot(y+z)=x\cdot y+x\cdot z,\quad x,y,z\in A.$$

Si la operación $\cdot$ es conmutativa se dirá que el anillo es conmutativo, y si posee elemento neutro respecto del producto, se dirá que A es un anillo con identidad.

Ejemplo 1.3.1 $(\mathbb{Z},+,\cdot)$ es un anillo conmutativo con identidad. En $\mathbb{Z}$ se satisface además la siguiente propiedad: si $pq=0$, entonces, o bien $p=0$ o $q=0$. Un tal anillo se dice que es un dominio de integridad.

Ejemplo 1.3.2 . Considérese el conjunto $\mathbb{H}=\{(q_0,q_1,q_2,q_3)\mid q_i\in\mathbb{Z}\}$ con las operaciones:

$$(q_0,q_1,q_2,q_3)+(q_0',q_1',q_2',q_3')=(q_0+q_0',q_1+q_1',q_2+q_2',q_3+q_3'),$$

$$(q_0,q_1,q_2,q_3)\cdot(q_0',q_1',q_2',q_3')=(q_0q_0'-q_1q_1'-q_2q_2'-q_3q_3',q_2q_3'-q_3q_2',q_3q_1'-q_1q_3',q_1q_2'-q_2q_1').$$

$\mathbb{H}$ es un anillo con identidad pero no es conmutativo. $\mathbb{H}$ no es un dominio de integridad.

Definición 1.3.2 *Un subconjunto $B\subset A$ de un anillo $(A,+,\cdot)$ se dirá que es un subanillo si*
i. $a-b\in B,\ \forall a,b\in B,$
ii. $a\cdot b\in B,\ \forall a,b\in B.$

Denotamos por $-b$ el inverso de b respecto a la operación $+$. Se desprende de la definición que todo subanillo es un anillo.

Proposición 1.3.2 *Si B es un subanillo de $\mathbb{Z}$ existe un número natural $m\in\mathbb{N}$ tal que $B=m\mathbb{Z}=\{mp\mid p\in\mathbb{Z}\}$.*

Demostración. Si $B=\{0\}$, sea $m=0$. Si no, el conjunto de elementos mayores que cero en B no puede ser vacío. Tomemos m el mínimo de ellos. Por ser B subanillo $m\mathbb{Z}\subset B$. Si $p\in B$, aplicamos el algoritmo de la división por m (ver Teorema 1.3.5) y obtenemos que existe $0\le r<m$ tal que $p=qm+r$, pero entonces $r=p-qm\in B$ y r es positivo y menor que m. QED

Nota. Es suficiente suponer que $m-n\in B$ para todo $m,n\in B\subset\mathbb{Z}$. Tal conjunto es automáticamente un subanillo.

En lo que sigue discutiremos exclusivamente anillos conmutativos con identidad (aunque no exigiremos tal propiedad a los posibles subanillos). La identidad será denotada por 1 o 1_A si hubiera peligro de confusión.

Los elementos invertibles de un anillo A se llaman unidades. El conjunto $U(A)=\{x\in A\mid \exists x^{-1}\in A\}$ es un grupo llamado el grupo de unidades de A.

Definición 1.3.3 *Ideales. Un ideal de un anillo A es un subanillo I que además satisface $xy\in I$ para todo $x\in I$, $y\in A$.*

Corolario 1.3.1 *Los ideales de $\mathbb{Z}$ son de la forma $m\mathbb{Z}$ para algún $m\in\mathbb{Z}$.*

Ejemplo 1.3.3 El anillo de los polinomios $\mathbb{Z}[x]$.

Sea x un símbolo abstracto (podríamos tomar por ejemplo en su lugar un cuadro "abstracto" o el logotipo de una compañía comercial) y considérese el conjunto cuyos elementos son objetos de la forma $a_0+a_1x+a_2x^2+\cdots+a_nx^n$, $a_i\in\mathbb{Z}$, $i=1,\ldots,n$, $n\in\mathbb{N}$. Los símbolos $x^2,x^3,\ldots,x^n$ representan xx,xxx, etc. Los elementos de este conjunto se denominan polinomios, los denotaremos por $P(x),Q(x)$, etc. y al conjunto de todos ellos lo denotaremos por $\mathbb{Z}[x]$ y lo denominaremos el anillo de los polinomios con coeficientes enteros. Definimos en este conjunto las operaciones $+$ y $\cdot$ como sigue:

Si $P(x) = a_0 + a_1x + \cdots + a_nx^n$, $Q(x) = b_0 + b_1x + \cdots + b_mx^m$, y $n \geq m$, entonces $P(x) + Q(x) = (a_0 + b_0) + (a_1 + b_1)x + (a_2 + b_2)x^2 + \cdots + (a_m + b_m)x^m + a_{m+1}x^{m+1} + \cdots + a_nx^n$, y $P(x) \cdot Q(x) = a_0b_0 + (a_1b_0 + a_0b_1)x + (a_2b_0 + a_1b_1 + a_0b_2)x^2 + \cdots + a_nb_mx^{n+m}$. Utilizando una notación más compacta podemos escribir $P(x) = \sum_{i=0}^n a_ix^i$, $Q(x) = \sum_{j=0}^m b_jx^j$, y

$$P(x) + Q(x) = \sum_{k=0}^{\max(n,m)} (a_k + b_k)x^k, \quad P(x) \cdot Q(x) = \sum_{k=0}^{n+m} \left(\sum_{i+j=k} a_ib_j \right) x^k,$$

y en la suma $b_k = 0$, para todo $k > m$.

$\mathbb{Z}[x]$ es un anillo conmutativo con identidad. Cada número entero p define un polinomio cuyo único término es el $a_0 = p$. Los enteros se convierten de este modo en un subanillo de $\mathbb{Z}[x]$ pero no forman un ideal.

Considérese por el contrario conjuntos como $B = \{P(x)(1 + x) \mid P(x) \in \mathbb{Z}[x]\} = (1 + x)\mathbb{Z}[x]$ o $C = \{P(x)(1 + x^2) \mid P(x) \in \mathbb{Z}[x]\} = (1 + x^2)\mathbb{Z}[x]$. Tanto B como C son subanillos y además son ideales.

$\mathbb{Z}[x]$ es un dominio de integridad.

1.3.2 Divisibilidad y factorización de números enteros

Un número entero p se dice que divide (o que es un divisor) de otro entero q si existe un entero r tal que $q = pr$. También diremos que q es un múltiplo de p. Si p divide a q lo indicaremos por $p \mid q$. Es evidente que todos los múltiplos de p son los enteros de la forma rp, $r \in \mathbb{Z}$, que es un ideal de $\mathbb{Z}$ denotado por $p\mathbb{Z}$ y también (p). Nótese que todo número entero tiene al menos cuatro divisores $\pm p$, ± 1 que llamaremos divisores impropios.

Un número entero p se dice que es primo si no posee más divisores que los impropios. Si p es primo $-p$ también lo es. Por esta razón habitualmente se consideran únicamente los primos positivos y mayores que 1.

Teorema 1.3.3 *Teorema fundamental de la aritmética. Todo número entero p se puede escribir como un producto de números primos. Además dicha escritura es única excepto por el orden de los factores.*

Demostración. Ver al final de esta sección.

Por esta razón se dice que $\mathbb{Z}$ es un dominio de factorización única (y también se llama un anillo factorial).

Teorema 1.3.4 *Teorema de Euclides. El conjunto de los primos es infinito.*

Demostración. Supongamos que el conjunto P de los números primos fuera finito, digamos $P = \{p_1, p_2, \ldots, p_N\}$. Entonces el número $p_1p_2 \cdots p_N + 1$ no está en P y en consecuencia no es primo. Pero entonces por el teorema fundamental de la aritmética este número debe ser divisible por alguno de los primos p_i de P lo cual es imposible. QED

Dados dos números enteros p, q, consideremos el conjunto S de todos los enteros de la forma $rp + sq$, con $r, s \in \mathbb{Z}$. Claramente dicho conjunto es un subanillo (de hecho es un ideal). Por tanto hemos visto que $S = m\mathbb{Z}$ para algún $m \in \mathbb{Z}$. Dicho m se llamará el máximo común divisor de p y q y se denotará o bien m.c.d. (p, q) o simplemente $m = (p, q)$.

Ejercicio 1.3.2 Probar que si p es un número primo tal que $p \mid ab$ entonces $p \mid a$ o $p \mid b$.

Ejercicio 1.3.3 Probar que si $m \mid p$ y $m \mid q$, entonces $m \mid (p, q)$. Probar que si $p \mid p'$ y $q \mid q'$, entonces $(p, q) \mid (p', q')$.

Diremos que dos números enteros p y q son primos entre si $(p, q) = 1$. Nótese que esto es equivalente a que existan dos números enteros r, s tales que $rp + sq = 1$.

Ejercicio 1.3.4 Pruébese que la ecuación $px + qy = r$, $p, q, r \in \mathbb{Z}$, tiene soluciones enteras si y sólo si $(p, q) \mid r$.

Teorema 1.3.5 *Algoritmo de la división. Sean $p, q \in \mathbb{Z}$, $q > 0$, entonces existen $d, r \in \mathbb{Z}$ tales que*

$$p = qd + r; 0 \leq r < q,$$

y además d, r son únicos.

Demostración. Consideremos el conjunto $S = \{p - dq \mid d \in \mathbb{Z}, p - dq \geq 0\}$. Claramente $S \neq \emptyset$ (tómese $d = -p^2$). Entonces S tendrá un mínimo (teorema 1.3.1) que denotaremos por r. Necesariamente $0 \leq r < q$ ya que si $r \geq q$, entonces $r = q + r'$, $0 \leq r' < r$ y $r' = p - (d+1)q \in S$ lo cual es absurdo.

Unicidad. Supongamos que d', r' son dos enteros tales que $p = qd' + r'$ y $0 \leq r' < q$. Entonces $q(d - d') = r' - r$. Supongamos que $r' > r$ por tanto $q(d - d') > 0$, esto es $d > d'$ y por tanto $d = d' + d_0$, $d_0 > 0$. Entonces $p = dq + r = q(d' + d_0) + r$ y por tanto $qd_0 + r = r'$, que implica que $r' > q$. Si suponemos que $r > r'$ obtendremos que $r > q$ por tanto la única posibilidad es que $r = r'$ y por tanto $d = d'$. QED

Teorema 1.3.6 *Algoritmo de Euclides. Dados dos números enteros $p, q \in \mathbb{Z}$ podemos calcular su máximo común divisor a través del siguiente algoritmo:*

$$\begin{aligned} p &= qd_0 + r_0, 0 \leq r_0 < q, \\ q &= r_0 d_1 + r_1, 0 \leq r_1 < r_0, \\ r_0 &= r_1 d_2 + r_2, 0 \leq r_2 < r_1, \\ & \cdots\cdots \\ r_{n-2} &= r_{n-1} d_n + r_n, 0 \leq r_n < r_{n-1}, \\ r_{n-1} &= r_n d_{n+1}, r_{n+1} = 0. \end{aligned}$$

Entonces $r_n = (p, q)$.

Demostración. En efecto, si $d \mid p$ y $d \mid q$, entonces $d \mid r_0$ ya que $r_0 = p - qd_0$, por tanto, $(p, q) \mid r_0$, pero $d \mid q$ y $d \mid r_0$ implica que $d \mid r_1$, y así sucesivamente, hasta que $(p, q) \mid r_n$.

Recíprocamente, está claro que $r_n \mid r_{n-1}$, pero $r_{n-2} = r_n d_{n+1} + r_n$, por tanto $r_n \mid r_{n-2}$, etc. hasta que $r_n \mid p$ y $r_n \mid q$, por tanto $r_n \mid (p, q)$. Por tanto $r_n = (p, q)$. QED

1.3.3 Congruencias de números enteros

En el conjunto de los números enteros $\mathbb{Z}$ introducimos una relación de equivalencia como sigue: fijemos un número entero positivo n; diremos que p es congruente con q módulo n si $n \mid p - q$, esto es si $\exists r \in \mathbb{Z}$ tal que $p - q = rn$, o todavía de otro modo, si $q < n$ como $p = rn + q$, q es el resto de dividir p por n.

Si p es congruente con q módulo n, escribiremos $p \equiv q \pmod{n}$.

Ejercicio 1.3.5 Probar que la relación anterior es efectivamente una relación de equivalencia.

La clase de equivalencia que contiene a p se denotará por $[p]$. Claramente $[p] = \{p + nr \mid r \in \mathbb{Z}\}$. Efectivamente si $p' \in [p]$, entonces $p' \equiv p \pmod{n}$, esto es $\exists s \in \mathbb{Z}$ tal que $p' - p = ns$. Obsérvese también que $[p] = [p + n]$, por tanto las clases de equivalencia diferentes de números enteros congruentes módulo n son:

$$[0] = \{sn \mid s \in \mathbb{Z}\}, [1] = \{1 + sn \mid s \in \mathbb{Z}\}, \ldots, [n-1] = \{n - 1 + sn \mid s \in \mathbb{Z}\},$$

esto es, $[0]$ es el conjunto de múltiplos de n, $[1]$ es el conjunto de múltiplos de n más 1, etc. El conjunto de clases de congruencia módulo n se denotará por $\mathbb{Z}_n$, así

$$\mathbb{Z}_n = \{[0], [1], [2], \ldots, [n-1]\}.$$

En $\mathbb{Z}_n$ se definen dos operaciones $+$ y $\cdot$ como sigue:

i. $[r] + [s] = [r + s]$,

ii. $[r] \cdot [s] = [rs]$, $r, s = 0, 1, \ldots, n-1$.

Ejercicio 1.3.6 Probar que las operaciones están bien definidas.

$(\mathbb{Z}_n, +, \cdot)$ es un anillo conmutativo con identidad $[1]$.

Ejemplo 1.3.4 El anillo $\mathbb{Z}_2$ posee dos elementos $\{[0],[1]\}$. En el anillo $\mathbb{Z}_3 = \{[0],[1],[2]\}$ todo elemento posee inverso ya que $[2][2] = [1]$. El anillo $\mathbb{Z}_4 = \{[0],[1],[2],[3]\}$ no es un dominio de integridad ya que $[2][2] = [0]$.

Ejercicio 1.3.7 Sea p primo, probar que todo elemento no nulo en $\mathbb{Z}_p$ tiene inverso.

1.4 Cuerpos

1.4.1 El cuerpo de los números racionales

Los únicos números enteros que poseen inverso respecto a la multiplicación son ± 1. Hay un procedimiento standard para construir a partir de un dominio de integridad un nuevo anillo donde todos sus elementos poseen inverso. Ilustraremos esta construcción con el dominio de integridad de los números enteros $\mathbb{Z}$.

Sea el conjunto $\mathbb{Z} \times \mathbb{Z}^*$, donde $\mathbb{Z}^* = \mathbb{Z} - \{0\}$. Consideremos la siguiente relación de equivalencia $(p,q) \sim (r,s) \Leftrightarrow ps = qr$. Las propiedades reflexiva y simétrica son evidentes. Con respecto a la transitiva, si $(p,q) \sim (r,s)$, y $(r,s) \sim (t,u)$, entonces $ps = qr$, $ru = st$, y por tanto $psu = qru = qst$, esto es $(pu - qt)s = 0$. Como $s \neq 0$, $pu = qt$ lo que quiere decir que $(p,q) \sim (t,u)$.

La clase de equivalencia que contiene al par (p,q) se denotará por p/q o pq^{-1}. El conjunto $\mathbb{Z} \times \mathbb{Z}^*/\sim$ se denotará por $\mathbb{Q}$ y sus elementos se llamaran números racionales (o fraccionarios). Así $\mathbb{Q} = \{p/q : p \in \mathbb{Z}, q \in \mathbb{Z}^*\}$. En $\mathbb{Q}$ definimos dos operaciones $+$, $\cdot$ como sigue:

i. $\dfrac{p}{q} + \dfrac{r}{s} = \dfrac{ps+qr}{st}$, ii. $\dfrac{p}{q} \cdot \dfrac{r}{s} = \dfrac{pr}{qs}$, $\forall \dfrac{p}{q}, \dfrac{r}{s} \in \mathbb{Q}$.

Ejercicio 1.4.1 Probar que $(\mathbb{Q}, +, \cdot)$ es un anillo conmutativo con identidad. Además es un dominio de integridad, pero no es un dominio de factorización única.

Todo elemento $p/q \in \mathbb{Q}$ con $p \neq 0$ tiene inverso. En efecto $(p/q)^{-1} = q/p$ ya que $p/q \cdot q/p = pq/qp = 1/1 = 1$.

En cada clase de equivalencia $p/q \in \mathbb{Q}$ hay un único representante p'/q' tal que que $(p', q') = 1$.

Notas.

1. El conjunto $\mathbb{Q}$ se obtiene en forma análoga a como hicimos en la construcción de los números enteros resolviendo la ecuación $qx = p$, $q \neq 0$. Las raíces de esta ecuación se pueden escribir en una notación obvia como pq^{-1}. Los detalles del análisis se completan de manera trivial. ¿Cómo obtendríamos la suma de números racionales siguiendo esta línea de razonamiento?

2. La misma construcción se puede aplicar al dominio de integridad de los polinomios con coeficientes en $\mathbb{Z}$. El conjunto que se obtiene es el cuerpo de las funciones racionales con coeficientes en $\mathbb{Z}$.

3. La construcción anterior se puede realizar en cualquier anillo A utilizando un sistema multiplicativo S. Un sistema multiplicativo es un subconjunto tal que el producto de cualesquiera par de elementos pertenece al conjunto. En el conjunto de pares $A \times S$ se introduce una relación de equivalencia como anteriormente, esto es $(x,s) \sim (y,t) \leftrightarrow xt = ys$, $x, y \in A$, $s, t \in S$. El conjunto cociente se denota por $S^{-1}A$ y es un anillo (anillo de fracciones de A por S).

Definición 1.4.1 *Un conjunto* $\mathbb{K}$ *dotado de dos operaciones binarias internas* $+$, $\cdot$ *se dirá que es un cuerpo si* $(\mathbb{K}, +, \cdot)$ *es un anillo con identidad y todo elemento* $x \neq 0$ *tiene inverso* x^{-1} *respecto del producto.*

Si $(\mathbb{K}, \cdot)$ es un semigrupo conmutativo el cuerpo $\mathbb{K}$ se dirá conmutativo. En lo sucesivo y salvo especificación contraria trataremos exclusivamente con cuerpo conmutativos.

Definición 1.4.2 *Un cuerpo* $\mathbb{K}$ *se dirá que tiene característica* $n \in \mathbb{N} \cup \{0\}$ *si* $n \cdot 1 = 0$*, donde* 0 *es el elemento neutro de la suma* $(+)$ *y* 1 *el elemento unidad del producto* $(\cdot)$*.*

Ejemplo 1.4.1 $(\mathbb{Q}, +, \cdot)$ es un cuerpo conmutativo de característica 0.

Ejercicio 1.4.2 Probar que si p es primo, $(\mathbb{Z}_p, +, \cdot)$ es un cuerpo. ¿Cuál es su característica?

1.4.2 El cuerpo de los números reales

Un subconjunto $\mathbb{F} \subset \mathbb{K}$ del cuerpo $\mathbb{K}$ se dirá que es un subcuerpo de $\mathbb{K}$, si es un subanillo y si para todo elemento $x \in \mathbb{F}$, $x \neq 0$, entonces $x^{-1} \in \mathbb{F}$. Automáticamente $\mathbb{F}$ es un cuerpo. También se dirá que $\mathbb{K}$ es una extensión de $\mathbb{F}$.

Ejercicio 1.4.3 Probar que $\mathbb{Z}_p$ no posee ningún subcuerpo propio (es decir, distinto de $\{0\}$, $\mathbb{Z}_p$).

Un cuerpo que no posee subcuerpos propios se llama primo.

Ejercicio 1.4.4 Probar que $\mathbb{Q}$ es un cuerpo primo.

El problema de las extensiones de $\mathbb{Q}$

Sea $P(x) = a_0 + a_1 x + \cdots + a_n x^n \in \mathbb{Z}[x]$, diremos que p/q es una raíz de $P(x)$ si $a_0 + a_1 p/q + \cdots + a_n (p/q)^n = 0$, esto es si $q^n a_0 + q^{n-1} p a_1 + \cdots + p^n a_n = 0$. Es claro que los números racionales p/q son las raíces de los polinomios de primer grado $-p + qx$, $q \neq 0$. Es también evidente que hay ecuaciones de segundo grado que no tienen raíces racionales, por ejemplo, $x^2 - 2 = 0$, o $x^2 + 1 = 0$. Podemos plantearnos por tanto si existen cuerpos $\mathbb{K}$, extensiones de $\mathbb{Q}$ donde las ecuaciones anteriores tengan solución. El problema es que hay muchísimos. Existe un cuerpo óptimo, extensión de $\mathbb{Q}$ en el que toda ecuación polinómica tiene alguna raíz: el cuerpo de los números complejos $\mathbb{C}$. La construcción de $\mathbb{C}$ se hace en dos etapas. Una primera que no es algebraica y en la que se construye un cuerpo, llamado de los números reales $\mathbb{R}$, caracterizado por una propiedad topológica (completitud) y una segunda etapa algebraica que describiremos posteriormente.

La construcción de los números reales a partir de los números racionales $\mathbb{Q}$ se realiza utilizando conceptos no algebraicos como ya hemos indicado y que por tanto no reproduciremos aquí (consultar un curso de análisis matemático elemental). Como es bien sabido, números como $\sqrt{2}$, $\sqrt{3}$, $\sqrt{5}+\sqrt{17}$ son números reales al igual que números como π, e, $e^{\sqrt{2}}$, ... y otros números no racionales como 0.01012012301234012345..., etc. Las operaciones de suma y producto de números reales se denotarán con los signos convencionales $+$, $\cdot$. $\mathbb{R}$ es un cuerpo de característica cero. Los números reales $x \in \mathbb{R}$ que son raíces de ecuaciones polinómicas con coeficientes en $\mathbb{Z}$ se llamarán números algebraicos. Todos los racionales son algebraicos, pero también lo son números irracionales como $\sqrt{2}$, $\sqrt{3}$, $\sqrt{2}+\sqrt{5}$, etc.

Teorema 1.4.1 *e y π no son algebraicos.*

1.4.3 Números Gaussianos

Sea $d \in \mathbb{N}$ tal que $\sqrt{d} \notin \mathbb{Q}$. Consideremos el conjunto $\mathbb{Q}(\sqrt{d})$ definido como $\mathbb{Q}(\sqrt{d}) = \{a + b\sqrt{d} \mid a, b \in \mathbb{Q}\}$. Es fácil observar que $\mathbb{Q}(\sqrt{d})$ es un cuerpo con las operaciones:

$$(a + b\sqrt{d}) + (a' + b'\sqrt{d}) = (a + a') + (b + b')\sqrt{d},$$

$$(a + b\sqrt{d}) \cdot (a' + b'\sqrt{d}) = (aa' + bb'd) + (ab' + a'b)\sqrt{d}.$$

Claramente la identidad para el producto es 1 y

$$(a + b\sqrt{d})^{-1} = \frac{a}{a^2 - db^2} + \frac{-b}{a^2 - db^2}\sqrt{d},$$

si a ó $b \neq 0$. Nótese que en este caso $a^2 - db^2 \neq 0$ ya que si $a^2 = db^2$, entonces $\sqrt{d} = a/b \in \mathbb{Q}$ en contra de la hipótesis. Obsérvese que $(a + b\sqrt{d})(a - b\sqrt{d}) = a^2 - db^2$ y así

$$\frac{1}{a + b\sqrt{d}} = \frac{a - b\sqrt{d}}{(a + b\sqrt{d})(a - b\sqrt{d})} = \frac{a - b\sqrt{d}}{a^2 - db^2}.$$

El conjunto $\mathbb{Z}(\sqrt{d}) = \{p + q\sqrt{d} \mid p, q \in \mathbb{Z}\}$ es un anillo contenido naturalmente en $\mathbb{Q}(\sqrt{d})$. El anillo $\mathbb{Z}(\sqrt{d})$ es un dominio de integridad pero no es un dominio de factorización única. Por ejemplo, $(5 - 2\sqrt{5})(5 + 2\sqrt{5}) = 5 = \sqrt{5}\sqrt{5}$, $(6 + 3\sqrt{3})(6 - 3\sqrt{3}) = 9 = 3 \cdot 3$.

Claramente los números $a + b\sqrt{d}$, $a - b\sqrt{d}$ son las raíces del polinomio entero

$$x^2 - 2ax + (a^2 - b^2 d) = 0.$$

Nótese que podemos permitir $d < 0$ en toda la discusión anterior.

1.4.4 El cuerpo de los números complejos

Consideremos el conjunto $\mathbb{Z} \times \mathbb{Z}$ con las operaciones:

$$(m, n) + (r, s) = (m + r, n + s),$$

$$(m, n) \cdot (r, s) = (mr - ns, ms + nr), \quad m, n, r, s \in \mathbb{Z}.$$

Con estas operaciones $\mathbb{Z} \times \mathbb{Z}$ es un anillo conmutativo con identidad. El elemento neutro respecto de la suma es $(0, 0)$ y la identidad respecto al producto es $(1, 0)$.

Proposición 1.4.1 *El elemento* $(0, 1)$ *es una raíz del polinomio* $x^2 + 1 \in \mathbb{Z}[x]$.

Demostración. En efecto $(0, 1)^2 = (0, 1) \cdot (0, 1) = (-1, 0) = -1$. QED

Si denotamos $(0, 1)$ como $\sqrt{-1}$, vemos inmediatamente que $\mathbb{Z} \times \mathbb{Z}$ se puede identificar con $\mathbb{Z}(\sqrt{-1}) = \{m + n\sqrt{-1} \mid m, n \in \mathbb{Z}\}$. Análogamente se pueden definir los enteros Gaussianos $\mathbb{Z}(\sqrt{-d})$, $d > 0$.

El anillo $\mathbb{Z}(\sqrt{-1})$ está obviamente contenido en el cuerpo $\mathbb{Q}(\sqrt{-1})$ y éste a su vez en el cuerpo $\mathbb{R}(\sqrt{-1}) = \{a + b\sqrt{-1} \mid a, b \in \mathbb{R}\}$.

Definición 1.4.3 *Llamaremos cuerpo de los números complejos* $\mathbb{C}$ *al cuerpo* $\mathbb{R}(\sqrt{-1})$ *extensión del cuerpo de los números reales* $\mathbb{R}$. *El elemento* $\sqrt{-1}$ *se denota tradicionalmente por* i, *así que un elemento* z *de* $\mathbb{C}$ *se escribirá como* $z = a + ib$, $a, b \in \mathbb{R}$. *El número* a *se llama la parte real de* a *y se denota por* $\operatorname{Re} z$, *asimismo el número* b *se llama la parte imaginaria de* z *y se denota por* $\operatorname{Im} z$. *Las operaciones de suma y producto en* $\mathbb{C}$ *quedan por tanto escritas como: si* $z = a + ib$, $z' = a' + ib'$,

$$z + z' = (a + a') + i(b + b'), \quad z \cdot z' = (aa' - bb') + i(ab' + a'b).$$

El elemento neutro para la suma es 0 y el neutro para el producto es 1. Tal y como ya indicamos en los cuerpos gaussianos, el inverso de un número complejo no nulo z, tiene la expresión:

$$z^{-1} = \frac{a}{a^2 + b^2} - i\frac{b}{a^2 + b^2}.$$

Un automorfismo de un cuerpo $\mathbb{K}$ es un isomorfismo $\varphi \colon \mathbb{K} \to \mathbb{K}$, esto es, $\varphi(x + y) = \varphi(x) + \varphi(y)$, $\varphi(xy) = \varphi(x)\varphi(y)$, $\forall x, y \in \mathbb{K}$.

Definimos $\varphi \colon \mathbb{C} \to \mathbb{C}$ como $\varphi(a + ib) = a - ib$. Habitualmente se denota $\varphi(z) = \bar{z}$, (o también z^*). $\bar{z}$ se llama complejo conjugado de z. Claramente φ es un automorfismo de $\mathbb{C}$. Nótese que $\varphi(i) = -i$.

Ejercicio 1.4.5 $\mathbb{C}$ tiene solamente dos automorfismos, el automorfismo identidad y φ. Lo mismo ocurre con los cuerpos de números Gaussianos $\mathbb{Q}(\sqrt{d})$.

Ejercicio 1.4.6 Un cuerpo primo no posee más automorfismos que la identidad.

Llamaremos módulo de un número complejo al único número real positivo $|z|$ tal que $|z|^2 = z\bar{z}$.

Representaciones de los números complejos

Un número complejo $z = a + ib$ se puede representar como el punto (a, b) del plano real $\mathbb{R}^2$, definimos de esta manera una aplicación $\mathbb{C} \to \mathbb{R}^2$, $z \mapsto (\operatorname{Re} z, \operatorname{Im} z)$. El módulo del número complejo z se corresponde con la norma del vector de coordenadas (a, b) ya que $\|(a, b)\| = \sqrt{a^2 + b^2} = |z|$. La suma de números complejos corresponde a la suma de vectores en el plano. El producto por contra tiene una interpretación geométrica menos evidente.

Representación polar de los números complejos. Un punto del plano $(a, b) \neq (0, 0)$ queda unívocamente determinado por su norma y el ángulo que forma con un eje arbitrario, por ejemplo con el eje OX. Así $r = |z|$, y $\tan\theta = b/a$; θ se llama argumento de z y se denotará $\theta = \arg z$, $\theta \in [0, 2\pi)$. En otras palabras

$$r^2 = (\operatorname{Re} z)^2 + (\operatorname{Im} z)^2, \quad \tan\theta = \frac{\operatorname{Im} z}{\operatorname{Re} z},$$

y la inversa

$$\operatorname{Re} z = r\cos\theta, \quad \operatorname{Im} z = r \operatorname{sen}\theta.$$

Nótese de nuevo que la correspondencia $z \mapsto (r, \theta)$ no está bien definida para todo z (para $z = 0$ no lo está).

Representación trigonométrica de los números complejos. Hemos obtenido así una nueva representación de los números complejos ($\neq 0$),

$$z = r\cos\theta + ir\operatorname{sen}\theta.$$

Si $w = s\cos\phi + is\operatorname{sen}\phi$, tenemos $z \cdot w = (rs)\cos(\theta + \phi) + i(rs)\operatorname{sen}(\theta + \phi)$. Por tanto en la representación polar, la multiplicación de números complejos corresponde al producto de sus módulos y la suma de sus argumentos (módulo 2π).

Estas propiedades y expresiones se vuelven más transparentes si utilizamos la función exponencial. La definición precisa de la función exponencial requiere nociones de análisis que no corresponden a este curso. En cualquier caso definimos e^z como:

1. $e^z = \lim_{n\to\infty} \left(1 + \frac{z}{n}\right)^n$.
2. $e^z = \sum_{n=0}^{\infty} \frac{z^n}{n!}$.
3. $e^z = e^x(\cos y + i\operatorname{sen} y)$ donde $z = x + iy$.

Las propiedades mas importantes de la función exponencial son:

1. $e^z \cdot e^w = e^{z+w}$.
2. $e^0 = 1$.
3. $\overline{e^z} = e^{\bar{z}}$.

Ejercicio 1.4.7 Probar las propiedades 1, 2 y 3 utilizando la definición 3 de la función exponencial.

Consecuencia inmediata de lo anterior son las siguientes fórmulas:

$$e^{i\theta} = \cos\theta + i\operatorname{sen}\theta, \quad z = |z|e^{i\arg z}.$$

Por tanto $z \cdot w = |z| \cdot |w| e^{i(\arg z + \arg w)}$. En particular $z^n = r^n e^{in\theta}$. Nótese que

$$e^{2\pi i} = 1,$$

y en general, $e^{2\pi in} = 1$. Más todavía, $e^{i\theta} = 1$ si y sólo si $\cos\theta = 1$ y $\operatorname{sen}\theta = 0$, esto es, si y sólo si $\theta = 2\pi n$, $n \in \mathbb{Z}$.

1.4.5 Raíces n-ésimas de la unidad

Tratemos de resolver la ecuación $z^n = 1$ en el cuerpo de los números complejos. Una solución a dicha ecuación es ciertamente todo número complejo $z_0 \neq 0$ tal que $z_0^n = 1$. Por tanto si $z = re^{\theta}$ tenemos que

$$z_0^n = r^n e^{in\theta},$$

es decir $r^n = 1$ y $e^{in\theta} = 1$. Por lo tanto $r = 1$ y $n\theta = 2\pi k$, $k \in \mathbb{Z}$, $\theta = 2\pi k/n$, $k = 0, \pm 1, \pm 2, \ldots$. El argumento θ de z_0 tiene que estar en $[0, 2\pi)$ y así los valores de k para los que esto ocurre son $k = 0, 1, 2, \ldots, n-1$. Las soluciones de la ecuación $z^n - 1 = 0$ serán por tanto:

$$z_0 = 1, \quad z_1 = e^{2\pi i/n}, \quad z_2 = e^{4\pi i/n}, \ldots, \quad z_{n-1} = e^{2\pi(n-1)/n}.$$

Dichas raíces determinan un polígono regular de n lados inscrito en el círculo de radio 1 ya que todas ellas tienen módulo 1.

Ejercicio 1.4.8 Hallar las raíces del polinomio $1 + x + x^2 + \cdots + x^n$

Si multiplicamos dos raíces n-ésimas entre sí obtenemos:

$$z_k \cdot z_l = e^{2\pi(k+l)i/n}.$$

Por tanto vemos que si $k + l \geq n$, entonces $k + l = rn + s$, $0 \leq s < n$ y $z_k \cdot z_l = e^{2\pi is/n}$ y así $z_k \cdot z_l = z_s$ con $s \equiv k + l \pmod{n}$. Es por tanto conveniente etiquetar las raíces n-ésimas de la unidad con las clases de congruencia de números enteros módulo n, $[0], [1], \ldots, [n-1]$ obteniendo así la hermosa fórmula:

$$z_{[k]} \cdot z_{[l]} = z_{[k+l]},$$

que nos dice que la aplicación $[k] \mapsto z_{[k]}$ es un isomorfismo entre el grupo $(\mathbb{Z}_n, +)$ y el grupo de raíces n-ésimas de la unidad.

Ejercicio 1.4.9 Probar que si $n \mid m$, el conjunto de raíces n-ésimas de la unidad es un subgrupo del conjunto de raíces m-ésimas de la unidad.

Ejercicio 1.4.10 Probar que el conjunto de números complejos de módulo 1 es un grupo con respecto al producto. ¿Tiene este grupo otros subgrupos aparte de las raíces n-ésimas de la unidad?

1.5 Polinomios

1.5.1 El anillo de los polinomios

Al igual que construimos el anillo de los polinomios con coeficientes enteros $\mathbb{Z}[x]$ en el ejemplo 1.3.3, es posible extender tal construcción a un anillo cualquiera A y utilizarlo como anillo de coeficientes de los polinomios. De ahora en adelante "anillo" significará "anillo conmutativo con identidad". Al igual que entonces en el ejemplo 1.3.3, x denotará un símbolo abstracto y $x^2 = xx$, $x^3 = xxx$, etc. El conjunto $A[x]$ está formado por las expresiones $P(x) = a_0 + a_1x + a_2x^2 + \cdots + a_nx^n$, donde $a_i \in A$, $i = 1, \ldots, n$. Cada elemento de $A[x]$ se llamará polinomio en x con coeficientes en A. Dado un polinomio $P(x) = \sum_{k=0}^n a_nx^n$, llamaremos grado del polinomio P al número natural $n = \max\{k \mid a_k \neq 0\}$ y lo denotaremos ∂P. Un polinomio constante no nulo tiene grado cero. El término a_nx^n tal que $n = \partial P$ se llama dominante. Un polinomio se llama mónico (o unitario) si el coeficiente del término dominante es 1.

Proposición 1.5.1 *Propiedades del grado. Si A es un dominio de integridad se verifica:*
i. $\partial(PQ) = \partial P + \partial Q$.
ii. $\partial(P + Q) \leq \max(\partial P, \partial Q)$.

En $A[x]$ se definen de manera natural dos operaciones $+$, $\cdot$ como en el ejemplo 1.3.3: si $P(x) = \sum_i a_ix^i$, $Q(x) = \sum_j b_jx^j$, entonces,

$$P(x) + Q(x) = \sum_k (a_k + b_k)x^k, \quad P(x)Q(x) = \sum_k \left(\sum_{i+j=k} a_ib_j \right) x^k.$$

El anillo $A[x]$ posee como subanillo al propio anillo A (los elementos de A se identifican con los polinomios de grado cero).

Ejercicio 1.5.1 Considérese el conjunto S de sucesiones infinitas $(a_0, a_1, \ldots, a_n, \ldots)$ de elementos de A tales que todos sus términos excepto un número finito son 0. En este conjunto introducimos dos operaciones

$$(a_0, a_1, \ldots, a_n, \ldots) + (b_0, b_1, \ldots, b_n, \ldots) = (a_0 + b_0, a_1 + b_1, \ldots, a_n + b_n, \ldots),$$

$$(a_0, a_1, \ldots, a_n, \ldots) \cdot (b_0, b_1, \ldots, b_n, \ldots) = (a_0 b_0, a_0 b_1 + a_1 b_0, \ldots, a_n b_0 + a_{n-1} b_1 + \cdots + a_0 b_n, \ldots).$$

$(S, +, \cdot)$ es un anillo. Probar que la correspondencia $(a_0, a_1, \ldots, a_n, \ldots) \mapsto P(x) = \sum_{k \geq 0} a_k x^k$ es un isomorfismo de anillos. Nótese que $(0, 1, 0, \ldots) \mapsto x$.

Ejemplo 1.5.1 De acuerdo con lo anterior, además de $\mathbb{Z}[x]$, tenemos los anillos $\mathbb{Q}[x]$, $\mathbb{R}[x]$, $\mathbb{C}[x]$, así como $\mathbb{Z}_n[x]$, etc.

Ejercicio 1.5.2 Si A es un dominio de integridad, entonces $A[x]$ es un dominio de integridad.

1.5.2 Divisibilidad en el anillo de polinomios

La noción de grado de un polinomio permite extender la teoría de divisibilidad de números enteros a los anillos de polinomios. Un anillo poseyendo una aplicación $\delta \colon A \to \mathbb{Z}^+$ con las propiedades del grado descritas en la proposición 1.5.1 se llama un dominio Euclídeo. Nos concentraremos en las propiedades de divisibilidad del anillo de polinomios sobre un cuerpo $\mathbb{K}$.

Sean $P, Q \in \mathbb{K}[x]$, diremos que P divide a Q si existe $R \in \mathbb{K}[x]$ tal que $Q = PR$. Las unidades del anillo $\mathbb{K}[x]$, esto es sus elementos invertibles, son los polinomios constantes no nulos. Un polinomio se dirá irreducible si sus únicos divisores son las unidades de $\mathbb{K}[x]$ y él mismo multiplicado por unidades. La noción de irreducible es equivalente a la noción de número primo.

Ejercicio 1.5.3 Probar que en el anillo $\mathbb{Z}_4[x]$ hay polinomios invertibles no constantes.

El anillo $\mathbb{K}[x]$ posee la propiedad de factorización única, esto es, todo polinomio $P(x)$ se escribe de manera única como

$$P(x) = a P_1(x) \ldots P_r(x),$$

donde $a \in IK^*$, $P_i(x)$, $i = 1, \ldots, r$ son polinomios irreducibles.

Para establecer este resultado, probaremos en primer lugar la extensión al anillo $\mathbb{K}[x]$ del algoritmo de la división.

Teorema 1.5.1 *Algoritmo de la división. Para todo par de polinomios $P(x), Q(x) \in \mathbb{K}[x]$, $Q(x) \neq 0$, existen dos polinomios $D(x)$, $R(x)$ tales que*

$$P(x) = D(x)Q(x) + R(x),$$

con $\partial R(x) < \partial Q(x)$. Además dichos polinomios son únicos.

Demostración. Existencia. Si $Q(x)$ divide a $P(x)$ el resultado es inmediato. Supongamos que no es así. Consideremos el conjunto de los números enteros positivos,

$$S = \{\partial(P(x) - D(x)Q(x)) \mid D(x) \in \mathbb{K}[x]\}.$$

El conjunto S es no vacío y por el teorema 1.3.1 existe $r = \min S$. Sea entonces $D(x)$ un polinomio tal que $\partial(P(x) - D(x)Q(x)) = r$. Entonces $P(x) = D(x)Q(x) + R(x)$ y $\partial R(x) = r$. Necesariamente $r < \partial Q(x)$ ya que si $r \geq \partial Q(x)$, entonces el término dominante de $R(x)$ será de la forma ax^r y el de $Q(x)$, bx^m con $r \geq m$. Pero el polinomio $\hat{D}(x) = b^{-1} a x^{r-m}$ es tal que $R(x) - \hat{D}(x)Q(x)$ tiene grado menor que r ya que su término de orden r sería $ax^r - b^{-1} a b x^{r-m} x^m = 0$. Por tanto $P(x) - D(x)Q(x) - \hat{D}(x)Q(x) = P(x) - (D(x) + \hat{D}(x))Q(x) = R(x) - \hat{D}(x)Q(x)$ tiene grado menor que r.

Unicidad. Supongamos ahora que $D(x)$ y $R(x)$ no son únicos, esto es, existen $\hat{D}(x)$ y $\hat{R}(x)$ tales que $P(x) = \hat{D}(x)Q(x) + \hat{R}(x)$, $\partial \hat{R}(x) < \partial Q(x)$. Entonces, $(D(x) - \hat{D}(x))Q(x) = \hat{R}(x) - R(x)$, pero entonces,

$$\partial(D - \hat{D}) + \partial Q = \partial(\hat{R} - R) \leq \max(\partial \hat{R}, \partial R) < \partial Q.$$

Lo cual es imposible a no ser que $D - \hat{D} = 0$. En cuyo caso además $R = \hat{R}$. QED

Máximo común divisor de polinomios

Repetiremos gran parte de la línea argumental que desarrollamos al definir el mínimo común múltiplo y máximo común divisor de números enteros.

Proposición 1.5.2 *Para todo ideal $I \neq 0$ de $\mathbb{K}[x]$ existe un polinomio $P(x)$ tal que $I = \{Q(x)P(x) \mid Q(x) \in \mathbb{K}[x]\} = (P)$.*

Demostración. Sea $I \neq 0$ un ideal de $\mathbb{K}[x]$ y $S = \{r = \partial R(x) \in \mathbb{Z} \mid R(x) \in I\}$. Es evidente que $S \neq \emptyset$ y por tanto tomemos el elemento mínimo $r_0 \geq 0$ de dicho conjunto. Si $r_0 = 0$, entonces hay un polinomio constante $R(x) = a$ en I, pero $a \neq 0$, y por tanto $R(x)$ es invertible y entonces $I = \mathbb{K}[x]$. Por tanto $I = (1)$. Supongamos por tanto que $r_0 > 0$. Sea $P_0(x) \in I$ tal que $\partial P_0 = r_0$. Supongamos que existe $Q(x) \in I$ tal que $Q \neq RP_0$, entonces por el algoritmo de la división existe $D(x)$ y $R(x)$ tal que $Q = DP_0 + R$ con $0 \leq \partial R < \partial P_0 = r_0$. Pero entonces $P = Q - DP_0 \in I$ y su grado es menor que r_0, lo que es absurdo. QED

Dados dos polinomios P, Q, consideremos todos los polinomios de la forma $MP + NQ$, $M, N \in \mathbb{K}[x]$, denotemos tal conjunto por J. Claramente J es un ideal y por la proposición anterior sabemos que existe un polinomio D tal que $J = (D)$. Diremos que D es el máximo común divisor de P y Q y se denotará por (P, Q) o m.c.d.(P, Q). Una consecuencia inmediata de la definición de máximo común divisor es la siguiente propiedad.

Corolario 1.5.1 *Sean $P, Q \in \mathbb{K}[x]$ y $D = (P, Q)$, entonces existen dos polinomios $M, N \in \mathbb{K}[x]$ tales que $D = MP + NQ$.*

Ejercicio 1.5.4 Si $D \mid P$ y $D \mid Q$ entonces $D \mid (P, Q)$. Concluir de aquí que si P es irreducible y P no divide a Q entonces $(P, Q) = 1$.

Un ideal de un anillo formado por los múltiplos de un elemento dado se llama principal. Un anillo tal que todos sus ideales están formados por los múltiplos de un elemento se llama anillo de ideales principales. Si además es un dominio de integridad se llama un dominio de ideales principales. Tanto $\mathbb{Z}$ como $\mathbb{K}[x]$ son por tanto dominios de ideales principales.

Ejercicio 1.5.5 Probar el algoritmo de Euclides para polinomios. Esto es, si $P(x), Q(x) \in \mathbb{K}[x]$, procedemos iterativamente y construimos:

$$\begin{aligned} P(x) &= D_0(x)Q(x) + R_0(x); \partial R_0 < \partial Q, \\ D_0(x) &= D_1(x)R_0(x) + R_1(x); \partial R_1 < \partial R_0, \\ D_1(x) &= D_2(x)R_1(x) + R_2(x); \partial R_2 < \partial R_1, \\ &\cdots \\ D_{n-1}(x) &= D_n(x)R_{n-1}(x), \end{aligned}$$

y $R_n = 0$. Entonces R_{n-1} es el m.c.d.(P, Q).

La unicidad de la factorización de un polinomio en factores irreducibles se sigue del siguiente Lema.

Lema 1.5.1 *Si $P(x) \in \mathbb{K}[x]$ es un polinomio irreducible y $P(x) \mid A(x)B(x)$ entonces o $P(x) \mid A(x)$ o bien $P(x) \mid B(x)$.*

Demostración. Supongamos que $P \mid AB$ y P no divide ni a A ni a B. Como P no divide a A y P es irreducible entonces $(P, A) = 1$, por tanto existen polinomios M, N tales que $PM + AN = 1$. Multiplicamos la anterior ecuación por B y obtenemos que

$$PMB + ANB = B,$$

y como $P \mid AB$, entonces $P \mid B$ lo cual es absurdo. QED

Teorema 1.5.2 *Todo polinomio $P(x) \in \mathbb{K}[x]$ con $\partial P \geq 1$ posee una descomposición única en producto de factores irreducibles salvo producto por unidades.*

Demostración. Probémoslo por inducción sobre el grado del polinomio. Supongamos que $\partial P = 1$. Entonces $P(x) = a + bx$ y P es irreducible. Supongamos a continuación que la hipótesis es cierta para todo k menor que n y probémoslo para $k = n$. Sea por tanto P un polinomio de grado n. Si P no posee divisores no triviales, es irreducible y ya está. Si P posee un divisor no trivial D_1 tendremos $P = P_1 D_1$ y $\partial D_1 < n$, $\partial P_1 < n$, por tanto por hipótesis de inducción, tanto D_1 como P_1 factorizan como producto de factores irreducibles. Por tanto P factoriza como producto de factores irreducibles.

Unicidad. Supongamos que $P(x) = P_1(x) \cdots P_r(x) = Q_1(x) \cdots Q_s(x)$ son dos descomposiciones en factores irreducibles de $P(x)$. Tomemos un factor P_i de la primera descomposición, entonces $P_i \mid Q_1 \cdots Q_s$ y por tanto por el Lema 1.5.1 P_i debe dividir a algún factor Q_j, pero P_i es irreducible y por tanto $P_i = Q_j$ excepto posiblemente una unidad. Repitiendo el proceso para todos los P_i se completa la demostración. QED

1.5.3 Raíces de polinomios y completitud algebraica

Sea $P(x) \in \mathbb{K}[x]$ un polinomio arbitrario $P(x) = a_0 + a_1 x + \cdots + a_n x^n$, $a_i \in \mathbb{K}$. Un elemento $\alpha \in \mathbb{K}$ se dirá que es una raíz de $P(x)$ si $a_0 + a_1\alpha + a_2\alpha^2 + \cdots + a_n\alpha^n = 0$, esto es si $P(\alpha) = 0$.

Teorema 1.5.3 *α es una raíz de $P(x)$ si y sólo si $(x - \alpha) \mid P(x)$.*

Demostración. Sea α una raíz de $P(x)$. Dividamos $P(x)$ por $(x - \alpha)$. Entonces $P(x) = Q(x)(x - \alpha) + R(x)$ y $0 \leq \partial R(x) < 1$ y por tanto $R(x)$ debe ser constante o cero. Por otro lado evaluando la anterior igualdad en α obtenemos que $R(\alpha) = 0$ y por tanto $R(x) = 0$. QED

Consideremos un cuerpo $\mathbb{K}$ y un subcuerpo $\mathbb{F} \subset \mathbb{K}$. Un elemento $\alpha \in \mathbb{K}$ se dirá algebraico sobre $\mathbb{F}$ si es raíz de algún polinomio $P(x) \in \mathbb{F}[x]$. Un elemento α se dirá transcendente sobre $\mathbb{F}$ si no es algebraico. Un cuerpo $\mathbb{K}$ se dirá algebraicamente cerrado si todos los elementos algebraicos sobre $\mathbb{K}$ en una extensión cualquiera de $\mathbb{K}$ están en $\mathbb{K}$. Los cuerpos $\mathbb{Q}$ y $\mathbb{R}$ no son algebraicamente cerrados.

Teorema 1.5.4 *Todo polinomio $P(x) \in \mathbb{C}[x]$ de grado mayor o igual a 1 posee al menos una raíz.*

Demostración. La demostración de este teorema no es puramente algebraica.

Teorema 1.5.5 *Todo polinomio $P(x) \in \mathbb{C}[x]$ factoriza como producto de factores de grado 1, esto es:*

$$P(x) = a(x - \alpha_1)(x - \alpha_2) \cdots (x - \alpha_n),$$

donde $a, \alpha_i \in \mathbb{C}$, $i = 1, \ldots, n = \partial P$ son las raíces de $P(x)$.

Demostración. Por el teorema de factorización de polinomios, Teorema 1.5.2, todo polinomio $P(x) \in \mathbb{C}[x]$ factoriza como producto de polinomios irreducibles. Veamos que todo polinomio irreducible sobre $\mathbb{C}$ es de grado 1. Supongamos que $P(x)$ es irreducible y de grado ≥ 1, entonces por el Teorema 1.5.4 $P(x)$ posee una raíz α, pero entonces $P(x) = (x - \alpha)Q(x)$ por el teorema 1.5.3 y $P(x)$ no es irreducible. QED

Ejercicio 1.5.6 Determinar si es cierta o falsa la siguiente proposición. Si $\mathbb{K}$ es algebraicamente cerrado y $P(x) \in \mathbb{K}[x]$ es irreducible, entonces $\partial P(x) = 1$.

Corolario 1.5.2 *Teorema fundamental del álgebra. $\mathbb{C}$ es algebraicamente cerrado.*

Demostración. Sea $P(x)$ un polinomio sobre $\mathbb{C}$. Por el teorema anterior, Teorema 1.5.5, factoriza como producto de factores de grado uno. Por lo tanto todos los elementos algebraicos sobre $\mathbb{C}$, esto es, raíces de polinomios sobre $\mathbb{C}$ están en $\mathbb{C}$. QED

Tema 2

Espacios vectoriales

Espacios vectoriales. Subespacios. Sistemas de generadores. Dependencia e independencia lineal. Bases. Matrices.

2.1 Definiciones

Veamos algunos ejemplos para introducir la noción de espacio vectorial.

Ejemplo 2.1.1 Consideremos el conjunto de los números reales. En él hay definidas dos operaciones, la suma, respecto de la cual es un grupo, y el producto.

Ejemplo 2.1.2 Sea ahora V el conjunto de los pares de números reales, (x, y), donde $x, y \in \mathbb{R}$. Podemos definir la suma de dos elementos de este conjunto en la manera usual:

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

Definimos también el producto de un número real por un par:

$$\lambda(x, y) = (\lambda x, \lambda y)$$

Ejemplo 2.1.3 Sea $V = \mathbb{K}^n$, donde $\mathbb{K}$ es un cuerpo, es decir, el espacio de n-uplas de escalares en $\mathbb{K}$. Con las operaciones obvias de suma y producto por escalares (ver ejemplo anterior), este espacio tiene propiedades similares a las que exhiben los ejemplos anteriores.

Ejemplo 2.1.4 Sea $C[0, 1]$ el conjunto de las funciones continuas definidas en el intervalo $[0, 1]$ de la recta real con valores en $\mathbb{R}$. La suma de funciones continuas es una función continua. La función que se obtiene al multiplicar una función continua por un número real, es de nuevo una función continua.

Ejemplo 2.1.5 Sea $\mathbb{R}[x]$ el conjunto de polinomios en la variable x. Como la suma de dos polinomios es otro polinomio, y el producto de un número real por un polinomio es también un polinomio, estamos en una situación similar a la de los ejemplos anteriores. De hecho, $\mathbb{R}[x]$ es, en cierto sentido, un subconjunto de $C(\mathbb{R})$.

Ejemplo 2.1.6 Sea $\mathbb{R}_n[x]$ el conjunto de polinomios en la variable x de grado menor o igual a $n \in \mathbb{N}$. Está claro que las mismas propiedades que vimos en los ejemplos anteriores aparecen de nuevo aquí.

Ejemplo 2.1.7 Sea V el conjunto de funciones continuas en $\mathbb{R}$ tales que $f(0) = 1$. Es fácil ver que estamos en un caso diferente. Ahora la suma de dos funciones en V no está en V. Si $f, g \in V$, $f(0)+g(0) = 2$, luego $f + g$ no está en V.

Ejemplo 2.1.8 Supongamos ahora que el conjunto V es el formado por las funciones $f(x)$ que verifican la siguiente ecuación:

$$\frac{d^2 f}{dx^2} = f(x) \operatorname{sen} x$$

En este caso no tenemos, al menos por ahora, una idea clara de cuales son los elementos de este conjunto. En los ejemplos precedentes podíamos construir de manera explícita elementos del conjunto en cuestión. Aquí solo sabemos que se trata de funciones que se pueden derivar dos veces (digamos que están en el conjunto $C^2(\mathbb{R})$), y que su segunda derivada es el producto de la función seno por la función de partida. Pues bien, a pesar de esta falta de información, la suma de dos de estas funciones verifica la ecuación, y el producto por un número real de cualquier función de V es también una función de V.

Los anteriores ejemplos son casos particulares de una situación general que pasamos a definir con precisión.

Definición 2.1.1 *Un espacio vectorial sobre un cuerpo $\mathbb{K}$ (los elementos de $\mathbb{K}$ se llamarán escalares) es un conjunto V (cuyos elementos se llamarán vectores) dotado de dos operaciones. Una de ellas interna (suma):*

$$+ \colon V \times V \longrightarrow V$$

respecto de la que V es un grupo conmutativo. Y una operación externa, el producto por escalares:

$$\cdot \colon \mathbb{K} \times V \longrightarrow V$$

que verifica:

1. $(\lambda + \mu)v = \lambda v + \mu v$,
2. $\lambda(u + v) = \lambda u + \lambda v$,
3. $\lambda(\mu v) = (\lambda\mu)v$,
4. $1v = v$, *donde* $u, v \in V$, $\lambda, \mu \in \mathbb{K}$ *y* 1 *es la unidad en* $\mathbb{K}$.

Es muy sencillo comprobar que todos los ejemplos anteriores son espacios vectoriales reales (sobre el cuerpo $\mathbb{R}$), salvo el ejemplo 2.1.7. La mayor parte de sus propiedades se derivan de propiedades similares sobre los números reales. Se tiene:

Teorema 2.1.1 *Todo cuerpo es un espacio vectorial sobre sí mismo.*

Demostración. En efecto, las dos operaciones son las que tiene el cuerpo y el producto por escalares se confunde con la propia operación interna de multiplicación del cuerpo. QED

> Nota. El mismo concepto se puede definir sobre un anillo. Se dice en este caso que se tiene un módulo. Debido a que el anillo no es conmutativo en general, es preciso especificar si la multiplicación externa es por la derecha o por la izquierda. Debido a que, en general, no tendremos un elemento inverso respecto a la multiplicación, las propiedades de los módulos son distintas de las de los espacios vectoriales. En este curso no insistiremos en la idea de módulo.

Ejemplo 2.1.9 Consideremos la ecuación que describe a un oscilador armónico (por ejemplo un muelle que verifica la ley de Hooke, con constante de recuperación k). El movimiento de la masa m sujeta al muelle viene descrito por una función $x(t)$ que da la posición en función del tiempo. De las leyes de la dinámica newtoniana se deduce inmediatamente que $x(t)$ verifica lo que se llama una ecuación diferencial lineal de segundo orden:

$$\frac{d^2 x}{dt^2} + \omega^2 x = 0$$

con $\omega^2 = k/m$. Las soluciones de esta ecuación forman un espacio vectorial, por un razonamiento semejante al que hicimos en el ejemplo 2.1.8. Desde un punto de vista del movimiento, lo que estamos diciendo es que la superposición (lineal) de dos movimientos del oscilador armónico es otro movimiento de este tipo. Todos los movimientos del oscilador armónico se obtienen por superposición de dos básicos, los dados por las funciones $\operatorname{sen} \omega t$ y $\cos \omega t$.

Los modelos lineales como el anterior son fundamentales en Física. No todo fenómeno que ocurre en la Naturaleza es lineal, y de hecho, los no lineales constituyen una clase muy importante. Pero incluso en estos casos, las aproximaciones lineales proporcionan muchas veces información valiosa sobre el fenómeno en cuestión.

Consecuencia de las operaciones definidas en un espacio vectorial es la siguiente, una herramienta fundamental en el estudio de los espacios vectoriales.

Definición 2.1.2 *Sean $x_1, \dots, x_n$ elementos de un espacio vectorial y $\lambda_1, \dots, \lambda_n$ escalares del cuerpo $\mathbb{K}$. Se llama combinación lineal de los vectores $x_1, \dots, x_n$ con coeficientes $\lambda_1, \dots, \lambda_n$ al vector:*

$$\sum_{i=1}^{n} \lambda_i x_i = \lambda_1 x_1 + \ldots + \lambda_n x_n$$

Obviamente, toda combinación lineal está contenida en el espacio. Téngase en cuenta que una combinación lineal es una suma finita con coeficientes en el cuerpo. Posibilidades de sumas con infinitos sumandos, o de otras con un número finito de coeficientes no nulos, aunque en cantidad variable, llevan a conceptos más avanzados de álgebra en los que no entraremos (sumas y productos directos con un número arbitrario de factores).

Ejemplo 2.1.10 Supongamos en $\mathbb{R}^3$ los vectores $v_1 = (1,0,0)$, $v_2 = (0,1,0)$ y $v_3 = (0,0,1)$. Una combinación lineal de estos tres vectores con coeficientes $\lambda_1, \lambda_2, \lambda_3 \in \mathbb{R}$ es:

$$\lambda_1(1,0,0) + \lambda_2(0,1,0) + \lambda_3(0,0,1) = (\lambda_1, \lambda_2, \lambda_3)$$

de donde resulta que cualquier vector de $\mathbb{R}^3$ se puede poner como combinación lineal de estos tres vectores, hecho que determinará buena parte de las propiedades de este espacio vectorial.

Ejemplo 2.1.11 Si en el espacio de los polinomios en una variable con coeficientes reales, seleccionamos cualquier familia finita del tipo $1, x, x^2, \dots, x^n$, las combinaciones lineales de estos elementos no cubren todo el espacio, por muy grande que hagamos n.

2.2 Subespacios

Hemos visto en el ejemplo 2.1.5, como los polinomios formaban un espacio vectorial real, y como las funciones continuas (en todo $\mathbb{R}$) son también un espacio vectorial. Puesto que los polinomios se pueden interpretar como funciones continuas, tenemos un espacio vectorial contenido en otro. La situación se presenta con mucha frecuencia y se encuentra descrita en la siguiente definición:

Definición 2.2.1 *Sea V un espacio vectorial sobre un cuerpo $\mathbb{K}$ y sea W un subconjunto de V no vacío. Se dice que W es un subespacio vectorial de V si:*
i. $u - v \in W$, $\forall u, v \in W$,
ii $\lambda u \in W$, $\forall \lambda \in \mathbb{K}$.

Proposición 2.2.1 *Si W es un subespacio vectorial de V entonces el conjunto W con las operaciones $+$ y $\cdot$, inducidas de la suma y el producto por escalares de V, es un espacio vectorial.*

Ejercicio 2.2.1 Probar la proposición anterior 2.2.1

Ejemplo 2.2.1 Consideremos ahora el conjunto de los números complejos. Al ser $\mathbb{C}$ un cuerpo, es un espacio vectorial sobre sí mismo. El conjunto de los números reales es un subconjunto de $\mathbb{C}$. La pregunta es obvia. ¿Es $\mathbb{R}$ un subespacio vectorial de $\mathbb{C}$? La respuesta no lo es tanto. En efecto, como sabemos, $\mathbb{R}$ es un espacio vectorial sobre $\mathbb{R}$. Pero aquí estamos hablando de $\mathbb{R}$ como un subconjunto de $\mathbb{C}$, es decir, como los números complejos que tienen parte imaginaria igual a cero. La suma de dos de estos números es otro número del mismo tipo. Pero el producto de un número complejo arbitrario (un escalar de $\mathbb{C}$) por un número complejo de parte imaginaria cero (es decir, un número real) no es en general un número real. Por tanto $\mathbb{R}$ no es un subespacio vectorial del espacio vectorial complejo $\mathbb{C}$.

Ejemplo 2.2.2 El conjunto de los números complejos es un espacio vectorial real. No es difícil probar que se cumplen todas las propiedades del caso. Los reales siguen siendo un subconjunto de $\mathbb{C}$ que ahora es un subespacio vectorial (por supuesto real).

Ejemplo 2.2.3 Consideremos el espacio tridimensional $\mathbb{R}^3$. Se trata de un espacio vectorial sobre $\mathbb{R}$ cuyos elementos son las ternas de números reales:

$$(x_1, x_2, x_3),\ x_i \in \mathbb{R}, i = 1, 2, 3$$

El siguiente subconjunto es un subespacio vectorial de $\mathbb{R}^3$:

$$W = \{(x_1, x_2, 0) \mid x_1, x_2 \in \mathbb{R}\}$$

Pero el subconjunto de $\mathbb{R}^3$:

$$A = \{(x_1, x_2, 1) \mid x_1, x_2 \in \mathbb{R}\}$$

no lo es. Tampoco es un subespacio el conjunto:

$$S^2 = \{(x_1, x_2, x_3) \mid x_1^2 + x_2^2 + x_3^2 = 1,\ x_i \in \mathbb{R},\ i = 1, 2, 3\}$$

Se trata de la esfera unidad en $\mathbb{R}^3$. Los conjuntos como éste se llaman variedades, y, aunque no presenten características lineales, su estudio local implica la consideración de espacios vectoriales (en este caso de planos).

En todo espacio vectorial hay siempre dos subespacios, el espacio total y el vector cero (el elemento neutro del conjunto considerado como grupo abeliano). Pero puede no haber más.

Ejemplo 2.2.4 Los dos únicos subespacios del espacio vectorial real $\mathbb{R}$ son $\{0\}$ y $\mathbb{R}$. La demostración es la siguiente. Sea W un subespacio vectorial de $\mathbb{R}$. Entonces, si $0 \neq x \in W$, $yx \in W$ para todo número real y. Como x tiene inverso, se tiene:

$$F = \{xy \mid y \in \mathbb{R}\} = \mathbb{R}$$

Como $F \subset W$, se tiene: $W = \mathbb{R}$. Si en W solo tenemos el elemento 0, entonces: $W = \{0\}$.

Ejemplo 2.2.5 El conjunto de polinomios de grado menor o igual que $n \in \mathbb{N}$ es un subespacio propio (es decir distinto del $\{0\}$ y el total) del espacio vectorial de todos los polinomios.

Los subespacios vienen determinados de varias maneras. Hemos visto alguna de ellas, concretamente, en espacios del tipo $\mathbb{K}^n$, en el que una o varias de las componentes de los vectores son iguales a cero. Pero se puede hacer de otras formas.

Ejemplo 2.2.6 En $\mathbb{C}^n$ se considera el conjunto de vectores tales que:

$$\{(x_1, \dots, x_n) \mid \sum_{i=1}^{n} x_i = 0\}$$

Se trata de un subespacio propio de $\mathbb{C}^n$. También el subconjunto:

$$\{(x_1, \dots, x_n) \mid \sum_{i=1}^{n} x_i = 0, x_1 + x_n = 0\}$$

es otro subespacio, contenido en el anterior.

Esta forma de dar subespacios se suele llamar implícita. Pero se podrían definir de una forma explícita, es decir, dando las componentes de los vectores. Por ejemplo:

$$\{(x_1, \dots, x_n) \mid x_1 = \lambda_1, x_2 = \lambda_1 + \lambda_2, x_n = 0,\ \lambda_1, \lambda_2 \in \mathbb{C}\}$$

Como iremos viendo, las posibilidades son muchas.

2.3 Operaciones con subespacios

La familia de subespacios de un espacio vectorial admite una serie de operaciones que pasamos a detallar.

Teorema 2.3.1 *La intersección de subespacios de un espacio vectorial es un subespacio vectorial.*

Demostración. Sean W_1 y W_2 dos subespacios de V. Si x, y son dos elementos de la intersección, $W_1 \cap W_2$, ambos están en cada subespacio, luego la suma pertenece a ambos y por tanto a la intersección. El mismo argumento se aplica al producto por escalares. Nótese que la intersección de subespacios vectoriales nunca es vacía, pues al menos el vector 0 está en todos ellos. QED

Ejemplo 2.3.1 Consideremos el espacio vectorial real de las funciones continuas definidas en $\mathbb{R}$ con valores en $\mathbb{R}$, $C(\mathbb{R})$. El conjunto de polinomios con grado menor o igual a n (n un número natural fijado) es un subespacio vectorial como ya hemos dicho. El conjunto de las funciones continuas que se anulan en $x = 0$ es un subespacio vectorial de $C(\mathbb{R})$. La intersección de ambos, es decir, el conjunto de polinomios de grado menor o igual que n que se anulan en $x = 0$, es un subespacio vectorial de $C(\mathbb{R})$.

Sin embargo, la unión de subespacios vectoriales no es en general un subespacio vectorial. Pero podemos construir un subespacio de la siguiente forma.

Definición 2.3.1 *Sea S un subconjunto de un espacio vectorial V. Se llama espacio vectorial generado por S al menor de los subespacios de V que contienen a S.*

Está claro que dicho subespacio será la intersección de todos los subespacios que contienen a S. La intersección no puede ser vacía, pues S está en todos ellos, y al menos hay un subespacio que contiene a S que es el espacio total. Pero no es sencillo, en principio, calcular explícitamente este subespacio.

Ejemplo 2.3.2 Sea el subconjunto del conjunto de polinomios en la variable x: $S = \{x\}$, que obviamente no es un espacio vectorial. Pero está claro que $W = \{\lambda x \mid \lambda \in \mathbb{R}\}$ sí es un subespacio vectorial y contiene a S.

Definición 2.3.2 *Sea S un subconjunto de un espacio vectorial. La envolvente lineal de S, $\mathrm{lin}(S)$, es el conjunto de combinaciones lineales que se pueden formar con los elementos de S.*

Se tiene:

Teorema 2.3.2 *La envolvente lineal de un subconjunto de un espacio vectorial V es un espacio vectorial (subespacio del espacio vectorial V).*

La demostración es evidente.

Teorema 2.3.3 *El subespacio generado por un subconjunto de un espacio vectorial es la envolvente lineal $\mathrm{lin}(S)$, de este subconjunto.*

Demostración. Claramente S está contenido en $\mathrm{lin}(S)$. Sea W un subespacio que contiene a S. Entonces, debe contener también a la envolvente lineal, pues es un subespacio. Por lo tanto $\mathrm{lin}(S) \subset W$ para todo W subespacio que contiene a S. De donde $\mathrm{lin}(S) \subset \cap W$ donde la intersección se refiere a todos los subespacios que contiene a W. De aquí se concluye que el espacio generado por S es la envolvente lineal de S. QED

Ejemplo 2.3.3 El conjunto de matrices 2×2 con elementos complejos, es un espacio vectorial complejo. El subespacio generado por los elementos:

$$A = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad B = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

es la envolvente lineal de estos elementos:

$$\begin{pmatrix} \alpha & \beta \\ \beta & -\alpha \end{pmatrix}, \quad \alpha, \beta \in \mathbb{C}$$

Nótese que es un subespacio propio del espacio de matrices del que hablamos. Sin embargo, existen casos en los que la envolvente lineal de una familia es el espacio total.

Definición 2.3.3 *Se dice que el subconjunto S del espacio vectorial V es un sistema de generadores de V si la envolvente lineal de los elementos de S (es decir, el espacio generado por S) es el espacio total V.*

Ejemplo 2.3.4 En el espacio vectorial de polinomios, la familia $S = \{1, x, x^2, x^3, \ldots\}$ es un sistema de generadores.

Ejemplo 2.3.5 En el espacio de matrices 2×2 con coeficientes complejos, la familia:

$$S = \left\{ \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} i & i \\ i & -i \end{pmatrix} \right\}$$

es un sistema de generadores.

Está claro que todo subconjunto es un sistema de generadores de su envolvente lineal.

Con estas nociones definiremos la suma de subespacios. Como hemos dicho, la unión de subespacios no es necesariamente un subespacio.

Ejemplo 2.3.6 Consideremos en $\mathbb{R}^2$, los subespacios: $W_1 = \{(a, 0) \mid a \in \mathbb{R}\}$ y $W_2 = \{(0, a) \mid a \in \mathbb{R}\}$. La unión es el conjunto:

$$W_1 \cup W_2 = \{(a, 0), (0, b) \mid a, b \in \mathbb{R}\}$$

Pero esto no es un espacio vectorial. Pues si sumamos $(1, 0)$ y $(0, 1)$ que están en la unión, obtenemos $(1, 1)$ que no pertenece a la unión.

Definición 2.3.4 *Se define la suma de dos subespacios W_1 y W_2 de un espacio vectorial como la envolvente lineal de la unión de ambos subespacios:*

$$W_1 + W_2 = \mathrm{lin}(W_1 \cup W_2)$$

La definición anterior no es muy útil en muchos casos. Sin embargo, se tiene lo siguiente:

Teorema 2.3.4 *La suma de dos subespacios de un espacio vectorial es:*

$$W_1 + W_2 = \{x_1 + x_2 \mid x_1 \in W_1, x_2 \in W_2\}$$

Demostración. Puesto que $x_1 \in W_1$ y $x_2 \in W_2$, ambos están en la unión y por lo tanto su suma está en la envolvente lineal. De aquí se tiene la mitad de la igualdad:

$$\{x_1 + x_2 \mid x_1 \in W_1, x_2 \in W_2\} \subset W_1 + W_2$$

Además, cualquier vector de la envolvente es una combinación lineal de elementos de la unión. Por tanto, podemos separar los vectores que forman la combinación lineal y que pertenecen a W_1 por un lado y los que pertenecen a W_2 por otro. Como ambos W_1 y W_2 son subespacios vectoriales, llegamos a que cualquier elemento de la envolvente se puede poner como suma de un elemento de W_1 más otro de W_2. QED

En general, los elementos de la suma se pueden poner de varias formas como suma de un vector de W_1 y otro de W_2. Dicho de otra manera, la descomposición no es única. Pero a veces sí lo es.

Definición 2.3.5 *Se dice que la suma de dos subespacios vectoriales de un espacio vectorial es directa si cada elemento de la suma admite una única descomposición como suma de un elemento del primer subespacio más un elemento del segundo. Se escribe entonces: $W_1 \oplus W_2$*

La caracterización de sumas directas se puede hacer también de la forma siguiente:

Teorema 2.3.5 *Sean W_1 y W_2 dos subespacios del espacio vectorial V. Entonces, la suma de W_1 y W_2 es directa si y solo si $W_1 \cap W_2 = \{0\}$*

Demostración. La parte "si" se deduce fácilmente. Si la intersección es el vector 0, y tenemos dos descomposiciones para un vector v de la suma, $v = x_1 + x_2 = y_1 + y_2$, entonces: $x_1 - y_1 = y_2 - x_2$. Pero el primer vector está en W_1 y el segundo en W_2, luego ambos (que son iguales) están en la intersección, luego son cero. De aquí se deduce que $x_1 = y_1$ y $x_2 = y_2$, luego la descomposición es única y la suma es directa. El "solo si" se demuestra por: si v está en la intersección, está en ambos subespacios. Pero eso quiere decir que $v = v + 0$ es una descomposición válida y que $v = 0 + v$ también lo es. Como la descomposición es única al ser la suma directa, se concluye que $v = 0$. QED

Ejemplo 2.3.7 Los subespacios $W_1 = \{(a, 0) \mid a \in \mathbb{R}\}$ y $W_2 = \{(0, a) \mid a \in \mathbb{R}\}$ tienen como suma el espacio total $\mathbb{R}^2$ y además la suma es directa.

Los conceptos de suma y suma directa se pueden extender a más de dos subespacios, imponiendo la unicidad de la descomposición de cualquier vector de la suma en suma de elementos de cada subespacio. Las condiciones para la suma directa de más de dos subespacios son más complicadas de lo que uno podría suponer:

Teorema 2.3.6 *La suma de los subespacios W_i, $i = 1, 2, 3$ del espacio vectorial V es directa si y solo si se cumplen las relaciones*

$$W_1 \cap (W_2 + W_3) = \{0\}, \quad W_2 \cap (W_3 + W_1) = \{0\}, \quad W_3 \cap (W_1 + W_2) = \{0\}$$

Demostración. Supongamos que la suma es directa. Sea x un vector en la intersección $W_1 \cap (W_2 + W_3)$. Entonces, $x \in W_1$ y $x = x_2 + x_3$, por estar en la suma $W_2 + W_3$. Pero como la descomposición es única: $x = x + 0 + 0$ y $x = 0 + x_2 + x_3$ deben ser la misma, luego $x = 0$. De forma similar demostraríamos las otras intersecciones. Ahora suponemos que las tres intersecciones mencionadas en el teorema son iguales al vector 0. Sean $x = x_1 + x_2 + x_3 = y_1 + y_2 + y_3$ dos descomposiciones de un vector x. De manera similar a como hicimos la demostración en el caso de dos subespacios, ponemos:

$$x_1 - y_1 = y_2 - x_2 + y_3 - x_3$$

Pero el vector de la izquierda está en W_1 y el de la derecha en $W_2 + W_3$, luego están en la intersección. Como la intersección es el vector 0 concluimos que $x_1 = x_2$ y también: $x_2 + x_3 = y_2 + y_3$. Podemos repetir el razonamiento con otra pareja:

$$x_2 - y_2 = y_1 - x_1 + y_3 - x_3$$

con lo que al estar ambos en W_2 y $W_1 + W_3$, son iguales a 0 y por tanto, $x_2 = y_2$. De la misma forma se demuestra que $x_3 = y_3$, y la descomposición es única. QED

Si la suma directa de varios subespacios es el espacio total, se dice que este último se descompone en suma directa de los subespacios. Cada vector del espacio admite una descomposición en suma de vectores, perteneciente cada uno de ellos a un subespacio.

Asimismo, el concepto de suma directa se puede extender a espacios vectoriales, no necesariamente subespacios de un mismo espacio vectorial.

Definición 2.3.6 *Dados dos espacios vectoriales V_1 y V_2 definidos sobre un cuerpo $\mathbb{K}$, se define la suma directa de estos dos espacios como el conjunto de expresiones de la forma $v_1 + v_2$ (el signo suma tiene un sentido formal aquí, nótese que los vectores son de espacios distintos).*

La suma y el producto por escalares se definen de forma natural:

$$(v_1 + v_2) + (w_1 + w_2) = (v_1 + w_1) + (v_2 + w_2)$$

$$\lambda(v_1 + v_2) = \lambda v_1 + \lambda v_2$$

Nótese que se puede introducir también como el conjunto de pares, es decir, como los elementos del producto cartesiano de V_1 y V_2. También se puede extender a un número arbitrario de factores (finito, el caso infinito requiere un análisis más cuidadoso).

La última operación con espacios vectoriales que vamos a considerar es el espacio cociente. La idea es clasificar los vectores de un espacio vectorial en clases siguiendo un criterio establecido por un subespacio vectorial elegido. La construcción en el caso de espacios vectoriales solo añade la forma de hacer la clasificación. Las relaciones de equivalencia, pues de eso se trata aquí, aparecen en conjuntos arbitrarios como ya se habrá estudiado en otros lugares.

Definición 2.3.7 *Sea V un espacio vectorial y W un subespacio de V. Dados dos vectores $x, y \in V$, se dice que están en la misma clase (respecto de W), si:*

$$x - y \in W$$

Se trata de una relación de equivalencia, como se puede demostrar fácilmente. Cada clase se escribirá como:

$$[x] = x + W = \{x + y \mid y \in W\}$$

y se dice que x (que es un elemento cualquiera de la clase) es el representante de esa clase. El conjunto de clases se designa por V/W y tiene una estructura de espacio vectorial, definida de la forma siguiente:

$$\begin{array}{ccc} V/W \times V/W & \longrightarrow & V/W \\ (x + W, y + W) & \mapsto & (x + y) + W \end{array}$$

$$\begin{array}{ccc} \mathbb{K} \times V/W & \longrightarrow & V/W \\ (\lambda, x + W) & \mapsto & (\lambda x) + W \end{array}$$

En las cuestiones relativas a clases de equivalencia es necesario prestar atención al representante elegido. Es decir, si: $x + W = x' + W$ e $y + W = y' + W$, las clases $x + y + W$ y $x' + y' + W$ deberían coincidir (es decir, $x + y$ y $x' + y'$ deberían estar en la misma clase). Lo que es muy sencillo de comprobar. De la misma forma para el producto por escalares.

La idea de espacio vectorial cociente es sin duda ligeramente más complicada que las anteriores. Veremos unos ejemplos para tratar de aclarar su construcción y utilidad.

Ejemplo 2.3.8 Consideremos el espacio vectorial real $V = \mathbb{R}^3$ y el subespacio $W = \{(0, 0, z) \mid z \in \mathbb{R}\}$. Gráficamente podemos pensar en V como el conjunto de vectores en el espacio con origen en el origen de coordenadas, y en W como el eje z. Los elementos del espacio cociente V/W son las clases:

$$(x, y, z) + W$$

pero podemos elegir un representante sencillo para cada clase: (x, y, z) y (x', y', z') están en la misma clase si su diferencia está en W, es decir, si $x = x'$ e $y = y'$. La tercera coordenada es arbitraria, es decir, en una clase toma todos los valores. El más sencillo es obviamente el valor 0, y por lo tanto:

$$V/W = \{[(x, y, 0)] \mid x, y \in \mathbb{R}\}$$

Si identificamos vectores (con origen en el origen de coordenadas) con los puntos donde está su extremo, este espacio es el plano xy. Con más precisión, cada punto del plano xy está en una clase diferente (y en cada clase hay un punto del plano xy). Si en un problema dado la coordenada z no aparece, este espacio cociente, o el plano al que es isomorfo (en un sentido que precisaremos más adelante) resulta más sencillo de utilizar.

Ejemplo 2.3.9 Supongamos ahora que V es el espacio de polinomios en una variable x. Y que W es el subespacio de constantes: $W = \{\lambda \mid \lambda \in \mathbb{R}\}$. En este caso, dos polinomios son equivalentes (están en la misma clase) si su diferencia es una constante. El representante más sencillo de cada clase es el que tiene el término de grado cero igual a cero. Como ejemplo de aplicación, la derivada es constante en cada clase, es decir, las derivadas de dos polinomios que estén en la misma clase son iguales. Y si dos polinomios están en diferentes clases, sus derivadas son distintas. Considerada la derivada como una aplicación del espacio de polinomios en sí mismo, es inmediato ver que no es inyectiva. Pero si se toma como espacio inicial este espacio cociente, la derivada (definida como la derivada de cualquier elemento de la clase) es inyectiva. Aplicaciones de este resultado aparecerán más tarde. Aquí solo diremos que la derivada se anula en W (y que si la derivada de un polinomio es cero, ese polinomio está en W).

2.4 Sistemas de generadores, rango y bases

Ya hemos indicado anteriormente lo que es un sistema de generadores de un espacio vectorial. Con un sistema de este tipo podemos construir todos los elementos del espacio vectorial mediante combinaciones lineales. Sin embargo, es posible que un vector pueda expresarse como varias combinaciones lineales diferentes.

Ejemplo 2.4.1 Sea $V = \mathbb{R}^2$, y el sistema de generadores:

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

Aunque aún no hemos visto un método para saber si un sistema de vectores es sistema de generadores de un espacio vectorial, admitamos que éste lo es. Por otra parte no es muy difícil comprobarlo usando directamente la definición. Un vector como por ejemplo el $(1,-1)$ se puede expresar de muchas formas mediante una combinación lineal de estos vectores. Por ejemplo:

$$(1,-1) = (1,0) - (0,1), \quad (1,-1) = (1,1) - 2(0,1)$$

Esto es debido a que este sistema de generadores no es linealmente independiente, concepto que introducimos a continuación:

Definición 2.4.1 *Sea V un espacio vectorial. Se dice que una familia de vectores es linealmente independiente (l.i.), si toda combinación lineal de vectores de la familia igualada a cero, tiene necesariamente todos los coeficientes iguales a cero.*

Ejemplo 2.4.2 Es muy sencillo demostrar que la familia del ejemplo anterior no es linealmente independiente. Por ejemplo la siguiente combinación lineal es igual a 0 y sin embargo los coeficientes no son iguales a cero:

$$(1,0) - (0,1) - (1,-1) = 0$$

Cuando una familia de vectores no es linealmente independiente, se dice que es linealmente dependiente (l.d.). Es decir, una familia de vectores de un espacio lineal es linealmente dependiente cuando es posible encontrar una combinación lineal de vectores de esa familia igual a cero, y en la que no todos los coeficientes son nulos.

Ejemplo 2.4.3 En todo espacio vectorial, toda familia de vectores que contenga al vector 0 es l.d. En efecto, la combinación lineal trivial: $\lambda.0$ es cero para cualquier λ.

Una consecuencia interesante de la d.l. es la siguiente:

Teorema 2.4.1 *Si los vectores $\{x_1,\dots,x_n\}$, $n > 1$ del espacio vectorial V son l.d., alguno de estos vectores se puede poner como combinación lineal de los demás.*

Demostración. . Si este conjunto de vectores es l.d., existe una combinación lineal:

$$\sum_{i=1}^{n} \lambda_i x_i = 0$$

en la que algún λ_i es distinto de cero. Sea por ejemplo $\lambda_k \neq 0$, para algún k entre 1 y n. Entonces:

$$\lambda_k x_k + \sum_{i=1, i\neq k}^{n} \lambda_i x_i = 0 \Rightarrow x_k = -\frac{1}{\lambda_k} \sum_{i=1, i\neq k}^{n} \lambda_i x_i$$

debido a que existe el inverso de λ_k. QED

Este teorema nos conduce a la siguiente definición:

Definición 2.4.2 *Se dice que el vector $x \in V$ depende linealmente de S (subconjunto de V), si x puede expresarse como una combinación lineal de vectores de S.*

Debido al teorema anterior, la envolvente lineal de una familia de vectores puede considerarse generada por menos vectores de los que en un principio podría suponerse.

Teorema 2.4.2 *Sea S una familia de vectores del espacio vectorial V, y supongamos que S es l.d. Sea x un vector de S que depende linealmente de los demás vectores de S. Entonces la envolvente lineal de S es igual a la envolvente lineal de $S \setminus \{x\}$.*

La demostración es una consecuencia del teorema sobre dependencia lineal y del hecho de que cada vez que x aparezca en una combinación lineal de un vector de $\text{lin}(S)$, podemos sustituirlo por la combinación lineal de otros vectores de S según hemos visto en el teorema anterior.

El teorema lleva inmediatamente a la siguiente conclusión:

Teorema 2.4.3 *Si S es un sistema de generadores de un espacio vectorial V, y el vector x depende linealmente de los otros vectores de S, entonces, $S \setminus \{x\}$ es también un sistema de generadores de V.*

La demostración es evidente.

Definición 2.4.3 *El rango de una familia de vectores es el número máximo de vectores linealmente independientes que se pueden encontrar en la familia.*

Veremos más adelante como estudiar el rango y como ampliar este concepto a matrices.

Estamos en condiciones de definir lo que es una base de un espacio vectorial. Esto nos permitirá relacionar los espacios vectoriales con unos espacios tipo y simplificará los cálculos en muchas ocasiones al poder hablar de coordenadas sin necesidad de usar los objetos abstractos del espacio.

Definición 2.4.4 *Se dice que la familia de vectores del espacio vectorial V, $\mathcal{B}$, es una base, si es un sistema de generadores de V y es l.i.*

Ejemplo 2.4.4 La familia estudiada en un ejemplo anterior $\{(1,0),(0,1),(1,-1)\}$ no es un base de $\mathbb{R}^2$ pues no es l.i. Sin embargo, la familia $\{(1,0),(0,1)\}$ sí es una base. Nótese que se obtiene de la primera eliminando un vector que se podía poner como combinación lineal de los otros dos, lo que hace que siga siendo un sistema de generadores de acuerdo con el teorema demostrado antes. Además es linealmente independiente, como se comprueba sin más que aplicar la definición:

$$\lambda_1(1,0) + \lambda_2(0,1) = (0,0) \Rightarrow (\lambda_1, \lambda_2) = (0,0) \Rightarrow \lambda_1 = \lambda_2 = 0$$

Ejemplo 2.4.5 Consideremos el conjunto de polinomios en una variable x con coeficientes reales. Como ya hemos visto es un espacio vectorial sobre $\mathbb{R}$. El conjunto:

$$S = \{1, x, x^2, \ldots\}$$

es una base de este espacio. Cualquier polinomio es una combinación lineal de elementos de este conjunto. Además, el conjunto S es l.i. Cualquier combinación igualada a cero obliga a que todos los coeficientes sean 0:

$$\lambda_1 x^{n_1} + \lambda_2 x^{n_2} + \cdots + \lambda_k x^{n_k} = 0$$

con todos los naturales $n_i, \ i = 1, \ldots k$ distintos entre sí, implica $\lambda_i = 0, \ i = 1, \ldots k$. Nótese que las combinaciones lineales son sumas de productos de escalares por vectores con un número finito de sumandos.

En los dos ejemplos anteriores la situación es muy diferente. En el primero, dos vectores formaban una base. En el segundo, la base está formada por un número infinito de vectores, pero al menos es numerable. Si consideramos el espacio de funciones continuas en $\mathbb{R}$, la existencia de una base, con un número finito o infinito (numerable o no) de vectores, no resulta fácil de establecer. En este curso nos limitaremos a bases con un número finito de elementos, aunque en lo referente a otros aspectos, aparecerán ejemplos de espacios que no tienen este tipo de bases. Usando conceptos de teoría de conjuntos (axioma de elección) es posible probar que todo espacio vectorial posee una base.

Un espacio vectorial tiene en principio muchas bases. Dada una de ellas es posible hacer combinaciones lineales de sus elementos, y si los vectores que resultan son linealmente independientes, forman otra base distinta de la anterior. Estudiaremos esta situación con más detalle más adelante.

Por ahora, nos limitamos a demostrar el siguiente resultado:

Teorema 2.4.4 *Sea V un espacio vectorial. Todas las bases de V tienen el mismo cardinal.*

Este teorema es fundamental. Permite relacionar las bases de un espacio vectorial, y asignar a este espacio un número natural cuando el cardinal anterior es finito. Desde el punto de vista de las propiedades algebraicas del espacio, este número proporciona toda la información que necesitamos.

Definición 2.4.5 *Se llama dimensión de un espacio vectorial V sobre un cuerpo $\mathbb{K}$ al cardinal común de las bases de V.*

Los espacios $\mathbb{K}^n$, de los que hemos visto varios ejemplos, nos dan los prototipos de los espacio vectoriales de dimensión finita sobre el cuerpo $\mathbb{K}$. Cuando la dimensión es infinita hay que prestar atención a otras cuestiones, pero no entraremos en esos detalles aquí. La demostración la haremos en un espacio vectorial que admita una base con un número finito de elementos.

Demostración. Supongamos que el espacio vectorial V tiene una base: $\mathcal{B} = \{v_1, \ldots, v_n\}$, con n elementos. Probaremos que cualquier conjunto de vectores l.i. tiene como máximo n elementos. Sea $S = \{u_1, \ldots, u_m\}$ un conjunto de vectores l.i. Entonces: $u_1 = \sum_{i=1}^n \lambda_i v_i$, y alguno de los coeficientes no es cero. Si es, por ejemplo, $\lambda_1 \neq 0$, podemos sustituir v_1 por u_1 y obtener otra base, ya que será un sistema de generadores (al poder despejar v_1 en función de u_1 y $v_2, \ldots, v_n$) y además es l.i. Si

$$\mu_1 u_1 + \sum_{i=2}^n \mu_i v_i = 0$$

entonces, $\mu_1 = 0$ implica que los demás son cero, ya que son l.i. Si $\mu_1 \neq 0$, u_1 sería combinación lineal del resto, lo que es contradictorio con $\lambda_1 \neq 0$. Siguiendo este proceso (cambiando el orden si es necesario) construiríamos una base de V: $\{u_1, \ldots, u_k, v_{k+1}, \ldots, v_n\}$. En esta base, mediante el mismo razonamiento, podríamos sustituir uno de los v_j, digamos, v_{k+1}, por u_{k+1} si el coeficiente de v_{k+1} en el desarrollo de u_{k+1} en esta base es no nulo (alguno de los coeficientes de los vectores v_j es no nulo por razones de independencia lineal de los vectores u_i). Si seguimos sustituyendo está claro que en cada paso tendremos una base de V, y el número de vectores de S no puede ser mayor que n. QED

También podemos enunciar el siguiente resultado:

Teorema 2.4.5 *En un espacio vectorial de dimensión finita n no hay conjuntos de vectores l.i. con más de n vectores. Si un conjunto de vectores l.i. en V tiene n elementos linealmente independientes, es una base.*

La demostración es inmediata de lo anterior.

Teorema 2.4.6 *El rango de un sistema de vectores de un espacio vectorial es la dimensión de la envolvente lineal de ese sistema*

Demostración. El rango es el número máximo de vectores l.i. La dimensión es el cardinal de una base. Del sistema de vectores podemos retirar los que dependen linealmente de los demás hasta quedarnos con un conjunto l.i., que sigue generando la envolvente. Luego la dimensión es igual al rango. QED

Ejemplo 2.4.6 En el espacio complejo $\mathbb{C}$, el vector 1 es una base. La dimensión es 1. Cualquier otro número complejo diferente de cero es una base. Siempre se tiene este resultado, la dimensión de un cuerpo considerado como un espacio vectorial sobre sí mismo es 1.

Ejemplo 2.4.7 Si se considera a $\mathbb{C}$ como un espacio vectorial sobre $\mathbb{R}$, una base es por ejemplo, $\{1, i\}$. Pero $\{1, -1\}$ no lo es. La dimensión de $\mathbb{C}$ sobre los reales es 2.

Ejemplo 2.4.8 La dimensión del espacio $\mathbb{K}^n$ (producto cartesiano de $\mathbb{K}$ por sí mismo n veces) sobre $\mathbb{K}$ es justamente n. Podemos elegir la llamada base canónica:

$$\{(1, 0, \ldots, 0), (0, 1, 0, \ldots, 0), \ldots, (0, \ldots, 0, 1)\}$$

que es un conjunto l.i., pues de la combinación lineal:

$$\lambda_1(1, 0, \ldots, 0) + \lambda_2(0, 1, 0, \ldots, 0) + \cdots + \lambda_n(0, \ldots, 0, 1) = (0, \ldots, 0)$$

se deduce:

$$(\lambda_1, \lambda_2, \ldots, \lambda_n) = (0, \ldots, 0)$$

y por tanto todos los coeficientes son cero. Además, cualquier elemento de este espacio se puede poner como combinación lineal de los vectores de este conjunto:

$$(\lambda_1, \lambda_2, \ldots, \lambda_n) = \lambda_1(1, 0, \ldots, 0) + \lambda_2(0, 1, 0, \ldots, 0) + \cdots + \lambda_n(0, \ldots, 0, 1)$$

luego es una base, y la dimensión es n.

Definición 2.4.6 *Dada una base $\mathcal{B} = \{u_1, \ldots, u_n\}$ de un espacio vectorial V (de dimensión finita) y un vector $x \in V$, existe una sola combinación lineal de los vectores de la base que sea igual al vector dado. Se llaman coordenadas del vector x en la base $\mathcal{B}$ a los escalares $\lambda_1, \lambda_2, \ldots, \lambda_n$ tales que:*

$$x = \sum_{i=1}^{n} \lambda_i u_i$$

Como hemos dicho, las coordenadas están unívocamente determinadas. Por supuesto si cambiamos la base, cambiarán.

La correspondencia que se puede establecer, fijada un base, entre un espacio de dimensión finita n y $\mathbb{K}^n$, asignando a cada vector sus coordenadas en esa base, es biyectiva y tiene unas propiedades que serán estudiadas más adelante.

Dado un espacio vectorial V, se puede hablar de la dimensión de sus subespacios, pues éstos son a su vez espacios vectoriales sobre el mismo cuerpo. Un resultado importante es el siguiente:

Teorema 2.4.7 *Sea V un espacio vectorial de dimensión finita y W un subespacio de V. Se verifica:*

$$\dim W \leq \dim V$$

Demostración. En efecto, sea e_1 un vector de W, no nulo (si W tiene solo el vector nulo, el resultado es trivial). Si $\text{lin}\{e_1\} \neq W$, existirá un segundo vector en W, e_2, linealmente independiente con el anterior. Si $\text{lin}\{e_1, e_2\} \neq W$, habrá un tercero, etc. Como V es de dimensión finita, el proceso se acaba. En cada uno de los pasos, la dimensión de W es menor o igual que la de V, lo que demuestra el teorema. QED

Los espacios de dimensión finita son más sencillos de estudiar que los de dimensión infinita y a ellos estará dedicada la mayor parte del curso.

La construcción de bases no siempre es fácil, pero veremos un resultado que ayuda.

Teorema 2.4.8 *Teorema de prolongación de la base*

Sea V un espacio vectorial de dimensión finita igual a n, y W un subespacio de V. Si $\mathcal{B}_W = \{w_1, \dots, w_k\}$ es una base de W, se pueden encontrar vectores $\{u_{k+1}, \dots, u_n\}$ tales que el conjunto $\{w_1, \dots, w_k, u_{k+1}, \dots, u_n\}$ es una base de V.

Demostración. Sea $\{w_1, \dots, w_m\}$ una base de W. Si $W \neq V$, existirá un vector en V, tal que $\mathcal{B} \cup \{u_{k+1}\}$ es un conjunto l.i. (si no lo fuera, u_{k+1} estaría en W). Consideremos el espacio $W_{k+1} = \text{lin}(\mathcal{B} \cup \{u_{k+1}\})$. Si este espacio es igual a V, la demostración está acabada. Si no lo es, habrá otro vector u_{k+2} con el que se podrá razonar como antes. Como la dimensión de V es finita, el proceso acaba. Los vectores añadidos forman junto con los de la base de W inicial, la base ampliada. QED

Como consecuencia del anterior teorema podemos probar:

Teorema 2.4.9 *Sea V un espacio vectorial de dimensión finita. Si W es un subespacio de V, existe un subespacio U de V tal que:*

$$V = W \oplus U$$

Demostración. Sea $\mathcal{B}_W$ un base de W. Por el teorema de prolongación de la base, podemos construir una base de V, añadiendo vectores a $\mathcal{B}_W$. Sea $\mathcal{B}_V = \{w_1, \dots, w_k, u_{k+1}, \dots, u_n\}$ esta base. Definiendo $U = \text{lin}(\{u_{k+1}, \dots, u_n\})$, no es difícil probar que los espacios W y U tienen intersección igual a $\{0\}$ y su suma es V. QED

Se dice que W y U son subespacios suplementarios. Dado W la elección de U no es única.

Las dimensiones de los subespacios construidos a partir de otros por las operaciones estudiadas anteriormente están relacionadas a través de los teoremas siguientes.

Teorema 2.4.10 *Sea V un espacio vectorial de dimensión finita, y W_1, W_2 dos subespacios de V. Entonces,*

$$\dim W_1 + \dim W_2 = \dim(W_1 + W_2) + \dim(W_1 \cap W_2)$$

Demostración. Sea $\mathcal{B}_{W_1 \cap W_2} = \{a_1, \dots, a_k\}$ una base del espacio $W_1 \cap W_2$. Como este espacio está contenido en W_1, podemos ampliar la base y obtener otra de W_1: $\mathcal{B}_{W_1} = \{a_1, \dots, a_k, b_1, \dots, b_m\}$. Y como también está contenido en W_2, la podemos ampliar a una base de este subespacio: $\mathcal{B}_{W_2} = \{a_1, \dots, a_k, c_1, \dots, c_r\}$. Consideremos el conjunto de vectores de V:

$$\mathcal{B}_W = \{a_1, \dots, a_k, b_1, \dots, b_m, c_1, \dots, c_r\}$$

y probemos que es una base del espacio $W_1 + W_2$. En primer lugar es l.i. Construimos una combinación lineal de estos vectores y la igualamos a cero:

$$\sum_{i=1}^{k} \alpha_i a_i + \sum_{i=1}^{m} \beta_i b_i + \sum_{i=1}^{r} \gamma_i c_i = 0$$

Sean $v = \sum_{i=1}^{k} \alpha_i a_i$, $v_1 = \sum_{i=1}^{m} \beta_i b_i$ y $v_2 = \sum_{i=1}^{r} \gamma_i c_i$. Entonces, $v \in W_1 \cap W_2$, $v_1 \in W_1$ y $v_2 \in W_2$. Como la suma es cero, $v_2 \in W_1$, luego $v_2 \in W_1 \cap W_2$. Este vector debe poder expresarse como una combinación lineal de la base de este subespacio. Por tanto $v_2 = 0$. Debido a la independencia lineal de cada uno de los tres conjuntos de vectores a_i, b_i, c_i concluimos que todos los coeficientes son cero, y por tanto el conjunto construido es l.i.

Probemos ahora que generan el espacio $W_1 + W_2$. Esto es más sencillo. Cualquier vector de este espacio es suma de un vector de W_1 más un vector de W_2. Basta examinar las bases ampliadas de estos subespacios para ver que cualquier vector de la suma se puede poner como combinación lineal de los vectores del conjunto l.i. determinado anteriormente. Además todos estos vectores están en la suma. Luego es una base. Por tanto, la dimensión de $W_1 + W_2$ es $k + m + r$ lo que demuestra el teorema. QED

Como consecuencia se tiene:

Teorema 2.4.11 *Si la suma $W_1 \oplus W_2$ es directa, se tiene:*

$$\dim(W_1 \oplus W_2) = \dim W_1 + \dim W_2$$

La demostración es inmediata de $\dim\{0\} = 0$.
Además:

Teorema 2.4.12 *Sea V un espacio vectorial de dimensión finita, y W un subespacio de V. Entonces,*

$$\dim V = \dim W + \dim(V/W)$$

La demostración es consecuencia del siguiente teorema.
De aquí se deduce que la dimensión del espacio cociente V/W es igual a la dimensión de un subespacio suplementario de W. Como veremos más adelante, el espacio cociente es isomorfo al suplementario de W (a cualquiera de ellos).
Se puede precisar aún más.

Teorema 2.4.13 *Sea V un espacio vectorial de dimensión finita, y W un subespacio de V. Fijada una base de W, $\mathcal{B}_W = \{w_1, \ldots, w_k\}$, la ampliamos a una base de V: $\mathcal{B}_V = \{w_1, \ldots, w_k, v_1, \ldots, v_m\}$, donde $\dim V = m + k$. Entonces, el conjunto de vectores de V/W:*

$$\mathcal{B}_{V/W} = \{v_1 + W, \ldots, v_m + W\}$$

es una base de V/W.

Demostración. El conjunto $\mathcal{B}_{V/W}$ es l.i. Tomamos una combinación lineal e igualamos a cero:

$$\sum_{i=1}^{m} \lambda_i(v_i + W) = 0$$

Operando, obtenemos: $(\sum_{i=1}^m \lambda_i v_i) + W = 0$, es decir, la clase cuyo representante es $\sum_{i=1}^m \lambda_i v_i$, es la clase cero, o sea, $\sum_{i=1}^m \lambda_i v_i \in W$. Pero los vectores v_i no están en W, sino en el suplementario, por lo tanto: $\sum_{i=1}^m \lambda_i v_i = 0$, y como son l.i. se concluye que los coeficientes λ_i son cero. Veamos ahora que son un sistema de generadores. Sea $x + W$ un elemento cualquiera del espacio cociente. Como $x \in V$, se tiene:

$$x = \sum_{i=1}^{k} x_i w_i + \sum_{i=1}^{m} y_i v_i$$

Por tanto:

$$x + W = \sum_{i=1}^{m} y_i v_i + W = \sum_{i=1}^{m} y_i(v_i + W)$$

que es lo queríamos demostrar. QED

2.5 Cambios de base. Matrices

En lo que sigue el espacio vectorial V es de dimensión finita.

Según hemos visto en la sección anterior, este espacio admite una base, con un número de vectores igual a la dimensión del espacio, y podremos expresar los vectores en función de esta base. Si tenemos otra base, las coordenadas cambiarán. Este es el inconveniente de usar bases en un espacio vectorial, las expresiones de los vectores cambian al cambiar la base y hay que prestar mucha atención a la hora de hablar de coordenadas en vez de vectores. Sin embargo, el uso de bases presenta otras muchas ventajas, por lo que trataremos de establecer como cambian las coordenadas al cambiar las bases.

Sean $\mathcal{B} = \{u_1, \dots, u_n\}$ y $\mathcal{B}' = \{u'_1, \dots, u'_n\}$ dos bases del espacio vectorial V. Como todo vector puede expresarse como combinación lineal de los vectores de una base, este resultado es cierto para los vectores de la base $\mathcal{B}$ en función de los vectores de la base $\mathcal{B}'$:

$$\begin{aligned} u_1 &= a_{11}u'_1 + a_{21}u'_2 + \cdots + a_{n1}u'_n \\ u_2 &= a_{12}u'_1 + a_{22}u'_2 + \cdots + a_{n2}u'_n \\ &\vdots \\ u_n &= a_{1n}u'_1 + a_{2n}u'_2 + \cdots + a_{nn}u'_n \end{aligned}$$

es decir:

$$u_i = \sum_{j=1}^{n} a_{ji}u'_j$$

Por lo tanto, si $x \in V$, y su expresión en la base $\mathcal{B}$ es:

$$x = \sum_{i=1}^{n} x_i u_i$$

su expresión en la base $\mathcal{B}'$ será:

$$x = \sum_{i=1}^{n} x_i \sum_{j=1}^{n} a_{ji}u'_j = \sum_{i,j=1}^{n} a_{ji}x_i u'_j$$

Si ahora pensamos que x también puede escribirse en la base $\mathcal{B}'$:

$$x = \sum_{i=1}^{n} x'_i u'_i$$

llegamos a la igualdad:

$$\sum_{j=1}^{n} x'_j u'_j = \sum_{i,j=1}^{n} a_{ji}x_i u'_j$$

Pero la expresión en una base es única, por lo tanto, las coordenadas son iguales y se tiene:

$$x'_i = \sum_{i=1}^{n} a_{ij}x_j, \ i = 1, \dots n$$

Es decir, conociendo la relación entre las bases podemos saber como cambian las coordenadas. El cambio inverso, es decir pasar de las coordenadas en la base $\mathcal{B}'$ a las coordenadas en la base $\mathcal{B}$ también es sencillo de expresar. Basta repetir el razonamiento anterior, cambiando el papel de las bases $\mathcal{B}$ y $\mathcal{B}'$:

$$\begin{aligned} u'_1 &= b_{11}u_1 + b_{21}u_2 + \cdots + b_{n1}u_n \\ u'_2 &= b_{12}u_1 + b_{22}u_2 + \cdots + b_{n2}u_n \\ &\vdots \\ u'_n &= b_{1n}u_1 + b_{2n}u_2 + \cdots + b_{nn}u_n \end{aligned}$$

o:

$$u'_i = \sum_{j=1}^{n} b_{ji} u_j$$

Repitiendo el proceso llegamos a:

$$x_i = \sum_{j=1}^{n} b_{ij} x'_j, \; i = 1, \ldots n$$

Está claro que los escalares a_{ij} y b_{ij} no pueden ser independientes. La relación que los liga es:

$$x'_i = \sum_{j=1}^{n} a_{ij} x_j = \sum_{i=1}^{n} a_{ij} \sum_{k=1}^{n} b_{jk} x'_k$$

y como esta relación se debe cumplir para todo vector del espacio, o si se quiere para valores arbitrarios de x'_i, se tiene:

$$\sum_{j=1}^{n} a_{ij} b_{jk} = \delta_{ik}$$

donde δ_{ij} es un símbolo (delta de Kronecker) que representa un conjunto de valores de la forma siguiente:

$$\delta_{ij} = \begin{cases} 1 & \text{si} \quad i = j \\ 0 & \text{si} \quad i \neq j \end{cases}$$

La expresión anterior es un conjunto de ecuaciones (n^2) que relacionan los coeficientes a_{ij} con los b_{ij}.

Los cálculos con estos coeficientes son bastante complicados de escribir. Incluso el cálculo de los coeficientes a_{ij} en función de los b_{ij} no parece sencillo, aunque en principio es un problema lineal. Un elemento esencial para estas operaciones que facilita enormemente los cálculos es la matriz.

Ejemplo 2.5.1 Consideremos en $\mathbb{R}_2[x]$ dos bases:

$$\mathcal{B} = \{1, x, x^2\}, \quad \mathcal{B}' = \{1, x, \frac{1}{2}(3x^2 - 1)\}$$

y estudiemos como se transforman las coordenadas de un polinomio $p(x) = \lambda_1 + \lambda_2 x + \lambda_3 x^2$. Expresado en la primera base, las coordenadas son los tres números reales: $(\lambda_1, \lambda_2, \lambda_3)$. Para calcularlas en la segunda base, veamos como se escriben los vectores de la primera base en función de los de la segunda:

$$\begin{aligned} 1 &= 1 \\ x &= x \\ x^2 &= \frac{1}{3} + \frac{2}{3}\left(-\frac{1}{2} + \frac{3}{2}x^2\right) \end{aligned}$$

luego los escalares a_{ij} son:

$$\begin{array}{lll} a_{11} = 1 & a_{21} = 0 & a_{31} = 0 \\ a_{12} = 0 & a_{22} = 1 & a_{32} = 0 \\ a_{13} = \frac{1}{3} & a_{23} = 0 & a_{33} = \frac{2}{3} \end{array}$$

y por lo tanto las coordenadas en la segunda base en función de las coordenadas en la primera son:

$$\begin{aligned} \lambda'_1 &= \lambda_1 + \frac{1}{3}\lambda_3 \\ \lambda'_2 &= \lambda_2 \\ \lambda'_3 &= \frac{2}{3}\lambda_3 \end{aligned}$$

Los coeficientes b_{ij} se calculan también fácilmente:

$$\begin{aligned} 1 &= 1 \\ x &= x \\ \frac{1}{2}(3x^2-1) &= -\frac{1}{2}+\frac{3}{2}x^2 \end{aligned}$$

$$\begin{array}{lll} b_{11}=1 & b_{21}=0 & b_{31}=0 \\ b_{12}=0 & b_{22}=1 & b_{32}=0 \\ b_{13}=-\frac{1}{2} & b_{23}=0 & b_{33}=\frac{3}{2} \end{array}$$

y el cambio de coordenadas inverso es:

$$\begin{aligned} \lambda_1 &= \lambda'_1 - \frac{1}{2}\lambda'_3 \\ \lambda_2 &= \lambda'_2 \\ \lambda_3 &= \frac{3}{2}\lambda'_3 \end{aligned}$$

No es difícil comprobar que los coeficientes a_{ij} y b_{ij} satisfacen las relaciones estudiadas anteriormente. Las dos expresiones del polinomio $p(x)$ en estas dos bases son:

$$p(x) = \lambda_1 + \lambda_2 x + \lambda_3 x^2 = \left(\lambda_1 + \frac{1}{3}\lambda_3\right) + \lambda_2 x + \frac{2}{3}\lambda_3 \frac{1}{2}(3x^2-1)$$

Aún en un ejemplo tan sencillo, los cálculos son tediosos de escribir. El lenguaje de matrices permite un mayor aprovechamiento.

2.5.1 Matrices

Definición 2.5.1 *Una matriz es una colección de objetos dispuestos en forma rectangular con un cierto número de filas y columnas.*

En lo que sigue, las matrices estarán formadas por escalares, pero se les puede encontrar muchas otras aplicaciones. Los elementos de la matriz se designan por dos subíndices que indican la posición que ocupan: el primero la fila y el segundo la columna. Así, el elemento a_{23} de una matriz está en la fila 2 y columna 3. La matriz $A = (a_{ij})$ es la formada por los elementos a_{ij} en las posiciones correspondientes.

Ejemplo 2.5.2 La siguiente disposición es una matriz 2×3, es decir, con dos filas y tres columnas:

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

El elemento 22 es: $3i$

Teorema 2.5.1 *El conjunto de matrices $n \times m$ con coeficientes en un cuerpo $\mathbb{K}$, $M_{n\times m}(\mathbb{K})$ es un espacio vectorial sobre $\mathbb{K}$ de dimensión nm*

Demostración. La suma de matrices se define elemento a elemento, es decir la matriz suma tiene como elementos la suma de los elementos que ocupan la misma posición en cada uno de los sumandos. Y el producto por escalares consiste en multiplicar cada elemento de la matriz por el escalar. Las propiedades de espacio vectorial son claras.

En cuanto a la dimensión, basta encontrar una base con nm elementos. La más sencilla es: $\mathcal{B} = \{E_{ij} \mid i = 1, \dots n,\ j = 1 \dots, m\}$ donde las matrices E_{ij} tienen un 0 en todas las posiciones, salvo en la fila i columna j donde tiene un 1. No es difícil ver que es un sistema l.i. y que es un sistema de generadores. QED

Existen muchas operaciones con matrices que iremos estudiando poco a poco. Por ejemplo, se puede definir la transpuesta de una matriz, que permite construir a partir de una matriz de dimensión $n \times m$ otra de dimensión $m \times n$:

Definición 2.5.2 *Sea $A \in M_{n\times m}(\mathbb{K})$, con elementos (a_{ij}). Se define la matriz transpuesta de A, A^t, como la matriz en $M_{m\times n}(\mathbb{K})$ con elementos (b_{ij}), tal que:*

$$b_{ij} = a_{ji},\ i = 1,\dots m,\ j = 1,\dots,n$$

Es decir, se intercambian las filas por las columnas.

Las matrices se usan en álgebra lineal constantemente, con distinto significado. Hay pues que poner cuidado en su interpretación.

Se puede definir otra operación entre matrices, el producto. Sin embargo, no siempre es posible multiplicar dos matrices. Para ello el número de columnas de la primera debe ser igual al de filas de la segunda y la matriz resultante tiene tantas filas como filas tenía la primera matriz y tantas columnas como columnas tenía la segunda. Se ve que no se trata de una operación en $M_{n\times m}(\mathbb{K})$, sino:

$$\cdot : M_{n\times m}(\mathbb{K}) \times M_{m\times k}(\mathbb{K}) \longrightarrow M_{n\times k}(\mathbb{K})$$

El método de multiplicación, que es asociativo y distributivo respecto a la suma, consiste en lo siguiente. El elemento ij de la matriz producto es la suma de los productos de los elementos de la fila i de la primera matriz por los elementos de la columna j de la segunda. En función de los elementos de las matrices que se multiplican se tiene la fórmula:

$$c_{ik} = \sum_{j=1}^{m} a_{ij} b_{jk}$$

donde $A = (a_{ij}),\ B = (b_{ij}),\ C = AB = (c_{ij})$ y los índices se mueven en el rango adecuado.

Ejemplo 2.5.3 Sean las matrices A y B:

$$A = \begin{pmatrix} 0 & -1 & 0 \\ 1+i & 2 & -i \end{pmatrix}, \quad B = \begin{pmatrix} -1 & 3 & i & 0 \\ 2+2i & -1 & 1 & 0 \\ 0 & -1 & 0 & 1 \end{pmatrix}$$

Su producto es:

$$AB = \begin{pmatrix} -2-2i & 1 & -1 & 0 \\ 3+4i & 1+4i & 1+i & -i \end{pmatrix}$$

Obviamente no se puede cambiar el orden de los factores, no porque se obtenga otro resultado, sino porque en la mayoría de los casos ni siquiera se podrá hacer el producto.

Pero cuando las matrices tienen igual el número de filas y de columnas (en este caso se llaman cuadradas), la operación anterior es una operación interna en el espacio vectorial de matrices. Con ella, este conjunto se transforma en un anillo respecto de las operaciones suma y producto, un anillo no conmutativo con elemento unidad (la matriz identidad, unos en la diagonal, es decir cuando $i = j$, y ceros en las demás posiciones), que sin embargo no es un cuerpo, porque no siempre existe el inverso. Esta estructura que mezcla la de espacio vectorial con la de anillo, se llama un álgebra, en este caso no conmutativa con elemento unidad (respecto a la multiplicación). Dentro de ella se pueden seleccionar las matrices que sí poseen inverso, y construir un grupo multiplicativo, como hemos hecho con los anillos. El cálculo del inverso de matrices cuadradas (cuando éste existe) no es sencillo. Se dirá que una matriz es regular cuando tiene inverso.

Existe otra forma de ver el producto de matrices. Supongamos que A es una matriz con coeficientes en $\mathbb{K}$, de dimensión $n \times m$ y B otra de dimensión $m \times k$, de manera que existe el producto AB. Supongamos que escribimos la matriz B como una colección de matrices de m filas y una columna (vectores columna):

$$B = (B_1, B_2, \dots, B_k)$$

Entonces, el producto AB se puede leer como una matriz cuyos vectores columna son la multiplicación de A por los vectores columna de la matriz B:

$$AB = (AB_1, AB_2, \dots, AB_k)$$

La demostración es evidente del producto de matrices. Un resultado similar se obtiene con las filas: Si la matriz A se escribe como una colección de vectores fila (una fila y m columnas):

$$A = \begin{pmatrix} A_1 \\ A_2 \\ \vdots \\ A_n \end{pmatrix}$$

el producto AB se puede escribir como:

$$AB = \begin{pmatrix} A_1 B \\ A_2 B \\ \vdots \\ A_n B \end{pmatrix}$$

2.5.2 Operaciones elementales con matrices

Aunque su motivación no sea excesivamente clara en este punto, vamos a estudiar una serie de manipulaciones formales con las filas y las columnas de una matriz.

Una operación elemental de filas en una matriz consiste en una de las tres transformaciones siguientes:

1. Cambiar entre sí dos filas
2. Multiplicar una fila por un escalar no nulo
3. Multiplicar una fila por un escalar no nulo y sumarla a otra fila

De manera similar se definen las operaciones elementales entre columnas.

Ejemplo 2.5.4 Sea la siguiente matriz con coeficientes en $\mathbb{C}$:

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

La siguiente matriz se obtiene de ésta mediante una operación elemental:

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

en la que hemos sumado la tercera fila multiplicada por -1 a la cuarta. La matriz que se obtiene es claramente distinta de la primera. No estamos diciendo que las operaciones elementales dejen invariantes las matrices. La siguiente matriz también se obtiene de la primera, mediante el intercambio de la tercera y cuarta columnas:

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

El uso de operaciones elementales permite simplificar las matrices llegando a formas más sencillas. Otra cosa es la utilidad, debido a que aún no hemos visto qué relación, desde el punto de vista de las aplicaciones de las matrices al álgebra, existe entre matrices obtenidas mediante operaciones elementales.

Teorema 2.5.2 *Sea A una matriz en $M_{n\times m}(\mathbb{K})$. La operación elemental que consiste en intercambiar la fila i por la fila j permite obtener una matriz igual a $F_{ij}A$, donde F_{ij} es una matriz cuadrada de dimensión n, que tiene ceros en todas las posiciones salvo en las ij y ji y en las kk para todo $k \neq i, j$ donde tiene un 1.*

Por ejemplo, para $n = 3$, $m = 8$, la matriz F_{13} es la siguiente:

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

Las matrices F_{ij} tienen inverso. Concretamente el inverso coincide con ella misma:

$$F_{ij}F_{ij} = I,$$

donde I es la matriz identidad en dimensión n. Resultado nada sorprendente si pensamos que si hacemos esta operación elemental dos veces, la matriz no cambia. Que se verifica el teorema es fácil de ver. Cuando multiplicamos la matriz F_{ij} por A tomamos una fila de F_{ij} y actuamos sobre una columna de A. Si la fila es distinta de la i o la j, no se produce ningún cambio, luego se obtiene la misma fila de la matriz A. Sin embargo, cuando usamos la fila i, al multiplicar por una columna cualquiera de A no se obtiene el elemento i de esa columna sino el j. Es decir la fila i de la matriz A es sustituida por la fila j.

Ejemplo 2.5.5 La matriz F_{24} para $n = 4$ es:

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

Su producto por la matriz A del ejemplo anterior es:

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

Con las columnas ocurre una situación parecida pero ahora las multiplicaciones son por la derecha:

Teorema 2.5.3 *Sea A una matriz en $M_{n\times m}(\mathbb{K})$. La operación elemental que consiste en intercambiar la columna i por la columna j nos da una matriz igual a AF_{ij}, donde F_{ij} es la matriz descrita en el teorema precedente, pero ahora en dimensión m.*

Ejemplo 2.5.6 El producto de la matriz A del ejemplo anterior por la matriz F_{24} en dimensión 6:

$$F_{24} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

es:

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

donde las columnas segunda y cuarta.

La segunda operación elemental es multiplicar una fila (o columna) por un escalar no nulo. Se tiene el teorema siguiente:

Teorema 2.5.4 *Sea A una matriz en $M_{n\times m}(\mathbb{K})$. La operación elemental que consiste en multiplicar la fila (o columna) i por un escalar $\lambda \neq 0$ nos da una matriz igual a $K_i(\lambda)A$ (o $AK_i(\lambda)$ para columnas), donde $K_i(\lambda)$ es la matriz de dimensión n (o m para columnas) que tiene ceros en todas las posiciones salvo en la diagonal, donde tiene 1 excepto en la posición ii que tiene λ.*

La demostración es evidente de las reglas del producto de matrices. Estas matrices $K_i(\lambda)$ tienen inverso, que es la matriz $K_i(\lambda^{-1})$.

Finalmente la tercera operación se describe de manera similar.

Teorema 2.5.5 *Sea A una matriz en $M_{n\times m}(\mathbb{K})$. La operación elemental que consiste en multiplicar la fila i por un escalar λ y sumarla a la fila j, nos proporciona una matriz igual a $L_{ij}(\lambda)A$, donde $L_{ij}(\lambda)$ es la matriz de dimensión n que tiene ceros en todas las posiciones y unos en la diagonal con la excepción siguiente: en la posición ji aparece λ*

Un resultado similar se obtiene para columnas. La matriz a emplear es ahora la transpuesta de $L_{ij}(\lambda)$ y suma a la columna j la i multiplicada por λ.

La demostración es también evidente de las reglas del producto. Pero podemos interpretarla de la forma siguiente. La matriz A se considera como una matriz de filas:

$$A = \begin{pmatrix} A_1 \\ A_2 \\ \vdots \\ A_n \end{pmatrix}$$

El producto es ahora:

$$L_{ij}(\lambda)A = L_{ij}(\lambda)\begin{pmatrix} A_1 \\ \vdots \\ A_i \\ \vdots \\ A_j \\ \vdots \\ A_n \end{pmatrix} = \begin{pmatrix} A_1 \\ \vdots \\ A_i \\ \vdots \\ \lambda A_i + A_j \\ \vdots \\ A_n \end{pmatrix}$$

En cuanto a las operaciones con columnas, se considera la matriz A como una matriz de columnas $A = (A_1, \dots, A_m)$ y se tiene:

$$AL_{ij}^t(\lambda) = (A_1, \dots, A_i, \dots, A_j, \dots A_m)L_{ij}^t(\lambda) = (A_1, \dots, A_i, \dots, \lambda A_i + A_j, \dots A_m)$$

No es difícil probar que estas matrices tienen también inverso (aunque λ sea cero). Concretamente:

$$L_{ij}(\lambda)L_{ij}(-\lambda) = I$$

Como hemos dicho, el uso de operaciones elementales sobre una matriz permite simplificar la forma de ésta. Se tiene el siguiente resultado.

Definición 2.5.3 *Una matriz escalón reducida por filas es una matriz en la que en cada fila, el primer elemento no nulo está en una columna situada a la derecha de la columna de la fila anterior en la que está el primer elemento no nulo de esa fila*

Teorema 2.5.6 *Sea A una matriz rectangular $m \times n$ sobre un cuerpo $\mathbb{K}$. Esta matriz se puede reducir mediante operaciones elementales con filas y columnas a una matriz escalón.*

Demostración. Consiste simplemente en la descripción de un algoritmo que permite llegar a ese resultado en un número finito de pasos (cosa natural dado el carácter finito de la matriz).

Mediante operaciones elementales del tipo cambios de filas, podemos conseguir que el primer elemento no nulo de la primera fila esté en la primera columna no nula de la matriz. Es decir las columnas anteriores son todas cero. Usando ese primer elemento no nulo en la primera fila, podemos hacer nulos los elementos situados debajo de él, utilizando la operación elemental de multiplicar una fila por un escalar y sumar a otra. Una vez conseguido esto, colocamos en la segunda fila aquella (descontando la primera) que tiene el primer elemento no nulo en la columna con el menor índice posible. Mediante operaciones elementales podemos conseguir que todos los elementos por debajo de éste sean nulos. Etc. QED

Ejemplo 2.5.7

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

Vamos a reducirla a la forma escalón. La primera fila tiene un elemento no nulo en la primera columna. Utilizamos éste para hacer cero el resto de elementos de la primera columna. Multiplicando por -1 y sumando a la tercera, y multiplicando por 1 y sumando a la cuarta, obtenemos:

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

La segunda fila sirve a nuestros propósitos. Y la tercera también. Con lo que la matriz ya está en forma escalón. Utilizando otras operaciones elementales, podríamos conseguir que los primeros elementos no nulos de cada fila fueran iguales a 1. Multiplicando la segunda y cuarta fila por $1/2$ y la tercera por $1/3$, tenemos:

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

Aún podemos obtener una forma más simplificada. Usamos operaciones elementales para eliminar los elementos de las columnas que no son cero ni son el primero de la fila no nulo. Este paso no siempre se puede hacer, debido a que los primeros elementos no nulos de cada fila no están escalonados como aquí, de uno en uno. Multiplicando la tercera fila por 1 y sumando a la primera, la cuarta por -1 y sumando a la primera, la tercera por $-3/2$ y sumando a la segunda, la cuarta por $-5/2$ y sumando a la segunda y la cuarta por 1 y sumando a la tercera se llega a:

$$\begin{pmatrix} 1 & 0 & 0 & 0 & 1/3 & 1/2 \\ 0 & 1 & 0 & 0 & -1/2 & -5/4 \\ 0 & 0 & 1 & 0 & 1/3 & 1/2 \\ 0 & 0 & 0 & 1 & 0 & 1/2 \end{pmatrix}$$

No se pueden hacer más ceros mediante operaciones elementales de filas. Nótese que las operaciones elementales de filas no tienen porqué conmutar. Si uno recuerda que en definitiva no son más que productos por la izquierda de las matrices explicadas antes, es claro que en general éstas no conmutan. La matriz final se obtiene de la inicial mediante un producto por una matriz que es el producto de las correspondientes a las operaciones elementales hechas. Recordando todos los pasos, se tiene:

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & -5/2 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & -3/2 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & -1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

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

Obsérvese que el orden de multiplicación es el inverso (claro está) del orden en el que se han hecho las operaciones elementales. El producto es:

$$P = \begin{pmatrix} 1/6 & 0 & 1/3 & -1/2 \\ -3/4 & 1/2 & -1/2 & -5/4 \\ 1/6 & 0 & 1/3 & 1/2 \\ 1/2 & 0 & 0 & 1/2 \end{pmatrix}$$

Esta es una matriz cuadrada con inverso, que corresponde a una sucesión de operaciones elementales. Su obtención es ciertamente laboriosa. Pero existe una forma mucho más sencilla de obtenerla. Supongamos que sometemos a la matriz identidad a las mismas operaciones elementales que a A. Al final obtenemos las matrices anteriores multiplicando a la matriz identidad: $PI = P$. Por tanto la matriz P se obtiene fácilmente de esta forma. Cada vez que hacemos una operación elemental en la matriz A, hacemos la misma en la matriz I.

Como hemos visto, no es posible simplificar más la matriz haciendo operaciones con filas. Sin embargo, operando con las columnas podemos simplificar aún más. Consideremos la matriz:

$$\begin{pmatrix} 1 & 0 & 0 & 0 & 1/3 & 1/2 \\ 0 & 1 & 0 & 0 & -1/2 & -5/4 \\ 0 & 0 & 1 & 0 & 1/3 & 1/2 \\ 0 & 0 & 0 & 1 & 0 & 1/2 \end{pmatrix}$$

La primera columna por $-1/3$ sumada a la quinta, y por $-1/2$ sumada a la sexta da:

$$\begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & -1/2 & -5/4 \\ 0 & 0 & 1 & 0 & 1/3 & 1/2 \\ 0 & 0 & 0 & 1 & 0 & 1/2 \end{pmatrix}$$

La segunda columna por $1/2$ sumada a la quinta, y por $5/4$ sumada a la sexta da:

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

La tercera columna por $-1/3$ sumada a la quinta, y por $-1/2$ sumada a la sexta da:

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

Finalmente, la cuarta columna por $-1/2$ sumada a la sexta da:

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

Ahora tenemos una matriz que multiplica a A por la derecha:

$$Q = \begin{pmatrix} 1 & 0 & 0 & 0 & -1/3 & -1/2 \\ 0 & 1 & 0 & 0 & 1/2 & 5/4 \\ 0 & 0 & 1 & 0 & -1/3 & -1/2 \\ 0 & 0 & 0 & 1 & 0 & -1/6 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

que es el producto de las operaciones elementales que hemos hecho (ahora en el mismo orden en que las hacemos). Entonces, decimos que hemos obtenido (PAQ) la forma más sencilla posible desde el punto de vista de transformaciones elementales.

Cálculo de la inversa de una matriz.

Sea A una matriz cuadrada, de la que suponemos tiene inversa. Haciendo operaciones elementales con filas podemos llegar a la forma más sencilla posible según lo dicho anteriormente. Es decir: PA es una matriz reducida, concretamente la matriz identidad (si no es así, A no puede tener inversa, discutiremos esto más adelante). Pero, la matriz P es entonces A^{-1}, la matriz inversa de A, y se obtiene aplicando a la matriz A las mismas operaciones elementales que nos permitieron pasar de A a la identidad. Luego $PI = P$ es la matriz inversa de A.

Ejemplo 2.5.8 Sea:

$$A = \begin{pmatrix} 2 & -1 & 1 \\ 0 & 1 & 0 \\ -4 & 0 & 1 \end{pmatrix}$$

que suponemos tiene inversa. Calculándola mediante operaciones elementales:

$$A = \left(\begin{array}{rrr|rrr} 2 & -1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ -4 & 0 & 1 & 0 & 0 & 1 \end{array}\right)$$

$$\left(\begin{array}{rrr|rrr} 2 & -1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & -2 & 3 & 2 & 0 & 1 \end{array}\right)$$

$$\left(\begin{array}{rrr|rrr} 2 & -1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 3 & 2 & 2 & 1 \end{array}\right)$$

$$\left(\begin{array}{rrr|rrr} 2 & 0 & 1 & 1 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 3 & 2 & 2 & 1 \end{array}\right)$$

$$\left(\begin{array}{rrr|rrr} 2 & 0 & 0 & 1/3 & 1/3 & -1/3 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 3 & 2 & 2 & 1 \end{array}\right)$$

$$\left(\begin{array}{rrr|rrr} 1 & 0 & 0 & 1/6 & 1/6 & -1/6 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 3 & 2 & 2 & 1 \end{array}\right)$$

$$\left(\begin{array}{rrr|rrr} 1 & 0 & 0 & 1/6 & 1/6 & -1/6 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 2/3 & 2/3 & 1/3 \end{array}\right)$$

luego la matriz inversa es:

$$\begin{pmatrix} 1/6 & 1/6 & -1/6 \\ 0 & 1 & 0 \\ 2/3 & 2/3 & 1/3 \end{pmatrix}$$

como se puede comprobar fácilmente.

2.5.3 La matriz del cambio de base

Veamos una primera aplicación de matrices a la teoría de espacios vectoriales. Como hemos visto, cuando tenemos dos bases en un espacio vectorial (de dimensión finita), podemos hallar las coordenadas de un vector en una base cuando conocemos las coordenadas en la otra. Si la relación entre los vectores de las bases $\mathcal{B} = \{u_i\}$ y $\mathcal{B}' = \{u'_i\}$ es:

$$u_i = \sum_{j=1}^{n} a_{ji} u'_j, \tag{2.1}$$

podemos definir una matriz:

$$P = (a_{ij})$$

cuyos elementos sean las coordenadas de los vectores de la primera base en función de los de la segunda, cuidando el orden de filas y columnas. Sea $X \in \mathbb{K}^n$ el vector (columna) de coordenadas correspondiente a $x \in V$ en la base $\mathcal{B}$ y $X' \in \mathbb{K}^n$ el correspondiente en la base $\mathcal{B}'$. Es decir, en la notación utilizada anteriormente:

$$X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, \quad X' = \begin{pmatrix} x'_1 \\ x'_2 \\ \vdots \\ x'_n \end{pmatrix}$$

donde:

$$x = \sum_{i=1}^{n} x_i u_i, \quad x = \sum_{i=1}^{n} x'_i u'_i$$

Teniendo en cuenta las ecuaciones que estudiamos para el cambio de coordenadas:

$$x'_i = \sum_{i=1}^{n} a_{ij} x_j, \; i = 1, \ldots n$$

vemos que se pueden escribir, utilizando matrices, como:

$$X' = PX$$

El cambio inverso es:

$$x_i = \sum_{i=1}^{n} b_{ij} x'_j, \; i = 1, \ldots n$$

luego, si $P' = (b_{ij})$, se tiene:

$$X = P'X'$$

Como sabemos, los escalares a_{ij} y b_{ij} no son independientes, sino que verifican:

$$\sum_{j=1}^{n} a_{ij} b_{jk} = \delta_{ik}$$

Pero esto no es más que la igualdad del producto de matrices P y P' con la matriz identidad:

$$PP' = I$$

es decir, la matriz de cambio de base es una matriz que posee inverso y este inverso es justamente la matriz de cambio de base en sentido contrario. Veamos un ejemplo de lo dicho.

Ejemplo 2.5.9 Se define la traza de una matriz cuadrada como la suma de los elementos de la diagonal. El conjunto de matrices 2×2 con coeficientes en $\mathbb{C}$ de traza nula es un espacio vectorial complejo. La suma de matrices de traza nula es una matriz de traza nula y el producto de escalares por matrices de traza nula es también una matriz de traza nula. La dimensión de este espacio es tres (la dimensión del espacio total es 4 como ya sabemos, y la condición de traza nula selecciona un subespacio con dimensión 3, como detallaremos en la próxima sección). Una base es:

$$h = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad e = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad f = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$$

Cualquier matriz de traza nula es combinación lineal de estas tres matrices:

$$A = \begin{pmatrix} \alpha & \beta \\ \gamma & -\alpha \end{pmatrix} = \alpha h + \beta e + \gamma f$$

Seleccionemos otra base en este espacio, que tendrá también tres elementos.

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

y calculemos la matriz de cambio de base, que será claramente una matriz 3×3. Para ello expresemos los elementos de la base $\{h, e, f\}$ en función de los elementos de la base $\{\sigma_1, \sigma_2, \sigma_3\}$:

$$h = \sigma_3, \quad e = \frac{1}{2}(\sigma_1 + i\sigma_2), \quad f = \frac{1}{2}(\sigma_1 - i\sigma_2)$$

por lo que la matriz de cambio de base es:

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

y por tanto, si A es cualquier matriz de traza nula, con coordenadas en la base $\{h, e, f\}$, dadas por x_1, x_2, x_3, y coordenadas en la base $\{\sigma_1, \sigma_2, \sigma_3\}$, y_1, y_2, y_3, se tiene:

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

es decir:

$$\begin{aligned} y_1 &= \frac{1}{2}(x_2 + x_3) \\ y_2 &= \frac{i}{2}(x_2 - x_3) \\ y_3 &= x_1 \end{aligned}$$

Es sencillo comprobar que el resultado es correcto:

$$A = \begin{pmatrix} x_1 & x_2 \\ x_3 & -x_1 \end{pmatrix} = \frac{1}{2}(x_2 + x_3) \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} + \frac{i}{2}(x_2 - x_3) \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} + x_1 \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

2.6 Ecuaciones de subespacios

Como dijimos anteriormente, el rango de una familia de vectores en un espacio vectorial es el número máximo de vectores l.i. que se pueden encontrar en esa familia. Cuando el espacio es de dimensión finita, el rango es un número finito, porque, como hemos visto, no hay conjuntos de vectores l.i. con más de n elementos, donde n es la dimensión del espacio. Ampliamos este concepto a matrices.

Definición 2.6.1 *Sea A una matriz con coeficientes en $\mathbb{K}$ con n filas y m columnas. Se define el rango de A como el rango de sus vectores fila (considerados como vectores del espacio $\mathbb{K}^n$).*

Ejemplo 2.6.1 El rango de la matriz:

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

es igual a 2, pues es fácil ver que solo hay dos vectores fila l.i. (la tercera fila es igual a dos veces la segunda menos la primera).

El rango por filas de una matriz es muy fácil de calcular si está en la forma escalón. En efecto, dada la forma que allí tienen los vectores fila, basta restar del número total de filas, las filas formadas por ceros. Nótese que desde este punto de vista, lo que estamos haciendo al reducir una matriz a su forma reducida es establecer combinaciones lineales de vectores (que generan la misma envolvente lineal que los vectores originales, debido a las exigencias que se han hecho sobre las operaciones elementales), y por tanto, una vez llegados a la forma final, basta excluir los vectores que son cero.

No parece que las filas hayan de jugar un papel más importante que las columnas. Podríamos haber definido el rango de una matriz como el rango del sistema de vectores formado por sus vectores columnas. Pero ocurre que ambos rangos son iguales.

Teorema 2.6.1 *El rango del sistema de vectores fila de una matriz y el rango del sistema de sus vectores columna son iguales, y es el rango de la matriz por definición.*

Demostración. Sea A una matriz $n \times m$ sobre un cuerpo $\mathbb{K}$ y supongamos que r_f y r_c son sus rangos por filas y columnas respectivamente. Por la definición de rango por filas, existen r_f filas l.i. que podemos suponer que son las r_f primeras. Las demás filas dependen linealmente de estas primeras:

$$F_k = \sum_{i=1}^{r_f} \lambda_{ki} F_i, \ k = r_f + 1, \ldots, n$$

siendo F_i los vectores fila de la matriz A. En componentes:

$$a_{kj} = \sum_{i=1}^{r_f} \lambda_{ki} a_{ij}, \ k = r_f + 1, \ldots, n, \ j = 1, \ldots m$$

Por lo tanto, las columnas $j = 1, \ldots, m$ se puedes poner como: a_{kj} arbitrarios cuando $k = 1, \ldots r_f$ y $a_{kj} = \sum_{i=1}^{r_f} \lambda_{ki} a_{ij}$ cuando $k = r_f + 1, \ldots, n$. Definiendo una colección de vectores en $\mathbb{K}^m$:

$$(1, 0, \ldots, 0, \lambda_{r_f+1,1}, \ldots, \lambda_{n,1}), \quad (0, \ldots, 0, 1, \lambda_{r_f+1,r}, \ldots, \lambda_{n,r})$$

vemos que la columna j es combinación lineal de ellos (con coeficientes: $a_{1j}, \ldots, a_{rj}$). Por lo tanto el número de columnas l.i. es menor o igual que r_f. Empezando de manera similar por las columnas obtendríamos: $r_c \leq r_f$. Así, el rango por filas es igual al rango por columnas. Obtendremos este resultado más tarde al estudiar la relación entre determinantes y rangos. QED

El rango de una matriz puede cambiar al multiplicarlo por otra. Se tiene el siguiente resultado.

Teorema 2.6.2 *Sean $A \in M_{n\times m}(\mathbb{K})$ y $B \in M_{m\times k}(\mathbb{K})$ dos matrices. Se tienen las desigualdades:*

$$r(AB) \leq r(A), \quad r(AB) \leq r(B)$$

Demostración. No es nada sorprendente que así sea. Como ya hemos visto, multiplicar una matriz (como A) por la derecha por otra matriz, no es más que construir otra matriz cuyos vectores columna son combinaciones lineales de los vectores columna de la inicial. Con estas operaciones uno no puede conseguir más vectores l.i. de los que había. Como mucho tendrá los mismos, luego el rango no puede aumentar. El mismo razonamiento se aplica a los vectores fila cuando multiplicamos por la izquierda. QED

Pero si una de las dos matrices tiene inverso (por supuesto es cuadrada), entonces el rango de la otra no varía:

Teorema 2.6.3 *Sea $A \in M_{n\times m}(\mathbb{K})$ y un matriz regular $B \in M_{m\times m}(\mathbb{K})$. Entonces:*

$$r(AB) = r(A)$$

Si consideramos $C \in M_{n\times n}(\mathbb{K})$ regular, se tiene también:

$$r(CA) = r(A)$$

Demostración. La razón es evidente del teorema anterior. Al multiplicar por una matriz cuadrada lo que estamos haciendo, desde otro punto de vista, es un cambio de base. La dimensión de la envolvente lineal no varía, es decir, el rango permanece constante. QED

Los subespacios de un espacio vectorial (de dimensión finita), se pueden definir de varias maneras, como ya hemos adelantado en otro punto. La forma implícita consiste en escribir los vectores en una base (del espacio total), y someter a las coordenadas a unas ecuaciones lineales homogéneas, es decir igualadas a cero.

Teorema 2.6.4 *Consideremos el espacio vectorial de dimensión n, $\mathbb{K}^n$. Los vectores $x = (x_1, x_2, \ldots, x_n)$ de este espacio que satisfacen las m ecuaciones:*

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= 0 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= 0 \\ &\cdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= 0 \end{aligned}$$

forman un subespacio vectorial de $\mathbb{K}^n$.

La demostración es inmediata debido a la linealidad de las ecuaciones.

El sistema anterior se puede escribir como una ecuación con matrices. Sean las matrices:

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}, \quad X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

La ecuación se puede escribir como:

$$AX = 0$$

y de aquí es inmediato probar que el conjunto de soluciones es un espacio vectorial, subespacio de $\mathbb{K}^n$. La dimensión de este subespacio es fácil de establecer. La matriz A se puede transformar en una matriz escalón reducida por filas mediante operaciones elementales. Como éstas son equivalentes a multiplicar la matriz por la izquierda por matrices regulares, el sistema de ecuaciones tiene las mismas soluciones:

$$PAX = 0 \Leftrightarrow AX = 0$$

De esta forma, el número de ecuaciones que nos queda es justamente el rango de la matriz A.

Teorema 2.6.5 *La dimensión del subespacio definido por la ecuación $AX = 0$ es igual a la dimensión del espacio ($X \in \mathbb{K}^n$) menos el rango de la matriz $A \in M_{m\times n}(\mathbb{K})$.*

Para un espacio vectorial arbitrario (no necesariamente $\mathbb{K}^n$), la situación es la misma. Basta elegir una base y emplear coordenadas para encontrarnos en una situación igual a la descrita anteriormente. Volveremos a estudiar estos aspectos con más detalle cuando definamos las aplicaciones lineales.

La otra forma de definir un subespacio es como la envolvente de una familia de vectores. En este caso la dimensión es clara, es justamente el rango de esa familia de vectores, es decir el número máximo de vectores l.i. que podemos encontrar en esa familia. Cualquier vector del subespacio viene dado como una combinación lineal de los vectores de la familia que genera el subespacio:

$$x \in V \Leftrightarrow x = \sum_{i=1}^{k} \lambda_i v_i$$

donde $S = \{v_1, \ldots, v_k\}$ es la familia generadora de W. Téngase en cuenta que el vector x no determina unívocamente los coeficientes λ_i. Pero de entre los vectores de S podemos seleccionar un conjunto maximal de vectores l.i. Este conjunto, como ya hemos dicho muchas veces, genera W. Y no solo eso, es una base de W. De modo que en función de estos vectores las coordenadas sí son únicas.

Una manera práctica de calcular esta base es la siguiente. Supongamos que tenemos una base en el espacio vectorial V de partida, $\mathcal{B} = \{u_1, \ldots, u_n\}$ y que los vectores v_i que generan el subespacio tienen en esta base unas coordenadas:

$$\begin{aligned}
v_1 &= b_{11}u_1 + b_{21}u_2 + \cdots + b_{n1}u_n \\
v_2 &= b_{12}u_1 + b_{22}u_2 + \cdots + b_{n2}u_n \\
&\vdots \\
v_k &= b_{1k}u_1 + b_{2k}u_2 + \cdots + b_{nk}u_n
\end{aligned}$$

Cualquier vector del subespacio es:

$$x \in W \Rightarrow x = \sum_{i=1}^{k} \lambda_i v_i = \sum_{j=1}^{n} \left(\sum_{i=1}^{k} b_{ji}\lambda_i \right) u_j$$

es decir, si $x = \sum_{i=1}^{n} x_i u_i$, se tiene:

$$x_j = \sum_{i=1}^{k} b_{ji}\lambda_i$$

que es la expresión que deben tener las coordenadas de x para que este vector esté en W y en la que λ_i toman valores arbitrarios en $\mathbb{K}$. Estas expresiones son las ecuaciones paramétricas de W. En forma matricial, la ecuación es:

$$X = B\Lambda$$

donde las matrices X, B, Λ son respectivamente:

$$X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, \quad B = \begin{pmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2n} \\ \vdots & \vdots & & \vdots \\ b_{m1} & b_{m2} & \cdots & b_{mn} \end{pmatrix}, \quad \Lambda = \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_k \end{pmatrix}$$

Estas ecuaciones tienen cierto parecido con las que dimos anteriormente en forma implícita (de hecho para $\mathbb{K}^n$, pero válidas en cualquier espacio de dimensión n una vez que definamos una base). ¿Cómo pasamos de unas a otras? El proceso se conoce como eliminación de parámetros yendo hacia la primera ecuación ($X = B\Lambda \Rightarrow AX = 0$) o resolución del sistema yendo de la segunda la primera ($AX = 0 \Rightarrow X = B\Lambda$). Ambos procesos son ya conocidos y no insistiremos en ellos. La segunda ecuación tiene parámetros redundantes en general, debido a que los vectores que generan el subespacio no tiene porqué ser l.i. Y la primera puede tener ecuaciones redundantes como hemos dicho ya. En ambos casos la dimensión del subespacio es:

1. De $AX = 0$ se deduce $\dim W = n - r(A)$.

2. De $X = B\Lambda$ se deduce $\dim W = r(B)$

Ya veremos en otra ocasión nuevas interpretaciones de estos resultados en relación con las aplicaciones lineales.

Tema 3

Aplicaciones lineales

Aplicaciones lineales. Núcleo e Imagen. Representación matricial. Cambios de base. Espacios de aplicaciones lineales. Rango. Sistemas de ecuaciones lineales. Determinantes.

A lo largo de este capítulo $V, W, \ldots$, etc., denotarán espacios vectoriales sobre el cuerpo $\mathbb{K}$ (de característica diferente a 2, p.e., $\mathbb{R}$, $\mathbb{C}$).

3.1 Generalidades sobre aplicaciones lineales

Las aplicaciones más notables entre espacios vectoriales son aquellas que preservan sus estructuras. Tales aplicaciones se denominan lineales y sirven a una gran variedad de propósitos. En este capítulo vamos a estudiar algunas de sus propiedades y aplicaciones.

3.1.1 Definiciones

Definición 3.1.1 *Una aplicación $f\colon V \to W$ entre dos espacios vectoriales se dirá que es lineal si,*
i. $f(x+y) = f(x) + f(y)$, $\forall x, y \in V$,
ii. $f(\lambda x) = \lambda f(x)$, $\forall \lambda \in \mathbb{K}$.

En otras palabras f es un homomorfismo de grupos abelianos $(V, +)$, $(W, +)$ y conmuta con el producto por escalares.

Si f es inyectiva diremos que es un monomorfismo de espacios vectoriales, si es suprayectiva, diremos que es un epimorfismo y si es biyectiva diremos que f es un isomorfismo de espacios vectoriales.

Si $f_1, \ldots, f_r$ son aplicaciones lineales de V en W y $\lambda_1, \ldots, \lambda_r$ son escalares, definimos la combinación lineal $\lambda_1 f_1 + \cdots + \lambda_r f_r$ como una aplicación de V en W dada por

$$(\lambda_1 f_1 + \cdots + \lambda_r f_r)(x) = \lambda_1 f_1(x) + \cdots + \lambda_r f_r(x), \forall x \in V.$$

Proposición 3.1.1 *La combinación lineal de aplicaciones lineales es una aplicación lineal. Lo mismo ocurre con la composición de aplicaciones lineales.*

Demostración. Efectivamente, si f, g son aplicaciones lineales de V en W y λ, μ dos elementos del cuerpo $\mathbb{K}$, hemos definido $(\lambda f + \mu g)(x) = \lambda f(x) + \mu g(x)$, $\forall x \in V$. Entonces, $(\lambda f + \mu g)(x+y) = \lambda f(x+y) + \mu g(x+y) = \lambda(f(x)+f(y)) + \mu(g(x)+g(y)) = \lambda f(x) + \mu g(x) + \lambda f(y) + \mu g(y) = (\lambda f + \mu g)(x) + (\lambda f + \mu g)(y)$.

Análogamente, si $f\colon V \to W$, $g\colon W \to U$, son dos aplicaciones lineales, la composición $g \circ f\colon V \to U$ es lineal. $g \circ f(x+y) = g(f(x+y)) = g(f(x) + f(y)) = g(f(x)) + g(f(y)) = g \circ f(x) + g \circ f(y)$, y de la misma forma con el producto por escalares. QED

Nota. En la clase $\mathcal{V}_{\mathbb{K}}$ de todos los espacios vectoriales de dimensión finita sobre el cuerpo $\mathbb{K}$, se puede establecer una relación de equivalencia como sigue: $V \approx W$ si y sólo si existe un isomorfismo $f: V \to W$ de espacios vectoriales. Es un ejercicio sencillo comprobar que dicha relación es de equivalencia.

Desde este punto de vista dos espacios vectoriales isomorfos se pueden considerar idénticos y el problema de la clasificación de espacios vectoriales consiste en describir el conjunto de clases de equivalencia $\mathcal{V}_{\mathbb{K}}/\approx$. Como veremos inmediatamente tal conjunto es $\mathbb{N} \cup 0$.

Dentro de la clase de equivalencia $[V]$ de un espacio vectorial dado V, se hallan todos aquellos isomorfos a él. Desde un punto de vista abstracto, las realizaciones concretas de un espacio vectorial son irrelevantes pero no así desde un punto de vista práctico.

3.1.2 Algunos ejemplos

Ejemplo 3.1.1 Sea $V = \mathbb{K}$, y $\lambda \in \mathbb{K}$ entonces $f_\lambda(x) = \lambda x$, es una aplicación lineal.

Ejercicio 3.1.1 Probar que toda aplicación lineal de $\mathbb{K}$ en sí mismo es de la forma descrita en el ejemplo anterior 3.1.1.

Ejemplo 3.1.2 Sea $V = \mathbb{C}$. Si consideramos $\mathbb{C}$ como un espacio vectorial sobre el cuerpo de los números reales $\mathbb{R}$, la aplicación $f(z) = \bar{z}$ es lineal, pero no si consideramos $\mathbb{C}$ como un espacio vectorial sobre $\mathbb{C}$.

Ejemplo 3.1.3 Sea $V = \mathbb{R}^2$. $f(x, y) = (x \cos\alpha - y \operatorname{sen}\alpha, x \operatorname{sen}\alpha + y \cos\alpha)$, $\alpha \in \mathbb{R}$. f denota una rotación de ángulo α ($f = e^{i\alpha}$ en notación compleja).

Ejemplo 3.1.4 Consideremos ahora $V = \mathbb{K}[x]$, $f(P) = P'$, $\forall P \in \mathbb{K}[x]$. Denotaremos la aplicación lineal anterior ("tomar la derivada") por el símbolo D (o también ∂_x), entonces D^2, $D^3, \ldots$, etc. son aplicaciones lineales debido a la proposición 3.1.1 así como cualquier combinación lineal de ellas, por tanto

$$L = D^n + \lambda_1 D^{n-1} + \lambda_2 D^{n-2} + \cdots + \lambda_{n-1} D + \lambda_n,$$

es una aplicación lineal. Un tal objeto se denomina un operador diferencial lineal en $\mathbb{K}[x]$.

Ejemplo 3.1.5 Consideremos de nuevo $V = \mathbb{K}[x]$, y la aplicación $f: \mathbb{K}[x] \to \mathbb{K}[x]$, $f(P) = \int P(x)dx$ donde el símbolo $\int \cdot dx$ denota la primitiva con término constante 0. La aplicación lineal $\int P(x)dx$ también se denota por D^{-1}, esto es, $D^{-1}P = \int P(x)dx$. Cualquier potencia de esta aplicación lineal también es lineal D^{-2}, D^{-3}, etc. Una combinación lineal de los operadores D^k, $k \in \mathbb{Z}$, se denominará un operador pseudodiferencial en $\mathbb{K}[x]$.

Ejemplo 3.1.6 Sea $V = M_n(\mathbb{K})$. En el espacio vectorial de las matrices cuadradas la aplicación $f: M_n(\mathbb{K}) \to M_n(\mathbb{K})$, $f(A) = A^t$ es lineal. Si $B \in M_n(\mathbb{K})$, $f(A) = BA$ es lineal.

La siguiente proposición proporciona un método sistemático y eficaz para la construcción de aplicaciones lineales "a la carta".

Proposición 3.1.2 *Construcción de aplicaciones lineales. Sea V un espacio vectorial y $\mathcal{B} = \{e_1, \ldots, e_n\}$ una base de V. Sea W un espacio vectorial. Asociamos a cada elemento e_i de $\mathcal{B}$ un vector arbitrario $u_i \in W$. Definimos entonces $f: V \to W$ como sigue: Si $x = \sum_{i=1}^n x^i e_i$, $f(v) = \sum_{i=1}^n x^i u_i$. Entonces la aplicación f es lineal.*

Demostración. Es inmediato comprobar que f es lineal. En efecto, si $x = \sum_i x^i e_i$, $y = \sum_i y^i e_i$, entonces $x + y = \sum_i (x^i + y^i) e_i$ y por tanto $f(x+y) = \sum_i (x^i + y^i) u_i = \sum_i x^i u_i + \sum_i y^i u_i = f(x) + f(y)$. Análogamente para $f(\lambda x)$. QED

3.1.3 Algunas propiedades de las aplicaciones lineales

Definición 3.1.2 *Llamaremos núcleo de la aplicación lineal* $f\colon V \to W$ *al subconjunto* $\ker f = \{v \in V \mid f(v) = 0\} = f^{-1}(0)$. *La imagen de* f *se denotará por* $\operatorname{im} f$ *o bien* $f(V)$.

Proposición 3.1.3 $\ker f$ *e* $\operatorname{im} f$ *son subespacios vectoriales.*

Ejercicio 3.1.2 Probar la proposición anterior 3.1.3.

Ejercicio 3.1.3 Probar que $f^2 = 0$ si y sólo si $\operatorname{im} f \subset \ker f$.

Proposición 3.1.4 *Una aplicación lineal* $f\colon V \to W$ *es un monomorfismo si y sólo si* $\ker f = 0$*;* f *es un epimorfismo si y sólo si* $f(V) = W$

Ejercicio 3.1.4 Probar la proposición anterior 3.1.4.

Ejemplo 3.1.7 $V = \mathbb{K}[x]$, $D\colon \mathbb{K}[x] \to \mathbb{K}[x]$. $\operatorname{im} D = \mathbb{K}[x]$, y $\ker D =$ polinomios de grado cero.

Ejemplo 3.1.8 $V = M_n(\mathbb{K})$, $f(A) = [A, B] = AB - BA$. $\mathbb{K} = \mathbb{C}$, $n = 2$, $B = \begin{pmatrix} 1 & 0 \\ 1 & -1 \end{pmatrix}$; $\ker f = \left\{ \begin{pmatrix} c+d & 0 \\ c & d \end{pmatrix} \mid c, d \in \mathbb{C} \right\}$, $\operatorname{im} f = \left\{ \begin{pmatrix} a & -2a \\ b & -a \end{pmatrix} \mid a, b \in \mathbb{C} \right\}$.

Ejemplo 3.1.9 Sea V el espacio vectorial $V = \{0\}$. Hay una única aplicación $f\colon \{0\} \to W$, y es $f(0) = 0_W$. Esta aplicación se denotará habitualmente por $0 \to W$. Hay también una única aplicación $f\colon W \to \{0\}$, es la aplicación trivial $f(u) = 0$, $\forall u \in W$. Tal aplicación se denota habitualmente $W \to 0$.

Proposición 3.1.5 *1. Si* $W \subset V$ *es un subespacio de* V *y* $f\colon V \to U$ *es lineal, entonces* $f(W)$ *es un subespacio de* U.

2. Si S *es un subconjunto no vacío de* V, $\operatorname{lin}(f(S)) = f(\operatorname{lin} S)$.

Ejercicio 3.1.5 Probar la proposición anterior 3.1.5.

Proposición 3.1.6 *1. Si* $f\colon V \to U$ *es un monomorfismo y* S *es un sistema linealmente independiente, entonces* $f(S)$ *es linealmente independiente.*

2. Si S *es un sistema generador de* V *y* f *es suprayectiva, entonces* $f(S)$ *es un sistema generador de* U.

Ejercicio 3.1.6 Probar la proposición anterior 3.1.6.

Proposición 3.1.7 *Si* $f\colon U \to V$ *es un isomorfismo y* $\mathcal{B}$ *es una base de* V*, entonces* $f(\mathcal{B})$ *es una base de* U.

Demostración. Si f es un isomorfismo, entonces es un monomorfismo. Si $\mathcal{B}$ es una base cualquiera de V, entonces por la proposición 3.1.6 $f(\mathcal{B})$ es l.i. Por otro lado, f es también un epimorfismo, y por la proposición 3.1.6 $f(\mathcal{B})$ es un sistema generador. Por lo tanto $f(\mathcal{B})$ es una base. QED

Podemos concluir esta cadena de razonamientos con el siguiente teorema.

Teorema 3.1.1 *Una aplicación lineal* $f\colon V \to U$ *es un isomorfismo si y solo si para alguna base* $\mathcal{B}$ *de* V*,* $f(\mathcal{B})$ *es una base de* U.

Demostración. El "sólo si" es el enunciado de la proposición 3.1.7. El "si" se prueba fácilmente como sigue. Sea $\mathcal{B} = \{e_1, \ldots, e_n\}$ una base de V tal que $f(\mathcal{B})$ es una base de U.

Supongamos que $x \in \ker f$. Entonces $f(x) = 0$, pero $x = \sum x^i e_i$, y por tanto $f(x) = \sum x^i f(e_i) = 0$. Como los elementos de $f(\mathcal{B})$ son l.i., entonces $x^i = 0$, $i = 1, \ldots, n$ y por tanto $x = 0$. Como $\ker f = 0$, f es un monomorfismo.

Sea $y \in U$. Como $f(\mathcal{B})$ es un sistema generador de U, existen y^i, $i = 1, \ldots, n$ tales que $y = \sum y^i f(e_i)$, por lo tanto $y = f(\sum y^i e_i)$ y $y \in \operatorname{im} f$, por tanto f es suprayectiva y por tanto es un epimorfismo. QED

Del resultado anterior se desprende la siguiente caracterización de espacios vectoriales isomorfos.

Corolario 3.1.1 *Dos espacios vectoriales de dimensión finita U y V son isomorfos si y solo si* $\dim V = \dim U$.

3.2 Teoremas de isomorfía de espacios vectoriales

3.2.1 Primer teorema de isomorfía de espacios vectoriales

Teorema 3.2.1 *Sea $f\colon V \to W$ una aplicación lineal, entonces:*
i. Existe un isomorfismo $\bar{f}\colon V/\ker f \to f(V)$,
ii. existe un monomorfismo $i\colon f(V) \to W$,
iii. existe un epimorfismo $\pi\colon V \to V/\ker f$,
tales que, $f = i \circ \bar{f} \circ \pi$.

$$\begin{tikzcd} V \arrow[r, "f"] \arrow[d, "\pi"'] & W \\ V/\ker f \arrow[r, "\bar{f}"'] & \operatorname{im} f \arrow[u, leftarrow, "i"] \end{tikzcd}$$

Demostración. El monomorfismo i es la inclusión canónica, $i(w) = w$, $\forall w \in f(V)$.
El epimorfismo π es la proyección canónica, $\pi(v) = v + \ker f$, $\forall v \in V$.
El isomorfismo $\bar{f}$ se define como sigue:

$$\bar{f}(v + \ker f) = f(v), \quad \forall v \in V.$$

Debemos comprobar en primer lugar que $\bar{f}$ está bien definida. En efecto, si $v + \ker f = v' + \ker f$, entonces $v - v' \in \ker f$. Por lo tanto $f(v) = f(v')$, y $\bar{f}(v + \ker f) = \bar{f}(v' + \ker f)$.

Debemos probar además que $\bar{f}$ es lineal, suprayectiva e inyectiva. La prueba de la linealidad de $\bar{f}$ es rutinaria y la suprayectividad es evidente.

Calculemos por ejemplo, $\ker \bar{f}$. Si $v + \ker f \in \ker \bar{f}$, entonces $f(v) = 0$, y por lo tanto $v \in \ker f$, y $v + \ker f = \ker f$ que es el cero del espacio cociente.

Finalmente calculemos $i \circ \bar{f} \circ \pi(v) = i \circ \bar{f}(v + \ker f) = i(f(v)) = f(v)$. QED

Corolario 3.2.1 $\dim V = \dim \ker f + \dim f(V)$.

Demostración. Efectivamente, como $\bar{f}$ es un isomorfismo, tenemos que $\dim V/\ker f = \dim f(V)$, pero $\dim V/\ker f = \dim V - \dim \ker f$. QED

La composición de dos aplicaciones $f\colon U \to V$, $g\colon V \to W$ se dirá exacta si $\ker g = \operatorname{im} f$. Se tienen las siguientes propiedades elementales:

Ejercicio 3.2.1 Probar las siguientes afirmaciones.

1. $0 \to U \xrightarrow{f} V$ es exacta $\iff f$ es un monomorfismo.
2. $U \xrightarrow{f} V \to 0$ es exacta $\iff f$ es un epimorfismo.
3. $0 \to U \to V \to W \to 0$ es exacta $\iff V/U \cong W$.

3.2.2 Otros teoremas de isomorfía

Además del teorema de isomorfía anterior existen otros teoremas que permiten identificar espacios vectoriales construidos a partir de operaciones suma y cociente. Citaremos dos.

Teorema 3.2.2 *Sean $W \subset U \subset V$ un espacio vectorial y dos subespacios contenidos el uno en el otro. Entonces:*

$$\frac{V/W}{U/W} \cong V/U.$$

Demostración. En primer lugar notemos que U/W es un subespacio de V/W ya que $u+W \in V/W$, $\forall u \in U$. Definamos la aplicación $f: V/W \to V/U$ por $f(v+W) = v+U$, $\forall v \in V$. Es claro que esta aplicación está bien definida ya que $W \subset U$. Por otro lado la aplicación es suprayectiva y además $\ker f = \{v+W \mid v \in U\} = U/W$. Entonces por el primer teorema de isomorfía, teorema 3.2.1,

$$\frac{V/W}{U/W} = \frac{V/W}{\ker f} \cong f(V/W) = V/U.$$

QED

Teorema 3.2.3 *Sean U, V dos subespacios de un espacio vectorial. Entonces se verifica:*

$$\frac{U}{U \cap V} \cong \frac{U+V}{V}.$$

Demostración. Definamos la aplicación $f: U \to (U+V)/V$ por $f(u) = u+V$. La aplicación f es suprayectiva. En efecto, si $x+V \in (U+V)/V$, entonces $x = u+v$, con $u \in U$, $v \in V$. Por tanto $f(u) = x+V$. Si calculamos el núcleo de f tenemos, $u \in \ker f$ si $f(u) = \mathbf{0}$, por tanto $u \in V$, entonces $\ker f = U \cap V$. Aplicamos el primer teorema de isomorfía a f y obtenemos el enunciado. QED

3.3 Representación matricial y cambios de base

3.3.1 Representación matricial de una aplicación lineal

Sea $f: V \to W$ una aplicación lineal entre los espacios vectoriales V, W (ambos sobre $\mathbb{K}$). Sea $\mathcal{B}_V = \{e_j\}_{j=1}^n$ una base de V ($\dim V = n$) y $\mathcal{B}_W = \{u_i\}_{i=1}^m$ una base de W ($\dim W = m$). La imagen del vector e_j por f, $f(e_j)$, será una combinación lineal de los vectores de $\mathcal{B}_W$, esto es:

$$f(e_j) = A_{1j}u_1 + A_{2j}u_2 + \cdots + A_{mj}u_m, \quad j = 1, \ldots, n. \tag{3.1}$$

Los coeficientes A_{ij}, $i = 1, \ldots, m$, $j = 1, \ldots, n$, pueden organizarse como una matriz $m \times n$,

$$A = \begin{pmatrix} A_{11} & A_{12} & \cdots & A_{1n} \\ A_{21} & A_{22} & \cdots & A_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ A_{m1} & A_{m2} & \cdots & A_{mn} \end{pmatrix}.$$

La primera columna está formada por los coeficientes de la imagen de e_1 en la base u_i,..., la j-ésima columna está formada por las coordenadas de la imagen del vector e_j en la base u_i, etc. Si colocamos los vectores u_i formando una matriz $1 \times m$, $(u_1, \ldots, u_m)$ podemos escribir:

$$(f(e_1), \ldots, f(e_n)) = (u_1, \ldots, u_m) \cdot A.$$

Llamaremos a la matriz A la representación matricial de f en las bases $\mathcal{B}_V$ y $\mathcal{B}_W$ y en ocasiones por motivos de precisión en la notación escribiremos también $A(f; \mathcal{B}_V, \mathcal{B}_W)$ con indicación expresa de las bases respecto de las cuales está definida.

Nota. Es evidente que en la definición de la matriz asociada a una aplicación lineal en dos bases dadas hay una serie de convenciones arbitrarias, por ejemplo en la ecuación (3.1), podríamos haber etiquetado los coeficientes en el desarrollo como sigue:

$$f(e_j) = A_{j1}u_1 + A_{j2}u_2 + \cdots + A_{jm}u_m, \quad j = 1, \ldots, n.$$

Podemos mantener la definición de A o cambiar filas por columnas. En uno u otro caso algunas de las expresiones que aparecerán a continuación adoptarán formas diferentes. En cualquier caso el conjunto de convenciones que hemos adoptado son consistentes con la notación tensorial comúnmente aceptada en nuestros días (aunque no siempre ha sido así) y a ella nos atendremos en todo lo que sigue.

La utilidad de la representación matricial de A de f se aprecia mejor si describimos como se transforman las coordenadas de un vector x al tomar su imagen $f(x)$. Si denotamos por x^j las coordenadas de x en la base e_j, $x = \sum_{j=1}^n x^j e_j$, y si denotamos por y^i las coordenadas de $f(x)$ en la base u_i, $f(x) = \sum_{i=1}^m y^i u_i$, tendremos,

$$f(x) = f\left(\sum_{j=1}^n x^j e_j\right) = \sum_{j=1}^n x^j f(e_j) = \sum_{j=1}^n \sum_{i=1}^m x^j A_{ij} u_i,$$

por tanto $\sum_{i=1}^m y^i u_i = \sum_{i=1}^m \left(\sum_j^n x^j A_{ij}\right) u_i$ lo que implica que $y^i = \sum_{j=1}^n A_{ij} x^j$. Denotando por X el vector columna $X = \begin{pmatrix} x^1 \\ \vdots \\ x^n \end{pmatrix}$ y análogamente con $Y = \begin{pmatrix} y^1 \\ \vdots \\ y^m \end{pmatrix}$ tenemos:

$$Y = A \cdot X, \tag{3.2}$$

o escribiéndolo explícitamente:

$$\begin{pmatrix} y^1 \\ \vdots \\ y^m \end{pmatrix} = \begin{pmatrix} A_{11} & A_{12} & \cdots & A_{1n} \\ A_{21} & A_{22} & \cdots & A_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ A_{m1} & A_{m2} & \cdots & A_{mn} \end{pmatrix} \cdot \begin{pmatrix} x^1 \\ \vdots \\ x^n \end{pmatrix}.$$

La ecuación anterior $Y = AX$ describe como actúa f sobre las coordenadas de los vectores en unas bases dadas, pero las bases no son únicas. Estudiaremos a continuación como cambia la ecuación anterior cuando cambiamos de bases.

3.3.2 Representación matricial de la composición de aplicaciones

Consideremos ahora dos aplicaciones lineales $f\colon U \to V$ y $g\colon V \to W$. La composición de las aplicaciones f y g es de nuevo una aplicación lineal $g \circ f\colon U \to W$ (proposición 3.1.1). Si fijamos bases $\mathcal{B}_U = \{u_i\}$, $\mathcal{B}_V = \{v_j\}$ y $\mathcal{B}_W = \{w_k\}$ de los espacios U, V y W respectivamente, tendremos una representación matricial A para la aplicación f en las bases $\mathcal{B}_U$ y $\mathcal{B}_V$, una representación matricial B para la aplicación lineal g en las bases $\mathcal{B}_U$ y $\mathcal{B}_W$ y una representación matricial C para $g \circ f$ en las bases $\mathcal{B}_U$ y $\mathcal{B}_W$. La pregunta que nos hacemos es ¿qué relación existe entre A, B y C?

Notemos que por definición de representación matricial, ecuación (3.1), tenemos para los tres casos:

$$\begin{aligned} f(u_i) &= A_{1i}v_1 + A_{2i}v_2 + \cdots + A_{mi}v_m, \quad i = 1, \ldots, n && (3.3) \\ g(v_j) &= B_{1j}w_1 + B_{2j}w_2 + \cdots + B_{rj}w_r, \quad j = 1, \ldots, m && (3.4) \\ g \circ f(u_i) &= C_{1i}w_1 + C_{2i}w_2 + \cdots + C_{ri}w_r, \quad i = 1, \ldots, n. && (3.5) \end{aligned}$$

Por lo tanto, desarrollando $g \circ f(u_i)$ en la ecuación (3.5) tenemos,

$$\begin{aligned} g \circ f(u_i) &= g(f(u_i)) = g(A_{1i}v_1 + A_{2i}v_2 + \cdots + A_{mi}v_m) \\ &= \sum_{l=1}^m A_{li} g(v_l) = \sum_{l=1}^m A_{li} \left(\sum_{k=1}^r B_{kl} w_k\right) = \sum_{l=1,k=1}^{m,r} A_{li} B_{kl} w_k, \end{aligned}$$

y comparando con el segundo miembro de la ecuación (3.5) tendremos

$$\sum_{k=1}^{r} C_{ki} w_k = \sum_{l=1,k=1}^{m,r} A_{li} B_{kl} w_k.$$

Como los vectores w_k forman una base, tendremos por tanto,

$$C_{ki} = \sum_{l=1}^{m} B_{kl} A_{li},$$

que en lenguaje de matrices, corresponde a la formula:

$$C = BA.$$

Hemos concluido, demostrando así el siguiente teorema.

Teorema 3.3.1 *La matriz correspondiente a la composición de dos aplicaciones lineales en bases dadas se obtiene multiplicando las correspondientes matrices de cada una de las aplicaciones en el orden contrario a su composición.*

> Notas. 1. Este teorema justifica "a posteriori" la definición del producto de matrices. El producto de matrices no es por tanto una operación exótica que tiene interesantes (y sorprendentes) aplicaciones sino que no es más que una manera especial de escribir la composición de aplicaciones y de ello emanan todas sus propiedades.
>
> 2. La no conmutatividad del producto de matrices simplemente refleja el hecho de que la composición de aplicaciones no es en general conmutativa.

Ejercicio 3.3.1 Escribir la matriz que representa a las aplicaciones lineales D y L del ejemplo 3.1.4 en la base $\{1, x, \ldots, x^n, \ldots\}$.

Ejercicio 3.3.2 Sea V un espacio vectorial de dimensión n y α una permutación de n elementos. Considérese la aplicación lineal $f_\alpha \colon V \to V$ definida por $f_\alpha(e_i) = e_{\alpha(i)}$ donde $\mathcal{B} = \{e_i\}_{i=1}^n$. Escribir la matriz A_α asociada a f_α en la base $\mathcal{B}$. Probar que $A_\alpha A_\beta = A_{\alpha\beta}$, $\forall \alpha, \beta \in S_n$.

3.3.3 Cambios de base

Punto de vista pasivo

Este es el punto de vista que adoptamos en el capítulo precedente, sección 2.5.3. Sea V un espacio vectorial en el que cambiamos las bases; la base $\mathcal{B}_V = \{u_1, \ldots, u_n\}$ se cambia a la nueva base $\mathcal{B}'_V = \{u'_1, \ldots, u'_n\}$. Los vectores $x \in V$ no son alterados, pero sus coordenadas variarán:

$$x = \sum_{i=1}^{n} x^i u_i = \sum_{i=1}^{n} x'^i u'_i.$$

El vector columna $X = (x^i)$ es el vector de las coordenadas antiguas y $X' = (x'^i)$ es el vector columna de las coordenadas nuevas. La relación entre ambas está proporcionada por

$$X' = P \cdot X, \tag{3.6}$$

con $u_i = \sum_{j=1}^n P_{ji} u'_j$ como en la ecuación (2.1), esto es, P es la matriz del cambio de base,

$$(u_1, \ldots, u_n) = (u'_1, \ldots, u'_n) \cdot P.$$

Si escribimos los vectores de la nueva base $\mathcal{B}'_V$ en función de los de la antigua base $\mathcal{B}_V$, tendremos

$$u'_i = \sum_{j=1}^{n} Q_{ji} u_j,$$

con Q la matriz inversa de P, y entonces

$$X' = Q^{-1} X.$$

Nota. El punto de vista pasivo es el más habitual cuando se trata de describir principios de invariancia relativista en Física. En efecto los vectores de un espacio vectorial representan habitualmente "estados" de un sistema físico y las leyes de la Física no dependen de la base que escojamos para escribirlas, esto es, son independientes del "sistema de referencia" que utilicemos para describirlas.

Hemos de notar que un cambio de base define un isomorfismo ϕ del espacio vectorial V a través de la fórmula (ver proposición 3.1.2)

$$\phi(u_i) = u'_i, \quad \forall i = 1, \dots n.$$

Debemos interpretar que este isomorfismo no está modificando los vectores de V sino solamente los observadores, esto es, las bases utilizadas para describir los vectores en coordenadas.

Nótese también que esta correspondencia entre cambios de bases e isomorfismos es biunívoca una vez que fijamos una base dada.

Punto de vista activo

A veces resulta conveniente adoptar otro punto de vista para discutir los cambios de bases en espacios vectoriales. Imaginemos ahora que la base $\mathcal{B}_V$ está fijada pero tenemos una transformación lineal $\phi\colon V \to V$ que cambia los vectores, $x \mapsto \phi(x)$. Esta transformación lineal enviará los vectores u_i de la base $\mathcal{B}_V$ a los de un nuevo sistema u'_i, $\phi(u_i) = u'_i$. Si la aplicación ϕ es un isomorfismo, los vectores u'_i serán una base $\mathcal{B}'_V$ de V. La representación matricial de ϕ en la base $\mathcal{B}_V$ estará dada por:

$$\phi(u_i) = u'_i = \sum_{j=1}^{n} \phi_{ji} u_j.$$

Pero ahora lo que nos importa no es la nueva base, sino el cambio de los vectores. Así, queremos obtener las coordenadas del nuevo vector $x' = \phi(x)$ respecto de la base $\mathcal{B}_V$, esto es:

$$\phi(x) = \sum_{i=1}^{n} x'_i u_i.$$

Se obtiene que $\phi(x) = \sum_i x^i \phi(u_i) = \sum_{i,j} x^i \phi_{ji} u_j$ y por tanto, $x'^j = \sum_i \phi_{ji} x^i$, o matricialmente,

$$X' = \Phi X,$$

donde Φ es la matriz con coeficientes ϕ_{ij}.

Nota. El punto de vista activo se utiliza cuando estamos interesados en estudiar el efecto de transformaciones en los estados de un sistema físico y sus propiedades de simetría. En lo que sigue, cuando hablemos de cambios de base y cambios de coordenadas estaremos asumiendo el punto de vista pasivo.

3.3.4 Representación matricial en bases diferentes

Sea $f\colon V \to W$ una aplicación lineal y $\mathcal{B}_V$, $\mathcal{B}_W$ bases respectivamente de V y W. Sea $\phi_V\colon V \to V$ un isomorfismo en V definiendo una nueva base de V, $\phi_V(\mathcal{B}_V) = \mathcal{B}'_V$ y $\phi_W\colon W \to W$ un isomorfismo en W definiendo una nueva base $\phi_W(\mathcal{B}_W) = \mathcal{B}'_W$ de W. Denotaremos por v_i los vectores de $\mathcal{B}_V$, esto es $\mathcal{B}_V = \{v_i\}$; análogamente $\mathcal{B}'_V = \{v'_i\}$, $\mathcal{B}_W = \{w_j\}$ y $\mathcal{B}'_W = \{w'_j\}$. Además

$$v'_i = \sum_{j=1}^{n} P_{ji} v_j, \quad w'_l = \sum_{k=1}^{m} Q_{kl} w_k. \tag{3.7}$$

La representación matricial de f en las bases $\mathcal{B}_V$ y $\mathcal{B}_W$ vendrá dada por una matriz A definida por

$$f(v_i) = \sum_{k=1}^{m} A_{ki} w_k,$$

y en las bases $\mathcal{B}'_V$ y $\mathcal{B}'_W$,

$$f(v'_i) = \sum_{l=1}^{m} A'_{li} w'_l. \tag{3.8}$$

Por tanto, usando la ecuación (3.7) en la ecuación (3.8) tendremos,

$$f(v'_i) = f(\sum_{j=1}^{n} P_{ji} v_j) = \sum_{j=1}^{n} P_{ji} f(v_j) = \sum_{j=1}^{n} P_{ji} \left(\sum_{k=1}^{m} A_{kj} w_k \right),$$

y análogamente,

$$\sum_{k=1}^{m} A'_{ki} w'_k = \sum_{k=1}^{m} A'_{ki} \left(\sum_{l=1}^{m} Q_{la} w_l \right),$$

por tanto

$$\sum_{k=1,l=1}^{m,m} A'_{ki} Q_{lk} w_l = \sum_{j=1}^{n} \sum_{l=1}^{m} P_{ji} A_{lj} w_l,$$

y la independencia lineal de los vectores w_l implica que

$$\sum_{k=1}^{m} A'_{ki} Q_{lk} = \sum_{j=1}^{n} P_{ji} A_{lj}.$$

En otras palabras, utilizando notación matricial, tenemos,

$$QA' = AP,$$

y despejando A' en el primer miembro de la ecuación (esto es, multiplicando por Q^{-1} por la izquierda),

$$A' = Q^{-1} AP. \tag{3.9}$$

Ejercicio 3.3.3 Con los isomorfismos ϕ_V y ϕ_W podemos construir una nueva aplicación lineal $\tilde{f} = \phi_W^{-1} \circ f \circ \phi_V : V \to W$ tal y como nos indica el diagrama adjunto.

$$\begin{array}{ccc} V & \xrightarrow{\tilde{f}f} & W \\ {\scriptstyle \phi_V}\downarrow & & \downarrow{\scriptstyle \phi_V} \\ Vf & \xrightarrow[f]{} & W \end{array}$$

Probar que la representación matricial de $\tilde{f}$ en las bases $\mathcal{B}_V$ y $\mathcal{B}_W$ es precisamente A'.

Ejercicio 3.3.4 Probar que el rango de la matriz asociada a una aplicación lineal no depende de las bases en que se escriba.

Si particularizamos la situación anterior al caso en que f es un endomorfismo, esto es, una aplicación lineal $f: V \to V$, podemos fijar la misma base $\mathcal{B}_V$ en su dominio y en su rango. Cuando cambiemos de base, podemos realizar simultáneamente el cambio en su rango y su dominio con lo que tendremos que las matrices P y Q de la discusión anterior coincidirán y la fórmula para el cambio de base de una realización matricial A de f, resultará:

$$A' = P^{-1} AP. \tag{3.10}$$

En el próximo capítulo discutiremos el problema de determinar la expresión matricial más sencilla para un endomorfismo.

3.4 Espacios de aplicaciones lineales

3.4.1 El espacio dual de un espacio vectorial

La proposición 3.1.1 nos enseñó que las combinaciones lineales de aplicaciones lineales son de nuevo aplicaciones lineales. Este hecho nos conduce a considerar como candidatos a nuevos espacios vectoriales conjuntos cuyos elementos son aplicaciones lineales ya que podemos sumarlas y multiplicar por escalares. En particular si consideramos $\mathbb{K}$ como un espacio vectorial de dimensión 1 sobre el propio cuerpo $\mathbb{K}$, las aplicaciones lineales $f\colon V \to \mathbb{K}$ se llaman covectores o formas lineales sobre V y forman un espacio vectorial.

Proposición 3.4.1 *Sea V un espacio vectorial sobre $\mathbb{K}$. El conjunto de aplicaciones lineales $f\colon V \to \mathbb{K}$ forman un espacio vectorial sobre $\mathbb{K}$ denotado por V^* llamado el espacio dual de V. Además* $\dim V = \dim V^*$.

Demostración. Definimos la suma y el producto por escalares de aplicaciones lineales de la manera habitual (ver proposición 3.1.1). Con ellas V^* se convierte en un espacio vectorial sobre $\mathbb{K}$ tras una comprobación rutinaria de las propiedades de la suma y el producto por escalares.

Si $\mathcal{B} = \{e_1, \dots, e_n\}$ es una base de V, un covector $f\colon V \to \mathbb{K}$ tiene la forma $f(x) = \sum_i x^i \lambda_i$, donde $\lambda_i = f(e_i)$. Definamos ahora una familia de covectores $e^i\colon V \to \mathbb{K}$, $i = 1, \dots, n$, como sigue:

$$e^i(e_j) = \delta^i_j, \quad \forall i, j = 1, \dots, n,$$

equivalentemente $e^i(x) = x^i$, donde $x = \sum_i x^i e_i$. Probemos que el conjunto $\mathcal{B}^* = \{e^1, \dots, e^n\}$ es una base de V^*.

Si $f \in V^*$, sea $\lambda_i = f(e_i)$, entonces $f = \sum_i \lambda_i e^i$. En efecto, $\sum_i \lambda_i e^i(x) = \sum_i \lambda_i x^i = \sum_i f(x^i e_i) = f(x)$, y $\mathcal{B}^*$ es un sistema generador.

Probemos que $\mathcal{B}^*$ es libre. Supongamos que $\sum_j \mu_j e^j = 0$, entonces $\sum_j \mu_j e^j(x) = 0$ para todo $x \in V$. Tomemos $x = e_i$, entonces $0 = \sum_j \mu_j e^j(e_i) = \sum_j \mu_j \delta^j_i = \mu_i$, y $\mathcal{B}^*$ es libre. QED

Ejercicio 3.4.1 Probar que si $f(x) = 0$ para todo $f \in V^*$, entonces $x = 0$.

> Nota. Los espacios vectoriales V y V^* tienen la misma dimensión, por tanto son isomorfos de acuerdo con el corolario 3.1.1, pero no hay ningún isomorfismo canónico entre ambos. Para cada elección de una base $\mathcal{B}$ en V tenemos el isomorfismo proporcionado por la aplicación $\phi\colon V \to V^*$, $\phi(e_i) = e^i$.

Ejemplo 3.4.1 El espacio $\mathbb{K}^*$ se puede identificar con $\mathbb{K}$ escogiendo como base el covector f que envía 1 en 1.

Si consideramos el espacio vectorial de los polinomios $\mathbb{K}[x]$, cada elemento a de $\mathbb{K}$ define un covector f_a como sigue:

$$f_a(P) = P(a) \in \mathbb{K}.$$

Ejercicio 3.4.2 Probar que el conjunto de covectores f_{a_i}, $a_i \neq a_j$ si $i \neq j$, $i = 1, \dots, n+1$ son l.i.

Ejemplo 3.4.2 El espacio $(V^*)^*$ es canónicamente isomorfo a V. Por un lado $\dim(V^*)^* = \dim V^* = \dim V$, luego $(V^*)^*$ es isomorfo a V. Además hay una aplicación natural $\phi\colon V \to (V^*)^*$ definida por $\phi(x)(f) = f(x)$, $\forall f \in V^*$, $x \in V$.

Esta aplicación es un monomorfismo ya que si $\phi(x) = 0$, entonces $f(x) = 0$, para todo $f \in V^*$, por tanto $x = 0$. Por tanto ϕ es un isomorfismo.

3.4.2 Endomorfismos de un espacio vectorial

Una aplicación lineal f de un espacio vectorial V en sí mismo se denominará un endomorfismo de V. El conjunto de aplicaciones lineales de V, esto es, de endomorfismos de V, se denotará por $\mathrm{End}(V)$.

Al igual que ocurre con el espacio dual de V tenemos la siguiente proposición.

Proposición 3.4.2 $\mathrm{End}(V)$ *es un espacio vectorial de dimensión* $(\dim V)^2$.

Demostración. La demostración de que $\mathrm{End}(V)$ es un espacio vectorial es una repetición del caso del espacio dual. Construyamos una base de $\mathrm{End}(V)$. Sea $\mathcal{B} = \{e_i\}_{i=1}^n$, una base de V. Denotemos por $e_i^j : V \to V$ las aplicaciones definidas por

$$e_i^j(e_k) = \delta_k^j e_i, \quad \forall i, j, k = 1, \ldots, n,$$

esto es, si $x = \sum_i x^i e_i$, entonces $e_i^j(x) = x^j e_i$. La familia de aplicaciones $\tilde{\mathcal{B}} = \{e_i^j \mid i, j = 1, \ldots, n\}$ es una base de $\mathrm{End}(V)$.

Probemos que es un sistema generador. Sea $f \in \mathrm{End}(V)$. Entonces $f(e_i) = \sum_k \lambda_{ki} e_k$, $i = 1, \ldots, n$. Definamos la aplicación

$$\sum_{i,j=1}^{n} \lambda_{ji} e_j^i.$$

Entonces, $\sum_{i,j=1}^n \lambda_i^j e_j^i(x) = \sum_{i,j} \lambda_{ji} x^i e_j = \sum_i x^i (\sum_j \lambda_{ji} e_j) = \sum_i x^i f(e_i) = f(x)$.

El sistema $\tilde{\mathcal{B}}$ es l.i. Si $\sum_{i,j} \mu_j^i e_i^j = 0$, actuando sobre el vector e_k tendremos, $\sum_i \mu_k^i e_i = 0$, y por tanto $\mu_k^i = 0$, para todo i, k. QED

> Nota. De manera análoga se puede definir el espacio vectorial de aplicaciones lineales $V \to V^*$, $V^* \to V$, $V^* \to V^*$, etc. Todos estos espacios vectoriales tienen la misma dimensión n^2, si $n = \dim V$, pero son diferentes. Todos ellos son ejemplos de espacios de tensores de tipo $(0, 2)$, $(2, 0)$, $(1, 1)$ respectivamente como se verá más adelante.

3.4.3 Otros espacios de aplicaciones lineales

Denotaremos por $\mathrm{Hom}(V, W)$ el conjunto de aplicaciones lineales $f : V \to W$.

Proposición 3.4.3 *El conjunto* $\mathrm{Hom}(V, W)$ *es un espacio vectorial de dimensión* $\dim V \dim W$.

La demostración es análoga a la de la proposición 3.4.2.

Ejercicio 3.4.3 Calcular la dimensión de los espacios vectoriales: $\mathrm{Hom}(V^*, W^*)$, $\mathrm{Hom}(V, \mathrm{Hom}(V, U))$ y $\mathrm{Hom}(\mathrm{End}(V), U)$.

Ejercicio 3.4.4 Probar que existe un isomorfismo canónico entre los espacios vectoriales:

$$\mathrm{Hom}(V, \mathrm{Hom}(W, U)), \ \mathrm{Hom}(\mathrm{Hom}(V^*, W), U)$$

3.5 Rango de una aplicación

Retomamos el concepto de rango ya introducido en la sección 2.4 para familias de vectores de un espacio vectorial V y para matrices en 2.6. Relacionaremos dicho concepto con propiedades de aplicaciones lineales.

Definición 3.5.1 *Si $f : V \to W$ es una aplicación lineal, llamaremos rango de f a la dimensión de $f(V)$, o en otras palabras al número de vectores independientes en la imagen de una base cualquiera de V. El rango de una aplicación se denotará por $r(f)$ o más explícitamente rango(f).*

De la definición se deduce inmediatamente que si $\mathcal{B}_V$ es una base de V, entonces $r(f) = r(f(\mathcal{B}_V))$, ya que $f(V) = \operatorname{lin} f(\mathcal{B}_V)$, y por tanto $\dim f(V) = r(f(\mathcal{B}_V))$.

La relación entre el rango de una matriz A y el rango de una aplicación lineal está dada por la siguiente proposición.

Proposición 3.5.1 *Si A es una representación matricial de $f\colon V \to W$, entonces:*

$$r(f) = r_c(A).$$

Demostración. Recordemos que el rango por columnas de A, $r_c(A)$ es el número de vectores columna l.i. Es fácil ver que los vectores columna A^j, A^k son l.i. si y sólo si $f(e_j)$ y $f(e_k)$ son l.i. En efecto, $\lambda f(e_j) + \mu f(e_k) = 0$ es equivalente a $(\lambda A_{lj} + \mu A_{lk})e_l = 0$, lo cual es cierto si y sólo si $\lambda A_{lj} + \mu A_{lk} = 0$. El argumento anterior nos indica que los vectores columna $A^{j_1}, \ldots, A^{j_r}$, serán l.i. si y sólo si $f(e_{j_1}), \ldots, f(e_{j_r})$ son l.i., y por tanto la proposición queda probada. QED

Corolario 3.5.1 *Si A' es equivalente a A, entonces*

$$r_c(A) = r_c(A').$$

Demostración. En efecto, la $\dim f(V)$ no depende de las bases que escojamos. QED

> Nota. El teorema 2.6.1 nos mostró que el rango de filas y el rango de columnas de una matriz coinciden. Por tanto los resultados anteriores se pueden enunciar directamente utilizando el rango de la matriz representando a la aplicación f. En cualquier caso es significativo que en las demostraciones es el rango por columnas el que aparece de manera natural. ¿Cómo aparece el rango por filas en este contexto?

En primer lugar observaremos que rango de filas de A = rango de columnas de A^t donde $(A^t)_{ia} = A_{ai}$. Nos interesa hallar una aplicación lineal relacionada con f tal que su representación matricial sea A^t.

Definición 3.5.2 *Sea $f\colon V \to W$ es una aplicación lineal; llamaremos aplicación transpuesta o dual de f y se denotará por f^* a la aplicación $f^*\colon W^* \to V^*$ definida por*

$$(f^*(\alpha))(v) = \alpha(f(v)), \quad \forall \alpha \in W^*, v \in V.$$

Proposición 3.5.2 *Si A es la representación matricial de f en las bases $\mathcal{B}_V$, $\mathcal{B}_W$ entonces A^t es la representación matricial de f^* en las bases duales $\mathcal{B}_V^*$, $\mathcal{B}_W^*$.*

Demostración. En efecto, si $\mathcal{B}_V = \{e_1, \ldots, e_n\}$, la base dual $\mathcal{B}_V^* = \{e^1, \ldots, e^n\}$ se define como $e^j(e_i) = \delta_i^j$ tal y como vimos en la proposición 3.4.1 y de manera análoga para $\mathcal{B}_W = \{u_1, \ldots, u_m\}$ y su base dual $\mathcal{B}_W^* = \{u^1, \ldots, u^m\}$. Calculemos $f^*(u^i) = \sum_{j=1}^n A_{ji}^* e^j$, pero por otro lado $f^*(u^i)(e_i) = u^i(f(e_i)) = u^i(\sum_{j=1}^m A_{ji} u_j) = A_{ji}$. Si evaluamos $\sum_{j=1}^n A_{ji}^* e^j$ en e_i, obtenemos $A_{ij}^* = A_{ji}$, y por tanto $A^* = A^t$. QED

Por tanto $r_f(A) = r(f^*)$.

Proposición 3.5.3 $r(f) = \dim V - \dim \ker f$.

Demostración. En efecto, $f\colon V \to W$ y por el primer teorema de isomorfía, $V/\ker f \cong f(V)$, entonces $\dim V - \dim \ker f = \dim f(V) = r(f)$. QED

Podemos por tanto ofrecer otra demostración del teorema 2.6.1 ligada a las propiedades y estructura de las aplicaciones lineales.

Teorema 3.5.1 $r_f(A) = r_c(A)$.

Demostración. Tenemos $r_f(A) = r_c(A^t) = r(f^*) = \dim W^* - \dim \ker f^*$.

Calculemos $\ker f^*$. Si $\alpha \in \ker f^*$, entonces $f^*(\alpha) = 0$. Por tanto $f^*(\alpha)(v) = 0$, $\forall v \in V$, y entonces $\alpha(f(v)) = 0$ para todo v, y así $\alpha(f(V)) = 0$.

Sea $\mathcal{B}_W$ una base de W adaptada a $f(V)$, esto es $\mathcal{B}_W = \{u_1, \ldots, u_r, u_{r+1}, \ldots, u_m\}$, y $\{u_1, \ldots, u_r\}$ es una base de $f(V)$ (notemos que entonces rango $f = r$). Por tanto la condición $\alpha(f(V)) = 0$ implica que $\alpha = \lambda_{r+1}u^{r+1} + \cdots + \lambda_m u^m$. En efecto, un elemento general de W^* tendrá la forma $\alpha = \lambda_1 u^1 + \cdots + \lambda_m u^m$, pero $u_i \in f(V)$, $i = 1, \ldots, r$, por tanto $\alpha(u_i) = 0$, $i = 1, \ldots, r$, y por tanto $\lambda_1 = \cdots = \lambda_r = 0$. Concluimos que $\dim \ker f^* = m - r$, y así rango $f^* = m - (m - r) = r$. QED

3.6 Sistemas de ecuaciones lineales

En muchas ocasiones los textos de álgebra lineal comienzan con una discusión de sistemas de ecuaciones lineales. Hemos pospuesto deliberadamente retrasar tal discusión hasta este momento y tratar los sistemas de ecuaciones lineales como una mera aplicación de la teoría de aplicaciones lineales.

En cualquier caso, un sistema de m ecuaciones lineales con n incógnitas $x_1, \ldots, x_n$ consiste en una familia de ecuaciones

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2 \\ &\vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= b_m \end{aligned}$$

con $a_{ij} \in \mathbb{K}$, $b_i \in \mathbb{K}$. El problema consiste en determinar cuando existen y cuáles son los valores de $x_1, \ldots, x_n$ en $\mathbb{K}$ que satisfacen dichas ecuaciones.

Podemos escribir el sistema anterior en forma matricial. Si A denota la matriz $m \times n$ definida por $A = (a_{ij})$, $i = 1, \ldots, m$, $j = 1, \ldots, n$ y X, B son las matrices columnas $n \times 1$ y $m \times 1$ respectivamente,

$$X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \quad B = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix},$$

tenemos:

$$A \cdot X = B.$$

Si pensamos que X y B son vectores en $\mathbb{K}^n$ y $\mathbb{K}^m$ respectivamente, y A denota una aplicación lineal $f\colon \mathbb{K}^n \to \mathbb{K}^m$ en las bases canónicas, tenemos que el problema de resolver el sistema anterior, es equivalente a determinar el conjunto $f^{-1}(B)$.

El siguiente teorema resuelve todas estas cuestiones. Denotaremos por $(A \mid B)$ la matriz que se obtiene añadiendo el vector columna B a la matriz A y se llamará matriz extendida de A por B.

Teorema 3.6.1 *Rouché–Frobenius. Dado un sistema de ecuaciones $A \cdot X = B$, de m ecuaciones con n incógnitas, el sistema posee solución si y sólo si $r(A) = r(A \mid B)$. Además si X_0 es una solución, el conjunto de todas las soluciones es $X_0 + \ker A$.*

Demostración. $\Leftarrow$) Supongamos que $r(A) = r(A \mid B)$. Entonces $r_c(A) = r_c(A \mid B)$, por tanto $\dim(\mathrm{lin}\{A^1, A^2, \ldots, A^n\}) = \dim(\mathrm{lin}\{A^1, \ldots, A^n, B\})$ lo que quiere decir que $B \in \mathrm{lin}\{A^1, \ldots, A^n\}$, esto es existen números x_{0i} tales que $B = x_{01}A^1 + \cdots + x_{0n}A^n = A \cdot X_0$, y el vector X_0 es una solución.

$\Rightarrow$) Si X_0 es una solución de $A \cdot X = B$, entonces $A \cdot X_0 = B$ y desarrollando el producto por columnas tenemos $x_{01}A^1 + x_{02}A^2 + \cdots + x_{0n}A^n = B$, luego $B \in \mathrm{lin}\{A^1, \ldots, A^n\}$, $\Rightarrow \dim(\mathrm{lin}\{A^1, \ldots, A^n\}) = \dim(\mathrm{lin}\{A^1, \ldots, A^n, B\}) \Rightarrow r(A) = r(A \mid B)$.

Finalmente si X_0 es una solución, el conjunto de soluciones es $X_0 + \ker A$. En efecto, si X_1 es otra solución, $A \cdot (X_1 - X_0) = 0 \Rightarrow X_1 - X_0 \in \ker A$. QED

Definición 3.6.1 *El sistema $A \cdot X = B$ se dirá compatible si $r(A) = r(A \mid B)$ e incompatible en caso contrario.*

Si el sistema $A \cdot X = B$ es compatible, dado que $r(A) = \dim \mathbb{K}^n - \dim \ker A$, el sistema tendrá una única solución si $\ker A = 0$, esto es, si $r(A) = \dim \mathbb{K}^n = n =$ número de incógnitas.

Podemos resumir la situación como:
– Si $r(A) \neq r(A \mid B) \Rightarrow$ Incompatible.
– Si $r(A) = r(A \mid B) \Rightarrow$ Compatible

$$\begin{cases} \text{Si } r(A) = n = \text{número de incógnitas } \Rightarrow \text{Compatible determinado} \\ \text{Si } r(A) < n \Rightarrow \text{Compatible indeterminado} \end{cases}$$

Ejemplo 3.6.1 Consideremos un sistema homogéneo, esto es un sistema tal que $B = 0$. Es obvio que el sistema siempre tiene una solución, la solución trivial $X = 0$. El sistema siempre es compatible. El sistema será compatible determinado cuando $r(A) =$ número de incógnitas, en cuyo caso sólo habrá una solución, la solución trivial. Si hay tantas ecuaciones como incógnitas, esto es equivalente a que la matriz A sea invertible. El sistema será compatible indeterminado cuando $r(A) < n$ y las soluciones serán todos los vectores en el núcleo de A.

3.7 Determinantes

El concepto de determinante es tan importante que aunque un tratamiento preciso de esta noción requiere adentrarnos en el ámbito del álgebra tensorial, está justificado el esfuerzo adicional que vamos a realizar en las próximas secciones por el gran rédito que nos ha de reportar en futuras empresas.

3.7.1 Aplicaciones multilineales

Definición 3.7.1 *Si V es un espacio vectorial sobre $\mathbb{K}$, diremos que una aplicación $f : V \times \overset{m}{\cdots} \times V \to \mathbb{K}$ es m-multilineal si,*
i. $f(x_1, \ldots, x_i + y_i, \ldots, x_m) = f(x_1, \ldots, x_i, \ldots, x_m) + f(x_1, \ldots, y_i, \ldots, x_m)$,
ii. $f(x_1, \ldots, \lambda x_i, \ldots, x_m) = \lambda f(x_1, \ldots, x_i, \ldots, x_m)$, $\forall i = 1, \ldots, m$, $x_1, \ldots, x_m, y_1, \ldots y_m \in V$, $\lambda \in \mathbb{K}$.

Nota. Es fácil comprobar que el conjunto de aplicaciones multilineales es un espacio vectorial de dimensión $(\dim V)^n$.

Una aplicación m–lineal f se dirá que es antisimétrica si,

$$f(x_{\alpha(1)}, \ldots, x_{\alpha(i)}, \ldots, x_{\alpha(m)}) = \epsilon(\alpha) f(x_1, \ldots, x_i, \ldots, x_m),$$

donde $\alpha \in S_m$ y $\epsilon(\alpha)$ es la signatura de la permutación.

Definición 3.7.2 *Una aplicación m–multilineal antisimétrica se llama una m–forma lineal.*

Si f es una aplicación m–multilineal y $\mathcal{B} = \{e_i\}$ es una base de V, f queda determinada por sus valores sobre las familias de vectores $e_{i_1}, \ldots, e_{i_m}$, esto es, si $x_k \in V$, tendremos que $x_k = \sum_{i_k} x_k^{i_k} e_{i_k}$, y entonces,

$$f(x_1, \ldots, x_m) = \sum_{i_1, \ldots, i_m = 1}^{n} x_1^{i_1} \ldots x_m^{i_m} f(e_{i_1}, \ldots, e_{i_m}).$$

Los números $\lambda_{i_1 i_2 \ldots i_m} = f(e_{i_1}, \ldots, e_{i_m})$ son las coordenadas de f en una cierta base.

Ejercicio 3.7.1 Si f es una m–forma lineal, se verifica $f(\ldots, x, \ldots, x, \ldots) = 0$, para todo $x \in V$.

Proposición 3.7.1 *Una m–forma lineal f en V queda determinada dando sus valores sobre las familias de vectores $e_{i_1}, \ldots, e_{i_m}$ tales que $1 \leq i_1 < i_2 < \ldots < i_m \leq n$, donde $\mathcal{B} = \{e_i\}_{i=1}^n$ es una base de V.*

Demostración. En efecto, si tomamos una familia cualquiera de m vectores en una base $\mathcal{B} = \{e_i\}$, $e_{j_1}, e_{j_2}, \ldots, e_{j_m}$ existe una permutación $\alpha \in S_m$ tal que $1 \leq \alpha(j_1) = i_1 < \alpha(j_2) = i_2 < \cdots \alpha(j_m) = i_m \leq n$. En efecto, no hay más que tomar la permutación que es la identidad en el complementario del conjunto $\{j_1, j_2, \ldots, j_m\}$, y la permutación que envía j_1 al mínimo de $j_1, j_2, \ldots, j_m$; que envía j_2 al mínimo del conjunto anterior excepto la imagen de j_1, etc. Entonces es evidente debido a la antisimetría de f que

$$f(e_{j_1}, \ldots, e_{j_m}) = \epsilon(\alpha) f(e_{i_1}, \ldots, e_{i_m}).$$

Por tanto f queda determinada sobre familias de vectores satisfaciendo la condición del enunciado. QED

Ejercicio 3.7.2 Probar que el número de familias de vectores que determinan una m–forma lineal es $\binom{n}{m}$.

Teorema 3.7.1 *Si* $\dim V = n$ *todas las* n*–formas lineales son proporcionales.*

Demostración. En efecto, según el ejercicio 3.7.2, el número de familias que determinan una n–forma en un espacio de dimensión n es $\binom{n}{n} = 1$. Por lo tanto si f es la aplicación definiendo tal forma, basta especificar el número

$$\lambda = f(e_1, e_2, \ldots, e_n),$$

donde $\mathcal{B} = \{e_i\}$ es una base cualquiera del espacio V. QED

Observemos que si f es una n–forma en un espacio de dimensión n, tendremos para cualquier familia de vectores $x_1, \ldots, x_n$, que,

$$f(x_1, \ldots, x_n) = \sum_{i_1, \ldots, i_n = 1}^{n} x_{1i_1} \cdots x_{ni_n} f(e_{i_1}, \ldots, e_{i_n}),$$

pero como f se anula sobre dos vectores idénticos, en la suma anterior índices i_k repetidos no contribuirán y sólo quedarán los términos en que todos los $i_1, \ldots, i_n$ sean diferentes, esto es los etiquetados por las permutaciones de $1, \ldots, n$, y por tanto:

$$f(x_1, \ldots, x_n) = \sum_{\alpha \in S_n} x_{1\alpha(1)} \cdots x_{n\alpha(n)} f(e_{\alpha(1)}, \ldots, e_{\alpha(n)}).$$

Por otro lado debido a la antisimetría de f, podemos reorganizar los vectores $e_{\alpha(1)}, \ldots, e_{\alpha(n)}$, en su argumento y llevarlos al orden natural $e_1, \ldots, e_n$. Para hacer esto podemos utilizar una descomposición cualquiera en transposiciones de α (lo cual es siempre posible debido a la proposición 1.2.2). Dado que cada transposición contribuye con un signo menos al valor de f, tras aplicar las transposiciones que convierte α en la identidad obtendremos un factor que será la paridad de la permutación α (recordar la proposición 1.2.3). Por tanto tendremos la siguiente fórmula para f,

$$f(x_1, \ldots, x_n) = \sum_{\alpha \in S_n} \epsilon(\alpha) x_{1\alpha(1)} \cdots x_{n\alpha(n)} f(e_1, \ldots, e_n). \tag{3.11}$$

Si g es otra n–forma tendremos aplicando la fórmula (3.11)

$$g(x_1, \ldots, x_n) = \sum_{\alpha \in S_n} \epsilon(\alpha) x_{1\alpha(1)} \cdots x_{n\alpha(n)} g(e_1, \ldots, e_n),$$

y por tanto si $g(e_1, \ldots, e_n) = \mu$ y $f(e_1, \ldots, e_n) = \lambda \neq 0$, tendremos que

$$g(x_1, \ldots, x_n) = \frac{\mu}{\lambda} f(x_1, \ldots, x_n),$$

confirmando de nuevo el resultado del teorema 3.7.1.

Definición 3.7.3 *Una* n*-forma no nula* Ω *en un espacio vectorial* V *de dimensión* n *se llama un volumen en* V*.*

Fijémonos que si seleccionamos una base $\mathcal{B} = \{e_1, \ldots, e_n\}$ en V podemos definir un volumen asociado a esta base a través de la fórmula

$$\Omega_{\mathcal{B}}(e_1, \ldots, e_n) = 1.$$

Notemos que aplicando la fórmula (3.11) obtenemos para tal volumen,

$$\Omega_{\mathcal{B}}(x_1, \ldots, x_n) = \sum_{\alpha \in S_n} \epsilon(\alpha) x_{1\alpha(1)} \cdots x_{n\alpha(n)}, \tag{3.12}$$

donde x_{ij} son las coordenadas del vector x_i en la base $\mathcal{B}$. Llamaremos a esta expresión el determinante de los vectores $x_1, \ldots, x_n$ respecto de la base $\mathcal{B}$. Notemos que el cambio de base sólo afectaría el valor de dicha expresión en un factor global según el teorema 3.7.1.

Ejercicio 3.7.3 Calcular dicho factor para las bases $\mathcal{B}'$, $\mathcal{B}$ con matriz de cambio de base P.

Ejercicio 3.7.4 Considérense en $\mathbb{R}^2$ el volumen Ω definido en la base canónica i, j por $\Omega(i, j) = 1$. Si u_1, u_2 son dos vectores, probar que el área del paralelogramo definido por ellos está dada por $\Omega(u_1, u_2)$.

Ejercicio 3.7.5 Considérese en $\mathbb{R}^3$ el volumen Ω definido en la base canónica i, j, k por $\Omega(i, j, k) = 1$. Pruébese que el volumen del paralelepípedo definido por los vectores u_1, u_2, u_3 es $\Omega(u_1, u_2, u_3)$.

Estos dos ejercicios muestran la razón de llamar volumen a una n–forma en un espacio vectorial. Podemos insistir en el hecho de que todas son proporcionales y por lo tanto multiplicándolas por un factor podemos convertir unas en otras; así dada una base podemos siempre normalizar un volumen con respecto a la citada base haciendo que $\Omega(e_1, e_2, \ldots, e_n) = 1$.

Notas. 1. Las fórmulas para el volumen en dimensión 2 y 3 obtenidas en los ejercicios 3.7.4, 3.7.5, son también válidas en cualquier dimensión. En efecto, si definimos en $\mathbb{R}^n$ la n–forma Ω tal que en la base canónica $e_1, \ldots, e_n$ toma el valor 1, entonces si $u_1, \ldots, u_n$ son n vectores en $\mathbb{R}^n$ el volumen (en el sentido de la medida del conjunto con respecto a la medida de Lebesgue en $\mathbb{R}^n$) del paralelepípedo definido por ellos (esto es, la envolvente convexa de $0, u_1, \ldots, u_n$) es precisamente $\Omega(u_1, \ldots, u_n)$. La demostración de este hecho utiliza las propiedades de transformación de la medida habitual en $\mathbb{R}^n$ que se estudian en un curso de cálculo avanzado.

2. Puede parecer extraño que el volumen de un conjunto de vectores pueda ser negativo. Tal posibilidad está asociada a la orientación de la familia de vectores que utilicemos. Los volúmenes en un espacio vectorial real definen una orientación en dicho espacio como sigue: Diremos que dos n–formas Ω, Ω' son equivalentes si existe un número real positivo $\lambda \in \mathbb{R}^+$ tal que $\Omega' = \lambda\Omega$. Tal relación es claramente una relación de equivalencia con exactamente dos clases. Llamaremos una orientación en nuestro espacio vectorial a cada una estas dos clases que denotaremos por $[+]$ y $[-]$. Supongamos que $\Omega \in [+]$, entonces diremos que una base $\mathcal{B} = \{v_1, \ldots, v_n\}$ está orientada positivamente si $\Omega(v_1, \ldots, v_n) > 0$ y negativa en caso contrario. Nótese que si una base está orientada negativamente basta intercambiar dos de sus vectores para obtener una orientada positivamente.

Si utilizamos la base $\mathcal{B}$ para identificar V con $\mathbb{K}^n$, la n–forma definida en $\mathbb{K}^n$ (en la base canónica) la llamaremos volumen canónico en $\mathbb{K}^n$ o también determinante a secas. Si consideramos entonces los vectores $a_1, \ldots, a_n$ en $\mathbb{K}^n$ y los identificamos con los vectores columna de una matriz llamaremos determinante de la matriz al determinante de los vectores $a_1, \ldots, a_n$, esto es si la matriz $n \times n$ A está definida por

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

entonces

$$\det A = \sum_{\alpha \in S_n} \epsilon(\alpha) a_{\alpha(1)1} \cdots a_{\alpha(n)n} \tag{3.13}$$

.

3.7.2 Determinante de una aplicación lineal

El determinante de una matriz definido por la fórmula (3.13) en la sección anterior debería ser tomado como la noción del determinante de una aplicación lineal (o como la expresión en unas bases dadas de dicho concepto). En realidad el concepto de determinante de una aplicación lineal surge de un modo ligeramente diferente y directamente relacionado con el ejercicio 3.7.3.

Sea $f\colon V \to W$ una aplicación lineal entre dos espacios vectoriales V y W de la misma dimensión n. Fijemos un volumen Ω_V en V y otro Ω_W en W. Definamos una n–forma $f^*\Omega_W$ en V como sigue:

$$(f^*\Omega_W)(x_1, \ldots, x_n) = \Omega_W(f(x_1), \ldots, f(x_n)), \quad \forall x_1, \ldots, x_n \in V.$$

(Nótese que si las dimensiones de los espacios vectoriales fueran diferentes $f^*\Omega_W$ no sería un volumen en V). Por el teorema 3.7.1 tenemos que los volúmenes Ω_V y $f^*\Omega_W$ son proporcionales. Llamaremos determinante de f a dicho número.

Definición 3.7.4 *Se llama determinante de f respecto de los volúmenes Ω_V y Ω_W al número $\lambda \in \mathbb{K}$ tal que*

$$f^*\Omega_W = \lambda\Omega_V$$

y se denotará por $\det(f; \Omega_V, \Omega_W)$.

Si f es un endomorfismo de V, llamaremos determinante de f al determinante respecto de cualquier volumen en V y se denotará simplemente por $\det f$.

Nótese que si $f\colon V \to V$ es lineal y Ω_V es un volumen, entonces la ecuación,

$$f^*\Omega_V = \det(f\Omega_V), \tag{3.14}$$

define el determinante de f. Si cambiamos el volumen Ω_V es lo mismo que multiplicar los dos miembros de esta ecuación por un mismo número y el factor $\det f$ no varía. Dado que $\det f$ no depende de la base escogida cuando f es un endomorfismo, ¿cuánto vale $\det f$?

Proposición 3.7.2 *Sea $f\colon V \to V$ una aplicación lineal y $\mathcal{B} = \{e_i\}_{i=1}^n$ una base de V. Si $A = (a_{ij})$ es la representación matricial de f en dicha base, entonces,*

$$\det f = \sum_{\alpha \in S_n} \epsilon(\alpha) a_{\alpha(1)1} \cdots a_{\alpha(n)n}.$$

Demostración. En efecto calculemos $\det f$ utilizando la fórmula (3.14). Calculemos en primer lugar $f^*\Omega_V$, para ello todo lo que tenemos que hacer es calcular $f^*\Omega_V(e_1, \ldots, e_n)$. Pero esto es, por definición,

$$\begin{aligned} f^*\Omega_V(e_1, \ldots, e_n) &= \Omega_V(f(e_1), \ldots, f(e_n)) \\ &= \Omega_V\left(\sum_{i_1} a_{i_1 1} e_{i_1}, \ldots, \sum_{i_n} a_{i_n n} e_{i_n}\right) = \sum_{\alpha \in S_n} \epsilon(\alpha) a_{\alpha(1)1} \cdots a_{\alpha(n)n} \Omega_V(e_1, \ldots, e_n). \end{aligned}$$

QED

Nótese que hemos obtenido que el determinante de f es simplemente el determinante de una representación matricial A de f en una base arbitraria y donde el determinante de la matriz A está dado por la fórmula (3.13).

1. En la próxima sección veremos que este hecho es evidente a partir de las propiedades de los determinantes y de las leyes de transformación de las representaciones matriciales de endomorfismos.
2. En la definición de determinante de una matriz A se podrían haber intercambiado columnas por filas y escribir dicho determinante exactamente igual que la fórmula (3.12). Veremos a continuación que tal convención es irrelevante porque el determinante de una matriz y su traspuesta coinciden. El resultado más general que no probaremos aquí es que el determinante de una aplicación lineal $f\colon V \to V$ y su dual $f^*\colon V^* \to V^*$ coinciden.

3.7.3 Determinantes de matrices y sus propiedades

Se puede desarrollar directamente la teoría de determinantes y sus aplicaciones tomando como definición la fórmula (3.13). A pesar de ello y como veremos a continuación las propiedades de los determinantes resultan transparentes (casi triviales) a partir de la fundamentación conceptual desarrollada en las secciones previas.

Definición 3.7.5 *Sea A una matriz cuadrada $n \times n$ sobre el cuerpo $\mathbb{K}$. Llamaremos determinante de A y lo denotaremos* $\det A$ *(o a veces también $|A|$) al número*

$$\det A = \sum_{\alpha \in S_n} \epsilon(\alpha) A_{\alpha(1)1} \cdots A_{\alpha(n)n}. \tag{3.15}$$

En los casos $n = 2$ y $n = 3$ es fácil recordar las expresiones de los determinantes, así:

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21},$$

y

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

Es claro a partir de la definición que el desarrollo del determinante de una matriz $n \times n$ tiene $n!$ términos. Por tanto la complejidad de su cálculo aumenta extraordinariamente con el orden de A.

Propiedades de los determinantes

Proposición 3.7.3 *La función $A \mapsto \det A$ definida en el conjunto $M_n(\mathbb{K})$ tiene las propiedades siguientes:*

i. $\det A$ *es una función multilineal de las columnas de A.*
ii. $\det A$ *es una función antisimétrica de las columnas de A.*
iii. $\det I_n = 1$.

Demostración. Si utilizamos los resultados de la sección anterior, en el sentido de que el determinante de A no es más que el volumen de los vectores columnas de A entonces las propiedades (i) y (ii) son triviales. La propiedad (iii) es inmediata a partir de la definición del determinante de una aplicación lineal (3.14) ya que el determinante de la aplicación identidad es 1.

A pesar de ello las propiedades anteriores se pueden probar directamente.

i. Supongamos que $A = (A^1 \dots A^i + B^i \dots A^n)$, entonces:

$$\begin{aligned} |\, A \,| &= \det(A^1 \dots A^i + B^i \dots A^n) = \sum_{\alpha \in S_n} \epsilon(\alpha) A_{\alpha(1)1} \cdots (A_{\alpha(i)i} + B_{\alpha(i)i}) \cdots A_{\alpha(n)n} \\ &= \sum_{\alpha \in S_n} \epsilon(\alpha) A_{\alpha(1)1} \cdots A_{\alpha(i)i} \cdots A_{\alpha(n)n} + \sum_{\alpha \in S_n} \epsilon(\alpha) A_{\alpha(1)1} \cdots B_{\alpha(i)i} \cdots A_{\alpha(n)n} \\ &= \det(A^1 \dots A^i \dots A^n) + \det(A^1 \dots B^i \dots A^n). \end{aligned}$$

ii. Denotemos por $\tilde{A} = (A^1 \dots A^j \dots A^i \dots A^n)$ la matriz que se obtiene de la A intercambiando la columna i con la j. Entonces,

$$\begin{aligned} \det \tilde{A} &= \sum_{\alpha \in S_n} \epsilon(\alpha) A_{\alpha(1)1} \cdots A_{\alpha(i)j} \cdots A_{\alpha(j)i} \cdots A_{\alpha(n)n} \\ &= \sum_{\alpha \in S_n} \epsilon(\alpha) A_{\alpha\circ\tau_{ij}(1)1} \cdots A_{\alpha\circ\tau_{ij}(j)j} \cdots A_{\alpha\circ\tau_{ij}(i)i} \cdots A_{\alpha\circ\tau_{ij}(n)n} \\ &= \sum_{\beta \in S_n} \epsilon(\beta \circ \tau_{ij}) A_{\beta(1)1} \cdots A_{\beta(j)j} \cdots A_{\beta(i)i} \cdots A_{\beta(n)n} \\ &= -\sum_{\beta \in S_n} \epsilon(\beta) A_{\beta(1)1} \cdots A_{\beta(j)j} \cdots A_{\beta(i)i} \cdots A_{\beta(n)n} = -\det A. \end{aligned}$$

iii. Es evidente a partir de la definición 3.15. QED

El siguiente teorema es una reformulación en coordenadas de nuestro teorema fundamental 3.7.1.

Teorema 3.7.2 *Sea $D: M_n(\mathbb{K}) \to \mathbb{K}$ una función con las siguientes propiedades:*
i. D es una función lineal de las columnas de $A \in M_n(\mathbb{K})$.
ii. D es una función antisimétrica en las columnas de $A \in M_n(\mathbb{K})$.
Entonces existe una constante λ tal que

$$D(A) = \lambda \det A.$$

Demostración. Es evidente que la función D define una n–forma en el espacio vectorial $\mathbb{K}^n$. Por lo tanto por el teorema 3.7.1 D es proporcional a la n-forma definida por $\det A$. QED

Proposición 3.7.4 $\det A = \det A^t$.

Demostración. Se puede probar directamente de: $\det A^t = \sum_{\alpha \in S_n} \epsilon(\alpha) A_{\alpha(1)1} \cdots A_{\alpha(n)n}$ y como $A_{\alpha(1)1} \cdots A_{\alpha(n)n} = A_{1\alpha^{-1}(1)} \cdots A_{n\alpha^{-1}(n)}$ al ser $\alpha \mapsto \alpha^{-1}$ una aplicación biyectiva de S_n el sumatorio sobre α es igual al sumatorio sobre α^{-1}. En cualquier caso podemos probarlo utilizando la caracterización dada por el teorema 3.7.2.

Definamos una aplicación $D(A) = \det A^t$. Probemos que D es lineal en las columnas. Esto es, si $A' = (A^1, \ldots, A^i + B^i, \ldots A^n)$, entonces

$$\begin{aligned} \det A' &= \sum_{\alpha \in S_n} \epsilon(\alpha) A_{\alpha(1)1} \cdots (A_{\alpha(i)i} + B_{\alpha(i)i}) \cdots A_{\alpha(n)n} \\ &= \sum_{\alpha \in S_n} \epsilon(\alpha) A_{\alpha(1)1} \cdots A_{\alpha(i)i} \cdots A_{\alpha(n)n} + \sum_{\alpha \in S_n} \epsilon(\alpha) A_{\alpha(1)1} \cdots B_{\alpha(i)i} \cdots A_{\alpha(n)n} \\ &= D(A^1 \ldots A^i \ldots A^n) + D(A^1 \ldots B^i \ldots A^n). \end{aligned}$$

Del mismo modo que probamos que el determinante es antisimétrico en las columnas de A, proposición 3.7.3 (ii), probamos que D es antisimétrica en las columnas de A.

Finalmente es inmediato probar que $D(I_n) = 1$, y por tanto el coeficiente $\lambda = 1$. Según el teorema anterior 3.7.2, $D(A) = \det A$ y por tanto $\det A^t = \det A$. QED

Proposición 3.7.5 *Desarrollo de un determinante por filas y por columnas. Si $A = (a_{ij})$ es una matriz $n \times n$, entonces:*

$$\det A = \sum_{j=1}^{n} (-1)^{i+j} a_{ij} M_{ij}(A) = \sum_{i=1}^{n} (-1)^{i+j} a_{ij} M_{ij}(A), \tag{3.16}$$

donde $M_{ij}(A)$ es el determinante de la matriz que se obtiene al quitar la fila i y la columna j de la matriz A. $M_{ij}(A)$ se denomina el menor (i,j) de la matriz A o también el complemento del elemento a_{ij}.

Demostración. La fórmula anterior se sigue del desarrollo:

$$\begin{aligned} \det A &= \sum_{\alpha \in S_n} \epsilon(\alpha) a_{1\alpha(1)} \cdots a_{n\alpha(n)} = \sum_{\substack{\alpha \in S_n \\ \alpha(i) = 1}} \epsilon(\alpha) a_{i1} (a_{1\alpha(1)} \cdots \hat{a}_{i\alpha(i)} \cdots a_{n\alpha(n)}) \\ &+ \sum_{\substack{\alpha \in S_n \\ \alpha(i) = 2}} \epsilon(\alpha) a_{i2} (a_{1\alpha(1)} \cdots \hat{a}_{i\alpha(i)} \cdots a_{n\alpha(n)}) + \cdots \\ &+ \sum_{\substack{\alpha \in S_n \\ \alpha(i) = n}} \epsilon(\alpha) a_{in} (a_{1\alpha(1)} \cdots \hat{a}_{i\alpha(i)} \cdots a_{n\alpha(n)}) \end{aligned}$$

QED

Producto de determinantes e inversa de una matriz

Teorema 3.7.3 *Dadas dos matrices $n \times n$ A y B con coeficientes en $\mathbb{K}$, se verifica:*

$$\det(AB) = \det A \det B. \tag{3.17}$$

Demostración. Podemos demostrarlo fácilmente utilizando la definición 3.14 del determinante. En efecto, si f es un endomorfismo representado por A y g es un endomorfismo representado por B (en la misma base), entonces

$$\det BA\Omega = (g \circ f)^*\Omega = f^* \circ g^*\Omega = f^*(g^*\Omega) = f^*(\det B\Omega) = \det B f^*\Omega = \det B \det A\Omega.$$

Se puede probar también directamente utilizando la definición 3.13 y manipulando las sumas adecuadamente. Otra demostración se puede realizar aplicando el Teorema 3.7.2 a la aplicación $D(A) = \det(AB)$ con B fija. QED

Ejercicio 3.7.6 Probar la fórmula (3.17) utilizando la definición 3.13 del determinante de una matriz.

Teorema 3.7.4 *La matriz A es invertible si y sólo si $\det A \neq 0$. Además $\det(A^{-1}) = (\det A)^{-1}$.*

Demostración. Si A es invertible, entonces existe una matriz A^{-1} tal que $AA^{-1} = I_n$. Por tanto, $1 = \det I_n = \det(AA^{-1}) = \det A \det A^{-1}$ y así $\det(A^{-1}) = (\det A)^{-1}$.

Recíprocamente si $\det A \neq 0$, construimos la matriz cuyo elemento (i, j) es:

$$B_{ij} = \frac{1}{\det A}(-1)^{i+j} M_{ji}(A).$$

Calculemos AB. Tenemos,

$$(AB)_{ij} = \sum_{k=1}^{n} A_{ik}B_{kj} = \sum_{k=1} \frac{(-1)^{k+j}}{\det A} A_{ik}M_{jk}(A).$$

Si $i = j$, utilizando el desarrollo del determinante por columnas (3.16), obtenemos,

$$(AB)_{ii} = \frac{1}{\det A}\sum_{k=1}^{n}(-1)^{i+k}A_{ik}M_{ik}(A) = \frac{\det A}{\det A} = 1.$$

Si $i \neq j$, entonces $(AB)_{ij} = 0$ ya que es el desarrollo por columnas del determinante de la matriz A con la columna j reemplazada por la i, esto es con dos columnas iguales, y por tanto 0. QED

La demostración anterior nos ha proporcionado de paso una expresión explícita de la inversa de una matriz:

$$(A^{-1})_{ij} = \frac{1}{\det A}(-1)^{i+j} M_{ji}(A). \tag{3.18}$$

Proposición 3.7.6 *Cálculo del rango de una matriz. El rango de una matriz A es igual al máximo de los órdenes de sus menores no nulos, donde llamamos menor de una matriz al determinante de una submatriz cuadrada y orden del menor al orden de la submatriz.*

Ejercicio 3.7.7 Probar la proposición anterior.

Proposición 3.7.7 *Regla de Cramer. Dado el sistema lineal de ecuaciones $AX = B$, si A es invertible su solución es única y es $X = A^{-1}B$. Explícitamente:*

$$x_i = \frac{\begin{vmatrix} a_{11} & \cdots & b_1 & \cdots & a_{1n} \\ a_{21} & \cdots & b_2 & \cdots & a_{2n} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ a_{n1} & \cdots & b_n & \cdots & a_{nn} \end{vmatrix}}{\begin{vmatrix} a_{11} & \cdots & a_{1n} \\ a_{21} & \cdots & a_{2n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{vmatrix}},$$

donde el vector B se halla en la columna i del determinante del numerador.

Ejercicio 3.7.8 Probar la proposición anterior 3.7.7.

Tema 4

Formas canónicas de endomorfismos

Diagonalización. Autovectores y autovalores. Subespacios invariantes. Ecuación característica. Endomorfismos nilpotentes. Formas canónicas de endomorfismos. Teorema de Cayley-Hamilton. Polinomio mínimo

4.1 Diagonalización

Sea V un $\mathbb{K}$-espacio vectorial de dimensión finita, $\mathbb{K} = \mathbb{R}, \mathbb{C}$. Un endomorfismo de V, $f \in \mathrm{End}(V)$ es una aplicación lineal de V en V.

Sabemos que dada un base de V, $\mathcal{B} = \{u_1, \ldots, u_n\}$, se asocia a un endomorfismo f una matriz A, construida de la forma siguiente (ver 3.3.1):

$$f(u_i) = \sum_{j=1}^{n} a_{ji} u_j, \quad A = (a_{ij}) = A(f, \mathcal{B})$$

Las columnas de A son las coordenadas de los vectores imágenes de la base $\mathcal{B}$, expresadas en esta base.

Si cambiamos de base:

$$\mathcal{B} = \{u_i\} \longrightarrow \mathcal{B}' = \{u'_i\}$$

con matriz cambio de base P:

$$u_i = \sum_{j=1}^{n} P_{ji} u'_j, \quad P = (P_{ij}), \quad \det P \neq 0,$$

la matriz A cambia a una matriz $A' = M(f, \mathcal{B}')$ dada por:

$$A' = PAP^{-1}. \tag{4.1}$$

La fórmula anterior corresponde a la fórmula (3.9) cuando la aplicamos al mismo cambio de bases tanto en el dominio como en el rango de f.

Nótese que si ambas bases están referidas a una tercera (como por ejemplo, en el caso de $\mathbb{R}^n$, con las bases $\mathcal{B}$ y $\mathcal{B}'$ escritas en la base canónica):

$$u_i = \sum_{j=1}^{n} q_{ji} e_j, \quad u'_i = \sum_{j=1}^{n} q'_{ji} e_j, \; j = 1, \ldots, n$$

se tienen dos matrices:

$$Q = \begin{pmatrix} q_{11} & \cdots & q_{1n} \\ \vdots & & \vdots \\ q_{n1} & \cdots & q_{nn} \end{pmatrix}, \quad Q' = \begin{pmatrix} q'_{11} & \cdots & q'_{1n} \\ \vdots & & \vdots \\ q'_{n1} & \cdots & q'_{nn} \end{pmatrix},$$

en las que las columnas son las coordenadas de los vectores de las bases $\mathcal{B}$ y $\mathcal{B}'$ en la tercera base $\{e_i\}$. La matriz de cambio de base es ahora muy sencilla: el cambio de base de $\mathcal{B}$ a $\{e_i\}$ viene dado por Q y el de $\{e_i\}$ a $\mathcal{B}'$ por: $(Q')^{-1}$, luego el de $\mathcal{B}$ a $\mathcal{B}'$ es $P = (Q')^{-1}Q$

De acuerdo con estas ideas, el objetivo de este tema es encontrar una base $\mathcal{B}$ en la cual la matriz $A(f, \mathcal{B})$ sea lo más sencilla posible.

Es muy fácil persuadirse que si utilizamos cambios de bases diferentes en el dominio y en el rango de la aplicación f es posible hallar una expresión para f de la forma

$$\left(\begin{array}{c|c} I_r & 0 \\ \hline 0 & 0 \end{array} \right), \tag{4.2}$$

donde r es el rango de f. En efecto, basta tomar una base $\{u_1, \ldots, u_s\}$ del núcleo de f y extenderla a todo V, así $\mathcal{B} = \{v_1, \ldots, v_r, u_1, \ldots, u_s\}$, con $r = r(f)$ y $r + s = \dim V$. Tomemos los vectores $f(v_1), \ldots, f(v_r)$ que forman una base de $\operatorname{im} f$ y completemos dicha base hasta obtener una base $\mathcal{B}'$ de V. Es obvio que la matriz asociada a f en estas dos bases es la descrita en (4.2).

Por el contrario si estamos describiendo nuestro endomorfismo f desde una base dada, nos interesará averiguar como cambia la forma de sus matrices asociadas bajo las transformaciones (4.1).

Diremos que dos matrices A y A' son equivalentes o conjugadas si existe una matriz invertible P tal que $A' = PAP^{-1}$. Dicha relación es de equivalencia. Desde un punto de vista asociado a la teoría de grupos la relación anterior corresponde a la conjugación por el grupo general lineal $GL(n, \mathbb{K})$ en el conjunto de matrices $M_n(\mathbb{K})$. El problema que estamos planteando consiste en buscar un elemento lo más sencillo posible en la órbita de A bajo la acción por conjugación del grupo general lineal.

El problema de determinar formas canónicas de endomorfismos consiste en describir el espacio cociente, esto es las clases de equivalencia, del conjunto de matrices con respecto a la relación de equivalencia anterior.

4.1.1 Matrices diagonales

Definición 4.1.1 *Una matriz diagonal $A \in M_n(\mathbb{K})$ tiene todos sus elementos cero salvo los de la diagonal:*

$$A = \begin{pmatrix} a_{11} & & \\ & \ddots & \\ & & a_{nn} \end{pmatrix}.$$

Una matriz diagonal A queda definida por las fórmulas:

$$A_{ij} = \left\{ \begin{array}{ccc} 0 & \text{si} & i \neq j \\ a_{ii} & \text{si} & i = j \end{array} \right. ,$$

o equivalentemente $A_{ij} = a_{ii}\delta_{ij}$, y la representaremos habitualmente como $A = \operatorname{diag}(a_{11}, \ldots, a_{nn})$.

Aceptaremos que ésta es la forma más sencilla de escribir un endomorfismo en una base adecuada. Sin embargo no siempre es posible encontrar un base en la cual el endomorfismo en cuestión venga representado por una matriz diagonal. En este caso, nos veremos obligados a contentarnos con una representación también sencilla (forma canónica de Jordan) pero no diagonal.

Definición 4.1.2 *Diremos que un endomorfismo f es diagonalizable si existe una base del espacio vectorial tal que la matriz asociada a f en dicha base es diagonal.*

De manera análoga podemos decir que una matriz es diagonalizable si existe una matriz equivalente a ella que es diagonal.

Ejemplo 4.1.1 Consideremos la matriz $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$. Estudiemos como cambia bajo conjugación. Si $P = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ es una matriz invertible $ad - bc \neq 0$, entonces $A' = P^{-1}AP$ es,

$$A' = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \frac{1}{ad - bc} \begin{pmatrix} ad + dc - bc & d^2 \\ -c^2 & ad - bc - cd \end{pmatrix}.$$

Por tanto A será diagonalizable solamente si $c = d = 0$ lo cual es imposible ya que P ha de ser invertible.

Ejercicio 4.1.1 Probar que el operador D en el espacio de polinomios no es diagonalizable. Probar que cualquier operador diferencial en el espacio de polinomios no es diagonalizable.

Ejercicio 4.1.2 Probar que si f es diagonalizable cualquier potencia de f también lo es.

4.2 Autovalores y autovectores

En la descripción de un endomorfismo juegan un papel crucial los vectores cuya imagen por f es proporcional a ellos mismos.

Definición 4.2.1 *Sea $f \in \mathrm{End}(V)$. Se dice que $\lambda \in \mathbb{K}$ es un autovalor (valor propio) de f si existe un vector $v \in V$, $v \neq 0$ tal que:*

$$f(v) = \lambda v. \tag{4.3}$$

En este caso, se dice que v es un autovector (vector propio) de f con autovalor λ.

Es evidente que cuantos más autovectores encontremos para un endomorfismo, más fácil resultará describirlo. Así para el endomorfismo identidad todos los vectores son autovectores con autovalor 1.

Definición 4.2.2 *El espectro de f es el conjunto de sus autovalores:*

$$\sigma(f) = \{\lambda \in \mathbb{K} \mid \lambda \ \text{autovalor de} \ f\}.$$

Nótese que dado un autovector, existe un único autovalor asociado a él (obviamente), pero a cada autovalor puede corresponderle más de un autovector.

Ejemplo 4.2.1 Considérese el endomorfismo f definido en un espacio vectorial V de dimensión 3 a través de la asignación:

$$f(v_1) = v_1 + v_2 + v_3, \quad f(v_2) = v_2 + v_3, \quad f(v_3) = v_3,$$

donde v_1, v_2, v_3 forman una base de V. Si resolvemos la ecuación $f(u) = \lambda u$, encontramos que necesariamente $\lambda = 1$ y $u = v_3$. Por tanto $\sigma(f) = \{1\}$.

Ejercicio 4.2.1 Probar que si $f^r = 0$ para algún $r > 0$, entonces $\sigma(f) = \{0\}$.

Ejercicio 4.2.2 Probar que si $f^2 = f$ y $f \neq 0$, entonces $\sigma(f) = \{1, -1\}$.

Dado un endomorfismo f diremos que un subespacio W es invariante si $f(W) \subset W$.

Proposición 4.2.1 *Sea $f \in \mathrm{End}(V)$. Para cada autovalor λ, se define el conjunto:*

$$V_\lambda = \{v \in V \mid f(v) = \lambda v\}.$$

Se tiene que V_λ es un subespacio de V invariante bajo f. El espacio V_λ se puede definir como:

$$V_\lambda = \ker(f - \lambda 1_V),$$

donde 1_V es la aplicación identidad en V.

Demostración. La demostración es evidente. La combinación lineal de autovectores correspondientes a un mismo autovalor es un autovector de ese autovalor. Y por supuesto, $f(V_\lambda) \subset V_\lambda$ ya que si $v \in V_\lambda$, $f(f(v)) = f(\lambda v) = \lambda f(v)$. QED

La proposición anterior (4.2.1), nos dice que los espacios de autovectores son invariantes. No todo subespacio invariante de V es de este tipo.

El siguiente resultado establece la relación que existe entre los subespacios invariantes V_λ.

Proposición 4.2.2 *Si $\lambda_1, \dots, \lambda_r$ son autovalores distintos de $f \in \operatorname{End}(V)$, la suma de los subespacios V_{λ_i}, $i = 1, \dots, r$ es directa.*

Demostración. Basta probar que si un vector pertenece a la suma de estos subespacios, se puede escribir de forma única como suma de vectores cada uno en un subespacio. Sea $v \in V_{\lambda_1} + \cdots + V_{\lambda_r}$, $v = v_1 + \cdots + v_r$, con $v_i \in V_{\lambda_i}$. Probar la unicidad de la descomposición es equivalente a probar que si $v = 0$, cada uno de los vectores v_i es cero. Lo hacemos por inducción en r. Sea $r = 1$. El resultado es inmediato. Supongamos que es cierto para $r - 1$. Tenemos, (para r):

$$v_1 + \cdots + v_r = 0$$

Aplicando f a los dos miembros de esta igualdad:

$$f(v_1) + \cdots + f(v_r) = 0,$$

y como v_i es un autovector de autovalor λ_i:

$$\lambda_1 v_1 + \cdots + \lambda_r v_r = 0,$$

de donde, restando la ecuación anterior multiplicada por λ_r:

$$(\lambda_1 - \lambda_r)v_1 + \cdots + (\lambda_{r-1} - \lambda_r)v_{r-1} = 0.$$

Pero ahora estamos en las condiciones del caso $r - 1$, por lo tanto:

$$(\lambda_i - \lambda_r)v_i = 0, \quad i = 1, \dots, r - 1.$$

Al ser todos los autovalores λ_i distintos, los vectores v_i son cero ($i = 1, \dots, r$), lo que implica que $v_r = 0$. Luego la suma es directa (lo que lleva, en particular, a que las intersecciones entre estos subespacios se reduzcan a $\{0\}$). QED

Nótese que el subespacio correspondiente al autovalor 0 es el núcleo de f. Por tanto un endomorfismo f será invertible si y sólo si el cero no está en su espectro.

4.3 Subespacios invariantes y matrices

Si se tiene un subespacio invariante de un endomorfismo (del tipo V_λ o cualquier otro), en bases adaptadas a este subespacio las matrices que representan al endomorfismo tienen una forma especial.

Proposición 4.3.1 *Sea $f \in \operatorname{End}(V)$ y $W \subset V$ un subespacio de V invariante bajo f. Entonces, existe un base de V, $\mathcal{B}$, en la que la matriz de f tiene la forma:*

$$A(f, \mathcal{B}) = \left(\begin{array}{c|c} A & B \\ \hline 0 & C \end{array}\right).$$

Demostración. Sea $\mathcal{B}_W$ una base de W que se amplía a una base de V. En esta base, la matriz es la dada en la proposición, porque las imágenes de los vectores de la base $\mathcal{B}_W$ están contenidas en el subespacio W que es invariante, por tanto sus coordenadas sobre el resto de la base $\mathcal{B}$ son cero. QED

Cuando se dispone de un subespacio invariante bajo f, es posible definir una aplicación lineal obtenida a partir de f del espacio cociente V/W en sí mismo:

$$\begin{array}{rccc} \tilde{f}\colon & V/W & \longrightarrow & V/W \\ & v + W & \longmapsto & f(v) + W \end{array}$$

La aplicación $\tilde{f}$ está bien definida y es lineal (debido al carácter de subespacio invariante de W). Sea $\mathcal{B}_W = \{w_1, \dots, w_r\}$ una base de W y $\mathcal{B} = \{w_1, \dots, w_r, u_1, \dots, u_s\}$ la base ampliada de V. Como hemos visto en temas anteriores, una base del espacio cociente V/W es: $\mathcal{B}_{V/W} = \{u_i + W \mid i = 1, \dots, s\}$. En la base $\mathcal{B}$, la matriz de la aplicación f es la dada por el teorema, es decir:

$$
\begin{aligned}
f(w_i) &= \sum_{j=1}^{r} a_{ji} w_j, \quad i = 1, \dots, r \\
f(u_k) &= \sum_{j=1}^{r} b_{jk} w_j + \sum_{j=1}^{s} c_{jk} u_j, \quad k = 1, \dots, s
\end{aligned}
$$

con lo que la matriz de la aplicación $\tilde{f}$ en la base $\mathcal{B}_{V/W}$ es:

$$
\tilde{f}(u_k + W) = f(u_k) + W = \sum_{j=1}^{r} b_{jk} w_j + \sum_{j=1}^{s} c_{jk} u_j + W = \sum_{j=1}^{s} c_{jk} u_j + W = \sum_{j=1}^{s} c_{jk} (u_j + W)
$$

Por lo tanto, la matriz de $\tilde{f}$ en la base $\mathcal{B}_{V/W}$ es igual a C.

Si el subespacio W en el teorema anterior tiene dimension r, entonces $A \in M_r(\mathbb{K})$, $C \in M_{n-r}(\mathbb{K})$ y $B \in M_{r\times(n-r)}(\mathbb{K})$. Si W es un subespacio invariante puede ocurrir que exista un suplementario U que también sea invariante. En tal caso la proposición anterior nos dice que existe una base tal que la matriz asociada a f tiene la forma:

$$
\left(\begin{array}{c|c} A & 0 \\ \hline 0 & C \end{array} \right).
$$

Todavía mas. Si V se puede descomponer como una suma directa de subespacios invariantes W_i, $i = 1, \dots, N$, $V = W_1 \oplus \cdots \oplus W_N$, $f(W_i) \subset W_i$, entonces podemos construir una base tal que la matriz asociada a f tiene la forma:

$$
A = \begin{pmatrix} A_1 & & & \\ & A_2 & & \\ & & \ddots & \\ & & & A_N \end{pmatrix},
$$

y el orden de la matriz A_i es la dimensión de W_i. Tales matrices se dirá que son diagonales por cajas.

4.3.1 Diagonalización de endomorfismos y matrices

El siguiente teorema establece una condición necesaria y suficiente para la existencia de una base en la que el endomorfismo f viene representado por una matriz diagonal, es decir, una condición para que f sea diagonalizable.

Teorema 4.3.1 *Sea $f \in \mathrm{End}(V)$. f es diagonalizable si y solo si existe una base de V formada por autovectores de f.*

Demostración. Si existe una base de autovectores: $\mathcal{B} = \{u_1, \dots, u_n\}$, sus imágenes mediante f son $f(u_i) = \lambda_i u_i$, $i = 1, \dots, n$, por lo que la matriz asociada es:

$$
A(f, \mathcal{B}) = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}.
$$

Y en sentido contrario es igualmente sencillo. Si la matriz asociada es diagonal, los elementos de la diagonal son justamente los autovalores, y los vectores de la base los vectores correspondientes. QED

4.4 La ecuación característica

4.4.1 Cálculo de autovalores y autovectores

Sea $f \in \mathrm{End}(V)$, y $\mathcal{B}$ una base de V. Sea A la matriz asociada a f en la base $\mathcal{B}$. Los autovalores y autovectores de f se pueden calcular en la base $\mathcal{B}$ en la forma siguiente:

$$f(v) = \lambda v,$$

implica que

$$A\hat{v} = \lambda \hat{v},$$

donde $\hat{v}$ es el vector de $\mathbb{K}^n$ que representa a $v \in V$ en la base $\mathcal{B}$. Resolver esta segunda ecuación es equivalente a resolver el sistema homogéneo de ecuaciones:

$$(A - \lambda I)\hat{v} = 0. \tag{4.4}$$

Este sistema poseerá soluciones no triviales si y solo si

$$\det(A - \lambda I) = 0, \tag{4.5}$$

tal y como mostramos en el capítulo anterior.

La ecuación anterior, (4.5), se denomina la ecuación de autovalores o ecuación característica, y nos permite encontrar los autovalores de un endomorfismo como raíces del polinomio $\det(A - \lambda I)$. Para cada solución de esta ecuación, se calcula el (o los) autovector correspondiente usando de nuevo la ecuación (4.4).

4.4.2 El polinomio característico de un endomorfismo

Sea $A \in M_n(\mathbb{K})$. Se define el polinomio característico de A como:

$$p_A(\lambda) = \det(A - \lambda I) \tag{4.6}$$

Es un polinomio de grado n y el coeficiente del término de mayor grado es $(-1)^n$.

Sea $f \in \mathrm{End}(V)$. Se define el polinomio característico de f como el polinomio característico de la matriz de f en cualquier base y lo denotaremos por p_f. En efecto, es muy sencillo demostrar que el polinomio no depende de la base ya que:

$$\det(A' - \lambda I) = \det(PAP^{-1} - \lambda I) = \det(P(A - \lambda I)P^{-1}) = \det(A - \lambda I)$$

donde A' es la matriz de f en otra base y P es la matriz de cambio de base.

De acuerdo con la ecuación de autovalores se tiene:

Proposición 4.4.1 $\lambda \in \mathbb{K}$ *es autovalor de f si y sólo si λ es raíz del polinomio característico, es decir,* $p_f(\lambda) = 0$.

Ejercicio 4.4.1 Probar que si el polinomio característico no posee término independiente el endomorfismo no es invertible.

Ejemplo 4.4.1 Notemos que si f es un endomorfismo de un espacio vectorial V complejo tal que en una cierta base su matriz asociada A tiene coeficientes reales, $A \in M_n(\mathbb{R})$, entonces si λ es un autovalor, también lo será $\bar{\lambda}$.

Dada una raíz del polinomio característico, existen dos números asociados a ella: uno es la multiplicidad algebraica como raíz de ese polinomio. El otro es la dimensión del espacio invariante V_λ. A este último lo llamaremos multiplicidad geométrica. Es decir, la multiplicidad geométrica de una raíz del polinomio característico es la dimensión de $\ker(f - \lambda 1_V)$. En general estos dos números son distintos. Pero se tiene:

Proposición 4.4.2 *Sea $f \in \text{End}(V)$, $\lambda \in \mathbb{K}$ raíz del polinomio característico de f, $p_f(\lambda)$. Entonces, la multiplicidad algebraica de λ es mayor o igual que la multiplicidad geométrica de λ.*

Demostración. Sea $\lambda_0 \in \mathbb{K}$ una raíz del polinomio característico, $p_f(\lambda_0) = 0$ y sea V_{λ_0} el subespacio invariante asociado a λ_0. Construimos una base de V_{λ_0} y la ampliamos a una base de V. La matriz de f en esta base es, como ya sabemos, Prop. (4.3.1):

$$\left(\begin{array}{c|c} A & B \\ \hline 0 & C \end{array}\right).$$

Es fácil probar que el polinomio característico de f es el producto de los polinomios característicos de A y C, debido a las propiedades de los determinantes:

$$p_f(\lambda) = \det(A - \lambda I)\det(C - \lambda I)$$

Pero A es una matriz diagonal (porque la base de V_{λ_0} está formada por autovectores de f), y su polinomio característico es: $(\lambda_0 - \lambda)^s$, donde $s = \dim V_{\lambda_0}$, que es la multiplicidad geométrica de λ_0:

$$p_f(\lambda) = (\lambda_0 - \lambda)^s \det(C - \lambda I)$$

Por tanto, la multiplicidad algebraica de λ_0 es mayor o igual que la geométrica ($= s$). QED

Consecuencia de estos resultados es el siguiente teorema, que da un criterio suficiente para la diagonalización de un endomorfismo (o una matriz):

Teorema 4.4.1 *Si $f \in \text{End}(V)$, $\dim V = n$, tiene polinomio característico con n raíces distintas, entonces f es diagonalizable.*

Demostración. El espectro de f es: $\sigma(f) = \{\lambda_1, \dots \lambda_n\}$, con todos los autovalores λ_i distintos. Los autovectores correspondientes son l.i., pues están en subespacios invariantes distintos, luego forman una base de V, y por tanto f es diagonalizable. En este caso:

$$V = V_{\lambda_1} \oplus \cdots \oplus V_{\lambda_n}.$$

Las multiplicidades algebraica y geométrica coinciden para cada autovalor y son iguales a 1. QED

La condición anterior no es una condición necesaria para la diagonalización. En efecto, la matriz $\text{diag}(\lambda, \dots, \lambda)$ es diagonal y todos sus autovalores coinciden.

4.5 Formas canónicas de endomorfismos nilpotentes

Como paso previo al estudio de las formas canónicas de endomorfismos de un espacio vectorial V de dimensión finita sobre $\mathbb{C}$, estudiaremos en primer lugar las de los endomorfismos nilpotentes. La razón está en que el estudio de estos endomorfismos se puede hacer sobre los subespacios invariantes de V, asociados a un autovalor (aunque no estén formados únicamente por autovectores), y en ellos, los endomorfismos $(f - \lambda 1_V)$ son nilpotentes. El limitarse a $\mathbb{C}$ viene dado por la propiedad de ser un cuerpo algebraicamente cerrado (propiedad que no tiene $\mathbb{R}$, recordad 1.5.3). Esta propiedad hace que la suma de las multiplicidades algebraicas de las raíces del polinomio característico sea igual al grado de este polinomio es decir a la dimensión del espacio V, lo que será decisivo en la construcción de las formas canónicas que nos proponemos estudiar.

Definición 4.5.1 *Sea $f \in \text{End}(V)$. Se dice que f es nilpotente de grado r, si $f^r = 0$ y $f^{r-1} \neq 0$.*

Los autovalores de un operador nilpotente son todos iguales a cero:

$$f(v) = \lambda v \Rightarrow 0 = f^r(v)v = \lambda^r v \Rightarrow \lambda = 0$$

por lo que un operador nilpotente no es diagonalizable, a no ser que sea igual a 0. Sin embargo, para cada operador nilpotente existe una base en la cual éste adopta una forma particularmente sencilla. Antes de discutir la situación general estudiemos brevemente que ocurre con un endomorfismo nilpotente de grado 2, esto es, $f^2 = 0$.

Ejemplo 4.5.1 Si $f^2 = 0$, resulta evidente que $\operatorname{im} f \subset \ker f$. En efecto, si $v = f(u)$, entonces $f(v) = f^2(u) = 0$. Supongamos que el rango de f es r. Entonces, $\dim\ker f = n-r$, donde $n = \dim V$, y $r \leq n-r$. Sea $\{u_1, \ldots, u_r\}$ una base de $\operatorname{im} f$. Ampliemos esta base hasta obtener una base de $\ker f$, esto es añadimos los vectores $u_{r+1}, \ldots, u_{n-r}$. Tomemos vectores anti-imágenes de los $u_1, \ldots, u_r$ que denotaremos por v_i, esto es $f(v_1) = u_1, \ldots, f(v_r) = u_r$. El conjunto $\{u_1, \ldots, u_r, u_{r+1}, \ldots, u_{n-r}, v_1, \ldots, v_r\}$ es una base de V. En dicha base la matriz A que representa a f tiene la forma:

$$A = \left(\begin{array}{c|c|c} 0 & 0 & I_r \\ \hline 0 & 0 & 0 \\ \hline 0 & 0 & 0 \end{array}\right).$$

Es posible reordenar los vectores de la base como $\{u_1, v_1, \ldots, u_r, v_r, u_{r+1}, \ldots, u_{n-r}\}$ y entonces la expresión de la matriz asociada es:

$$A = \begin{pmatrix} 0 & 1 & & & & & & \\ 0 & 0 & & & & & & \\ & & \ddots & & & & & \\ & & & 0 & 1 & & & \\ & & & 0 & 0 & & & \\ & & & & & 0 & & \\ & & & & & & \ddots & \\ & & & & & & & 0 \end{pmatrix}.$$

Un endomorfismo f tal que $f^2 = 0$ e $\operatorname{im} f = \ker f$, se dice que es un endomorfismo vertical.

Teorema 4.5.1 *Todo operador nilpotente $f \in \operatorname{End}(V)$ induce una descomposición del espacio V en subespacios invariantes. En cada uno de ellos se puede encontrar una base en la que el endomorfismo restringido a ese subespacio tiene como matriz la siguiente:*

$$\begin{pmatrix} 0 & 1 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 1 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & 0 & \cdots & 1 \\ 0 & 0 & 0 & 0 & \cdots & 0 \end{pmatrix}$$

Demostración. Sea $f \in \operatorname{End}(V)$, $f^r = 0$. Construimos los espacios imágenes de los operadores f^k, $k = 0, 1, \ldots, r$:

$$U_k = \operatorname{im} f^k = f^k(V)$$

es decir:

$$U_0 = V,\ U_1 = f(V), \ldots, U_{r-1} = f^{r-1}(V),\ U_r = \{0\},$$

que están contenidos unos en otros formando un cadena de subespacios:

$$\{0\} = U_r \subset U_{r-1} \subset \cdots \subset U_1 \subset U_0 = V.$$

Nótese que $f(U_{i-1}) = U_i$ y que, por tanto, $f(U_{r-1}) = \{0\}$, es decir:

$$U_{r-1} \subset \ker f.$$

Sin embargo, no tienen porque ser iguales.

Construimos una base del subespacio más pequeño no trivial, U_{r-1} y la ampliamos a una base del subespacio siguiente y así sucesivamente. Sea $d_{r-1} = \dim U_{r-1}$ y una base de U_{r-1}:

$$\{u_1^{(r-1)}, \ldots, u_{d_{r-1}}^{(r-1)}\}$$

Todos estos vectores son anulados por f:

$$f(u_i^{(r-1)}) = 0, i = 1, \ldots, d_{r-1}$$

Al ser el subespacio U_{r-1} la imagen mediante f del subespacio U_{r-2}, para cada vector de esta base de U_{r-1} existe un original (o varios) en U_{r-2}, es decir, existen vectores:

$$u_1^{(r-2)}, \ldots, u_{d_{r-1}}^{(r-2)} \in U_{r-2}$$

tales que:

$$f(u_i^{(r-2)}) = u_i^{(r-1)},\ i = 1, \ldots, d_{r-1}$$

Podemos demostrar que todos estos vectores están en U_{r-2} (pues $U_{r-1} \subset U_{r-2}$) y que son linealmente independientes. Para ello construyamos una combinación lineal e igualémosla a cero.

$$\sum_{i=1}^{d_{r-1}} (\alpha_i u_i^{(r-1)} + \beta_i u_i^{(r-2)}) = 0$$

y aplicando f:

$$\sum_{i=1}^{d_{r-1}} (\alpha_i f(u_i^{(r-1)}) + \beta_i f(u_i^{(r-2)})) = 0$$

Como los vectores $u_i^{(r-1)} \in U_{r-1} \in \ker f$ y $f(u_i^{(r-2)}) = u_i^{(r-1)}$, se tiene:

$$\sum_{i=1}^{d_{r-1}} \beta_i u_i^{(r-1)} = 0$$

que es una combinación lineal de vectores de una base igual a cero, luego los coeficientes son nulos:

$$\beta_i = 0,\ i = 1, \ldots d_{r-1}$$

De manera inmediata se prueba que también los coeficientes α_i son todos iguales a cero.

Este conjunto de vectores linealmente independientes en U_{r-2} se puede ampliar a una base de este subespacio:

$$\{u_1^{(r-1)}, \ldots, u_{d_{r-1}}^{(r-1)}, u_1^{(r-2)}, \ldots, u_{d_{r-1}}^{(r-2)}, v_{d_{r-1}+1}^{(r-2)}, \ldots, v_{s_{r-2}}^{(r-2)}\}$$

donde $s_{r-2} = d_{r-2} - d_{r-1}$.

En principio, los vectores $v_i^{(r-2)}$ se pueden elegir con cierta arbitrariedad. Podemos usar ésta para escogerlos en el núcleo de f. Las imágenes de estos vectores $v_i^{(r-2)}$ están en U_{r-1}, luego se pueden escribir en una base de este espacio:

$$f(v_k^{(r-2)}) = \sum_{i=1}^{d_{r-1}} \mu_{ik} u_i^{(r-1)},\ k = d_{r-1} + 1, \ldots, s_{r-2}$$

con lo que los vectores:

$$u_k^{(r-2)} = v_k^{(r-2)} - \sum_{i=1}^{d_{r-1}} \mu_{ik} u_i^{(r-2)},\ k = d_{r-1} + 1, \ldots, s_{r-2}$$

están en el núcleo de f. En efecto:

$$f(u_k^{(r-2)}) = f(v_k^{(r-2)}) - \sum_{i=1}^{d_{r-1}} \mu_{ik} f(u_i^{(r-2)}) = 0,\ k = d_{r-1} + 1, \ldots, s_{r-2}$$

pues: $f(u_i^{(r-2)}) = u_i^{(r-1)}$. No es difícil comprobar que estos vectores son también l.i. con el resto y que por tanto tenemos una base de U_{r-2}:

$$\begin{array}{lllll} u_1^{(r-1)}, & \dots, & u_{d_{r-1}}^{(r-1)}, & & \\ u_1^{(r-2)}, & \dots, & u_{d_{r-1}}^{(r-2)}, & \dots, & u_{s_{r-2}}^{(r-2)} \end{array}$$

que verifica:

$$f(u_i^{(r-1)}) = 0,\ f(u_i^{(r-2)}) = u_i^{(r-1)},\ f(u_j^{(r-2)}) = 0,\ i = 1, \dots, d_{r-1},\ j = d_{r-1}+1, \dots, s_{r-2}.$$

Esta misma construcción se puede hacer en U_{r-3}. Como $U_{r-2} \subset U_{r-3}$ y $f(U_{r-3}) = U_{r-2}$, existen vectores $u_i^{(r-3)} \in U_{r-3}$ tales que:

$$f(u_i^{(r-3)}) = u_i^{(r-2)}, i = 1, \dots, s_{r-2}$$

Estos vectores forman con los anteriores un sistema de vectores de U_{r-3} que es linealmente independiente. Se amplia a una base de U_{r-3} usando vectores en $\ker f$. Todas estas cuestiones se demuestran de la misma forma que se hizo en el paso anterior.

Y así sucesivamente hasta acabar la cadena. De esta forma, construimos una base del espacio V que podemos ordenar como:

$$\begin{array}{cccccccccc} u_1^{(r-1)}, & \dots, & u_{d_{r-1}}^{(r-1)}, & & & & & & \\ u_1^{(r-2)}, & \dots, & u_{d_{r-1}}^{(r-2)}, & \dots, & u_{s_{r-2}}^{(r-2)}, & & & & \\ u_1^{(r-3)}, & \dots, & u_{d_{r-1}}^{(r-3)}, & \dots, & u_{s_{r-2}}^{(r-3)}, & \dots, & u_{s_{r-3}}^{(r-3)}, & & \\ \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots & & \ddots \\ u_1^{(0)}, & \dots, & u_{d_{r-1}}^{(0)}, & \dots, & u_{s_{r-2}}^{(0)}, & \dots, & u_{s_{r-3}}^{(0)}, & \dots, u_{s_0}^{(0)} & \end{array}$$

Las propiedades más importantes de esta base son: en cada columna, $f(u_k^{(j)}) = u_k^{(j+1)}$. Por tanto, los vectores de cada columna generan un subespacio de V invariante bajo f. El espacio total V es suma directa de todos ellos. El primer vector de cada columna está en $\ker f$, es decir, es un autovector de f (los demás no son autovectores):

$$V = V_1 \oplus \cdots \oplus V_{s_0}$$

Como todos los espacios son invariantes, la matriz está formada por cajas (cuadradas) en la diagonal:

$$A = \begin{pmatrix} A_1 & 0 & 0 & \cdots & 0 \\ 0 & A_2 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & A_{s_0} \end{pmatrix}$$

Cada caja A_i tiene la siguiente forma. La base de V_i es:

$$\{u_i^{(r-k)}, u_i^{(r-k-1)}, \dots, u_i^{(0)}\}$$

y como:

$$f(u_i^{(j)}) = u_i^{(j+1)}, \quad f(u_i^{(r-k)}) = 0$$

la caja i (correspondiente al subespacio $=V_i$) es:

$$\begin{pmatrix} 0 & 1 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 1 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & 0 & \cdots & 1 \\ 0 & 0 & 0 & 0 & \cdots & 0 \end{pmatrix}$$

como se decía en el enunciado del teorema. El orden de cada caja coincide con la dimensión de los subespacios V_i. QED

Nótese que todas las cajas corresponden al mismo autovalor, 0, el único que tiene el endomorfismo nilpotente. El orden de la primera caja es r, el orden de nilpotencia del endomorfismo. El orden de las demás cajas es menor o igual a r y hay que calcularlo en cada caso.

Esta forma de la matriz del endomorfismo nilpotente f se llama **forma canónica de Jordan** de f.

Ejemplo 4.5.2 Sea $f \in \mathrm{End}(V)$, $V = \mathbb{R}^4$. Supongamos que la matriz de f en la base canónica es:

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

y que deseamos hallar una base en la cual f tenga la forma canónica del teorema anterior. Calculemos las potencias sucesivas de esta matriz:

$$A^2 = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad A^3 = 0$$

Se trata de un endomorfismo nilpotente de orden 3.

La cadena de subespacios que aparece en el teorema es:

$$U_3 = \{0\} \subset U_2 = \mathrm{im}\, f^2 = f(U_1) \subset U_1 = \mathrm{im}\, f = f(U_0) \subset U_0 = \mathbb{R}^4$$

Calculemos U_1:

$$\begin{pmatrix} 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} = \begin{pmatrix} z+t \\ t \\ t \\ 0 \end{pmatrix}$$

luego:

$$U_1 = f(\mathbb{R}^4) = \mathrm{lin}\left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \end{pmatrix} \right\}$$

El espacio U_2 es:

$$\begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} = \begin{pmatrix} t \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

luego:

$$U_2 = f(U_1) = \mathrm{lin}\left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} \right\}$$

De acuerdo con el teorema, seleccionamos una base en el espacio $U_{r-1} = U_2$. Escogemos el vector calculado antes:

$$u_1^{(2)} = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

y ahora calculamos una base de U_1. El primer vector de la base es $u_1^{(1)}$, tal que $f(u_1^{(1)}) = u_1^{(2)}$. Entonces:

$$\begin{pmatrix} z+t \\ t \\ t \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

por ejemplo:

$$u_1^{(1)} = \begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \end{pmatrix}$$

Ahora deberíamos ampliar este conjunto $\{u_1^{(2)}, u_1^{(1)}\}$ a una base de U_1. Pero ya lo es. Solo nos queda U_0. Buscamos un vector de $U_0 = \mathbb{R}^4$, $u_1^{(0)}$ tal que $f(u^{(0)}) = u_1^{(1)}$:

$$\begin{pmatrix} z+t \\ t \\ t \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \end{pmatrix}$$

por ejemplo:

$$u_1^{(0)} = \begin{pmatrix} 0 \\ 0 \\ -1 \\ 1 \end{pmatrix}$$

y completamos la base con un vector de $\ker f$, l.i. con los anteriores. Las ecuaciones de $\ker f$ son $z = t = 0$. Elijamos:

$$u_2^{(0)} = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}$$

y por la tanto, la base de todo el espacio es:

$$u_1^{(2)} = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$
$$u_1^{(1)} = \begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \end{pmatrix}$$
$$u_1^{(0)} = \begin{pmatrix} 0 \\ 0 \\ -1 \\ 1 \end{pmatrix} \quad u_2^{(0)} = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}$$

Hay dos subespacios invariantes:

$$V_1 = \mathrm{lin}\left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ -1 \\ 1 \end{pmatrix} \right\}, \quad V_2 = \left\{ \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} \right\}.$$

Las cajas correspondientes en la matriz son:

$$V_1 \to \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}, \quad V_2 \to (0)$$

y la forma canónica de la matriz es:

$$J = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

La matriz de cambio de base (de la encontrada a la inicial) está formada por los vectores de la base:

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

y se tiene por tanto:

$$A = PJP^{-1}$$

4.6 Formas canónicas de endomorfismos

Veamos ahora como podemos encontrar para un endomorfismo cualquiera un forma canónica similar a la anterior, la **forma canónica de Jordan** de un endomorfismo.

Para ello, lo primero es descomponer el espacio en una serie de subespacios, a cada uno de los cuales se asocia un endomorfismo nilpotente cuya forma canónica conocemos.

Definición 4.6.1 *Sea* $f \in \mathrm{End}(V)$*,* $\lambda \in \sigma(f) \subset \mathbb{K}$*. Se dice que el vector* $v \in V$*,* $v \neq 0$ *es un vector propio generalizado de* V *si existe un entero positivo* r *tal que:*

$$(f - \lambda 1_V)^r v = 0$$

Los vectores propios generalizados no son en general autovectores, aunque todo autovector es un vector propio generalizado (con $r = 1$).

Definición 4.6.2 *Se definen los espacios invariantes generalizados como:*

$$N_\lambda = \{v \in V \mid \exists r \geq 0,\ (f - \lambda 1_V)^r v = 0\}$$

Se tiene el siguiente resultado sobre las propiedades de estos espacios.

Proposición 4.6.1 *Con las notaciones anteriores,*
i. N_λ *es un subespacio vectorial de* V
ii. $f - \lambda 1_V$ *es nilpotente en* N_λ*:* $\left[(f - \lambda 1_V)|_{N_\lambda}\right]^r = 0$ *para algún entero positivo* r*.*
iii. N_λ *es invariante bajo* f*.*

Demostración. La primera propiedad es muy sencilla de probar. Si v_1 y v_2 son dos vectores de N_λ que son anulados por $f - \lambda 1_V$ elevado a las potencias r_1 y r_2 respectivamente, cualquier combinación lineal de estos vectores es anulada por ese operador elevado al mayor de r_1 y r_2. En cuanto a la segunda, basta considerar los exponentes que se necesitan para anular los vectores de una base de N_λ y coger el mayor de todos ellos. Finalmente, la tercera se prueba como sigue. Sea $v \in N_\lambda$,con:

$$(f - \lambda 1_V)^r v = 0$$

para algún entero positivo r. Como $f - \lambda 1_V$ conmuta con f, se tiene:

$$f((f - \lambda 1_V)^r v) = 0 \Rightarrow (f - \lambda 1_V)^r f(v) = 0$$

y por tanto, N_λ es invariante bajo f. QED

El punto más importante es que estos espacios invariantes generalizados forman una suma directa. Y no sólo eso. Si el cuerpo es algebraicamente cerrado (por ejemplo $\mathbb{C}$, o si se trata de $\mathbb{R}$, si todas las raíces del polinomio característico están en $\mathbb{R}$) la suma directa de estos espacios cuando se consideran todos los autovalores es igual al espacio total.

Probaremos un resultado preliminar.

Proposición 4.6.2 *Si $\mu \in \mathbb{K}$, $\mu \neq \lambda$, entonces la restricción de $f - \mu 1_V$ al subespacio N_λ, $(f - \mu 1_V)|_{N_\lambda}$ tiene inverso.*

Demostración. El subespacio N_λ es invariante bajo $f - \mu 1_V$. Veamos que el núcleo de $(f - \mu 1_V)|_{N_\lambda}$ es igual a $\{0\}$, o lo que es lo mismo, $(f - \mu 1_V) \cap N_\lambda = \{0\}$. Sea $v \in N_\lambda$ tal que $(f - \mu 1_V)v = 0$. Aplicando $f - \lambda 1_V$ a v:

$$(f - \lambda)(v) = f(v) - \lambda v = (\mu - \lambda)v$$

Si $v = 0$ la aplicación es inyectiva. Si $v \neq 0$, entonces es un autovector de $f - \lambda 1_V$ con autovalor $\mu - \lambda$. Pero $f - \lambda 1_V$ es nilpotente en N_λ, de donde $\mu = \lambda$ en contra de la hipótesis. QED

Como ocurría con los subespacios V_λ, los subespacios N_λ asociados a autovalores distintos, forman una suma directa.

Proposición 4.6.3 *Sean $\lambda_1, \ldots, \lambda_m$ autovalores de f distintos. Entonces, los subespacios $N_{\lambda_1}, \ldots, N_{\lambda_m}$ forman una suma directa.*

Demostración. Como en el teorema para los subespacios V_λ, lo haremos por inducción en m. Para $m = 1$ es trivialmente cierto. Supongamos que es correcto para $m - 1$. Para el caso m, consideremos la suma:

$$v_1 + \cdots + v_m = 0, \ v_i \in N_{\lambda_i}, \ i = 1, \ldots, m$$

y demostremos que cada v_i es igual a 0. Para ello, como $v_m \in N_{\lambda_m}$, existe un entero positivo s tal que:

$$(f - \lambda_m 1_V)^s v_m = 0$$

Aplicando a la suma de v_i este operador:

$$(f - \lambda_m 1_V)^s v_1 + \cdots + (f - \lambda_m 1_V)^s v_{m-1} = 0$$

que es el caso $m - 1$ (recordando que los espacios N_{λ_i} son invariantes). Por la hipótesis de inducción:

$$(f - \lambda_m 1_V)^s v_i = 0, \ i = 1, \ldots, m - 1$$

Pero hemos demostrado que el operador $f - \lambda_m 1_V$ era inyectivo en N_{λ_i}, con $i = 1, \ldots, m - 1$ por ser los autovalores distintos. Por tanto, $v_i = 0$, $i = 1, \ldots, m - 1$, lo que implica que también $v_m = 0$. QED

Hemos visto que la multiplicidad geométrica de un autovalor, la dimensión del subespacio V_λ, era siempre menor o igual que la algebraica. Para los subespacios N_λ se tiene el siguiente resultado.

Teorema 4.6.1 *Si n_{λ_0} es la multiplicidad algebraica de $\lambda_0 \in \sigma(f)$,*

$$\dim N_{\lambda_0} = n_{\lambda_0}$$

Demostración. Al ser N_{λ_0} un subespacio invariante bajo f, podemos definir la restricción de f a este subespacio, $\hat{f} = f|_{N_{\lambda_0}}$ y la aplicación inducida en el espacio cociente V/N_{λ_0}, $\tilde{f}$. Hemos estudiado la forma que toma la matriz de f en una base adaptada a al subespacio N_{λ_0}, lo que implica que los polinomios característicos de estas tres aplicaciones están relacionados por:

$$p_f(\lambda) = p_{\hat{f}}(\lambda) p_{\tilde{f}}(\lambda)$$

El grado de $p_{\hat{f}}(\lambda)$ es la dimensión del subespacio N_{λ_0}, por tanto, si:

$$\dim N_{\lambda_0} < n_{\lambda_0}$$

entonces, λ_0 es raíz de $p_{\tilde{f}}(\lambda)$, es decir, autovalor de $\tilde{f}$, de donde existe un autovector $v_0 + N_{\lambda_0} \in V/N_{\lambda_0}$:

$$\tilde{f}(v_0 + N_{\lambda_0}) = f(v_0) + N_{\lambda_0} = \lambda_0 v_0 + N_{\lambda_0}$$

es decir:

$$(f - \lambda 1_V)(v_0) \in N_{\lambda_0}$$

Esto quiere decir que existe un entero positivo, s tal que:

$$(f - \lambda 1_V)^{s+1}(v_0) = 0$$

y por lo tanto,

$$v_0 \in N_{\lambda_0} \Rightarrow v_0 + N_{\lambda_0} = 0$$

lo que es contradictorio con el carácter de autovector. Por lo tanto, al ser $\dim N_{\lambda_0} \leq n_{\lambda_0}$, como ya habíamos demostrado anteriormente, concluimos que ambas son iguales. QED

En cada uno de los espacios N_λ, f es igual a un endomorfismo nilpotente más la aplicación identidad por λ:

$$f|_{N_\lambda} = g_\lambda + \lambda 1_{N_\lambda}$$

Como hemos demostrado la existencia de un base donde g_λ toma la forma canónica de Jordan, está claro que en esa base f será la forma canónica de Jordan de un endomorfismo nilpotente más la aplicación identidad por el autovalor correspondiente. Esta será la forma de Jordan del endomorfismo f. Para acabar, sólo nos queda probar que la suma de los subespacios N_λ cubre todo el espacio V. Esto sólo es cierto si las raíces del polinomio característico están en el cuerpo $\mathbb{K}$ (los factores irreducibles del polinomio tienen grado 1). El resultado es cierto en $\mathbb{C}$ siempre y en $\mathbb{R}$ si no hay raíces complejas. Enunciamos el teorema para $\mathbb{C}$.

Teorema 4.6.2 *Sea V un $\mathbb{C}$-espacio vectorial de dimensión finita n. Sea $f \in \mathrm{End}(V)$ y $\sigma(f) = \{\lambda_1, \ldots, \lambda_m\}$ el espectro de f. Existe una base de V en la cual f tiene la forma canónica de Jordan: diagonal por cajas y cada caja del tipo:*

$$\begin{pmatrix} \lambda & 1 & 0 & 0 & \cdots & 0 \\ 0 & \lambda & 1 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \lambda & \cdots & 1 \\ 0 & 0 & 0 & 0 & \cdots & \lambda \end{pmatrix}$$

Téngase en cuenta que a un sólo autovalor pueden estar asociadas varias cajas.

Demostración. Por ser $\mathbb{C}$ un cuerpo algebraicamente cerrado:

$$n = n_1 + \cdots + n_m$$

donde n_i es la multiplicidad algebraica de λ_i. Los subespacios invariantes generalizados $N_{\lambda_1}, \ldots, N_{\lambda_m}$ asociados a los autovalores de f forman una suma directa. Pero

$$\dim(N_{\lambda_1} \oplus \cdots \oplus N_{\lambda_m}) = n_1 + \cdots n_m = n$$

luego:

$$V = N_{\lambda_1} \oplus \cdots \oplus N_{\lambda_m},$$

y en cada subespacio se tiene el resultado demostrado anteriormente. QED

En espacios vectoriales reales puede ocurrir que: $n_1 + \cdots + n_m < n$ y no se pueda poner la matriz en la forma canónica de Jordan. Sin embargo, si se pasa a un espacio vectorial complejo, es posible hacerlo.

Si los endomorfismos nilpotentes $f - \lambda 1_V$ son cero, el endomorfismo es diagonalizable. En caso contrario no lo es.

4.7 El teorema de Cayley-Hamilton

Si $q(\lambda)$ es un polinomio con coeficientes en $\mathbb{K}$ y $f \in \mathrm{End}(V)$, donde V es un $\mathbb{K}$-espacio vectorial de dimensión finita, se puede definir el endomorfismo $q(\lambda)$:

$$q(\lambda) = a_m \lambda^m + \cdots + a_1 \lambda + a_0 \longrightarrow q(f) = a_m f^m + \cdots + a_1 f + a_0 1_V$$

Los operadores $1_V, f, \dots, f^m$ no pueden ser todos independientes (para m suficientemente grande ya que $\dim \mathrm{End}(V) = n^2$) y por tanto existe una combinación lineal, con coeficientes no todos nulos, igual a cero. Es decir, existe $q(\lambda)$ tal que $q(f) = 0$.

El teorema de Cayley-Hamilton establece la existencia de un polinomio de grado $n = \dim V$ que anula al endomorfismo.

Teorema 4.7.1 *Sea V un $\mathbb{C}$-espacio vectorial de dimensión finita y $f \in \mathrm{End}(V)$. Entonces, el polinomio característico de f anula a f:*

$$p_f(f) = 0$$

Demostración. Si $n = \dim V$ y $f = g + \lambda 1_V$, donde g es un endomorfismo nilpotente de una caja, el polinomio característico es: $(\lambda_0 - \lambda)^n$, que anula a f:

$$p_f(f) = (\lambda_0 1_V - f)^n = (\lambda_0 1_V - g - \lambda_0 1_V)^n = (-g)^n = 0$$

Sea ahora f un endomorfismo arbitrario de V. De acuerdo con la forma canónica de Jordan, en cada caja f tiene la forma anterior, y el polinomio característico de f es el producto de los polinomios característicos asociados a cada caja. Si $f_i = f|_{N_i}$, el polinomio característico en N_i anula a f_i:

$$p_{f_i}(f_i) = 0$$

lo que implica que el polinomio característico de f anula también a las restricciones f_i y por lo tanto a f (al ser suma directa). QED

El resultado es también cierto en $\mathbb{R}$, incluso aunque el polinomio característico tenga raíces complejas y no exista una forma canónica de Jordan.

4.8 Polinomio mínimo

De acuerdo con el teorema de Cayley-Hamilton, el polinomio característico anula a f. Sin embargo, no es, en general, el polinomio de menor grado entre los que anulan a f.

Proposición 4.8.1 *El conjunto $\mathcal{I}_f = \{q \in \mathbb{K}[\lambda] \mid q(f) = 0\}$ es un ideal en $\mathbb{K}[\lambda]$.*

La demostración es inmediata. $\mathcal{I}_f$ es no vacío al contener al polinomio característico.

Todos los ideales en $\mathbb{K}[\lambda]$ son principales, por lo que existe un polinomio de grado mínimo en $\mathcal{I}_f$ y todo otro polinomio del ideal es múltiplo de éste.

Definición 4.8.1 *Se llama polinomio mínimo de $f \in \mathrm{End}(f)$ al polinomio de menor grado entre los que anulan a f. Se elige con el coeficiente de mayor grado igual a 1.*

Veamos ahora cual es el polinomio mínimo de los endomorfismos nilpotentes. Sea $f \in \mathrm{End}(V)$ un endomorfismo nilpotente y $n = \dim V$. El polinomio característico de f es $p(\lambda) = (-\lambda)^n$, pero si f es de grado de nilpotencia r, $1 \le r \le n$, el polinomio mínimo es:

$$m_f(\lambda) = \lambda^r$$

Hay que señalar que si $r = 1$ el endomorfismo f es cero (y trivialmente diagonalizable), mientras que si $r > 1$ no es diagonalizable.

De acuerdo con lo demostrado para la forma canónica de Jordan, si f es un endomorfismo nilpotente de varias cajas, el orden de nilpotencia de f es la dimensión de la mayor de las cajas, número que coincide, como acabamos de ver, con el grado del polinomio mínimo:

$$n = n_1 + \cdots + n_k,\ n_1 \ge \cdots \ge n_k \ge 1, \quad m_f(\lambda) = \lambda^{n_1}$$

Para un endomorfismo cualquiera, en cada subespacio invariante generalizado N_{λ_0}, el operador $f - \lambda_0 1_V$ es nilpotente de grado n_0, donde n_0 es la dimensión de la mayor de las cajas de este endomorfismo en la forma canónica de Jordan. Por lo tanto, el polinomio mínimo de $(f - \lambda_0 1_V)|_{N_\lambda}$ es: $(\lambda - \lambda_0)^{n_0}$.

Para que f sea diagonalizable en N_{λ_0}, las cajas deben tener dimensión 1 y por lo tanto el polinomio mínimo debe ser $\lambda - \lambda_0$. Hemos demostrado el siguiente resultado:

Proposición 4.8.2 *Sea V un $\mathbb{C}$-espacio vectorial de dimensión finita y $f \in \mathrm{End}(V)$. El endomorfismo f es diagonalizable si y sólo si las raíces del polinomio mínimo tienen multiplicidad igual a 1.*

Nótese que las raíces del polinomio mínimo, según hemos visto al analizar la forma canónica de Jordan coinciden con los autovalores, es decir con las raíces del polinomio característico. Los polinomios mínimo y característico tienen la mismas raíces pero en general distintas multiplicidades.

Tema 5

Espacios con producto escalar

El espacio dual. Formas bilineales. Diagonalización. Ortogonalidad. Formas cuadráticas. Formas sesquilineales. Producto escalar.

El producto escalar aparece como una estructura adicional en la teoría de espacios vectoriales. Aunque los resultados expuestos se consideran solo en espacios de dimensión finita, muchos de ellos pueden ser aplicados a espacios de dimensión infinita. Se estudian primero las formas bilineales, para pasar después a las simétricas definidas positivas (en el caso de espacios reales) o sesquilineales en el caso de espacios complejos).

5.1 El espacio dual

Repasaremos en primer lugar nociones del espacio dual ya introducidas en el Tema 3.

5.1.1 Introducción

Sea V un $\mathbb{K}$-espacio vectorial de dimensión finita, con $\mathbb{K} = \mathbb{R}, \mathbb{C}$. Sea $\mathcal{L}(V, \mathbb{K})$ el espacio vectorial de los homomorfismos de V en $\mathbb{K}$. Su dimensión es igual a la dimensión de V (pues $\dim \mathbb{K} = 1$).

Definición 5.1.1 *Se llama a $V^* = \mathcal{L}(V, \mathbb{K})$ el espacio dual de V. Los elementos de V^* se llaman formas lineales:*

$$\omega \in V^*, \qquad \omega : V \to \mathbb{K}, \quad \text{lineal}$$

Proposición 5.1.1 $\dim V^* = \dim V$.

Es una consecuencia inmediata de la definición del espacio dual. Introducimos ahora una base especial en el espacio dual: la base dual.

Proposición 5.1.2 *Sea $\mathcal{B} = \{u_1, \dots, u_n\}$ una base de V. El conjunto de formas que verifican:*

$$u_i^*(u_j) = \delta_{ij}, \quad i, j = 1, \dots n$$

es una base de V^*, llamada la base dual de $\mathcal{B}$.

Demostración. Las formas lineales quedan definidas al dar las imágenes de los vectores de una base. Veamos que son linealmente independientes. Sea:

$$\sum_{i=1}^{n} \lambda_i u_i^* = 0$$

Al aplicar la forma lineal del miembro izquierdo de esta ecuación a los vectores de la base $\mathcal{B}$ se obtiene:

$$\sum_{i=1}^{n} \lambda_i u_i^*(u_j) = \sum_{i=1}^{n} \lambda_i \delta_{ij} = \lambda_j = 0, \quad j = 1, \ldots, n$$

luego son linealmente independientes. Al ser n formas ($n = \dim V^*$), son una base. QED

Dado un vector del espacio V, sus coordenadas en una base $\mathcal{B}$ se calculan haciendo actuar las formas de la base dual correspondiente sobre el vector:

$$x \in V, \quad x = \sum_{i=1}^{n} x_i u_i, \quad u_j^*(x) = \sum_{i=1}^{n} x_i u_j^*(u_i) = \sum_{i=1}^{n} x_i \delta_{ij} = x_j$$

Una notación muy empleada para la acción de las formas sobre los vectores es:

$$\langle x, \omega \rangle$$

en la que se pone de manifiesto el carácter lineal de la actuación de ω, y a la vez las propiedades de espacio lineal de V^*

5.1.2 El espacio bidual

Se define el espacio bidual de un espacio vectorial V como el dual de su dual:

$$V^{**} = (V^*)^* = \mathcal{L}(V^*, \mathbb{K})$$

es decir, los elementos del bidual son las aplicaciones lineales de V^* en $\mathbb{K}$.

Existe un isomorfismo natural entre un espacio y su bidual (en dimensión finita), definido de la forma siguiente:

$$\begin{array}{rccl} \phi\colon & V & \to & V^{**} \\ & x & \mapsto & \phi(x)\colon \begin{array}{ccc} V^* & \to & \mathbb{K} \\ \omega & \mapsto & \phi(x)(\omega) \end{array} \end{array}.$$

Ahora bien:

$$\begin{array}{rccc} \omega\colon & V & \to & \mathbb{K} \\ & x & \mapsto & \omega(x) \end{array},$$

lo que sugiere la definición de ϕ como:

$$\phi(x)(\omega) = \omega(x)$$

o, en la segunda notación:

$$\langle \omega, \phi(x) \rangle = \langle x, \omega \rangle$$

Veamos que ϕ es un isomorfismo. La aplicación está bien definida. Además, es lineal:

$$\phi(x + y)(\omega) = \omega(x + y) = \omega(x) + \omega(y) = \phi(x)(\omega) + \phi(y)_(\omega)$$

lo que es cierto para toda forma ω. Por tanto:

$$\phi(x + y) = \phi(x) + \phi(y)$$

De la misma forma:

$$\phi(\lambda x)(\omega) = \omega(\lambda x) = \lambda \omega(x) = \lambda \phi(x)(\omega)$$

es decir:

$$\phi(\lambda x) = \lambda \phi(x)$$

También se tiene que ϕ es biyectiva. Demostremos que su núcleo es trivial.

$$\phi(x)(\omega) = 0 \Rightarrow \omega(x) = 0, \ \forall \omega \in V^*$$

pero eso quiere decir que $x = 0$. Por tanto, ϕ es inyectiva. Como las dimensiones del espacio inicial (V) y final (V^{**}) coinciden, la aplicación es biyectiva y tenemos un isomorfismo.

No existe un isomorfismo natural (como éste) entre un espacio y su dual. Más adelante estudiaremos como definir tal isomorfismo cuando V está dotado de un producto escalar.

5.1.3 Anulador

Sea S un subconjunto de V.

Definición 5.1.2 *El anulador de S es un subespacio de V^* dado por:*

$$S^0 = \{\omega \in V^* \mid \omega(x) = 0, \ \forall x \in S\}$$

Es fácil ver que S^0 es un subespacio:

$$\omega, \omega' \in S^0, \ (\omega + \omega')(x) = \omega(x) + \omega'(x) = 0$$

$$\omega \in S^0, \lambda \in \mathbb{K}, \ (\lambda\omega)(x) = \lambda(\omega(x)) = 0$$

Si $S^* \subset V^*$, el anulador de este subconjunto estaría en V^{**}, que hemos visto que es isomorfo de una forma natural a V.

$$(S^*)^0 = \{\alpha \in V^{**} \mid \alpha(\omega) = 0, \ \forall \omega \in S^*\}$$

Usando el isomorfismo, se suele identificar V con V^{**} y definir el anulador de S^* como:

$$(S^*)^0 = \{x \in V \mid \omega(x) = 0, \ \forall \omega \in S^*\}$$

Si W es un subespacio de V, el anulador del anulador de W coincide con W, como es fácil deducir de las definiciones anteriores. Además se tiene el siguiente resultado:

Proposición 5.1.3 *Si W es un subespacio de V, entonces:*

$$\dim W^0 = \dim V - \dim W$$

Demostración. Sea $\mathcal{B}_W = \{u_1, \ldots, u_k\}$ una base de W, y ampliemos esta base a una de V:

$$\mathcal{B} = \{u_1, \ldots, u_k, u_{k+1}, \ldots, u_n\}$$

Construyamos la base dual: $\mathcal{B}^* = \{u_1^*, \ldots, u_k^*, u_{k+1}^*, \ldots, u_n^*\}$.

Demostremos ahora que el conjunto: $\mathcal{B}_{W^0}^* = \{u_{k+1}^*, \ldots, u_n^*\}$ es una base de W^0. Cada elemento de este conjunto está en W^0, pues:

$$u_j^*(u^i) = 0, \ j = k+1, \ldots, n, \ i = 1, \ldots, k$$

al ser bases duales. Además, sea $\omega \in W^0$. Entonces:

$$\omega(u_i) = 0, \ i = 1, \ldots, k$$

Como ω es un elemento del dual, se puede escribir en la base $\mathcal{B}^*$:

$$\omega = \sum_{i=1}^{n} \lambda_i u_i^*$$

y usando el que ω es un elemento de V^*:

$$\omega(u_i) = \lambda_i = 0, \ i = 1, \ldots, k$$

Por lo tanto ω es una combinación lineal de los elementos de $\mathcal{B}_{W^0}^*$, que forman una base. La relación entre las dimensiones es ahora inmediata. QED

5.1.4 La aplicación transpuesta

Las aplicaciones entre espacios vectoriales se pueden llevar a sus duales. Sea $f: V \to V'$ un homomorfismo de espacios vectoriales. Sean V^* y V'^* los espacios duales de V y V' respectivamente.

Definición 5.1.3 *La aplicación transpuesta de f es:*

$$f^t: V'^* \to V$$

donde:

$$f^t(\omega')(x) = \omega'(f(x)), \quad \forall \omega' \in V^*, \ \forall x \in V$$

También podemos escribir:

$$\langle x, f^t(\omega')\rangle = \langle f(x), \omega'\rangle$$

La aplicación f^t está bien definida y es lineal. En efecto, dado ω', $f^t(\omega')$ es una aplicación lineal de V en $\mathbb{K}$:

$$f^t(\omega'_1 + \omega'_2)(x) = (\omega'_1 + \omega'_2)(f(x)) = \omega'_1 f(x) + \omega'_2(f(x)) = f^t(\omega'_1)(x) + f^t(\omega'_2)(x)$$

$$f^t(\lambda\omega')(x) = (\lambda\omega')(f(x)) = \lambda(\omega' f(x)) = \lambda f^t(\omega')(x)$$

5.1.5 La matriz de la aplicación transpuesta

Veamos que relación existe entre las matrices de una aplicación y su transpuesta. Sean $\mathcal{B}$ y $\mathcal{B}'$ bases de V y V' y $\mathcal{B}^*$, $\mathcal{B}'^*$ sus bases duales respectivamente. Sean $n = \dim V$, $n' = \dim V'$.

Sean $A_f = (a_{ij})$ y $A_{f^t} = (b_{ij})$ las matrices de f en las bases $\mathcal{B}, \mathcal{B}'$ y de f^t en las bases duales $\mathcal{B}'^*, \mathcal{B}^*$. Entonces:

$$f(u_i) = \sum_{j=1}^{n'} a_{ji} u'_j, \quad f^t(u_i'^*) = \sum_{j=1}^{n} b_{ji} u_j^*$$

Los elementos de las matrices que representan a f y f^t se pueden calcular de la formas siguiente:

$$u_j'^*(f(u_i)) = u_j'^*(\sum_{k=1}^{n'} a_{ki} u'_k) = \sum_{k=1}^{n'} a_{ki} u_j'^*(u'_k) = \sum_{k=1}^{n'} a_{ki}\delta_{jk} = a_{ji}$$

De forma similar:

$$f^t(u_j'^*)(u_i) = \sum_{k=1}^{n} b_{kj} u_k^*(u_i) = \sum_{k=1}^{n} b_{kj}\delta_{ki} = b_{ij}$$

Pero, por la definición de aplicación transpuesta:

$$f^t(u_j'^*)(u_i) = u_j'^*(f(u_i))$$

y por tanto:

$$a_{ji} = b_{ij}$$

y las matrices (en las bases duales) son transpuestas una de la otra:

$$A_{f^t} = A_f^t$$

Si usamos notación vectorial para las aplicaciones lineales, sea:

$$X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \quad A_f = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n'1} & \cdots & a_{n'n} \end{pmatrix}, \quad \Omega = \begin{pmatrix} \omega_1 \\ \vdots \\ \omega_n \end{pmatrix}, \quad A_{f^t} = \begin{pmatrix} b_{11} & \cdots & b_{1n'} \\ \vdots & & \vdots \\ b_{n1} & \cdots & b_{nn'} \end{pmatrix}$$

donde X son las coordenadas de $x \in V$ en una base $\mathcal{B}$ y Ω son las coordenadas de ω en la base dual $\mathcal{B}^*$. Entonces,

$$\omega(x) = \Omega^t X$$

en el sentido de producto de matrices. La acción de los homomorfismos f y f^t se escribe con matrices:

$$X' = A_f X, \quad \Omega = A_{f^t}\Omega'$$

Por tanto, usando nuevamente la definición de la aplicación transpuesta, se tiene:

$$(A_{f^t}\Omega')^t X = (\Omega')^t A_f X$$

es decir:

$$(\Omega')^t A^t_{f^t} X = (\Omega')^t A_f X$$

y como el resultado debe ser cierto para todo vector x y toda forma ω', se tiene el resultado anterior:

$$A^t_{f^t} = A_f$$

Entre los núcleos e imágenes de f y f_t existen las siguientes relaciones:

$$\ker f^t = (\operatorname{Im} f)^0, \quad \operatorname{Im} f^t = (\ker f)^0$$

En efecto, probemos la primera de ellas. Si $\omega' \in \ker f^t$, se tiene $f^t(\omega') = 0$, es decir:

$$\omega'(f(x)) = 0, \ \forall x \in V$$

lo que quiere decir:

$$\omega'(x') = 0, \ \forall x' \in \operatorname{Im} f$$

o sea:

$$\omega' \in (\operatorname{Im}\ f)^0$$

Supongamos ahora que $\omega' \in (\operatorname{Im} f)^0$. Siguiendo el camino anterior a la inversa, concluimos que $\omega' \in \ker f^t$ y por tanto la primera relación queda demostrada. La segunda se demuestra de forma similar. Sea $\omega \in \operatorname{im}\ f^t$. Entonces, existe $\omega' \in V'^*$ tal que: $f^t(\omega') = \omega$. Por tanto:

$$\omega(x) = \omega'(f(x)).$$

Si $x \in \ker f$, $\omega(x) = 0$, y por tanto $\omega \in (\ker f)^0$. Es decir,

$$\operatorname{im}\ f^t \subset (\ker f)^0.$$

Pero:

$$\dim \operatorname{im}\ f^t = n - \dim\ker f^t = n - \dim(\operatorname{im}\ f)^0 = n - (n' - \dim\operatorname{im}\ f) = n - \dim\ker f = \dim(\ker f)^0,$$

lo que prueba la igualdad de ambos subespacios.

Ejemplo 5.1.1 Sea $V = \mathbb{R}_2[x] = \operatorname{lin}\{1, x, x^2\}$. La base dual viene dada por las tres formas:

$$E_1(p) = p(0), \quad E_2(p) = p'(0), \quad E_3(p) = \frac{1}{2}p''(0), \quad \forall p \in V$$

Está claro que son tres elementos de dual, y que:

$$E_i(p_j(x)) = \delta_{ij}, \quad i, j = 1, 2, 3$$

donde:

$$p_1(x) = 1, \quad p_2(x) = x, \quad p_3(x) = x^2$$

Cualquier polinomio se escribe en esta base como:

$$p(x) = E_1(p) + E_2(p)x + E_3(p)x^2$$

Se considera ahora la aplicación derivada en V:

$$\begin{array}{cccc} D\colon & V & \to & V \\ & p(x) & \mapsto & p'(x) \end{array}$$

que tiene como matriz en la base dada:

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

La aplicación transpuesta verifica:

$$D^t(\omega')(p) = \omega'(Dp)$$

Sustituyendo ω' por los elementos de la base:

$$\begin{aligned} D^t(E_1)(p) &= E_1(Dp) = p'(0) = E_2(p) \\ D^t(E_2)(p) &= E_2(Dp) = p''(0) = 2E_3(p) \\ D^t(E_3)(p) &= E_3(Dp) = p'''(0)/2 = 0 \end{aligned}$$

por lo que la matriz de la aplicación transpuesta en la base dual es:

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

5.2 Formas bilineales

Se estudian en esta sección las formas bilineales, especialmente las simétricas y se introducen las aplicaciones multilineales en forma general.

5.2.1 Aplicaciones multilineales

Definición 5.2.1 *Sean $V_1, \ldots, V_n, W$ $\mathbb{K}$-espacios vectoriales. Se dice que la aplicación:*

$$\varphi\colon V_1 \times \ldots \times V_n \to W$$

es multilineal si es lineal en cada variable separadamente:

$$\begin{aligned} &\varphi(x_1, \ldots, x_i + x_i', \ldots, x_n) = \varphi(x_1, \ldots, x_i, \ldots, x_n) + \varphi(x_1, \ldots, x_i', \ldots, x_n) \\ &\varphi(x_1, \ldots, \lambda x_i, \ldots, x_n) = \lambda\varphi(x_1, \ldots, x_i, \ldots, x_n) \end{aligned}$$

El conjunto de las aplicaciones multilineales forma un espacio vectorial: $\mathcal{L}(V_1, \ldots, V_n; W)$.

5.2.2 Formas bilineales

Definición 5.2.2 *Una forma bilineal es una aplicación multilineal de $V \times V$ en $\mathbb{K}$, donde V es un $\mathbb{K}$-espacio vectorial*

Una forma bilineal es pues:

$$\varphi\colon V \times V \to \mathbb{K}$$

tal que:

$$\begin{aligned} \varphi(x + y, z) &= \varphi(x, z) + \varphi(y, z) \\ \varphi(x, y + z) &= \varphi(x, y) + \varphi(x, z) \\ \varphi(\lambda x, y) &= \lambda\varphi(x, y) \\ \varphi(x, \lambda y) &= \lambda\varphi(x, y) \end{aligned}$$

para $x, y, z \in V$, $\lambda \in \mathbb{K}$. El conjunto de formas bilineales es un espacio vectorial, y cuando la dimensión de V es finita, la de $\mathcal{L}(V, V; \mathbb{K}) = \mathcal{L}_2(V)$ es igual a la de V elevada al cuadrado.

Proposición 5.2.1 *Sea V un $\mathbb{K}$-espacio vectorial de dimensión n. Sea:*

$$\mathcal{L}_2(V) = \{\varphi: V \times V \to \mathbb{K},\ \varphi \text{ bilineal}\}$$

Entonces, $\mathcal{L}_2(V)$ es un $\mathbb{K}$-espacio vectorial. Si $\mathcal{B} = \{u_1, \ldots, u_n\}$ es una base de V y $\mathcal{B}^* = \{u_1^*, \ldots, u_n^*\}$ su base dual, entonces, el conjunto de formas bilineales:

$$\varphi_{ij}(x, y) = \langle x, u_i^* \rangle \langle y, u_j^* \rangle, \quad i, j = 1, \ldots, n$$

es una base de $\mathcal{L}_2(V)$ que por tanto, tiene dimensión n^2.

Demostración. Es muy sencillo probar que φ_{ij} es una forma bilineal. Probemos que son l.i. Sea:

$$\sum_{i,j=1}^{n} \lambda_{ij} \varphi_{ij} = 0$$

Aplicando esta expresión a los elementos de la base de V:

$$\left(\sum_{i,j=1}^{n} \lambda_{ij} \varphi_{ij}\right)(u_k, u_l) = \sum_{i,j=1}^{n} \lambda_{ij} \langle u_k, u_i^* \rangle \langle u_l, u_j^* \rangle = \sum_{i,j=1}^{n} \lambda_{ij} \delta_{ki} \delta_{lj} = \lambda_{kl}$$

luego:

$$\lambda_{ij} = 0, \quad i, j = 1, \ldots n$$

Veamos ahora que generan todo el espacio $\mathcal{L}_2(V)$: Sea $\varphi \in \mathcal{L}_2(V)$. Esta forma queda fijada calculando sus valores sobre una base de V (al ser bilineal). Es decir: $\varphi(u_i, u_j) = a_{ij}$. En efecto:

$$\varphi(x, y) = \varphi\left(\sum_{i=1}^{n} x_i u_i, \sum_{j=1}^{n} y_j u_j\right) = \sum_{i,j=1}^{n} a_{ij} x_i y_j$$

Construimos ahora la forma bilineal:

$$\tilde{\varphi}(x, y) = \sum_{i,j=1}^{n} a_{ij} \langle x, u_i^* \rangle \langle y, u_j^* \rangle$$

que es una combinación lineal de las formas φ_{ij}. Es inmediato ver que es igual a φ. En efecto, calculando los valores sobre una base:

$$\tilde{\varphi}(u_k, u_l) = \sum_{i,j=1}^{n} a_{ij} \langle u_k, u_i^* \rangle \langle u_l, u_j^* \rangle = \sum_{i,j=1}^{n} a_{ij} \delta_{ki} \delta_{lj} = a_{kl}$$

QED

5.2.3 Matriz de una forma bilineal

Como hemos visto antes, los escalares $a_{ij} = \varphi(u_i, u_j)$ determinan de manera única a la forma bilineal.

Definición 5.2.3 *Se llama matriz de una forma bilineal en una base $\mathcal{B} = \{u_1, \ldots, u_n\}$ a la matriz cuyos elementos son los escalares $\varphi(u_i, u_j)$.*

De esta manera, los valores que toma la forma bilineal sobre vectores x, y, de coordenadas en la base $\mathcal{B}$:

$$X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \quad Y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}, \quad \mathcal{A} = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix},$$

vienen dados por:

$$\varphi(x, y) = \sum_{ij} a_{ij} x_i y_j = X^t \mathcal{A} Y$$

Como la correspondencia entre formas bilineales y matrices $n \times n$ es biunívoca, tenemos establecido un isomorfismo entre estos dos espacios vectoriales. Toda matriz cuadrada representa una forma bilineal en una base dada de V. Nótese que esta es otra utilización de las matrices aparte de la ya considerada de representar aplicaciones lineales.

Evidentemente, cuando cambiamos de base, la matriz de la forma bilineal cambia (como ocurría con las matrices que representaban homomorfismos). Veamos como es este cambio.

Sean $\mathcal{B}, \mathcal{B}'$ dos bases del espacio vectorial V:

$$\mathcal{B} = \{u_1, \ldots, u_n\}, \quad \mathcal{B}' = \{u'_1, \ldots, u'_n\}$$

con la matriz de cambio de base P:

$$u'_i = \sum_{j=1}^{n} P_{ji} u_j, \quad i = 1, \ldots, n$$

En las coordenadas, el cambio de base es:

$$x_i = \sum_{j=1}^{n} P_{ij} x'_j, \quad X = PX'$$

por tanto:

$$\varphi(x, y) = X^t \mathcal{A} Y = (PX')^t \mathcal{A} Y' = (X')^t P^t \mathcal{A} P Y' = (X')^t \mathcal{A}' Y'$$

y se tiene la relación:

$$\mathcal{A}' = P^t \mathcal{A} P$$

Hay que señalar la diferencia que aparece con respecto al caso de matrices que representan homomorfismos (o en particular endomorfismos). Allí es la inversa de la matriz P la que aparece en esta relación, mientras que aquí es la transpuesta la que juega ese papel.

5.2.4 Formas bilineales simétricas y antisimétricas

Definición 5.2.4 *Sea $\varphi: V \times V \to \mathbb{K}$ una forma bilineal. Se dice que φ es simétrica si:*

$$\varphi(x, y) = \varphi(y, x), \quad \forall x, y \in V$$

Se dice que φ es antisimétrica si:

$$\varphi(x, y) = -\varphi(y, x), \quad \forall x, y \in V.$$

Proposición 5.2.2 *Si φ es una forma bilineal simétrica, la matriz que representa a φ en cualquier base es una matriz simétrica.*

Demostración. Sea $\mathcal{B}$ una base de V. Entonces:

$$a_{ij} = \varphi(u_i, u_j) = \varphi(u_j, u_i) = a_{ji}$$

es decir:

$$\mathcal{A}^t = \mathcal{A}$$

QEDEs también evidente que si la matriz asociada a una forma bilineal es simétrica en una base, lo es en todas y la forma bilineal correspondiente es simétrica.

De forma similar se concluye que una forma bilineal es antisimétrica si y solo si su matriz en cualquier base es antisimétrica. Las formas simétricas forman un subespacio vectorial del espacio de las formas bilineales. Las formas antisimétricas forman también un subespacio vectorial

El espacio $\mathcal{L}_2(V)$ se descompone en suma directa de formas bilineales simétricas y antisimétricas:

$$\mathcal{L}_2(V) = \mathcal{A}(V) \oplus \mathcal{S}(V)$$

y las dimensiones de estos dos subespacios son:

$$\dim \mathcal{A}(V) = \frac{1}{2}n(n-1), \quad \dim \mathcal{S}(V) = \frac{1}{2}n(n+1)$$

5.2.5 Formas bilineales simétricas regulares

Sea $\varphi: V \times V \to \mathbb{K}$ una forma bilineal simétrica.

Definición 5.2.5 *Se define el radical de φ como el subespacio de V dado por:*

$$\operatorname{rad} \varphi = \{x \in V \mid \varphi(x, y) = 0,\ \forall y \in V\}$$

Es inmediato probar que $\operatorname{rad} \varphi$ así definido es un subespacio de V.

Definición 5.2.6 *Se dice que φ es regular (no degenerada) si su radical es el vector* 0.

Definición 5.2.7 *Se llama rango de la forma bilineal simétrica φ a:*

$$\operatorname{ran}(\varphi) = \dim V - \dim \operatorname{rad} \varphi$$

Proposición 5.2.3 *Sea φ una forma bilineal simétrica de rango r. Sea $\mathcal{A}$ la matriz de φ en una base $\mathcal{B}$. Entonces:*

$$\operatorname{ran} \varphi = \operatorname{ran} \mathcal{A}$$

Demostración. Sea W un subespacio complementario a $\operatorname{rad} \varphi$:

$$V = W \oplus \operatorname{rad} \varphi$$

y sea $\mathcal{B}$ una base adaptada a esta descomposición:

$$\mathcal{B} = \{u_1, \dots, u_r, u_{r+1}, \dots, u_n\}$$

La matriz de φ en la base $\mathcal{B}$ es:

$$\begin{pmatrix} C & 0 \\ 0 & 0 \end{pmatrix}$$

donde C es una matriz $r \times r$ que probaremos que tiene determinante distinto de cero.

Construyamos una combinación lineal de las r columnas de C igual a cero:

$$\lambda_1 \begin{pmatrix} \varphi(u_1, u_1) \\ \vdots \\ \varphi(u_r, u_1) \end{pmatrix} + \cdots + \lambda_r \begin{pmatrix} \varphi(u_1, u_r) \\ \vdots \\ \varphi(u_r, u_r) \end{pmatrix} = 0$$

es decir:

$$\lambda_1\varphi(u_i,u_1)+\cdots+\lambda_r\varphi(u_i,u_r)=0,\quad i=1,\ldots,r$$

o, usando las propiedades de bilinealidad de φ:

$$\varphi(u_i,\lambda_1u_1+\cdots+\lambda_ru_r)=0,\quad i=1,\ldots,r$$

Como consecuencia, el vector $v=\lambda_1u_1+\cdots+\lambda_ru_r$ está en el radical, pues:

$$\varphi(u_i,v)=0,\ i=1,\ldots,n$$

Sin embargo, $v\in W$ por lo que $v=0$. Como los vectores $u_1,\ldots,u_r$ son l.i., se concluye que $\lambda_1=\cdots=\lambda_r=0$ y las columnas de C son l.i. lo que asegura que el rango de A es igual a r. QED

El cambio de base no modifica el rango de la matriz (pues $\det P\neq 0$), por lo que la dimensión del radical (y por tanto el rango de la forma bilineal) es igual al rango de la matriz que representa a φ en cualquier base.

Veámoslo de otra forma. Sea $r=\mathrm{ran}(\mathcal{A})$, siendo $\mathcal{A}$ la matriz de φ en una base arbitraria. Las columnas de la matriz $\mathcal{A}$ tienen $n-r$ relaciones lineales linealmente independientes:

$$\begin{aligned}\lambda_{11}\varphi(u_i,u_1)+\cdots+\lambda_{1n}\varphi(u_i,u_n) &= 0\\ &\vdots\\ \lambda_{n-r,1}\varphi(u_i,u_1)+\cdots+\lambda_{n-r,n}\varphi(u_i,u_n) &= 0\end{aligned}$$

con $i=1,\ldots,n$. Aplicando las propiedades de φ:

$$\begin{aligned}\varphi(u_i,\lambda_{11}u_1+\cdots+\lambda_{1n}u_n) &= 0\\ &\vdots\\ \varphi(u_i,\lambda_{n-r,1}u_1+\cdots+\lambda_{n-r,n}u_n) &= 0\end{aligned}$$

Por tanto, los vectores: $\{\lambda_{11}u_1+\cdots+\lambda_{1n}u_n,\ldots,\lambda_{n-r,1}u_1+\cdots+\lambda_{n-r,n}u_n\}$ son una base del radical de φ que tiene dimensión $n-r$ (son linealmente independientes, anulan a una base y no hay más vectores l.i. con estas propiedades). QED

Con este resultado es inmediato demostrar la siguiente proposición:

Proposición 5.2.4 *Sea $\varphi: V\times V\to \mathbb{K}$ una forma bilineal simétrica. Entonces:*

$$\varphi\quad\text{regular}\ \Leftrightarrow \det\mathcal{A}\neq 0$$

donde $\mathcal{A}$ es la matriz de φ en una base de V.

5.2.6 Ortogonalidad

Sea φ una forma bilineal simétrica en un $\mathbb{K}$-espacio vectorial V.

Definición 5.2.8 *Se dice que $x,y\in V$ son ortogonales ($x\perp y$) respecto φ si $\varphi(x,y)=0$.*

Definición 5.2.9 *Se dice que $x\in V$ es isótropo (siempre respecto φ) si $\varphi(x,x)=0$.*

Definición 5.2.10 *Si U es un subespacio de V, el subespacio ortogonal a U es:*

$$U^\perp=\{x\in V\mid \varphi(x,y)=0,\ \forall y\in U\}$$

Veamos dos propiedades de la relación de ortogonalidad.

Proposición 5.2.5 *Sea φ una forma bilineal simétrica regular en V y U un subespacio de V. Entonces:*

$$\dim V = \dim U + \dim U^\perp.$$

Demostración. Sea $\mathcal{B}$ una base de V (de dimensión n), y $\mathcal{B}_U = \{u_1, \dots, u_m\}$ una base de U. Las ecuaciones de $U^\perp$ son:

$$\begin{pmatrix} b_{11} & \cdots & b_{1n} \\ \vdots & & \vdots \\ b_{m1} & \cdots & b_{mn} \end{pmatrix} \mathcal{A} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = 0$$

donde b_{ij} son las coordenadas de los vectores de la base $\mathcal{B}_U$ en la base $\mathcal{B}$ y $\mathcal{A}$ es la matriz de la forma bilineal φ en la base $\mathcal{B}$.

Tenemos un sistema de ecuaciones que define $U^\perp$: $B\mathcal{A}X = 0$. El rango de la matriz $B\mathcal{A}$ es igual al rango de la matriz B, ya que la matriz $\mathcal{A}$ es regular. Por lo tanto, la dimensión del espacio de soluciones de este sistema es el número de incógnitas, que es n, menos el de ecuaciones que es m, siendo m justamente la dimensión de U. QED

Proposición 5.2.6 *Sea φ una forma bilineal simétrica regular en V y U un subespacio de V. Si $\varphi|_U$ es regular, entonces:*

$$V = U \oplus U^\perp$$

Nota: Aunque φ sea un forma regular en V, esto no quiere decir que lo sea en cualquier subespacio.

Demostración. Veamos que la intersección de U y $U^\perp$ es igual al vector 0.

Si $x \in U \cap U^\perp$, entonces $\varphi(x, y) = 0, \forall y \in U$, ya que x está en $U^\perp$. Como x también pertenece a U, se deduce que x está en el radical de la forma φ restringida a U (pues está en U y anula a todos los vectores de U):

$$x \in \operatorname{rad} \varphi|_U$$

Pero esta forma es regular, luego su radical es el vector nulo y por tanto $x = 0$. Es decir:

$$U \cap U^\perp = \{0\}$$

Como además se tiene por la proposición anterior:

$$\dim V = \dim U + \dim U^\perp$$

concluimos que:

$$V = U \oplus U^\perp.$$

QED

5.2.7 Diagonalización de formas bilineales simétricas

Como ya hemos visto, al cambiar la base la matriz asociada a una forma bilineal cambia. Podemos intentar buscar una base en la cual la matriz sea lo más sencilla posible. Demostraremos a continuación que, para las formas bilineales simétricas, siempre es posible encontrar una base en la cual la matriz sea diagonal.

Proposición 5.2.7 *Sea V un $\mathbb{K}$-espacio vectorial de dimensión n y $\varphi: V \times V \to \mathbb{K}$ una forma bilineal simétrica de rango r. Entonces, existe una base de V, $\mathcal{B} = \{u_1, \dots, u_n\}$ en la cual se tiene:*

$$\begin{aligned} \varphi(u_i, u_j) &= 0, \qquad i \neq j \\ \varphi(u_i, u_i) &= c_i, \qquad i = 1, \dots, n \end{aligned}$$

Además:

$$c_1, \dots, c_r \neq 0, \quad c_{r+1} = \dots = c_n = 0$$

En el caso complejo, la base se puede escoger de forma que:

$$c_1 = \cdots = c_r = 1$$

En el caso real, existe un número entero mayor o igual que cero, p, tal que:

$$c_1 = \cdots = c_p = 1, \quad c_{p+1} = \cdots = c_r = -1$$

En este caso, se llama signatura de la forma bilineal φ al par (p, q), con $q = r - p$.

Demostración. El problema se puede reducir al caso regular. Sea:

$$V = W \oplus \operatorname{rad} \varphi$$

En una base adaptada a esta descomposición, la matriz de la forma bilineal es:

$$\mathcal{A} = \begin{pmatrix} B & 0 \\ 0 & 0 \end{pmatrix}$$

con $\det B \neq 0$ y por tanto $\varphi|_W$ regular.

Supongamos entonces que $\varphi\colon V \times V \to \mathbb{K}$ es una forma bilineal simétrica regular. Sea u_1 un vector de V tal que: $\varphi(u_1, u_1) = c_1 \neq 0$. Al menos existe un vector con estas características. Si no fuera así, la forma bilineal sería cero. En efecto, es inmediato probar que:

$$\varphi(x, y) = \frac{1}{2}\left(\varphi(x + y, x + y) - \varphi(x, x) - \varphi(y, y)\right)$$

lo que permite expresar una forma bilineal en términos de sus valores sobre la diagonal de $V \times V$ (el conjunto de pares con las dos componentes iguales).

Sea $W_1 = \operatorname{lin}\{u_1\}$. Como φ es regular, (y $\varphi|_{W_1}$ también) se tiene:

$$V = W_1 \oplus W_1^{\perp}$$

Al ser $\varphi|_{W_1}$ regular, se puede probar que $\varphi|_{W_1^{\perp}}$ también es regular.

Si no fuera así, existiría un vector $x \in W_1^{\perp}$, tal que $\varphi(x, y) = 0,\ \forall y \in W_1^{\perp}$. En este caso, $\varphi(x, y) = 0,\ \forall y \in V$, ya que:

$$\varphi(x, y) = \varphi(x, y_1) + \varphi(x, y_2) = 0$$

donde $y_1 \in W_1,\ y_2 \in W_1^{\perp}$, y φ no sería regular en V.

Consideremos ahora un segundo vector u_2 en el espacio $W_1^{\perp}$, tal que: $\varphi(u_2, u_2) = c_2 \neq 0$ y construyamos el subespacio:

$$W_2 = \operatorname{lin}\{u_1, u_2\}.$$

La forma bilineal φ es también regular en W_2 (pues su matriz asociada en la base $\{u_1, u_2\}$ es diagonal y los valores en la diagonal son c_1, c_2 que son distintos de cero). Por tanto:

$$V = W_2 \oplus W_2^{\perp}.$$

Como el espacio es de dimensión finita, podemos seguir este proceso hasta construir una base de V: $\{u_1, \ldots, u_n\}$, que satisface las condiciones del teorema.

Si $\mathbb{K} = \mathbb{C}$, podemos tomar otra base:

$$\left\{ \frac{1}{\sqrt{c_1}} u_1, \ldots, \frac{1}{\sqrt{c_n}} u_n \right\},$$

que hace que la matriz asociada tenga todos los elementos de la diagonal iguales a 1.

Si estamos en $\mathbb{R}$, se puede tomar como base:

$$\left\{ \frac{1}{\sqrt{|c_1|}} u_1, \ldots, \frac{1}{\sqrt{|c_n|}} u_n \right\},$$

y en ella los elementos de la diagonal de la matriz asociada son iguales a ± 1. QED

En el caso real, el número de $+1$ y -1 en la diagonal (para una matriz diagonal, lo que hemos llamado signatura de la forma bilineal) no depende de la base elegida (hay muchas bases en las cuales la matriz es diagonal y está formada por ± 1 y 0). Es claro que el número de ceros, que es igual al rango de la forma bilineal, no cambia al cambiar de base. No es tan evidente que la signatura sea un invariante característico de la forma bilineal. Se tiene el siguiente teorema:

Teorema 5.2.1 *Teorema de Sylvester. (Ley de inercia). La signatura de una forma bilineal simétrica real es un invariante.*

Demostración. Sean $\{u_1, \dots u_n\}$, $\{u'_1, \dots u'_n\}$ dos bases de V con las siguientes propiedades:

$$\begin{array}{llll} \varphi(u_i, u_j) = 0, & i \neq j & \varphi(u'_i, u'_j) = 0, & i \neq j \\ \varphi(u_i, u_i) = 1, & i = 1, \dots, p & \varphi(u'_i, u'_i) = 1, & i = 1, \dots, p' \\ \varphi(u_i, u_i) = -1, & i = p+1, \dots, r & \varphi(u'_i, u'_i) = -1, & i = p'+1, \dots, r \\ \varphi(u_i, u_i) = 0, & i = r, \dots, n & \varphi(u'_i, u'_i) = 0, & i = r, \dots, n \end{array}$$

Demostremos que $p = p'$. Para ello, supongamos primero que $p' > p$. Sea $\{u_1'^*, \dots, u_n'^*\}$ la base dual de $\{u'_1, \dots u'_n\}$. El sistema de ecuaciones:

$$\langle x, u_k'^* \rangle = 0, \quad k = p'+1, \dots, r$$

tiene soluciones no triviales cuando nos restringimos a $x \in \text{lin}\{u_{p+1}, \dots, u_r\}$. En efecto, se trata de un sistema con $r - p'$ ecuaciones y $r - p$ incógnitas, y $r - p' < r - p$ al ser $p' > p$. Sea x una de esas soluciones. Si escribimos x en la base $\{u_i\}$, resulta ser una combinación lineal de los vectores: $\{u_{p+1}, \dots, u_r\}$. Si lo hacemos en la base $\{u'_i\}$, es una combinación lineal de: $\{u'_1, \dots, u'_{p'}, u'_{r+1}, \dots, u'_n\}$, pues verifica el sistema anterior (basta recordar como se escribían las coordenadas de un vector en una base usando la base dual). Es decir:

$$x = y + z, \quad y \in \text{lin}\{u'_1, \dots, u'_{p'}\}, \quad z \in \text{lin}\{u'_{r+1}, \dots, u'_n\}.$$

El valor que toma la forma bilineal φ sobre x es:

$$\varphi(x, x) = \varphi(y, y) + \varphi(z, z) + 2\varphi(y, z) = \varphi(y, y) \geq 0.$$

Sin embargo,

$$\varphi(x, x) < 0$$

pues $x \in \text{lin}\{u_{p+1}, \dots, u_r\}$. Por lo tanto, $p \leq p'$. De manera similar probaríamos que $p' \leq p$ y concluiríamos que $p = p'$. QED

En el caso real, existen tipos de formas bilineales simétricas particularmente interesantes.

Definición 5.2.11 *Sea φ una forma bilineal simétrica definida sobre un espacio vectorial real de dimensión n. Se dice que φ es definida positiva si*

$$\varphi(x, x) > 0, \quad \forall x \in V, \ x \neq 0$$

En este caso, el rango de la forma bilineal es n y la signatura es $(n, 0)$. Se dice que φ es definida negativa si

$$\varphi(x, x) < 0, \quad \forall x \in V, \ x \neq 0$$

En este caso, el rango de la forma bilineal es también n y la signatura es $(0, n)$.

Se puede hablar también de formas semidefinidas positivas o negativas, cuando las desigualdades no son estrictas.

5.2.8 Ortonormalización de Gram-Schmidt

Del apartado anterior se deduce inmediatamente que una forma bilineal simétrica real definida positiva tiene como matriz asociada en una base apropiada, la matriz identidad. Se dice entonces que la base correspondiente es ortonormal. El proceso de ortonormalización de Gram-Schmidt es una forma particular de construir bases ortonormales a partir de una base cualquiera para formas bilineales definidas positivas.

Proposición 5.2.8 (Gram-Schmidt.) *Sea V un $\mathbb{R}$-espacio vectorial de dimensión finita n. Sea $\varphi \in \mathcal{S}_2(V)$ definida positiva y $\mathcal{B} = \{e_1, \ldots, e_n\}$ una base de V. Entonces, existe una base de V, $\mathcal{B}' = \{u_1, \ldots, u_n\}$ tal que:*
i.- $\text{lin}\{e_1, \ldots, e_r\} = \text{lin}\{u_1, \ldots, u_r\}$,
ii.- $\varphi(u_i, u_j) = \delta_{ij},\ i, j = 1, \ldots, n$, *es decir, $\mathcal{B}'$ es una base ortonormal.*

Demostración. Se define el primer vector de la base $\mathcal{B}'$ como:

$$u_1 = \frac{1}{\sqrt{\varphi(e_1, e_1)}} e_1$$

que cumple las dos condiciones del teorema, pues φ es definida positiva. El segundo vector se obtiene como una combinación lineal de u_1 y e_2. Sea:

$$u_2' = e_2 - \lambda_{21} u_1$$

donde λ_{21} es un número real a fijar. Para ello imponemos que u_2' sea ortogonal a u_1:

$$\varphi(u_1, e_2 - \lambda_{21} u_1) = \varphi(u_1, e_2) - \lambda_{21} \varphi(u_1, u_1)$$

de donde se deduce:

$$\lambda_{21} = \varphi(u_1, e_2)$$

Si ahora definimos:

$$u_2 = \frac{1}{\sqrt{\varphi(u_2', u_2')}} u_2'$$

los vectores u_1, u_2 verifican las condiciones del teorema.

Supongamos que de esta forma hemos construido los vectores $u_1, \ldots, u_r$ que verifican las condiciones del proceso de ortonormalización de Gram-Schmidt. Buscamos u_{r+1} como una combinación lineal de e_{r+1} y de los ya construidos $u_1, \ldots, u_r$:

$$u_{r+1}' = e_{r+1} - \sum_{j=1}^{r} \lambda_{r+1,j} u_j$$

Imponiendo las condiciones $\varphi(u_{r+1}', u_i) = 0,\ i = 1, \ldots, r$, se obtienen los parámetros λ_{ij}:

$$0 = \varphi(e_{r+1}, u_i) - \sum_{j=1}^{r} \lambda_{r+1,j} \varphi(u_j, u_i) = \varphi(e_{r+1}, u_i) - \sum_{j=1}^{r} \lambda_{r+1,j} \delta_{ji}$$

lo que nos da la solución:

$$\lambda_{r+1,i} = \varphi(e_{r+1}, u_i)$$

es decir:

$$u_{r+1}' = e_{r+1} - \sum_{i=1}^{r} \varphi(e_{r+1}, u_i) u_i$$

Normalizando el vector u_{r+1}' (es decir, dividiendo por la raíz cuadrada de $\varphi(u_{r+1}', u_{r+1}')$) obtenemos el vector u_{r+1}. Este vector cumple las condiciones del teorema como es fácil observar (nótese que e_{r+1} no depende linealmente de $u_1, \ldots, u_r$).

El proceso termina al ser V un espacio de dimensión finita. QED

5.3 Formas Cuadráticas

Dada una forma bilineal simétrica, φ, se verifican las identidades siguientes (identidades de polarización) que permiten expresar los valores que toma φ sobre cualquier par de vectores en función de los que toma sobre la diagonal (como ya hemos usado anteriormente).

$$\begin{aligned}\varphi(x,y) &= \frac{1}{2}[\varphi(x+y,x+y)-\varphi(x,x)-\varphi(y,y)]\\ \varphi(x,y) &= \frac{1}{4}[\varphi(x+y,x+y)-\varphi(x-y,x-y)]\end{aligned}$$

Su demostración es inmediata sin más que desarrollar.

Definición 5.3.1 *Sea V un $\mathbb{K}$-espacio vectorial. Una aplicación: $Q\colon V\to\mathbb{K}$ es una forma cuadrática si:*

$$a)\quad Q(\lambda x)=\lambda^2 Q(x)$$

y la aplicación $\varphi_Q\colon V\times V\to\mathbb{K}$, definida por:

$$b)\quad \varphi_Q(x,y)=\frac{1}{2}[Q(x+y)-Q(x)-Q(y)]$$

es una forma bilineal simétrica.

A cada forma cuadrática le corresponde una forma bilineal simétrica, y viceversa, dada una forma bilineal simétrica podemos construir un forma cuadrática mediante:

$$Q_\varphi(x)=\varphi(x,x)$$

Se tiene:

$$Q\to\varphi_Q\to Q_{\varphi_Q}=Q$$

Se dice que una forma cuadrática es definida positiva, etc, si la correspondiente forma bilineal lo es. Si $\mathcal{A}$ es la matriz de la forma bilineal φ_Q en una base $\mathcal{B}$, se dice que $\mathcal{A}$ es la matriz de la forma cuadrática en esa base. La expresión de $Q(x)$ es entonces:

$$Q(x)=X^t\mathcal{A}X$$

y por supuesto, $\mathcal{A}^t=\mathcal{A}$. Es decir, la matriz asociada a una forma cuadrática en cualquier base es simétrica.

5.3.1 Diagonalización de formas cuadráticas

El problema de reducir una forma cuadrática a una suma de cuadrados (en $\mathbb{C}$) o a una suma y diferencia de cuadrados (en $\mathbb{R}$) es el de encontrar una base en la cual la forma bilineal simétrica asociada tenga una matriz diagonal con ±1 y 0 en la diagonal. Ya hemos visto que eso es siempre posible.

Veamos ahora un método práctico de hacerlo, el método de Lagrange. La forma cuadrática Q se escribe en una base dada como:

$$Q(x)=\sum_{i,k=1}^{n}a_{ik}x_ix_k$$

donde $a_{ik}=a_{ki}$. La idea del método es la de completar cuadrados. Supongamos que existe un elemento de la diagonal no nulo, $a_{jj}\neq0$. Entonces, Q se puede escribir como:

$$Q(x)=\frac{1}{a_{jj}}\left(\sum_{i=1}^{n}a_{ji}x_i\right)^2+Q_1(x),$$

donde Q_1 es otra forma cuadrática. Lo importante es que el primer sumando es el cuadrado de una suma (salvo un factor que se puede incluir extrayendo su raíz cuadrada, o la de su valor absoluto si estamos en $\mathbb{R}$) y que Q_1 no depende de x_j. Basta desarrollar el cuadrado y restarlo de $Q(x)$ para comprobarlo. De esta forma podemos seguir el procedimiento con $Q_1(x)$ que depende de $n-1$ variables, hasta acabar el desarrollo. Podría ocurrir que en algún momento, ninguno de los elementos de la diagonal fuera distinto de cero.

Supongamos entonces que $a_{ii} = 0,\ i = 1, \ldots, n$. Existirá un elemento $a_{jh} \neq 0$. La descomposición que podemos hacer ahora es:

$$Q(x) = \frac{1}{2a_{jh}} \left(\sum_{k=1}^{n} (a_{jk} + a_{hk}) x_k \right)^2 - \frac{1}{2a_{jh}} \left(\sum_{k=1}^{n} (a_{jk} - a_{hk}) x_k \right)^2 + Q_2(x)$$

donde $Q_2(x)$ es una forma cuadrática que no depende de x_j, x_h, y las formas lineales $\sum_{k=1}^{n}(a_{jk} + a_{hk})x_k$, $\sum_{k=1}^{n}(a_{jk} - a_{hk})x_k$ son linealmente independientes. Basta desarrollar para comprobar estas afirmaciones, pero es fácil darse cuenta que no se trata más que de una generalización de la siguiente (y evidente) identidad:

$$2xy = \frac{1}{2}(x+y)^2 - \frac{1}{2}(x-y)^2$$

Al descomponer en suma de cuadrados (o suma y diferencia), las formas lineales que aparecen (elevadas al cuadrado) en la descomposición son linealmente independientes (en el primero de los supuestos es trivial, pues dependen de un número de variables distinto cada vez; en el segundo se puede comprobar como se ha dicho anteriormente). Esta formas lineales dan el cambio de base, o al menos parte de él, pues la forma puede no ser regular (con radical no nulo) y aparecer menos de n cuadrados en la suma. En este último caso no es difícil completarlas con otras formas l.i. hasta tener la expresión de la nueva base en la que la forma cuadrática es diagonal.

5.3.2 Formas cuadráticas definidas

Si una forma cuadrática (o una forma bilineal simétrica) está escrita en una base arbitraria, no es posible deducir si es definida positiva o no de una inspección de los elementos de la matriz, como es el caso en el que esta matriz es diagonal. Sin embargo se puede dar un criterio sencillo que permite averiguar esta propiedad mediante el cálculo de los menores principales (los determinantes de las matrices que están construidas sobre la diagonal, tomando los elementos de las r primeras filas y r primeras columnas hasta hacer una matriz cuadrada $r \times r$).

Proposición 5.3.1 *Sea Q una forma cuadrática definida sobre un espacio vectorial real, y $\mathcal{A}$ su matriz en una base $\mathcal{B}$. Entonces, Q es definida positiva si y solo si los menores principales de $\mathcal{A}$, $D_1, D_2, \ldots, D_n$ son todos mayores que cero.*

Proposición 5.3.2 *Sea Q una forma cuadrática definida sobre un espacio vectorial real, y $\mathcal{A}$ su matriz en una base $\mathcal{B}$. Entonces, Q es definida negativa si y solo si los menores principales de $\mathcal{A}$, verifican:*

$$D_1 < 0, D_2 > 0, \ldots, (-1)^n D_n > 0$$

No demostraremos estas propiedades.

5.4 Producto escalar

De entre todas las formas bilineales simétricas, las definidas positivas presentan unas propiedades particulares que las hacen apropiadas para las aplicaciones en Física (junto con las pseudodefinidas Lorentzianas).

5.4.1 Producto escalar en un espacio real

Definición 5.4.1 *Un producto escalar en un espacio vectorial real V es una forma bilineal simétrica definida positiva. Es decir:*

$$(\ ,\ ): V \times V \to \mathbb{R}$$

con las propiedades:
i) $(x, y) = (y, x), \quad x, y \in V$
ii) $(x + y, z) = (x, z) + (y, z), \quad (\lambda x, y) = \lambda(x, y), \quad x, y, z \in V,\ \lambda \in \mathbb{R}$
iii) $(x, x) \geq 0, \quad \forall x \in V, \quad (x, x) = 0 \Leftrightarrow x = 0$

5.4.2 Formas sesquilineales

Esta definición no puede extenderse a espacios complejos, pues el concepto de forma bilineal simétrica *definida positiva* no se puede establecer allí. Sin embargo, una aplicación similar a ésta puede definirse en espacios complejos, sustituyendo la propiedad de bilineal por otra.

Definición 5.4.2 *Sea V un espacio vectorial complejo. Una forma sesquilineal es una aplicación:*

$$\phi: V \times V \to \mathbb{C}$$

que verifica:
i) $\phi(x, y) = \overline{\phi(y, x)}, \quad x, y \in V$
ii) $\phi(x, y + z) = \phi(x, y) + \phi(x, z), \quad \phi(x, \lambda y) = \lambda\phi(x, y), \quad x, y, z \in V,\ \lambda \in \mathbb{C}$

Debido a la primera propiedad se tiene:

$$\phi(\lambda x, y) = \bar{\lambda}\phi(x, y),$$

es decir, la aplicación no es bilineal. Solo es lineal en la segunda variable, pero no en la primera (se trata de una convención, en otros lugares la definición se hace de modo que la aplicación es lineal en la primera variable).

La teoría de formas sesquilineales es muy similar a la de formas bilineales simétricas reales. Si el espacio es de dimensión finita, podemos escribir la aplicación en una base dada. Es fácil ver (siguiendo en todo la teoría de las formas bilineales), que la expresión es:

$$\phi(x, y) = X^+ \mathcal{A} Y$$

donde X, Y son los vectores de $\mathbb{C}^n$ que representan a $x, y \in V$ en esa base y X^+ representa la transpuesta conjugada de una matriz. Debido a la primera propiedad de las formas sesquilineales, la matriz $\mathcal{A}$ verifica:

$$\mathcal{A}^+ = \mathcal{A}$$

(las matrices con esta propiedad se llaman hermíticas). En efecto:

$$\overline{\phi(x, y)} = \overline{X^+ \mathcal{A} Y} = (X^+ \mathcal{A} Y)^+ = Y^+ \mathcal{A}^+ X = Y^+ \mathcal{A} X, \quad \forall X, Y \in \mathbb{C}^n$$

Al igual que en el caso real las matrices simétricas estaban asociadas a las formas bilineales reales, en el caso complejo las matrices hermíticas están asociadas a las formas sesquilineales. Si el espacio es real, una forma sesquilineal es simplemente una forma bilineal simétrica (al ser el conjugado de un número real igual a sí mismo).

Si se cambia la base, la matriz de una forma sesquilineal cambia. Como se puede comprobar fácilmente (siempre teniendo como guía las formas bilineales simétricas), si P es la matriz de cambio de base (es decir la que expresa los vectores de la segunda base en función de los de la primera) se tiene:

$$\mathcal{A}' = P^+ \mathcal{A} P$$

Lo que hace particularmente interesantes a las formas sesquilineales es que los valores que toman sobre la diagonal (es decir sobre los pares (x, x)), son reales (basta usar la primera propiedad y comprobar que $\phi(x, x)$ es igual a su complejo conjugado). Debido a esto, se puede establecer para las formas sesquilineales la propiedad de ser definida positiva (o negativa).

Definición 5.4.3 *Sea ϕ una forma sesquilineal sobre un espacio complejo. Se dice que ϕ es definida positiva si*

$$\phi(x,x) \geq 0,\ \forall x \in V, \quad \phi(x,x) = 0 \Leftrightarrow x = 0$$

5.4.3 Producto escalar complejo

Ahora podemos definir un producto escalar en un espacio vectorial complejo.

Definición 5.4.4 *Sea V un espacio vectorial complejo. Un producto escalar en V es una aplicación sesquilineal definida positiva. Es decir,*

$$(\ ,\ ) : V \times V \to \mathbb{C}$$

con las propiedades:
i) $(x,y) = \overline{(y,x)}, \quad x,y \in V$
ii) $(x+y,z) = (x,z) + (y,z), \quad (\lambda x, y) = \lambda(x,y), \quad x,y,z \in V,\ \lambda \in \mathbb{C}$
iii) $(x,x) \geq 0, \quad \forall x \in V, \quad (x,x) = 0 \Leftrightarrow x = 0$

El producto escalar real puede considerarse como un caso particular del complejo, por lo que en los siguientes apartados, nos referiremos de forma sistemática al caso complejo, considerando al real incluido en nuestras afirmaciones.

5.4.4 Norma en un espacio vectorial

Una norma en un espacio vectorial permite asignar a cada vector una *longitud.*

Definición 5.4.5 *Una norma en un espacio vectorial V es una aplicación:*

$$\|\cdot\| : V \to \mathbb{R}$$

que verifica:
a) $\|x\| \geq 0,\ \forall x \in V, \quad \|x\| = 0 \Leftrightarrow x = 0$
b) $\|\lambda x\| = |\lambda|\,\|x\|, \quad \lambda \in \mathbb{C},\ x \in V$
c) $\|x+y\| \leq \|x\| + \|y\|, \quad x,y \in V$

La tercera propiedad se conoce como desigualdad triangular. La definición es la misma en el caso real. En el caso complejo, se toma el módulo de λ, que es un número real, y en el caso real, el valor absoluto de λ (que es un número real).

Una norma no está relacionada, en principio, con un producto escalar. Sin embargo, dado un producto escalar siempre es posible definir una norma a partir de él. Es decir, un producto escalar nos permite definir la longitud de un vector. La norma asociada a un producto escalar se define como:

$$\|x\| = \sqrt{(x,x)}, \quad x \in V$$

Proposición 5.4.1 *Si $(\ ,\ )$ es un producto escalar, la aplicación $\|\cdot\|$ definida anteriormente, es una norma en V.*

Demostración. La propiedad a) de las normas es inmediata a consecuencia de ser definido positivo el producto escalar (como forma bilineal simétrica en el caso real o como forma sesquilineal en el caso complejo).

La propiedad b) es:

$$\|\lambda x\|^2 = (\lambda x, \lambda x) = \lambda\bar{\lambda}(x,x) = |\lambda|^2\|x\|^2$$

de donde:

$$\|\lambda x\| = |\lambda|\|x\|.$$

La tercera propiedad, la desigualdad triangular es algo más difícil de demostrar. Veremos que es una consecuencia de la desigualdad de Cauchy-Schwarz que demostramos a continuación.

Desigualdad de Cauchy-Schwarz:

$$|(x,y)| \leq \sqrt{(x,x)(y,y)}, \quad x,y \in V$$

Para demostrarlo, consideremos el producto escalar del vector $\lambda x - \mu y$ por sí mismo. Se obtiene un número real mayor o igual que cero:

$$(\lambda x - \mu y, \lambda x - \mu y) = |\lambda|^2(x,x) + |\mu|^2(y,y) - \bar{\lambda}\mu(x,y) - \lambda\bar{\mu}(y,x) \geq 0$$

Como la desigualdad es cierta para todo los escalares λ, μ, elegimos:

$$\lambda = (x,y), \quad \mu = (x,x)$$

$$|(x,y)|^2(x,x) + (x,x)^2(y,y) - 2(x,x)|(x,y)|^2 \geq 0$$

es decir:

$$|(x,y)|^2 \leq (x,x)(y,y)$$

lo que demuestra la desigualdad de Cauchy-Schwarz. Esta desigualdad es estricta si y sólo si los vectores x, y son linealmente independientes.

La desigualdad triangular es ahora inmediata:

$$\begin{aligned} \|x+y\|^2 &= (x+y, x+y) = (x,x) + (y,y) + (x,y) + \overline{(x,y)} = (x,x) + (y,y) + (x,y) + \overline{(x,y)} \\ &\leq \|x\|^2 + \|y\|^2 + 2\|x\|\,\|y\| \end{aligned}$$

de donde se deduce:

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

QED

5.4.5 Ortogonalidad

La relación de ortogonalidad definida para formas bilineales simétricas, se extiende a formas sesquilineales sin ningún problema.

Definición 5.4.6 *Sea V un espacio vectorial (real o complejo) y $(\ ,\ )$ un producto escalar en V. Se dice que dos vectores son ortogonales con respecto a este producto escalar si $(x,y) = 0$.*

Al ser un producto escalar, la matriz asociada en cualquier base es regular y el único vector que es ortogonal a todo el espacio es el vector 0.

Dada una base cualquiera en un espacio vectorial (de dimensión finita) con un producto escalar, es posible construir una base ortonormal (es decir, $(u_i, u_j) = \delta_{ij},\ i,j = 1, \ldots n$), utilizando, por ejemplo, el método de Gram-Schmidt (el hecho de ser una forma sesquilineal no afecta para nada al desarrollo. Simplemente hay que prestar atención a los complejos conjugados).

Sea $\mathcal{B} = \{u_1, \ldots, u_n\}$ una base ortonormal en un espacio V de dimensión finita dotado de un producto escalar. Los coeficientes de cualquier vector en esta base se pueden calcular fácilmente:

$$x \in V, \quad x = \sum_{i=1}^{n} x_i u_i$$

Haciendo el producto escalar de u_j por x se tiene:

$$(u_j, x) = (u_j, \sum_{i=1}^{n} x_i u_i) = \sum_{i=1}^{n} x_i (u_j, u_i) = \sum_{i=1}^{n} x_i \delta_{ji}$$

es decir:

$$x_i = (u_i, x), \quad \forall i = 1, \ldots, n$$

y por tanto:

$$x = \sum_{i=1}^{n} (u_i, x) u_i$$

En una base ortonormal, la matriz asociada a un producto escalar es la matriz identidad, es decir:

$$(x, y) = \sum_{i=1}^{n} \bar{x}_i y_i$$

En lo que se refiere a subespacios, se tiene:

Proposición 5.4.2 *Sea V un espacio vectorial de dimensión finita sobre $\mathbb{R}$ o $\mathbb{C}$. Sea W un subespacio de V y $W^\perp$ su ortogonal (definido de la forma ya establecida en el estudio de formas bilineales simétricas). Entonces:*

$$V = W \oplus W^\perp.$$

Se dice que $W^\perp$ es el complemento ortogonal de W. Las propiedades de ortogonalidad son fundamentales en la descripción de espacio dotados de un producto escalar. Nótese que en bases ortogonales todos los productos escalares son el mismo.

5.4.6 Proyección ortogonal

Sea V un espacio vectorial (real o complejo) con un producto escalar. Sea W un subespacio propio de V, y $W^\perp$ su complemento ortogonal. Todo vector de V se puede poner de manera única como la suma de dos vectores ortogonales entre sí:

$$x = y + z, \quad x \in V,\ y \in W,\ z \in W^\perp.$$

Debido a que y, z están definidos unívocamente por x podemos definir las siguientes aplicaciones:

$$\begin{array}{rccc} P_1\colon & V & \to & V \\ & x & \mapsto & y \end{array} \qquad \begin{array}{rccc} P_2\colon & V & \to & V \\ & x & \mapsto & z \end{array}.$$

Las aplicaciones P_1 y P_2, que son claramente lineales, se llaman las proyecciones ortogonales sobre W y $W^\perp$ respectivamente. Estas aplicaciones verifican las siguientes propiedades:

Proposición 5.4.3 *Si P_1 y P_2 son las proyecciones ortogonales sobre los espacios W y $W^\perp$ se tiene:*
a) $P_1 + P_2 = 1_V$
b) $P_1^2 = P_1, \quad P_2^2 = P_2, \quad P_1P_2 = P_2P_1 = 0$
c) $(P_1x, x') = (x, P_1y), \quad (P_2x, x') = (x, P_2x'), \quad x, x' \in V$

Demostración. Si $x = y + z$, de acuerdo con la descomposición $V = W \oplus W^\perp$:

$$y = P_1(x), \quad z = P_2(x)$$

y por tanto:

$$x = y + z = P_1(x) + P_2(x) = (P_1 + P_2)(x), \quad \forall x \in V$$

es decir la suma de los proyectores ortogonales es igual a la identidad en V.

Además, de $P_1(x) = y \in W$, se deduce:

$$P_1^2(x) = P_1(y) = y = P_1(x) \Rightarrow P_1^2 = P_1$$

De la misma manera se prueba para P_2

También es fácil de probar la otra propiedad:

$$P_1P_2 = P_1(1_V - P_1) = P_1 - P_1^2 = 0$$

y de igual forma $P_2P_1 = 0$.

Finalmente:

$$(P_1x, x') = (y, y' + z') = (y, y') = (y + z, P_1(x')) = (x, P_1(x'))$$

y de la misma forma para P_2. QED

Veamos ahora como se escribe la proyección ortogonal en una base ortonormal adaptada a la descomposición de subespacios. Sea V un espacio vectorial de dimensión finita n, y W un subespacio propio de dimensión r, Sea $\mathcal{B} = \{e_1, \dots, e_n\}$ una base ortonormal de V en la cual queremos calcular las matrices que representan a P_1 y P_2. Sea $\mathcal{B}' = \{u_1, \dots, u_n\}$ una base ortonormal de V, de forma que $\{u_1, \dots, u_r\}$ sea una base (también ortonormal) de W. El teorema de ampliación de la base permite obtener este resultado. Aunque en principio está establecido para bases no necesariamente ortonormales, el procedimiento de Gram-Schmidt nos asegura que la situación anterior es correcta. Como consecuencia, el resto de los vectores de la base $\mathcal{B}'$, es decir: $\{u_{r+1}, \dots, u_n\}$ son un base (también ortonormal) de $W^\perp$. Supongamos que los vectores de la base $\mathcal{B}'$ tienen como coordenadas en la base $\mathcal{B}$ los vectores columna de $\mathbb{C}^n$ (o $\mathbb{R}^n$ si el espacio es real):

$$U_1, \dots, U_n \in \mathbb{C}^n$$

Sea x un vector de V y $X \in \mathbb{C}^n$ sus coordenadas en la base $\mathcal{B}$. Entonces:

$$y = P_1(x) = \sum_{i=1}^r (u_i, x)u_i = \sum_{i=1}^r (U_i^+ X)u_i$$

es decir, en coordenadas:

$$Y = \sum_{i=1}^r (U_i^+ X)U_i = (\sum_{i=1}^r U_iU_i^+)X$$

y por tanto, la matriz asociada a P_1 es:

$$\sum_{i=1}^r U_iU_i^+$$

Téngase en cuenta que ambas bases son ortonormales para poder deducir este resultado. La matriz que representa a P_2 en esta misma base es:

$$\sum_{i=r+1}^n U_iU_i^+$$

y se tiene:

$$\sum_{i=1}^n U_iU_i^+ = I_n$$

siendo I_n la matriz identidad en dimensión n.

Un resultado de gran utilidad en muchos campos (en particular en espacios de dimensión infinita) es el teorema de la proyección ortogonal. La pregunta que se puede uno plantear es: dado un vector de un espacio vectorial con un producto escalar cuál es el vector de entre todos los de un subespacio que mejor aproxima a este vector. La respuesta es que es la proyección ortogonal del vector sobre el subespacio.

Teorema 5.4.1 *Sea V un espacio vectorial de dimensión finita y $x \in V$. Sea W un subespacio de W. La norma del vector $x - w$, donde $w \in W$, toma su valor mínimo cuando w es la proyección ortogonal de x sobre W.*

Demostración. De acuerdo con el resultado sobre descomposición ortogonal de V referida al subespacio W, $x = y + z$ donde $y \in W$, $z \in W^\perp$. Entonces

$$\|x - w\| = \|y + z - w\| = \|y - w\| + \|z\|$$

pues $y - w \in W$, $z \in W^\perp$. Esta expresión es mínima cuando el primer sumando es cero: $\|y - w\| = 0$, y por lo tanto:

$$w = P_W(x).$$

QED

5.4.7 La propiedad del paralelogramo

Hemos visto anteriormente que todo producto escalar da lugar a una norma asociada a él. La pregunta que surge es si toda norma deriva de un producto escalar. La respuesta es no. Existe una propiedad sencilla que caracteriza a las normas que provienen de un producto escalar.

Teorema 5.4.2 *1) Sea $\|\cdot\|$ una norma que proviene de un producto escalar, es decir:*

$$\|x\| = \sqrt{(x,x)}$$

Entonces:

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

2) Sea $\|\cdot\|$ una norma que verifica:

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

Entonces, $\|\cdot\|$ deriva de un producto escalar, que se escribe como:

$$(x,y) = \frac{1}{4}\left(\|x+y\|^2 - \|x-y\|^2 + i(\|x-iy\|^2 - \|x+iy\|^2)\right)$$

La demostración de la primera parte es inmediata. Basta escribir la definición de la norma en términos del producto escalar. La demostración de la segunda parte es más complicada y necesita argumentos de continuidad en números reales. No la haremos aquí.

5.4.8 El teorema de Riesz-Fréchet

Como hemos tenido ocasión de estudiar, no existe un isomorfismo canónico entre un espacio (de dimensión finita) y su dual. Sin embargo, si el espacio tiene un producto escalar, podemos establecer una correspondencia (que es un isomorfismo cuando el cuerpo es $\mathbb{R}$) entre ambos espacios usando este producto escalar asignando a cada forma lineal un vector del espacio original.

Sea V un espacio vectorial (sobre $\mathbb{K} = \mathbb{R}$ o $\mathbb{K} = \mathbb{C}$) de dimensión finita, con un producto escalar. La aplicación:

$$\begin{array}{rccc} \omega_x\colon & V & \to \mathbb{K} & \\ & y & \mapsto & (x,y) \end{array}$$

es una aplicación lineal (con reales o complejos), es decir un elemento del dual, $\omega_x \in V^*$. El teorema de Riesz-Fréchet asegura que el resultado inverso es también cierto.

Teorema 5.4.3 *Dada una forma $\omega \in V^*$ existe un único vector $x_\omega \in V$ tal que:*

$$\omega(y) = (x_\omega, y), \quad \forall y \in V$$

Demostración. Sea $\{u_1, \ldots, u_n\}$ una base ortonormal de V. Entonces:

$$\omega(y) = \omega(\sum_{i=1}^n (u_i,y)u_i) = \sum_{i=1}^n (u_i,y)\omega(u_i) = \sum_{i=1}^n (\overline{\omega(u_i)}u_i, y)$$

y por lo tanto, el vector:

$$x_\omega = \sum_{i=1}^n \overline{\omega(u_i)}u_i$$

verifica el teorema. Veamos que es único, aunque en la expresión anterior parezca depender de la base elegida. Supongamos que

$$\omega(y) = (x,y) = (x',y), \quad \forall y \in V$$

Entonces:

$$(x-x',y) = 0, \quad \forall y \in V$$

Pero el único vector ortogonal a todo el espacio es el vector 0, por tanto, $x = x'$. QED

La correspondencia:

$$\begin{array}{rccc} \psi\colon & V & \to & V \\ & x & \mapsto & \omega_x \end{array}$$

es un isomorfismo de espacios vectoriales reales:

$$\psi(x+y)(z) = (z, x+y) = (z, x) + (z, y) = \psi(x)(z) + \psi(y)(z)$$

Además:

$$\psi(\lambda x)(z) = (z, \lambda x) = \lambda(z, x) = \lambda\psi(x)(z)$$

En espacios complejos aparece un conjugado.

Tema 6

Operadores en espacios con producto escalar

Operadores en espacios complejos. Operador adjunto. Operadores autoadjuntos y unitarios. Proyectores ortogonales. Teorema espectral. Operadores en espacios reales. Operadores simétricos y ortogonales.

Al igual que cuando hablamos de diagonalización de endomorfismos, haremos aquí una distinción entre el caso real y complejo. Como veremos, el caso complejo es más simple que el real, y muchos de los resultados que obtengamos en este caso serán aplicables en el real.

6.1 Operadores en espacios complejos con producto escalar

En toda esta sección V será un espacio vectorial complejo de dimensión finita, dotado de un producto escalar.

6.1.1 El operador adjunto

Llamaremos operadores (lineales) a los endomorfismos de V. Para cada operador en V introducimos un operador asociado.

Definición 6.1.1 *Sea $A\colon V \to V$ un operador. El operador adjunto se define como un operador $A^+\colon V \to V$ que verifica:*

$$(x, Ay) = (A^+x, y), \forall x, y \in V.$$

Veamos que tal operador existe. Para cada $x \in V$, definimos la siguiente forma lineal:

$$V \quad \to \quad \mathbb{C}y \quad \to \quad (x, Ay)$$

Por el teorema de Riesz-Fréchet, existe un único vector $z \in V$ tal que:

$$(x, Ay) = (z, y), \quad \forall y \in V$$

La correspondencia $x \mapsto z$ de V en V es lineal. Sean $x, x' \in V$ y consideremos la forma lineal:

$$y \mapsto (x + x', Ay) = (x, Ay) + (x', Ay)$$

Existe un único $\tilde{z} \in V$ tal que:

$$(x + x', Ay) = (\tilde{z}, y), \quad \forall y \in V$$

es decir:

$$(x, Ay) + (x', Ay) = (z, y) + (z', y) = (z + z', y)$$

de donde:

$$\tilde{z} = z + z'$$

En cuanto al producto por escalares:
Sean $x \in V$, $\lambda \in \mathbb{C}$ y consideremos:

$$y \mapsto (\lambda x, Ay) = \bar{\lambda}(x, Ay)$$

Existe un único $\tilde{z} \in V$ tal que:

$$(\lambda x, Ay) = (\tilde{z}, y), \quad \forall y \in V$$

y por tanto:

$$\bar{\lambda}(x, Ay) = \bar{\lambda}(z, y) = (\lambda z, y)$$

de donde:

$$\tilde{z} = \lambda z$$

La operación de tomar adjuntos (pasar de A a A^+) tiene las siguientes propiedades, que se pueden demostrar fácilmente:
1) $(A^+)^+ = A$
2) $(A + B)^+ = A^+ + B^+$
3) $(\lambda A)^+ = \bar{\lambda} A^+$
4) $(AB)^+ = B^+ A^+$
Por ejemplo, la propiedad 1):

$$(x, Ay) = (A^+ x, y) = \overline{(y, A^+ x)} = \overline{((A^+)^+ y, x)} = (x, (A^+)^+ y)$$

relación que debe ser cierta para todo $x, y \in V$, lo que demuestra 1). Las propiedades 2) y 3) son inmediatas. En cuanto a 4):

$$(x, ABy) = ((AB)^+ x, y) = (A^+ x, By) = (B^+ A^+ x, y)$$

6.1.2 Representación matricial del operador adjunto

Veamos como obtener la representación matricial del operador adjunto a partir de la del operador original. Sea V un espacio vectorial complejo de dimensión finita dotado de un producto escalar y $\mathcal{B} = \{u_1, \ldots, u_n\}$ una base ortonormal de V. Sea $\mathcal{A}$ la matriz de A en la base $\mathcal{B}$, es decir, $\mathcal{A} = (a_{ij})$:

$$Au_i = \sum_{i=1}^{n} a_{ji} u_j, \quad i = 1, \ldots, n$$

Sea $\mathcal{A}'$ la matriz del operador adjunto, $\mathcal{A}' = (a'_{ij})$. Se tiene:

$$(x, Ay) = (A^+ x, y), \quad \forall x, y \in V$$

En particular:

$$(u_i, Au_j) = (A^+ u_i, u_j), \quad \forall i, j = 1, \ldots, n$$

y sustituyendo las expresiones de Au_j y $A^+ u_i$:

$$(u_i, \sum_{k=1}^{n} a_{kj} u_k) = (\sum_{k=1}^{n} a'_{ki} u_k, u_j)$$

$$\sum_{k=1}^{n} a_{kj} (u_i, u_k) = \sum_{k=1}^{n} \bar{a}'_{ki} (u_k, u_j)$$

$$\sum_{k=1}^{n} a_{kj} \delta_{ik} = \sum_{k=1}^{n} \bar{a}'_{ki} \delta_{kj}$$

$$a_{ij} = \bar{a}'_{ji}$$

es decir:

$$\mathcal{A}' = \mathcal{A}^+$$

la matriz del operador adjunto es la matriz transpuesta conjugada (la matriz adjunta) del operador de partida (el símbolo $^+$ significará indistintamente operador adjunto o matriz transpuesta conjugada, dependiendo a quién esté aplicado). Nótese que este resultado es solo cierto cuando la base en la que están escritos los operadores es ortonormal. Si no es así, la relación es más complicada y hace intervenir la matriz del producto escalar. En estas bases que no son ortonormales, la matriz de A^+ no es $\mathcal{A}^+$, lo que puede inducir a cierta confusión si no se presta atención.

Recordando como se calculaban las coordenadas de un vector en una base ortonormal, podemos encontrar una expresión de los elementos de matriz de un operador en bases de este tipo. En efecto,

$$(u_i, Au_j) = (u_i, \sum_{k=1}^{n} a_{kj} u_k) = \sum_{k=1}^{n} a_{kj}(u_i, u_k) = \sum_{k=1}^{n} a_{kj}\delta_{ik} = a_{ij}$$

es decir:

$$a_{ij} = (u_i, Au_j)$$

6.1.3 Operadores normales, autoadjuntos y unitarios

Teniendo en cuenta las relaciones entre A y A^+ se pueden definir clases especiales de operadores. Sea $(V, (\ ,\ ))$ un espacio vectorial complejo con producto escalar, y A un operador en V.

Definición 6.1.2 *Se dice que el operador A es normal si conmuta con su adjunto:*

$$AA^+ = A^+A$$

Definición 6.1.3 *Se dice que el operador A es autoadjunto si coincide con su adjunto:*

$$A^+ = A$$

Definición 6.1.4 *Se dice que el operador A es unitario si:*

$$AA^+ = A^+A = 1_V$$

Los operadores autoadjuntos verifican:

$$(x, Ay) = (Ax, y),$$

y en bases ortonormales vienen representados por matrices hermíticas (o autoadjuntas):

$$\mathcal{A}^+ = \mathcal{A}.$$

Los operadores unitarios verifican:

$$(Ax, Ay) = (x, y),$$

y en bases ortonormales, sus matrices son unitarias, es decir:

$$\mathcal{A}\mathcal{A}^+ = \mathcal{A}^+\mathcal{A} = I_n.$$

Es inmediato comprobar que los operadores autoadjuntos y unitarios son normales. Existen operadores normales que no son autoadjuntos ni unitarios. Nuestro interés se centra en los operadores autoadjuntos y unitarios. Sin embargo, es más conveniente, y no implica ningún esfuerzo adicional, estudiar los operadores normales y luego restringirnos a estos dos casos.

6.1.4 Teorema espectral para operadores normales

Nuestro objetivo es probar que los operadores normales son diagonalizables, es decir existe una base del espacio formada por autovectores, y además esta base es ortonormal. Para ello demostraremos unos lemas previos que se verifican para operadores más generales.

Proposición 6.1.1 *Sea V un espacio vectorial complejo de dimensión finita y A, B dos endomorfismos de V, tales que $AB = BA$. Entonces, existe un vector no nulo $y \in V$ que es autovector de A y B.*

Demostración. Al ser V un espacio complejo de dimensión finita, el polinomio característico del endomorfismo A tiene al menos una raíz (teorema fundamental del álgebra). Es decir, existe al menos un autovector de A, $x \in V$:

$$Ax = \lambda x, \quad x \in V, \ x \neq 0, \ \lambda \in \mathbb{C}$$

Como A y B conmutan, los vectores $Bx, B^2x, \ldots$ son también autovectores de A con el mismo autovalor:

$$AB^kx = B^kAx = \lambda B^kx, \quad k = 0, 1, 2, \ldots$$

Consideremos la sucesión de autovectores $x, Bx, B^2x, \ldots$. No pueden ser todos linealmente independientes, pues el espacio es de dimensión finita (podría ocurrir que hubiera $n = \dim V$ vectores l.i. Entonces el vector x se dice que es un vector cíclico para A. La teoría de vectores cíclicos es muy importante sobre todo en el caso de espacios de dimensión infinita, pero no la trataremos aquí). Supongamos pues que $B^{r+1}x$ depende linealmente de los anteriores. El subespacio que generan $\{x, Bx, \ldots, B^rx\}$ es un subespacio invariante bajo B, formado por autovectores de A de autovalor λ. Restringiendo B a este subespacio, vemos que existe en él un autovalor de B, que por lo anterior también lo será de A. QED

La siguiente proposición trata con espacios con producto escalar.

Proposición 6.1.2 *Sea V un espacio complejo de dimensión finita, con producto escalar. Sea A un operador en V. Entonces, si S es un subespacio de V invariante bajo A ($AS \subset S$), el subespacio ortogonal $V^\perp$ es invariante bajo A^+:*

$$A^+(S^\perp) \subset S^\perp$$

Demostración. Sea $y \in S^\perp$. Por la definición de operador adjunto:

$$(x, A^+y) = (Ax, y)$$

Si $x \in S$, al ser S invariante, $Ax \in S$, luego:

$$(x, A^+y) = (Ax, y) = 0$$

de donde $A^+y \in S^\perp$. QED

Enunciemos ahora el teorema espectral:

Teorema 6.1.1 *Sea V un espacio vectorial complejo de dimensión finita dotado de un producto escalar. Sea A un operador normal en V. Entonces, existe una base ortonormal de V formada por autovectores de A.*

Demostración. Al ser A normal, A conmuta con su adjunto, luego por la primera de las dos proposiciones demostradas previamente, A y A^+ tienen un autovector común:

$$Ax = \lambda_1 x, \quad A^+x = \mu x$$

Como $(x, Ay) = (Ax, y)$, se tiene:

$$(x, Ax) = (x, \lambda_1 x) = \lambda_1(x, x) = (A^+x, x) = (\mu x, x) = \bar{\mu}(x, x),$$

es decir:

$$\mu = \bar{\lambda}_1.$$

Sea

$$u_1 = \frac{x}{||x||}$$

un autovector de norma 1 (vector unitario) de A y A^+. El subespacio $S_1 = \text{lin}\{u_1\}$ es invariante bajo A y bajo A^+. Por lo tanto, haciendo uso de la segunda proposición, $S_1^\perp$ es invariante bajo A (y bajo A^+). Consideramos la restricción de A a $S_1^\perp$. Buscamos allí un autovector común a A y A^+ (supongamos que de norma 1):

$$Au_2 = \lambda_2 u_2, \quad A^+ u_2 = \bar{\lambda}_2 u_2$$

Además:

$$(u_1, u_2) = 0$$

Se construye el subespacio $S_2 = \text{lin}\{u_1, u_2\}$ y se continua el proceso, que debe acabar al ser el espacio de dimensión finita. De esta forma se obtiene una base formada por autovectores de A que son ortonormales. Nótese que es también una base de autovectores de A^+. QED

Estudiaremos ahora una serie de resultados relacionados con este teorema espectral. Lo primero que demostraremos es que la existencia de bases ortonormales caracteriza a los operadores normales.

Proposición 6.1.3 *Sea V un espacio vectorial complejo de dimensión finita dotado de producto escalar. Sea A un operador en V y sea $\mathcal{B} = \{u_1, \ldots, u_n\}$ una base ortonormal de V formada por autovectores de A. Entonces A es un operador normal.*

Demostración. Calculemos las coordenadas de la imagen de u_i mediante el operador A^+ en la base $\mathcal{B}$:

$$A^+ u_i = \sum_{j=1}^n (u_j, A^+ u_i) u_j = \sum_{j=1}^n (Au_j, u_i) u_j = \sum_{j=1}^n (\lambda_j u_j, u_i) u_j = \sum_{j=1}^n \bar{\lambda}_j (u_j, u_i) u_j = \sum_{j=1}^n \bar{\lambda}_j \delta_{ji} u_j = \bar{\lambda}_i u_i$$

luego u_i es también un autovector de A^+ con autovalor $\bar{\lambda}_i$. Entonces:

$$AA^+ u_i = |\lambda_i|^2 u_i = A^+ A u_i, \quad i = 1, \ldots, n$$

luego $AA^+ = A^+ A$ al coincidir en una base. QED

El resultado sobre la existencia de una base ortonormal formada por autovectores se puede enunciar en términos de matrices.

Proposición 6.1.4 *Sea $\mathcal{A}$ una matriz compleja $n \times n$, normal, es decir:*

$$\mathcal{A}\mathcal{A}^+ = \mathcal{A}^+ \mathcal{A}$$

Entonces, existe una matriz unitaria $\mathcal{U}$ tal que:

$$\mathcal{U}^+ \mathcal{A} \mathcal{U}$$

es una matriz diagonal

Demostración. Se considera a $\mathcal{A}$ como la matriz de un cierto operador en una base ortonormal $\{e_1, \ldots, e_n\}$. Entonces A^+ viene representado en esa base por la matriz $\mathcal{A}^+$. Por tanto, por las hipótesis del teorema, A es un operador normal. Al existir una base ortonormal formada por autovectores de A, $\{u_1, \ldots, u_n\}$ la matriz de cambio de base es:

$$u_i = \sum_{j=1}^n \mathcal{U}_{ji} e_i$$

Por tanto, la matriz de A en la base $\{u_i\}$ se obtiene de la matriz $\mathcal{A}$ mediante la expresión:

$$\mathcal{U}^{-1} \mathcal{A} \mathcal{U}$$

y es obviamente diagonal, al estar formada la nueva base por autovectores de A.

La matriz de cambio de base es unitaria. Esto es cierto siempre que se pasa de una base ortonormal a otra:

$$\mathcal{U}\mathcal{U}^+ = \mathcal{U}^+\mathcal{U} = I_n$$

En efecto:

$$\delta_{ij} = (u_i, u_j) = (\sum_{k=1}^n \mathcal{U}_{ik} e_k \sum_{l=1}^n \mathcal{U}_{jl} e_l) = \sum_{k,l=1}^n \bar{\mathcal{U}}_{ik}\mathcal{U}_{jl}(e_k, e_l) = \sum_{k,l=1}^n \bar{\mathcal{U}}_{ik}\mathcal{U}_{jl}\delta_{kl} = \sum_{k=1}^n \bar{\mathcal{U}}_{ik}\mathcal{U}_{jk}$$

es decir, la matriz $\mathcal{U}$ es unitaria, y:

$$\mathcal{U}^+\mathcal{A}\mathcal{U}$$

es una matriz diagonal. QED

El teorema espectral se puede establecer diciendo que todo operador normal se puede llevar a una forma diagonal mediante una transformación unitaria.

6.1.5 Teorema espectral para operadores autoadjuntos

Puesto que un operador autoadjunto es normal, los teoremas anteriores se aplican a este tipo de operadores. Sin embargo, podemos decir algo más de ellos.

Proposición 6.1.5 *Los autovalores de un operador autoadjunto son reales.*

Demostración.
Sea λ un autovalor de A con autovector x. Entonces:

$$(Ax, x) = \bar{\lambda}(x, x) = (x, Ax) = \lambda(x, x)$$

de donde

$$\bar{\lambda} = \lambda.$$

QED

Podemos establecer el siguiente teorema espectral:

Teorema 6.1.2 *Sea V un espacio vectorial complejo de dimensión finita con producto escalar. Sea A un operador autoadjunto en V. Entonces, los autovalores de A son reales y existe una base ortonormal de V formada por autovectores de A.*

El resultado inverso es:

Teorema 6.1.3 *Sea V un espacio vectorial complejo de dimensión finita con producto escalar. Sea A un operador en V con autovalores reales tal que existe una base ortonormal de V formada por autovectores de A. Entonces A es autoadjunto.*

Demostración.
De acuerdo con el teorema espectral para operadores normales, al existir una base ortonormal de V formada por autovectores de A, A es normal. Al tener los autovalores reales:

$$A^+u_k = \bar{\lambda}_k u_k = \lambda_k u_k = Au_k, \quad k = 1, \ldots, n$$

luego $A = A^+$ y A es autoadjunto. QED

El resultado para matrices es:

Proposición 6.1.6 *Sea $\mathcal{A}$ una matriz hermítica (autoadjunta). Entonces $\mathcal{A}$ es diagonalizable mediante una matriz unitaria, y la matriz diagonal es real.*

La demostración sigue las líneas del caso normal.

6.1.6 Teorema espectral para operadores unitarios

Un operador unitario es también normal. Sus autovalores tienen módulo unidad.

Proposición 6.1.7 *Los autovalores de un operador unitario tienen módulo 1.*

Demostración. Sea λ autovalor de A con autovector x. Entonces:

$$Ax = \lambda x \Rightarrow A^+Ax = \lambda A^+x = |\lambda|^2 x$$

y por tanto:

$$|\lambda| = 1$$

QED

Teorema 6.1.4 *Sea V un espacio vectorial complejo de dimensión finita con producto escalar. Sea A un operador unitario en V. Entonces, los autovalores de A tienen módulo igual a 1 y existe una base ortonormal de V formada por autovectores de A.*

El resultado inverso es:

Teorema 6.1.5 *Sea V un espacio vectorial complejo de dimensión finita con producto escalar. Sea A un operador en V con autovalores de módulo unidad tal que existe una base ortonormal de V formada por autovectores de A. Entonces A es unitario.*

Demostración. El operador A es normal. Al tener los autovalores de módulo igual a 1:

$$A^+Au_k = \bar{\lambda}_k\lambda_k u_k = |\lambda_k|^2 u_k = u_k, \quad k = 1, \dots, n$$

luego $AA^+ = A^+A = 1_V$ yA es unitario. QED

El resultado para matrices es:

Proposición 6.1.8 *Sea $\mathcal{A}$ una matriz unitaria. Entonces $\mathcal{A}$ es diagonalizable mediante una matriz unitaria, y la matriz diagonal tiene elementos de módulo 1 en la diagonal.*

Los operadores unitarios relacionan bases ortonormales entre sí. Conservan longitudes y ángulos.

De la relación $\mathcal{U}\mathcal{U}^+ = I_n$ se deduce que el determinante de una matriz unitaria (y por lo tanto de un operador unitario) tiene módulo igual a 1:

$$1 = \det(\mathcal{U}\mathcal{U}^+) = \det\mathcal{U}\det\mathcal{U}^+ = |\det\mathcal{U}|^2$$

pues $\det\mathcal{U}^t = \det\mathcal{U}$ y $\det\bar{\mathcal{U}} = \overline{\det\mathcal{U}}$.

Los operadores unitarios forman un grupo respecto a la composición de operadores. Se le llama el grupo unitario ($U(n)$). El subconjunto de operadores unitarios con determinante igual a 1 es un subgrupo de éste, ($SU(n)$). Por ejemplo, $U(1)$ son los números complejos de módulo 1 (es decir, la circunferencia unidad).

6.2 Proyectores ortogonales

Sea V un espacio vectorial complejo de dimensión finita con producto escalar. Sea A un operador normal en V, con espectro

$$\sigma(A) = \{\lambda_1, \dots, \lambda_r\}, \quad \lambda_i \neq \lambda_j,\ i \neq j,\ i, j = 1, \dots, r,\ r \leq n$$

Los subespacios invariantes:

$$V_i = \ker(A - \lambda_i 1_V), \quad i = 1, \dots, r$$

son ortogonales entre sí:

$$V_i \perp V_j, \quad i \neq j$$

como consecuencia del teorema espectral, pues corresponden a autovalores distintos. Además su suma es el espacio total:

$$V = V_1 \oplus \cdots \oplus V_r.$$

De esta forma podemos definir los proyectores sobre cada uno de los subespacios invariantes, generalizando los proyectores ortogonales que estudiamos en la descomposición de V en suma de un subespacio y su ortogonal:

$$x \in V, \quad x = x_1 + \cdots + x_r$$

$$\begin{array}{rcl} P_i\colon & \to & V \\ x & \mapsto & x_i \end{array}$$

La familia de proyectores P_i verifica una serie de propiedades que se conocen como el teorema de descomposición espectral.

Teorema 6.2.1 *Los proyectores ortogonales $P_1, \ldots, P_r$ asociados a un operador normal A en un espacio vectorial complejo V de dimensión finita con producto escalar, verifican las siguientes propiedades:*
1) $P_i^+ = P_i$
2) $P_i P_j = \delta_{ij} P_i$
3) $P_1 + \cdots + P_r = 1_V$
4) $\lambda_1 P_1 + \cdots + \lambda_r P_r = A$

Demostración. Los proyectores ortogonales son idempotentes y autoadjuntos:

$$P_i^2 x = P_i x_i = x_i$$

$$(x, P_i y) = (x, y_i) = (x_i, y_i) = (x_i, y) = (P_i x, y)$$

Además:

$$P_i P_j(x) = P_i(x_j) = 0, \quad i \neq j$$

En cuanto a su suma:

$$(P_1 + \cdots + P_r)x = P_1 x + \cdots + P_r x = x_1 + \cdots + x_r = x, \quad \forall x \in V$$

luego es la identidad en V.

Finalmente:

$$(\lambda_1 P_1 + \cdots + \lambda_r P_r)x = \lambda_1 P_1 x + \cdots + \lambda_r P_r x =$$

$$= \lambda_1 x_1 + \cdots + \lambda_r x_r = A x_1 + \cdots + A x_r = A(x_1 + \cdots + x_r) = Ax$$

QED

La expresión:

$$A = \lambda_1 P_1 + \cdots + \lambda_r P_r$$

se conoce como la descomposición espectral del operador A.

La descomposición espectral de un operador permite caracterizarlo de la forma siguiente.

Proposición 6.2.1 *Sea V un espacio vectorial complejo de dimensión finita dotado de un producto escalar, y A un operador en V. Supongamos que existe una familia de proyectores ortogonales (idempotentes y autoadjuntos), $\{P_1, \ldots, P_r\}$ y un conjunto de escalares (distintos), $\{\lambda_1, \ldots, \lambda_r\}$ que verifican:*
1) $P_i P_j = 0,\ i \neq j$
2) $P_1 + \cdots + P_r = 1_V$
3) $\lambda_1 P_1 + \cdots + \lambda_r P_r = A$
entonces, A es normal.

Demostración.

$$AA^+ = (\sum_{i=1}^r \lambda_i P_i)(\sum_{j=1}^r \bar{\lambda}_j P_j^+) = \sum_{i,j=1}^r \lambda_i \bar{\lambda}_j P_i P_j = \sum_{i,j=1}^r \lambda_i \bar{\lambda}_j \delta_{ij} P_i = \sum_{i=1}^r |\lambda_i|^2 P_i = A^+ A$$

Los escalares λ_i son los autovalores de A, y los autovectores son de la forma $P_i x$ con $x \in V$. En efecto:

$$AP_i(x) = \sum_{j=1}^r \lambda_j P_j P_i x = \sum_{j=1}^r \lambda_j \delta_{ij} P_i x = \lambda_i P_i x$$

Veamos que no hay otros autovalores. Sea $\lambda \in \mathbb{C}$ tal que existe $x \in V$ distinto de cero y:

$$Ax = \lambda x$$

$$Ax = \sum_{i=1}^r \lambda_i P_i x = \lambda \sum_{i=1}^r P_i(x)$$

es decir:

$$\sum_{i=1}^r (\lambda_i - \lambda) P_i x = 0$$

Aplicando P_k:

$$\sum_{i=1}^r (\lambda_i - \lambda) P_k P_i x = \sum_{i=1}^r (\lambda_i - \lambda) \delta_{ik} P_i x = (\lambda_k - \lambda) P_k x = 0$$

Por tanto, $P_k x = 0$ o bien $\lambda = \lambda_k$. En el segundo caso, λ es uno de los escalares que teníamos. En el primero, si la relación es cierta para todo $k = 1, \ldots, r$, entonces $x = 0$. QED

Los operadores autoadjuntos y unitarios se caracterizan de forma similar:

Proposición 6.2.2 *En las condiciones de la proposición anterior, si los escalares $\{\lambda_1, \ldots, \lambda_r\}$ son reales, A es autoadjunto.*

Demostración. De acuerdo con la primera proposición A es normal. Al tener todos sus autovalores reales es autoadjunto. QED

Proposición 6.2.3 *Si los escalares $\{\lambda_1, \ldots, \lambda_r\}$ son de módulo 1, A es unitario.*

6.2.1 Cálculo de proyectores ortogonales

Ya hemos estudiado como calcular un proyector conociendo una base ortonormal del subespacio sobre el que proyecta. Si $\{e_1, \ldots, e_n\}$ es una base ortonormal del espacio V y $\{u_1, \ldots, u_n\}$ es una base ortonormal de autovectores del operador A, los espacios V_i, $i = 1, \ldots, n$ estarán generados por los vectores:

$$V_1 = \operatorname{lin}\{u_1, \ldots, u_{n_1}\} \quad (6.1)$$

$$V_2 = \operatorname{lin}\{u_{n_1+1}, \ldots, u_{n_1+n_2}\} \quad (6.2)$$

$$\ldots \quad (6.3)$$

$$V_r = \operatorname{lin}\{u_{n_1+\cdots+n_{r-1}+1}, \ldots, u_n\} \quad (6.4)$$

donde n_i es la dimensión del espacio V_i. Por tanto, el proyector P_i sobre el subespacio

$$V_i = \operatorname{lin}\{u_{n_1+\cdots+n_{i-1}+1}, \ldots, u_{n_1+\cdots+n_{i-1}+n_i}\},$$

se puede calcular como:

$$P_i = \sum_{k=n_1+\cdots+n_{i-1}+1}^{n_1+\cdots+n_i} U_k U_k^+,$$

donde U_l son las coordenadas del vector u_l en la base $\{e_i\}$. Nótese que la acción de un proyector es:

$$P_i(x) = \sum_{k=n_1+\cdots+n_{i-1}+1}^{n_1+\cdots+n_i} (u_k, x)u_k.$$

En esta notación se tiene:

$$I_n = \sum_{i=1}^{n} U_i U_i^+,$$

la descomposición espectral de la identidad, y

$$\mathcal{A} = \sum_{i=1}^{n} \lambda_i U_i U_i^+,$$

la descomposición espectral del operador A (o de su matriz $\mathcal{A}$ en la base $\{e_i\}$).

Existen otras formas de calcular proyectores ortogonales. Veremos una de ellas que usa los polinomios interpoladores de Lagrange.

Proposición 6.2.4 *Sea V un espacio vectorial complejo de dimensión finita con un producto escalar. Sea A un operador normal en V y $A = \sum_{i=1}^{n} \lambda_i P_i$ su descomposición espectral. Entonces, los proyectores ortogonales P_i son iguales a:*

$$P_i = \varphi_i(A), \quad i = 1, \ldots, r$$

donde los polinomios $p(\lambda)$ vienen definidos por:

$$\varphi_i(\lambda) = \frac{(\lambda - \lambda_1)\cdots(\lambda - \lambda_{i-1})(\lambda - \lambda_{i+1})\cdots(\lambda - \lambda_r)}{(\lambda_i - \lambda_1)\cdots(\lambda_i - \lambda_{i-1})(\lambda_i - \lambda_{i+1})\cdots(\lambda_i - \lambda_r)}, \quad i = 1, \ldots, r$$

Demostración. Se tiene:

$$\varphi_i(\lambda) = \prod_{j\neq i} \frac{\lambda - \lambda_j}{\lambda_i - \lambda_j}, \quad \varphi_i(A) = \prod_{j\neq i} \frac{A - \lambda_j 1_V}{\lambda_i - \lambda_j}$$

Pero no es difícil calcular el valor de estos polinomio sobre A:

$$\varphi_i(A) = \varphi_i(\sum_{j=1}^{n} \lambda_j P_j) = \sum_{j=1}^{n} \varphi_i(\lambda_j) P_j$$

debido a las propiedades de los proyectores ortogonales.

Como:

$$\varphi_i(\lambda_j) = \delta_{ij}$$

se obtiene el resultado buscado. QED

De la demostración anterior se deduce que para cualquier polinomio $p(\lambda)$, el valor que toma sobre el operador A es:

$$p(A) = \sum_{j=1}^{n} p(\lambda_j) P_j$$

Se puede extender a funciones analíticas (es decir, que admiten un desarrollo en serie que es convergente en un entorno del punto considerado):

$$f(A) = \sum_{j=1}^{n} f(\lambda_j) P_j$$

lo que permite calcular valores de funciones sobre operadores (normales).

6.3 Operadores en espacios vectoriales reales con producto escalar

Las mismas cuestiones que se suscitaron en relación con los espacios complejos dotados de un producto escalar serán estudiadas aquí. Como veremos, las dificultades principales provienen del hecho que no todo operador en un espacio real posee autovalores (se entiende reales). La extensión del cuerpo base (noción que se puede definir rigurosamente) a los números complejos, permitiría aligerar esta sección. Sin embargo nos mantendremos en el campo real en todo lo posible.

6.3.1 El operador transpuesto

En analogía con el operador adjunto, definiremos aquí el operador transpuesto. Es este un nombre ya usado en relación con el espacio dual. De hecho, utilizando el teorema de Riesz-Fréchet, ambos conceptos coinciden. Sin embargo, para evitar problemas de interpretación el operador transpuesto se entenderá en la forma que sigue.

Definición 6.3.1 *Sea V un espacio vectorial real con producto escalar, y A un operador en V. Se define el operador transpuesto de A, A^t como el único operador que verifica:*

$$(x, Ay) = (A^t x, y)$$

Para demostrar la existencia y unicidad de este operador, basta aplicar el teorema de Riesz-Fréchet, tal y como hicimos en el caso complejo para el operador adjunto.

Las propiedades del operador transpuesto son similares a las del adjunto:

1) $(A^t)^t = A$
2) $(A+B)^t = A^t + B^t$
3) $(\lambda A)^t = \lambda A^t$
4) $(AB)^t = B^t A^t$

6.3.2 Representación matricial del operador transpuesto

Queremos obtener la representación matricial del operador transpuesto A^t dada la del operador A. Sea V un espacio vectorial real de dimensión finita dotado de un producto escalar y $\mathcal{B} = \{u_1, \ldots, u_n\}$ una base ortonormal de V. Sea $\mathcal{A}$ la matriz de A en la base $\mathcal{B}$, es decir, $\mathcal{A} = (a_{ij})$:

$$Au_i = \sum_{i=1}^{n} a_{ji} u_j, \quad i = 1, \ldots, n$$

De lo estudiado para la expresión de los elementos de matriz del operador A se tiene:

$$a_{ij} = (u_i, Au_j)$$

y si $\mathcal{A}'$ es la matriz del operador transpuesto,

$$a'_{ij} = (u_i, A^t u_j)$$

como

$$(u_i, Au_j) = (A^t u_i, u_j) = (u_j, A^t u_i)$$

se concluye:

$$a_{ij} = a'_{ji}$$

es decir:

$$\mathcal{A}' = \mathcal{A}^t$$

la matriz del operador transpuesto es la matriz transpuesta del operador de partida cuando la base es ortonormal. El símbolo ^{t} denota tanto el operador transpuesto como la matriz transpuesta.

6.3.3 Operadores normales, simétricos y ortogonales

En el caso real, los operadores más interesantes serán los simétricos (análogos a los autoadjuntos) y ortogonales (análogos a los unitarios).

Definición 6.3.2 *Se dice que el operador A es normal si conmuta con su transpuesto:*

$$AA^t = A^tA.$$

Definición 6.3.3 *Se dice que el operador A es simétrico si coincide con su transpuesto:*

$$A^t = A.$$

Definición 6.3.4 *Se dice que el operador A es ortogonal si:*

$$AA^t = A^tA = 1_V.$$

Los operadores simétricos verifican:

$$(x, Ay) = (Ax, y)$$

y en bases ortonormales vienen representados por matrices simétricas:

$$\mathcal{A}^t = \mathcal{A}$$

Los operadores ortogonales verifican:

$$(Ax, Ay) = (x, y)$$

En bases ortonormales, sus matrices son ortogonales, es decir:

$$\mathcal{A}\mathcal{A}^t = \mathcal{A}^t\mathcal{A} = I_n$$

Es inmediato comprobar que los operadores simétricos y ortogonales son normales. Sin embargo en el caso real no estudiaremos los operadores normales. La razón es que los operadores simétricos tienen todos sus autovalores reales y su estudio es muy similar al caso complejo. Pero los ortogonales no los tienen reales en general (con más precisión no tienen autovalores en general)y por lo tanto requerirán un estudio especial.

6.3.4 Teorema espectral para operadores simétricos

El principal problema que surge en relación con el caso complejo es probar que los autovalores de un operador simétrico son reales (aunque no sea muy precisa, utilizaremos esta terminología para expresar el que las raíces del polinomio característico pueden no ser reales, en cuyo caso no son autovalores del operador).

Proposición 6.3.1 *Sea V un espacio vectorial real de dimensión finita con producto escalar. Sea A un operador normal en V. Si $x \in V$ es un autovector de A con autovalor λ, entonces x es autovector de A^t con el mismo autovalor.*

Demostración.

$$||(A^t - \lambda 1_V)x||^2 = ((A^t - \lambda 1_V)x, (A^t - \lambda 1_V)x) = ((A - \lambda 1_V)(A^t - \lambda 1_V)x, x) = ((A^t - \lambda 1_V)(A - \lambda 1_V)x, x) = 0$$

y por tanto:

$$(A^t - \lambda 1_V)x = 0$$

QED

Al igual que en el caso complejo, si un subespacio es invariante bajo un operador A, su complemento ortogonal lo es bajo el operador transpuesto.

El teorema espectral para operadores simétricos se puede enunciar como sigue:

Teorema 6.3.1 *Sea V un espacio vectorial real de dimensión finita con producto escalar. Sea A un operador simétrico en V. Entonces, existe una base ortonormal de V formada por autovectores de A.*

Demostración. Demostramos en primer lugar que las raíces del polinomio característico de A son todas reales. Para ello, consideremos el polinomio mínimo de A, $m(\lambda)$. Los factores irreducibles de este polinomio que tiene los coeficientes reales, son de grado 1 o 2. Demostremos que no pueden ser de grado 2. Si tuviera un factor irreducible de este grado:

$$m(\lambda) = [(\lambda - a)^2 + b^2]m_1(\lambda)$$

donde $b \neq 0$. El polinomio mínimo anula al operador, es decir, $m(A) = 0$. Entonces $\forall x \in V$ se tiene:

$$m(A)x = 0 \Rightarrow [(A - a1_V)^2 + b^2 1_V]m_1(A)x = 0$$

Sea $y = m_1(A)x$. Entonces:

$$[(A - a1_V)^2 + b^2 1_V]y = 0$$

Calculemos el producto escalar:

$$\begin{aligned}([(A - a1_V)^2 + b^2 1_V]y, y) &= ((A - a1_V)^2 y, y) + (b^2 y, y) = ((A - a1_V)y, (A - a1_V)y) + b^2(y, y) = \\ &= ||(A - a1_V)y||^2 + b^2||y||^2 = 0.\end{aligned}$$

Como $b \neq 0$, $||y|| = 0$, es decir $y = 0$. En consecuencia, el operador $m_1(A)$ es cero, y por tanto $m(\lambda)$ no sería el polinomio mínimo. No hay factores de grado 2, y por consiguiente los autovalores son reales.

El argumento es ahora igual que en el caso complejo. Se toma un autovector de A y se construye su subespacio ortogonal, que es invariante bajo $A^t = A$. En este espacio ortogonal se construye otro autovector y se sigue el proceso hasta tener una base ortonormal de V formada por autovectores de A. QED

El resultado inverso es también cierto.

Teorema 6.3.2 *Sea V un espacio vectorial real de dimensión finita con un producto escalar. Sea A un operador en V, tal que existe una base ortonormal de V formada por autovectores de A. Entonces A es simétrico.*

Demostración. Sea $\{u_1, \ldots, u_n\}$ la base ortonormal. Entonces:

$$Au_k = \lambda_k u_k$$

Ahora bien,

$$A^t u_k = \sum_{i=1}^n (u_i, A^t u_k)u_i = \sum_{i=1}^n (Au_i, u_k)u_i = \sum_{i=1}^n \lambda_i (u_i, u_k)u_i = \sum_{i=1}^n \lambda_i \delta_{ik} u_i = \lambda_k u_k$$

y por lo tanto,

$$A^t = A$$

QED

Las matrices que intercambian las bases ortonormales son ortogonales (la demostración es idéntica a la hecha en el caso complejo). Se tiene el resultado siguiente para matrices:

Teorema 6.3.3 *Toda matriz simétrica (real) es diagonalizable por una transformación ortogonal.*

Demostración. Una matriz simétrica se puede considerar como la matriz de un operador simétrico en una base ortonormal. Pasando a la base ortonormal de autovectores la matriz que representa al operador es ahora diagonal y la matriz de cambio de base es ortogonal:

$$\mathcal{P}^t \mathcal{A} \mathcal{P}$$

es diagonal. QED

6.3.5 Descomposición espectral de operadores simétricos

Al igual que los operadores normales, los operadores simétricos admiten una descomposición espectral.

Sea A un operador simétrico en un espacio vectorial real de dimensión finita dotado de un producto escalar.

Sea $\sigma(A) = \{\lambda_1, \ldots, \lambda_r\}$ el espectro de A. Sean $V_i = \ker(A - \lambda_i 1_V)$ los subespacios invariantes. Entonces:

$$V = V_1 \oplus \cdots \oplus V_r$$

y los subespacios son ortogonales entre sí, al corresponder a autovalores distintos.

Existe entonces una familia de proyectores ortogonales (simétricos e idempotentes) que además verifican:

1) $P_i P_j = 0,\ i \neq j$
2) $P_1 + \cdots + P_r = 1_V$
3) $\lambda_1 P_1 + \cdots + \lambda_r P_r = A$

El cálculo de proyectores ortogonales se hace igual que en el caso complejo, bien en una base dada, o empleando polinomios interpoladores.

La descomposición espectral permite identificar a los operadores simétricos:

Si dado un operador A en un espacio vectorial real de dimensión finita con un producto escalar, existe una familia de proyectores ortogonales (simétricos e idempotentes) que verifican las anteriores propiedades para una familia de escalares (reales) distintos, entonces A es simétrico y esos escalares son sus autovalores. La multiplicidad del autovalor λ_i es la dimensión del espacio $P_i V$.

6.4 Operadores ortogonales

El último tipo de operadores en espacios con producto escalar que vamos a estudiar son los operadores ortogonales. En este caso, al no ser (en general) todas las raíces del polinomio característico reales, no podremos encontrar una base formada por autovectores. Sin embargo, existen bases en las que estos operadores adoptan formas sencillas. Son estas formas canónicas las que vamos a discutir. Comenzaremos por el caso de dimensión 2 y veremos como los demás se reducen a éste. La razón es que los factores irreducibles del polinomio característico (un polinomio con coeficientes reales) son de grado 1 o 2. Recordemos que los operadores ortogonales vienen representados por matrices ortogonales en bases ortonormales:

$$\mathcal{A}^t \mathcal{A} = I_n$$

De esta relación se deduce que el determinante de una matriz ortogonal (y por lo tanto de un operador ortogonal) es igual a ± 1. Los operadores ortogonales con determinante igual a $+1$ se llaman rotaciones.

El conjunto de operadores ortogonales forma un grupo respecto a la composición de operadores. Se le llama el grupo ortogonal ($O(n)$). El subconjunto de operadores ortogonales con determinante igual a 1 es un subgrupo de éste, ($SO(n)$).

6.4.1 Operadores ortogonales en un espacio de dimensión 2

Sea V un espacio vectorial real de dimensión 2 con producto escalar y A un operador ortogonal en este espacio:

$$AA^t = A^t A = 1_V$$

Proposición 6.4.1 *En las condiciones anteriores, existe una base ortonormal de V en la que el operador A tienen como matriz una de las siguientes:*

$$1) \begin{pmatrix} \cos\theta & -\operatorname{sen}\theta \\ \operatorname{sen}\theta & \cos\theta \end{pmatrix}, \quad 2) \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

En el primer caso, el determinante de la matriz es 1. En el segundo es -1.

Demostración. Sea

$$\mathcal{A} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

la matriz de A en una base ortonormal. Entonces:

$$\mathcal{A}^t\mathcal{A} = I_2$$

y operando:

$$\begin{pmatrix} a & c \\ b & d \end{pmatrix}\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} a^2+c^2 & ab+cd \\ ab+cd & b^2+d^2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

es decir:

$$a^2 + c^2 = b^2 + d^2 = 1 \quad (6.5)$$

$$ab + cd = 0 \quad (6.6)$$

Si $a^2 + c^2 = 1$ podemos encontrar un número real θ en el intervalo $[0, 2\pi)$ tal que:

$$a = \cos\theta, \quad c = \operatorname{sen}\theta$$

Razonando igual con $b^2 + d^2 = 1$:

$$b = \cos\theta', \quad d = \operatorname{sen}\theta'$$

Sustituyendo estos valores en la tercera ecuación:

$$cos\theta \operatorname{sen}\theta' + \operatorname{sen}\theta\cos\theta' = 0$$

$$\operatorname{sen}(\theta + \theta') = 0$$

es decir,

$$\theta' = \theta \quad \text{o} \quad \theta' = \pi - \theta$$

En el primer caso, la matriz del operador es:

$$\mathcal{A} = \begin{pmatrix} \cos\theta & -\operatorname{sen}\theta \\ \operatorname{sen}\theta & \cos\theta \end{pmatrix}$$

es decir, el tipo 1).

En el segundo caso:

$$\mathcal{A} = \begin{pmatrix} \cos\theta & \operatorname{sen}\theta \\ \operatorname{sen}\theta & -\cos\theta \end{pmatrix}$$

En el primer caso no hay ningún vector invariante (salvo, obviamente el vector cero). Se trata de una rotación de ángulo θ en sentido antihorario (positivo). Su determinante es igual a 1. Toda matriz ortogonal 2×2 de determinante igual a 1 tiene esta forma (con distintos valores de θ) en cualquier base ortogonal.

Sin embargo, la segunda matriz tiene un vector invariante con autovalor igual a 1 (o el operador correspondiente):

$$\det \begin{pmatrix} \cos\theta - \lambda & \operatorname{sen}\theta \\ \operatorname{sen}\theta & -\cos\theta - \lambda \end{pmatrix} = \lambda^2 - 1 = 0$$

tiene como soluciones $\lambda = \pm 1$. Sea u_1 el autovector de autovalor 1 y norma 1:

$$Au_1 = u_1$$

Escojamos un vector unitario, u_2, ortogonal a u_1. Estos dos vectores son l.i. y forman una base ortogonal en V. En ella la matriz de A tiene la forma:

$$\begin{pmatrix} 1 & \alpha \\ 0 & \beta \end{pmatrix}$$

debido a que u_1 es autovector de autovalor 1. Pero esta matriz debe ser ortogonal (y tener determinate igual a -1). Por lo tanto, $\beta = -1$ y $\alpha = 0$. Es decir, el segundo vector (elegido como ortogonal a u_1) es justamente el autovector de autovalor -1. En este caso, tenemos una base ortonormal de autovectores del operador A. Esta es la segunda de las formas canónicas de las que habla la proposición. En la base $\{u_1, u_2\}$ se observa que este operador representa una reflexión. Hay una recta que permanece invariante (punto a punto) la asociada al autovector u_1. El resto de vectores sufre una reflexión con respecto a esta recta (en particular el u_2 que simplemente cambia de signo). Toda matriz ortogonal (2×2) de determinante negativo es el producto de una matriz que representa una rotación (con determinante positivo) por la matriz de la segunda forma canónica que aparece en la proposición. QED

6.4.2 Subespacios invariantes de un operador ortogonal

Sea V un espacio real de dimensión finita dotado de un producto escalar, y A un operador ortogonal en V.

Proposición 6.4.2 *Los autovalores de A (en caso de que existan) son iguales a ± 1.*

Demostración. Si $\lambda \in \mathbb{R}$ es un autovalor de A, sea x un autovector con ese autovalor. Entonces:

$$(Ax, Ax) = \lambda^2 (x, x) \Rightarrow \lambda^2 = 1$$

QED

Pero podría ocurrir que no tuviera ningún autovalor. Al ser los autovalores las raíces del polinomio característico, un polinomio con coeficientes reales, las raíces complejas aparecen siempre a pares (una y su compleja conjugada). Por tanto, en dimensión impar siempre existen autovalores reales (al menos uno, igual a $+1$ ó -1). En este caso, siempre hay un subespacio invariante de dimensión 1. En general se tiene el resultado siguiente:

Proposición 6.4.3 *Sea V un espacio vectorial real de dimensión finita, con un producto escalar y A un operador ortogonal. Entonces, existe un subespacio invariante de V de dimensión 1 ó 2.*

Demostración. Se considera el polinomio característico de A que se puede descomponer en factores irreducibles (algunos posiblemente repetidos) de grados 1 ó 2:

$$p(\lambda) = p_1(\lambda) \cdots p_r(\lambda)$$

Como $p(A) = 0$, consideremos un vector $x \in V$ tal que existe un número $j \in \{1, \ldots, r\}$ y;

$$p_1(A) \cdots p_j(A)x = 0, \quad p_1(A) \cdots p_{j-1}(A)x \neq 0$$

Sea $y = p_1(A) \cdots p_{j-1}(A)x$. Entonces:

$$p_j(A)y = 0$$

Si el grado de $p_j(\lambda)$ es igual a 1:

$$(A + b1_V)y = 0 \Rightarrow Ay = -by$$

y existe un subespacio de dimensión 1, generado por y, invariante bajo A.

Pero si el grado de $p_j(\lambda)$ es igual a 2, entonces:

$$(A^2 + aA + b1_V)y = 0$$

y $\{y, Ay\}$ generan un subespacio invariante de dimensión 2. En efecto:

$$Ay \in \text{lin}\{y, Ay\}, \quad A(Ay) = A^2 y = -aAy - by \in \text{lin}\{y, Ay\}$$

QED

6.4.3 Forma canónica de un operador ortogonal

Con lo dicho anteriormente podemos construir una base de V donde un operador ortogonal adopta una forma sencilla. Basta considerar un subespacio invariante (que siempre existe, de dimensión 1 o 2) y descomponer el espacio en una suma de ese subespacio más su complementario ortogonal (que también será invariante).

Teorema 6.4.1 *Sea V un espacio vectorial real de dimensión finita, dotado de un producto escalar y A un operador ortogonal en V. Entonces, existe una base ortonormal de V en la que la matriz que representa al operador tiene la forma:*

$$\begin{pmatrix} 1 & & & & & & & & & & \\ & \ddots & & & & & & & & & \\ & & 1 & & & & & & & & \\ & & & -1 & & & & & & & \\ & & & & \ddots & & & & & & \\ & & & & & -1 & & & & & \\ & & & & & & \cos\theta_1 & -\operatorname{sen}\theta_1 & & & \\ & & & & & & \operatorname{sen}\theta_1 & \cos\theta_1 & & & \\ & & & & & & & & \ddots & & \\ & & & & & & & & & \cos\theta_r & -\operatorname{sen}\theta_r \\ & & & & & & & & & \operatorname{sen}\theta_r & \cos\theta_r \end{pmatrix}$$

Es decir, el espacio se descompone en la suma de subespacios invariantes de dimensión 1 ó 2.

El número de autovalores $+1$ y -1 coincide con sus respectivas multiplicidades. En cada bloque 2×2 no hay autovalores, pero las raíces del polinomio característico son $e^{\pm i\theta_j}$ y el número de estos bloques para un mismo valor de θ_j coincide con la multiplicidad de estas raíces.

Por ejemplo, en dimensión 3 siempre hay un autovalor real (al menos) que es igual a $+1$ o a -1. Si el determinante es $+1$, hay un autovalor igual a $+1$. Se trata de una rotación, cuyo eje (conjunto de vectores invariantes) viene dado por el autovector correspondiente al autovalor 1. Hay un segundo subespacio de dimensión 2 (salvo casos degenerados) invariante bajo la rotación, y que no contiene ningún autovector. Es el (único) espacio ortogonal al eje.

Tema 7

Tensores

Aplicaciones multilineales. Coordenadas contravariantes y covariantes. Producto tensorial. Cambios de base. Tensores y grupos de transformaciones. Tensores sobre espacios con producto escalar.

El tema que se expone a continuación contiene una introducción a tensores. Su dificultad es superior a la media de los temas anteriores y solo presenta los tensores desde un punto de vista algebraico, perdiéndose por consiguiente una visión más geométrica (y quizás más intuitiva) de este concepto, debido a la no inclusión de campos tensoriales.

7.1 Una justificación

Ejemplo 7.1.1 Las ecuaciones de la dinámica newtoniana

El movimiento de un punto material (con $m = 1$) viene descrito en la mecánica newtoniana por un sistema de ecuaciones diferenciales:

$$\begin{aligned} \frac{d^2x}{dt^2} &= f_x(x,y,z,t) \\ \frac{d^2y}{dt^2} &= f_y(x,y,z,t) \\ \frac{d^2z}{dt^2} &= f_z(x,y,z,t) \end{aligned}$$

La razón de reunirlas en una notación vectorial proviene de sus propiedades de transformación frente al cambio de sistema de referencia. Un cambio de este tipo mezcla las coordenadas y las componentes de la fuerza en la manera en que lo hacen los cambios de base de un espacio vectorial:

$$\frac{d^2\vec{r}}{dt^2} = \vec{F}(\vec{r},t)$$

Ejemplo 7.1.2 Las ecuaciones de Maxwell

Las ecuaciones de Maxwell del campo electromagnético relacionan los campos eléctrico ($\vec{E}$) y magnético ($\vec{B}$) con las cargas (ρ) y corrientes ($\vec{j}$). Una forma de escribirlas en una notación compacta es la siguiente (tomamos $c = 1$):

$$\begin{aligned} \partial_\mu F^{\mu\nu} &= -4\pi j^\nu \\ \partial_\mu F_{\nu\rho} + \partial_\rho F_{\mu\nu} + \partial_\nu F_{\rho\mu} &= 0 \end{aligned}$$

donde $F^{\mu\nu}$ es el llamado tensor del campo electromagnético. Sus componentes, es decir, los valores de $F^{\mu\nu}$ (y $F_{\mu\nu}$, que son distintos como estudiaremos más adelante), cuando los índices $\mu\nu$ recorren los valores $0,1,2,3$ (nótese que empezamos a numerar en 0), son los siguientes:

$$F^{00} = F^{11} = F^{22} = F^{33} = F_{00} = F_{11} = F_{22} = F_{33} = 0$$

$$
\begin{aligned}
F^{01} &= -F^{10} = -F_{01} = F_{10} = E_x \\
F^{02} &= -F^{20} = -F_{02} = F_{20} = E_y \\
F^{03} &= -F^{30} = -F_{03} = F_{30} = E_z \\
F^{12} &= -F^{21} = F_{12} = -F_{21} = B_z \\
F^{31} &= -F^{13} = F_{31} = -F_{13} = B_y \\
F^{23} &= -F^{32} = F_{23} = -F_{32} = B_x
\end{aligned}
$$

que se pueden disponer como una matriz:

$$
F^{\mu\nu} = \begin{pmatrix} 0 & E_x & E_y & E_z \\ -E_x & 0 & B_z & -B_y \\ -E_y & -B_z & 0 & B_x \\ -E_z & B_y & -B_x & 0 \end{pmatrix}, \quad F_{\mu\nu} = \begin{pmatrix} 0 & -E_x & -E_y & -E_z \\ E_x & 0 & B_z & -B_y \\ E_y & -B_z & 0 & B_x \\ E_z & B_y & -B_x & 0 \end{pmatrix}
$$

Además, $\partial_0 = \partial_t, \partial_1 = \partial_x, \partial_2 = \partial_y, \partial_3 = \partial_z$ y $j^0 = \rho, j^1 = j_x, j^2 = j_y, j^3 = j_z$. En la notación $\vec{E}, \vec{B}, \rho, \vec{j}$, la primera de las ecuaciones de Maxwell se escribe como:

$$
\begin{aligned}
\nabla \vec{E} &= 4\pi\rho \\
\frac{\partial \vec{E}}{\partial t} - \nabla \times \vec{B} &= -4\pi\vec{j}
\end{aligned}
$$

La segunda es:

$$
\begin{aligned}
\nabla \vec{B} &= 0 \\
\frac{\partial \vec{B}}{\partial t} + \nabla \times \vec{E} &= 0
\end{aligned}
$$

Las ventajas de utilizar el objeto $F^{\mu\nu}$ son, en parte, las de manejar de forma sencilla los cambios de sistemas de referencia. Como veremos, $F^{\mu\nu}$ será un tensor de segundo orden, 2 veces contravariante, y esto determina de forma unívoca sus propiedades de transformación bajo cambios de base (concretamente los asociados a la teoría de la relatividad).

Ejemplo 7.1.3 Energía cinética de rotación de un sólido rígido

Consideremos un sólido rígido que supondremos descrito por un sistema de n partículas de masas m_i con posiciones $\vec{r}_i$, y velocidades $\vec{v}_i = \dot{\vec{r}}_i$. La energía cinética es la suma de las energías cinéticas de cada partícula , es decir:

$$
T = \frac{1}{2} \sum_{i=1}^{n} m_i \vec{v}_i^2
$$

Si el sólido gira en torno a un eje (supongamos que no hay movimiento de traslación) de vector unitario $\vec{e}$ (constante), podemos definir una velocidad angular, $\vec{\omega}$ (en la dirección del eje de rotación) tal que:

$$
\vec{v}_i = \vec{\omega} \times \vec{r}_i, \quad i = 1, \ldots, n
$$

El momento angular de una de las partículas es $m_i \vec{r}_i \times \vec{v}_i$ y el total es la suma:

$$
\vec{L} = \sum_{i=1}^{n} m_i \vec{r}_i \times \vec{v}_i
$$

Sustituyendo $\vec{v}_i$ de la expresión anterior:

$$
\vec{L} = \sum_{i=1}^{n} m_i \vec{r}_i \times (\vec{\omega} \times \vec{r}_i)
$$

y operando:

$$\vec{L} = \sum_{i=1}^{n} m_i(|\vec{r}_i|^2\vec{\omega} - (\vec{r}_i.\vec{\omega})\vec{r}_i)$$

Escribamos las coordenadas de $\vec{r}_i$ y $\vec{\omega}$ en un sistema de referencia ortonormal como:

$$\vec{r}_i = (x_{i1}, x_{i2}, x_{i3}), \quad \vec{\omega} = (\omega_1, \omega_2, \omega_3)$$

Entonces, el momento angular se escribe en coordenadas como:

$$L_k = \sum_{j=1}^{3}\sum_{i=1}^{n} m_i(\omega_k|\vec{r}_i|^2 - x_{ik}x_{ij}\omega_j) = \sum_{j=1}^{3}\omega_j \sum_{i=1}^{n} m_i(\delta_{kj}|\vec{r}_i|^2 - x_{ik}x_{ij}), \quad k = 1,2,3$$

y si definimos:

$$I_{kj} = \sum_{i=1}^{n} m_i(\delta_{kj}|\vec{r}_i|^2 - x_{ik}x_{ij})$$

podemos escribir:

$$L_k = \sum_{j=1}^{3} I_{kj}\omega_j$$

Se llama tensor de inercia a un objeto cuyas componentes en el sistema de referencia en el que estamos trabajando son I_{ij}. Los valores explícitos que toman estas componentes son:

$$I_{11} = \sum_{i=1}^{n} m_i(x_{i2}^2 + x_{i3}^2) \qquad I_{22} = \textstyle\sum_{i=1}^{n} m_i(x_{i1}^2 + x_{i3}^2) \qquad I_{33} = \displaystyle\sum_{i=1}^{n} m_i(x_{i1}^2 + x_{i2}^2)$$

$$I_{12} = I_{21} = -\sum_{i=1}^{n} m_i x_{i1}x_{i2} \quad I_{13} = I_{31} = -\textstyle\sum_{i=1}^{n} m_i x_{i1}x_{i3} \quad I_{23} = I_{32} = -\displaystyle\sum_{i=1}^{n} m_i x_{i2}x_{i3}$$

que corresponden a los momentos de inercia respecto a los ejes coordenados y a los opuestos de los productos de inercia.

La energía cinética de rotación del sistema es:

$$T = \frac{1}{2}\sum_{i=1}^{3} L_i\omega_i = \frac{1}{2}\sum_{i,j=1}^{3} I_{ij}\omega_i\omega_j$$

que resulta ser una forma cuadrática en la velocidad angular.

Los cambios de coordenadas en este sistema vienen dados por transformaciones entre bases ortonormales (en el producto escalar usual de $\mathbb{R}^3$) que preservan la orientación, es decir por rotaciones (además de por traslaciones, que no son lineales y no serán tratadas aquí). Uno puede imaginar transformaciones de coordenadas más generales, pero, dado que trabajamos con un sólido rígido, éstas deben conservar ángulos, distancias y orientación.

Si R es la matriz de la rotación:

$$x'_{ij} = \sum_{k=1}^{3} R_{jk}x_{ik}$$

Las nuevas componentes del tensor de inercia son:

$$I'_{kj} = \sum_{i=1}^{n} m_i(\delta_{kj}|\vec{r}'_i|_2 - x'_{ik}x'_{ij})$$

y sustituyendo:

$$\begin{aligned} I'_{kj} &= \sum_{i=1}^{n} m_i(\delta_{kj}|\vec{r}_i|_2 - \sum_{l,s=1}^{3} R_{kl}R_{js}x_{il}x_{is})) = \sum_{l,s=1}^{3} R_{kl}R_{js}\left(\sum_{i=1}^{n} m_i(\delta_{kj}|\vec{r}_i|_2 - x_{il}x_{is}\right) \\ &= \sum_{l,s=1}^{3} R_{kl}R_{js}I_{ls} \end{aligned}$$

Esta regla de trasformación para el tensor de inercia es la que justifica el nombre de tensor dado a este conjunto de cantidades I_{ij}.

7.2 Aplicaciones multilineales

Consideremos los espacios vectoriales $E_1, \ldots, E_n, F$ sobre un cuerpo $\mathbb{K}$, y una aplicación φ:

$$\varphi\colon E_1 \times \cdots \times E_n \to F$$

que verifica:

$$\varphi(v_1, \ldots, \lambda v_i + \lambda' v_i', \ldots, v_n) = \lambda \varphi(v_1, \ldots, v_i, \ldots, v_n) + \lambda' \varphi(v_1, \ldots, v_i', \ldots, v_n)$$

para cualquier índice $i = 1, \ldots, n$, $\lambda, \lambda' \in \mathbb{K}$.

Se dice que φ es una aplicación multilineal.

Supongamos que los espacios E_i, F, $i = 1, \ldots, n$ son de dimensión finita y que $\mathcal{B}_i = \{u_1^{(i)}, \ldots, u_{d_i}^{(i)}\}$ son bases de E_i respectivamente y $\mathcal{B} = \{u_1, \ldots, u_d\}$ es una base de F. Veamos como se puede escribir la aplicación multilineal referida a estas bases:

$$\varphi(v_1, \ldots, v_n) = \varphi(\sum_{i_1=1}^{d_1} x_{i_1}^{(1)} u_{i_1}^{(1)}, \ldots, \sum_{i_n=1}^{d_n} x_{i_n}^{(n)} u_{i_n}^{(n)}) = \sum_{i_1=1}^{d_1} \cdots \sum_{i_n=1}^{d_n} x_{i_1}^{(1)} \ldots x_{i_n}^{(n)} \varphi(u_{i_1}^{(1)}, \ldots, u_{i_n}^{(n)})$$

Es decir, la aplicación multilineal queda determinada por el valor que toma sobre las bases de los espacios E_i. La expresión $\varphi(u_{i_1}^{(1)}, \ldots, u_{i_n}^{(n)})$ representa un vector de F, que se podrá escribir en la base de F dada:

$$\varphi(u_{i_1}^{(1)}, \ldots, u_{i_n}^{(n)}) = \sum_{i=1}^{d} t_{i_1 \ldots i_n i} u_i$$

y por tanto:

$$\varphi(v_1, \ldots, v_n) = \sum_{i_1=1}^{d_1} \cdots \sum_{i_n=1}^{d_n} \sum_{i=1}^{d} t_{i_1 \ldots i_n i} x_{i_1}^{(1)} \ldots x_{i_n}^{(n)} u_i$$

Estos valores $t_{i_1 \ldots i_n i}$ son un conjunto de $d_1 \ldots d_n d$ escalares, que, obviamente, dependen de las bases elegidas.

Si cambiamos las bases, la expresión de φ cambiará. Sean $P_1, \ldots, P_n$ las matrices de cambio de base en cada uno de los espacios vectoriales considerados, que pasan de $\mathcal{B}_i$ a $\mathcal{B}_i'$. Es decir:

$$x_{j_i}^{(i)} = \sum_{k_i=1}^{d_i} (P_i^{-1})_{j_i k_i} x_{k_i}'^{(i)}, \quad i = 1, \ldots, n$$

y también, en el espacio F, el cambio de base de $\mathcal{B}$ a $\mathcal{B}'$ es:

$$u_i = \sum_{j=1}^{d} P_{ji} u_j'$$

Sustituyendo en la expresión de φ:

$$\varphi(v_1, \ldots, v_n) = \sum_{i_1=1}^{d_1} \cdots \sum_{i_n=1}^{d_n} \sum_{i=1}^{d} t_{i_1 \ldots i_n i} x_{i_1}^{(1)} \ldots x_{i_n}^{(n)} u_i \quad =$$

$$\sum_{i_1=1}^{d_1} \cdots \sum_{i_n=1}^{d_n} \sum_{i=1}^{d} t_{i_1 \ldots i_n i} \sum_{j_1=1}^{d_1} (P_1^{-1})_{i_1 j_1} x_{j_1}'^{(1)} \ldots \sum_{j_n=1}^{d_n} (P_n^{-1})_{i_n j_n} x_{j_n}'^{(n)} \sum_{j=1}^{d} P_{ji} u_j' \quad =$$

$$\sum_{j_1=1}^{d_1} \cdots \sum_{j_n=1}^{d_n} \sum_{j=1}^{d} \left(\sum_{i_1=1}^{d_1} \cdots \sum_{i_n=1}^{d_n} \sum_{i=1}^{d} (P_1^{-1})_{i_1 j_1} \ldots (P_n^{-1})_{i_n j_n} P_{ji} t_{i_1 \ldots i_n i} \right) x_{j_1}'^{(1)} \ldots x_{j_n}'^{(n)} u_j'$$

y por lo tanto:

$$t'_{j_1\dots j_n j} = \sum_{i_1=1}^{d_1} \dots \sum_{i_n=1}^{d_n} \sum_{i=1}^{d} (P_1^{-1})_{i_1 j_1} \dots (P_n^{-1})_{i_n j_n} P_{ij} t_{i_1\dots i_n i}$$

Como se ve, la transformación guarda un cierto parecido con la encontrada para el tensor de inercia de nuestro primer ejemplo. Concretamente, si tenemos dos espacios E_i y el espacio F es igual al cuerpo $\mathbb{K}$, la transformación es:

$$t'_{j_1 j_2} = \sum_{i_1=1}^{d_1} \sum_{i_2=1}^{d_2} (P_1^{-1})_{i_1 j_1} (P_2^{-1})_{i_2 j_2} t_{i_1 i_2}$$

Las diferencias que se observan serán explicadas más adelante.

7.3 Coordenadas

7.3.1 Coordenadas contravariantes y covariantes

En muchas de las aplicaciones de la teoría de tensores se trabaja con espacios en los que hay definido un producto escalar. Además, todos los espacios E_i son iguales a un espacio vectorial V. Supongamos además que el espacio F es el cuerpo $\mathbb{K}$. Con esto, las aplicaciones multilineales (formas multilineales) de las que hablábamos en la sección anterior, son ahora:

$$\varphi : V \times \cdots \times V \to \mathbb{K}$$

y en una base de V, $\mathcal{B} = \{u_1, \dots u_n\}$, se escriben como:

$$\varphi(v_1, \dots, v_n) = \sum_{i_1=1}^{d_1} \cdots \sum_{i_n=1}^{d_n} t_{i_1\dots i_n} x_{(1)}^{i_1} \dots x_{(n)}^{i_n}$$

Nótese que hemos colocado en alto los índices de las coordenadas de los vectores v_i, es decir:

$$v_i = \sum_{j=1}^{d_i} x^j u_j = x^j u_j$$

expresión en la que hemos introducido una simplificación (convenio de Einstein). Suprimimos los signos de suma, y suponemos que cuando dos índices son iguales en una expresión y uno está arriba y otro abajo, hay una suma implícita sobre el rango correspondiente:

$$\varphi(v_1, \dots, v_n) = t_{i_1\dots i_n} x_{(1)}^{i_1} \dots x_{(n)}^{i_n}$$

Mientras nos acostumbramos a este convenio, usaremos el signo de suma de vez en cuando.

Como hemos dicho, suponemos que en V hay un producto escalar, de manera, que en, por ejemplo, la base que estamos considerando, se tiene:

$$(u_i, u_j) = g_{ij}, \quad i, j = 1, \dots, n$$

pudiendo escribir g_{ij} como una matriz, cuyo determinante es distinto de cero. Asociada a esta base, vamos a introducir otra, la base recíproca, $\mathcal{B}_r = \{u^1, \dots, u^n\}$, que verifica:

$$(u^i, u_j) = \delta^i{}_j, \quad i, j = 1, \dots, n$$

La base recíproca está determinada unívocamente por la base de partida. En efecto, escrita en la base de partida:

$$u^i = \sum_{k=1}^{n} \lambda^{ki} u_k$$

y multiplicando por u_k se tiene:

$$(u^i, u_j) = \sum_{k=1}^{n} \lambda^{ki}(u_k, u_j) = \sum_{k=1}^{n} \lambda^{ki} g_{kj} = \delta^i{}_j$$

tenemos pues un sistema en el que las incógnitas son λ^{ij} y la matriz de coeficientes es g_{ij}, luego el sistema tiene solución única. Desde el punto de vista de representación matricial, la matriz de los coeficientes λ^{ij} es la inversa (transpuesta) de la matriz correspondiente a g_{ij}, y la llamaremos:

$$\lambda^{ij} = g^{ij}.$$

Se tiene la identidad:

$$g^{ki} g_{kj} = \delta^i{}_j$$

Buscamos ahora la relación que existe entre las coordenadas en una base y en su recíproca. Este es un problema trivial de cambios de base. Sea $v \in V$, tal que:

$$v = \sum_{i=1}^{n} x^i u_i = \sum_{i=1}^{n} x_i u^i$$

Entonces, si $u^i = g^{ji} u_j$:

$$x_i = g_{ik} x^k, \quad x^i = g^{ik} x_k$$

Las coordenadas en la base de partida se llaman coordenadas contravariantes del vector. Las coordenadas en la base recíproca se llaman coordenadas covariantes. Obviamente, hay que fijar una base para empezar a discutir estos términos.

Si los vectores de la base tienen norma unidad, las coordenadas contravariantes se obtienen trazando las paralelas a estos vectores (en la forma usual en que se descompone un vector como suma de otros dos), mientras que las coordenadas covariantes se obtienen trazando las perpendiculares a los vectores que forman la base.

> Nota. Si la base de partida es una base ortonormal, $g_{ij} = \delta_{ij}$, la base recíproca coincide con la de partida y las coordenadas contravariantes y covariantes son iguales.

Supongamos ahora que cambiamos la base $\mathcal{B}$ a una base $\mathcal{B}'$. La base recíproca cambia también, de $\mathcal{B}_r$ a $\mathcal{B}'_r$. Sea:

$$u_i = P^j{}_i u'_j$$

el cambio de base. Para la base recíproca se tendrá:

$$u^i = g^{ji} u_j = g^{ji} P^k{}_j u'_k = g^{ji} P^k{}_j g'_{mk} u'^m = Q_m{}^i u'^m$$

donde g'_{ij} es la matriz del producto escalar en la base $\mathcal{B}'$:

$$Q_m{}^i = g'_{mk} P^k{}_j g^{ji}$$

Pero no es difícil darse cuenta que estas dos matrices P y Q son transpuestas inversas una de la otra:

$$\delta^l{}_m = (u^l, u_m) = (Q_i{}^l u'^i, P^k{}_m u'_k) = Q_i{}^l P^k{}_m (u'^i, u'_k) = Q_i{}^l P^k{}_m \delta^i{}_k$$

es decir,

$$Q_i{}^l P^i{}_m = \delta^l{}_m$$

Todo lo anterior se puede establecer en términos de matrices. La única justificación para emplear índices aquí es la facilidad con que se extienden estas nociones al caso de tensores arbitrarios (con más de uno o dos índices). Sean G, G^{-1}, G', G'^{-1} las matrices del producto escalar en las bases $\mathcal{B}, \mathcal{B}_r, \mathcal{B}', \mathcal{B}'_r$ respectivamente. Sea P la matriz del cambio de base de $\mathcal{B}$ a $\mathcal{B}'$ (es decir, las columnas de P son las coordenadas de los vectores de la base $\mathcal{B}$ en la base $\mathcal{B}'$) y Q la de $\mathcal{B}_r$ a $\mathcal{B}'_r$. De acuerdo con lo que hemos visto, G es también la matriz de cambio de base de $\mathcal{B}$ a $\mathcal{B}_r$ y G' la del cambio de $\mathcal{B}'$ a $\mathcal{B}'_r$. El siguiente diagrama puede contribuir a aclarar la situación:

$$\begin{tikzcd} B \arrow[r, "P"] \arrow[d, "G"'] & B' \arrow[d, "G'"] \\ B_r \arrow[r, "Q"'] & B'_r \end{tikzcd}$$

de donde se deduce que la matriz:

$$G'P = QG$$

es la de cambio de base de $\mathcal{B}$ a $\mathcal{B}'_r$. Por tanto, como al mismo tiempo la matriz G es la de una forma bilineal, su cambio es:

$$G' = (P^{-1})^t G P^{-1}$$

es decir:

$$G'P = (P^{-1})^t G = QG$$

de donde:

$$Q = (P^{-1})^t$$

como queríamos probar.

Las coordenadas contravariantes y covariantes no cambian de la misma manera al cambiar la base (se entiende, como hemos visto, que solo hay un cambio de base, de $\mathcal{B}$ a $\mathcal{B}'$. Los cambios en las bases recíprocas son consecuencia de este):

$$x'^i = P^i{}_j x^j, \quad x'_i = (P^{-1})^j{}_i x_j$$

> Nota. En todo lo anterior, lo fundamental ha sido el hecho de que el producto escalar es una forma bilineal no degenerada. Por ejemplo, no hemos hecho uso de la propiedad de ser definido positivo (salvo cuando hemos hablado de bases ortonormales). Por tanto lo anterior se puede hacer también con formas bilineales no degeneradas, lo que permite usar estos conceptos cuando no hay un producto escalar estricto, como en relatividad especial.

Veamos un ejemplo en $\mathbb{R}^2$ para ilustrar estas ideas.

Ejemplo 7.3.1 Sea la base de $\mathbb{R}^2$ (con el producto escalar usual):

$$\mathcal{B} = \left\{ u_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix},\ u_2 = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \right\}$$

que no es ortonormal. La matriz del producto escalar (y su inversa) en esta base es:

$$G = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}, \quad G^{-1} = \begin{pmatrix} 1 & -1 \\ -1 & 2 \end{pmatrix}$$

es decir:

$$g_{11} = 2,\ g_{12} = g_{21} = 1,\ g_{22} = 1, \quad g^{11} = 1,\ g^{12} = g^{21} = -1,\ g^{22} = 2$$

La base recíproca se obtiene imponiendo la condición: $(u^i, u_j) = \delta^i{}_j$:

$$\mathcal{B}_r = \left\{ u^1 = \begin{pmatrix} 0 \\ 1 \end{pmatrix},\ u^2 = \begin{pmatrix} 1 \\ -1 \end{pmatrix} \right\}$$

o, directamente:

$$u^1 = g^{11} u_1 + g^{21} u_2, \quad u^2 = g^{12} u_1 + g^{22} u_2$$

La relación entre coordenadas contravariantes y covariantes es:

$$\left\{ \begin{array}{lcl} x^1 & = & x_1 - x_2 \\ x^2 & = & -x_1 + 2x_2 \end{array} \right.$$

Si ahora pasamos a una nueva base:

$$\mathcal{B}' = \left\{ u_1' = \begin{pmatrix} -1 \\ 1 \end{pmatrix}, \ u_2' = \begin{pmatrix} 1 \\ 2 \end{pmatrix} \right\}, \quad \mathcal{B}_r' = \left\{ u'^1 = \begin{pmatrix} -2/3 \\ 1/3 \end{pmatrix}, \ u'^2 = \begin{pmatrix} 1/3 \\ 1/3 \end{pmatrix} \right\}$$

La matriz del producto escalar en estas bases es:

$$G' = \begin{pmatrix} 2 & 1 \\ 1 & 5 \end{pmatrix}, \quad (G')^{-1} = \begin{pmatrix} 5/9 & -1/9 \\ -1/9 & 2/9 \end{pmatrix}$$

y la matriz de cambio de base (de $\mathcal{B}$ a $\mathcal{B}'$, y de $\mathcal{B}_r$ a $\mathcal{B}_r'$) es:

$$P = \begin{pmatrix} -1/3 & -2/3 \\ 2/3 & 1/3 \end{pmatrix}, \quad Q = \begin{pmatrix} 1 & -2 \\ 2 & -1 \end{pmatrix}$$

siendo Q la inversa transpuesta de P. Se pueden comprobar las relaciones:

$$G'P = QG, \quad G' = (P^{-1})^t G P^{-1}$$

Como ejemplo concreto de coordenadas de un vector, sea:

$$x = \begin{pmatrix} -1 \\ 1 \end{pmatrix} \in \mathbb{R}^2$$

Sus coordenadas en las diferentes bases usadas, aparecen en la siguiente tabla:

	$\mathcal{B}$	$\mathcal{B}_r$	$\mathcal{B}'$	$\mathcal{B}_r'$
Contravariantes	$2,1$		$-4/3, 5/3$	
Covariantes		$5,3$		$-1,7$

7.3.2 Coordenadas en relatividad especial

Según los postulados de la relatividad especial, la velocidad de la luz es una constante en todos los sistemas inerciales, y el intervalo fundamental, la cantidad que se conserva al cambiar de sistema inercial, es:

$$c^2 t^2 - \vec{r}^2$$

donde c es la velocidad de la luz, que escogeremos igual a 1.

Supongamos un sistema unidimensional. Para describir sus coordenadas espacio-temporales, usaremos un vector de $\mathbb{R}^2$, $x = (x^0, x^1)$, dado por sus coordenadas en una base $\{e_0, e_1\}$. Las transformaciones de coordenadas admisibles en este espacio, desde un punto de vista de la teoría que estamos estudiando son las que dejan invariantes la forma bilineal simétrica no degenerada:

$$\varphi(x, y) = x^0 y^0 - x^1 y^1$$

que, sin embargo, no es un producto escalar, pues no es definida positiva. Esta forma bilineal tiene en la base que estamos considerando, la siguiente representación matricial:

$$g_{00} = 1, \ g_{01} = g_{10} = 0, \ g_{11} = -1$$

y el intervalo fundamental es:

$$\varphi(x, y) = g_{\mu\nu} x^\mu y^\nu$$

Diremos que una base es ortonormal respecto esta forma bilineal si la representación matricial es la dada (diagonal $(1, -1)$).

Calculemos las transformaciones lineales (homogéneas) que dejan invariantes la forma bilineal (es decir, que cambian las bases ortonormales entre sí).

Sea:

$$\Lambda = \begin{pmatrix} \Lambda^0{}_0 & \Lambda^0{}_1 \\ \Lambda^1{}_0 & \Lambda^1{}_1 \end{pmatrix}$$

la matriz de una transformación lineal en $\mathbb{R}^2$ en la base canónica. Si la forma bilineal anterior es invariante bajo esta transformación, entonces:

$$\Lambda^t K \Lambda = K$$

donde:

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

es la matriz de la forma bilineal.

Las ecuaciones que verifica Λ son:

$$({\Lambda^0}_0)^2 - ({\Lambda^1}_0)^2 = 1, \quad ({\Lambda^1}_1)^2 - ({\Lambda^0}_1)^2 = 1, \quad {\Lambda^0}_0{\Lambda^0}_1 - {\Lambda^1}_0{\Lambda^1}_1 = 0$$

Podemos parametrizar los coeficientes de la matriz por:

$${\Lambda^0}_0 = \epsilon \cosh t,\ {\Lambda^1}_0 = \operatorname{senh} t,\ {\Lambda^0}_1 = \eta \operatorname{senh} t,\ {\Lambda^1}_1 = \eta\epsilon \cosh t$$

donde $\epsilon = \pm 1,\ \eta = \pm 1$.

Este conjunto de matrices forma un grupo, $\mathcal{L}$, que llamaremos el grupo de Lorentz en $1+1$ (una dimensión temporal y otra espacial). Como se ve $\det \Lambda = \pm 1$ y $|{\Lambda^0}_0| \geq 1$. Se divide en cuatro subconjuntos según los valores del determinante y de la componente ${\Lambda^0}_0$:

$$\begin{aligned}
\mathcal{L}_+^\uparrow &= \{\Lambda \in \mathcal{L} \mid \det \Lambda = 1, {\Lambda^0}_0 \geq 1\} \\
\mathcal{L}_+^\downarrow &= \{\Lambda \in \mathcal{L} \mid \det \Lambda = 1, {\Lambda^0}_0 \leq 1\} \\
\mathcal{L}_-^\uparrow &= \{\Lambda \in \mathcal{L} \mid \det \Lambda = -1, {\Lambda^0}_0 \geq 1\} \\
\mathcal{L}_-^\downarrow &= \{\Lambda \in \mathcal{L} \mid \det \Lambda = -1, {\Lambda^0}_0 \leq 1\}
\end{aligned}$$

El primer subconjunto, $\mathcal{L}l_+^\uparrow$ es un subgrupo de $\mathcal{L}$, el subgrupo ortocrono propio, y consiste en las transformaciones que dejan invariantes el sentido del tiempo y no cambian la orientación del espacio. Al segundo pertenece la transformación:

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

que cambia la coordenada espacial por su negativa (paridad). Al cuarto pertenece:

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

que cambia la coordenada temporal por su negativa (inversión temporal).

Este tipo de transformaciones serán los cambios de base (los cambios de sistema de referencia) admisibles. Las leyes de la física, de acuerdo con el principio de relatividad tendrán la misma forma en todos los sistemas relacionados por las matrices del grupo $\mathcal{L}$ (sistemas inerciales).

Aunque no dispongamos de un producto escalar, podemos definir una base recíproca de la $\mathcal{B} = \{u_0, u_1\}$:

$$\mathcal{B}_r = \{u^0, u^1\}$$

con las propiedades:

$$\varphi(u^i, u_j) = \delta^i_j,\ i, j = 0, 1$$

De acuerdo con lo visto en el caso del producto escalar, la relación entre las coordenadas en la base inicial, $x^\mu = (x^0, x^1)$ (coordenadas contravariantes) y en la base recíproca $x_\mu = (x_0, x_1)$ (coordenadas covariantes) es:

$$x_\mu = g_{\mu\nu} x^\nu$$

es decir:

$$x_0 = x^0, \quad x_1 = -x^1$$

El vector posición es un vector contravariante de forma natural, escrito en la base de partida. Sin embargo consideremos el siguiente objeto:

$$\partial_\mu = (\partial_0, \partial_1)$$

donde:

$$\partial_\mu = \frac{\partial}{\partial x^\mu}$$

El colocar los índices como subíndices viene dado por la siguiente ley de transformación. Supongamos que cambiamos de sistema inercial de acuerdo con la expresión:

$$x'^\mu = \Lambda^\mu{}_\nu x^\nu$$

donde Λ es una de las transformaciones admisibles de las que hablamos antes. Entonces:

$$\frac{\partial}{\partial x'^\mu} = \frac{\partial x^\nu}{\partial x'^\mu}\frac{\partial}{\partial x^\nu} = (\Lambda^{-1})^\nu{}_\mu \frac{\partial}{\partial x'^\nu}$$

que es, como hemos dicho antes, la transformación de las coordenadas covariantes.

Esta expresión se puede escribir también de la siguiente forma. De acuerdo con la restricción que verifica Λ:

$$\Lambda^t K \Lambda = K$$

de donde:

$$(\Lambda^{-1})^t = K\Lambda K^{-1}$$

o, en coordenadas:

$$(\Lambda^{-1})^\nu{}_\mu = g_{\mu\rho}\Lambda^\rho_\sigma g^{\rho\sigma} = \Lambda_\mu{}^\nu$$

y, por tanto, la transformación es:

$$\frac{\partial}{\partial x'^\mu} = \Lambda_\mu{}^\nu \frac{\partial}{\partial x'^\nu}$$

De esta forma, podemos decir que el vector posición se transforma de manera contravariante, mientras el vector gradiente lo hace de forma covariante.

Volveremos más tarde a este ejemplo.

7.4 Espacios vectoriales y sus duales

No siempre dispondremos de un producto escalar en un espacio vectorial para poder definir coordenadas contravariantes y covariantes. Pero es muy frecuente en el estudio de lo que llamaremos tensores, la aparición de formas lineales (es decir, de aplicaciones del espacio vectorial en el cuerpo). Supongamos que tenemos una aplicación multilineal de un producto $V \times V \times \cdots \times V \times V^* \times V^* \times \cdots \times V^*$ en el cuerpo $\mathbb{K}$ donde está construido V, siendo V^* el espacio dual de V.

> Nota. El espacio final puede ser otro espacio vectorial y no necesariamente el cuerpo $\mathbb{K}$. Pero aquí nos limitaremos a este caso.

Según hemos visto antes, seleccionando bases en los espacios vectoriales que forman el producto cartesiano, es posible definir un conjunto de coordenadas asociada a la aplicación multilineal (o forma multilineal, si se quiere). En este caso particular que estamos tratando, consideremos una base $\mathcal{B} = \{u_1, \ldots, u_n\}$ de V y su base dual en V^*: $\mathcal{B}^* = \{u^{*1}, \ldots, u^{*n}\}$ con la propiedad ya conocida:

$$u^{*i}(u_j) = \delta^i{}_j$$

similar a la que usamos para definir las bases recíprocas. Nótese que aquí las bases $\mathcal{B}$ y $\mathcal{B}^*$ lo son de diferentes espacios.

Supongamos que hacemos simultáneamente los cambios de base:

$$u_i = \sum_{j=1}^n P^j{}_i u'_j, \quad u^{*i} = \sum_{j=1}^n (P^{-1})^i{}_j u'^{*j}$$

lo que asegura que las nuevas bases también son duales una de la otra, como ya sabemos:

$$\begin{aligned} u'^{*i}(u'_j) &= \sum_{k=1}^n P^i{}_k u^{*k}(\sum_{l=1}^n (P^{-1})^l{}_j u_l) = \sum_{k=1}^n P^i{}_k \sum_{l=1}^n (P^{-1})^l{}_j u^{*k}(u_l) \\ &= \sum_{k=1}^n \sum_{l=1}^n P^i{}_k (P^{-1})^l{}_j \delta^k_l = \sum_{k=1}^n P^i{}_k (P^{-1})^k{}_j = \delta^i{}_j \end{aligned}$$

Por tanto, las coordenadas contravariantes (asociadas a vectores del espacio inicial) se transforman con P mientras que las covariantes (asociadas a vectores del espacio dual) se transforman con la transpuesta inversa de P.

¿Cómo podemos relacionar esta definición con la dada anteriormente para las bases recíprocas? Supongamos que tenemos en V un producto escalar (necesario para poder definir la base recíproca). Sea $\mathcal{B}_r = \{u^1, \ldots, u^n\}$ la base recíproca de $\mathcal{B}$. Según el teorema de Riesz-Fréchet, si ω es una forma lineal, existe un único vector x_ω de V tal que:

$$\omega(y) = (x_\omega, y), \quad \forall y \in V$$

Dada una forma de la base dual, u^{*i}, veamos cual es su correspondiente vector en V:

$$u^{*i}(y) = (v^i, y), \quad \forall y \in V$$

Usando $y = u_k$:

$$u^{*i}(u_k) = (v^i, u_k) = \delta^i{}_k$$

que es justamente la definición de la base recíproca, por lo que la correspondencia (el isomorfismo entre el espacio dual y el espacio de partida proporcionado por el teorema de Riesz-Fréchet) es:

$$u^{*i} \longrightarrow u^i$$

Es decir, las coordenadas (covariantes) de una forma en el espacio dual, son las coordenadas covariantes del vector correspondiente (según Riesz-Fréchet) en el espacio de partida (en la base recíproca). Si hay un producto escalar, ambos conceptos coinciden. En el caso de que no lo haya, se entenderán las coordenadas covariantes como las de las formas en la base dual. Como hemos dicho antes, este resultado se extiende al caso en el que tengamos una forma bilineal simétrica no degenerada (como en relatividad).

7.5 Producto tensorial

Introducimos en esta sección una definición formal de producto tensorial. Los conceptos que aquí aparecen son de dificultad superior al resto de los temas, por lo que pueden suprimirse en una primera lectura.

7.5.1 Definición de producto tensorial

En esta sección discutiremos la relación entre aplicaciones multilineales y tensores, basándonos en una definición más rigurosa del concepto de tensor.

Sean $V_1, \ldots, V_n$ espacios vectoriales de dimensión finita sobre un cuerpo $\mathbb{K}$. Se consideran las aplicaciones multilineales:

$$\varphi: V_1 \times \cdots \times V_n \to W$$

donde W es otro espacio vectorial sobre $\mathbb{K}$. Es decir, se tienen los pares (φ, W) formados por una aplicación multilineal y el espacio de llegada.

Dadas dos parejas (φ, W) y (φ', W'), se puede estudiar si tiene solución el siguiente problema: encontrar una aplicación lineal f de W en W' tal que: $f \circ \varphi = \varphi'$:

$$\begin{array}{ccc} V_1 \times \cdots \times V_n & \xrightarrow{\varphi} & W \\ & \searrow & \downarrow \\ & & W' \end{array}$$

La respuesta es que no siempre es así. Sin embargo, se puede demostrar que existe una pareja $(\psi\ \mathcal{T})$, que verifica: dada cualquier aplicación multilineal, $\varphi: V_1 \times \cdots \times V_n \to W$, existe una única aplicación lineal, $\varphi_*: \mathcal{T} \to W$, tal que:

$$\varphi_* \circ \psi = \varphi$$

A esta aplicación multilineal, ψ (o al espacio $\mathcal{T}$) se le llama producto tensorial de los espacios $V_1, \ldots, V_n$. Aunque no entraremos en detalles, el espacio tensorial es único (salvo isomorfismo).

La correspondencia entre las aplicaciones multilineales φ y las aplicaciones lineales φ_* es única. Por eso, el espacio producto tensorial sustituye en cierto sentido al espacio producto cartesiano y transforma las aplicaciones multilineales de $V_1 \times \cdots \times V_n$ en W en aplicaciones lineales de $\mathcal{T}$ en W.

7.5.2 Construcción del producto tensorial

En esta sección construiremos explícitamente un modelo de producto tensorial. Para ello, se considera el espacio vectorial libre sobre el conjunto de generadores $V_1 \times \cdots \times V_n$. Es decir, todos los elementos de este conjunto se consideran linealmente independientes y, usando el cuerpo $\mathbb{K}$, se construyen sus combinaciones lineales para dar lugar a un espacio vectorial, que llamaremos M. Para aclarar la situación, si $(x_1, \ldots, x_n)$ e $(y_1, \ldots, y_n)$ son elementos de $V_1 \times \cdots \times V_n$, entonces, $(x_1, \ldots, x_n)$, $(y_1, \ldots, y_n)$ y $(x_1 + y_1, \ldots, x_n + y_n)$ son elementos de M, sin que el último sea la suma de los dos primeros, siendo por tanto diferente de $(x_1, \ldots, x_n) + (y_1, \ldots, y_n)$.

Dado que no utilizaremos esta notación en el futuro, no insistiremos aquí sobre ella. Baste decir que tendremos que indicar si trabajamos en el producto cartesiano o en M, para saber como son las propiedades de sus elementos, ya que los escribimos igual.

Consideremos en M el subespacio vectorial generado por todos los elementos de M de la forma:

$$(x_1, \ldots, x_i + x_i', \ldots, x_n) - (x_1, \ldots, x_i, \ldots, x_n) - (x_1, \ldots, x_i', \ldots, x_n)$$

$$(x_1, \ldots, \lambda x_i, \ldots, x_n) - \lambda(x_1, \ldots, x_i, \ldots, x_n)$$

y el espacio vectorial cociente M/N.

La inclusión $i: V_1 \times \cdots \times V_n \to M$ no es una aplicación lineal, pero la composición de i con la proyección: $\pi: M \to M/N$, $\psi = \pi \circ i$ es una aplicación multilineal:

$$\psi(x_1, \ldots, x_n) = [x_1, \ldots, x_n] \in M/N$$

En efecto:

$$\begin{aligned} \psi(x_1, \ldots, x_i + x_i', \ldots, x_n) &= [x_1, \ldots, x_i + x_i', \ldots, x_n] = [x_1, \ldots, x_i, \ldots, x_n] + [x_1, \ldots, x_i', \ldots, x_n] \\ &= \psi(x_1, \ldots, x_i, \ldots, x_n) + \psi(x_1, \ldots, x_i', \ldots, x_n) \end{aligned}$$

$$\psi(x_1, \ldots, \lambda x_i, \ldots, x_n) = [x_1, \ldots, \lambda x_i, \ldots, x_n] = \lambda[x_1, \ldots, x_i, \ldots, x_n] = \lambda\psi(x_1, \ldots, x_i, \ldots, x_n)$$

Ya tenemos un par $(\psi M/N\ )$ de los que estamos estudiando. Consideremos otro par cualquiera (φ, W) y busquemos una aplicación lineal $\varphi_*: M/N \to W$, que compuesta con ψ dé φ. La aplicación multilineal φ asigna a cada elemento del espacio producto cartesiano $V_1 \times \cdots \times V_n$ un elemento del espacio W. Por tanto, como los elementos del producto cartesiano son un sistema de generadores del espacio M, podemos definir una aplicación h de $V_1 \times \cdots \times V_n$ (como sistema de generadores de M) en W, con valores iguales a los dados por φ, y extenderla de forma lineal a todo M. Evidentemente:

$$h \circ i = \varphi.$$

La aplicación h, siendo lineal, vale cero sobre los elementos de N, y por consiguiente se puede factorizar a través del espacio cociente M/N, por una aplicación, también lineal, φ_*:

$$h = \varphi_* \circ \pi$$

de manera única.

Como $\psi = \pi \circ i$, se tiene:

$$\varphi = h \circ i = (\varphi_* \circ \pi) \circ i = \varphi_* \circ \psi$$

que es lo que queríamos probar (la unicidad de φ_* proviene del hecho de que la imagen de ψ genera todo el espacio cociente M/N).

Se llama a M/N el espacio producto tensorial de los espacios vectoriales $V_1, \ldots, V_n$:

$$V_1 \otimes \cdots \otimes V_n$$

y a los elementos de este espacio se les llama tensores y se escribe:

$$\psi(x_1, \ldots, x_n) = x_1 \otimes \cdots \otimes x_n.$$

Nótese que no todos los elementos del espacio producto tensorial se escriben de esta forma. Lo que es cierto es que los tensores de este tipo generan (linealmente) todo el espacio producto tensorial, con lo que un tensor arbitrario es combinación lineal de elementos de este tipo, por ejemplo expresiones como:

$$x \otimes y + z \otimes v$$

Las propiedades más elementales de este símbolo ($\otimes$) son:

$$x \otimes (y + z) = x \otimes y + x \otimes z$$

$$(x + y) \otimes z = x \otimes z + y \otimes z$$

$$\lambda(x \otimes y) = (\lambda x) \otimes y = x \otimes (\lambda y)$$

7.5.3 Propiedades del producto tensorial

Si $V_1, \ldots, V_n$ son espacios vectoriales de dimensiones $d_1, \ldots, d_n$, entonces el producto tensorial es un espacio vectorial de dimensión:

$$d_1 \times \cdots \times d_n$$

Nótese la diferencia con el producto cartesiano, donde la dimensión es la suma de las dimensiones.

La demostración se basa en un estudio del producto tensorial con el cuerpo $\mathbb{K}$. Sea V un espacio vectorial sobre $\mathbb{K}$ y construyamos el producto tensorial: $V \otimes \mathbb{K}$. A cada elemento de V se le hace corresponder un elemento de $\mathbb{K}$:

$$x \in V \longrightarrow x \otimes 1$$

Se trata de una aplicación lineal que es además un isomorfismo. Una base de $V \otimes \mathbb{K}$ está formada por los tensores $u_i \otimes 1$ donde los vectores de V, u_i, forman una base de V. Por lo tanto la dimensión de $V \otimes \mathbb{K}$ es igual a la dimensión de V.

La generalización de este resultado es la fórmula anterior para el cálculo de la dimensión de un producto tensorial.

Si $\mathcal{B}_k = \{u_i^{(k)}\}$ es una base de V_k, entonces:

$$\{u_{i_1}^{(1)} \otimes \cdots \otimes u_{i_n}^{(n)}\}$$

es una base del producto tensorial. Por tanto, cualquier tensor puede ponerse como:

$$t = \sum_{i_1=1,\ldots,i_n=1}^{d_1,\ldots,d_n} t^{i_1\ldots i_n} u_{i_1}^{(1)} \otimes \cdots \otimes u_{i_n}^{(n)}$$

Si se hace un cambio de base en cada uno de los espacios V_i dado por:

$$u_i^{(k)} = \sum_{j=1}^{d_k} P_{ji}^{(k)} u_j'^{(k)}$$

el cambio en las coordenadas del tensor es:

$$t'^{i_1 \dots i_n} = \sum_{j_1=1,\dots,j_n=1}^{d_1,\dots,d_n} P_{i_1 j_1}^{(1)} \dots P_{i_n j_n}^{(n)} t^{j_1 \dots j_n}$$

La propiedad que más nos interesa aquí para relacionar los tensores con las aplicaciones multilineales es la siguiente. Sean V_1, V_2, V_3 tres espacios vectoriales sobre un cuerpo $\mathbb{K}$. Se tiene:

$$\mathcal{L}(V_1, \mathcal{L}(V_2, V_3)) \approx \mathcal{L}^2(V_1, V_2; V_3) \approx \mathcal{L}(V_1 \otimes V_2, V_3)$$

donde $\mathcal{L}$ son las aplicaciones lineales entre los dos espacios que se muestran y $\mathcal{L}^2$ son las aplicaciones bilineales del producto cartesiano de los dos primeros espacios en el tercero. Demostremos esta propiedad.

La primera propiedad es muy sencilla. Sea:

$$\varphi: V_1 \times V_2 \to V_3$$

una aplicación bilineal. Si fijamos $x \in V_1$, podemos construir una aplicación lineal:

$$\tilde{\varphi}: V_1 \to \mathcal{L}(V_2, V_3), \quad \tilde{\varphi}(x) = \varphi_x$$

mediante:

$$\varphi_x: V_2 \to V_3, \quad \varphi_x(y) = \varphi(x, y)$$

Entonces, a cada aplicación bilineal φ le hacemos corresponder la aplicación lineal $\tilde{\varphi}$. Y viceversa, consideremos una aplicación lineal:

$$\tilde{\phi}: V_1 \to \mathcal{L}(V_2, V_3)$$

Definimos un aplicación bilineal de $V_1 \times V_2$ en V_3:

$$\phi: V_1 \times V_2 \to V_3, \quad \phi(x, y) = \tilde{\phi}(x)(y)$$

Esta es la correspondencia inversa de la anterior y por tanto, se da el isomorfismo del que se hablaba.

En cuanto al otro isomorfismo, el que relaciona las aplicaciones bilineales con aplicaciones lineales del producto tensorial en un espacio vectorial, se trata simplemente de asociar a cada aplicación bilineal de $V_1 \times V_2$ en V_3 la aplicación lineal dada por el carácter universal del producto tensorial:

$$\begin{array}{ccc} V_1 \times V_2 & \longrightarrow & V_1 \otimes V_2 \\ & \searrow & \downarrow \\ & & V_3 \end{array}$$

Evidentemente los resultados se generalizan a un número arbitrario de espacios y a aplicaciones multilineales.

7.6 Tensores y aplicaciones multilineales

Hemos visto en las secciones precedentes como a cada aplicación lineal del producto tensorial en un espacio vectorial se le puede asignar una aplicación multilineal del producto cartesiano en el espacio en cuestión. Es decir, los tensores no son aplicaciones multilineales, ni se puede establecer un isomorfismo entre estos dos espacios vectoriales de forma inmediata.

Sin embargo, supongamos que el espacio de llegada de las aplicaciones multilineales en cuestión, es el cuerpo $\mathbb{K}$. Lo que realmente se tiene es un isomorfismo entre las formas multilineales y las formas lineales del espacio producto tensorial. Como los espacios vectoriales (siempre de dimensión finita) y sus espacios duales son isomorfos (aunque no sea de forma canónica) se puede establecer una relación entre tensores y formas multilineales. Si además tenemos un producto escalar, el teorema de Riesz-Fréchet permite establecer este isomorfismo de forma canónica (lo que es también cierto aunque la forma bilineal que genera el producto escalar no sea definida positiva, aunque sí no degenerada).

Dado un tensor $x_1 \otimes \cdots \otimes x_n$ en el espacio producto tensorial $V_1 \otimes \cdots \otimes V_n$, y una forma bilineal simétrica no degenerada en cada uno de los espacios V_k, φ_k, con coordenadas en las bases $\mathcal{B}_k$ de V_k dadas por:

$$\varphi_k(u_i^{(k)}, u_j^{(k)}) = g_{ij}^{(k)}$$

definimos una forma multilineal:

$$\phi\colon V_1 \times \cdots \times V_n \to \mathbb{K}$$

asociada a este tensor mediante:

$$\phi(y_1, \ldots, y_n) = \varphi_1(x_1, y_1) \cdots \varphi(x_n, y_n)$$

y podemos trabajar indistintamente con tensores o formas multilineales. Veamos la relación entre las coordenadas de ϕ y el tensor $t = x_1 \otimes \cdots \otimes x_n$ en unas bases dadas de $V_1, \ldots, V_n$.

Supongamos que elegimos como tensores t los de una base del espacio tensorial:

$$t_{(i_1 \ldots i_n)} = u_{i_1}^{(1)} \otimes \cdots \otimes u_{i_n}^{(n)}$$

La forma bilineal correspondiente aplicada a n vectores de las bases anteriores es:

$$\phi_{(i_1 \ldots i_n)}(u_{j_1}^{(1)}, \ldots, u_{j_n}^{(n)}) = \varphi_1(u_{i_1}^{(1)}, u_{j_1}^{(1)}) \cdots \varphi(u_{i_n}^{(n)}, u_{j_n}^{(n)}) = g_{i_1 j_1}^{(1)} \ldots g_{i_n j_n}^{(n)}$$

7.7 Cambios de base

Supongamos que en cada uno de los espacios V_i hacemos un cambio de base como dijimos antes:

$$u_i^{(k)} = \sum_{j=1}^{d_k} P_{ji}^{(k)} u_j'^{(k)}$$

Veamos como cambian las formas multilineales y como cambian los tensores. Sea:

$$\varphi\colon V_1 \times \cdots \times V_n\colon \to \mathbb{K}$$

una forma multilineal, y sea su expresión en la base de partida:

$$\varphi(x_1, \ldots, x_n) = \sum_{i_1, \cdots, i_n = 1}^{d_1, \ldots, d_n} \varphi_{i_1 \ldots i_n} x_{i_1}^{(1)} \cdots x_{i_n}^{(n)}$$

donde:

$$x_k = \sum_{i_k=1}^{d_k} x_{i_k}^{(k)} u_{i_k}^{(k)}, \quad \varphi(u_{i_1}^{(1)}, \ldots, u_{i_n}^{(n)}) = \varphi_{i_1 \ldots i_n}$$

Bajo el cambio de base,

$$\varphi(x_1, \ldots, x_n) = \sum_{i_1, \cdots, i_n = 1}^{d_1, \ldots, d_n} \varphi_{i_1 \ldots i_n} x_{i_1}^{(1)} \cdots x_{i_n}^{(n)} = \sum_{i_1, \cdots, i_n = 1}^{d_1, \ldots, d_n} \varphi_{i_1 \ldots i_n} \sum_{j_1=1}^{d_1} (P^{(1)})^{-1}_{i_1 j_1} x_{j_1}'^{(1)} \cdots \sum_{j_n=1}^{d_n} (P^{(n)})^{-1}_{i_n j_n} x_{j_n}'^{(n)}$$

y, por lo tanto:

$$\varphi'_{i_1 \ldots i_n} = \sum_{i_1, \cdots, i_n = 1}^{d_1, \ldots, d_n} \varphi_{i_1 \ldots i_n} (P^{(1)})^{-1}_{j_1 i_1} \cdots (P^{(n)})^{-1}_{j_n i_n} \varphi^{(n)}_{j_1 \ldots j_n}$$

Si recordamos como cambian las coordenadas de un tensor $t \in V_1 \otimes \cdots \otimes V_n$:

$$t'^{i_1 \dots i_n} = \sum_{j_1,\dots,j_n=1}^{d_1,\dots,d_n} P^{(1)}_{i_1 j_1} \dots P^{(n)}_{i_n j_n} t^{j_1 \dots j_n}$$

vemos como unas lo hacen con la matriz de cambio de base y otras con la inversa traspuesta, como era de esperar dada la dualidad entre estos dos objetos. Si las matrices son ortogonales, el cambio es el mismo (pues $(P^{-1})^t = P$). En este sentido, las aplicaciones multilineales (con valores en el cuerpo) son una de las formas posibles de representar los tensores.

7.8 Deﬁnición de tensores bajo transformaciones

Definición 7.8.1 (Tensores en un espacio vectorial V) *Un tensor de rango $p = r + s$, r veces contravariante y s veces covariante, definido sobre un $\mathbb{K}$-espacio vectorial V de dimensión n, es un objeto con n^{r+s} componentes referidas a una base de V, $\mathcal{B}$ y a la base dual en V^*, que se escribirá como:*

$$t^{i_1 \dots i_r}_{j_1 \dots j_s}$$

y que están relacionadas con las componentes en otra base mediante la expresión:

$$t'^{i_1 \dots i_r}_{j_1 \dots j_s} = P^{i_1}_{k_1} \cdots P^{i_r}_{k_r} (P^{-1})^{l_1}_{j_1} \cdots (P^{-1})^{l_s}_{j_s} t^{k_1 \dots k_r}_{l_1 \dots l_s}$$

donde se emplea el convenio de Einstein de la suma.

La matriz P es la matriz de cambio de base, de $\mathcal{B}$ a una nueva base $\mathcal{B}'$ (y las duales en las coordenadas covariantes).

El orden de los índices es fundamental (salvo en casos de simetría como veremos más tarde). Suponemos que los índices contravariantes van antes que los covariantes, cuando los escribimos como hemos hecho arriba. Sin embargo, cuando hablemos de subir y bajar índices, hay que prestar atención a la posición relativa de unos y otros.

Ejemplo 7.8.1 Los tensores de rango 1 y de tipo $(1, 0)$ (es decir, una vez contravariante) son simplemente los vectores del espacio V. La ley de transformación es la que conocemos al cambiar de base:

$$t'^i = P^i_k t^k$$

Los tensores de rango 1 y tipo $(0, 1)$ (es decir 1 vez covariante) son los elementos del dual:

$$t'_i = (P^{-1})^k_i t_k$$

Los tensores de tipo $(0, 2)$ son formas cuadráticas (del tipo del tensor de inercia) y su regla de transformación es:

$$t'_{j_1 j_2} = (P^{-1})^{l_1}_{j_1} (P^{-1})^{l_2}_{j_2} t_{l_1 l_2}$$

Si representamos el tensor $t_{j_1 j_2}$ por una matriz T, la ley de transformación es simplemente:

$$T' = (P^{-1})^t T (P^{-1})$$

o, llamando $Q = P^{-1}$:

$$T' = Q^t T Q$$

una expresión bien conocida.

Los tensores de tipo $(1, 1)$ son, en un sentido que se puede precisar completamente, isomorfos a endomorfismos. Aquí nos limitaremos a mostrar la ley de transformación:

$$t'^i_j = P^i_k (P^{-1})^l_j t^k_l$$

que, escrita como antes en términos de matrices, se lee:

$$T' = P T P^{-1}$$

también conocida.

Ejemplo 7.8.2 En todo espacio de tensores de tipo $(1,1)$ existe un tensor definido en cualquier base por:

$$\begin{cases} \delta^{i_1}_{i_2} = 1 & i_1 = i_2 \\ \delta^{i_1}_{i_2} = 0 & i_1 \neq i_2 \end{cases}$$

Veamos que es un tensor:

$$\delta'^{i_1}_{i_2} = P^{i_1}_{j_1}(P^{-1})^{j_2}_{i_2}\delta^{j_1}_{j_2} = P^{i_1}_{j_1}(P^{-1})^{j_1}_{i_2} = \delta^{i_1}_{i_2}$$

Se llama el tensor de Kronecker (delta de Kronecker) y aparece constantemente en las operaciones con tensores.

Sea $\mathcal{T}^r_s$ el conjunto de tensores de rango $p = r + s$, r veces contravariante y s veces covariantes. Se suele escribir también:

$$\mathcal{T}^r_s = V \otimes \cdots \otimes V \otimes V^* \otimes \cdots \otimes V^*$$

y se llama a este conjunto el producto tensorial de r veces el espacio V y s veces el espacio V^*. El símbolo $\otimes$ se llama producto tensorial, como ya hemos visto.

Pues bien, $\mathcal{T}^r_s$ es un espacio vectorial sobre $\mathbb{K}$ de dimensión n^{r+s}. Las combinaciones lineales de tensores de tipo (r,s) (donde la suma se entiende componente a componente y el producto por escalares multiplicando cada componente por el escalar) son también tensores de este tipo (como es fácil de demostrar). En cuanto a la dimensión, basta considerar tensores que tienen en una base dada todas sus componentes igual a cero salvo una igual a 1. Hay evidentemente n^{r+s} tensores de este tipo, son linealmente independientes y generan todo el espacio $\mathcal{T}^r_s$.

Sea $\mathcal{B} = \{u_1, \ldots, u_n\}$ una base de V y $\mathcal{B}^* = \{u^{*1}, \ldots, u^{*n}\}$ su base dual. Obviamente, el escalar $1 \in \mathbb{K}$ es una base de $\mathcal{T}^0_0 \approx \mathbb{K}$. La base $\mathcal{B}$ lo es de $\mathcal{T}^1_0 \approx V$ y $\mathcal{B}^*$ de $\mathcal{T}^0_1 \approx V^*$.

En un espacio $\mathcal{T}^r_s$ la base, que contiene n^{r+s} vectores (tensores de rango $p = r + s$), se designa por:

$$u_{i_1} \otimes \cdots u_{i_r} \otimes u^{*j_1} \otimes \cdots \otimes u^{*j_s}$$

y por tanto, un tensor se escribirá como:

$$t = t^{i_1 \ldots i_r}_{j_1 \ldots j_s} u_{i_1} \otimes \cdots u_{i_r} \otimes u^{*j_1} \otimes \cdots \otimes u^{*j_s}$$

De esta forma, la ley de transformación se explica como un cambio de base simultáneo en cada una de las copias de los espacios V y V^*.

7.9 Propiedades de los tensores

Estudiaremos en esta sección tres propiedades de los tensores de rango arbitrario en un espacio vectorial V.

7.9.1 Tensores simétricos y antisimétricos

Sea $t^{i_1 \ldots i_r}_{j_1 \ldots j_s}$ un tensor de tipo (r,s), y consideremos el objeto que se define a partir de él mediante una permutación de los índices contravariantes entre sí (o de los covariantes entre sí):

$$\hat{t}^{i_1 \ldots i_r}_{j_1 \ldots j_s} = t^{\sigma(i_1 \ldots i_r)}_{j_1 \ldots j_s}$$

donde σ es un elemento del grupo de permutaciones S_r.:

$$\sigma(i_1 \ldots i_r) = (i_{\sigma(1)} \ldots i_{\sigma(r)})$$

Este objeto $\hat{t}$ es también un tensor del mismo tipo que t, y en general distinto de t. La demostración consiste simplemente en comprobar sus propiedades de transformación, aunque es bastante laboriosa de escribir. Veámosla en un caso sencillo, para tensores de tipo $(2,0)$, y en el caso no trivial, cuando

$$\sigma(1) = 2, \ \sigma(2) = 1$$

Entonces:

$$\hat{t}^{i_1 i_2} = t^{i_{\sigma(1)} i_{\sigma(2)}} = t^{i_2 i_1}$$

y la regla de transformación es:

$$\hat{t}'^{i_1 i_2} = t'^{i_2 i_1} = P^{i_2}_{j_2} P^{i_1}_{j_1} t^{j_2 j_1} = P^{i_1}_{j_1} P^{i_2}_{j_2} \hat{t}^{j_1 j_2}$$

lo que demuestra que es un tensor de tipo $(2,0)$.

Igual se hace para índices covariantes. Lo que no se puede hacer es permutar un índice contravariante con otro covariante. Estos índices están referidos a distintas bases (una y su dual) en diferentes espacios y no tiene sentido la operación, no obteniéndose un tensor.

Esta operación da lugar a las siguientes definiciones.

Definición 7.9.1 *Se dice que el tensor $t^{i_1 \cdots i_r}$ de tipo $(r,0)$ es un tensor simétrico, si para toda operación del grupo S_r sobre sus índices contravariantes, el tensor resultante es igual al original.*

Evidentemente, la misma definición se puede establecer para un tensor de tipo $(0,s)$, e incluso para tensores mixtos (tipo (r,s)), refiriéndose a cada conjunto de índices contravariantes y covariantes.

La suma de tensores simétricos de tipo $(r,0)$ y el producto por escalares de estos tensores lleva a un tensor del mismo tipo. El conjunto de tensores simétricos es un subespacio vectorial del espacio de tensores. Si la dimensión del espacio de tensores $\mathcal{T}^r$ era n^r, no es difícil probar que la dimensión del subespacio de tensores simétricos $\mathcal{S}^r$ es justamente:

$$\binom{n+r-1}{r}$$

Basta simplemente contar las coordenadas independientes.

Ejemplo 7.9.1 Consideremos los tensores simétricos de tipo $(2,0)$. La ecuación que verifican es:

$$t^{i_1 i_2} = t^{i_2 i_1}$$

Si el espacio es $\mathbb{R}^3$, la dimensión de $\mathcal{T}^2$ es 9, y la de estos tensores simétricos es 6.

Los tensores simétricos de tipo $(3,0)$ verifican:

$$t^{i_1 i_2 i_3} = t^{i_2 i_1 i_3} = t^{i_3 i_2 i_1} = t^{i_1 i_3 i_2} = t^{i_3 i_1 i_2} = t^{i_2 i_3 i_1}$$

En este caso, si el espacio es $\mathbb{R}^3$ forman un subespacio de dimensión 10. Y los de tipo $(4,0)$ tienen dimensión 15.

Las formas bilineales simétricas son un ejemplo de tensores de tipo $(0,2)$ simétricos (en particular, el producto escalar).

De forma similar se definen los tensores totalmente antisimétricos.

Definición 7.9.2 *Se dice que el tensor $t^{i_1 \cdots i_r}$ de tipo $(r,0)$ es un tensor totalmente antisimétrico (o alternado), si, para toda operación del grupo S_r sobre sus índices contravariantes, se tiene:*

$$t^{i_{\sigma(1)} \cdots i_{\sigma(r)}} = (-1)^{\varepsilon(\sigma)} t^{i_1 \ldots i_r}$$

donde $\varepsilon(\sigma)$ es la paridad de la permutación.

Los tensores antisimétricos forman también un espacio vectorial y la dimensión es:

$$\binom{n}{r}$$

Veamos algunos ejemplos.

Ejemplo 7.9.2 El conjunto de tensores antisimétricos de tipo $(2,0)$ verifican:

$$t^{i_1 i_2} = -t^{i_2 i_1}$$

En $\mathbb{R}^3$ tienen dimensión 3. Nótese que $3+6=9$. De hecho se tiene en cualquier espacio V:

$$\mathcal{T}^2 = \mathcal{S}^2 \oplus \mathcal{A}^2$$

La dimensión de los tensores antisimétricos de tipo $(3,0)$ en $\mathbb{R}^3$ es 1. Es decir, salvo un factor, solo hay un tensor de tipo $(3,0)$ totalmente antisimétrico. Esto también es generalizable. En un espacio de dimensión n sólo existe (salvo un factor) un tensor de tipo $(n,0)$ totalmente antisimétrico. Y no existe ningún tensor de tipo $(r,0)$, con $r>n$, totalmente antisimétrico en un espacio de dimensión n.

Veamos como se transforma un tensor antisimétrico de orden n:

$$t^{i_1 \ldots i_n} = a\epsilon^{i_1 \ldots i_n}$$

donde $a \in \mathbb{K}$ y $\epsilon^{i_1 \ldots i_n}$ verifica:

$$\epsilon^{1 \ldots n} = -1$$

en una base dada. Veamos como cambia este tensor ϵ al cambiar la base.

$$\epsilon'^{i_1 \ldots i_n} = P^{i_1}_{j_1} \cdots P^{i_n}_{j_n} \epsilon^{j_1 \ldots j_n}$$

y por tanto,

$$\epsilon'^{1 \ldots n} = P^{1}_{j_1} \cdots P^{n}_{j_n} \epsilon^{j_1 \ldots j_n} = -\det P$$

Por lo tanto, estos tensores se transforman con el determinante. Veremos más adelante una aplicación de este resultado.

Nótese también que el caso de tensores de segundo orden es el único en el que se verifica que todo tensor es suma de un tensor simétrico más otro antisimétrico. En órdenes mayores esto es incorrecto. La razón es que existen tensores con simetrías intermedias (por ejemplo, simétricos en dos índices, pero no en el resto,y verificando relaciones más complicadas entre las coordenadas). El espacio total se descompone en suma directa de estos tensores de simetrías intermedias en los que no entraremos aquí, pero que juegan un papel fundamental, por ejemplo en la teoría de representaciones de grupos.

7.9.2 Contracción de índices

La segunda operación que vamos a considerar afecta a tensores de tipo mixto, y contrariamente a la primera no actúa en el espacio de tensores de tipo (r,s) sino que pasa de $\mathcal{T}^r_s$ a $\mathcal{T}^{r-1}_{s-1}$. Desde un punto de vista más riguroso, habría que considerar el conjunto:

$$\mathcal{T} = \bigoplus_{r,s=0}^{\infty} \mathcal{T}^r_s$$

el álgebra tensorial, pero no entraremos aquí en estos detalles.

Sea $t^{i_1 \ldots i_r}_{j_1 \ldots j_s}$ un tensor de tipo (r,s) y definamos:

$$\hat{t}^{i_1 \ldots i_{r-1}}_{j_1 \ldots j_{s-1}} = t^{i_1 \ldots i_{k-1} i i_k \ldots i_{r-1}}_{j_1 \ldots j_{l-1} i j_l \ldots j_{s-1}}$$

Se dice que el tensor $\hat{t}$ se ha obtenido por contracción de dos de los índices del tensor t. Obviamente la operación se puede extender a más índices y contraerlos sucesivamente. Nótese que la contracción se hace con índices contravariantes y covariantes. No se contraen índices del mismo tipo. Desde el punto de vista del espacio y su dual lo que se está haciendo es actuar con uno de los espacios V^* sobre uno de los V. Que el resultado de esta operación es un tensor es también consecuencia de la ley de transformación.

Ejemplo 7.9.3 Consideremos un tensor de tipo $(1,1)$, t^i_j. Contrayendo sus dos índices, obtenemos otro tensor, de tipo $(0,0)$ es decir, un escalar:

$$\operatorname{tr} t = t^i_i$$

Se le llama la traza del tensor t. En particular, para el tensor de Kronecker, la traza es la dimensión del espacio V.

Se dice que un tensor de tipo $(1,1)$ es de traza nula si $\operatorname{tr} t = 0$. Todo tensor de tipo $(1,1)$ se puede escribir de la siguiente forma (de manera única):

$$t^i_j = a^i_j + (\operatorname{tr} t)\delta^i_j$$

donde a^i_j es un tensor de traza nula.

Ejemplo 7.9.4 Sea el tensor $R^\lambda_{\mu\nu\rho}$ de tipo $(1,3)$. Un tensor contraído a partir de él es:

$$r_{\mu\nu} = R^\lambda_{\lambda\mu\nu}$$

Nótese que, en principio:

$$R^\lambda_{\lambda\mu\nu} \neq R^\lambda_{\mu\lambda\nu}$$

Sin embargo, si el tensor R es totalmente simétrico en sus índices covariantes, se tiene:

$$R^\lambda_{\lambda\mu\nu} = R^\lambda_{\mu\lambda\nu} = R^\lambda_{\mu\nu\lambda}$$

7.9.3 Producto tensorial

La tercera operación con tensores a estudiar es el producto tensorial. Con ella, a partir de dos tensores de tipos (r,s) y (r',s') podemos construir un tercer tensor de tipo $(r+r', s+s')$. Sean:

$$a \in \mathcal{T}^r_s, \quad b \in \mathcal{T}^{r'}_{s'}$$

Entonces, el objeto definido por sus componentes en una base como:

$$c^{i_1\dots i_r i_{r+1}\dots i_{r+r'}}_{j_1\dots j_s j_{s+1}\dots j_{s+s'}} = a^{i_1\dots i_r}_{j_1\dots j_s} b^{i_{r+1}\dots i_{r+r'}}_{j_{s+1}\dots j_{s+s'}}$$

es un tensor que se llama el producto tensorial de los tensores a y b. La demostración es otra vez la ley de transformación al cambiar de base.

Esta operación permite establecer una aplicación entre el producto tensorial $\mathcal{T}^r_s \otimes \mathcal{T}^{r'}_{s'}$ y el espacio de tensores: $\mathcal{T}^{r+r'}_{s+s'}$; que esta aplicación es lineal e inyectiva no será desarrollado en detalle aquí.

Ejemplo 7.9.5 Consideremos dos tensores sobre un espacio vectorial V, a^i de tipo $(1,0)$ y b_i de tipo $(0,1)$. Construimos el producto tensorial de estos dos tensores, que es un tensor de tipo $(1,1)$:

$$t^i_j = a^i b_j$$

Si ahora contraemos los dos índices del tensor t, obtenemos un escalar:

$$c = a^i b_i$$

Según la idea de contracción que hemos introducido, no es posible contraer dos tensores de tipo $(1,0)$. Sin embargo, consideremos un tensor de tipo $(0,2)$, g_{ij}. Podemos hacer el producto tensorial de g por dos tensores de tipo $(1,0)$, x^i e y^i, es decir por dos elementos del espacio V de partida:

$$g_{ij} x^k y^l$$

Si ahora contraemos el índice i con el k y el j con el l:

$$g_{ij} x^i y^j$$

obtenemos un escalar. Si g es simétrico y definido positivo (es decir, la matriz que representa a este tensor de segundo orden es definida positiva), tenemos un producto escalar en V.

Dada una forma cuadrática no degenerada de determinante g, la ley de transformación para g es:

$$g' = (\det P)^2 g$$

o sea:

$$\sqrt{|g'|} = |\det P|\sqrt{|g|}$$

Por tanto, $\sqrt{|g|}$ no es un escalar. Tampoco es un tensor alternado de orden n para transformaciones generales, pero sí para aquellas que tienen el determinante positivo, pues entonces:

$$\sqrt{|g'|} = \det P\sqrt{|g|}$$

En este caso, $\sqrt{|g|}(e_1 \otimes e_2 \cdots \otimes e_n)_a$ es el elemento de volumen correspondiente a la métrica g, donde el subíndice a significa la antisimetrización de este tensor.

Finalmente consideremos el caso de un tensor de tipo $(1,1)$ multiplicado tensorialmente por otro de tipo $(1,0)$:

$$t^i_j x^k$$

y hagamos la única contracción posible (índices jk):

$$y^i = t^i_j x^j$$

que es de nuevo un tensor de tipo $(1,0)$. Esto sugiere la relación entre los tensores de tipo $(1,1)$ y los endomorfismos de V de la que hablábamos antes. Se tiene el isomorfismo de espacios vectoriales:

$$\mathcal{L}(V) \approx \mathcal{T}^1_1$$

Nótese que, usando el producto tensorial y tomando combinaciones lineales, uno puede construir cualquier tipo de tensor a partir de tensores de tipo $(1,0)$ y $(0,1)$ y escalares.

7.10 Tensores covariantes antisimétricos: formas

Los tensores covariantes totalmente antisimétricos se pueden interpretar como aplicaciones lineales del espacio vectorial de tensores contravariantes (del mismo orden) en el cuerpo $\mathbb{K}$. Si $t_{j_1\ldots j_s}$ es uno de estos tensores, al aplicarlo a un tensor como $x^{i_1\cdots i_s}$, es decir, al hacer el producto tensorial y luego contraer todos los índices, se obtiene un escalar:

$$t_{i_1\ldots i_s}x^{i_1\ldots i_s} \in \mathbb{K}$$

El conjunto de tensores covariantes de orden s totalmente antisimétricos es un subespacio vectorial del espacio $\mathcal{T}_s$ y se llama Λ_s, conjunto de formas de orden s. Un tensor de este tipo se llama una s-forma. La dimensión de estos espacios ya la hemos calculado antes, y también hemos visto cómo no hay s-formas con $s > n$. La suma directa de espacios vectoriales:

$$\bigoplus_{s=0}^{n} \Lambda_s$$

se llama álgebra exterior de V. Se toma $\Lambda_0 = \mathbb{K}$.

Cuando se trabaja con tensores alternados, se usa el símbolo $\wedge$ para describir el producto tensorial alternado.

Dentro de las relaciones que hemos visto hay entre aplicaciones multilineales y tensores, las formas son aplicaciones multilineales alternadas. La única forma de dimensión n (linealmente independiente) que hay, se le llama elemento de volumen (el determinante de la teoría de matrices) y se escribe:

$$e_1 \wedge \cdots \wedge e_n$$

En el álgebra exterior es posible definir otra operación , que relaciona los tensores alternados de tipo $(0,k)$ y los de tipo $(0,n-k)$ definidos sobre un espacio con una métrica g_{ij}. Si $\epsilon_{i_1,\ldots i_n}$ es el único tensor alternado de orden n, con valor $\epsilon_{12\ldots n}=-1$, se define:

$$(\star t)_{i_{k+1}\ldots i_n} = \frac{1}{k!}\sqrt{|g|}\epsilon_{i_1\ldots i_n} t^{i_1\ldots i_k}$$

Se tiene:

$$\star(\star t) = (-1)^{k(n-k)}\,\mathrm{sgn}(g) t$$

7.11 Tensores y grupos de transformaciones

Según hemos visto hasta ahora, los elementos que necesitamos para la construcción de tensores (al menos, de los que estamos hablando en estas últimas secciones), son: un espacio vectorial de dimensión finita y los cambios de base en dicho espacio. Las transformaciones que pasan de unas bases a otras, son las que forman el grupo general lineal del espacio V, $GL(V)$, es decir cualquier automorfismo de V. En términos de matrices, se trata de las matrices regulares $n\times n$, que forman el grupo general lineal $GL(n,\mathbb{K})$ (obviamente isomorfo al anterior $GL(V)$). Sin embargo, muchas veces, no interesa hacer cambios de base tan generales. Por ejemplo si tenemos un producto escalar, puede ser útil hacer cambios de base que conserven la orientación o los ángulos que forman los vectores de la base entre sí. O, en relatividad especial, las transformaciones que resultan aceptables son aquellas que conservan el intervalo espacio-temporal constante, pues son las que relacionan los sistemas inerciales entre sí. En general, el conjunto de transformaciones que se usan en cada caso, forman un grupo (pues si no, no se podrá hablar de transformaciones inversas etc.) y se dice que un tensor lo es bajo ese grupo de transformaciones. Por ejemplo, consideremos el espacio $\mathbb{R}^n$ con su producto escalar usual y las transformaciones ortogonales con determinante positivo. Los tensores de este espacio (es decir los objetos que cambian adecuadamente al cambiar la base mediante una rotación) son los tensores cartesianos de este espacio. Es posible que si empleamos otra transformación que no sea una rotación, el tensor no se comporte como tal.

Por ejemplo, pensemos en el tensor de Kronecker δ^i_j, cuyas componentes son las mismas en cualquier base. Si queremos definir un tensor con esa misma propiedad, pero de tipo $(0,2)$, es decir g_{ij}, con componentes iguales a 1 si los índices son iguales y a 0 si los índices son distintos, vemos que podemos hacerlo para cualquier base:

$$g'_{ij} = (P^{-1})^k_i (P^{-1})^l_j g_{kl}$$

o escritos como matrices (supongamos que el índice contravariante numera a las filas):

$$G' = (P^{-1})^t G (P^{-1})$$

Si queremos que tanto G como G' sean la matriz identidad, las matrices que permiten pasar de unas bases a otras, verifican:

$$P^tP = PP^t = I$$

es decir, son las matrices ortogonales. Podemos decir que este tensor simétrico lo es bajo el grupo ortogonal, pero no bajo el grupo general lineal.

Pensemos ahora en un tensor en $\mathbb{R}^2$ de tipo $(0,2)$, antisimétrico y escojamos $\epsilon_{12}=-1$ (lo que fija el tensor, pues $\epsilon_{11}=\epsilon_{22}=0, \epsilon_{21}=-\epsilon_{12}$). Si ahora cambiamos las base con una transformación P, tendremos como antes:

$$\epsilon_{ij} = (P^{-1})^k_i (P^{-1})^l_j \epsilon_{kl}$$

para que no varíe al cambiar de base. Por tanto:

$$P^tJP = J$$

donde:

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

Curiosamente, puede comprobarse como las transformaciones que tienen esta propiedad son las del grupo general lineal que tienen determinante igual a 1, es decir, el grupo especial lineal, $SL(2,\mathbb{R})$.

Teniendo en cuenta lo que hemos dicho sobre la interpretación de volumen de esta 2-forma, no es extraño que se diga que las transformaciones del grupo $SL(2,\mathbb{R})$ conservan el volumen (en este caso un área al ser en dimensión 2).

7.12 Espacios con producto escalar

Si el espacio vectorial V está dotado de un producto escalar, esto quiere decir que existe un tensor de segundo orden de tipo $(0,2)$, simétrico definido positivo, y que es invariante, teniendo las mismas componentes en cualquier base. Por tanto, como ya hemos dicho, las transformaciones permisibles son las que verifican:

$$P^t G P = G$$

El producto escalar nos permite hacer uso del teorema de Riesz-Fréchet, identificando de forma canónica el espacio dual con el espacio V, a través del tensor métrico g_{ij}. En efecto:

$$\omega \in V \longrightarrow v \in V$$

tal que:

$$\omega(x) = (v, x)$$

o, en términos de tensores:

$$\omega_i x^i = g_{ij} v^i x^j$$

es decir:

$$\omega_i = g_{ij} v^j$$

Podemos interpretar esto de dos formas. O bien, como el isomorfismo entre el espacio V y su dual, o como la expresión de un mismo vector de V en bases distintas, en realidad las bases recíprocas de las que ya habíamos hablado al introducir las coordenadas covariantes y contravariantes. Haciéndolo así, ω es lo mismo que v pero escrito en otra base y escribiremos:

$$v_i = g_{ij} v^j$$

Esta operación se denomina bajar índices y se hace a través del tensor métrico. Como también hemos dicho, no es necesario tener una forma bilineal definida positiva para hacer esto. Basta con que sea no degenerada.

La operación inversa, subir índices, se realiza con el inverso del tensor métrico:

$$g^{ik} g_{kj} = \delta^i{}_j$$

y es:

$$v^i = g^{ij} v_j$$

Esta operación se puede hacer todas las veces que se quiera en tensores de tipo arbitrario. Cada vez que se baje o suba algún índice aparece un tensor métrico g (o su inverso).

Nótese que el producto escalar de dos vectores de V se escribe simplemente como:

$$g_{ij} x^i y^j = x_i y^i = x^i y_i$$

7.13 Aplicaciones entre espacios producto tensorial

Sean V, W, V', W' espacios vectoriales sobre un cuerpo $\mathbb{K}$. Sean f y g aplicaciones lineales:

$$f\colon V \to V', \quad g\colon W \to W'$$

Si construimos los productos tensoriales $V \otimes W$ y $V' \otimes W'$, es posible definir una aplicación lineal entre estos dos espacios, que se llamará producto tensorial de las aplicaciones f y g:

$$f \otimes g\colon V \otimes W \to V' \otimes W'$$

de la siguiente forma:

$$(f \otimes g)(v \otimes w) = f(v) \otimes g(w)$$

para cualquier par de vectores $v \in V$, $w \in W$. La aplicación se extiende linealmente a todos los elementos de $V \otimes W$. También se puede definir para productos tensoriales de más de dos espacios.

Veamos como se relaciona la representación matricial de las aplicaciones producto tensorial con las matrices de las aplicaciones que forman el producto. Lo haremos en el caso sencillo del producto de dos aplicaciones.

Sean $\mathcal{B}_V = \{v_1, \dots, v_n\}$, $\mathcal{B}_W = \{w_1, \dots, w_m\}$, $\mathcal{B}_{V'} = \{v'_1, \dots, v'_{n'}\}$, $\mathcal{B}_{W'} = \{w'_1, \dots, w'_{m'}\}$, bases de V, W, V', W' respectivamente. Sea $A = (a^i{}_j)$ la matriz que representa a la aplicación lineal f en las bases $\mathcal{B}_V$ y $\mathcal{B}_{V'}$ y $B = (b^i{}_j)$ la que representa a g en las bases $\mathcal{B}_W$ y $\mathcal{B}_{W'}$. Es decir:

$$f(v_i) = \sum_{j=1}^{n'} a^j{}_i v'_j,\ i = 1, \dots, n, \quad g(w_i) = \sum_{j=1}^{m'} b^j{}_i w'_j,\ i = 1, \dots, m$$

De acuerdo con la definición de producto tensorial, una base de $V \otimes W$ es:

$$\mathcal{B}_{V\otimes W} = \{v_i \otimes w_j,\ i = 1, \dots, n,\ j = 1, \dots, m\}$$

y una base de $V' \otimes W'$ es:

$$\mathcal{B}_{V'\otimes W'} = \{v'_i \otimes w'_j,\ i = 1, \dots, n',\ j = 1, \dots, m'\}$$

Lo único que tenemos que decidir es como ordenar esta base. Hay dos formas naturales de hacerlo, bien fijando el índice del vector v_i y dejando variar el de w_j o viceversa. Supongamos que el orden es:

$$\{v_1 \otimes w_1, v_1 \otimes w_2, \dots, v_1 \otimes w_m, v_2 \otimes w_1, \dots, v_2 \otimes w_m, \dots, v_n \otimes w_1, \dots, v_n \otimes w_m\}$$

Según la definición de la aplicación $f \otimes g$:

$$(f \otimes g)(v_i \otimes w_j) = f(v_i) \otimes g(w_j)$$

y sustituyendo las expresiones de estos vectores, se tiene:

$$(f \otimes g)(v_i \otimes w_j) = \left(\sum_{k=1}^{n'} a^k{}_i v'_k\right) \otimes \left(\sum_{l=1}^{m'} b^l{}_j w'_l\right)$$

Aplicando las propiedades del producto tensorial

$$(f \otimes g)(v_i \otimes w_j) = \sum_{k=1}^{n'}\sum_{l=1}^{m'} a^k{}_i b^l{}_j v'_k \otimes w'_l$$

La aplicación lineal $f \otimes g$ viene dada en las bases $\mathcal{B}_{V\otimes W}$ y $\mathcal{B}_{V'\otimes W'}$ por la "matriz" $a^i{}_j b^k{}_l$ cuyas propiedades tensoriales estudiaremos más adelante, pero que, en cualquier caso, puede escribirse como una verdadera matriz, adoptando el orden anteriormente elegido para las bases de estos espacios tensoriales. Se tiene que la matriz de $f \otimes g$, la cual llamaremos producto tensorial de las matrices A y B es:

$$A \otimes B = \begin{pmatrix} a^1{}_1 B & a^1{}_2 B & \cdots & a^1{}_n B \\ a^2{}_1 B & a^2{}_2 B & \cdots & a^2{}_n B \\ \vdots & \vdots & & \vdots \\ a^{n'}{}_1 B & a^{n'}{}_2 B & \cdots & a^{n'}{}_n B \end{pmatrix}$$

que tiene las dimensiones y propiedades correctas como se puede comprobar fácilmente. El objeto $a^i{}_j b^k{}_l$ es obviamente un tensor de cuarto orden, dos veces contravariante y dos veces covariante, pues es el producto tensorial de dos tensores de segundo orden. Por tanto sus propiedades de transformación son las adecuadas a estos tensores. Restringiéndonos al caso $V' = V$, $W' = W$, éstas son:

$$a'^{i'}{}_{j'} b'^{k'}{}_{l'} = P^{i'}{}_i (P^{-1})^j{}_{j'} Q^{k'}{}_k (Q^{-1})^l{}_{l'} a^i{}_j b^k{}_l$$

siendo P la matriz de cambio de base en V y Q la matriz de cambio de base en W.

Ejemplo 7.13.1 El siguiente ejemplo proporciona una interesante aplicación en el estudio de las partículas elementales. Cuando se tiene un sistema compuesto de dos partículas de spin s y s' el sistema presenta un spin que depende de los de las partículas que lo forman. De acuerdo con la teoría cuántica, los valores que pueden tomar s y s' son números enteros o semienteros, positivos $(0, 1/2, 1, 3/2, \ldots)$. Sea V el espacio de los estados de spin de la partícula 1 y W el de los estados de la partícula 2. El espacio de estados del sistema $1+2$ es justamente el producto tensorial de los espacios V y W y los operadores de spin en este espacio se obtienen de la forma siguiente:

$$S_3^{V\otimes W} = S_3^V \otimes I_W + I_V \otimes S_3^W$$

donde S_3^V, S_3^W son los operadores (tercera componente) de spin en cada uno de los espacios V y W.

Consideremos bases en las que estos operadores son diagonales (es decir, bases de autovectores):

$$u_{s_3^V} = |s^V, s_3^V\rangle, \quad u_{s_3^W} = |s^W, s_3^W\rangle$$

donde los índices s_3^V, s_3^W varían en los conjuntos:

$$\{s^V, s^V - 1, \ldots - s^V + 1, -s^V\}, \quad \{s^W, s^W - 1, \ldots - s^W + 1, -s^W\}$$

respectivamente. Estos índices son, además, los autovalores de los operadores S_3^V, S_3^W que por tanto se escriben como:

$$S_3^V = \begin{pmatrix} s_3^V & & & & & \\ & s_3^V - 1 & & & & \\ & & s_3^V - 2 & & & \\ & & & \ddots & & \\ & & & & -s_3^V + 1 & \\ & & & & & -s_3^V \end{pmatrix}$$

$$S_3^W = \begin{pmatrix} s_3^W & & & & & \\ & s_3^W - 1 & & & & \\ & & s_3^W - 2 & & & \\ & & & \ddots & & \\ & & & & -s_3^W + 1 & \\ & & & & & -s_3^W \end{pmatrix}$$

El producto tensorial es ahora muy sencillo, puesto que los operadores son diagonales (identificamos operadores con matrices para mayor sencillez de notación):

$$S_3^V \otimes I_W = \begin{pmatrix} s_3^V & & & & & & & & \\ & \ddots & & & & & & & \\ & & s_3^V & & & & & & \\ & & & s_3^V - 1 & & & & & \\ & & & & \ddots & & & & \\ & & & & & s_3^V - 1 & & & \\ & & & & & & \ddots & & \\ & & & & & & & -s_3^V & \\ & & & & & & & & \ddots & \\ & & & & & & & & & -s_3^V \end{pmatrix}$$

$$I_V \otimes S_3^W \quad = \quad \begin{pmatrix} s_3^W & & & & & & & & & \\ & \ddots & & & & & & & & \\ & & -s_3^W & & & & & & & \\ & & & s_3^W & & & & & & \\ & & & & \ddots & & & & & \\ & & & & & -s_3^W & & & & \\ & & & & & & \ddots & & & \\ & & & & & & & s_3^W & & \\ & & & & & & & & \ddots & \\ & & & & & & & & & -s_3^W \end{pmatrix}$$

y por tanto:

$$S_3^{V\otimes W} = S_3^V \otimes I_W + I_V \otimes S_3^W = \begin{pmatrix} s_3^V + s_3^W & & \\ & \ddots & \\ & & -s_3^V - s_3^W \end{pmatrix}$$

y los números intermedios dependen de las dimensiones de los espacios V y W. Hay que tener en cuenta que la dimensión de V es $2s^V + 1$ y la de W es $2s^W + 1$, con lo que la dimensión de $V \otimes W$ es el producto $(2s^V + 1) \times (2s^W + 1)$. No es difícil demostrar que:

$$(2s^V + 1) \times (2s^W + 1) = \sum_{s=|s^V - s^W|}^{s^V + s^W} 2s + 1$$

Mediante un cambio de base es posible reordenar los valores de s_3 que corresponden a un mismo spin. Los elementos de la matriz de cambio de base se conocen como coeficientes de Clebsch-Gordan.

Ejemplo 7.13.2 Veamos un ejemplo sencillo. Supongamos que tenemos $V = W = \mathbb{C}^2$ y por tanto $s^V = s^W = 1/2$. En este caso, el producto tensorial tiene dimensión 4. El operador de la tercera componente de spin es:

$$S_3^{(1/2)} = \begin{pmatrix} 1/2 & 0 \\ 0 & -1/2 \end{pmatrix}$$

y el correspondiente al producto tensorial es:

$$\begin{aligned} S_3^{(1/2)\otimes(1/2)} &= \begin{pmatrix} 1/2 & 0 \\ 0 & -1/2 \end{pmatrix} \otimes \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \otimes \begin{pmatrix} 1/2 & 0 \\ 0 & -1/2 \end{pmatrix} \\ &= \begin{pmatrix} 1/2 & & & \\ & 1/2 & & \\ & & -1/2 & \\ & & & -1/2 \end{pmatrix} + \begin{pmatrix} 1/2 & & & \\ & -1/2 & & \\ & & 1/2 & \\ & & & -1/2 \end{pmatrix} \\ &= \begin{pmatrix} 1 & & & \\ & 0 & & \\ & & 0 & \\ & & & -1 \end{pmatrix} \end{aligned}$$

La matriz de coeficientes de Clebsch-Gordan, que no describiremos aquí, pasa de esta matriz a la forma:

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

que tiene dos cajas una correspondiente a spin 0 y otra a spin 1.

Ejemplo 7.13.3 Otro ejemplo, en el que aparece un producto de spins distintos, es el caso 1/2 por 1. Ahora tenemos $V = \mathbb{C}^2, W = \mathbb{C}^3$. El producto tensorial tiene dimensión 6. Los operadores de la tercera componente de spin son:

$$S_3^{(1/2)} = \begin{pmatrix} 1/2 & 0 \\ 0 & -1/2 \end{pmatrix}, \quad S_3^{(1)} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix}$$

El producto tensorial es:

$$\begin{aligned} S_3^{(1/2)\times(1)} &= \begin{pmatrix} 1/2 & 0 \\ 0 & -1/2 \end{pmatrix} \otimes \begin{pmatrix} 1 & & \\ & 1 & \\ & & 1 \end{pmatrix} + \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \otimes \begin{pmatrix} 1 & & \\ & 0 & \\ & & -1 \end{pmatrix} \\ &= \begin{pmatrix} 1/2 & & & & & \\ & 1/2 & & & & \\ & & 1/2 & & & \\ & & & -1/2 & & \\ & & & & -1/2 & \\ & & & & & -1/2 \end{pmatrix} + \begin{pmatrix} 1 & & & & & \\ & 0 & & & & \\ & & -1 & & & \\ & & & 1 & & \\ & & & & 0 & \\ & & & & & -1 \end{pmatrix} \\ &= \begin{pmatrix} 3/2 & & & & & \\ & 1/2 & & & & \\ & & -1/2 & & & \\ & & & 1/2 & & \\ & & & & -1/2 & \\ & & & & & -3/2 \end{pmatrix} \end{aligned}$$

De nuevo, la matriz de coeficientes de Clebsch-Gordan pasa a la forma:

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

que tiene dos cajas, una correspondiente a spin 1/2 y otra a spin 3/2.

Tema 8

El espacio afín

El espacio afín. Sistemas de referencia. Transformaciones afines. Espacios euclidianos. Isometrías. Cónicas.

8.1 Introducción

En este breve capítulo introduciremos algunas nociones elementales del espacio afín, junto con la clasificación de cónicas. Usaremos en este tema flechas para designar a los elementos del espacio vectorial y distinguirlos de los puntos (aunque también trataremos de usar minúsculas y mayúsculas con este fin). Además, el producto escalar de dos vectores $\vec{u}$ y $\vec{v}$ se notará por $\vec{u} \cdot \vec{v}$ y su producto vectorial, en $\mathbb{R}^3$, será el usual, denotado por $\vec{u} \times \vec{v}$.

Definición 8.1.1 *Sea V un espacio vectorial sobre un cuerpo $\mathbb{K}$. Un par (X, φ) formado por un conjunto X y una aplicación φ:*

$$\varphi: X \times X \longrightarrow V$$

es un espacio afín sobre V si se verifican las siguientes propiedades:
1) $\varphi(P, R) + \varphi(Q, P) + \varphi(R, Q) = 0, \quad P, Q, R \in X$
2) $\forall P \in X,\ \forall \vec{v} \in V,\ \exists Q \in X$ tal que $\varphi(P, Q) = \vec{v}$

Ejemplo 8.1.1 El ejemplo más inmediato de espacio afín es el que se obtiene tomando $X = V$ y definiendo φ como:

$$\varphi(\vec{x}, \vec{y}) = \vec{y} - \vec{x}$$

El espacio afín puede considerarse en este caso como un espacio vectorial sin un origen preferido.

A los elementos de X se les llama puntos y a los de V vectores. Denotaremos en general:

$$\overrightarrow{PQ} = \varphi(P, Q)$$

8.2 Sistemas de referencia

Supongamos que V es un espacio vectorial de dimensión finita. Un sistema de referencia en un espacio afín (X, φ) con espacio vectorial base V de dimensión finita n, es un conjunto de puntos (elementos de X) $\{P_0, P_1, \ldots, P_n\}$ tal que los vectores $\overrightarrow{P_0P_i}$ son una base de V. Se dice que P_0 es el origen del sistema de referencia.

De esta forma, cualquier punto P del espacio X se puede escribir referido al origen P_0 como:

$$\overrightarrow{P_0P} = \sum_{i=1}^{n} \lambda_i \overrightarrow{P_0P_i}$$

Dos sistemas de referencia están relacionados por una traslación (que pasa de un origen a otro) y por un cambio de base en V. Sean:

$$\{P_0, P_1, \dots, P_n\}, \quad \{Q_0, Q_1, \dots, Q_n\}$$

dos sistemas de referencia en el espacio afín X. El vector $\overrightarrow{Q_0P_0}$ se puede escribir en la base:

$$\{\overrightarrow{Q_0Q_1}, \dots \overrightarrow{Q_0Q_n}\}$$

como:

$$\overrightarrow{Q_0P_0} = \sum_{i=1}^{n} a_i \overrightarrow{Q_0Q_i}$$

mientras que los vectores de la primera base se escriben en función de los de la segunda:

$$\overrightarrow{P_0P_i} = \sum_{j=1}^{n} a_{ji} \overrightarrow{Q_0Q_j}$$

El cambio de base viene especificado por una matriz en $\mathcal{M}_n(\mathbb{K})$, (a_{ij}), la matriz de cambio de base en V y un vector en $\mathbb{K}^n$, (a_i) que da la relación entre los orígenes.

Dado un punto P referido al sistema con origen en P_0:

$$\overrightarrow{P_0P} = \sum_{i=1}^{n} \lambda_i \overrightarrow{P_0P_i} = \sum_{i=1}^{n} \lambda_i \sum_{j=1}^{n} a_{ji} \overrightarrow{Q_0Q_j} = \sum_{j=1}^{n} \left(\sum_{i=1}^{n} a_{ji} \lambda_i \right) \overrightarrow{Q_0Q_j} = \sum_{j=1}^{n} \lambda'_j \overrightarrow{Q_0Q_j}$$

Entonces,

$$\overrightarrow{Q_0P} = \overrightarrow{Q_0P_0} + \overrightarrow{P_0P}$$

con lo que nos queda la expresión:

$$\overrightarrow{Q_0P} = \sum_{j=1}^{n} a_j \overrightarrow{Q_0Q_j} + \sum_{i=1}^{n} \lambda'_i \overrightarrow{Q_0Q_i} = \sum_{i=1}^{n} \left(a_i + \sum_{j=1}^{n} a_{ij} \lambda_j \right) \overrightarrow{Q_0Q_i}$$

8.3 Transformaciones afines

Dados dos espacios afines, (X_1, φ_1) y (X_2, φ_2) sobre los espacios vectoriales V_1 y V_2 respectivamente, consideremos una aplicación entre X_1 y X_2:

$$f : X_1 \longrightarrow X_2$$

Esta aplicación induce otra entre V_1 y V_2 definida de la forma siguiente. Fijado P en X_1, consideremos su imagen $f(P) \in X_2$. Para todo $Q \in X_1$, al vector $\overrightarrow{PQ} \in V_1$ se le hace corresponder el vector $\overrightarrow{f(P)f(Q)} \in V_2$. Por la segunda propiedad de los espacios afines, esta aplicación esta bien definida para todo vector de V_1:

$$\tilde{f} : V_1 \longrightarrow V_2$$

Si $\vec{x} \in V_1$, existe un único $Q \in X_1$ tal que:

$$\vec{x} = \overrightarrow{PQ}$$

Entonces:

$$\tilde{f}(\vec{x}) = \overrightarrow{f(P)f(Q)}$$

Teorema 8.3.1 *Si $\tilde{f}$ es lineal, no depende del punto P elegido para definirla.*

Se dice en este caso que f es una aplicación afín. Si el espacio afín inicial y final coinciden y la aplicación afín φ es biyectiva se llama transformación afín.

Definición 8.3.1 *Una traslación de vector $\vec{x} \in V$ transforma un punto $P \in X$ en otro $Q \in X$ tal que:*

$$\vec{x} = \overrightarrow{PQ}$$

Se tiene el siguiente teorema de estructura de las transformaciones afines

Teorema 8.3.2 *Cualquier transformación afín se puede escribir como un producto de una transformación afín que deja invariante un punto dado (que se puede elegir arbitrariamente) y una traslación.*

Las transformaciones afines tienen en sistemas de referencia afines una expresión matricial dada por:

$$\begin{pmatrix} \mathcal{A} & a \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} I_n & a \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \mathcal{A} & 0 \\ 0 & 1 \end{pmatrix}$$

cuando las coordenadas del punto P se escriben como:

$$\begin{pmatrix} x_1 \\ \vdots \\ x_n \\ 1 \end{pmatrix}$$

La matriz $\mathcal{A}$ es una matriz $n \times n$ de determinante distinto de cero. El vector $\vec{a} \in V$ representa una traslación y $\mathcal{A}$ la transformación afín:

$$\begin{pmatrix} \mathcal{A} & 0 \\ 0 & 1 \end{pmatrix}$$

que deja una punto fijo.

8.4 Espacios euclidianos

Definición 8.4.1 *Un espacio euclidiano es un espacio afín con un espacio vectorial real dotado de un producto escalar.*

El producto escalar en V permite definir una distancia entre los puntos de X:

$$P, Q \in X, \quad d(P, Q) = \|\overrightarrow{PQ}\|$$

Las bases ortonormales de V llevan a sistemas de referencia ortonormales en X. Las transformaciones que pasan de unos a otros están formadas por traslaciones y transformaciones ortogonales.

8.4.1 Isometrías en espacios euclidianos

Definición 8.4.2 *Una aplicación afín es una isometría si la aplicación lineal asociada conserva el producto escalar (es decir, es ortogonal).*

No es difícil probar que las isometrías conservan la distancia, son biyectivas y en sistemas de referencia ortonormales vienen representadas por traslaciones y matrices ortogonales.

Se dice que una isometría es un movimiento del espacio euclidiano si la parte ortogonal de la transformación tiene determinante igual a 1, es decir es una rotación. La isometría es el producto de una traslación por una transformación ortogonal. En un sistema de referencia ortonormal, la isometría viene determinada por una matriz:

$$\left(\begin{array}{c|c} \mathcal{A} & a \\ \hline 0 & 1 \end{array}\right)$$

donde:

$$\mathcal{A}\mathcal{A}^t = I_n, \quad a \in V$$

Si (x_i) son las coordenadas del punto en este sistema de referencia, las del punto transformado son:

$$\begin{pmatrix} x'_1 \\ \vdots \\ x'_n \\ 1 \end{pmatrix} = \left(\begin{array}{c|c} & a_1 \\ \mathcal{A} & \vdots \\ & a_n \\ \hline 0 & 1 \end{array}\right) = \begin{pmatrix} x_1 \\ \vdots \\ x_n \\ 1 \end{pmatrix}$$

Una vez elegido el punto fijo, la descomposición en rotación y traslación es única.

8.5 El plano euclidiano

Sea $\mathbb{R}^2$ dotado del producto escalar usual (es decir, la base canónica es ortonormal) y el espacio afín $X = \mathbb{R}^2$.

8.5.1 Rectas en $\mathbb{R}^2$

Supongamos fijado un origen, P_0, en el espacio afín $X = \mathbb{R}^2$. Una recta es el conjunto de puntos de $\mathbb{R}^2$, P, que verifican la ecuación:

$$\overrightarrow{P_0P} = \overrightarrow{P_0P_1} + \lambda\vec{v}, \lambda \in \mathbb{R}$$

El vector $\vec{v}$ da la dirección de la recta mientras que P_1 es un punto dado de la recta. Supongamos que tenemos un sistema de referencia afín: $\{P_0, \vec{u}, \vec{w}\}$ (con $\vec{u} = \overrightarrow{P_0P_1}$ y $\vec{w} = \overrightarrow{P_0P_2}$). Si (x, y) son las coordenadas de un punto Q (es decir, $\overrightarrow{P_0Q} = x\vec{u} + y\vec{w}$), las ecuaciones paramétricas de la recta se pueden escribir como:

$$\begin{aligned} x &= x_0 + \lambda v_1 \\ y &= y_0 + \lambda v_2 \end{aligned}$$

El parámetro λ se puede eliminar de las ecuaciones y obtener la forma implícita:

$$\frac{x - x_0}{v_1} = \frac{y - y_0}{v_2}$$

Dados dos puntos del plano, existe una única recta que pasa por ellos. Sean Q_1 y Q_2 los puntos de coordenadas (x_1, y_1) y (x_2, y_2). La recta que pasa por esos dos puntos es:

$$\frac{x - x_1}{x_1 - x_2} = \frac{y - y_1}{y_1 - y_2}$$

Por un punto pasa un haz de rectas, de ecuación:

$$\frac{x - x_0}{y - y_0} = k, \quad k \in \mathbb{R}$$

además de la recta $y = y_0$.

8.5.2 Distancia de un punto a una recta

Sea la recta

$$r \equiv \vec{v} = \vec{v}_0 + \lambda\vec{t}$$

y el punto P, en un sistema de referencia afín ortonormal, con origen en P_0. Queremos calcular la distancia del punto a la recta, entendida como la mínima distancia del punto P a los puntos de la recta. La distancia se define a partir de la norma derivada del producto escalar. La distancia entre dos puntos P, Q, de coordenadas en el sistema de referencia afín ortonormal dadas por: (x_1, y_1), (x_2, y_2) es:

$$d(P, Q) = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2}$$

En lenguaje vectorial, sea $w = \overrightarrow{P_0P}$ y $v = \overrightarrow{P_0Q}$ los vectores del punto P y un punto cualquiera Q de la recta r respectivamente. La distancia es entonces:

$$\|\vec{w} - \vec{v}\| = \|\vec{w} - \vec{v}_0 - \lambda\vec{t}\|$$

Sea n un vector unitario perpendicular al vector que define la recta (solo hay dos, elijamos cualquiera). Los vectores $\vec{t}/\|\vec{t}\|$ y $\vec{n}$ forman una base ortonormal en el plano. Por tanto, el vector $\vec{w} - \vec{v}_0$ se puede escribir en esta base como:

$$\vec{w} - \vec{v}_0 = \vec{t} \cdot (\vec{w} - \vec{v}_0)\frac{1}{\|\vec{t}\|^2}\vec{t} + \vec{n} \cdot (\vec{w} - \vec{v}_0)\vec{n} = a\vec{t} + b\vec{n}$$

Por tanto la distancia (al cuadrado) es:

$$d(P,Q)^2 = \|(a - \lambda)\vec{t} + b\vec{n}\|^2 = |a - \lambda|^2 + |b|^2$$

y será mínima cuando:

$$\lambda = a = \frac{1}{\|\vec{t}\|^2}\vec{t} \cdot (\vec{w} - \vec{v}_0)$$

Para este valor de λ la distancia es simplemente:

$$d(P,Q) = |b| = \|\vec{n} \cdot (\vec{w} - \vec{v}_0)\|$$

es decir, como ya sabíamos, la proyección sobre el vector normal de un vector que une el punto P con un punto cualquiera de la recta. Si las coordenadas de $\vec{t}$ en la base en la que estamos trabajando son (t_1, t_2), el vector $\vec{n}$ se puede tomar como:

$$\vec{n} = \frac{1}{\sqrt{t_1^2 + t_2^2}}(-t_2, t_1)$$

y por tanto, la distancia es:

$$d(P,Q) = \frac{1}{\sqrt{t_1^2 + t_2^2}}\|(-t_2, t_1) \cdot (\vec{w} - \vec{v}_0)\|$$

Si la recta se expresa como:

$$ax + by + c = 0$$

el vector dirección es: $(b, -a)$ y un punto sobre ella tiene como coordenadas $(0, -c/b)$ (si $b \neq 0$). Por tanto, la distancia es:

$$d(P,r) = \frac{|ax_1 + by_1 + c|}{\sqrt{a^2 + b^2}}$$

Nótese que el vector normal (unitario) a una recta escrita en la forma anterior es $\vec{n} = (a,b)/\sqrt{a^2 + b^2}$, con lo que la ecuación de la recta se puede escribir como:

$$\vec{n} \cdot \vec{v} = k$$

y el valor absoluto de k resulta ser la distancia al origen.

Una recta es un subespacio afín de dimensión 1. El equivalente en espacios vectoriales es el núcleo de una forma lineal no nula.

8.5.3 Isometrías en el plano

De acuerdo con lo visto anteriormente, las isometrías en $\mathbb{R}^2$ se descomponen en el producto de una rotación (propia si det $= 1$ o impropia si det $= -1$) y una traslación. Por lo tanto la clasificación de las isometrías es la siguiente.

Teorema 8.5.1 *Las isometrías en $\mathbb{R}^2$ se clasifican en los siguientes tipos:*

1. ***Traslaciones.*** *Eligiendo adecuadamente el sistema de referencia (ortonormal), todas las traslaciones son del tipo:*

$$\left\{ \begin{array}{lcl} x' & = & x+a \\ y' & = & y \end{array} \right.$$

donde $a \in \mathbb{R}$, correspondiendo $a=0$ a la identidad.

2. ***Rotaciones propias.***

$$\left\{ \begin{array}{lcl} x' & = & x\cos\theta - y \operatorname{sen}\theta \\ y' & = & x \operatorname{sen}\theta + y\cos\theta \end{array} \right.$$

con $0 \leq \theta < 2\pi$, que dejan invariante el origen de coordenadas. También $\theta = 0$ corresponde a la identidad.

3. ***Reflexiones respecto una recta.*** *Como hemos visto, toda rotación impropia podía ser llevada a la forma: $\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$. Por tanto:*

$$\left\{ \begin{array}{lcl} x' & = & x \\ y' & = & -y \end{array} \right.$$

4. ***Reflexiones respecto una recta y traslación en la dirección de esa recta.***

$$\left\{ \begin{array}{lcl} x' & = & x+a \\ y' & = & -y \end{array} \right.$$

con $a \neq 0$.

Los tipos 1,2,3,4 no son equivalentes y cualquier isometría puede ser llevada a uno de ellos definiendo adecuadamente el sistema de referencia. Si son distintos de la identidad, el tipo 1 no tiene puntos fijos y es un movimiento (conserva la orientación), el 2 tiene un un punto fijo (y es también un movimiento). Los tipos 3 y 4 no son movimientos. El 3 tiene una recta fija (punto a punto) y el 4 no tiene puntos fijos.

Demostración. La matriz asociada a una isometría es:

$$\left(\begin{array}{cc|c} m & n & a \\ p & q & b \\ \hline 0 & 0 & 1 \end{array} \right)$$

donde la matriz

$$\begin{pmatrix} m & n \\ p & q \end{pmatrix}$$

es ortogonal. Por tanto podemos elegir una base ortonormal en $\mathbb{R}^2$, de forma que esta matriz se pueda poner como una de las dos formas siguientes:

$$\begin{pmatrix} \cos\theta & -\operatorname{sen}\theta \\ \operatorname{sen}\theta & \cos\theta \end{pmatrix}, \quad \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

dependiendo de su determinante (± 1). En el primer caso (rotación propia) se tiene:

$$\left(\begin{array}{cc|c} \cos\theta & -\operatorname{sen}\theta & a \\ \operatorname{sen}\theta & \cos\theta & b \\ \hline 0 & 0 & 1 \end{array} \right)$$

Si $\theta = 0$:

$$x' = x + a, \quad y' = y + b$$

y pasando a otro sistema de referencia con:

$$u = x\cos\alpha - y \operatorname{sen}\alpha, \quad v = x \operatorname{sen}\alpha + y\cos\alpha$$

con $c = \sqrt{a^2 + b^2}$, $\tan \alpha = -b/a$ se obtiene el tipo 1:

$$u' = u + c, \quad v' = v$$

La matriz de la isometría es:

$$\left(\begin{array}{cc|c} 1 & 0 & c \\ 0 & 1 & 0 \\ \hline 0 & 0 & 1 \end{array}\right)$$

y no hay puntos fijos ($c \neq 0$).

Si $a = b = 0$, no hay traslación y se tiene una rotación propia con punto fijo en el origen:

$$\left(\begin{array}{cc|c} \cos\theta & -\operatorname{sen}\theta & 0 \\ \operatorname{sen}\theta & \cos\theta & 0 \\ \hline 0 & 0 & 1 \end{array}\right)$$

Supongamos ahora que $\theta \neq 0$ y $(a, b) \neq (0, 0)$. En este caso podemos estudiar la existencia de puntos fijos:

$$\begin{array}{rcl} x\cos\theta - y\operatorname{sen}\theta + a & = & x \\ x\operatorname{sen}\theta + y\cos\theta + b & = & y \end{array}$$

El determinante de la matriz de coeficientes de este sistema lineal es:

$$\det \begin{pmatrix} \cos\theta - 1 & -\operatorname{sen}\theta \\ \operatorname{sen}\theta & \cos\theta - 1 \end{pmatrix} = 2(1 - \cos\theta)$$

es decir, si $\theta \neq 0$ existe un único punto fijo, de coordenadas:

$$\left(\frac{a}{2} - \frac{b\operatorname{sen}\theta}{2(1-\cos\theta)}\right), \quad \left(\frac{b}{2} + \frac{a\operatorname{sen}\theta}{2(1-\cos\theta)}\right)$$

que se puede tomar como centro de la rotación, trasladando el origen de coordenadas. De esta forma la traslación desaparece y tenemos nuevamente una isometría de tipo 2. Si $\theta = 0$ se obtiene nuevamente una traslación. No hay puntos fijos.

Cuando la rotación es impropia, existe un sistema de referencia en el que la matriz es:

$$\left(\begin{array}{cc|c} 1 & 0 & a \\ 0 & -1 & b \\ \hline 0 & 0 & 1 \end{array}\right)$$

Si $a = b = 0$ no hay traslación y obtenemos una reflexión respecto a una recta que es invariante (tipo 3). Si $(a, b) \neq (0, 0)$, podemos eliminar b mediante una traslación:

$$u = x, \quad v = y - \frac{b}{2}$$

y obtenemos una transformación de tipo 4:

$$u' = u + a, \quad v' = -v$$

si $a \neq 0$ (si $a = 0$ es de tipo 3). QED

8.5.4 Transformaciones de puntos y rectas bajo isometrías

Las rotaciones propias giran los puntos del plano un ángulo θ. Si $P = (x, y)$ es uno de esos puntos, su punto transformado $P' = (x', y')$ viene dado por:

$$x' = x\cos\theta - y\operatorname{sen}\theta, \quad y' = x\operatorname{sen}\theta + y\cos\theta$$

Por tanto, dado un vector $\vec{v} = \overrightarrow{P_1P_2}$, (con $P_1 = (x_1, y_2),\ P_2 = (x_2, y_2)$ el vector transformado $\vec{v}' = \overrightarrow{P_1'P_2'}$ es:

$$\vec{v}' = (x_2' - x_1', y_2' - y_1') = (x_2 - x_1)\cos\theta - (y_2 - y_1)\operatorname{sen}\theta, x_2 - x_1)\operatorname{sen}\theta + (y_2 - y_1)\cos\theta)$$

es decir, el vector $\vec{v}$ se transforma con una rotación R en el espacio vectorial:

$$\vec{v}' = R\vec{v}$$

Como consecuencia de estas transformaciones, una recta que pase por el punto $P = (x_0, y_0) = \vec{v}_0$ y que tenga como vector $\vec{t}$, se transforma en otra recta:

$$\vec{v} = \vec{v}_0 + \lambda\vec{t} \longrightarrow \vec{v} = R\vec{v}_0 + \lambda R\vec{t}_0$$

Si la recta se representa en la forma $\vec{n} \cdot \vec{v} = c$, la ecuación transformada es:

$$(R\vec{n}) \cdot \vec{v} = c$$

es decir, el vector normal gira con la rotación y la distancia al origen es la misma.

En una reflexión respecto a una recta r, la imagen de un punto se puede calcular así. Sea $P = (x_0, y_0) = \vec{v}_0$ un punto del plano, y la recta: $\vec{n} \cdot \vec{v} = c$. El punto $P' = (x_0', y_0')$ simétrico de P respecto de la recta r, verifica que el punto A de coordenadas:

$$\frac{1}{2}(\vec{v}_0 + \vec{v}_0') = \left(\frac{x_0 + x_0'}{2}, \frac{y_0 + y_0'}{2}\right)$$

está sobre la recta r, es decir:

$$\vec{n} \cdot (\vec{v}_0 + \vec{v}_0') = 2c$$

y además el vector $\vec{v}_0 - \vec{v}_0'$ es paralelo a $\vec{n}$:

$$\vec{v}_0 - \vec{v}0' = \mu\vec{n}$$

De estas dos relaciones podemos despejar las coordenadas de P':

$$\vec{v}_0' = \vec{v}_0 - \mu\vec{n}, \quad \mu = \frac{2}{\|\vec{n}\|^2}(\vec{n}\vec{v}_0 - c)$$

llegando a la expresión:

$$\vec{v}_0' = \vec{v}_0 - \frac{2}{\|\vec{n}\|^2}(\vec{n} \cdot \vec{v}_0 - c)\vec{n}$$

Si la recta pasa por el origen, $c = 0$ y se tiene:

$$\vec{v}_0' = \vec{v}_0 - \frac{2\vec{n} \cdot \vec{v}_0}{\|\vec{n}\|^2}\vec{n}$$

Una traslación consiste simplemente en la suma de un vector constante. Por tanto una recta se convierte en otra recta con el mismo vector de dirección y sus puntos trasladados por ese vector constante.

Las isometrías también se pueden interpretar como cambios en el sistema de referencia, (al igual que ocurre con las aplicaciones lineales y los cambios de base en los espacios vectoriales).

8.6 El espacio euclidiano

El espacio afín que estudiaremos en esta sección es el que tiene como como conjunto de puntos y espacio vectorial a $\mathbb{R}^3$ y en el que está definido el producto escalar usual. Veamos como son las variedades (subespacios) afines en este espacio.

8.6.1 Rectas en el espacio

Una recta en $\mathbb{R}^3$ es una variedad afín de dimensión 1, determinada por tanto por un vector $\vec{v} \in \mathbb{R}^3$ y un punto P:

$$\vec{v} = \vec{v}_0 + \lambda\vec{t}, \quad \lambda \in \mathbb{R}$$

es decir:

$$x = x_0 + \lambda t_1, \quad y = y_0 + \lambda t_2, \quad z = z_0 + \lambda t_3$$

Eliminando λ de estas tres ecuaciones se obtiene:

$$\frac{x - x_0}{t_1} = \frac{y - y_0}{t_2} = \frac{z - z_0}{t_3}$$

Una recta viene determinada por dos ecuaciones. Como veremos es la intersección de dos subvariedades afines de dimensión 2.

Al igual que en el plano, una recta viene fijada por dos puntos distintos. La ecuación se escribe como:

$$\frac{x - x_1}{x_1 - x_2} = \frac{y - y_1}{y_1 - y_2} = \frac{z - z_1}{z_1 - z_2}$$

Además de las dos posiciones relativas de dos rectas en el plano (que se corten o sean paralelas), en el espacio existe una tercera posibilidad, que las rectas se crucen. Pero antes de estudiar estas posiciones, veamos como son las subvariedades de dimensión 2, los planos.

8.6.2 Planos en el espacio

Un plano es una subvariedad afín de dimensión 2. Viene determinado por dos vectores linealmente independiente y un punto:

$$\vec{v} = \vec{v}_0 + \lambda\vec{u} + \mu\vec{v}$$

es decir:

$$x = x_0 + \lambda u_1 + \mu v_1, \quad y = y_0 + \lambda u_2 + \mu v_2, \quad z = z_0 + \lambda u_3 + \mu v_3$$

Los parámetros λ y μ pueden ser eliminados en la manera usual. El vector $(x, y, z) - (x_0, y_0, z_0)$ depende linealmente de $\vec{u}, \vec{v}$ si el siguiente determinante es cero:

$$\det \begin{pmatrix} u_1 & v_1 & x - x_0 \\ u_2 & v_2 & y - y_0 \\ u_3 & v_3 & z - z_0 \end{pmatrix} = 0 \tag{8.1}$$

Tres puntos no alineados determinan un plano en el espacio. La ecuación se obtiene fácilmente de la anterior:

$$\det \begin{pmatrix} x_1 - x_0 & x_2 - x_0 & x - x_0 \\ y_1 - y_0 & y_2 - y_0 & y - y_0 \\ z_1 - z_0 & z_2 - z_0 & z - z_0 \end{pmatrix} = 0$$

Resolviendo la ecuación (8.1), la ecuación del plano se escribe como:

$$ax + by + cz = d$$

y el vector $\vec{n} = (a, b, c)$ es un vector normal al plano, pues es igual al producto vectorial de los vectores $\vec{u}$ y $\vec{v}$ como se comprueba fácilmente:

$$ax + by + cz = \det \begin{pmatrix} u_1 & v_1 & x \\ u_2 & v_2 & y \\ u_3 & v_3 & z \end{pmatrix} = \det \begin{pmatrix} u_1 & v_1 & x_0 \\ u_2 & v_2 & y_0 \\ u_3 & v_3 & z_0 \end{pmatrix} = d$$

8.6.3 Posiciones relativas de rectas

Volvemos en esta sección al estudio de las posiciones de rectas en $\mathbb{R}^3$. Dos rectas en el espacio se pueden cortar, ser paralelas o cruzarse (aparte de coincidir). Si se cortan o son paralelas, están contenidas en un único plano. Si se cruzan no existe ningún plano que las contenga. Sean las rectas:

$$\vec{v} = \vec{v}_0 + \lambda\vec{t}, \quad \vec{v}' = \vec{v}'_0 + \lambda'\vec{t}'$$

y consideremos un vector que vaya de un punto de una de ellas a un punto de la otra:

$$\vec{w} = \vec{v} - \vec{v}' = \vec{v}_0 - \vec{v}'_0 + \lambda\vec{t} - \lambda'\vec{t}'$$

Veamos si este vector puede ser perpendicular a ambas rectas:

$$\vec{w} \cdot \vec{t} = \vec{w} \cdot \vec{t}' = 0$$

obteniéndose un sistema para λ, λ':

$$\begin{array}{rcl} \lambda\|\vec{t}\|^2 - \lambda'\vec{t} \cdot \vec{t}' & = & -(\vec{v}_0 - \vec{v}'_0) \cdot \vec{t} \\ \lambda\vec{t} \cdot \vec{t}' - \lambda'\|\vec{t}\|^2 & = & -(\vec{v}_0 - \vec{v}'_0 \cdot)\vec{t}' \end{array}$$

El determinante de la matriz de coeficientes es

$$(\vec{t} \cdot \vec{t}')^2 - \|\vec{t}\|^2\|\vec{t}'\|^2 \leq 0$$

debido a la desigualdad de Cauchy-Schwarz. Es cero solo si $\vec{t}$ y $\vec{t}'$ son linealmente dependientes. Por tanto las posibilidades son:

1. $\vec{t}$, $\vec{t}'$ linealmente independientes. Existe una única solución
2. $\vec{t}$, $\vec{t}'$ linealmente dependientes. Puede haber infinitas soluciones o no haber ninguna. Sin embargo, el rango de la matriz ampliada es igual al de la de coeficientes, con lo que hay infinitas soluciones ($\vec{t}' = \alpha\vec{t}$):

$$\det \begin{pmatrix} (\vec{v} - \vec{v}') \cdot \vec{t} & \|\vec{t}\|^2 \\ \alpha(\vec{v} - \vec{v}') \cdot \vec{t} & \alpha\|\vec{t}\|^2 \end{pmatrix} = 0$$

   En este caso las rectas son paralelas.

En el primer caso, solo hay un vector (l.i.) que sea perpendicular a ambas rectas, que se define por

$$\vec{n} = \frac{1}{\|\vec{t} \times \vec{t}'\|}\vec{t} \times \vec{t}'$$

escogiéndolo unitario. Es por lo tanto paralelo a $\vec{v} - \vec{v}'$ cuando éste verifica las ecuaciones anteriores. El coeficiente de proporcionalidad es la proyección de cualquier vector que una dos puntos arbitrarios de ambas rectas:

$$\vec{w} = \vec{v} - \vec{v}' = (\vec{n} \cdot \vec{w}_0)\vec{n}$$

donde $\vec{w}_0 = \vec{v}_0 - \vec{v}'_0$. Aquí aparecen dos casos:

1. $(\vec{n} \cdot \vec{w}_0) = 0$. En este caso hay un plano que contiene a ambas rectas, de vector normal $\vec{n}$ y las rectas se cortan.
2. $(\vec{n} \cdot \vec{w}_0) \neq 0$. Ahora no hay ningún plano que las contenga y las rectas se cruzan. La distancia entre ambas rectas (la distancia mínima) es la longitud del vector $\vec{w}$ calculado anteriormente:

$$d = \frac{1}{\|\vec{t} \times \vec{t}'\|}|(\vec{v}_0 - \vec{v}'_0) \cdot (\vec{t} \times \vec{t}')|$$

8.6.4 Posiciones relativas de planos

Dos planos en $\mathbb{R}^3$ se cortan según una recta o son paralelos (o coincidentes). Si las ecuaciones de los planos son:

$$\vec{n}_1 \cdot \vec{v} = c_1, \quad \vec{n}_2 \cdot \vec{v} = c_2$$

los planos son paralelos si los vectores $\vec{n}_1, \vec{n}_2$ son paralelos. En caso contrario, su producto vectorial determina el vector dirección de la recta que forma su intersección:

$$\vec{v} = \vec{v}_0 + \lambda(\vec{n}_1 \times \vec{n}_2)$$

8.6.5 Distancia de un punto a un plano

El cálculo se puede hacer mediante una proyección. Si la ecuación del plano es:

$$ax + by + cz = d$$

y el punto es $P(x_0, y_0, z_0)$, tomando un punto cualquiera del plano: (x_1, y_1, z_1), y el vector: $(x_0 - x_1, y_0 - y_1, z_0 - z_1)$ al proyectarlo sobre la dirección (a, b, c) se tiene:

$$d(P, \pi) = \frac{1}{\sqrt{a^2 + b^2 + c^2}} |a(x_0 - x_1) + b(y_0 - y_1) + c(z_0 - z_1)| = \frac{1}{\sqrt{a^2 + b^2 + c^2}} |ax_1 + by_1 + cz_1 - d|$$

La distancia de un plano al origen es:

$$d(O, \pi) = \frac{|d|}{\sqrt{a^2 + b^2 + c^2}}$$

8.6.6 Isometrías

Se tiene el siguiente resultado sobre la clasificación de las isometrías en $\mathbb{R}^3$ que no demostraremos:

Teorema 8.6.1 *las isometrías en $\mathbb{R}^3$ se pueden clasificar en los siguientes tipos:*

1. *Traslaciones.*

$$\begin{aligned} x' &= x + a \\ y' &= y \\ z' &= z \end{aligned}$$

   con $a \in \mathbb{R}$.

2. *Rotaciones propias alrededor de una recta y una traslación en la dirección de la recta (que puede ser trivial).*

$$\begin{aligned} x' &= x + a \\ y' &= y \cos\theta - z \operatorname{sen}\theta \\ z' &= y \operatorname{sen}\theta + z \cos\theta \end{aligned}$$

3. *Rotación más una reflexión relativa a un plano que la rotación deja fijo.*

$$\begin{aligned} x' &= x \cos\theta - y \operatorname{sen}\theta \\ y' &= x \operatorname{sen}\theta + y \cos\theta \\ z' &= -z \end{aligned}$$

4. *Reflexión respecto a un plano seguido de una traslación no trivial.*

$$\begin{aligned} x' &= x + a \\ y' &= y \\ z' &= -z \end{aligned}$$

8.7 Clasificación de cónicas

Aunque no es un tema propio del álgebra lineal, la clasificación de cónicas (y cuádricas) tiene un gran interés en su aplicación a numerosos problemas (como el estudio de formas cuadráticas) así y como un ejemplo de la aplicación de transformaciones afines a objetos geométricos.

Una cónica es el conjunto de puntos del plano afín que verifican la ecuación:

$$a_{11}x^2 + a_{22}y^2 + 2a_{12}xy + 2a_{01}x + 2a_{02}y + a_{00} = 0, \quad x, y, a_{ij} \in \mathbb{R}$$

en un sistema de referencia dado. Esta ecuación se puede poner en forma matricial como:

$$(x\ y\ 1) \begin{pmatrix} a_{11} & a_{12} & a_{01} \\ a_{12} & a_{22} & a_{02} \\ a_{01} & a_{02} & a_{00} \end{pmatrix} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} = 0$$

Sean las matrices A, a y los vectores a_0, X:

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{01} \\ a_{12} & a_{22} & a_{02} \\ a_{01} & a_{02} & a_{00} \end{pmatrix}, \quad a = \begin{pmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{pmatrix}, \quad a_0 = \begin{pmatrix} a_{01} \\ a_{02} \end{pmatrix}, \quad X = \begin{pmatrix} x \\ y \\ 1 \end{pmatrix}$$

La ecuación de la cónica es:

$$X^t A X = 0$$

La parte homogénea de segundo orden, correspondiente a la matriz a, es una forma cuadrática en $\mathbb{R}^2$. El objetivo es hacer transformaciones en el plano afín como las que ya hemos estudiado:

$$X' = MX, \quad M = \begin{pmatrix} m_{11} & m_{12} & m_1 \\ m_{12} & m_{22} & m_2 \\ 0 & 0 & 1 \end{pmatrix}, \quad m = \begin{pmatrix} m_{11} & m_{12} \\ m_{12} & m_{22} \end{pmatrix}, \quad m_0 = \begin{pmatrix} m_1 \\ m_2 \end{pmatrix}$$

de manera que la cónica adopte la forma más sencilla posible. Además de isometrías, también permitiremos homotecias, es decir transformaciones afines dadas por matrices:

$$M = \begin{pmatrix} \lambda & & \\ & \mu & \\ & & 1 \end{pmatrix}$$

donde $\lambda, \mu \neq 0$, que no conservan ángulos ni distancias.

Una transformación afín convierte a la cónica en otra, de matriz:

$$X' = MX, \quad (X')^t A' X' = 0, \quad A' = M^t A M$$

y de manera más explícita:

$$\begin{pmatrix} m^t & 0 \\ m_0^t & 1 \end{pmatrix} \begin{pmatrix} a & a_0 \\ a_0^t & a_{00} \end{pmatrix} \begin{pmatrix} m & m_0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} m^t a m & m^t(a m_0 + a_0) \\ (m^t(a m_0 + a_0))^t & m_0^t a m_0 + 2m_0^t a_0 + a_{00} \end{pmatrix}$$

Como la matriz a es simétrica es posible diagonalizarla mediante una transformación de similaridad (como una forma cuadrática). Por tanto, la signatura no cambia al hacer esta transformación. De esta manera disponemos de cuatro invariantes, los rangos de las matrices A y a y la diferencia entre el número de $+1$ y -1 en la forma diagonal (como la matriz de la cónica está definida salvo un factor, podemos siempre escogerla de forma que el número de $+1$ sea mayor o igual que el de -1 en la forma diagonal). Se puede establecer el siguiente cuadro de clasificación en el que aparecen los nombres de las cónicas que se obtienen. Cuando el cuadro no contiene ninguna denominación es que tal posibilidad no puede darse.

	R	3		2		1	0
r	s/S	3	1	2	0	1	0
2	2	Ei	Er	2Rci			
	0		H		2Rir		
1	1		P	2Rpi	2Rpr	2Rcr	
0	0				1Rr	$\emptyset$	$\mathbb{R}^2$

donde: Ei= elipse imaginaria, Er= elipse real, H= hipérbola, P= parábola, 2Rci= dos rectas imaginarias conjugadas, 2Rpi=dos rectas paralelas imaginarias, 2Rir=dos rectas incidentes reales, 2Rpr=dos rectas paralelas reales, 2Rcr=dos rectas coincidentes reales , 1Rr=una recta real.

8.7.1 Formas canónicas

Veamos ahora cómo elegir el sistema de referencia para que las formas de las cónicas sean lo más sencillas posibles. Usaremos transformaciones afines para obtenerlas.

En primer lugar trataremos de anular los términos lineales.

$$m^t(am_0 + a_0) = 0$$

o, como m es regular,

$$am_0 + a_0 = 0$$

Distinguiremos dos casos

1. a regular. Entonces:

$$m_0 = -a^{-1}a_0$$

y como a es simétrica, tenemos:

$$A' = \begin{pmatrix} m^tam & 0 \\ 0 & a_{00} - a_0^t a^{-1} a_0 \end{pmatrix}$$

ii) a no regular. En este caso, no es posible anular en general los términos no diagonales.

De cualquier forma, la matriz a, que es simétrica, se puede diagonalizar mediante una transformación m^tam. Volvamos a los casos anteriores

1. La matriz de la cónica es:

$$A' = \begin{pmatrix} \lambda_1 & & \\ & \lambda_2 & \\ & & a'_{00} \end{pmatrix}$$

donde $\lambda_1 \neq 0$, $\lambda_2 \neq 0$. El rango de la matriz a es 2 ($r = 2$), y s puede ser igual a 2 ó 0. El rango de la matriz A es 3 si $a_{00} \neq a_0^t a^{-1} a_0$ o 2 si son iguales. En el primer caso, S puede valer 3 ó 1. En el segundo caso, 2 ó 0. Por supuesto si $R = 3$ y $S = 3$, entonces $s = 2$. Si $R = 2$ entonces $S = s$.

2. En el segundo caso, al menos uno de los términos de la diagonal de a' es 0:

$$A' = \begin{pmatrix} \lambda_1 & 0 & a'_{01} \\ 0 & 0 & a'_{02} \\ a'_{01} & a'_{02} & a'_{00} \end{pmatrix}$$

Dos situaciones se pueden presentar aquí:

(a) Aunque no exista a^{-1}, se puede encontrar m_0 tal que $a_0 = -am_0$. De esta forma, la matriz A' es:

$$A' = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & a'_{00} \end{pmatrix}$$

El rango de la matriz A debe ser 2, 1 ó 0, y el de la matriz a 1 ó 0.

(b) No existe m_0 tal que $a_0 = -am_0$. Entonces, la matriz A no es diagonalizable mediante este tipo de transformaciones. No podemos anular los términos lineales. La matriz a es igual a 0 o se puede escribir en forma diagonal con uno de los elemento de la diagonal no nulos.

i. Si $a = 0$, la matriz A es:

$$\begin{pmatrix} 0 & 0 & a_{01} \\ 0 & 0 & a_{02} \\ a_{01} & a_{02} & a_{00} \end{pmatrix}$$

Si $a_0 \neq 0$ podemos encontrar una matriz m tal que $a_0^t m = (1,0)$ y $a_{00} = -m_0^t a_0$, con lo que la matriz A se convierte en:

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

ii. Si

$$a' = \begin{pmatrix} 1 & \\ & 0 \end{pmatrix}$$

podemos hacer que $(a_0')^t = (0,1)$ y $a_{00} = 0$, con lo que la matriz A' es:

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

Veamos las formas canónicas:

1. Elipse imaginaria:

$$R = 3,\ r = 2\ S = 3,\ s = 2 \longrightarrow \begin{pmatrix} 1 & & \\ & 1 & \\ & & 1 \end{pmatrix}, \quad x^2 + y^2 + 1 = 0$$

2. Elipse real:

$$R = 3,\ r = 2\ S = 1,\ s = 2 \longrightarrow \begin{pmatrix} 1 & & \\ & 1 & \\ & & -1 \end{pmatrix}, \quad x^2 + y^2 - 1 = 0$$

3. Hipérbola:

$$R = 3,\ r = 2\ S = 1,\ s = 0 \longrightarrow \begin{pmatrix} 1 & & \\ & -1 & \\ & & 1 \end{pmatrix}, \quad x^2 - y^2 + 1 = 0$$

4. Dos rectas imaginarias conjugadas:

$$R = 2,\ r = 2\ S = 2,\ s = 2 \longrightarrow \begin{pmatrix} 1 & & \\ & 1 & \\ & & 0 \end{pmatrix}, \quad x^2 + y^2 = 0$$

5. Dos rectas reales incidentes:

$$R = 2,\ r = 2\ S = 0,\ s = 0 \longrightarrow \begin{pmatrix} 1 & & \\ & -1 & \\ & & 0 \end{pmatrix}, \quad x^2 - y^2 = 0$$

Se llaman cónicas con un centro único.

6. Dos rectas imaginarias paralelas:

$$R = 2,\ r = 1\ S = 2,\ s = 1 \longrightarrow \begin{pmatrix} 1 & & \\ & 0 & \\ & & 1 \end{pmatrix}, \quad x^2 + 1 = 0$$

7. Dos rectas reales paralelas:

$$R = 2,\ r = 1\ S = 0,\ s = 1 \longrightarrow \begin{pmatrix} 1 & & \\ & 0 & \\ & & -1 \end{pmatrix}, \quad x^2 - 1 = 0$$

8. Dos rectas reales coincidentes:

$$R = 1,\ r = 1\ S = 1,\ s = 1 \longrightarrow \begin{pmatrix} 1 & & \\ & 0 & \\ & & 0 \end{pmatrix}, \quad x^2 = 0$$

Se llaman cónicas con una recta de centros. Finalmente:

9. Parábola:

$$R = 3,\ r = 1\ S = 1,\ s = 1 \longrightarrow \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, \quad x^2 + 2y = 0$$

10. Una recta real:

$$R = 2,\ r = 0\ S = 0,\ s = 0 \longrightarrow \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \quad x = 0$$

y además:

11. Vacío:

$$R = 1,\ r = 0\ S = 1,\ s = 0 \longrightarrow \begin{pmatrix} 0 & & \\ & 0 & \\ & & 1 \end{pmatrix}, \quad 1 = 0$$

12. $\mathbb{R}^2$:

$$R = 0,\ r = 0\ S = 0,\ s = 0 \longrightarrow \begin{pmatrix} 0 & & \\ & 0 & \\ & & 0 \end{pmatrix}, \quad 0 = 0$$

La clasificación se puede hacer también en función de invariantes asociados a las matrices A y a:

$$K = \det A,\ J = \det a,\ I = \operatorname{tr} a,\ \tilde{J} = A_{11} + A_{22},\ A_{11} = \begin{pmatrix} a_{11} & a_{01} \\ a_{01} & a_{00} \end{pmatrix}\ A_{22} = \begin{pmatrix} a_{22} & a_{02} \\ a_{02} & a_{00} \end{pmatrix}$$

con lo que la clasificación es:

$$\begin{cases} K \neq 0 \text{ (CO)} & \begin{cases} J \neq 0 \text{ (CCU)} & \begin{cases} J > 0 \text{ (E)} & \begin{cases} KI > 0 \text{ (i)} \\ KI < 0 \text{ (r)} \end{cases} \\ J < 0 \text{ (H)} \end{cases} \\ J = 0 \text{ (P)} \end{cases} \\ K = 0 \text{ (CD)} & \begin{cases} J \neq 0 \text{ (CCU)} & \begin{cases} J > 0 \text{ (2Ric)} \\ J < 0 \text{ (2Rri)} \end{cases} \\ J = 0 & \begin{cases} I \neq 0 \text{ (CRC)} & \begin{cases} \tilde{J} \neq 0 \text{ (2Rp)} & \begin{cases} \tilde{J} > 0 \text{ (i)} \\ \tilde{J} < 0 \text{ (r)} \end{cases} \\ \tilde{J} = 0 \text{ (2Rrc)} \end{cases} \\ I = 0 \text{ (R)} \end{cases} \end{cases} \end{cases}$$

donde CO son cónicas ordinarias y CD cónicas degeneradas. CCU son cónicas con centro único y CRC cónicas con recta de centros. Es sencillo ver la equivalencia de ambas formas de clasificar las cónicas. Un estudio similar se podría hacer para las cuádricas, los conjuntos correspondientes en $\mathbb{R}^3$, pero no lo haremos aquí.

Las clasificaciones de cónicas y cuádricas adquieren una forma más simplificada cuando se considera el espacio proyectivo. Pero esta cuestión va más allá de lo que estas notas pretenden ser.

Problemas

Los problemas que aparecen en esta sección son parte de colecciones de problemas elaborados por numerosas personas. No se han incluido al final de cada sección debido a que muchos de ellos contienen conceptos que aparecen en más de un tema. Sin embargo se ha procurado mantener el orden correspondiente al resto de estas notas.

1. Sean X y X' dos conjuntos, $A, B \subset X$, $A', B' \subset X'$, y f una aplicación, $f\colon X \to X'$. Estudiar si son ciertas o no las siguientes afirmaciones:

   (a) $f(A \cup B) = f(A) \cup f(B)$
   (b) $f(A \cap B) \subset f(A) \cap f(B)$
   (c) $f^{-1}(A' \cup B') = f^{-1}(A') \cup f^{-1}(B')$
   (d) $f^{-1}(A' \cap B') \subset f^{-1}(A') \cap f^{-1}(B')$
   (e) $A \subset B \Rightarrow f(A) \subset f(B)$
   (f) $A' \subset B' \Rightarrow f^{-1}(A') \subset f^{-1}(B')$

   En caso de ser ciertas, demostrarlo. Si son falsas, construir un contraejemplo.

2. Sea $f\colon X \to Y$ una aplicación. Demostrar que f es inyectiva si y solo si: $f(A \cap B) = f(A) \cap f(B)$, para cualquier par $A, B \subset X$.

3. Sean X, Y y Z tres conjuntos, f, g, aplicaciones: $f\colon X \to Y$, $g\colon Y \to Z$. Estudiar si es cierto o no que las afirmaciones de las columnas segunda y tercera implican la de la cuarta en cada fila. (i=inyectiva, s=sobreyectiva, b=biyectiva).

	f	g	$g \circ f$
1	i	i	i
2	i	s	i
3	b	i	i
4	s	s	s
5	i	s	s

	f	g	$g \circ f$
6	s	b	s
7	b	b	b
8	b	i	b
9	i	b	s
10	i	s	b

   Demostrarlas en caso de ser ciertas y encontrar contraejemplos cuando no lo sean.

4. Sean X y X' dos conjuntos, $f\colon X \to X'$. Demostrar que si f es biyectiva, existe la aplicación inversa, $f^{-1}\colon X' \to X$, y que en este caso, la imagen inversa de un conjunto $A' \subset X'$, que llamaremos $f^{-1}(A')$, coincide con la imagen directa de A' mediante la aplicación f^{-1}.

5. Sea $f\colon X \to X'$ una aplicación sobreyectiva. Demostrar que existe una aplicación $\sigma\colon X' \to X$ tal que: $f \circ \sigma = 1'_X$ donde $1'_X$ es la aplicación identidad en X'. Sea $h\colon X \to X'$ una aplicación inyectiva. Demostrar que $\hat{h}\colon X \to h(X)$ definida por $\hat{h}(x) = h(x)$ es biyectiva y existe $\tau = \hat{h}^{-1}\colon h(X) \to X$ tal que: $\tau \circ h = 1_X$ donde 1_X es la aplicación identidad en X. ¿Existe una única aplicación $\alpha\colon X' \to X$ que verifique $\alpha \circ h = 1_X$?

6. Sean X, Y y Z tres conjuntos, f, g, aplicaciones: $f: X \to Y$, $g: Y \to Z$. Sea $h = g \circ f$. Estudiar los siguientes enunciados:

   (a) Si h es inyectiva entonces f es inyectiva.

   (b) Si h es inyectiva y f sobreyectiva, entonces g es inyectiva.

   (c) Si h es sobreyectiva, entonces g es sobreyectiva.

   (d) Si h es sobreyectiva y g es inyectiva, entonces f es sobreyectiva.

7. Sean A y B dos subconjuntos de E. Sea $\mathcal{P}(E)$ la familia de subconjuntos de E. Se considera la aplicación: $f: \mathcal{P}(E) \to \mathcal{P}(A) \times \mathcal{P}(B)$ definida por: $f(X) = (X \cap A, B \cap X)$, $X \in \mathcal{P}(E)$. Determinar una condición necesaria y suficiente para que a) f sea inyectiva. b) f sea sobreyectiva.

8. Sea E el plano $(\mathbb{R} \times \mathbb{R})$, y O un punto de E. Se define en E la relación:

$$P, Q \in E, \quad P\mathcal{R}Q \Longleftrightarrow O, P, Q \text{ están alineados}$$

   ¿Es una relación de equivalencia en E? ¿Es una relación de equivalencia en $E - \{O\}$? Hallar las clases de equivalencia en caso afirmativo.

9. Sea el conjunto $\mathbb{R} \times \mathbb{R}$, y la relación:

$$(x, y), (x', y') \in \mathbb{R} \times \mathbb{R}, \quad (x, y)\mathcal{R}(x', y') \Longleftrightarrow xy = x'y'$$

   ¿Es una relación de equivalencia en $\mathbb{R} \times \mathbb{R}$? Hallar las clases de equivalencia en caso afirmativo. Sea la relación $\mathcal{R}'$ definida por:

$$(x, y), (x', y') \in \mathbb{R} \times \mathbb{R}, \quad (x, y)\mathcal{R}(x', y') \Longleftrightarrow xy = x'y', xx' \geq 0$$

   ¿Es una relación de equivalencia en $\mathbb{R} \times \mathbb{R}$?

10. Sea $f: X \to Y$ una aplicación y $\mathcal{R}$ la siguiente relación en X:

$$x\mathcal{R}x' \Longleftrightarrow f(x) = f(x')$$

   Probar que es una relación de equivalencia. Sea $X/\mathcal{R}$ el conjunto de clases definido por $\mathcal{R}$ y $[x]$ una clase de X. Se define: $\hat{f}: X/\mathcal{R} \to Y$ por $\hat{f}([x]) = f(x)$. Demostrar que $\hat{f}$ es una aplicación y es inyectiva.

11. Sea $A = \{(a, x) \mid a, x \in \mathbb{Q},\ a \neq 0\}$. Se define en A una operación: $(a, x) \cdot (b, y) = (ab, bx + y)$. Demostrar que $(A, \cdot)$ es un grupo no conmutativo. Demostrar que $B = \{(1, x) \mid x \in \mathbb{Q}\}$ es un subgrupo de A.

12. Encontrar todos los subgrupos del grupo de alternaciones A_4.

13. Se considera el grupo cíclico de orden 5, G_5 con generador a. Sea $f: \mathbb{Z} \to G_5$ la aplicación definida por $f(n) = a^n$. Demostrar que es un homomorfismo de grupos y hallar el núcleo ($\ker f$) y la imagen. Calcular el grupo cociente $\mathbb{Z}/\ker f$ y probar que es isomorfo a G_5.

14. Se consideran las funciones $f_a: \mathbb{R} \to \mathbb{R}$, definidas por: $f_a(x) = x + a$. Probar que el conjunto $T = \{f_a \mid a \in \mathbb{R}\}$ es un grupo abeliano, respecto de la composición de funciones.

15. Calcular la tabla de multiplicación para las simetrías del tetraedro y estudiar su relación con algún subgrupo de un grupo de permutaciones.

16. Probar que el conjunto de matrices:

$$\left\{ \begin{pmatrix} a & -b \\ b & a \end{pmatrix} \mid a, b \in \mathbb{R},\ a^2 + b^2 \neq 0 \right\}$$

   con la operación de multiplicación de matrices, es un grupo abeliano. Demostrar que es isomorfo al grupo multiplicativo de los números complejos distintos de cero. Demostrar que (cuando se incluye la matriz nula) es también un cuerpo isomorfo al de los números complejos.

17. Demostrar que cualquier grupo de orden 4 es abeliano y construir todos los grupos de orden 4 no isomorfos.

18. Demostrar que las matrices cuadradas de orden 2 con coeficientes reales de determinante igual a 1 forman un grupo no conmutativo, con respecto al producto. Demostrar que las matrices de la forma:
$$\begin{pmatrix} x & -y \\ y & x \end{pmatrix}, \quad x^2+y^2=1$$
forman un grupo abeliano subgrupo del anterior. ¿Es un subgrupo normal?

19. Sean los grupos $G_1=\{a,b,c,d\}$, $G_2=\{x,y,z,t\}$ con operaciones definidas por las siguientes relaciones:
$$G_1:\left\{\begin{array}{l} a\cdot a=a,\ a\cdot b=b,\ a\cdot c=c,\ a\cdot d=d,\ b\cdot b=a, \\ b\cdot c=d,\ b\cdot d=c,\ c\cdot c=a,\ c\cdot d=b,\ d\cdot d=a \end{array}\right.$$
$$G_2:\left\{\begin{array}{l} x*x=x,\ x*y=y,\ x*z=z,\ x*t=t,\ y*y=x, \\ y*z=t,\ y*t=z,\ z*z=y,\ z*t=x,\ t*t=y \end{array}\right.$$
    (a) Calcular $d\cdot b$, $t*y$.
    (b) Estudiar si existe un homomorfismo de G_1 en G_2 y calcularlo en caso afirmativo.
    (c) Estudiar si G_1 y G_2 son isomorfos. Calcular un isomorfismo en caso afirmativo.

20. Calcular los subgrupos de los grupos $\mathbb{Z}_8$ y $\mathbb{Z}_6$.
    ¿Se podría construir un homomorfismo $f\colon \mathbb{Z}_8\to\mathbb{Z}_6$ siendo $f(1)=3$? ¿Y si fuese $f(2)=3$? En caso afirmativo construirlos explícitamente.

21. Resolver la ecuación $x^2+2x+1=0$ en $\mathbb{Z}_4$.

22. Sea $f\colon \mathbb{Z}_{16}\to\mathbb{Z}_{16}$ un homomorfismo de grupos. Probar que $f([2])\neq[3]$.

23. Sea $q\in\mathbb{Z}_4$ fijado. Considérese el conjunto $F_q=\{(a,b)\mid a,b\in\mathbb{Z}_4,(a,b)\neq([0],[0])\}$ con la operación asociativa,
$$(a,b)\star(c,d)=(ac+qbd,ad+bc).$$
    Hallar el elemento neutro y calcular, si existe, el inverso de $([2],[3])$.

24. En el anillo de polinomios $\mathbb{R}[x]$ se considera (para $a\in\mathbb{R}$ fijado) el conjunto $I_a=\{p(x)\in\mathbb{R}[x]\mid p(a)=0\}$. Demostrar que I_a es un ideal de $\mathbb{R}[x]$ y que $\mathbb{R}[x]/I_a\approx\mathbb{R}$ (como anillos).

25. Un elemento x de un anillo A se dice que es nilpotente si $x^r=0$ para algún $r\in\mathbb{N}$. Determinar:
    (a) Si el conjunto de los elementos nilpotentes de $\mathbb{Z}_8$ forman un ideal.
    (b) Lo mismo en el anillo de matrices reales 2×2.

26. Calcular el grupo de automorfismos de los cuerpos $\mathbb{Q}$, y $\mathbb{F}_2=\{a+b\sqrt{2}\mid a,b\in\mathbb{Q}\}$.

27. Se considera el conjunto de matrices con coeficientes reales:
$$\left\{\begin{pmatrix} a & -b & -c & -d \\ b & a & -d & c \\ c & d & a & -b \\ d & -c & b & a \end{pmatrix}\mid a,b,c,d\in\mathbb{R}\right\}$$
    Demostrar que es un cuerpo no conmutativo. Si:
$$I=\begin{pmatrix} 1&0&0&0\\0&1&0&0\\0&0&1&0\\0&0&0&1\end{pmatrix},$$

$$i = \begin{pmatrix} 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \end{pmatrix}, \; j = \begin{pmatrix} 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix}, \; k = \begin{pmatrix} 0 & 0 & 0 & -1 \\ 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}$$

Demostrar que $\{\pm I, \pm i, \pm j, \pm k\}$ es un grupo no conmutativo con respecto al producto de matrices y escribir la tabla del producto.

28. Demostrar que si I es un subanillo no nulo de $\mathbb{Z}$ entonces I es un ideal. Demostrar que el conjunto:

$$A = \left\{ \begin{pmatrix} a & b \\ 0 & d \end{pmatrix} \mid a, b, d \in \mathbb{Z} \right\}$$

es un subanillo del anillo $\mathcal{M}_2(\mathbb{Z})$ pero no es un ideal.

29. Estudiar la estructura del conjunto de funciones continuas: $C(\mathbb{R}) = \{f : \mathbb{R} \to \mathbb{R}\}$, respecto a las operaciones de suma y producto.

30. Formar las tablas de sumar y multiplicar del anillo $\mathbb{Z}_8$. Hallar todos los divisores de cero y el grupo de elementos invertibles.

31. Demostrar que la condición necesaria y suficiente para que un elemento de un anillo posea inverso es que no pertenezca a ningún ideal propio del anillo.

32. Demostrar que todo homomorfismo de cuerpos distinto del nulo es inyectivo.

33. Demostrar que si los puntos $z_1, \ldots, z_n$ del plano complejo están situados a un mismo lado de una recta que pase por cero entonces: $\sum_{i=1}^n z_i \neq 0$. Demostrar que si $\sum_{i=1}^n z_i^{-1} = 0$, los puntos z_i no pueden estar a un mismo lado de una recta que pase por 0.

34. Calcular las raíces de orden 8 de la unidad y comprobar que su suma es igual a 0. Generalizar el resultado para raíces de orden n.

35. Usar la fórmula de Moivre para calcular $\cos 5x$ en función de $\cos x$ y $\operatorname{sen} x$ y sus potencias.

36. Demostrar que si $z \in \mathbb{C}$ es una raíz del polinomio de coeficientes reales $p(x)$, entonces $\bar{z}$ es también raíz de ese polinomio.

37. Sea z_0 una raíz de la unidad de orden n, distinta de 1. Demostrar que z_0 es raíz del polinomio: $p(z) = z^{n-1} + z^{n-2} + \cdots + z + 1$.

38. Calcular todas las raíces primitivas de la unidad de orden 6. Demostrar que si p es primo, toda raíz de la unidad de orden p distinta de 1 es primitiva.

39. Calcular las raíces y factorizar el polinomio $P(x) = x^{10} - 2x^5 + 1$.

40. Si V es un espacio vectorial y $S \subset V$, probar que $\operatorname{lin} S$ es la intersección de todos los subespacios de V que contienen a S.

41. Si $S_1, \ldots, S_n$ son subconjuntos arbitrarios de un espacio vectorial V, probar que $\operatorname{lin}(S_1 \cup S_2 \cup \ldots \cup S_n) = \operatorname{lin} S_1 + \operatorname{lin} S_2 + \cdots + \operatorname{lin} S_n$. Deducir que $\operatorname{lin}(W_1 \cup W_2 \cup \cdots \cup W_n) = W_1 + W_2 + \cdots + W_n$, si W_i es subespacio, $1 \leq i \leq n$.

42. Probar que si $v_r \in \operatorname{lin}\{v_1, \ldots, v_{r-1}\}$ entonces $\operatorname{lin}\{v_1, \ldots, v_{r-1}\} = \operatorname{lin}\{v_1, \ldots, v_r\}$. Deducir que existe una base B de $\operatorname{lin}\{v_1, \ldots, v_r\}$ tal que $B \subset \{v_1, \ldots, v_r\}$. Demostrar que $\dim \operatorname{lin}\{v_1, \ldots, v_r\} \leq r$, y que $\dim \operatorname{lin}\{v_1, \ldots, v_r\} = r \iff \{v_1, \ldots, v_r\}$ es linealmente independiente.

43. Para cada uno de los siguientes subconjuntos de $\mathbb{C}^3$, estudiar si son o no subespacios vectoriales:

    (a) $\{x \in \mathbb{C}^3 \mid (1-i)x_1 + x_2 - ix_3 = 0\}$

    (b) $\{x \in \mathbb{C}^3 \mid x_1 + x_2 - x_3 + i = 0\}$

(c) $\{x \in \mathbb{C}^3 \mid x_1^2 + x_2^2 - x_3^2 = 0\}$

(d) $\{x \in \mathbb{C}^3 \mid e^{x_1+x_2-x_3} - 1 = 0\}$

Misma pregunta si consideramos los conjuntos c) y d) como subconjuntos de $\mathbb{R}^3$.

44. Decir cuáles de los siguientes subconjuntos de $V = M_n(\mathbb{K})$ son subespacios vectoriales:

$$\begin{aligned} W_1 &= \{A \in V \mid A \text{ es invertible}\} \\ W_2 &= \{A \in V \mid r(A) = n-1\} \\ W_3 &= \left\{A \in V \mid A^t = 2A\right\} \\ W_4 &= \left\{A \in V \mid A^2 - 2A = 0\right\} \\ W_5 &= \{(a_{ij})_{1\le i,j\le n} \in V \mid a_{11} - 2a_{1n} + a_{nn} = 0\}. \end{aligned}$$

45. Estudiar cuáles de los siguientes subconjuntos de $\mathcal{C}(\mathbb{R}, \mathbb{R})$ son subespacios vectoriales:

(a) $\{f \in \mathcal{C}(\mathbb{R}, \mathbb{R}) \mid f(t) = 0,\ \forall t \in \mathbb{Z}\}$

(b) $\{f \in \mathcal{C}(\mathbb{R}, \mathbb{R}) \mid f(0) = \sqrt{2}f(1)\}$

(c) $\{f \in \mathcal{C}(\mathbb{R}, \mathbb{R}) \mid f(0) = 1\}$

(d) $\{f \in \mathcal{C}(\mathbb{R}, \mathbb{R}) \mid f \text{ es derivable dos veces, y } f'' = 0\}$

(e) $\{f \in \mathcal{C}(\mathbb{R}, \mathbb{R}) \mid \exists m, n \in \mathbb{Z} \text{ t.q. } f(t) = mt + n\}$

(f) $\{f \in \mathcal{C}(\mathbb{R}, \mathbb{R}) \mid f^2(0) + f^2(1) = 0\}$

46. Si W es un subespacio vectorial *propio* de un espacio vectorial V, ¿cuál es el subespacio vectorial $\text{lin}(V - W)$ generado por $V - W$? (Nota: $V - W = \{x \in V \mid x \notin W\}$.)

47. Sean W_1 y W_2 dos subespacios vectoriales de un espacio vectorial V. Probar que $W_1 \cap W_2 \neq \{0\}$ si $\dim V < \dim W_1 + \dim W_2$.

48. Se consideran los siguientes subconjuntos de $\mathcal{C}(\mathbb{R}, \mathbb{R})$:

$$\mathcal{P} = \{f \in \mathcal{C}(\mathbb{R}, \mathbb{R}) \mid f(t) = f(-t), \forall t \in \mathbb{R}\}, \quad \mathcal{I} = \{f \in \mathcal{C}(\mathbb{R}, \mathbb{R}) \mid f(t) = -f(-t), \forall t \in \mathbb{R}\}$$

Demostrar que $\mathcal{I}$ y $\mathcal{P}$ son subespacios de $\mathcal{C}(\mathbb{R}, \mathbb{R})$, y que $\mathcal{C}(\mathbb{R}, \mathbb{R}) = \mathcal{P} \oplus \mathcal{I}$.

49. En el espacio vectorial $V = \mathcal{F}(\mathbb{R}, \mathbb{R})$, se consideran los subconjuntos

$$U = \{f \in V \mid f(1) = f(-1) = 0\}, \qquad W = \{f \in V \mid \exists a, b \in \mathbb{R} \text{ t.q. } f(x) = ax + b\}.$$

Probar que U y W son subespacios de V. ¿Se cumple la igualdad $U \oplus W = V$?

50. Sea $B = \{v_1, \dots v_n\}$ una base de un espacio vectorial V, y sea $u_i = \sum_{j=1}^n a_{ij} v_j$, $1 \le i \le r$. Si $A = (a_{ij})_{\substack{1\le i\le r \\ 1\le j\le n}}$, probar que $\dim(\text{lin}\{u_1, \dots u_r\}) = r(A)$.

51. Dados los subespacios $W_1 = \text{lin}\{(1, 2, -1, 0), (0, -i, 0, i)\}$ y $W_2 = \text{lin}\{(3, 1, 0, -1), (5, 6, -2, -2)\}$ en $\mathbb{C}^4$, hallar una base de $W_1 \cap W_2$.

52. Si $W = \text{lin}\{(1, 1, 1, 0,), (0, 1, 1, 0), (1, 1, 0, 1)\}$ y

$$v_1 = (2, -1, 3, 0), \quad v_2 = (1, 1, -1, -1), \quad v_3 = (2, -1, -3, 2),$$

decir cuáles de los vectores v_i pertenecen a W.

53. Dado el subespacio W de $\mathbb{R}^4$ de ecuaciones $x_1 + 2x_2 - x_3 + x_4 = 0$, $2x_1 + x_3 - 3x_4 = 0$, encontrar un subespacio complementario de W.

54. Estudiar la dependencia lineal de los siguientes subconjuntos del espacio vectorial de los polinomios en una indeterminada con coeficientes complejos:

$$\begin{aligned} S_1 &= \{1+it^2, 1+5i+(i-5)t-(1+5i)t^2, i+(1+i)t+(i-1)t^2\} \\ S_2 &= \{1, t, t^2, 1+t, (1+t)^2\} \\ S_3 &= \{1, a-t, (a-t)^2\} \\ S_4 &= \{1+t^2, t-i, t+i\} \\ S_5 &= \{t^2-i, it^2+1, it\}. \end{aligned}$$

55. Si $V = \mathbb{C}_4[x]$, se consideran los polinomios

$$p_1(x) = 3-2x+x^2+4x^3+x^4, \quad p_2(x) = 4-x+x^2+6x^3-2x^4, \quad p_3(x) = 7-8x+3x^2+ax^3+bx^4.$$

Determinar los valores de los parámetros $a, b \in \mathbb{C}$ para los cuáles $W = \operatorname{lin}\{p_1, p_2, p_3\}$ tiene dimensión dos. Para dichos valores de a y b, hallar una base de W y calcular las coordenadas de p_1, p_2 y p_3 en dicha base.

56. Estudiar la dependencia lineal de los siguientes subconjuntos de $\mathbb{C}^3$:

$$\begin{aligned} S_1 &= \{(1, i, 0), (1, 1+i, 0), (1-i, 1+i, 3)\} \\ S_2 &= \{(1, i, -1), (1+5i, -5+i, -1-5i), (i, 1+i, i-1)\} \end{aligned}$$

57. Probar que el subconjunto $S = \{(0,1,0), (1,1,0), (1,2,1), (-2,1,3), (5,2,-3), (1,0,1)\} \subset \mathbb{C}^3$ es un sistema de generadores de $\mathbb{C}^3$, y encontrar una base de $\mathbb{C}^3$ contenida en S.

58. Sean $\lambda_1, \lambda_2, \ldots, \lambda_n$ n números complejos *distintos*. Probar que las funciones $\{f_1, f_2, \ldots, f_n\}$ dadas por $f_i(z) = e^{\lambda_i z}$ son linealmente independientes en $\mathcal{C}(\mathbb{C}, \mathbb{C})$. Utilizar lo anterior para demostrar que los conjuntos $\{1, \cos x, \cos 2x, \ldots, \cos nx\}$ y $\{\operatorname{sen} x, \operatorname{sen} 2x, \ldots, \operatorname{sen} nx\}$ son linealmente independientes en $\mathcal{C}(\mathbb{R}, \mathbb{R})$.

59. Para cada uno de los siguientes pares de bases de $\mathbb{C}^3$, calcular la matriz del cambio de base de B a B':

(a) $B = \{(1,1,0), (-1,1,1), (0,1,2)\}$, $\quad B' = \{(2,1,1), (0,0,1), (-1,1,1)\}$

(b) $B = \{(3,2,1), (0,-2,5), (1,1,2)\}$, $\quad B' = \{(1,1,0), (-1,2,4), (2,-1,1)\}$

60. Sea $V = \mathbb{C}_2[x]$, y sea $a \in \mathbb{R}$ un número real fijo. Si definimos los polinomios

$$p_1(x) = 1, \quad p_2(x) = x+a, \quad p_3(x) = (x+a)^2,$$

probar que $\{p_1, p_2, p_3\}$ es una base de V. ¿Cuáles son las coordenadas de un elemento cualquiera de V en esta base? ¿Sugiere esto alguna generalización?

61. Si $p \in \mathbb{C}_n[x]$, hallar cuál es la condición necesaria y suficiente para que el conjunto $\{p, p', \ldots, p^{(n)}\}$ de las derivadas de p hasta orden n inclusive sea una base de $\mathbb{C}_n[x]$.

62. Dados los subespacios de $\mathbb{R}^4$

$$W_1 = \operatorname{lin}\{(1,0,0,0), (0,-1,0,1)\}, \quad W_2 = \operatorname{lin}\{(0,0,1,1)\}, \quad W_3 = \operatorname{lin}\{(0,1,1,0)\},$$

decir cuáles de las sumas $W_i + W_j$ $(i \neq j)$ y $W_1 + W_2 + W_3$ son sumas directas.

63. Si $W_1, \ldots, W_n$ son subespacios vectoriales de un espacio vectorial V, probar que

$$\begin{aligned} \dim(W_1 + \cdots + W_n) &= \dim W_1 + \cdots + \dim W_n - \dim(W_1 \cap W_2) \\ &\quad - \dim\big((W_1 + W_2) \cap W_3\big) - \cdots - \dim\big((W_1 + \cdots + W_{n-1}) \cap W_n\big). \end{aligned}$$

Deducir que $W_1 + \cdots + W_n$ es suma directa si y sólo si $\dim(W_1 + \cdots + W_n) = \dim W_1 + \cdots + \dim W_n$.

64. Dado el subespacio $U = \{(x_1, \ldots, x_n) \in \mathbb{C}^n \mid x_1 + x_2 + \cdots + x_n = 0\}$ de $\mathbb{C}^n$, estudiar si $W = \{(x_1, \ldots, x_n) \in \mathbb{C}^n \mid x_1 = x_2 = \cdots = x_n\}$ es un subespacio complementario de U.

65. Dados los subespacios

$$W_1 = \operatorname{lin}\{(0,1,-1,0,1),(1,1,-1,1,2)\}, \quad W_2 = \operatorname{lin}\{(-1,0,5,1,0),(a,1,1,-1,b)\}$$

de $\mathbb{C}^5$, dígase para qué valores de los parámetros a y b la suma $W_1 + W_2$ es suma directa.

66. Se consideran los subespacios $U = \operatorname{lin}\{p_1, p_2, p_3\}$ y $W = \operatorname{lin}\{q_1, q_2, q_3\}$ de $\mathbb{C}_4[x]$, siendo

$$\begin{array}{rcl} p_1(x) & = & 1 + 2x + 5x^2 + 3x^3 + 2x^4 \\ p_2(x) & = & 3 + x + 5x^2 - 6x^3 + 6x^4 \\ p_3(x) & = & 1 + x + 3x^2 + 2x^4 \\ q_1(x) & = & 2 + x + 4x^2 - 3x^3 + 4x^4 \\ q_2(x) & = & 3 + x + 3x^2 - 2x^3 + 2x^4 \\ q_3(x) & = & 9 + 2x + 3x^2 - x^3 - 2x^4. \end{array}$$

Hallar una base de $U + W$ y $U \cap W$.

67. Si W_1, W_2 y W_3 son subespacios vectoriales de un espacio vectorial V, estudiar si es cierta la fórmula $(W_1 + W_2) \cap W_3 = (W_1 \cap W_3) + (W_2 \cap W_3)$.

68. Dado el subespacio V_1 de $\mathbb{R}^4$ generado por los vectores:

$$v_1 = (1,1,0,m), \quad v_2 = (3,-1,n,-1), \quad v_3 = (-3,5,m,-4)$$

hallar m y n para que $\dim V_1 = 2$. Para estos m, n calculados, hallar las ecuaciones paramétricas e implícitas de otro subespacio V_2 tal que $V_1 \oplus V_2 = \mathbb{R}^4$.

69. Si W_1, W_2 y W_3 son subespacios vectoriales de un espacio vectorial V, probar que

$$\begin{array}{rcl} \dim(W_1 + W_2 + W_3) & \leq & \dim W_1 + \dim W_2 + \dim W_3 - \dim(W_1 \cap W_2) \\ & & - \dim(W_1 \cap W_3) - \dim(W_2 \cap W_3) + \dim(W_1 \cap W_2 \cap W_3). \end{array}$$

Comprobar que la desigualdad anterior es estricta si

$$V = \mathbb{R}^3, \quad W_1 = \operatorname{lin}\{(1,0,0)\}, \quad W_2 = \operatorname{lin}\{(0,1,0)\}, \quad W_3 = \{(x,y,z) \in \mathbb{R}^3 \mid x - y = 0\}.$$

70. Sea V un espacio vectorial de dimensión 4. Considérense dos subespacios W_1 y W_2 de V tales que $W_1 \cap W_2 = \operatorname{lin}\{v\}$, con $v \neq 0$ y $\dim W_1 = 3$ y $\dim W_2 = 2$. Hallar $W_1 + W_2$.

71. Dados los subespacios U y V de $\mathbb{C}^5$ definidos por los conjuntos de ecuaciones siguientes:

$$U: \left.\begin{array}{rcl} x_1 + x_2 + x_3 + x_5 & = & 0 \\ 4x_2 + 3x_3 - x_4 & = & 0 \\ 4x_1 + x_3 + x_4 + 4x_5 & = & 0 \end{array}\right\}, \quad V: \left.\begin{array}{rcl} x_1 & = & \lambda + \mu + \nu \\ x_2 & = & -\lambda - \nu \\ x_3 & = & 2\lambda + \mu + 2\nu \\ x_4 & = & 0 \\ x_5 & = & \mu \end{array}\right\}, \lambda, \mu, \nu \in \mathbb{C}.$$

Calcular la dimensión y una base de cada uno de ellos. Calcular una base de la suma y otra de la intersección de estos dos subespacios.

72. Se considera el subespacio W de $\mathbb{C}^3$ de ecuación: $ax_1 - ix_2 = 0$, donde $a \in \mathbb{C}$ es una constante. Calcular a de forma que la intersección de W con el subespacio S tenga dimensión 2, donde:

$$S = \operatorname{lin}\{(1,0,-i),(1-i,0,0),(-1-2i,0,3i)\}$$

73. Considérese el espacio de polinomios sobre $\mathbb{R}$ en la variable x de grado menor o igual a n y el subconjunto de polinomios una de cuyas raíces es 1. ¿Es un subespacio vectorial? En caso afirmativo determinar su dimensión y hallar una base.

74. Determinar razonadamente si son ciertas o falsas las proposiciones siguientes, siendo W_1, W_2, W_3 tres subespacios de un espacio vectorial de dimensión finita V.

    (a) $W_1 \cap W_3 = \{\mathbf{0}\}$, $W_2 \cap W_3 = \{0\} \Rightarrow (W_1 + W_2) \cap W_3 = \{0\}$.

    (b) $\dim W_1 \geq \dim V/2$, $\dim W_2 \geq \dim V/2$, $W_1 \cap W_2 = \{0\} \Rightarrow V = W_1 + W_2$.

    (c) $W_1 \cap W_2 \cap W_3 = \{0\}$, $W_1 + W_2 + W_3 = V \Rightarrow V = W_1 \oplus W_2 \oplus W_3$.

75. Sean V_1 y V_2 dos espacios vectoriales sobre un mismo cuerpo $\mathbb{K}$, y considérese el conjunto (producto cartesiano de V_1 y V_2)
$$V_1 \times V_2 = \{(x_1, x_2) \mid x_1 \in V_1, x_2 \in V_2\}.$$

    (a) Demostrar que $V_1 \times V_2$, con las operaciones definidas por
$$\begin{aligned}(x_1, x_2) + (y_1, y_2) &= (x_1 + y_1, x_2 + y_2)\\ \lambda(x_1, x_2) &= (\lambda x_1, \lambda x_2)\end{aligned}$$
    es un espacio vectorial sobre $\mathbb{K}$, y calcular su dimensión en función de las dimensiones de V_1 y V_2.

    (b) Sean ahora W_1 y W_2 subespacios vectoriales de un espacio vectorial V, y sea $T\colon W_1 \times W_2 \to W_1 + W_2$ la aplicación definida por $T(x_1, x_2) = x_1 + x_2$. Demostrar que la suma $W_1 + W_2$ es directa si y sólo si T es un isomorfismo.

76. Se considera el espacio vectorial V_4 sobre $\mathbb{C}$ y la variedad lineal L engendrada por los vectores:
$$L = \{(1, 2+i, 3-i, -i),\ (-1, 1-i, -2+i, 4+i),\ (1, 5+i, 4-i, 4-i)\}$$
    Se pide:

    (a) $\dim L$.

    (b) Ecuaciones paramétricas e implícitas de L.

    (c) Una base del espacio cociente V_4/L.

77. Sea V un espacio vectorial real de dimensión 5 y $\mathcal{B} = \{v_1, v_2, v_3, v_4, v_5\}$ una base de V. Calcular las ecuaciones implícitas (en la base $\mathcal{B}$) del subespacio $W = W_1 + W_2 + W_3$, siendo W_i los subespacios definidos por:
$$W_1 = \operatorname{lin}\{(1, 0, 0, 1, 0), (0, 0, -1, 0, 0)\}$$
$$W_2 : \begin{cases} x_1 + x_2 - x_3 = 0 \\ x_2 - x_5 = 0 \\ x_2 + 2x_5 = 0 \end{cases}$$
$$W_3 = \ker f, \quad f\colon V \to V$$
$$f(x) = (x_4 + x_5, -x_2 + x_3 - x_5, x_2 - x_3, x_3 - x_5, 2x_2 - 3x_3 + x_4 + 4x_5)$$
    Calcular los subespacios $W_1 \cap W_2$, $W_2 \cap W_3$, $W_1 \cap W_3$, $W_1 \cap W_2 \cap W_3$. ¿Es W suma directa de los subespacios W_1, W_2, W_3?

78. En un espacio vectorial V de dimensión 4 se tiene un sistema de generadores formado por los vectores $S = \{u_1, u_2, u_3, u_4, u_5\}$. Se sabe además que $u_1 - u_2 + u_3 = 0$.

    (a) Calcular una base de V a partir de los vectores de S.

(b) Hallar las coordenadas en la base anterior de una base de un subespacio W de V de dimensión 2, cuya intersección con $U = \text{lin}\{u_1, u_2, u_3\}$ es el vector 0 y tal que la suma con el subespacio R cuyas ecuaciones paramétricas en la base encontrada en el apartado a) son

$$-x_1 + 3x_2 - 3x_3 - 2x_4 = 0, \quad -x_1 - 3x_2 + 3x_3 + 4x_4 = 0, \quad 3x_1 + 2x_2 - 2x_3 - 5x_4 = 0,$$

sea directa.

79. Demostrar que toda matriz 2×2, $A = (a_{ij})_{1 \leq i,j \leq 2}$, es raíz del polinomio

$$P(t) = t^2 - (a_{11} + a_{22})t + (a_{11}a_{22} - a_{12}a_{21}).$$

Utilizar este resultado para obtener una expresión de la inversa de una matriz 2×2.

80. Probar que la traza del conmutador de dos matrices es cero. ¿Pueden existir dos matrices P, Q tales que $[P, Q] = iI$?

81. Si $A = \begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix}$, demostrar que $A^2 = 2A - I$, y calcular A^{100}.

82. Dada la matriz $A = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$, hallar A^n, $\forall n \in \mathbb{N}$.

83. Demostrar que las únicas matrices cuadradas de orden n que conmutan con todas las matrices cuadradas del mismo orden son de la forma λI, con $\lambda \in \mathbb{C}$.

84. Por definición, una matriz cuadrada M es *idempotente* si $M^2 = M$. Si A y B son matrices cuadradas tales que $AB = A$ y $BA = B$, probar que A y B son idempotentes. ¿Pueden ser A ó B invertibles?

85. Si A es una matriz cuadrada tal que $A^3 = 0$, definimos la matriz $M(\lambda) = I + \lambda A + \frac{1}{2}\lambda^2 A^2$, $\forall \lambda \in \mathbb{C}$. Probar que el conjunto $\{M(\lambda) \mid \lambda \in \mathbb{C}\}$ es un grupo abeliano respecto del producto de matrices, y calcular $M(\lambda)^{-1}$.

86. Probar que si A y B son matrices cuadradas de la misma dimensión y $AB = I$, entonces también se cumple que $BA = I$; en otras palabras, si B es una inversa por la derecha de A entonces $B = A^{-1}$.

87. Calcular el signo con que aparece el término $a_{n1}a_{n-1,2} \cdots a_{1n}$ en el desarrollo del determinante de la matriz $(a_{ij})_{1 \leq i,j \leq n}$.

88. Probar que si A es una matriz triangular superior, es decir si:

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ 0 & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & a_{nn} \end{pmatrix},$$

entonces $\det A = a_{11}a_{22} \cdots a_{nn}$. Deducir un resultado análogo para las matrices triangulares inferiores.

89. Demostrar que si A y B son dos matrices cuadradas de órdenes n y m, respectivamente, se tiene:

$$\det \begin{pmatrix} A & C \\ 0 & B \end{pmatrix} = \det A \cdot \det B, \qquad \det \begin{pmatrix} C & A \\ B & 0 \end{pmatrix} = (-1)^{nm} \det A \cdot \det B.$$

90. Utilizando las fórmulas del problema anterior, demostrar que $\det(AB) = \det A \cdot \det B$.

91. Calcular los determinantes siguientes:

$$D_1 = \det \begin{pmatrix} x & a & a & \dots & a \\ a & x & a & \dots & a \\ a & a & x & \dots & a \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a & a & a & \dots & x \end{pmatrix}, \qquad D_2 = \det \begin{pmatrix} x_1+a & x_2 & x_3 & \cdots & x_n \\ x_1 & x_2+a & x_3 & \cdots & x_n \\ x_1 & x_2 & x_3+a & \cdots & x_n \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ x_1 & x_2 & x_3 & \cdots & x_n+a \end{pmatrix}.$$

92. Calcular el determinante de Vandermonde

$$W(x_1,\dots,x_n) = \det \begin{pmatrix} 1 & 1 & 1 & \dots & 1 \\ x_1 & x_2 & x_3 & \dots & x_n \\ x_1^2 & x_2^2 & x_3^2 & \dots & x_n^2 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ x_1^{n-1} & x_2^{n-1} & x_3^{n-1} & \dots & x_n^{n-1} \end{pmatrix}$$

93. Calcular los determinantes de las matrices siguientes:

$$A = \begin{pmatrix} 1 & -1 & 1 & 1 \\ 1 & -1 & -1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & 1 & 1 & -1 \end{pmatrix}, \qquad B = \begin{pmatrix} 1 & 1 & 1 & 1 \\ a & b & c & d \\ a^2 & b^2 & c^2 & d^2 \\ a^4 & b^4 & c^4 & d^4 \end{pmatrix}, \qquad C = \begin{pmatrix} a^2 & ab & ab & b^2 \\ ab & a^2 & b^2 & ab \\ ab & b^2 & a^2 & ab \\ b^2 & ab & ab & a^2 \end{pmatrix}.$$

94. Calcular el determinante de orden n

$$\Delta_n = \det \begin{pmatrix} x_1+y_1 & x_1+y_2 & \cdots & x_1+y_n \\ x_2+y_1 & x_2+y_2 & \cdots & x_2+y_n \\ \vdots & \vdots & \vdots & \vdots \\ x_n+y_1 & x_n+y_2 & \cdots & x_n+y_n \end{pmatrix}.$$

95. Una matriz cuadrada es *antisimétrica* si $A^t = -A$. Probar que, si $2 \neq 0$, el determinante de una matriz antisimétrica de orden impar es cero.

96. Una matriz cuadrada $A \in M_n(\mathbb{C})$ se dice *unitaria* si $AA^\dagger = I$. Probar que si A es unitaria entonces $|\det A| = 1$.

97. (*Fórmulas de Schur*) Sean $A, B, C, D \in M_n(\mathbb{C})$, y sea Δ el determinante de la matriz $\begin{pmatrix} A & B \\ C & D \end{pmatrix}$. Probar que se cumplen las igualdades siguientes:

(a) $\Delta = \det(AD - ACA^{-1}B)$, si $\det A \neq 0$

(b) $\Delta = \det(AD - BD^{-1}CD)$, si $\det D \neq 0$

98. Determinar, en función del parámetro a, el rango de las siguientes matrices:

$$\begin{pmatrix} 1 & 1 & 0 & 1 & a \\ 3 & 2 & -1 & 3 & 2a \\ a & 3 & -2 & 0 & a(a-2) \\ -1 & 0 & -4 & 3 & -5a \end{pmatrix}, \qquad \begin{pmatrix} 1 & 1 & -1 & 2 & 0 \\ a & 1 & 1 & 1 & 1+a \\ 1 & -1 & 3 & -3 & 4 \\ 4 & 2 & 0 & a & 4 \end{pmatrix}.$$

99. Hallar los valores de a y b para los cuales el rango de la siguiente matriz es el mínimo posible:

$$\begin{pmatrix} 1 & 3 & -2 & -1 & 4 \\ -2 & 1 & 1 & 2 & -3 \\ 3 & -4 & 3 & 1 & -2 \\ 3 & 3 & 0 & a & 3 \\ 3 & 2 & -3 & -3 & b \end{pmatrix}.$$

100. Hallar el rango de la matriz compleja $A = \begin{pmatrix} 1 & i & 3i & 3 \\ 2+i & 1 & 1+2i & 4+i \\ -1+i & 1+i & 1+i & -1+i \end{pmatrix}$.

101. Demostrar que $r(AB) \leq \min\big(r(A), r(B)\big)$.

102. Probar que $A \in M_{m\times n}(\mathbb{C})$ tiene rango ≤ 1 si y sólo si existen $R \in M_{m\times 1}(\mathbb{C})$ y $S \in M_{1\times n}(\mathbb{C})$ tales que $A = RS$.

103. Sea A una matriz $m \times n$. Demostrar que:
   (a) Si $m > n$, A no tiene inversa por la derecha
   (b) Si $m < n$, hay dos posibilidades:
      i. Si $r(A) = m$, hay infinitas inversas por la derecha de A.
      ii. Si $r(A) < m$, A no tiene inversa por la derecha

104. Sabiendo que $\det \begin{pmatrix} a & 1 & d \\ b & 2 & e \\ c & -1 & f \end{pmatrix} = 1$, hallar el valor del determinante de:

$$\begin{pmatrix} 2a-d & a+d & 3-a \\ 2b-e & b+e & 6-b \\ 2c-f & c+f & -3-c \end{pmatrix}$$

105. Si A y B son matrices invertibles y λ es un escalar, expresar $\operatorname{cof}(\lambda A)$, $\det(\operatorname{cof}(A))$, $\operatorname{cof}(\operatorname{cof}(A))$ y $\operatorname{cof}(AB)$ en función de A, B y λ. ¿Qué ocurre si A ó B no son invertibles?

106. Dada la matriz $m \times n$

$$A = \begin{pmatrix} 1 & 2 & \dots & n-1 & n \\ n+1 & n+2 & \dots & 2n-1 & 2n \\ \vdots & \vdots & \dots & \vdots & \vdots \\ (m-1)n+1 & (m-1)n+2 & \dots & mn-1 & mn \end{pmatrix}$$

con $m, n > 1$, expresar a_{ij} en función de i y j, y calcular el rango de A.

107. Utilizando el método de Gauss–Jordan, determinar cuáles de las siguientes matrices son invertibles, y calcular la matriz inversa cuando esto sea posible:

$$A = \begin{pmatrix} 2 & 5 & -1 \\ 4 & -1 & 2 \\ 6 & 4 & 1 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & -1 & 2 \\ 3 & 2 & 4 \\ 0 & 1 & -2 \end{pmatrix}, \quad C = \begin{pmatrix} 1 & -2 & -1 & 3 \\ -1 & 0 & -2 & 3 \\ 0 & 2 & -1 & -1 \\ -2 & 3 & -1 & -1 \end{pmatrix}.$$

108. Utilizando la fórmula $A^{-1} = \operatorname{cof}(A)^t / \det A$, calcular la inversa de las siguientes matrices:

$$A = \begin{pmatrix} 2 & 3 & 4 \\ 2 & 1 & 1 \\ -1 & 1 & 2 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & 2 & 2 \\ 2 & -1 & 1 \\ 1 & 3 & 2 \end{pmatrix}, \quad C = \begin{pmatrix} 1 & -2 & 1 \\ -2 & 5 & -4 \\ 1 & -4 & 6 \end{pmatrix}.$$

109. Calcular la inversa de la matriz compleja $A = \begin{pmatrix} 1 & 2-i & -1+2i \\ 1 & 2 & -2+i \\ -1 & -2+i & 2-2i \end{pmatrix}$.

110. Si A es una matriz cuadrada cuyos elementos de matriz son números enteros, encontrar una condición necesaria y suficiente para que A^{-1} tenga esta misma propiedad.

111. Determinar cuáles de las siguientes aplicaciones $T: \mathbb{R}^2 \to \mathbb{R}^2$ son lineales:

$$\begin{array}{llll} a)\ T(x,y) = (y,x), & b)\ T(x,y) = (x,0), & c)\ T(x,y) = (x,-y), & d)\ T(x,y) = (x,x) \\ e)\ T(x,y) = (x^2,y^2), & f)\ T(x,y) = (e^x,e^y), & g)\ T(x,y) = (x+1,y+1), & h)\ T(x,y) = (x,1). \end{array}$$

112. Sea V el espacio vectorial de las matrices cuadradas de orden n, sea $Q \in V$ una matriz invertible fija, y considérense las aplicaciones de V en V definidas por:

$$A_1(X) = QX^{-1}, \quad A_2(X) = XX^t, \quad A_3 = X^t - QX, \quad A_4(X) = Q - X^t.$$

Decir cuáles de estas aplicaciones son operadores lineales.

113. Probar que para definir un operador lineal A basta con dar la imagen bajo A de los elementos de una base. Es decir: si $B = \{v_1, \ldots, v_n\}$ es una base de V_1, y $\{w_1, \ldots, w_n\} \subset V_2$, entonces existe un único operador lineal $A: V_1 \to V_2$ tal que $Av_i = w_i$, $1 \leq i \leq n$.

114. Sea V el espacio vectorial sobre $\mathbb{C}$ de todas las funciones de $\mathbb{R}$ en $\mathbb{C}$, y sean

$$f_1(t) = 1, \quad f_2(t) = e^{it}, \quad f_3(t) = e^{-it}.$$

(a) Probar que $B = \{f_1, f_2, f_3\}$ es un conjunto linealmente independiente.

(b) Si $W = \operatorname{lin} B$, sean $g_1(t) = 1$, $g_2(t) = \cos t$, $g_3(t) = \operatorname{sen} t$. Probar que $B' = \{g_1, g_2, g_3\}$ es base de W, y hallar la matriz del cambio de base de B a B'.

115. Si V y W son dos espacios vectoriales de dimensión finita tal que $\dim V > \dim W$, y $A: V \to W$ es un operador lineal, decir cuáles de las siguientes afirmaciones son siempre ciertas:

(a) A es biyectivo

(b) A es no degenerado

(c) A es sobre

(d) A es degenerado

116. Sea $V = \mathbb{C}_n[t]$ el espacio vectorial de los polinomios de grado $\leq n$ con coeficientes complejos, y sea $T: V \to V$ la aplicación definida por $(T \cdot p)(t) = p(t+1)$. Probar que T es lineal y determinar su núcleo e imagen.

117. Sea $V = \mathbb{C}_n[x]$ y $T: V \to V$ la aplicación dada por $(T \cdot p)(x) = p(x+1) + p(x-1) - 2p(x)$.

(a) Probar que T es lineal.

(b) Calcular $T(x^p)$, $0 \leq p \leq n$.

(c) Calcular $\ker(T)$ e $\operatorname{im}(T)$.

(d) Sea $q \in \operatorname{im}(T)$. Probar que existe un único $p \in V$ tal que $T(p) = q$, $p(0) = p'(0) = 0$.

118. Dada $A \in M_n(\mathbb{K})$, sea $F_A: M_n(\mathbb{K}) \to M_n(\mathbb{K})$ la aplicación definida por

$$F_A(X) = [A, X], \quad \forall X \in M_n(\mathbb{K}).$$

Probar que F_A es un operador lineal que cumple $F_A(XY) = F_A(X)Y + XF_A(Y)$.

119. Sea $T: \mathbb{R}^3 \to \mathbb{R}^2$ la aplicación lineal definida por $T(x_1, x_2, x_3) = (x_1 + x_2, x_1 + 2x_3)$. Si $B = \{u_1, u_2, u_3\}$ y $B' = \{v_1, v_2\}$, donde

$$u_1 = (1, 0, -1), \quad u_2 = (1, 1, 1), \quad u_3 = (1, 0, 0); \qquad v_1 = (0, 1), \quad v_2 = (1, 1),$$

hallar la matriz de T respecto de estas bases.

120. Sea $A: V_1 \to V_2$ un operador lineal, $S = \{v_1, \ldots, v_n\} \subset V_1$, y denotemos por $A(S)$ al conjunto $\{Av_1, \ldots, Av_n\}$. ¿Cuáles de las afirmaciones siguientes son verdaderas?

(a) S linealmente dependiente $\Rightarrow A(S)$ linealmente dependiente

(b) S linealmente independiente $\Rightarrow A(S)$ linealmente independiente

(c) $A(S)$ linealmente dependiente $\Rightarrow S$ linealmente dependiente

(d) $A(S)$ linealmente independiente $\Rightarrow S$ linealmente independiente

(e) A no degenerado $\Rightarrow \operatorname{lin} S$ y $\operatorname{lin} A(S)$ tienen la misma dimensión

121. Sea T el endomorfismo de $\mathbb{R}^3$ cuya matriz en la base canónica de $\mathbb{R}^3$ es $A = \begin{pmatrix} 1 & 2 & 1 \\ 0 & 1 & 1 \\ -1 & 3 & 4 \end{pmatrix}$. Calcular una base de la imagen y del núcleo de T.

122. Sea T el endomorfismo de $\mathbb{R}^2$ definido por $T(x, y) = (-y, x)$.

(a) Calcular la matriz de T en la base canónica de $\mathbb{R}^2$

(b) Calcular la matriz de T en la base $\{(1, 2), (1, -1)\}$

(c) Demostrar que para cada número real c, el operador lineal $T - cI$ es invertible.

123. Sea T el endomorfismo de $\mathbb{R}^3$ definido por $T(x, y, z) = (3x + z, -2x + y, -x + 2y + 4z)$.

(a) Calcular la matriz de T en la base $\{(1, 0, 1), (-1, 2, 1), (2, 1, 1)\}$

(b) Demostrar que T no es degenerado y calcular $T^{-1}(x, y, z)$.

124. Si $A: V \to V$ es un operador lineal que cumple la condición $\ker A = \operatorname{im} A$, ¿qué se puede decir de A^2?

125. Sea $A \in M_n(\mathbb{K})$ una matriz fija. Demostrar que las aplicaciones $L_A, R_A: M_n(\mathbb{K}) \to M_n(\mathbb{K})$ definidas por $L_A(X) = AX$, $R_A(X) = XA$, $\forall X \in M_n(\mathbb{K})$, son lineales. Si $n = 2$ y

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

hallar el determinante y la traza de L_A y R_A.

126. Sea $A: \mathbb{C}^3 \to \mathbb{C}^3$ un operador lineal, y $W = \operatorname{lin}\{(0, 1, 2), (1, -1, 1)\}$. Si $Aw = iw$, $\forall w \in W$, y $(0, 1, 1) \in \ker A$, calcular la matriz de A respecto de la base canónica de $\mathbb{C}^3$.

127. Si $V = M_n(\mathbb{K})$ y $A \in M_n(\mathbb{K})$ es una matriz fija, sea T_A el endomorfismo de V definido por $T_A(X) = XA - AX$. Demostrar, *sin calcular la matriz* de T_A, que $\det(T_A) = 0$

128. Si $V = \mathbb{C}_n[t]$ y A es el endomorfismo de V definido por $(A \cdot P)(t) = P'(t + 1)$, $\forall P \in V$, calcular $\ker(A), \operatorname{im}(A), \operatorname{tr}(A)$ y $\det(A)$.

129. Se dice que una matriz $A \in M_n(\mathbb{C})$ es *autoadjunta* si y sólo si $A = A^\dagger$. Si H es el conjunto de todas las matrices autoadjuntas de $M_n(\mathbb{C})$, comprobar que H es un espacio vectorial *real.* ¿Es H subespacio vectorial de $M_n(\mathbb{C})$? Sea $B \in M_n(\mathbb{C})$ una matriz fija; probar que si definimos $T_B(A) = BAB^\dagger$, $\forall A \in H$, entonces T_B es un endomorfismo de H.

130. Sea $V = M_2(\mathbb{C})$, y sea $T: V \to V$ la aplicación definida por $T(X) = X + X^t$, $\forall X \in V$. Calcular $\ker(T)$, $\operatorname{im}(T)$, $\operatorname{tr}(T)$ y $\det(T)$.

131. Sea V el espacio lineal de las funciones continuas de $[-\pi, \pi]$ en $\mathbb{R}$, y considérese el subconjunto $W \subset V$ formado por todos los elementos f de V que verifican las condiciones

$$\int_{-\pi}^{\pi} f(s) ds = \int_{-\pi}^{\pi} f(s) \cos s \, ds = \int_{-\pi}^{\pi} f(s) \operatorname{sen} s \, ds = 0.$$

(a) Demostrar que W es subespacio lineal de V.

(b) Demostrar que, si $n \geq 2$, W contiene a las funciones $f_n(t) = \operatorname{sen} nt$ y $g_n(t) = \cos nt$.

(c) ¿Es W espacio vectorial de dimensión finita?

(d) Sea $T\colon V \to V$ la aplicación dada por

$$(Tf)(t) = \int_{-\pi}^{\pi} [1 + \cos(t-s)] f(s) ds.$$

Demostrar que T es lineal.

(e) Demostrar que $\operatorname{im}(T)$ es de dimensión finita y hallar una base de $\operatorname{im}(T)$.

(f) Calcular el núcleo de T.

(g) Hallar todos los escalares $\lambda \in \mathbb{R} - \{0\}$ y todas las funciones $f \in V - \{0\}$ tales que $T \cdot f = \lambda f$.

132. Sea V un espacio vectorial y $f \in \operatorname{End}(V)$. Demostrar que si $\ker f \cap \operatorname{im} f = \{0\}$, entonces, $\forall x \in V$ existe un único vector $y \in \ker f$ tal que $x - y \in \operatorname{im} f$.

133. La matriz

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

representa a la aplicación lineal $f\colon V \to W$ en las bases $\mathcal{B}_V = \{e_1, e_2, e_3\}$ y $\mathcal{B}_W = \{u_1, u_2, u_3\}$. ¿Existe un cambio de bases en V y W tal que transforme la representación matricial A en la matriz

$$A' = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ -1 & -1 & 0 \end{pmatrix}?$$

Determinar bases del núcleo y la imagen de f.

134. Sea $f\colon V \to V'$ una aplicación lineal entre dos espacios vectoriales de dimensión finita. Sea W un subespacio de V tal que $V = W \oplus \ker f$. Demostrar que si $u, v \in W$ y $u \neq v$ entonces $f(u) \neq f(v)$.

135. Definir una aplicación lineal $f\colon\mathbb{C}^5 \to \mathbb{C}^3$ cuyo núcleo está dado por las ecuaciones:

$$\begin{aligned} x_1 + x_2 - x_3 - x_4 + x_5 &= 0 \\ x_2 + x_3 + x_4 - x_5 &= 0 \end{aligned}$$

y su imagen es el subespacio de $\mathbb{C}^3$ definido por

$$y_1 = \mu + \lambda,\ y_2 = \mu - \lambda,\ y_3 = 2\mu - 3\lambda,\ \lambda, \mu \in \mathbb{C}$$

Hallar una expresión matricial de f.

136. Si f es un endomorfismo del espacio vectorial V tal que $f^2 = 0$, estudiar la veracidad o falsedad de las siguientes proposiciones, probándolas si son ciertas o hallando un contraejemplo si son falsas.

(a) $\dim \ker f = \dim \operatorname{im} f$.

(b) f es diagonalizable.

(c) $f = 0$.

(d) $\dim V \leq 2 \dim \ker f$.

(e) Si A es la matriz asociada a f en una cierta base, la ecuación $AX = b$ tiene solución si $r(A) = \dim \ker f$ y $A \cdot b = 0$.

137. En $\mathbb{R}^5$ se tienen los subespacios:

$$W_1 = \operatorname{lin}\{(0,1,1,1,0), (0,-1,0,1,0), (0,-2,-1,0,0)\}$$

y W_2 definido por las ecuaciones implícitas:

$$x_1 - x_3 = 0, \quad x_1 + x_2 - x_3 + x_4 = 0$$

(a) Calcular los subespacios $V_1 = W_1 + W_2$ y $V_2 = W_1 \cap W_2$

(b) Calcular, si existe, un endomorfismo de $\mathbb{R}^5$ cuya imagen sea igual a V_1 y cuyo núcleo sea V_2.

138. Hallar una base del espacio vectorial V de los polinomios con grado menor o igual que 4 que se anulan en $x = 1$. Considérese el espacio vectorial W de los polinomios de grado menor o igual que 3 y la aplicación $D: V \to W$ definida por la derivada. Hallar una representación matricial de dicha aplicación, su núcleo y su imagen.

139. Sean V, W espacios vectoriales sobre un cuerpo $\mathbb{K}$, y $f: V \to W$, $g: W \to V$, aplicaciones lineales. Estudiar si son ciertas o falsas las siguientes equivalencias (en los dos sentidos):

(a) $\operatorname{im} f \subset \ker g \Leftrightarrow g \circ f = \{0\}$

(b) $\operatorname{im} f \cap \ker g = \{0\} \Leftrightarrow g \circ f$ isomorfismo

(c) $\operatorname{im} f \oplus \ker g = W \Leftrightarrow \dim \ker f + \dim \operatorname{im} g = \dim V$

140. (*Alternativa de Fredholm*) Considérese el sistema de ecuaciones lineales

$$AX = B, \qquad (*)$$

donde A es una matriz cuadrada. 1) Demostrar que $(*)$ tiene solución única para todo valor de B si y sólo si el sistema homogéneo $AX = 0$ no tiene más solución que la trivial. 2) Probar que si el sistema homogéneo tiene solución distinta de la trivial, siempre es posible escoger B de forma que $(*)$ sea incompatible.

141. Calcular mediante el método de eliminación de Gauss todas las soluciones del sistema

$$\begin{aligned} \frac{1}{3}x_1 + 2x_2 - 6x_3 &= 0 \\ -4x_1 + 5x_3 &= 0 \\ -3x_1 + 6x_2 - 13x_3 &= 0 \\ -\frac{7}{3}x_1 + 2x_2 - \frac{8}{3}x_3 &= 0. \end{aligned}$$

142. Hallar todas las soluciones del sistema cuya matriz ampliada es

$$A' = \begin{pmatrix} 2 & -3 & -7 & 5 & 2 & -2 \\ 1 & -2 & -4 & 3 & 1 & -2 \\ 2 & 0 & -4 & 2 & 1 & 3 \\ 1 & -5 & -7 & 6 & 2 & -7 \end{pmatrix}.$$

143. Hallar los valores de a, b y c para los que el sistema lineal

$$\begin{aligned} x_1 - 2x_2 + x_3 + 2x_4 &= a \\ x_1 + x_2 - x_3 + x_4 &= b \\ x_1 + 7x_2 - 5x_3 - x_4 &= c \end{aligned}$$

no tiene solución.

144. Considérese el sistema de ecuaciones cuya matriz ampliada es

$$A' = \begin{pmatrix} 1 & -1 & 2 & 1 \\ 2 & 0 & 2 & 1 \\ 1 & -3 & 4 & 2 \end{pmatrix}.$$

¿Es compatible dicho sistema? Si es así, calcular todas sus soluciones.

145. Si α es un número complejo arbitrario, estudiar y resolver el sistema lineal

$$\begin{aligned} x+\alpha y+\alpha^2 z &= 0\\ \overline{\alpha}x+y+\alpha z &= 0\\ \overline{\alpha}^2 x+\overline{\alpha}y+z &= 0. \end{aligned}$$

146. Si ω es una de las raíces cúbicas de la unidad (i.e. $\omega^3=1$), resolver el sistema lineal

$$\begin{aligned} x+y+z &= a\\ x+\omega y+\omega^2 z &= b\\ x+\omega^2 y+\omega z &= c. \end{aligned}$$

147. Resolver los siguientes sistemas de ecuaciones lineales:

$$\begin{aligned} x_1+x_2+x_3+x_4 &= 0\\ x_2+x_3+x_4+x_5 &= 0\\ x_1+2x_2+3x_3 &= 2\\ x_2+2x_3+3x_4 &= -2\\ x_3+2x_4+3x_5 &= 2 \end{aligned} \qquad \begin{aligned} x_1+x_2-3x_3 &= -1\\ 2x_1+x_2-2x_3 &= 1\\ x_1+x_2+x_3 &= 3\\ x_1+2x_2-3x_3 &= 1 \end{aligned}$$

$$\begin{aligned} x_1+x_2-3x_4-x_5 &= 0\\ x_1-x_2+2x_3-x_4 &= 0\\ 4x_1-2x_2+6x_3+3x_4-4x_5 &= 0\\ 2x_1+4x_2-2x_3+4x_4-7x_5 &= 0 \end{aligned} \qquad \begin{aligned} 2x_1-x_2+x_3-x_4 &= 1\\ 2x_1-x_2-3x_4 &= 2\\ 3x_1-x_3+x_4 &= -3\\ 2x_1+2x_2-2x_3+5x_4 &= -6. \end{aligned}$$

148. Discutir y resolver los siguientes sistemas lineales:

$$\begin{aligned} ax+by+z &= 1\\ x+aby+z &= b\\ x+by+az &= 1 \end{aligned} \qquad \begin{aligned} ax+by+2z &= 1\\ ax+(2b-1)y+3z &= 1\\ ax+by+(b+3)z &= 1 \end{aligned}$$

$$\begin{aligned} ax+by+t &= a+b\\ bx+ay+z &= a-b\\ y+az+bt &= a+1\\ x+bz+at &= a-1 \end{aligned} \qquad \begin{aligned} ax+y+z+t &= 1\\ x+ay+z+t &= b\\ x+y+az+t &= b^2\\ x+y+z+at &= b^3. \end{aligned}$$

149. Discutir y resolver, cuando sea posible, el sistema lineal

$$\begin{aligned} \alpha x_1+\alpha x_2+\cdots+\alpha x_{n-1}+\beta x_n &= a_n\\ \alpha x_1+\alpha x_2+\cdots+\beta x_{n-1}+\alpha x_n &= a_{n-1}\\ &\vdots\\ \alpha x_1+\beta x_2+\cdots+\alpha x_{n-1}+\alpha x_n &= a_2\\ \beta x_1+\alpha x_2+\cdots+\alpha x_{n-1}+\alpha x_n &= a_1. \end{aligned}$$

150. Si $A=\begin{pmatrix} 3 & -6 & 2 & -1\\ -2 & 4 & 1 & 3\\ 0 & 0 & 1 & 1\\ 1 & -2 & 1 & 0 \end{pmatrix}$, decir para qué valores de $B\in M_{4\times 1}(\mathbb{C})$ el sistema lineal $AX=B$ tiene solución.

151. Estudiar según los valores del parámetro a el siguiente sistema:

$$\begin{aligned} (a+1)x+y+z &= a^2+3a\\ x+(a+1)y+z &= a^3+3a^2\\ x+y+(a+1)z &= a^4+a^2 \end{aligned}$$

Resolverlo en los casos en que sea posible.

152. Calcular todas las raíces en $\mathbb{C}$ de los siguientes polinomios:

a) $x^4 - 4x^3 - 19x^2 + 46x + 120$,	*b*) $12x^5 - 16x^4 - 7x^3 - 2x^2 - 62x + 60$,
c) $x^5 - 10x^4 + 29x^3 - 10x^2 - 62x + 60$,	*d*) $x^3 - 7x^2 + 13x - 3$,
e) $x^5 - 4x^4 - 21x^3 - x^2 + 4x + 21$,	*f*) $x^4 - 12x^3 + 47x^2 - 72x + 36$,
g) $6x^5 - 11x^4 - 37x^3 - 51x^2 - 34x - 8$,	*h*) $72x^5 - 228x^4 - 22x^3 + 177x^2 + x - 30$

153. Calcular la multiplicidad de la raíz $x = 1$ de la ecuación $x^{2n} - nx^{n+1} + nx^{n-1} - 1 = 0$.

154. Sea f un polinomio, y supongamos que el operador lineal A es raíz de f, es decir, se cumple la ecuación $f(A) = 0$. Probar que si λ es un autovalor cualquiera de A entonces $f(\lambda) = 0$. Si μ es una raíz cualquiera de f ¿es necesariamente μ un autovalor de A?

155. Sea V el espacio vectorial de los polinomios de grado $\leq n$, y sea $A\colon V \to V$ el operador derivada, definido por

$$A \cdot P = \frac{dP}{dt}, \quad \forall P \in V.$$

Hallar los autovalores y autovectores de A.

156. Se considera el operador lineal T de $\mathbb{R}^3$ cuya matriz en la base canónica $B = \{e_1, e_2, e_3\}$ es:

$$\begin{pmatrix} a+2b & a-b-3c & a-b+3c \\ a-b+c & a+2b+c & a-b-2c \\ a-b-c & a-b+2c & a+2b-c \end{pmatrix}$$

(a) Sean $\hat{e}_1 = e_1 + e_2 + e_3$, $\hat{e}_2 = e_1 - e_2$, $\hat{e}_3 = e_1 - e_3$. Probar que $\hat{B} = \{\hat{e}_1, \hat{e}_2, \hat{e}_3\}$ es base de $\mathbb{R}^3$.

(b) Calcular la matriz del operador lineal T en esta base.

(c) Calcular los polinomios mínimo y característico de T.

(d) ¿Para qué valores de a, b, c es T diagonalizable?

157. Sea V el espacio vectorial de todas las funciones continuas de $\mathbb{R}$ en $\mathbb{R}$, y sea $T\colon V \to V$ el operador lineal definido por

$$(Tf)(t) = \int_0^t f(s)ds.$$

Probar que T no tiene valores propios.

158. Calcular los autovalores y autovectores del operador $\mathbb{C}^n \to \mathbb{C}^n$ cuya matriz en la base canónica de $\mathbb{C}^n$ es

$$\begin{pmatrix} 1 & 1 & \dots & 1 \\ 1 & 1 & \dots & 1 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & \dots & 1 \end{pmatrix}.$$

159. Probar que si $A\colon V \to V$ es un operador diagonalizable, entonces:

(a) $\operatorname{im} A$ es la suma directa de todos los subespacios propios de A asociados a autovalores distintos de cero

(b) $\ker A \oplus \operatorname{im} A = V$

160. Demostrar que toda matriz tiene los mismos autovalores que su matriz transpuesta. Si A es un endomorfismo invertible, probar que A y A^{-1} tienen los mismos autovectores, y hallar la relación existente entre sus autovalores.

161. De un operador lineal $A\colon \mathbb{C}^3 \to \mathbb{C}^3$ se sabe que los vectores $(0,1,1)$, $(1,-1,0)$ y $(1,0,-1)$ son vectores propios, y que la primera columna de A en la base canónica de $\mathbb{C}^3$ es $(1,2,3)^t$. Determinar la matriz de A en la base canónica de $\mathbb{C}^3$.

162. Sabiendo que el endomorfismo f del espacio vectorial real de dimensión finita V verifica

$$f^4 + f - 1_V = 0,$$

estudiar si f es un automorfismo.

163. Sea $f: V \to V$ un endomorfismo de V (espacio vectorial real de dimensión finita), tal que para un cierto $x \in V$ no existe ningún vector $y \in V$ tal que $f(y) = x$. Demostrar que f tiene un autovalor igual a 0.

164. Sea V un $\mathbb{C}$-espacio vectorial de dimensión 3 y $f \in \mathrm{End}(V)$. Se tiene una base de V, $\mathcal{B} = \{u_1, u_2, u_3\}$ y se sabe que:

$$f(u_1) = u_1 - u_2, \ f(u_3) = -u_1 + u_3.$$

Calcular la imagen del vector $v \in V$ cuyas coordenadas en la base $\mathcal{B}$ son: $(1+\sqrt{5}, 2, 0)$, si el subespacio $W = \mathrm{lin}\{u_1 + u_2 - u_3\}$ es invariante bajo f y $\det f = 1$.

165. Calcular una matriz P tal que $P^{-1}AP$ sea diagonal donde:

$$A = \begin{pmatrix} 3 & 0 & -2 & 2 \\ 0 & 1 & 0 & 0 \\ 4 & 0 & -3 & 1 \\ 0 & 0 & 0 & 2 \end{pmatrix}.$$

166. Sea $f: V \to V$, un endomorfismo de un espacio vectorial real de dimensión 3. Sea $\mathcal{B} = \{u_1, u_2, u_3\}$ una base de V. Se sabe que las ecuaciones del núcleo de f en la base $\mathcal{B}$ son: $\{x_1 + x_2 - x_3 = 0,\ x_2 + x_3 = 0\}$, y que los vectores $u_1 + u_2 - u_3,\ u_2 + u_3$ son autovectores de f con autovalores 1 y -1 respectivamente. Calcular la matriz de f en la base $\mathcal{B}$.

167. Para cada una de las matrices siguientes:

$$a) \begin{pmatrix} 5 & -3 & 2 \\ 6 & -4 & 4 \\ 4 & -4 & 5 \end{pmatrix}, \quad b) \begin{pmatrix} 7 & -12 & 6 \\ 10 & -19 & 10 \\ 12 & -24 & 13 \end{pmatrix}, \quad c) \begin{pmatrix} 4 & -5 & 7 \\ 1 & -4 & 9 \\ -4 & 0 & 5 \end{pmatrix},$$

$$d) \begin{pmatrix} 3 & -1 & 1 & -7 \\ 9 & -3 & -7 & -1 \\ 0 & 0 & 4 & -8 \\ 0 & 0 & 2 & -4 \end{pmatrix}, \quad e) \begin{pmatrix} 1 & 2 & 3 \\ 0 & 2 & 3 \\ 0 & 0 & 3 \end{pmatrix}, \quad f) \begin{pmatrix} 9 & -6 & -2 \\ 18 & -12 & -3 \\ 18 & -9 & -6 \end{pmatrix},$$

$$g) \begin{pmatrix} 3 & 2 & 1 & -1 \\ 2 & 2 & 1 & -1 \\ 1 & 1 & 1 & 0 \\ -1 & -1 & 0 & 0 \end{pmatrix}, \quad h) \begin{pmatrix} 4 & 10 & -19 & 4 \\ 1 & 6 & -8 & 3 \\ 1 & 4 & -6 & 2 \\ 0 & -1 & 1 & 0 \end{pmatrix}$$

responder a las siguientes cuestiones:

(a) Calcular el polinomio característico y los valores propios (suponiendo que el cuerpo de base es $\mathbb{C}$)

(b) Para cada valor propio, calcular los vectores propios correspondientes en $\mathbb{C}^n$

(c) Encontrar, cuando exista, una base de $\mathbb{C}^n$ formada por vectores propios

168. Determinar para qué valores de $a, b, c, d \in \mathbb{C}$ el operador $A: \mathbb{C}^2 \to \mathbb{C}^2$ definido por $A(x, y) = (ax + by, cx + dy)$ es diagonalizable. Considerar el mismo problema si $A: \mathbb{R}^2 \to \mathbb{R}^2$.

169. Sea $V = V_1 \oplus V_2$ y $A = A_1 \oplus A_2$, siendo, $A_i \in L(V_i, V_i)$, $i = 1, 2$. Demostrar las siguientes afirmaciones:

(a) $\sigma(A) = \sigma(A_1) \cup \sigma(A_2)$

(b) $V_\lambda = \ker(A_1 - \lambda I_{V_1}) \oplus \ker(A_2 - \lambda I_{V_2})$

(c) A diagonalizable $\Rightarrow A_1$ y A_2 son diagonalizables

170. Sea f un endomorfismo del espacio vectorial real V definido en la base $\mathcal{B} = \{u_1, u_2, u_3, u_4, u_5\}$ por las ecuaciones,

$$f(u_1) = u_1 + u + 2 + u_3 + u_4 + u_5,\ f(u_2) = au_2,\ f(u_3) = bu_3,\ f(u_4) = cu_4,$$

$$f(u_5) = u_1 + u + 2 + u_3 + u_4 + u_5,$$

con $a, b \neq 2$. Estudiar su espectro. ¿Es diagonalizable? ¿Es invertible?

171. Determinar si son ciertas o falsas las siguientes afirmaciones, probándolas en caso positivo o dando un contraejemplo en caso contrario.

(a) Todo polinomio mónico (esto es, cuyo coeficiente del término de mayor grado es 1) es el polinomio característico de algún endomorfismo.

(b) Un polinomio que sólo posee raíces reales ha de ser característico de un endomorfismo real.

(c) Si $p_A(\lambda) = \lambda^n - 1$ el endomorfismo es diagonalizable.

172. Determinar si son ciertas o falsas las siguientes proposiciones (A es un endomorfismo en un espacio vectorial V):

(a) Si λ_1, λ_2 son autovalores de A, entonces $\lambda_1 + \lambda_2$ es un autovalor de A.

(b) Si $\lambda \neq 0$ es un autovalor de A, entonces A no es nilpotente.

(c) Si A es invertible y $\lambda \neq 0$ es un autovalor de A, entonces λ^{-1} también es un autovalor de A.

(d) Si λ es un autovalor de A, entonces λ^n es un autovalor de A^n.

173. Sea la matriz:

$$A = \begin{pmatrix} 0 & 0 & 0 & a \\ 1 & 0 & 0 & b \\ 0 & 1 & 0 & c \\ 0 & 0 & 1 & d \end{pmatrix}$$

Estudiar si es cierta la siguiente afirmación:

$\forall a, b, c, d \in \mathbb{R}$, la matriz A tiene un autovalor con multiplicidad mayor que 1 si y solo si A no es diagonalizable.

174. Sea el endomorfismo de $\mathbb{R}^4$ cuya matriz en la base canónica es:

$$A = \begin{pmatrix} 1 & -1 & 0 & -1 \\ 1 & 1 & -2 & 1 \\ -1 & 1 & 4 & -1 \\ 0 & 2 & 2 & 2 \end{pmatrix}$$

(a) Calcular la forma canónica de Jordan, J, de A.

(b) Calcular una matriz P tal que $PAP^{-1} = J$

(c) Calcular la matriz $B = A^5 - 10A^4 + 40A^3 - 80A^2 + 80A + 32I$.

175. Dada la matriz:

$$A = \begin{pmatrix} 7 & -1 & -1 & 2 \\ 4 & 0 & 4 & 0 \\ -3 & 1 & 5 & -2 \\ -4 & 0 & 4 & 0 \end{pmatrix}$$

(a) Calcular la forma canónica de Jordan de A y la matriz P de cambio de base ($A = PJP^{-1}$).

(b) Hallar un subespacio de $\mathbb{R}^4$ invariante bajo A de dimensión 2.

(c) ¿Cuál es la dimensión de la imagen de la aplicación lineal de $\mathbb{R}^4$ en $\mathbb{R}^4$ cuya matriz es A? ¿Y la del núcleo?

176. Para cada uno de los operadores lineales cuyas matrices en la base canónica de $\mathbb{C}^n$ se dan a continuación, calcular su forma canónica de Jordan y hallar una base en la cual la matriz del operador se reduzca a dicha forma canónica:

$$a)\ \begin{pmatrix} 1 & 1 & 2 & 3 \\ 0 & 1 & 1 & 2 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 2 \end{pmatrix},\quad b)\ \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 2 & 0 & 0 \\ 3 & 0 & 0 & 0 \end{pmatrix},\quad c)\ \begin{pmatrix} -1 & -1 & -2 \\ 0 & -1 & 4 \\ 0 & 0 & 1 \end{pmatrix},$$

$$d)\ \begin{pmatrix} -10 & -9 & -3 & -5 \\ 5 & 4 & 1 & 3 \\ 2 & 2 & 0 & 1 \\ 6 & 6 & 3 & 2 \end{pmatrix},\quad e)\ \begin{pmatrix} 5 & 0 & 1 & 0 & 0 & 0 \\ 0 & 5 & 0 & 1 & 0 & 0 \\ 0 & 0 & 5 & 0 & 1 & 0 \\ 0 & 0 & 0 & 5 & 0 & 1 \\ 0 & 0 & 0 & 0 & 5 & 0 \\ 0 & 0 & 0 & 0 & 0 & 5 \end{pmatrix},$$

$$f)\ \begin{pmatrix} -1 & 0 & -1 & 1 & 0 \\ -4 & -1 & -3 & 2 & 1 \\ -2 & -1 & -2 & 1 & 1 \\ -3 & -1 & -3 & 2 & 1 \\ -8 & -2 & -7 & 5 & 2 \end{pmatrix},\quad g)\ \begin{pmatrix} 3 & -3 & 4 & -1 & -5 \\ 9 & -8 & 10 & -1 & -10 \\ 0 & 0 & 2 & -1 & -1 \\ 5 & -3 & 2 & 2 & 0 \\ -5 & 3 & -1 & -2 & 0 \end{pmatrix},$$

$$h)\ \begin{pmatrix} 3 & 1 & 4 & -3 & 2 \\ -2 & 2 & -4 & 1 & -2 \\ -3 & 0 & -5 & 4 & -3 \\ -2 & 1 & -4 & 2 & -2 \\ 0 & 0 & 0 & -2 & 1 \end{pmatrix},\quad i)\ \begin{pmatrix} -1 & 0 & 0 & 0 & 0 \\ 2 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 \end{pmatrix}$$

177. Se considera la matriz

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

(a) Calcular el rango de A.

(b) Calcular su polinomio característico y su polinomio mínimo.

(c) Calcular una matriz regular P tal que $P^{-1}AP = J_A$ donde J_A es la forma canónica de Jordan de A.

178. Encontrar los valores de $a, b \in \mathbb{R}$ para los que es diagonalizable la siguiente matriz:

$$\begin{pmatrix} 0 & a & 1 \\ 0 & 1 & 0 \\ 0 & 0 & b \end{pmatrix}$$

y diagonalizarla en esos casos.

179. Sea $E = \mathbb{R}_4[x]$ el espacio lineal real de los polinomios de grado menor o igual que 4 con coeficientes reales. Sea la aplicación:

$$\begin{aligned} \phi\colon\ & E \to E \\ & p \mapsto \phi(p) = (x^2 - \lambda^2)p' - 2(2x + \mu)p \end{aligned}$$

con $\lambda, \mu \in \mathbb{R}$ fijos.

(a) Probar que ϕ es una aplicación bien definida y lineal.

(b) Calcular la matriz de ϕ en la base canónica de E.

(c) Calcular, cuando $\lambda = 0$, los autovalores y autovectores de ϕ. ¿Forman estos últimos una base de E?

180. Calcular la exponencial, el seno y el coseno de las siguientes matrices:

$$\begin{pmatrix} 3 & 0 & 0 \\ 4 & 3 & 0 \\ 5 & 1 & -2 \end{pmatrix}, \quad \begin{pmatrix} 1 & 0 & 0 \\ 1 & 2 & 0 \\ 3 & 0 & -3 \end{pmatrix}$$

181. En $\mathbb{R}^3$ sean:

$$v_1 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \quad v_2 = \begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix}, \quad v_3 = \begin{pmatrix} -1 \\ -1 \\ 0 \end{pmatrix}$$

(a) Sea $\omega \in (\mathbb{R}^3)^*$ tal que $\omega(v_1) = 1$, $\omega(v_2) = -1$ y $\omega(v_3) = 3$. Calcular $\omega(x)$ para cualquier $x \in \mathbb{R}^3$.

(b) Describir explícitamente una forma lineal $\mu \in (\mathbb{R}^3)^*$ tal que

$$\mu(v_1) = \mu(v_2) = 0, \quad \mu(v_3) \neq 0$$

(c) Sea $\mu \in (\mathbb{R}^3)^*$ con las propiedades del apartado anterior. Probar que $\mu(x) \neq 0$ si:

$$x = \begin{pmatrix} 2 \\ 3 \\ -1 \end{pmatrix}$$

182. Sea $\mathcal{B} = \{e_1, e_2, e_3\}$ la base de $\mathbb{C}^3$ definida por:

$$e_1 = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}, \quad e_2 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \quad e_3 = \begin{pmatrix} 2 \\ 2 \\ 0 \end{pmatrix}$$

Hallar la base dual de $\mathcal{B}$.

183. Sea $\mathcal{I}$ el espacio lineal $\mathbb{R}_2[x]$ formado por todos los polinomios con coeficientes reales de grado menor o igual que 2. Se consideran las siguientes formas lineales en $\mathcal{I}$:

$$\omega^1(p) = \int_0^1 p(t)dt, \quad \omega^2(p) = \int_0^2 p(t)dt, \quad \omega^3(p) = \int_0^{-1} p(t)dt$$

(a) Probar que $\mathcal{B}^* = \{\omega^1, \omega^2, \omega^3\}$ es una base de $\updownarrow^*$.

(b) Calcular una base $\mathcal{B}$ de $\updownarrow$, que sea la base dual de $\mathcal{B}^*$.

(c) Encontrar $p \in \mathcal{I}$ tal que:

$$\omega^1(p) = a, \quad \omega^2(p) = b, \quad \omega^3(p) = c$$

siendo a, b, c números reales dados.

184. Sea W el subespacio de $\mathbb{R}^5$ generado por los vectores:

$$\begin{aligned} v_1 &= e_1 + 2e_2 + e_3 \\ v_2 &= e_2 + 3e_3 + 3e_4 + e_5 \\ v_3 &= e_1 + 4e_2 + 6e_3 + e_5 \end{aligned}$$

donde $\{e_1, \ldots, e_5\}$ es la base canónica de $\mathbb{R}^5$. Calcular una base del anulador de W.

185. Sea V un espacio de dimensión finita, n, sobre $\mathbb{C}$. Sean μ y ν dos formas lineales no nulas sobre V. Supóngase que no existe ningún escalar $k \in \mathbb{C}$, tal que $\mu = k\nu$. Probar que:

$$\dim(\ker \mu \cap \ker \nu) = n - 2$$

186. Sea $\omega \in (\mathbb{R}^2)^*$ definida por:

$$\omega \begin{pmatrix} x^1 \\ x^2 \end{pmatrix} = a_1 x^1 + a_2 x^2$$

Para cada uno de los siguientes operadores lineales T, calcular $\sigma = T^t \omega$:

$$1)\ T \begin{pmatrix} x^1 \\ x^2 \end{pmatrix} = \begin{pmatrix} x^1 \\ 0 \end{pmatrix}, \quad 2)\ T \begin{pmatrix} x^1 \\ x^2 \end{pmatrix} = \begin{pmatrix} -x^2 \\ x^1 \end{pmatrix}, \quad 3)\ T \begin{pmatrix} x^1 \\ x^2 \end{pmatrix} = \begin{pmatrix} x^1 - x^2 \\ x^1 + x^2 \end{pmatrix}$$

187. Sea $f: V \times V \to \mathbb{C}$ (V espacio vectorial de dimensión finita), una forma bilineal. Demostrar la siguiente equivalencia:

$$f(x, y) = f_1(x) f_2(y) \Longleftrightarrow \operatorname{rango} f = 1$$

donde $f_1, f_2: V \to \mathbb{C}$ son dos formas lineales no nulas.

188. Determinar cuáles de las siguientes funciones $f_i: \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$ son formas bilineales:

$$f_1(u, v) = u_1 v_2 + u_2 v_1, \qquad f_2(u, v) = u_2 - v_2,$$
$$f_3(u, v) = a,\ a = \text{constante} \qquad f_4(u, v) = -2u_1 u_2 + v_1 v_2$$

$$u = u_1 e_1 + u_2 e_2,\ v = v_1 e_1 + v_2 e_2$$

189. Si V es el espacio de polinomios $V = \{p(t) = p_0 + p_1 t + p_2 t^2,\ p_i \in \mathbb{R}\}$, calcular la matriz de la forma bilineal

$$g(p, q) = \int_0^1 p(t) q(t) dt$$

en la base $\{1, t, t^2\}$ ¿Qué vale $g(t^2 - 2, 2t + 4)$?

190. Si $g(u, v) = u_1 v_1 - u_1 v_2 + 3u_2 v_1 - u_2 v_2$ con $u = u_1 e_1 + u_2 e_2 + u_3 e_3,\ v = v_1 e_1 + v_2 e_2 + v_3 e_3$ en la base $\mathcal{B} = \{e_1, e_2, e_3\}$ de $\mathbb{R}^3$, hallar la matriz de g en dicha base. Calcular $g(x, y)$ si $x = 2e_1' + e_3',\ y = -e_2' + 2e_3'$ con $e_1' = e_1 + e_2 + e_3,\ e_2' = -e_2,\ e_3' = e_1 - e_3$.

191. Decir cuáles de las aplicaciones siguientes son formas bilineales:

$$g(A, B) = \operatorname{tr}(A^t B), \quad g(A, B) = \det(AB), \quad g(A, B) = (\operatorname{tr} A)(\operatorname{tr} B)$$
$$A, B \in M_3(\mathbb{R}), \quad (A^t)_{ij} = A_{ji}$$

192. Se considera el espacio $\mathbb{R}^4$ y en él la forma bilineal simétrica cuya matriz en la base canónica es:

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

Se pide:

(a) Estudiar si es definida positiva. En caso contrario calcular el radical y la signatura.

(b) Encontrar una base de $\mathbb{R}^4$ en la que esta forma bilineal esté en su forma canónica.

193. Calcular la matriz de $g(A, B) = \operatorname{tr}(A^t J B)$ en la base

$$E_{11} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad E_{12} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad E_{21} = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \quad E_{22} = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix},$$

con $J = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$.

194. Determinar cuales de las siguientes formas bilineales son equivalentes en $\mathbb{R}$ y en $\mathbb{C}$:

$$\begin{aligned} f_1(x,y) &= x_1y_1 - \frac{1}{2}x_1y_3 - \frac{1}{2}x_3y_1 \\ f_2(x,y) &= \frac{1}{2}x_1y_2 + \frac{1}{2}x_2y_1 - x_3y_3 \\ f_3(x,y) &= \frac{1}{2}x_1y_2 + \frac{1}{2}x_2y_1 + x_3y_3 \end{aligned}$$

195. Reducir a suma de cuadrados y determinar la signatura de la forma cuadrática: $q(v) = x^2 - 4xy + 6y^2 + 2yz - z^2$ ¿Puede ser la matriz asociada a dicha forma cuadrática la matriz de un producto escalar en $\mathbb{R}^3$?

196. Reducir a suma de cuadrados y determinar la signatura de las formas cuadráticas que en una cierta base $\mathcal{B}$ de $\mathbb{R}^3$ vienen representadas por las matrices:

$$a)\ q_{\mathcal{B}} = \begin{pmatrix} 1 & 1 & 2 \\ 1 & 3 & 3 \\ 2 & 3 & 5 \end{pmatrix},\ b)\ q_{\mathcal{B}} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 2 \\ 0 & 2 & -4 \end{pmatrix},\ c)\ q_{\mathcal{B}} = \begin{pmatrix} 1 & -2 & 0 \\ -2 & 2 & 3 \\ 0 & 3 & -1 \end{pmatrix},$$

$$d)\ q_{\mathcal{B}} = \begin{pmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & -2 \end{pmatrix}$$

197. Calcular el rango y la signatura de la forma cuadrática en $\mathbb{R}^n$:

$$q(u) = \sum_{i,j=1}^{n} (i^2 + ij + j^2)u_iu_j, u = u_1e_1 + u_2e_2 + \cdots + u_ne_n, \quad n \geq 3$$

Encontrar una base en la que q sea una suma de cuadrados.

198. Si $u = (u_1, u_2),\ v = (v_1, v_2)$ calcular los valores de a, b, c, d, e para los que:

$$(u, v) = au_1v_1 + bu_1v_2 + cu_2v_1 + du_2v_2 + eu_1v_2^2$$

es un producto escalar en $\mathbb{R}^2$.

199. Demostrar que la fórmula

$$(u, v) = 10u_1v_1 + 3(u_1v_2 + u_2v_1) + 2u_2v_2 + u_2v_3 + u_3v_2 + u_3v_3$$

define un producto escalar en $\mathbb{R}^3$. Hallar una base ortonormal respecto a dicho producto escalar.

200. Calcular la proyección ortogonal del vector de componentes $(1, 1, 0)$ respecto de una base ortonormal de $\mathbb{R}^3$, sobre el subespacio W de $\mathbb{R}^3$ definido por: $W = \{x \in \mathbb{R}^3 \mid x_1 + x_2 + x_3 = 0\}$.

201. Si $W = \text{lin}\,\{(1, 3, 0, 2), (3, 7, -1, 2), (2, 4, -1, 0)\}$ es un subespacio de $\mathbb{R}^4$ con el producto escalar usual, hallar una base ortonormal de $W^\perp$.

202. Sea $V = M_n(\mathbb{R})$.
   (a) Si $B \in M_n(\mathbb{R})$ es una matriz fijada, se define:

$$\begin{aligned} \omega_B\colon\ V &\to \mathbb{R} \\ A &\mapsto \omega_B(A) = \text{tr}(B^tA) \end{aligned}$$

   Probar que $\omega_B \in V^*$.
   (b) Demostrar que para cada $\omega \in V^*$, existe alguna matriz B tal que $\omega = \omega_B$.
   (c) Probar que $B \mapsto \omega_B$ es un isomorfismo de V en V^*.

203. Demostrar que si W es el subespacio de ecuaciones

$$\sum_{j=1}^{n} a_{ij}x_j = 0, \quad i = 1, \ldots, k$$

respecto a una base ortonormal $\{e_1, \ldots, e_n\}$, entonces $W^\perp$ está generado por los vectores:

$$v_i = \sum_{j=1}^{n} a_{ij}e_j, \quad i = 1, \ldots, k$$

.

204. Obtener una base ortonormal en el espacio de polinomios $V = \left\{p(t) = p_0 + p_1 t + p_2 t^2, p_i \in \mathbb{R}\right\}$ con el producto escalar:

$$g(p,q) = \int_{-1}^{1} p(t)q(t)dt$$

205. En $\mathbb{R}^3$ se define la forma cuadrática: $Q(x_1,x_2,x_3) = 2x_1^2 + x_3^2 + 4x_1x_2 + 2x_1x_3 + 2x_2x_3$.

(a) Diagonalizar la forma cuadrática. Sea $\varphi(x,y)$ una forma bilineal simétrica tal que $\varphi(x,x) = Q(x)$. ¿Cuál es su signatura?

(b) Escribir la matriz del cambio de base que diagonaliza φ.

(c) Encontrar una base del conjunto $\{(1,1,1)\}^\perp$.

206. Se considera la forma cuadrática en $\mathbb{R}^3$: $Q(x) = x_1x_2 - x_1x_3$.

(a) Estudiar si Q es definida positiva.

(b) Calcular el radical de la forma bilineal simétrica asociada.

(c) Diagonalizar Q usando el método de Lagrange y calcular su signatura.

207. Se considera el espacio vectorial $\mathbb{R}^3$ con el producto escalar definido por:

$$(x,y) = x_1y_1 + x_2y_1 + x_1y_2 + 2x_2y_2 + x_3y_2 + x_2y_3 + 2x_3y_3$$

(a) Calcular una base ortonormal del subespacio: $x_2 - 2x_3 = 0$.

(b) Calcular la distancia del vector $(0,1,-2)$ al subespacio anterior.

(c) Estudiar si el operador dado por la siguiente expresión es simétrico:

$$x_1' = x_1, \quad x_2' = x_3, \quad x_3' = x_2 + x_3$$

208. Sea V el espacio vectorial real de los polinomios de grado menor o igual que 2 y sean $\alpha, \beta \in \mathbb{R}$. Dada la aplicación $q\colon V \to \mathbb{R}$, definida por:

$$q(p) = p(\alpha)p(\beta), \quad p \in V$$

(a) Probar que q es una forma cuadrática en V.

(b) Hallar la matriz asociada a q respecto de la base $\{1, x, x^2\}$ y dar el rango y signatura para los distintos valores de α y β.

209. Sea V un espacio vectorial real de dimensión n y φ una forma bilineal simétrica de signatura (p,q) con $p \geq q > 0$, $p + q = n$. Se dice que un vector $x \in V$ es isótropo si $\varphi(x,x) = 0$. Estudiar si son ciertas o falsas las siguientes afirmaciones:

(a) En V existen vectores isótropos distintos de 0.

(b) El conjunto de todos los vectores isótropos forma un subespacio de V.

(c) Hay subespacios de V (distintos del trivial $\{0\}$) cuyos vectores son todos isótropos.

(d) La forma bilineal φ es igual a cero cuando se restringe a un subespacio formado por vectores isótropos.

(e) Existen subespacios de vectores isótropos con dimensión igual a q.

210. Considérese la forma bilineal en $\mathbb{R}^3$ definida por

$$\phi(x,y) = x_3y_1 + x_2y_2 + x_1y_3 + ax_3y_3.$$

Diagonalizarla. Si $a = 3$ hallar sus vectores isótropos (es decir $\phi(x,x) = 0$). ¿Forman un subespacio? ¿Existe algún a tal que ϕ sea definida positiva?

211. En el espacio vectorial $\mathbb{R}^3$ se considera la forma bilineal simétrica ϕ cuya forma cuadrática asociada es

$$q_\phi(x_1,x_2,x_3) = 3x_1^2 - 4x_1x_2 - 6x_1x_3 + 3x_2^2 + 4x_2x_3 + 4x_3^2$$

Comprobar, aplicando el método de Lagrange, que ϕ define un producto escalar y hallar una base ortonormal respecto de este producto escalar.

212. Sea V un espacio vectorial complejo de dimensión 4, $(\ ,\ )$ un producto escalar en V, $\mathcal{B} = \{u_1,u_2,u_3,u_4\}$ una base ortonormal de V y W el subespacio generado por los vectores w_1,w_2 cuyas coordenadas en la base $\mathcal{B}$ son $(1-i,0,1+i,0),(1,0,0,1)$ respectivamente. Sabiendo que w_1,w_2 son autovectores de autovalor -1 de un operador autoadjunto $\mathcal{A}$ en V cuyo otro autovalor (de multiplicidad 2) es 1, calcular una base ortonormal de V formada por autovectores de $\mathcal{A}$ y el proyector ortogonal sobre el subespacio W. ¿Es $\mathcal{A}$ unitario?

213. Sea V un espacio vectorial complejo de dimensión finita con un producto escalar y sea $A\colon V \to V$ un operador autoadjunto. Si $R = A + i1_V$, demostrar:

(a) $\|Ru\|^2 = \|Au\|^2 + \|u\|^2$, $\forall u \in V$.

(b) R es un operador inversible.

(c) $(A - i1_V)(A + i1_V)^{-1}$ es unitario.

214. Sea A el operador en $\mathbb{R}^3$ con el producto escalar usual, cuya matriz asociada respecto de la base canónica es

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

(a) Obtener una matriz ortogonal P tal que $P^t\mathcal{A}P$ sea diagonal.

(b) Dar la descomposición espectral del operador A.

215. Se considera la matriz en $\mathcal{M}(\mathbb{R}^4)$:

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

(a) Sea V un espacio vectorial real dotado de un producto escalar. Si A es la matriz de un endomorfismo f de V en una base ortonormal $\mathcal{B}$, calcular bases del núcleo y la imagen.

(b) En la situación descrita en a), calcular una base ortonormal de V formada por autovectores de f, y hallar su descomposición espectral. Encontrar una matriz ortogonal P, tal que P^tAP sea diagonal.

216. Sea $\mathcal{B} = \{u_1,u_2,u_3\}$ una base ortonormal de $\mathbb{R}^3$ respecto al producto escalar usual ($(x,y) = \sum_{i=1}^3 x_iy_i$ en la base canónica). Se define un operador lineal, T, mediante: $T(u_1) = 5u_1 + 2u_2 + 4u_3$, $\quad T(u_2) = 2u_1 + 8u_2 - 2u_3$, $\quad T(u_3) = 4u_1 - 2u_2 + 5u_3$.

(a) Encontrar una base ortonormal de $\mathbb{R}^3$, $\mathcal{B}' = \{v_1, v_2, v_3\}$, tal que: $Tv_1 = a_1 v_1$, $Tv_2 = a_2 v_2$, $Tv_3 = a_3 v_3$ y calcular $a_1, a_2, a_3 \in \mathbb{R}$.

(b) Calcular la descomposición espectral de T en la base $\mathcal{B}$.

217. Diagonalizar mediante una transformación ortogonal el operador que en una cierta base ortonormal $\mathcal{B}$ viene representado por la matriz:

$$A_B = \begin{pmatrix} 2 & -2 \\ -2 & 5 \end{pmatrix}$$

Utilizar dicha transformación para reducir a suma de cuadrados la forma cuadrática

$$q(v) = 2v_1^2 - 4v_1v_2 + 5v_2^2$$

218. Sea A un operador autoadjunto en el espacio vectorial $\mathbb{C}^n$, dotado del producto escalar usual: $(x, y) = \sum_{i=1}^n \bar{x}_i y_i$. Sean: $u = (1, 0, \dots, 0, i)$, $v = (1, 0, \dots, 0, 1)$ dos autovectores de A, con autovalores λ, μ respectivamente. Calcular λ en función de μ.

219. En $\operatorname{End}(V)$ se define el producto escalar $(A, B) = \operatorname{tr}(A^t B)$:

(a) Calcular el complemento ortogonal del subespacio de los operadores simétricos $S(V)^\perp$

(b) Si $V = \mathbb{R}^3$ describir la descomposición $\operatorname{End}(V) = S(V) \oplus S(V)^\perp$

220. Calcular, si existe, una base ortonormal de $\mathbb{C}^4$ (con el producto escalar usual), en la que sea diagonal el operador que en la base canónica tiene como matriz:

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

Calcular la descomposición espectral de este operador.

221. Calcular la proyección ortogonal del vector $v = e_1 + 2e_3$ de $\mathbb{R}^3$ sobre el subespacio

$$S = W^\perp, W = \operatorname{lin}\{e_1 - e_2\}$$

$\mathcal{B} = \{e_1, e_2, e_3\}$ es una base ortonormal de $\mathbb{R}^3$

222. Escribir la matriz que representa una rotación en el plano perpendicular al vector $(0, 1, 0)$.

223. Calcular un valor de $a \in \mathbb{R}$ para el que la transformación de $\mathbb{R}^3$, representada por la siguiente matriz en una base ortonormal, sea una rotación.

$$R = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & a \end{pmatrix}$$

Calcular en ese caso el eje y el ángulo de rotación.

224. Determinar las matrices A para las que e^{tA} es ortogonal.

225. ¿Cuántas rotaciones existen en $\mathbb{R}^3$ que lleven el vector $(1, 1, 1)$ en el $(0, 1, 1)$? ¿Y cuántas que lleven el vector $(1, 0, 0)$ en el $(0, 1, 0)$?

226. Encontrar los valores de $\lambda \in \mathbb{R}$ que hacen a las siguientes formas cuadráticas definidas positivas:

$$\begin{array}{ll} a) & 5x_1^2 + x_2^2 + \lambda x_3^2 + 4x_1x_2 - 2x_1x_3 - 2x_2x_3 \\ b) & 2x_1^2 + x_2^2 + 3x_3^2 + 2\lambda x_1x_2 + 2x_1x_3 \end{array}$$

y diagonalizar las formas definidas positivas por Gram-Schmidt.

227. Determinar si las aplicaciones siguientes son sesquilineales, calcular las formas hermíticas asociadas y diagonalizarlas:

$$\begin{array}{ll} a)\ f: \mathbb{C}^2 \times \mathbb{C}^2 \to \mathbb{C} & f(x,y) = \bar{x}_1 y_1 - i\bar{x}_2 y_1 + i\bar{x}_1 y_2 + 2\bar{x}_2 y_2 \\ b)\ g: \mathbb{C}^3 \times \mathbb{C}^3 \to \mathbb{C} & g(x,y) = -i\bar{x}_1 y_2 + i\bar{x}_2 y_1 - \bar{x}_3 y_1 - \bar{x}_1 y_3 + \bar{x}_2 y_3 + \bar{x}_3 y_2 \end{array}$$

¿Son definidas positivas? En caso afirmativo diagonalizarlas usando Gram-Schmidt.

228. Sean L_1 y L_2 dos subespacios de un espacio de Hilbert de dimensión finita, y $\dim L_1 < \dim L_2$. Probar que existe en L_2 un vector no nulo ortogonal a L_1.

229. Calcular el vector del subespacio de $\mathbb{R}^4$ dado por:

$$\left.\begin{array}{rcl} 2x_1 + x_2 + x_3 + 3x_4 & = & 0 \\ 3x_1 + 2x_2 + 2x_3 + x_4 & = & 0 \\ x_1 + 2x_2 + 2x_3 - 9x_4 & = & 0 \end{array}\right\}$$

que más se aproxima (en el sentido de la norma que deriva del producto escalar usual de $\mathbb{R}^4$) al vector $(7, -4, -1, 2)$.

230. Sea $\{u_1, u_2\}$ una base ortonormal del plano y la matriz de la aplicación lineal ϕ en la base $\{v_1 = u_1, v_2 = u_1 + u_2\}$:

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

Calcular la matriz de ϕ^t en la base $\{v_1, v_2\}$.

231. Sea $(E, (\ ,\ ))$ un espacio euclidiano, $x, y \in E$ dos vectores no nulos. Estudiar si son equivalentes las siguientes afirmaciones:

$$a)\ x \perp y, \quad b)\ \|x + \lambda y\| \geq \|x\|,\ \forall \lambda \in \mathbb{R}$$

232. Sea $(E, (\ ,\ ))$ un espacio euclidiano de dimensión 4, y $\mathcal{B} = \{e_1, e_2, e_3, e_4\}$ una base ortonormal. Describir todos los operadores ortogonales cuya matriz en la base $\mathcal{B}$ es cuasitriangular superior.

233. Sea $(E, (\ ,\ ))$ un espacio euclidiano de dimensión finita y A un operador lineal simétrico en E. Probar que si $A^k = I$ para algún entero positivo k, entonces $A^2 = I$.

234. Sea $(E, (\ ,\ ))$ un espacio euclidiano de dimensión $n \geq 2$ y sean v, w dos vectores no nulos. Sea $q: E \to \mathbb{R}$ la forma cuadrática definida por:

$$q(x) = (v, w)(x, x) - (v, x)(w, x)$$

(a) Calcular la forma bilineal simétrica $f_q: E \times E \to \mathbb{R}$ tal que: $q(x) = f_q(x, x)$

(b) Calcular el operador lineal simétrico $A_q: E \to E$ tal que: $q(x) = (x, A_q x)$.

(c) Suponiendo que $(v, w) = 0$, calcular $\ker A_q$.

235. Se considera el operador simétrico $T: \mathbb{R}^4 \to \mathbb{R}^4$ (dotamos a $\mathbb{R}^4$ del producto escalar usual) cuya matriz en la base canónica es:

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

Calcular una base ortonormal de $\mathbb{R}^4$ formada por vectores propios de T y encontrar una matriz ortogonal P tal que $P^t A P$ sea diagonal.

236. Sea E un espacio vectorial complejo de dimensión finita n y $(\ ,\ )$ un producto escalar en E. Sea $A\colon E \to E$ un operador autoadjunto con valores propios $\lambda_1 \geq \ldots \geq \lambda_n$. Considérese un vector $x \in E$ unitario. Probar que:

$$\lambda_n \leq (x, Ax) \leq \lambda_1$$

y deducir que:

$$\begin{aligned} (x, Ax) = \lambda_1 \quad &\Longleftrightarrow Ax = \lambda_1 x \\ (x, Ax) = \lambda_n \quad &\Longleftrightarrow Ax = \lambda_n x \end{aligned}$$

Si $T\colon E \to E$ es un operador lineal, probar que T^+T es autoadjunto y positivo:

$$(x, T^+Tx) \geq 0, \quad \forall x \in E$$

237. Se considera la matriz real simétrica:

$$A = \begin{pmatrix} -2 & -2 & -4 \\ -2 & 1 & -2 \\ -4 & -2 & -2 \end{pmatrix}$$

Calcular la descomposición espectral de A.

238. La siguiente matriz representa una rotación en $\mathbb{R}^3$:

$$\frac{1}{9}\begin{pmatrix} 8 & 1 & -4 \\ -4 & 4 & -7 \\ 1 & 8 & 4 \end{pmatrix}$$

Calcular la dirección del eje de rotación y el ángulo de giro.

239. Calcular una matriz ortogonal $P \in M_3(\mathbb{R})$, tal que P^tAP sea diagonal, siendo:

$$A = \begin{pmatrix} 6 & -2 & 2 \\ -2 & 5 & 0 \\ 2 & 0 & 7 \end{pmatrix}$$

240. Se considera la matriz:

$$A = \begin{pmatrix} 5 & 2 & 2 \\ 2 & 2 & -4 \\ 2 & -4 & 2 \end{pmatrix}$$

(a) Si A es la matriz de un operador lineal en $\mathbb{R}^3$ respecto a una base ortonormal, ¿de qué tipo es ese operador?

(b) Calcular una matriz ortogonal P tal que P^tAP sea diagonal.

(c) Descomponer A como combinación lineal de proyectores ortogonales (descomposición espectral).

241. Calcular una matriz ortogonal P tal que P^tAP sea diagonal, donde:

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

y calcular la descomposición espectral de A.

242. Sea $f\colon \mathbb{R}^3 \times \mathbb{R}^3 \to \mathbb{R}$, la forma bilineal simétrica cuyas ecuaciones referidas a la base canónica de $\mathbb{R}^3$ son:

$$f(x, y) = x^t \begin{pmatrix} 1 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{pmatrix} y$$

(a) Comprobar que f es definida positiva. Tómese f como un producto escalar en $\mathbb{R}^3$ y calcúlese una base ortonormal de $\mathbb{R}^3$ respecto a este producto escalar, aplicando el procedimiento de Gram-Schmidt a la base canónica de $\mathbb{R}^3$.

(b) Se considera la transformación lineal $T\colon \mathbb{R}^3 \to \mathbb{R}^3$ dada por sus ecuaciones en la base canónica:

$$T(x) = \begin{pmatrix} 3 & 0 & -2 \\ 2 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix} x$$

Comprobar que T es un operador simétrico en el espacio euclidiano $(\mathbb{R}^3, f)$.

(c) Calcular la descomposición espectral de T que deberá expresarse en la base canónica de $\mathbb{R}^3$.

243. Se considera la matriz simétrica:

$$A = \begin{pmatrix} 3 & 2 & 4 \\ 2 & 0 & 2 \\ 4 & 2 & 3 \end{pmatrix}$$

Calcular la descomposición espectral de esta matriz (se considera que A es la matriz en la base canónica de un operador simétrico de $\mathbb{R}^3$ con el producto escalar usual).

244. Dada la matriz:

$$A = \begin{pmatrix} 0 & 1 & -2 \\ -1 & 0 & -1 \\ 2 & 1 & 0 \end{pmatrix} \in M_3(\mathbb{C})$$

(a) Probar que es normal y calcular su espectro.

(b) Encontrar una matriz unitaria U, tal que UAU^+ sea diagonal.

245. Sea el operador cuya matriz en la base canónica de $\mathbb{C}^3$, con el producto escalar usual, es:

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

(a) Calcular una base ortonormal de autovectores.

(b) Calcular la descomposición espectral.

(c) Encontrar la distancia del vector $x = (1, 0, i)$ al subespacio lineal correspondiente al autovalor máximo.

(d) Calcular e^A.

246. Sea V un espacio vectorial real de dimensión 2, dotado de un producto escalar, y sea $\mathcal{B} = \{u_1, u_2\}$ una base de V. Sea A un operador simétrico en V respecto a ese producto escalar, tal que: $Au_1 = 2u_1 + 2u_2$, $Au_2 = u_1 - 2u_2$. Sabiendo que los vectores u_1 y u_2 tienen norma igual a 1, calcular el producto escalar de u_1 por u_2.

247. Sea el operador cuya matriz en la base canónica de $\mathbb{C}^3$ (con el producto escalar usual) es $A = \begin{pmatrix} 1 & i & 1 \\ -i & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}$.

(a) Decir qué tipo de operador es.

(b) Calcular los autovalores.

(c) Hallar, si existe, una base ortonormal de autovectores.

(d) Calcular la descomposición espectral.

(e) Calcular $\cos(\pi A)$.

248. Probar que el tensor $a\delta_{ik}\delta_{jl} + b\delta_{il}\delta_{jk} + c\delta_{ij}\delta_{kl}$ es invariante bajo transformaciones ortogonales.

249. Las componentes de un tensor 3 veces covariante referidas a una base ortonormal de $\mathbb{R}^2$ son todas 1. Calcular las componentes referidas a la base ortonormal girada 90º respecto a la primera.

250. Demostrar las siguientes relaciones en $\mathbb{R}^3$:

$$\epsilon^{kij}\epsilon_{jlm}x_i y^l z^m = x^i z_i y^k - x^i y_i z^k, \quad \epsilon^{ijk}\epsilon_{ilm}x_j y_k x^l y^m = x_i x^i y_j y^j - (x_i y^i)^2$$

251. Dados los vectores x e y en $\mathbb{R}^3$, escribir las componentes de $(x \wedge y)_i = \epsilon_{ijk}x^j y^k$ y decir si es tensor y de qué tipo.

252. Dados x e y, vectores de $\mathbb{R}^2$ definidos por $x^1 = 1$, $x^2 = -1$, $y^1 = 0$ e $y^2 = 2$ y el tensor métrico: $g_{11} = 1$, $g_{12} = g_{21} = 1/2$, $g_{22} = 2$, hallar: (a) $x^i x_i$, (b) $y^i y_i$ y (c) $y^i x_i$.

253. Sean v y w dos vectores de $\mathbb{R}^n$ de norma unidad y ortogonales entre sí. Hallar el valor del escalar:

$$\epsilon_{ijk}\delta^{kl}v_l v^j w^i + v^k\delta_{kl}w^l + \delta^{ij}\delta_{ji}v^k\delta_{kl}v^l$$

254. Se consideran las matrices $\gamma^\mu \in \mathcal{M}_4(\mathbb{C})$, $\mu = 0, 1, 2, 3$, que verifican: $\gamma^\mu\gamma^\nu + \gamma^\nu\gamma^\mu = 2g^{\mu\nu}I_4$. Supongamos que γ^μ son las componentes de un tensor de tipo $(1, 0)$ y $g^{\mu\nu}$ las de un tensor invariante de tipo $(2, 0)$, con valores: $g^{00} = -g^{11} = -g^{22} = -g^{33} = 1$, $g^{\mu\nu} = 0$, $\mu \neq \nu$. Calcular el número de componentes linealmente independientes (en el espacio $\mathcal{M}_4(\mathbb{C})$) de los tensores: $\gamma^\mu\gamma^\nu$, $\gamma^\mu\gamma^\nu\gamma^\rho$, $\gamma^\mu\gamma^\nu\gamma^\rho\gamma^\sigma$.

255. Sea V un espacio vectorial y los tensores A_μ, $T_{\mu\nu}$, $g_{\mu\nu}$, donde $T_{\mu\nu}$ es simétrico y $g_{\mu\nu}$ define un producto escalar en V. Construir el escalar más general que se puede formar con los tensores A_μ y $T_{\mu\nu}$ mediante una combinación lineal de productos tensoriales hasta orden 3 y contracciones.

256. En el espacio $\mathcal{M}_n(\mathbb{C})$ se considera el conjunto de matrices linealmente independientes, $\{X_1, \cdots, X_r\}$, que generan un subespacio W. Supongamos que el conmutador de dos matrices de W es una matriz de W, es decir:

$$[X_i, X_j] = c_{ijk}X_k, \; i, j, k = 1, \cdots, r$$

Demostrar que c_{ijk} es un tensor bajo transformaciones asociadas a cambios de base en W. ¿De qué tipo?

257. Sea $A_{\mu\nu}$ un tensor simétrico en el espacio $\mathbb{R}^3$ con tensor métrico $g_{\mu\nu}$. Sean λ_i, $i = 1, 2, 3$ los autovalores de $A^\mu{}_\nu$. Demostrar que:

$$\sum_{i=1}^{3}\lambda_i = A^\mu{}_\mu, \quad \sum_{i=1}^{3}\lambda_i^2 = A^{\mu\nu}A_{\mu\nu}, \quad \sum_{i=1}^{3}\lambda_i^3 = A^{\mu\nu}A_{\nu\rho}A^\rho{}_\mu$$

258. Sea V un $\mathbb{R}$-espacio vectorial con producto escalar y $\mathcal{B} = \{u_1, u_2, u_3\}$ una base de V, con tensor métrico: $g_{11} = 3$, $g_{22} = 2$, $g_{33} = 1$, $g_{12} = g_{21} = 1$, $g_{13} = g_{31} = 1$, $g_{23} = g_{32} = 0$.

   (a) Dado un vector x cuyas coordenadas covariantes son $x_1 = x_2 = x_3 = 1$, hallar sus coordenadas contravariantes. Si $y = u_1 + u_2 + u_3$, ¿cuáles son sus coordenadas covariantes?

   (b) Sea el tensor A una vez contravariante y 2 veces covariante cuyas componentes referidas a $\mathcal{B}$ son:

$$A^i{}_{jk} = \delta^i{}_j x_k + \delta^i{}_k x_j$$

   Calcular $A^{ijk}y_i y_j y_k$.

259. Considérese el espacio $\mathbb{R}^3$ y un tensor métrico que en cierta base $\mathcal{B}$ viene determinado por $g_{11} = g_{22} = -g_{33} = 1$, $g_{ij} = 0$ si $i \neq j$. Sean los siguientes tensores definidos por sus coordenadas:

$$x^1 = 1, \; x^2 = 2, \; x^3 = 3; \qquad a^1{}_1 = a^1{}_3 = a^2{}_1 = a^3{}_1 = a^3{}_2 = a^3{}_3 = 1, \; a^1{}_2 = a^2{}_2 = a^2{}_3 = 2$$

Calcular: a) $x_i x^i$. b) $a^{ij}x_i x_j$. c) $a^{ij}a_{ij}$. d) $\epsilon^{ijk}a_{ij}x_k$.

260. Las componentes de un tensor 3 veces covariante referidas a una base ortonormal de $\mathbb{R}^3$ son todas iguales a 1. Hallar sus componentes referidas a la base que resulta al girar la base dada un ángulo de $\pi/4$ respecto del primer eje.

261. En el espacio $\mathbb{R}^2$ se considera el producto escalar cuya matriz respecto a una base $\mathcal{B} = \{u_1, u_2\}$ es:

$$\begin{pmatrix} 4 & 2 \\ 2 & 2 \end{pmatrix}$$

(a) Hallar las coordenadas contravariantes y covariantes del vector $2u_1 + u_2$ en la base $\mathcal{B}' = \{u_1 + 2u_2, u_1 - u_2\}$.

(b) Estudiar si son ciertas las siguientes igualdades:

i. $(u_1 + u_2) \otimes (2u_1 - u_2) + (u_1 + 2u_2) \otimes u_2 = (2u_1 + u_2) \otimes u_1 + u_2 \otimes (u_1 + u_2)$

ii. $(u_1 - u_2) \otimes (u_1 + u_2) = (u_1 + u_2) \otimes (u_1 - u_2)$

(c) Dados los tensores cuyas componentes referidas a la base $\mathcal{B}$ son: $r_k^{ij} = i(2-k)$, $\quad s_{jk}^i = (i+1)j$, hallar las componentes respecto de $\mathcal{B}'$ del tensor cuyas componentes respecto de $\mathcal{B}$ son: $r_k^{ij} s_{il}^k$.

Soluciones

Las soluciones que aquí aparecen son simplemente resultados numéricos de los problemas o bien indicaciones escuetas de como resolverlos.

1. Todas son ciertas.
2. La inclusión $f(A \cap B) \subset f(A) \cap f(B)$ es siempre cierta.
3. 1) Sí; 2) No; 3) Sí; 4) Sí; 5) No; 6) Sí; 7) Sí; 8) No; 9) No; 10) No.
4. La imagen inversa de un subconjunto se puede definir aunque no exista la aplicación inversa.
5. La aplicación α que verifica $\alpha \circ h = 1_X$, no es única.
6. a) Sí; b) Sí; c) Sí; d) Sí.
7. La condición necesaria y suficiente es: a) $E = A \cup B$. b) $A \cap B = \emptyset$
8. No es una relación de equivalencia en E, pero sí lo es en $E - \{O\}$. Las clases de equivalencia en este segundo caso son las rectas que pasan por el origen (sin el origen).
9. El primer caso es una relación de equivalencia. Las clases son hipérbolas equiláteras con asíntotas en los ejes. En el segundo caso no se trata de una relación de equivalencia.
10. La aplicación f es constante sobre los elementos de una clase.
11. El elemento neutro es $(1, 0)$. El elemento inverso de (a, x) es $(a^{-1}, -xa^{-1})$. B es subgrupo conmutativo de A.
12. A_4 tiene 12 elementos. El único subgrupo de orden 12 es A_4. Solo hay un subgrupo de orden 4 (e es el elemento neutro): $\{e, (12)(34), (13)(24), (14)(23)\}$. De orden 3 hay los siguientes (que son isomorfos): $\{e, (123), (132)\}$, $\{e, (124), (142)\}$, $\{e, (134), (143)\}$, $\{e, (234), (243)\}$.

    De orden 2 (isomorfos): $\{e, (12)(34)\}$, $\{e, (13)(24)\}$, $\{e, (14)(23)\}$. De orden 1 solo hay un subgrupo: $\{e\}$
13. f es un homomorfismo: $a^{n+m} = a^n a^m$. El núcleo de f son los múltiplos de 5. La imagen es el grupo G_5. El grupo cociente, $\mathbb{Z}/\ker f$ está formado por los números congruentes módulo 5.
14. El elemento neutro es f_0 y el inverso de f_a es f_{-a}
15. El grupo del tetraedro T es isomorfo al grupo de alternaciones A_4.
16. El isomorfismo hace corresponder a la matriz dada por a, b el número complejo $z = a + ib$.
17. Solo hay dos grupos no isomorfos de orden 4 y son abelianos (el grupo de Klein (con $a^2 = b^2 = e$) y el cíclico de orden 4): $G_1 = \{e, a, b, ab\}$, $G_2 = \{e, r, r^2, r^3\}$.
18. No es un subgrupo normal. El conjunto cociente grupo/subgrupo (definidos en el problema) no es un grupo.
19. a) Son abelianos (orden 4). b) Sí. $f(a) = f(b) = f(c) = f(d) = x$. c) No.

20. $\mathbb{Z}_8$: $\{0\}, \{0,4\}, \{0,2,4,6\}, \mathbb{Z}_8$. $\mathbb{Z}_6$: $\{0\}, \{0,3\}, \{0,2,4\}, \mathbb{Z}_6$.
$f(0) = 0, f(1) = 3, f(2) = 0, f(3) = 3, f(4) = 0, f(5) = 3, f(6) = 0, f(7) = 3$. No.

21. $[1], [3]$.

22. No. $f([2n]) = [3n]$, $\forall n \in \mathbb{N}$.

23. Elemento neutro: $([1], [0])$. Si q tiene inverso: $([2], [3])^{-1} = (q^{-1}[2], q^{-1}[3])$.

24. Las clases se pueden caracterizar por el valor que toma el polinomio en $a \in \mathbb{R}$).

25. a) $R = \{[0], [2], [4], [6]\}$ es un ideal de $\mathbb{Z}_8$. b) No.

26. El cuerpo de los números racionales sólo tiene el automorfismo identidad. El cuerpo F_2 tiene dos automorfismos: la identidad y $\varphi(a + b\sqrt{2}) = a - b\sqrt{2}$

27. Se trata de una representación matricial de los cuaterniones.

28. Se supone que los ideales no contienen a la unidad (pues si no, son triviales).

29. Es un anillo con las operaciones dadas.

30. Los divisores de cero son: $\{[2], [4], [6]\}$ y los elementos invertibles: $\{[1], [3], [5], [7]\}$.

31. Si un elemento tiene inverso y está en un ideal, éste es igual al anillo.

32. En un cuerpo no hay ideales propios.

33. Suponer que la recta es el eje real.

34.
$$1, \quad \frac{1+i}{\sqrt{2}} \quad i, \quad \frac{-1+i}{\sqrt{2}}, \quad -1, \quad \frac{-1-i}{\sqrt{2}}, \quad -i, \quad \frac{1-i}{\sqrt{2}}$$

35. $\cos 5x = \cos^5 x - 10\cos^3 x \operatorname{sen}^2 x + 5\cos x \operatorname{sen}^4 x$, $\quad \operatorname{sen} 5x = 5\cos^4 x \operatorname{sen} x - 10\cos^2 x \operatorname{sen}^3 x + \operatorname{sen}^5 x$.

36. Si $p(z)$ tiene todos los coeficientes reales: $\overline{p(z)} = p(\overline{z})$.

37. $zp(z) = z^n + p(z) - 1 \Rightarrow (z-1)p(z) = z^n - 1$. Si $z_0^n = 1$ y $z_0 \neq 1$, entonces $p(z_0) = 0$

38. Raíces primitivas: $(1 \pm i\sqrt{3})/2$

39. $P(x) = (x-1)^5 = (x-1)^2(x - e^{2\pi i/5})^2(x - e^{4\pi i/5})^2(x - e^{6\pi i/5})^2(x - e^{8\pi i/5})^2$

40. La envolvente lineal de S es la intersección de todos los subespacios que contienen a S.

41. Usar para la suma la definición: $U + V = \{x + y \mid x \in U, y \in V\}$.

42. Si $x \in \operatorname{lin}\{v_1, \ldots, v_r\}$, $x = \sum_{i=1}^{r} \lambda_i v_i$. Como $v_r = \sum_{i=1}^{r-1} \mu_i v_i$, $x \in \operatorname{lin}\{v_1, \ldots, v_{r-1}\}$

43. a) Sí; b) No; c) No; d) No.

44. W_1 no; W_2 no; W_3 sí; W_4 no; W_5 sí.

45. a) Sí; b) Sí; c) No; d) Sí; e) No; f) Sí.

46. $V = \operatorname{lin}(V - W)$.

47. Usando la fórmula: $\dim W_1 + \dim W_2 = \dim(W_1 + W_2) + \dim(W_1 \cap W_2)$.

48. La igualdad $\mathcal{C}(\mathbb{R}, \mathbb{R}) = \mathcal{P} \oplus \mathcal{I}$ es consecuencia de la identidad:
$$f(t) = \frac{1}{2}(f(t) + f(-t)) + \frac{1}{2}(f(t) - f(-t))$$

49. La igualdad $V = U \oplus W$ es consecuencia de la identidad:

$$f(x) = \left(f(x) + \frac{b-a}{2}x - \frac{b+a}{2} \right) + \left(-\frac{b-a}{2}x + \frac{b+a}{2} \right)$$

donde: $a = f(1)$ y $b = f(-1)$.

50. Estudiar el rango de la matriz A, cuyas columnas son las coordenadas de los vectores u_i en la base B, y emplear el isomorfismo entre V y $\mathbb{K}^n$.

51. $W_1 \cap W_2 = \text{lin}\{(-2, -5, 2, 1)\}$.

52. Los dos primeros, no. El tercero, sí.

53. Base de W: $\{2, 0, 5, 3), (0, 1, 3, 1)\}$. Una base de un complementario es: $\{1, 0, 0, 0), (0, 1, 0, 0)\}$

54. S_1: l.i.; S_2: l.d.; S_3: l.i.; S_4: l.i.; S_5: l.d.;

55. $a = 8$, $b = 9$. Base de W: $\{p_1(x), p_2(x)\}$, y las coordenadas son: $p_1(x) \to (1, 0)$, $p_2(x) \to (0, 1)$, $p_3(x) \to (5, -2)$.

56. S_1 es l.i. y S_2 es l.d.

57. Los tres primeros vectores forman una base de $\mathbb{C}^3$.

58. Derivando $n-1$ veces en $\sum_{i=1}^n \mu_i e^{\lambda_i z} = 0$ se obtienen las n ecuaciones: $\sum_{i=1}^n \lambda_i^k \mu_i e^{\lambda_i z} = 0$, con $k = 0, \ldots, n-1$. Se tiene un sistema lineal de ecuaciones en μ_i, con determinante:

$$\exp(z \sum_{i=1}^n \lambda_i) \det \begin{pmatrix} 1 & \ldots & 1 \\ \lambda_1 & \ldots & \lambda_n \\ \vdots & & \vdots \\ \lambda_1^{n-1} & \ldots & \lambda_n^{n-1} \end{pmatrix}$$

que es el determinante de Vandermonde igual a (salvo un factor no nulo):

$$\prod_{i<j} (\lambda_i - \lambda_j) \neq 0$$

Para la segunda parte se usan las identidades:

$$\cos z = \frac{1}{2}(e^{iz} + e^{-iz}), \quad \text{sen}\, z = \frac{1}{2i}(e^{iz} - e^{-iz})$$

59.

$$\text{a)} \quad P = \frac{1}{3} \begin{pmatrix} 2 & 0 & 1 \\ -3 & 0 & 3 \\ 1 & 3 & 2 \end{pmatrix}, \quad \text{b)} \quad P = \frac{1}{15} \begin{pmatrix} 33 & -33 & 9 \\ 2 & 13 & 6 \\ 7 & 23 & 6 \end{pmatrix}$$

60.

$$p(x) = \sum_{i=0}^n a_i x^i = \sum_{i=0}^n \lambda_i (x+a)^i$$

luego λ_i son los coeficientes del desarrollo de Taylor en $x = -a$: $\lambda_k = p^{(i)}(-a)/k!$

61. El grado de $p(x)$ debe ser igual a n.

62. $W_1 \oplus W_2$, $W_2 \oplus W_3$, $W_3 \oplus W_1$ son sumas directas, pero $W_1 + W_2 + W_3$ no es suma directa.

63. Por inducción. Para $n = 2$ ya está demostrado. Si es cierto para $n-1$, para n basta escribir:

$$W_1 + \cdots + W_n = (W_1 + W_2) + W_3 \cdots + W_n$$

y aplicar la hipótesis de inducción y la fórmula para el caso $n = 2$. En el caso de suma directa, todas las intersecciones de la expresión anterior son iguales a cero.

64. Sí son complementarios: $U \cap W = \{0\}$ y la suma de dimensiones es n.

65. La suma no es directa para $a = -9/5$ y $b = -2/5$.

66. Base de U: $\{p_1(x), p_3(x)\}$. Base de W: $\{q_1(x), q_2(x)\}$
Base de $U \cap W$: $\{3p_3(x) - p_1(x)\}$. Base de $U + W$: $\{p_1(x), p_3(x), q_2(x)\}$

67. La única inclusión que es siempre cierta es: $(W_1 \cap W_3) + (W_2 \cap W_3) \subset (W_1 + W_2) \cap W_3$.

68. $m = -2, \quad n = 1$. $V_2 = \text{lin}(\{(1,0,0,0),(0,1,0,0)\}$.

69. La dimensión de cualquier intersección de la fórmula es 0. Entonces: $3 < 1 + 1 + 2$.

70. $\dim W_1 \cap W_2 = 1 \Rightarrow \dim(W_1 + W_2) = 4 \Rightarrow W_1 + W_2 = V$

71.

$$\dim U = 3,\ \mathcal{B}_U = \{(1,0,0,0,-1),(0,1,0,4,-1),(0,0,1,3,-1)\}$$

$$\dim V = 2,\ \mathcal{B}_V = \{(1,-1,2,0,0),(1,0,1,0,1)\}, \quad \dim U \cap V = 1,\ \mathcal{B}_{U\cap V} = \{(1,-3,4,0,-2)\}$$

$$\dim U + V = 4,\ \mathcal{B}_{U+V} = \{(1,0,0,0,-1),(0,1,0,4,-1),(0,0,1,3,-1),(1,-1,2,0,0)\}$$

72. $a = 0$

73. Sí. $\mathcal{B} = \{(x-1),(x-1)x,(x-1)x^2,\ldots,(x-1)x^{n-1}\}$

74. a) Falsa. b) Verdadera. c) Falsa.

75. $\dim(V_1 \times V_2) = \dim(V_1) + \dim(V_2)$. Para la segunda parte se usa la definición de suma directa.

76. $\dim L = 2$. Ecuaciones paramétricas: $x_1 = \lambda,\ x_2 = (2+i)\lambda + 3\mu,\ x_3 = (2-i)\lambda + \mu,\ x_4 = -i\lambda + 4\mu$. Ecuaciones implícitas: $(-7+4i)x_1 - x_2 + 3x_3 = 0,\ (-12+3i)x_1 + 4x_3 - x_4 = 0$. Base del espacio cociente: $\{(1,0,0,0)+L,(0,1,0,0)+L\}$

77. $x_2 = x_5 = 0$. $W_1 \cap W_2 = \text{lin}\{v_1 + v_3 + v_4\}$, $W_1 \cap W_3 = \{0\}$ $W_2 \cap W_3 = \{0\}$, $W_1 \cap W_2 \cap W_3 = \{0\}$. No es suma directa.

78. Base de V: $\{u_2, u_3, u_4, u_5\}$. base de W: $\{u_4, u_5\}$

79. Sustituir t por A en la expresión del polinomio.

80. $\text{tr}\, A = \text{tr}[B,C] = \text{tr}(BC - CB) = \text{tr}\, BC - \text{tr}\, CB = \text{tr}\, BC - \text{tr}\, BC = 0$. No: $\text{tr}\, iI = ni$.

81. Usando inducción se prueba que: $A^n = nA - (n-1)I$. Por tanto,

$$A^{100} = \begin{pmatrix} 1 & 0 \\ -100 & 1 \end{pmatrix}$$

82.

$$A^n = I + nN + \frac{1}{2}n(n-1)N^2 = \begin{pmatrix} 1 & n & \frac{1}{2}n(n-1) \\ 0 & 1 & n \\ 0 & 0 & 1 \end{pmatrix}, \quad N = A - I$$

83. Basta hacer la conmutación con los elementos de la base canónica E_{ij}.

84. $A^2 = ABAB = ABB = AB = A, \quad B^2 = BABA = BA^2 = B$. Sólo son invertibles si son iguales a la identidad.

85. Se tiene: $M(\lambda)M(\lambda') = M(\lambda + \lambda')$. El inverso de $M(\lambda)$ es $M(-\lambda)$.

86. Escribiendo la ecuación $AB = I$ como un sistema lineal, o considerando que cualquier matriz cuadrada con determinante distinto de cero tiene inversa.

87. $(-1)^{n(n-1)/2}$

88. Desarrollar por la primera columna todas las matrices que van apareciendo. Para matrices triangulares inferiores utilizar la transpuesta.

89. En el primer caso, por inducción en el orden de A. En el segundo caso, usar:

$$\begin{pmatrix} C & A \\ B & 0 \end{pmatrix} = \begin{pmatrix} A & C \\ 0 & B \end{pmatrix} \begin{pmatrix} 0 & I_n \\ I_m & 0 \end{pmatrix}$$

y probar que:

$$\det \begin{pmatrix} 0 & I_n \\ I_m & 0 \end{pmatrix} = (-1)^{nm}$$

90.
$$\det A \det B = \det \begin{pmatrix} A & -I \\ 0 & B \end{pmatrix} = \det \begin{pmatrix} A & -I \\ AB & 0 \end{pmatrix} = (-1)^{n^2+n} \det(AB) = \det(AB)$$

91. $D_1 = (x-a)^n(x+(n-1)a), \quad D_2 = a^{n-1}(a + \sum_{i=1}^n x_i)$.

92. Haciendo ceros en la primera columna (desde la fila 2 hasta la última, restando de cada fila la anterior multiplicada por x_1) se demuestra que: $W[x_1, \ldots, x_n] = \prod_{1<i}(x_i - x_1)W[x_2, \ldots, x_n]$. Con esta recurrencia se prueba que: $W[x_1, \ldots, x_n] = \prod_{j<i}(x_i - x_j)$

93. $\det(A) = -8, \quad \det(B) = (a-b)(a-c)(a-d)(b-c)(b-d)(c-d)(a+b+c+d), \quad \det(C) = (a^2-b^2)^4$.

94. $\Delta_n = 0,\ n \geq 3, \quad \Delta_2 = (x_1 - x_2)(y_2 - y_1)$.

95. $\det(A) = \det(A^t) = \det(-A) = (-1)^n \det(A)$, luego $\det(A) = 0$ si n es impar (y $2 \neq 0$).

96. $(\det A)(\det A^+) = (\det A)(\overline{\det A}) = |\det A|^2 = \det I = 1 \Rightarrow |\det A| = 1$.

97. Usar para a) y b) las siguientes igualdades:

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix} = \begin{pmatrix} A & 0 \\ C & D - CA^{-1}B \end{pmatrix} \begin{pmatrix} I & A^{-1}B \\ 0 & I \end{pmatrix},$$

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix} = \begin{pmatrix} I & BD^{-1} \\ 0 & I \end{pmatrix} \begin{pmatrix} A - BD^{-1}C & 0 \\ C & D \end{pmatrix}$$

98. 1) $a = -20 \Rightarrow r(A) = 3,\ a \neq -20 \Rightarrow r(A) = 4$. 2) $a = 3 \Rightarrow r(A) = 2,\ a \neq 3 \Rightarrow r(A) = 4$

99. $a = 1,\ b = 7 \Rightarrow r(A) = 3$.

100. $r(A) = 2$.

101. $r(B) = \dim(\mathrm{lin}\{B_1, \ldots, B_m\}),\ r(AB) = \dim(\mathrm{lin}\{AB_1, \ldots, AB_m\}) \Rightarrow r(AB) \leq r(B)$. De forma similar (usando transpuestas, por ejemplo) se prueba: $r(AB) \leq r(A)$.

102. Si $r(A)=0$ (suponiendo que la primera fila es no nula; en el caso en que todas lo sean la igualdad es evidente):

$$\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ a_{21} & \cdots & a_{2n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \lambda_2 a_{11} & \cdots & \lambda_2 a_{1n} \\ \vdots & & \vdots \\ \lambda_n a_{11} & \cdots & \lambda_n a_{1n} \end{pmatrix} = \begin{pmatrix} 1 \\ \lambda_2 \\ \vdots \\ \lambda_n \end{pmatrix} \begin{pmatrix} a_{11} & \cdots & a_{1n} \end{pmatrix}$$

La implicación en sentido contrario es consecuencia de un problema anterior:

$$r(RS) \leq \min(r(R), r(S)) \leq 1$$

103. a) $r(AB) \leq \min(r(A), r(B)) \leq n < m$. Pero, $r(I_m) = m$.

b) $AB = I_m \Rightarrow AB_i = e_i,\ i = 1, \dots m$. Si $r(A) = m$ entonces $r(A|e_i) = m$ para cualquier i, luego el sistema es compatible y como $n < m$, es indeterminado y hay infinitas soluciones. Si $r(A) < m$, $r(AB) < m$, luego no puede ser igual a I_m.

104. $\det(A) = -9$.

105. $\operatorname{cof}(\lambda A) = \lambda^{n-1} \operatorname{cof} A,\ \det(\operatorname{cof} A) = (\det A)^{n-1},\ \operatorname{cof}(\operatorname{cof} A) = (\det A)^{n-2} A,\ \operatorname{cof}(AB) = \operatorname{cof} A \operatorname{cof} B$.

106. $a_{ij} = j + n(i-1)$. El rango es 2 (restar a cada fila la anterior).

107. No existe inversa de A.

$$B^{-1} = \frac{1}{8} \begin{pmatrix} 8 & 0 & 8 \\ -6 & 2 & -2 \\ -3 & 1 & -5 \end{pmatrix}, \quad C^{-1} = \begin{pmatrix} 5 & -4 & -1 & 4 \\ 10 & -8 & -3 & 9 \\ 11 & -9 & -4 & 10 \\ 9 & -7 & -3 & 8 \end{pmatrix}$$

108.

$$A^{-1} = \begin{pmatrix} -1 & 2 & 1 \\ 5 & -8 & -6 \\ -3 & 5 & 4 \end{pmatrix}, \quad B^{-1} = \frac{1}{3} \begin{pmatrix} -5 & 2 & 4 \\ -3 & 0 & 3 \\ 7 & -1 & -5 \end{pmatrix}, \quad C^{-1} = \begin{pmatrix} 14 & 8 & 3 \\ 8 & 5 & 2 \\ 3 & 2 & 1 \end{pmatrix}$$

109.

$$A^{-1} = \begin{pmatrix} -i & 1+2i & i \\ 1 & -i & 1-i \\ 1 & 0 & 1 \end{pmatrix}$$

110. Sea A con elementos en $\mathbb{Z}$. Si $\det A = \pm 1$, la inversa tiene elementos enteros. Si A y A^{-1} tienen elementos enteros, los dos determinantes son enteros. Como uno es el inverso del otro y en $\mathbb{Z}$ las únicas unidades (elementos con inverso) son $1, -1$, el determinante es ± 1.

111. a) Sí. b) Sí. c) Sí. d) Sí. e) No. f) No. g) No. h) No.

112. 1) No. 2) No. 3) Sí. 4) No.

113. Se define, si $x = \sum_{i=1}^n x_i v_i \in V_1$, el operador A como: $Ax = \sum_{i=1}^n x_i w_i$, que verifica: $Av_i = w_i$ y es lineal. Además, es único (al ser $\mathcal{B}$ una base de V_1).

114. a) $\lambda_1 + \lambda_2 e^{it} + \lambda_3 e^{-it} = 0$ implica $\lambda_1 = \lambda_2 = \lambda_3 = 0$, por ejemplo derivando y poniendo $t = 0$.

b) Que $\mathcal{B}'$ es otra base se prueba como antes. La matriz de cambio es:

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

115. a) No. b) No. c) No d) Sí. (no degenerado=inyectivo).

116. T es claramente lineal. Además es biyectiva: $\ker T = \{0\}$ e $\operatorname{im} T = \mathbb{C}_n[t]$.

117. a) T es lineal.

b) $T(x^{2m}) = 2\sum_{k=0}^{m-1}\binom{2m}{2k}x^{2k}, \quad T(x^{2m+1}) = 2\sum_{k=0}^{m-1}\binom{2m+1}{2k+1}x^{2k+1}$

c) $\ker T = \mathbb{C}_1[t]$, $\operatorname{im} T = \mathbb{C}_{n-2}[t]$

d) Si $T\tilde{p}(t) = q(t)$, entonces $p(t) = \tilde{p}(t) - \tilde{p}'(0)t - \tilde{p}(0)$ verifica: $Tp(t) = q(t)$ y $p'(0) = p(0) = 0$. Además, p es único con estas propiedades.

118. F_A es lineal. $F_A(XY) = [A, XY] = X[A, Y] + [A, X]Y$.

119. $A' = QAP^{-1}$ con:

$$Q^{-1} = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}, \quad P^{-1} = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 0 \\ -1 & 1 & 0 \end{pmatrix}, \quad A' = \begin{pmatrix} -2 & 1 & 0 \\ 1 & 2 & 1 \end{pmatrix}$$

120. a) Sí. b) No. c) No. d) Sí. e) Sí.

121. $\ker T = \operatorname{lin}\{(1, -1, 1)\}, \quad \operatorname{im} T = \operatorname{lin}\{(1, 0, -1), (2, 1, 3)\}$.

122.

$$\text{a)}\ \mathcal{T} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, \quad \text{b)}\ P = \frac{1}{3}\begin{pmatrix} 1 & 1 \\ 2 & -1 \end{pmatrix}, \quad \mathcal{T}' = \frac{1}{3}\begin{pmatrix} -1 & 2 \\ -5 & 1 \end{pmatrix}$$

c) $\det T = 1 + c^2 \neq 0, \ \forall c \in \mathbb{R}$.

123. a)

$$\mathcal{T} = \begin{pmatrix} 3 & 0 & 1 \\ -2 & 1 & 0 \\ -1 & 2 & 4 \end{pmatrix}, \quad P = -\frac{1}{4}\begin{pmatrix} 1 & 3 & -5 \\ 1 & -1 & -1 \\ -2 & -2 & 2 \end{pmatrix}, \quad \mathcal{T}' = \frac{1}{4}\begin{pmatrix} 17 & 35 & 22 \\ -3 & 15 & -6 \\ -2 & -14 & 0 \end{pmatrix}$$

b)$\det T = 9, \quad T^{-1}(x, y, z) = (1/9)(4x + 2y - z, 8x + 13y - 2z, -3x - 6y + 3z)$.

124. $A^2 = 0$, pues si $x \in V$, $Ax \in \operatorname{im} V = \ker A \Rightarrow A(Ax) = 0$.

125. $\det L_A = \det R_A = 4, \quad \operatorname{tr} L_A = \operatorname{tr} R_A = 2$.

126.

$$A' = PAP^{-1} = \begin{pmatrix} 0 & 1 & 0 \\ 1 & -1 & 1 \\ 2 & 1 & 1 \end{pmatrix}\begin{pmatrix} i & 0 & 0 \\ 0 & i & 0 \\ 0 & 0 & 0 \end{pmatrix}\begin{pmatrix} 0 & 1 & 0 \\ 1 & -1 & 1 \\ 2 & 1 & 1 \end{pmatrix}^{-1} = \begin{pmatrix} i & 0 & 0 \\ -3i & -i & i \\ -3i & -2i & 2i \end{pmatrix}$$

127. $T_A(A) = 0$ luego T_A no es inyectiva y su determinante es cero.

128. A es la composición de la aplicación de un problema anterior y de la derivada: $A = D \circ T$. $\det A = 0, \ \operatorname{tr} A = 0$.

129. $(A + B)^\dagger = A^\dagger + B^\dagger, \quad (\lambda A)^\dagger = \lambda A^\dagger, \ \lambda \in \mathbb{R}$. No es un subespacio de $\mathcal{M}_n(\mathbb{C})$. T_B es lineal y: $(BAB^\dagger)^\dagger = BA^\dagger B^\dagger = BAB^\dagger$.

130. $\ker T$ son las matrices antisimétricas. $\operatorname{im} T$ son las matrices simétricas. Para $n = 2$, $\operatorname{tr} T = 6, \ \det T = 0$.

131. a) W es un subespacio lineal (de dimensión infinita).

b)

$$\int_{-\pi}^{\pi} \cos nsds = \int_{-\pi}^{\pi} \cos ns \cos s\, ds = \int_{-\pi}^{\pi} \cos ns \operatorname{sen} s\, ds \quad =$$
$$\int_{-\pi}^{\pi} \operatorname{sen} nsds = \int_{-\pi}^{\pi} \operatorname{sen} ns \cos s\, ds = \int_{-\pi}^{\pi} \operatorname{sen} ns \operatorname{sen} s\, ds = 0.$$

cuando $n \geq 2$. c) No, por el apartado b). d) La integral es lineal. e) Desarrollando el integrando: $\operatorname{im} T = \operatorname{lin}\{1, \cos t, \operatorname{sen} t\}$, $\dim(\operatorname{im} T) = 3$. f) $\ker T = W$. g) $\lambda = 2\pi,\ f(t) = 1, \quad \lambda = \pi,\ f(t) = a_1 \cos t + a_2 \operatorname{sen} t$

132. En este caso V es suma directa del núcleo y la imagen de f.

133. No. $r(A) = 2 \neq r(A') = 3$. $\mathcal{B}_{\ker f} = \{-3e_1 + e_2 + e_3\}$, $\mathcal{B}_{\operatorname{im} f} = \{u_1 + u_3, 2u_1 - u_2 + u_3\}$.

134. $V = \ker f \oplus W \Rightarrow \ker f \cap W = \{0\}$.

135. En las bases:

$$\mathcal{B}_{\mathbb{C}^5} = \left\{ u_1 = \begin{pmatrix} -2 \\ 1 \\ 0 \\ 0 \\ 1 \end{pmatrix},\ u_2 = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 1 \end{pmatrix},\ u_3 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 1 \end{pmatrix},\ u_4 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}.,\ u_5 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} \right\}$$

$$\mathcal{B}_{\mathbb{C}^3} = \left\{ v_1 = \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix},\ v_2 = \begin{pmatrix} 1 \\ -1 \\ -3 \end{pmatrix},\ v_3 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \right\}$$

$$\mathcal{M}(f, \mathcal{B}_{\mathbb{C}^5}, \mathcal{B}_{\mathbb{C}^3}) = \begin{pmatrix} 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

136. a) Falso. b) Falso. c) Falso. d) Cierto. e) Cierto.

137. $\mathcal{B}_{W_1+W_2} = \{(0,2,1,0,0), (0,1,0,-1,0), (1,0,1,0,0), (0,0,0,0,1)\}$. $\mathcal{B}_{W_1 \cap W_2} = \{(0,1,0,-1,0)\}$.

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

138. Base: $\{x-1, (x-1)^2, (x-1)^3, (x-1)^4\}$. Matriz:

$$\begin{pmatrix} 1 & -2 & 3 & -4 \\ 0 & 2 & -6 & 12 \\ 0 & 0 & 3 & -12 \\ 0 & 0 & 0 & 4 \end{pmatrix}$$

$\ker D = \{0\}$, $\operatorname{im} D = W$.

139. a) Cierta en los dos sentidos. b) Falsa hacia la derecha y cierta hacia la izquierda. c) Cierta hacia la derecha y falsa hacia la izquierda.

140. Si $AX = 0$ sólo tiene la solución trivial, A (como transformación lineal) es un isomorfismo, y $AX = B$ tiene solución única. Similar en el sentido opuesto. Si $AX = 0$ tiene soluciones distintas de la trivial, A no es sobreyectiva.

141. $x_1 = 5\lambda/4, \quad x_2 = 67\lambda/24, \quad x_3 = \lambda.$

142. $x_1 = 1 + 2\lambda - \mu, \quad x_2 = 2 - \lambda + \mu \quad , x_3 = \lambda, \quad x_4 = \mu, \quad x_5 = 1.$

143. Cuando $2a - 3b + c \neq 0$ el sistema no tiene solución.

144. Sí. $x_1 = -\lambda + 1/2, \quad x_2 = \lambda - 1/2, \quad x_3 = \lambda.$

145. Hay solución no trivial si $|\alpha| = 1$. $x_1 = -\alpha x_2 - \alpha^2 x_3$.

146. $x_1 = (a + b + c)/3, \quad x_2 = (a + \omega^2 b + \omega c)/3, \quad x_3 = (a + \omega b + \omega^2 c)/3.$

147. a) $(1, -1, 1, -1, 1)$. b) incompatible. c) $(-\lambda + 7\mu/6, \lambda + 5\mu/6, \lambda, \mu/3, \mu)$. d) $(0, 2, 5/3, -4/3)$.

148. a) Solución única: $a \neq 1, -2,\ b \neq 0$:

$$x = \frac{a - b}{(a - 1)(a + 2)}, \quad y = \frac{b(a + 1) - 2}{b(a - 1)(a + 2)}, \quad z = \frac{a - b}{(a - 1)(a + 2)}$$

$b = 0$, $(b \neq 1, a = 1)$, $(b \neq -2, a = -2)$ no hay solución.

$a = b = 1$: $x = 1 - z - y$. Si $a = b = -2$: $x = z = -1 - 2y$.

b) $a \neq 0,\ b \neq -1, 1$. Solución única: $(1/a, 0, 0)$.

$a = 0,\ b \neq -1, 1$, no hay solución.

$a = 0,\ b = 1$, solución: $(x, 1, 0)$

$a = 0,\ b = -1$, solución: $(x, 1/3, 2/3)$

$a \neq 0,\ b = 1$, solución: $((1 - y)/a, y, 0)$

$a \neq 0,\ b = -1$, solución: $((2 - 3z)/2a, z/2, z)$

c) Si $b \neq a+1, a-1, -a+1, -a-1$, solución única: $(1+a^3+b+ab-a^2b-b^2+ab^2-b^3, -1-2a+a^3-b-ab-3a^2b+b^2+ab^2+b^3, a(a+a^2+b-2ab+b^2), a(-2-a+a^2-b-2ab+b^2))/(a+b+1)(a-b+1)(a-b-1)$.

Si $b = a + 1,\ a \neq 0$, no hay solución. Si $b = 1,\ a = 0$, la solución es $(-1 - z, 1 - t, z, t)$.

Si $b = a - 1,\ a \neq 0$, no hay solución. Si $b = -1,\ a = 0$, la solución es $(-1 + z, 1 + t, z, t)$.

Si $b = -a + 1,\ a \neq 0, 1$, la solución es: $(-t + a - 1/2, -t + (2a^2 - a - 2)/2(a - 1), t + 1/2(a - 1), t)$. Si $b = 0,\ a = 1$ no hay solución.

Si $b = -a - 1,\ a \neq 0$ no hay solución.

d) Si $a \neq 1, -3$ solución única: $(-b^3 - b^2 - b + a + 2, -b^3 - b^2 + ab + 2b - 1, -b^3 + ab^2 + 2b^2 - b - 1, ab^3 + 2b^3 - b^2 - b - 1)/(a + 3)(a - 1)$.

Si $a = 1,\ b \neq 1$, no hay solución. Si $a = 1,\ b = 1$ la solución es $(1 - y - z - t, y, z, t)$. Si $a = -3,\ b \neq 1, i, -i$ no hay solución. Si $a = -3,\ b = -1$, la solución es $(t - 1/2, t, t - 1/2, t)$. Si $a = -3,\ b = i$, la solución es $(t - (1 + i)/4, t - i/2, t + (1 - i)/4, t)$. Si $a = -3,\ b = -i$, la solución es $(t + (-1 + i)/4, t + i/2, t + (1 + i)/4, t)$.

149. Si $\alpha = \beta$, el sistema no tiene solución a no ser que $a_1 = \cdots = a_n$. (el caso $\alpha = \beta = 0$ no lleva a un sistema). La solución, cuando existe, es: $(a_1 - \alpha \sum_{i=2}^n x_i, x_2, \ldots, x_n)$.

Si $\beta = (1 - n)\alpha \neq 0$, el sistema no tiene solución a no ser que $\sum_{i=1}^n a_i = 0$.

Si $\beta \neq \alpha, (1 - n)\alpha$, la solución es:

$$x_j = \frac{1}{n(\alpha - \beta)} \left(\frac{(2n - 1)\alpha}{(n - 1)\alpha + \beta} \sum_{i=1}^n a_i - n a_j \right),\ j = 1, \ldots, n$$

150. $B = (-\lambda + 3\mu, 3\lambda - 2\mu, \lambda, \mu),\ \lambda, \mu \in \mathbb{C}$.

151. Si $a \neq 0, -3$, solución única: $(-a^3 + a + 6, 5a^2 + 4a - 3, a^4 + 2a^3 - 2a - 3)/(a + 3)$. Si $a = 0$, solución: $(-y - z, y, z)$. Si $a = -3$ no hay solución.

152. a) $-2, -3, 4, 5$. b) $2, -3/2, 5/6, \pm i\sqrt{2}$. c) $2, 3, 5, \pm\sqrt{2}$. d) $3, 2 \pm \sqrt{3}$. e) $1, -3, 7, (-1 \pm i\sqrt{3})/2$. f) $1, 2, 3, 6$. g) $4, -1/2, -2/3, (-1 \pm i\sqrt{3})/2$. h) $3, 1/2, -1/2, -2/3, 5/6$.

153. Derivando: $p(1) = 0,\ p'(1) = 0,\ p''(1) = 0,\ p'''(1) = 2n^3 - 2n \neq 0$. La multiplicidad es 3.

154. Usar $f(A)v = f(\lambda)v$, si $Av = \lambda v$. Si μ es raíz de f, puede no ser autovalor de A. Por ejemplo, considerar la matriz identidad en 2×2 y el polinomio $f(t) = t^2 - 1$.

155. Autovalor $\lambda = 0$. Autovector: $p(t) = 1$.

156.
$$\text{a) } \det \begin{pmatrix} 1 & 1 & 1 \\ 1 & -1 & 0 \\ 1 & 0 & -1 \end{pmatrix} = 3 \neq 0, \quad \text{b) } \mathcal{M}(T, \hat{\mathcal{B}}) = \begin{pmatrix} 3a & 0 & 0 \\ 0 & 3b & -3c \\ 0 & 3c & 3b \end{pmatrix}$$

c) Si $c \neq 0$, $p_T(\lambda) = -(\lambda - 3a)(\lambda^2 - 6b\lambda + 9(b^2 + c^2)) = -m_T(\lambda)$. Si $c = 0$, $m_T = (\lambda - 3a)(\lambda - 3b)$.

d) T sólo es diagonalizable en $\mathbb{R}^3$ si $c = 0$. En $\mathbb{C}^3$ es siempre diagonalizable.

157. Las ecuaciones:
$$\int_0^t f(s)ds = \lambda f(t); \quad \lambda f'(t) = f(t),\ f(0) = 0$$
son equivalentes y sólo tienen la solución trivial $f(t) = 0$.

158. $\lambda = 0,\ (1, -1, 0, \ldots, 0), \ldots (0, \ldots, 0, 1, -1), \quad \lambda = n,\ (1, \ldots, 1)$.

159. a) Al ser diagonalizable, $\ker A$ corresponde al subespacio invariante de autovalor 0.

b) Por la misma razón que en el apartado b).

160. $\det(A - \lambda I) = \det(A - \lambda I)^t = \det(A^t - \lambda I)$. Si A tiene inversa, y $Av = \lambda v$, entonces $A^{-1}v = \lambda^{-1}v$.

161. En la base de vectores propios, (P es la matriz del cambio de base):
$$\mathcal{M}(A, \mathcal{B}') = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix}, \quad P^{-1} = \begin{pmatrix} 0 & 1 & 1 \\ 1 & -1 & 0 \\ 1 & 0 & -1 \end{pmatrix}$$
En la base canónica:
$$\mathcal{M}(A, \mathcal{B}) = \begin{pmatrix} 1 & -1 & 1 \\ 2 & 4 & 2 \\ 3 & 3 & 3 \end{pmatrix}$$

162. $\lambda = 0$ no es una raíz de $\lambda^4 + \lambda - 1$, f es un automorfismo.

163. f no es sobreyectiva. Al ser un endomorfismo, tampoco es inyectiva.

164. $f(v) = \frac{3-\sqrt{5}}{2}((1 + \sqrt{5})u_1 + 2u_2)$

165.
$$P = \begin{pmatrix} 1 & 8 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 2 & 7 & 0 & 1 \\ 0 & 3 & 0 & 0 \end{pmatrix}$$

166.
$$\frac{1}{6} \begin{pmatrix} 2 & 2 & -2 \\ 2 & -1 & -5 \\ -2 & -5 & -1 \end{pmatrix}$$

167. a) $1, 2, 3,\ \{(1, 2, 1), (1, 1, 0), (1, 2, 2)\}$.
 b) $-1, 1(2),\ \{(3, 5, 6), (-1, 0, 1), (2, 1, 0)\}$.
 c) $1, 2 + 3i, 2 - 3i,\ \{(1, 2, 1), (3 - 3i, 5 - 3i, 4), (3 + 3i, 5 + 3i, 4)\}$.
 d) $0,\ \{(1, 3, 0, 0), (5, 0, 6, 3)\}$.
 e) $1, 2, 3,\ \{(1, 0, 0), (2, 1, 0), (9, 6, 2)\}$.
 f) $-3,\ \{(1, 0, 6), (1, 2, 0)\}$.
 g) $-0.6, 0.7, 0.5, 5.4,\ (0.2, 0.4, -0.4, 1), (-1.2, 0.4, 2.8, 1), (2, -2.4, 0.8, 1), (-3, -2.4, -1.2, 1)$.
 h) $1,\ \{(3, 1, 1, 0), (2, -1, 0, 1)\}$.

168. 1.i) $(a - d)^2 > 4bc$, diagonalizable en $\mathbb{R}$ y $\mathbb{C}$.
 1.ii) $(a-d)^2 = 4bc$, $b = c = 0$, diagonalizable en $\mathbb{R}$ y $\mathbb{C}$ (diagonal). $b \neq 0$ o $c \neq 0$, no diagonalizable.
 2) $(a - d)^2 < 4bc$, diagonalizable en $\mathbb{C}$.

169. En un base adaptada a la descomposición, la matriz de A es diagonal por bloques.

170. $\sigma(f) = \{0, 2, a, b, c\}$. Siempre es diagonalizable. No es invertible.

171. a) Cierta. b) Falsa. c) Cierta.

172. a) Falsa. b) Cierta. b) Falsa. c) Cierta.

173. Cierta. El rango de $A - \lambda I$ es siempre 3.

174. a)

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

 b)

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

 c) $B = 64I$

175. a)

$$J = \begin{pmatrix} 4 & 0 & 0 & 0 \\ 0 & 4 & 1 & 0 \\ 0 & 0 & 4 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad P = \begin{pmatrix} 1 & 2 & 0 & 0 \\ 2 & 0 & 0 & 2 \\ 1 & -2 & 0 & 0 \\ 0 & -4 & 1 & 1 \end{pmatrix}$$

 b) $\mathrm{lin}\{P_1, P_2\}$. c) $\dim\ker A = 1,\ \dim\mathrm{im}\, A = 3$

176. En todos los casos, $A = PJP^{-1}$ donde J es la forma canónica de Jordan.
 a)

$$J = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 2 \end{pmatrix}, \quad P = \begin{pmatrix} 1 & 0 & 5 & 3 \\ 0 & 1 & 2 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}$$

 b)

$$J = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & -\sqrt{2} & 0 \\ 0 & 0 & 0 & \sqrt{2} \end{pmatrix}, \quad P = \begin{pmatrix} 0 & 1/3 & 0 & 0 \\ 0 & 0 & -1/\sqrt{2} & 1/\sqrt{2} \\ 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 0 \end{pmatrix}$$

c)

$$J = \begin{pmatrix} -1 & 1 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad P = \begin{pmatrix} 1 & 0 & -2 \\ 0 & -1 & 2 \\ 0 & 0 & 1 \end{pmatrix}$$

d)

$$J = \begin{pmatrix} -1 & 1 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 1 \\ 0 & 0 & 0 & -1 \end{pmatrix}, \quad P = \begin{pmatrix} -2 & -1/3 & -1 & 1/3 \\ 0 & 0 & 1 & 0 \\ 1 & 5/3 & 0 & -2/3 \\ 3 & 0 & 0 & 0 \end{pmatrix}$$

e)

$$J = \begin{pmatrix} 5 & 1 & 0 & 0 & 0 & 0 \\ 0 & 5 & 1 & 0 & 0 & 0 \\ 0 & 0 & 5 & 0 & 0 & 0 \\ 0 & 0 & 0 & 5 & 1 & 0 \\ 0 & 0 & 0 & 0 & 5 & 1 \\ 0 & 0 & 0 & 0 & 0 & 5 \end{pmatrix}, \quad P = \begin{pmatrix} 0 & 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 & 0 \end{pmatrix}$$

f)

$$J = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}, \quad P = \begin{pmatrix} 0 & -1 & 1 & 0 & 0 \\ 0 & -4 & 0 & 1 & 0 \\ -1 & -2 & 0 & 1 & 0 \\ -1 & -3 & 0 & 1 & 0 \\ -1 & -8 & 0 & 2 & 1 \end{pmatrix}$$

g)

$$J = \begin{pmatrix} -2 & 1 & 0 & 0 & 0 \\ 0 & -2 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}, \quad P = \begin{pmatrix} -1 & 1 & -1 & 0 & 1 \\ -3 & 2 & -1 & 1 & 2 \\ 0 & 0 & 1 & 1 & 1 \\ -1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 \end{pmatrix}$$

h)

$$J = \begin{pmatrix} -1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}, \quad P = \begin{pmatrix} -1 & -2 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 & 1 \\ 1 & 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & -2 & 0 & 0 \end{pmatrix}$$

i)

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

177. a) rango $(A) = 2$. b) $p(\lambda) = \lambda^2(\lambda^2 - 4), \quad m(\lambda) = \lambda(\lambda - 2)(\lambda + 2)$

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

178. Si $b = 0$ no es diagonalizable. Si $b \neq 1$ es diagonalizable para todo a:

$$J = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & b \end{pmatrix}, \quad P = \begin{pmatrix} 1 & a & 1 \\ 0 & 1 & 0 \\ 0 & 0 & b \end{pmatrix}, \quad A = PJP^{-1}$$

179. a) La imagen es un polinomio de grado 4.

b)

$$\begin{pmatrix} -2\mu & -\lambda^2 & 0 & 0 & 0 \\ -4 & -2\mu & -2\lambda^2 & 0 & 0 \\ 0 & -3 & -2\mu & -3\lambda^2 & 0 \\ 0 & 0 & -2 & -2\mu & -4\lambda^2 \\ 0 & 0 & 0 & -1 & -2\mu \end{pmatrix}$$

c) Autovalor: -2μ. Autovector: x^4.

180.

$$e^A = \begin{pmatrix} e^3 & 0 & 0 \\ 4e^3 & e^3 & 0 \\ -\frac{21}{25}e^{-2} + \frac{41}{25}e^3 & -\frac{1}{5}e^{-2} + \frac{1}{5}e^3 & e^{-2} \end{pmatrix}, \quad e^B = \begin{pmatrix} e & 0 & 0 \\ e^2 - e & e^2 & 0 \\ -\frac{3}{4}e^{-3} + \frac{3}{4}e & 0 & e^{-3} \end{pmatrix}$$

181. a) $\omega((x_1, x_2, x_3)) = -3x_2 + x_3$. b) $\mu = (c, 2c, -c),\ c \neq 0$. c) $\mu(2, 3, -1) = 9c$.

182. $e^1 = (1, -1, 0), \quad e^2 = (1, -1, 1), \quad e^3 = (-1/2, 1, -1/2)$

183.

$$a = \begin{pmatrix} 1 & 2 & -1 \\ 1/2 & 2 & 1/2 \\ 1/3 & 8/3 & -1/3 \end{pmatrix}, \det a = -2 \neq 0$$

$$p_1(x) = -\frac{3}{2}x^2 + x + 1, \quad p_2(x) = \frac{1}{2}x^2 - \frac{1}{6}, \quad p_3(x) = -\frac{1}{2}x^2 + x - \frac{1}{3}$$

$$p(x) = \left(-\frac{3a}{2} + \frac{b}{2} - \frac{c}{2}\right)x^2 + (a + c)x + \left(a - \frac{b}{6} - \frac{c}{3}\right)$$

184. $w^1 = (6, -3, 0, -1, 6), \quad w^2 = (-3, 2, -1, 0, 1)$

185. $(\mathbb{C}\mu)^\circ \cap (\mathbb{C}\nu)^\circ = (\mathbb{C}\mu + \mathbb{C}\nu)^\circ, \quad \dim(\mathbb{C}\mu + \mathbb{C}\nu) = 2 \Rightarrow \dim\left((\mathbb{C}\mu)^\circ \cap (\mathbb{C}\nu)^\circ\right) = n - 2$

186. 1) $\sigma = (a_1, 0)$. 2) $\sigma = (a_2, -a_1)$. 3) $\sigma = (a_1 + a_2, -a_1 + a_2)$.

187. Matriz de f en la base $\{u_1, \ldots u_2\}$:

$$\begin{pmatrix} f_1(u_1)f_2(u_1) & \cdots & f_1(u_1)f_2(u_n) \\ \vdots & & \vdots \\ f_1(u_n)f_2(u_1) & \cdots & f_1(u_n)f_2(u_n) \end{pmatrix}$$

188. Solo f_1 lo es.

189.

$$\begin{pmatrix} 1 & 1/2 & 1/3 \\ 1/2 & 1/3 & 1/4 \\ 1/3 & 1/4 & 1/5 \end{pmatrix}, \quad g(t^2 - 2, 2t + 4) = -\frac{49}{6}$$

190.

$$\begin{pmatrix} 2 & 2 & 4 \\ -2 & -1 & -3 \\ 0 & 1 & 1 \end{pmatrix}, \quad g(x, y) = 13$$

191. 1 y 3 son formas bilineales.

192. a) No es regular. $\text{rad} = \text{lin}\{(1, 0, -1, 0), (0, 1, 0, -1)\}$. $\text{sig} = (1, 1)$.

b) $\mathcal{B} = \{(1, 1, 0, 0)/\sqrt{2},\ (1, -1, 0, 0)/\sqrt{2},\ (1, 0, -1, 0),\ (0, 1, 0, -1)\}$.

193.

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

194. $\mathrm{ran}(f_1) = 2, \mathrm{sig}(f_1) = (1,1)$, $\quad \mathrm{ran}(f_2) = 3, \mathrm{sig}(f_2) = (1,2)$, $\quad \mathrm{ran}(f_3) = 3, \mathrm{sig}(f_3) = (2,1)$. f_2 y f_3 son equivalentes en $\mathbb{C}$ pero no en $\mathbb{R}$.

195. $q(v) = (x-2y)^2 + 2(y+z/2)^2 - 3z^2/2$. $\mathrm{sig}(q) = (2,1)$. No.

196.

$$\begin{aligned} q_1(u) &= \left(\frac{1}{\sqrt{6}}x\right)^2 + \left(-\frac{1}{\sqrt{30}}x + \frac{6}{\sqrt{30}}y\right)^2 + \left(\frac{2}{\sqrt{5}}x + \frac{3}{\sqrt{5}}y + \frac{5}{\sqrt{5}}z\right)^2 \\ q_2(u) &= x^2 - (y-2z)^2 \\ q_3(u) &= \left(\sqrt{\frac{7}{11}}x\right)^2 + \left(-\frac{2}{\sqrt{11}}x + \sqrt{11}y\right)^2 - (3y-z)^2 \\ q_4(u) &= \left(\sqrt{\frac{8}{5}}x\right)^2 + \left(-\sqrt{\frac{2}{5}}x + \sqrt{\frac{5}{2}}y\right)^2 - \left(\frac{1}{\sqrt{2}}y + \frac{1}{\sqrt{2}}z\right)^2 \end{aligned}$$

197.

$$q(u) = \left(\sum_{j=1}^{n} j u_j\right)^2 + \frac{1}{2}\left(\sum_{j=1}^{n}(j^2+1)u_j\right)^2 - \frac{1}{2}\left(\sum_{j=1}^{n}(j^2-1)u_j\right)^2$$

198. $e = 0, \quad b = c, \quad a > 0, \quad ad - b^2 > 0$

$$\left(\sqrt{a}x + \frac{b}{\sqrt{a}}y\right)^2 + \left(\frac{\sqrt{ad-b^2}}{a}y\right)^2$$

199.

$$\begin{pmatrix} 10 & 3 & 0 \\ 3 & 2 & 1 \\ 0 & 1 & 1 \end{pmatrix}, \ 10 > 0, \ 11 > 0, \ 1 > 0, \quad e_1 = \begin{pmatrix} 1 \\ -3 \\ 3 \end{pmatrix}, ; e_2 = \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}, \ e_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

200. $(1,1,-2)/3$.

201. $(3,-1,2,0)$, $(4,-2,0,1)$.

202. $\omega_B(\mu A + \nu C) = \mu\omega_B(A) + \nu\omega_B(C)$. $(A,B) = \mathrm{tr}(B^t A)$ es un producto escalar.

203.

$$Ax = 0, \quad A = \begin{pmatrix} u_1 \\ \vdots \\ u_k \end{pmatrix}$$

$$W = \{x \in V \mid (u_i, x) = 0, \ i = 1, \ldots k\}, \quad W^\perp = \{y \in V \mid (y,x) = 0, \ x \in W\}$$

204.

$$q_1(t) = \frac{1}{\sqrt{2}}, \quad q_2(t) = \sqrt{\frac{3}{2}}t, \quad q_3(t) = \sqrt{\frac{5}{8}}(1-3t^2)$$

205.

$$\text{a) } Q(x)=\frac{1}{2}(2x_1+2x_2+x_3)^2+\frac{1}{2}x_3^2-2x_2^2, \quad \operatorname{sig}(\varphi)=(2,1)$$

$$\text{b) } P=\begin{pmatrix} \sqrt{2} & \sqrt{2} & 1/\sqrt{2} \\ 0 & 0 & 1/\sqrt{2} \\ 0 & \sqrt{2} & 0 \end{pmatrix}, \quad \text{c) } \left\{ \begin{pmatrix} 3 \\ -5 \\ 0 \end{pmatrix}, \begin{pmatrix} 3 \\ 0 \\ -5 \end{pmatrix} \right\}$$

206. a) No. b) $(0,1,1)$

$$\text{c) } \frac{1}{4}(x_1+x_2-x_3)^2-\frac{1}{4}(x_1-x_2+x_3)^2, \quad \operatorname{sig}=(1,1)$$

207.

$$\text{a) } u_1=\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad u_2=\frac{1}{\sqrt{10}}\begin{pmatrix} -2 \\ 2 \\ 1 \end{pmatrix}, \quad \text{b) } \sqrt{\frac{5}{2}}, \quad \text{c) No}$$

208.

$$Q=\begin{pmatrix} 1 & \frac{1}{2}(\alpha+\beta) & \frac{1}{2}(\alpha^2+\beta^2) \\ \frac{1}{2}(\alpha+\beta) & \alpha\beta & \frac{1}{2}\alpha\beta(\alpha+\beta) \\ \frac{1}{2}(\alpha^2+\beta^2) & \frac{1}{2}\alpha\beta(\alpha+\beta) & \alpha^2\beta^2 \end{pmatrix}$$

Si $\alpha \neq \beta$, rango $=2$, signatura $=(1,1)$ Si $\alpha=\beta$, rango $=1$, signatura $=(1)$

209. a) Cierta. b) Falsa. c) Cierta. d) Cierta. e) Cierta.

210.

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{a+\sqrt{a^2+4}}{2} & 0 \\ 0 & 0 & \frac{a-\sqrt{a^2+4}}{2} \end{pmatrix}$$

Si $a=3$, vectores isótropos en la base que diagonaliza a ϕ: $(x_1')^2+(x_2')^2-(x_3')^2=0$. No forman un subespacio lineal. ϕ nunca es definida positiva.

211. $q(x)=\left(\sqrt{3}x_1-\frac{2}{\sqrt{3}}x_2-\sqrt{3}x_3\right)^2+\left(\sqrt{\frac{5}{3}}x_2\right)^2+x_3^2$ $(1/\sqrt{3},0,0)$, $(2/\sqrt{15},\sqrt{3/5},0)$, $(1,0,1)$

212. Ortonormalizando por Gram–Schmidt: $\left\{v_1=\frac{1-i}{2}(1,0,i,0),\ v_2=\frac{1}{\sqrt{6}}(1,0,-i,2)\right\}$. Proyector:

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

Base ortonormal: v_1, v_2, $v_3=\frac{1}{\sqrt{3}}(1,0,-i,-1)$, $v_4=(0,1,0,0)$. $\mathcal{A}$ es unitario.

213. a) $\|Ru\|^2=(Ru,Eu)$, b) i no es autovalor de a. c) $A+i1_V$ y $A-i1_V$ conmutan.

214.

$$a)\ P=\begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{3} & 1/\sqrt{6} \\ 0 & 1/\sqrt{3} & -2/\sqrt{6} \\ -1/\sqrt{2} & 1/\sqrt{3} & 1/\sqrt{6} \end{pmatrix}$$

$$b)\ P_3=\frac{1}{6}\begin{pmatrix} 5 & 2 & -1 \\ 2 & 2 & 2 \\ -1 & 2 & 5 \end{pmatrix}, \quad P_{-3}=\frac{1}{6}\begin{pmatrix} 1 & -2 & 1 \\ -2 & 4 & -2 \\ 1 & -2 & 1 \end{pmatrix}$$

215. a) $\ker f = \{1, 0, 0, -1)\}$, $\operatorname{im} f = \{(-1, 0, 0, -1), (0, 1, 3, 0), (0, 3, 1, 0)\}$
b) $\{(1, 0, 0, -1)/\sqrt{2},\ (1, 0, 0, 1)/\sqrt{2},\ (0, 1, -1, 0)/\sqrt{2},\ (0, 1, 1, 0)/\sqrt{2},\ \}$

$$P_0 = \frac{1}{2}\begin{pmatrix} 1 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 1 \end{pmatrix}, \quad P_{-2} = \frac{1}{2}\begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & -1 & 0 \\ 0 & -1 & 1 & 0 \\ 1 & 0 & 0 & 1 \end{pmatrix}, \quad P_4 = \frac{1}{2}\begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

$$P = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & -1 & 1 \\ -1 & 1 & 0 & 0 \end{pmatrix}$$

216.

$$\text{a) } v_1 = \frac{1}{3}(2, -1, -2), \quad v_2 = \frac{1}{\sqrt{5}}(1, 2, 0), \quad v_3 = \frac{1}{3\sqrt{5}}(4, -2, 5), \quad a_1 = 0,\ a_2 = a_3 = 9$$

$$\text{b) } T = a_1\frac{1}{9}\begin{pmatrix} 4 & -2 & -4 \\ -2 & 1 & 2 \\ -4 & 2 & 4 \end{pmatrix} + a_2\frac{1}{9}\begin{pmatrix} 5 & 2 & 4 \\ 2 & 8 & -2 \\ 4 & -2 & 5 \end{pmatrix}$$

217.

$$P^{-1}QP = \begin{pmatrix} 1 & 0 \\ 0 & 6 \end{pmatrix}, \quad P = \frac{1}{\sqrt{5}}\begin{pmatrix} 2 & -1 \\ 1 & 2 \end{pmatrix}$$

$$q(v) = \left(\frac{2}{\sqrt{5}}v_1 + \frac{1}{\sqrt{5}}v_2\right)^2 + 6\left(-\frac{1}{\sqrt{5}}v_1 + \frac{2}{\sqrt{5}}v_2\right)^2$$

218. $\lambda = \mu$.

219. $S^\perp(V) = \{A \in M_n(\mathbb{C}) \mid A^t = -A\}$.

$$\frac{1}{2}\begin{pmatrix} 2a_{11} & a_{12}+a_{21} & a_{13}+a_{31} \\ a_{12}+a_{21} & 2a_{22} & a_{23}+a_{32} \\ a_{13}+a_{31} & a_{23}+a_{32} & 2a_{33} \end{pmatrix} + \frac{1}{2}\begin{pmatrix} 0 & a_{12}-a_{21} & a_{13}-a_{31} \\ -(a_{12}-a_{21}) & 0 & a_{23}-a_{32} \\ -(a_{13}-a_{31}) & -(a_{23}-a_{32}) & 0 \end{pmatrix}$$

220. $u_1 = (1, 0, 0, 1)/\sqrt{2}$, $u_2 = (0, 1, 1, 0)/\sqrt{2}$, $u_3 = (1, -i, i, -1)/2$, $u_4 = (1, i, -i, -1)/2$

$$P_1 = \frac{1}{2}\begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 \end{pmatrix}, \quad P_2 = \frac{1}{4}\begin{pmatrix} 1 & -i & i & -1 \\ i & 1 & -1 & -i \\ -i & -1 & 1 & i \\ -1 & i & -i & 1 \end{pmatrix}, \quad P_3 = \frac{1}{4}\begin{pmatrix} 1 & i & -i & -1 \\ -i & 1 & -1 & i \\ i & -1 & 1 & -i \\ -1 & -i & i & 1 \end{pmatrix}$$

221. $(1/2, 1/2, 2)$

222.

$$R(t) = \begin{pmatrix} \cos t & 0 & \operatorname{sen} t \\ 0 & 1 & 0 \\ -\operatorname{sen} t & 0 & \cos t \end{pmatrix}$$

223. $a = -1$. Eje: $(1, 1, 0)/\sqrt{2}$. Ángulo: π.

224. $A = -A^t$.

225. a) Ninguna. b) Infinitas.

226. a) $\lambda > 0$m. b) $\lambda^2 < 5/3$

227.

$$\text{a)} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \text{b)} \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

228.

$$\dim L_1 + \dim L_1^{\perp} = \dim H \Rightarrow \dim L_2 + \dim L_1^{\perp} > \dim H$$

$$\dim L_2 + \dim L_1^{\perp} = \dim H + \dim(L_2 \cap L_1^{\perp}) \Rightarrow \dim(L_2 \cap L_1^{\perp}) > 0$$

229. $(5, -5, -2, -1)$.

230.

$$\begin{pmatrix} 3 & 6 \\ -1 & -3 \end{pmatrix}$$

231. $\|x\|^2 = (x, x)$.

232.

$$P = \begin{pmatrix} a_1 & 0 & 0 & 0 \\ 0 & a_2 & 0 & 0 \\ 0 & 0 & a_3 & 0 \\ 0 & 0 & 0 & a_4 \end{pmatrix}, \quad |a_1| = |a_2| = |a_3| = |a_4| = 1$$

233. $A^k = I$ implica que los autovalores son las raíces k-ésimas de la unidad. Como A es real y simétrico, sus autovalores son reales, ± 1. Luego $A^2 = I$ por ser diagonalizable.

234.

$$\text{a) } f_q(x, y) = (v, w)(x, y) - \frac{1}{2}((v, x)(w, y) + (v, y)(w, x))$$

$$\text{b) } A_q x = (v, w)x - \frac{1}{2}((x, w)v + (x, v)w), \quad \text{c) } \ker A_q = (\operatorname{lin}\{v, w\})^{\perp}$$

235.

$$P = \begin{pmatrix} 1/2 & 1/\sqrt{2} & 0 & 1/2 \\ -1/2 & 0 & 1/\sqrt{2} & 1/2 \\ -1/2 & 0 & -1/\sqrt{2} & 1/2 \\ -1/2 & 1/\sqrt{2} & 0 & -1/2 \end{pmatrix}$$

236. Base de autovectores: $\{u_1, \ldots, u_n\}$

$$x = \sum_{i=1}^{n} c_i u_i, \ \sum_{i=1}^{n} |c_i|^2 = 1, \quad (x, Ax) = \sum_{i=1}^{n} \lambda_i |c_i|^2, \quad (x, T^+ T x) = \|Tx\|^2$$

237.

$$P_{-7} = \frac{1}{9} \begin{pmatrix} 4 & 2 & 4 \\ 2 & 1 & 2 \\ 4 & 2 & 4 \end{pmatrix}, \quad P_2 = \frac{1}{9} \begin{pmatrix} 5 & -2 & -4 \\ -2 & 8 & -2 \\ -4 & -2 & 5 \end{pmatrix}$$

238. $(-3, 1, 1)$, $\quad \cos\varphi = 7/18$.

239.

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

240.

$$P = \begin{pmatrix} 1/3 & 2/\sqrt{5} & 2/(3\sqrt{5}) \\ -2/3 & 1/\sqrt{5} & -4/(3\sqrt{5}) \\ -2/3 & 0 & 5/(3\sqrt{5}) \end{pmatrix}$$

$$P_{-3} = \frac{1}{9}\begin{pmatrix} 1 & -2 & -2 \\ -2 & 4 & 4 \\ -2 & 4 & 4 \end{pmatrix}, \quad P_6 = \frac{1}{9}\begin{pmatrix} 8 & 2 & 2 \\ 2 & 5 & -4 \\ 2 & -4 & 5 \end{pmatrix}$$

241.

$$P = \begin{pmatrix} 1/\sqrt{3} & 1/\sqrt{2} & 1/\sqrt{6} \\ -1/\sqrt{3} & 1/\sqrt{2} & -1/\sqrt{6} \\ -1/\sqrt{3} & 0 & 2/\sqrt{6} \end{pmatrix}$$

$$P_3 = \frac{1}{3}\begin{pmatrix} 2 & 1 & 1 \\ 1 & 2 & -1 \\ 1 & -1 & 2 \end{pmatrix}, \quad P_{-3} = \frac{1}{3}\begin{pmatrix} 1 & -1 & -1 \\ -1 & 1 & 1 \\ -1 & 1 & 1 \end{pmatrix}$$

242.

$$\text{a) } 1 > 0,\ 1 > 0,\ 1 > 0, \quad u_1 = e_1,\ u_2 = e_1 + e_2,\ u_3 = e_1 + e_2 + e_3$$

$$\text{b) } P = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}, \quad P^{-1}TP = \begin{pmatrix} 1 & 2 & 0 \\ 2 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

$$\text{c) } P_{-1} = \frac{1}{2}\begin{pmatrix} 1 & -1 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad P_1 = \frac{1}{2}\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad P_3 = \frac{1}{2}\begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

En la base canónica:

$$P_{-1} = \frac{1}{2}\begin{pmatrix} 0 & 0 & 0 \\ -1 & 2 & -1 \\ 0 & 0 & 0 \end{pmatrix}, \quad P_1 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{pmatrix}, \quad P_3 = \frac{1}{2}\begin{pmatrix} 2 & 0 & -2 \\ 1 & 0 & -1 \\ 0 & 0 & 0 \end{pmatrix}$$

243.

$$P_{-1} = \frac{1}{9}\begin{pmatrix} 5 & -2 & 4 \\ -2 & 8 & -2 \\ 4 & -2 & 5 \end{pmatrix}, \quad P_8 = \frac{1}{9}\begin{pmatrix} 4 & 2 & 4 \\ 2 & 1 & 2 \\ 4 & 2 & 4 \end{pmatrix}$$

244.

$$\text{a) } \sigma(A) = \{0, \frac{i}{\sqrt{6}}, -\frac{i}{\sqrt{6}}\}, \quad \text{b) } U = \begin{pmatrix} \frac{1}{\sqrt{6}} & -\frac{2}{\sqrt{6}} & -\frac{1}{\sqrt{6}} \\ \frac{1+2i\sqrt{6}}{2\sqrt{15}} & \frac{-2+i\sqrt{6}}{2\sqrt{15}} & \frac{5}{2\sqrt{15}} \\ \frac{1-2i\sqrt{6}}{2\sqrt{15}} & \frac{-2-i\sqrt{6}}{2\sqrt{15}} & \frac{5}{2\sqrt{15}} \end{pmatrix}$$

245.

$$\text{a) } u_1 = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \quad u_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad u_3 = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}$$

$$\text{b) } P_0 = \frac{1}{2}\begin{pmatrix} 1 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 1 \end{pmatrix}, \quad P_2 = \frac{1}{2}\begin{pmatrix} 1 & 0 & -1 \\ 0 & 2 & 0 \\ -1 & 0 & 1 \end{pmatrix},$$

$$\text{c) } d = 1, \quad \text{d) } e^A = \frac{1}{2}\begin{pmatrix} 1+e^2 & 0 & 1-e^2 \\ 0 & 2e^2 & 0 \\ 1-e^2 & 0 & 1+e^2 \end{pmatrix}$$

246. $(u_1, u_2) = -1/4$

247. a) Autoadjunto. b) $\{-1, 0, 2\}$

$$\text{c) } u_1 = \begin{pmatrix} -1/\sqrt{3} \\ -i/\sqrt{3} \\ 1/\sqrt{3} \end{pmatrix}, \quad u_2 = \begin{pmatrix} 0 \\ i/\sqrt{2} \\ 1/\sqrt{2} \end{pmatrix}, \quad u_3 = \begin{pmatrix} 2/\sqrt{6} \\ -i/\sqrt{6} \\ 1/\sqrt{6} \end{pmatrix}$$

$$\text{d) } P_1 = \frac{1}{3}\begin{pmatrix} 1 & -i & -1 \\ i & 1 & -i \\ -1 & i & 1 \end{pmatrix}, \quad P_2 = \frac{1}{2}\begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & i \\ 0 & -i & 1 \end{pmatrix}, \quad P_3 = v_3^+ v_3 = \frac{1}{6}\begin{pmatrix} 4 & 2i & 2 \\ -2i & 1 & -i \\ 2 & i & 1 \end{pmatrix}$$

$$\text{e) } \cos(\pi A) = \frac{1}{3}\begin{pmatrix} 1 & 2i & 2 \\ -2i & 1 & 2i \\ 2 & -2i & 1 \end{pmatrix}$$

248. $x_i \longrightarrow x'_i = (P^{-1})_{i'i} x_{i'}$
$\delta_{ik}\delta_{jl} \longrightarrow (P^{-1})_{i'i}(P^{-1})_{j'j}(P^{-1})_{k'k}(P^{-1})_{l'l}\delta_{i'k'}\delta_{j'l'} = (P^{-1})_{i'i}(P^{-1})_{j'j}(P^{-1})_{i'k}(P^{-1})_{j'l} = \delta_{ik}\delta_{jl}$

249. Base ortonormal de $\mathbb{R}^2$: $\{u_1, u_2\}$. Base girada: $v_1 = u_2, \quad v_2 = -u_1$.

$$P = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \quad t'_{ijk} = (P^{-1})^{i'}{}_i (P^{-1})^{j'}{}_j (P^{-1})^{k'}{}_k t_{i'j'k'},$$

$$t'_{111} = -1, \; t'_{112} = 1, \; t'_{122} = -1, \; t'_{222} = 1$$

250. Usando: $\epsilon^{kij}\epsilon_{jlm} = \delta^k{}_l \delta^i{}_m - \delta^k{}_m \delta^i{}_l$

251. Tipo: $(0,1)$. $(x \wedge y)_1 = x^2y^3 - x^3y^2, \quad (x \wedge y)_2 = x^3y^1 - x^1y^3, \quad (x \wedge y)_3 = x^1y^2 - x^2y^1$

252. $x^i x_i = 2, \quad y^i y_i = 8, \quad x^i y_i = -3.$

253. $\epsilon^{ijl} w_i v_j v_l + v^k w_k + n v^k v_k = n.$

254. Las matrices γ^μ, I_4 son l.i.
$\gamma^0\gamma^1$, $\gamma^0\gamma^2$, $\gamma^0\gamma^3$, $\gamma^1\gamma^2$, $\gamma^1\gamma^3$, $\gamma^2\gamma^3$ son l.i.
$\gamma^0\gamma^1\gamma^2$, $\gamma^0\gamma^1\gamma^3$, $\gamma^0\gamma^2\gamma^3$, $\gamma^1\gamma^2\gamma^3$ son l.i.
$\gamma^0\gamma^1\gamma^2\gamma^3$ es distinta de cero.
En total hay 16 matrices l.i. que forman una base de $\mathcal{M}_4(\mathbb{C})$

255. $C = c_0 + c_1 T^\mu{}_\mu + c_2 A^\mu A_\mu + c_3 T^{\mu\nu} T_{\mu\nu} + c_4 T^{\mu\nu} A_\mu A_\nu + c_5 T^\mu{}_\mu A^\nu A_\nu.$

256. Tipo $(2,1)$. $X'_i = P_{ji} X_j, \quad c'_{ijs} = P_{ki} P_{lj} (P^{-1})_{sm} c_{klm}$

257. Si A tiene autovalores λ_i, los de A^2 son λ_i^2 y los de A^3, λ_i^3 con los mismos autovectores. Por tanto:

$$\sum_{i=1}^{3} \lambda_i = \operatorname{tr} A = A^\mu{}_\mu, \quad \sum_{i=1}^{3} \lambda_i^2 = \operatorname{tr} A^2 = A^\mu{}_\nu A^\nu{}_\mu = A^{\mu\nu} A_{\mu\nu},$$

$$\sum_{i=1}^{3} \lambda_i^3 = \operatorname{tr} A^3 = A^\mu{}_\nu A^\nu{}_\rho A^\rho{}_\mu = A^{\mu\nu} A_{\nu\rho} A^\rho{}_\mu$$

258.

$$\text{a) } x^1 = -\frac{1}{3}, \; x^2 = \frac{2}{3}, \; x^3 = \frac{4}{3}, \quad y_1 = 5, \; y_2 = 3, \; y_3 = 2, \quad \text{b) } A^{ijk} y_i y_j y_k = 60$$

259. a) -4. b) 12. c) 4. d) -4.

260. $t'_{111} = 1, \; t'_{112} = \sqrt{2}, \; t'_{122} = 2, \; t'_{222} = 2\sqrt{2}$. Las demás son cero.

261. a) Contravariantes: $x'^1 = 1, \; x'^2 = 1$. Covariantes: $x'_1 = 6, \; x'_2 = 4$ b) i) Cierta. ii) Falsa. c) $r^{ij}_k s^k_{il} = 10$.