

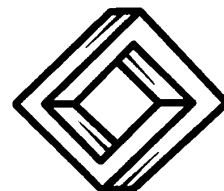
**Publicaciones Electrónicas
Sociedad Matemática Mexicana**

**Álgebra Lineal,
Álgebra Multilineal
y
K-Teoría Algebraica Clásica**

Emilio Lluis-Puebla

www.sociedadmatematicamexicana.org.mx

**Serie: Textos. Vol. 9 (2008)
ISBN 968-9161-31-8**



ALGEBRA LINEAL,
ALGEBRA MULTILINEAL
Y
K-TEORIA
ALGEBRAICA CLASICA

EMILIO LLUIS-PUEBLA

Universidad Nacional Autónoma de México

2008 Segunda Edición: Sociedad Matemática Mexicana,
Publicaciones Electrónicas

ISBN 968-9161-31-8 (versión en línea)

ISBN 968-9161-32-6 (versión en CD)

ISBN 968-9161-33-4 (versión en papel)

1997 Primera Edición: Sistemas Técnicos de Edición, S.A. de C.V.
San Marcos, 102. Tlalpan 14000 México, D.F.

©1996 Emilio Lluís-Puebla

Obra compuesta y formada en TEX por Flor de María Aceff Sánchez

Hecho en México.

ISBN 970-629-149-0

Sistemas Técnicos de Edición

ABCDEFGHIJKL-M 9987

INDICE GENERAL

Prefacio	v
Introducción	1
Capítulo I	
Conceptos Fundamentales	
I.1 Espacios vectoriales y funciones lineales	21
I.2 Subespacios vectoriales	27
I.3 Espacios vectoriales de dimensión finita	35
I.4 Aplicaciones	42
I.5 La matriz asociada a una función lineal	47
Capítulo II	
Vectores Característicos y Formas Canónicas	
II.1 Valores y vectores característicos	55
II.2 Teorema de Cayley-Hamilton	62
II.3 El polinomio mínimo	67
II.4 Forma canónica triangular	70
II.5 Forma canónica de Jordan	77

Capítulo III**Formas y Operadores**

III.1 Formas bilineales	87
III.2 Formas bilineales, simétricas, alternantes, cuadráticas y hermitianas	97
III.3 Producto escalar	103
III.4 Operadores adjuntos	109
III.5 El teorema espectral	114

Capítulo IV**Algebra Multilineal y K-Teoría Algebraica Clásica**

IV.1 Producto tensorial	121
IV.2 Producto exterior	128
IV.3 Estructuras algebraicas	135
IV.4 K_0 y K_1	141

Apéndice

Notas Históricas	151
-------------------------	-----

Bibliografía	173
---------------------	-----

Lista de Símbolos	175
--------------------------	-----

Índice Analítico	177
-------------------------	-----

PREFACIO

La mayoría de los textos de Algebra Lineal que se encuentran en nuestro país, en español, son traducciones de textos en inglés dirigidos a estudiantes de diversas disciplinas (Psicología, Música, Medicina, etc., para los cuales es obligatorio cursar Algebra Lineal). Por lo tanto, dichos textos no poseen el enfoque que deben tener para estudiantes de las carreras de Matemáticas o Física, por mencionar algunas. Los textos en otros idiomas dirigidos exclusivamente a estudiantes de Física o Matemática son escasos y en español lo son aún más. Es así que nuestra intención es la de proveer a los estudiantes de carreras científicas de un enfoque serio, fundamentado y moderno del Algebra Lineal. Hemos incluido el Algebra Multilineal, tema excluido de los programas usuales pero que nosotros pensamos que es de importancia fundamental, así como la notación que se utiliza en física para tensores. No se ha descuidado el aspecto computacional, al contrario, se incluyen diversos ejercicios de cálculo explícito. Sin embargo, han sido incluidos una gran cantidad de problemas interesantes que preparan al estudiante para realmente darle la oportunidad de crear matemáticas. Todos ellos se resuelven utilizando el material expuesto. Como consecuencia del trabajo de resolverlos, le brindan al estudiante la oportunidad de redactar matemáticas.

Suponemos que el lector está familiarizado con algunos de los temas básicos del Algebra Superior, es decir, que ya conoce y maneja estructuras algebraicas como \mathbf{N} , \mathbf{Q} , \mathbf{R} , \mathbf{C} , $\mathbf{C}[x]$, la definición de campo, y ha trabajado numéricamente con matrices, determinantes y sistemas de ecuaciones lineales. Es conveniente que haya conocido los espacios vectoriales \mathbf{R}^n sobre el campo \mathbf{R} .

Este libro está diseñado para un curso de un año (dos semestres, estudiando los capítulos I y II en el primero y los capítulos III y IV en el segundo; o bien, haciendo

modificaciones adecuadas, como por ejemplo, el capítulo I, II.1, II.2, II.3 y III.3 en el primer semestre y el resto en el segundo).

En la Introducción se presenta un panorama del presente texto. Obviamente no se pretende que el lector que no conozca el tema pueda comprender lo que en ella se expone. Al contrario, conforme el lector avance en el estudio, podrá regresar a ella y obtener una visión global del Algebra Lineal incluida en esta obra. Las referencias sin números romanos indican resultados del capítulo en consideración.

Hemos incluido un apéndice que contiene notas históricas sobre algunos de los conceptos definidos en el texto. Tiene como finalidad la de motivar el estudio del Algebra Lineal a través del análisis de las ideas que dieron lugar a dichos conceptos y del conocimiento de quienes contribuyeron a ellos. No está diseñado para una lectura continua, más bien, lo está para ser consultado conforme el lector encuentre los conceptos en cada sección. Organizamos el apéndice de acuerdo al orden de las secciones indicadas con una A antepuesta.

Durante varios años he impartido cursos basados en el material aquí incluido a alumnos de licenciatura, a quienes agradezco su atención y sus oportunos comentarios.

Del mismo modo, deseo agradecer a varios de mis colegas, entre ellos María Elena García y Mary Glazman el haber utilizado versiones preliminares en sus cursos; muy especialmente a mi padre, Emilio Lluís Riera por haber hecho importantes comentarios, sugerencias y correcciones al haber utilizado el texto varios años. También, mi mayor agradecimiento a mi esposa Flor de María Aceff, quien además de haberlo utilizado en sus cursos, hecho numerosas correcciones y proporcionado agradables conversaciones sobre el libro, realizó la formación y composición en el sistema $T_E X$ de las múltiples versiones preliminares. Sin embargo, cualquier falta u omisión es exclusivamente mía.

Igualmente, agradezco a María de Jesús Figueroa el haber escrito las notas históricas como consecuencia de su largo estudio de investigación para realizar su tesis doctoral y a Alejandro Garciadiego a quien agradecemos el asesoramiento experto para la realización de dichas notas.

Finalmente, agradezco a Rosa Quintana la captura preliminar del texto. Este libro, producto de varios años de trabajo, se elaboró durante su fase final bajo el auspicio del CONACYT a través de una Cátedra Patrimonial.

Emilio Lluís Puebla

PREFACIO (SEGUNDA EDICION)

Este libro cumple ya más de diez años de ser utilizado exitosamente como texto sobre la materia en diversas universidades del Continente Americano, incluyendo algunas universidades de Estados Unidos de Norteamérica y, desde luego, en México.

He tenido el gusto de ofrecer conferencias en muchas universidades de Centroamérica y Sudamérica donde me he encontrado con colegas, que llevan mi libro como texto. Me han solicitado una nueva edición pues la anterior es imposible de conseguir. Esta nueva edición, donde he corregido algunos errores tipográficos y atendido nuevas ideas o sugerencias que al través de los años me he hecho y han hecho mis propios alumnos (a quienes mucho agradezco), la he incluido dentro de las Publicaciones Electrónicas de la Sociedad Matemática Mexicana mostrando (como matemático y Editor Ejecutivo de las mismas) la confianza en este tipo de publicación. Éste tiene una excelente accesibilidad, así como un nulo costo, que de no ser así, resultaría elevado para los estudiantes y colegas de muchos lugares. Las Publicaciones Electrónicas de la Sociedad Matemática Mexicana tienen acceso libre en línea, pueden copiarse en el ordenador o imprimirse en papel para uso personal. Además, el lector podrá adquirir las publicaciones en CD o impresas en papel empastado.

Primavera de 2008

Emilio Lluis-Puebla

— |

| —

— |

| —

INTRODUCCION

Siempre que se habla de alguna rama de la Matemática se establecen los objetos de estudio. Los objetos que estudiaremos serán los espacios vectoriales (que son un caso especial de los módulos). Sea K un campo. Diremos que el conjunto V junto con la operación binaria $+$ y acción μ de K en V forman un *espacio vectorial sobre un campo K* si bajo $+$, V es un grupo abeliano, μ distribuye tanto a la suma de elementos de V como a la suma de elementos de K , la acción del producto de dos elementos de K es igual a uno de ellos por la acción del otro y finalmente, la acción del elemento unitario de K en V es trivial.

Los elementos de V se llaman *vectores*, los del campo K se llaman *escalares* y la acción μ se llama *multiplicación escalar*. K^n , $K[x]$ y el conjunto de las matrices de $m \times n$ con elementos en K denotado con $M_{m \times n}K$ son ejemplos de espacios vectoriales.

¿Cómo relacionamos dos espacios vectoriales sobre un campo K ? Así como a los conjuntos los podemos relacionar mediante funciones, a los espacios vectoriales los relacionaremos mediante funciones que preservan la estructura de espacio vectorial llamadas *homomorfismos* o *funciones lineales* (o *aplicaciones* o *transformaciones lineales*). Entonces, si U y V son espacios vectoriales sobre un campo K , $f: U \rightarrow V$ es un *homomorfismo* o *función lineal* si $f(u + v) = f(u) + f(v)$ y además $f(\alpha v) = \alpha f(v)$; $\alpha \in K$; $u, v \in U$. Resulta que la composición de homomorfismos es un homomorfismo. Se dice que una función lineal $f: U \rightarrow V$ es un *isomorfismo* si existe una función $g: V \rightarrow U$ tal que $g \circ f = 1_U$ y $f \circ g = 1_V$. Es un hecho el que si dicha g existe, es lineal, está determinada en forma única, se

denota con f^{-1} y se llama *inversa* de f . Si existe un isomorfismo entre dos espacios vectoriales U y V se dice que los espacios son *isomorfos* y escribimos $U \cong V$.

Si K es un campo, el conjunto de homomorfismos o funciones lineales de U en V lo denotamos con $\text{Hom}_K(U, V)$ (o con $L(U, V)$, $A(U, V)$). Le podemos dar una estructura de espacio vectorial sobre K . Una gran parte del Algebra Lineal consiste del estudio de este espacio vectorial.

Dada una función lineal $f: U \rightarrow V$ entre espacios vectoriales sobre un campo K podemos considerar su *núcleo*, es decir, el conjunto de todos los elementos de U cuya imagen es el cero de V , denotado con $\ker f$. Podemos considerar la *imagen* de f , es decir, el conjunto de los elementos de V que provienen de elementos de U , denotado con $\text{im } f$. También podemos considerar *subespacios* de un espacio vectorial. Estos son subconjuntos tales que el cero del espacio pertenece al subconjunto y este último es cerrado bajo la suma y multiplicación escalar.

Sucede lo esperado: la imagen bajo una función lineal de un subespacio es un subespacio; la imagen inversa de un subespacio bajo una función lineal es un subespacio; en particular, el núcleo y la imagen de una función lineal son subespacios (de donde deben serlo).

La *suma* de dos subespacios U y V de W , denotada $U + V$, es el conjunto de todas las sumas $u + v$ donde $u \in U$ y $v \in V$. Se dice que W es *suma directa interna* de U y V si todo elemento de W puede escribirse en forma única como suma de uno de U y uno de V y escribimos $W = U \oplus V$. Podemos definir la *suma directa externa* de espacios vectoriales $\{V_i\}_{i=1}^n$ sobre un campo K y la denotamos con $\bigoplus_{i=1}^n V_i$ como el espacio vectorial cuyos elementos son listas ordenadas de la forma (v_1, \dots, v_n) con las operaciones usuales de suma y multiplicación escalar. Si un espacio vectorial V es suma directa interna de subespacios de V entonces es isomorfo a la suma directa externa de los mismos subespacios. En vista de esto último hablaremos de *la* suma directa.

Un espacio vectorial es suma directa de dos subespacios si, y sólo si, es suma de ellos y su intersección es vacía. La suma directa posee la siguiente propiedad importante llamada *universal*: si $\varphi_j: V_j \rightarrow V$ son funciones lineales de espacios vectoriales e $\iota_j: V_j \rightarrow \bigoplus V_i$ son las inclusiones para $i \in I = \{1, \dots, n\}$, entonces existe una función lineal única $\varphi: \bigoplus_{i=1}^n V_i \rightarrow V$ tal que $\varphi \circ \iota_j = \varphi_j$, $j \in I$.

Esta propiedad caracteriza a la suma directa y podemos representarla mediante el diagrama

$$\begin{array}{ccc}
 & & V \\
 & \nearrow \varphi_j & \uparrow \varphi \\
 V_j & \xrightarrow{v_j} & \bigoplus_{i=1}^n V_i
 \end{array}$$

Decimos que un vector $v \in V$ es una *combinación lineal* de elementos de un subconjunto S de V si existe un número finito de elementos $\{v_i\}_{i=1}^n$ de S tal que $v = \sum_{i=1}^n \alpha_i v_i$, $\alpha_i \in K$. Las α_i se llaman *coeficientes*. El conjunto de todas las combinaciones lineales $\langle S \rangle$ de un subconjunto no vacío S de V es un subespacio que contiene a S y es el más pequeño de los subespacios de V que contiene a S . Dicho espacio se llama *subespacio generado por S* y es, por lo tanto, la intersección de todos los subespacios que contienen a S . Como caso particular, si $\langle S \rangle = V$, todo elemento de V es una combinación lineal de elementos de S y diremos que V está *generado* por el subconjunto S de V .

El siguiente resultado es fundamental y es consecuencia de la propiedad universal para la suma directa: considere $K^n = \bigoplus_{j=1}^n K_j$ donde cada $K_j \cong K$, K un campo fijo, $g: \{1, \dots, n\} \rightarrow K^n$ dada por $i \mapsto e_i$ y V un espacio vectorial sobre K . Entonces para toda función $f: \{1, 2, \dots, n\} \rightarrow V$ existe una función lineal única $\phi: \bigoplus_{j=1}^n K_j \rightarrow V$ tal que $f = \phi \circ g$, es decir, el siguiente diagrama *conmuta*:

$$\begin{array}{ccc}
 & & V \\
 & \uparrow \phi & \nwarrow f \\
 K^n \cong \bigoplus K_j & \xleftarrow[g]{} & \{1, 2, \dots, n\}
 \end{array}$$

La función g se llama *función canónica*.

Diremos que el conjunto $\{v_j\}_{j=1}^n$ de vectores de V es (i) *linealmente independiente* si ϕ es inyectiva, (ii) un conjunto de *generadores* si ϕ es suprayectiva y (iii) una *base* si ϕ es biyectiva.

En otras palabras, el conjunto $\{v_j\}$ es linealmente independiente si

$$\phi \left(\sum_{j=1}^n \alpha_j e_j \right) = \sum_{j=1}^n \alpha_j v_j = 0$$

implica que $\alpha_j = 0$ para toda $j = 1, \dots, n$; $\alpha_j \in K_j$. El que ϕ sea suprayectiva equivale a decir que todo elemento de V se puede escribir como $\sum_{j=1}^n \alpha_j v_j$, es decir,

como una combinación lineal. El que ϕ sea biyectiva quiere decir que todo elemento de V puede escribirse de una, y solamente una manera, en la forma $\sum_{j=1}^n \alpha_j v_j$. Es claro que el conjunto $\{e_j\}_{j=1}^n$ es una base de $\oplus_{j=1}^n K_j = K^n$ llamada *canónica*. También diremos que el conjunto $\{v_j\}_{j=1}^n$ de vectores de V es *linealmente dependiente* si no es linealmente independiente.

Es inmediato, de la definición de base, que todo espacio vectorial V sobre un campo K con base $\{v_j\}_{j=1}^n$ es isomorfo a K^n . Cualquier base de V posee la misma cardinalidad y los espacios K^n y K^m son isomorfos si, y sólo si, $n = m$. Esto nos permite definir el concepto de dimensión. Definimos la *dimensión* de un espacio vectorial V sobre un campo K , denotada $\dim V$, como el número de elementos de una de sus bases podemos poseer una teoría de la dimensión para los espacios vectoriales. Así, podemos decir que dos espacios vectoriales son isomorfos si, y sólo si, tienen la misma dimensión.

Un resultado que relaciona la dimensión de la suma de subespacios con la de cada uno de ellos es

$$\dim(U + V) = \dim U + \dim V - \dim(U \cap V)$$

donde U y V son subespacios de algún espacio W , y como consecuencia inmediata

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

Si $f: U \rightarrow V$ es una función lineal entonces se tiene que

$$\dim(U) = \dim(\text{im } f) + \dim(\text{ker } f).$$

Utilizaremos este hecho en lo subsecuente.

Existe una relación fundamental entre los homomorfismos y las matrices. Un sistema de m ecuaciones con n incógnitas puede escribirse en la forma

$$AX = B$$

donde

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix}, \quad X = {}^t(x_1, \dots, x_n) \text{ y } B = {}^t(b_1, \dots, b_m).$$

Entonces cualquier matriz A de $m \times n$ determina una función lineal $f = A: K^n \rightarrow K^m$ (por abuso de notación) dada por $v \mapsto Av$ donde los vectores

se llama *matriz de transición* de la base β en la base γ .

La matriz de transición N_{β}^{γ} puede verse como la matriz asociada a la función lineal $1_U: U \rightarrow U$ con respecto a las bases β y γ .

Tenemos los siguientes resultados:

- (i) si $f \in \text{Hom}_K(U, U)$ y N es la matriz de transición de la base $\beta = \beta'$ a la base $\gamma = \gamma'$ entonces $[f]_{\gamma'}^{\gamma'} = N^{-1}[f]_{\beta}^{\beta'} N$ y
- (ii) si $f \in \text{Hom}_K(U, V)$, N es la matriz de transición de la base β en la base γ de U y M es la matriz de transición de la base β' en la base γ' de V entonces $[f]_{\gamma'}^{\gamma'} = M^{-1}[f]_{\beta}^{\beta'} N$.

Finalmente, existe un isomorfismo

$$\text{Hom}_K(U, U) \cong M_n(K)$$

dado por $f \mapsto [f]_{\beta}^{\beta}$.

Pasamos ahora al estudio del espacio vectorial $\text{Hom}_K(U, U)$ cuando U es un espacio vectorial de dimensión finita sobre K . Sus elementos los llamaremos *operadores lineales*. Podemos definir otra operación binaria en $\text{Hom}_K(U, U)$ mediante $\rho_2 \rho_1(v) = \rho_2(\rho_1(v))$ la cual hace de $\text{Hom}_K(U, U)$ un álgebra sobre K . Así, $\text{Hom}_K(U, U)$ es un objeto provisto de varias estructuras que lo hace sumamente interesante. (En efecto, si \mathcal{A} es un álgebra con uno sobre un campo K entonces \mathcal{A} resulta ser isomorfa a una subálgebra de $\text{Hom}_K(U, U)$.)

El resultado (i) anterior nos lleva a definir lo siguiente: dos matrices cuadradas A y B son *similares* (o *semejantes*) si $A = N^{-1}BN$ con N una matriz invertible. De aquí que las matrices A y B representan al mismo operador lineal $f \in \text{Hom}_K(U, U)$ si, y sólo si, son similares una con la otra. La relación de similaridad o semejanza es una relación de equivalencia, de manera que las matrices asociadas a un operador lineal específico constituyen una clase de equivalencia. Para el caso de matrices, sumarlas y multiplicarlas cuando éstas son diagonales, es muy sencillo. Simplemente uno suma o multiplica los elementos correspondientes de las diagonales. Esta es una buena razón para saber cuales matrices son similares a una matriz diagonal.

Un operador lineal $f \in \text{Hom}_K(U, U)$ es *diagonalizable* si para alguna base de U , la matriz asociada es diagonal. También decimos que dicha base *diagonaliza* a f . Entonces podemos afirmar que f es diagonalizable si, y sólo si, existe una matriz invertible N tal que $N^{-1}BN$ es diagonal. Quisiéramos dar un criterio para saber cuándo un operador lineal f es diagonalizable. Para ello necesitamos definir algunos conceptos.

Sea $f \in \text{Hom}_K(U, U)$. Si existe un vector $u \in U$ distinto de cero y un escalar $\lambda \in K$ tal que $f(u) = \lambda u$ llamaremos a λ *valor característico* y a u *vector característico* correspondiente a λ . Resulta que el conjunto de vectores característicos U_λ correspondientes a un valor característico λ junto con el cero es un subespacio de U y es igual al núcleo del operador $\lambda I - f$. El siguiente teorema nos dice cuando un operador es diagonalizable:

TEOREMA. *Sea $f \in \text{Hom}_K(U, U)$. f es diagonalizable si, y sólo si, U posee una base que consta de vectores característicos de f .*

A continuación asociaremos unos polinomios a operadores lineales y matrices cuadradas que son de fundamental importancia.

Sea A una matriz cuadrada con coeficientes en K , $\rho \in \text{Hom}_K(U, U)$ y $f(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0$ un polinomio con coeficientes en K . Definimos

$$\begin{aligned} f(A) &= a_n A^n + a_{n-1} A^{n-1} + \cdots + a_1 A + a_0 I \quad \text{y} \\ f(\rho) &= a_n \rho^n + a_{n-1} \rho^{n-1} + \cdots + a_1 \rho + a_0 I \end{aligned}$$

donde ρ^n es la composición de ρ , n veces. Se dice que A o ρ es *raíz* del polinomio f si $f(A) = 0$ o si $f(\rho) = 0$. Si A es una matriz cuadrada de $m \times m$, la matriz cuadrada $\lambda I_m - A$ se llama *matriz característica*, el determinante de la matriz característica se llama *polinomio característico* y lo denotamos con

$$p_A(\lambda) = |\lambda I_m - A|.$$

Resulta que toda matriz cuadrada es raíz de su polinomio característico. Este es el famoso enunciado de Cayley-Hamilton.

Un criterio para saber si un escalar es un valor característico es la siguiente

PROPOSICION. *$\alpha \in K$ es un valor característico de A si, y sólo si, α es una raíz del polinomio característico $p_A(\lambda)$.*

Es un hecho el que si el polinomio característico de A es producto de factores lineales de la forma $(\lambda - a_1)(\lambda - a_2) \cdots (\lambda - a_n)$ con todas las a_i distintas entonces A es similar a una matriz diagonal con las a_i en la diagonal. Finalmente, si dos matrices A y B son similares entonces $p_A(\lambda) = p_B(\lambda)$.

Debido a ésto último decimos que la función $M_n(K) \rightarrow K[\lambda]$ que asigna a cada matriz de $n \times n$ su polinomio característico es un *invariante* bajo la relación de similaridad.

El polinomio característico es un polinomio para el cual $p_A(A) = 0$, A una matriz cuadrada. Puede haber otros polinomios para los cuales A sea raíz. El

polinomio mónico de menor grado tal que A es su raíz lo denotaremos con $m_A(\lambda)$ y se llama el *polinomio mínimo* de A . Dicho polinomio es único, divide al polinomio característico y posee los mismos factores irreducibles que el característico de A . También, λ es un valor característico de A si, y sólo si, λ es una raíz del polinomio mínimo de A . Aún más, una matriz cuadrada A es diagonalizable si, y sólo si, $m_A(\lambda) = (\lambda - \lambda_1)(\lambda - \lambda_2) \cdots (\lambda - \lambda_r)$ donde $\lambda_1, \dots, \lambda_r$ son los valores característicos distintos de A .

El concepto de similaridad de matrices se traduce en uno de *similaridad de operadores* el cual, al igual que el de matrices, es una relación de equivalencia. ¿Cómo determinamos si dos operadores son similares? o equivalentemente, ¿cómo podemos distinguir las clases de equivalencia? Para hacerlo, definiremos ciertas matrices llamadas formas canónicas, una para cada clase de equivalencia. Definimos un conjunto de *formas canónicas* para una relación de equivalencia \sim en un conjunto C como un subconjunto F de C que consiste de exactamente un elemento de cada clase de equivalencia de \sim . Así, una vez obtenidas, bastará comparar si son las mismas para cada operador. Existen varias formas canónicas, nosotros consideraremos en este texto únicamente la triangular y la de Jordan.

Sea U un subespacio de V . Denotamos con $v + U$ el conjunto de todas las expresiones de la forma $v + u$ donde u recorre todos los elementos de U . Dichos $v + U$ los llamaremos *clases laterales* de U en V . Es inmediato comprobar que cualesquiera dos clases laterales o son ajenas o son iguales. Denotamos con V/U el conjunto de todas las clases laterales de U en V . Si definimos $(u + U) + (w + U) = (v + w) + U$ y $\lambda(v + U) = \lambda v + U$ hacemos de V/U un espacio vectorial llamado *espacio cociente*. La función lineal $p: V \rightarrow V/U$ tal que $v \mapsto v + U$ se llama *proyección canónica*.

Si U es un subespacio de V tal que $\rho(U) \subset U$ para $\rho \in \text{Hom}_K(V, V)$ decimos que U es *invariante* bajo ρ . Dicho operador $\rho: V \rightarrow V$ induce un operador de U denotado $\rho|_U$. Se sabe que $\dim V = \dim U + \dim V/U$. Diremos que un operador $\rho \in \text{Hom}_K(V, V)$ puede representarse por una matriz triangular si su matriz asociada con respecto a alguna base lo es. Su polinomio característico se factoriza como producto de polinomios lineales. Lo inverso es cierto: si el polinomio característico de ρ se factoriza como producto de polinomios lineales entonces existe una base de V para la cual la matriz asociada es triangular. Traducido a matrices tenemos que si A es una matriz cuadrada cuyo polinomio característico se factoriza en polinomios lineales entonces A es similar a una matriz triangular.

Se dice que un operador $\rho \in \text{Hom}_K(V, V)$ es *descomponible como suma directa de operadores* $\rho|_{U_i}$ si $V = \bigoplus_{i=1}^s U_i$ con U_i invariante bajo ρ . Escribimos

$\rho = \bigoplus_{i=1}^s \rho|_{U_i}$. El siguiente resultado se conoce como el teorema de la *descomposición primaria*.

TEOREMA. Si $\rho \in \text{Hom}_K(V, V)$ posee el polinomio mínimo

$$m_\rho(\lambda) = f_1(\lambda)^{\eta_1} f_2(\lambda)^{\eta_2} \dots f_s(\lambda)^{\eta_s}$$

donde los $f_i(\lambda)$ son polinomios mónicos irreducibles distintos, entonces V es suma directa de los subespacios $\ker f_i(\rho)^{\eta_i}$ y éstos son invariantes bajo ρ . Aún más, $f_i(\lambda)^{\eta_i}$ es el polinomio mínimo de $\rho|_{\ker f_i(\rho)^{\eta_i}}$.

Como consecuencia se tiene que $\rho \in \text{Hom}_K(V, V)$ posee una matriz asociada diagonal si, y sólo si, su polinomio mínimo $m_\rho(\lambda)$ es producto de polinomios lineales distintos.

Un operador lineal $\rho \in \text{Hom}_K(V, V)$ se llama *nilpotente* si $\rho^n = 0$ para alguna $n > 0$. El entero r es el *índice de nilpotencia* de ρ si $\rho^r = 0$ pero $\rho^{r-1} \neq 0$. También diremos que una matriz cuadrada A es *nilpotente* si $A^n = 0$ y r es el *índice de nilpotencia* de A si $A^r = 0$ pero $A^{r-1} \neq 0$. Obsérvese que el polinomio mínimo de un operador nilpotente de índice r es $m_\rho(\lambda) = \lambda^r$ y su único valor característico es el cero. Entonces existe una base del espacio vectorial tal que la matriz asociada a ρ es triangular. El encontrar formas canónicas para dichos operadores nilpotentes nos permiten encontrar formas canónicas para cualquier operador que se factorice como producto de polinomios lineales. Así tenemos la siguiente

PROPOSICION. Si $\rho \in \text{Hom}_K(V, V)$ es de índice de nilpotencia r y $v \in V$ tal que $\rho^{r-1}(v) \neq 0$, entonces el conjunto $\{\rho^{r-1}(v), \rho^{r-2}(v), \dots, \rho(v), v\}$ es una base del subespacio que genera, cuya matriz asociada posee índice de nilpotencia r y es de la forma

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

Aún más, si $\rho \in \text{Hom}_K(V, V)$ es de índice de nilpotencia r , entonces ρ posee una matriz asociada diagonal por bloques que son de la forma de la matriz anterior.

Se sabe que existe al menos una matriz de orden r y que las otras son de órdenes $\leq r$. El número de matrices está determinado en forma única por ρ y el número de matrices de todos los órdenes es igual a la dimensión de $\ker \rho$.

Finalmente tenemos el siguiente

TEOREMA. *Sea $\rho \in \text{Hom}_K(V, V)$ tal que sus polinomios característico y mínimo se factorizan como producto de potencias de polinomios lineales. Entonces ρ posee una matriz asociada diagonal por bloques J , llamada forma canónica de Jordan de ρ cuyo bloques son de la forma*

$$J_{ij} = \begin{pmatrix} \lambda_i & 1 & 0 & \cdots & 0 & 0 \\ 0 & \lambda_i & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & \lambda_i & 1 \\ 0 & 0 & 0 & \cdots & 0 & \lambda_i \end{pmatrix}.$$

Así, dos operadores lineales cualesquiera son similares si, y sólo si, poseen la misma forma canónica de Jordan salvo el orden de los bloques.

Una función $f: U \times V \rightarrow W$ de espacios vectoriales sobre un campo K se llama *bilineal* si es lineal en cada variable cuando la otra se mantiene fija. El ser bilineal no quiere decir que sea lineal ni viceversa. Si $W = K$ diremos que f es una *forma bilineal* y denotamos con $L^2(U, V; K)$ el conjunto de formas bilineales de $U \times V$ en K . Si $U = V$, utilizaremos la notación $\text{Bil}(V)$ y le podemos dar a $\text{Bil}(V)$ una estructura de espacio vectorial sobre K .

Considere el espacio vectorial sobre K , $\text{Hom}_K(V, K)$. Sus elementos $f: V \rightarrow K$ se llaman *funcionales lineales* o *formas lineales*, se acostumbra denotar a $\text{Hom}_K(V, K)$ con $L^1(V; K)$ o simplemente V^* y se le llama *espacio dual* de V . Se tienen los siguientes resultados:

- (i) Sea $\{v_i\}_{i=1}^n$ una base de V y $\{f_i\}_{i=1}^n \in \text{Hom}_K(V, K) = V^*$ funcionales tales que $f_i(v_j) = \delta_{ij}$. Entonces $\{f_i\}_{i=1}^n$ es una base de V^* y $\dim V^* = n$.
- (ii) Si $\{f_i\}_{i=1}^n$ es una base de V^* entonces $\{f_{ij}\}_{i,j=1}^n$ dada por $f_{ij}(u, v) = f_i(u)f_j(v)$ es una base para $\text{Bil}(V)$ y $\dim \text{Bil}(V) = n^2$.
- (iii) Sea $\gamma = \{v_i\}_{i=1}^n$ una base de V y $f: V \times V \rightarrow K$ una forma bilineal. Si $u = \sum_{i=1}^n \alpha_i v_i$ y $v = \sum_{j=1}^n \beta_j v_j$ entonces

$$f(u, v) = \sum_{i,j=1}^n \alpha_i \beta_j f(v_i, v_j).$$

Si $A = (a_{ij})$ es la matriz cuadrada tal que $a_{ij} = f(v_i, v_j)$ entonces

$$f(u, v) = \sum_{i,j=1}^n \alpha_i \beta_j a_{ij} = {}^t[u]_\gamma A [v]_\gamma$$

y A se llama *matriz asociada a la forma bilineal f con respecto a la base γ* , también denotada $[f]_\gamma$. Se tienen los siguientes resultados para una forma bilineal $f: V \times V \rightarrow K$:

- (i) Si N es la matriz de transición de una base γ en otra γ' de V entonces la matriz B asociada a f con respecto a γ' es $B = {}^tNAN$.
- (ii) $Bil(V) \cong M_n(K)$ dado por $f \mapsto [f]_\gamma$.
- (iii) Si N es la matriz de transición de la base $\{u_i\}$ en la base $\{v_i\}$, entonces ${}^tN^{-1}$ es la matriz de transición de las bases duales $\{f_i\}$ en $\{g_i\}$.
- (iv) Sea $V^{**} = (V^*)^*$, entonces $V \cong V^{**}$.

Diremos que una forma bilineal de V es *simétrica* si $f(u, v) = f(v, u)$ para toda $u, v \in V$. Sucede que una forma bilineal es simétrica si, y sólo si, su matriz asociada es simétrica, i.e. es igual a su traspuesta. Si f posee una matriz asociada diagonal entonces f es simétrica.

La *forma cuadrática* asociada a f es la función $q: V \rightarrow K$ dada por $q(v) = f(v, v)$, $v \in V$. Sea A la matriz simétrica asociada a la forma bilineal simétrica f . Entonces $q(X) = f(X, X) = {}^tXAX = \sum_{i,j} a_{ij}x_ix_j$. Si A es diagonal, $q(X) = a_{11}x_1^2 + \dots + a_{nn}x_n^2$. La fórmula $f(u, v) = [q(u+v) - q(u) - q(v)]/2$ permite obtener f a partir de q .

Si f es una forma bilineal simétrica y K es de característica diferente de 2, entonces existe una base de V tal que f posee una matriz asociada diagonal. Una matriz simétrica B es *congruente* con una matriz simétrica A si existe una matriz no singular o invertible N tal que $B = {}^tNAN$. Así, si A es una matriz simétrica con elementos en un campo de característica diferente de 2, entonces A es congruente con una matriz diagonal.

Decimos que una forma bilineal f es *antisimétrica* si $f(u, v) = -f(v, u)$ para toda $u, v \in V$. Si V es de dimensión finita, f es antisimétrica si, y sólo si, su matriz asociada A es tal que $A = -{}^tA$, es decir, *antisimétrica*. También decimos que f es *alternante* si $f(v, v) = 0$ para toda $v \in V$. Se tiene que toda forma bilineal es suma de una simétrica y una antisimétrica.

Si consideramos el caso en que $K = \mathbb{C}$, una forma $f: V \times V \rightarrow \mathbb{C}$ se llama *hermitiana* si f es lineal en la primera variable y $f(u, v) = \overline{f(v, u)}$ para $u, v \in V$. La forma cuadrática $q: V \rightarrow \mathbb{R}$ asociada a f , dada por $q(v) = f(v, v)$ se llama *forma cuadrática hermitiana*.

Si $K = \mathbb{R}$, decimos que $f: V \times V \rightarrow \mathbb{R}$ está *definida positivamente* si $f(v, v) > 0$ para toda $v \in V$, $v \neq 0$.

Ahora consideraremos espacios vectoriales sobre \mathbb{R} o \mathbb{C} . En ellos podemos definir una forma bilineal simétrica o hermitiana definida positivamente $\langle \cdot, \cdot \rangle$ llamada *producto escalar* o *interno* sobre \mathbb{R} o \mathbb{C} . Este producto escalar permite definir los conceptos de longitud y ángulo. La *norma* o *longitud* $\|v\|$ de un vector v que pertenece a un espacio vectorial sobre $K = \mathbb{R}$ o \mathbb{C} se define como $\sqrt{\langle v, v \rangle}$. Dos vectores son *ortogonales* si $\langle v, w \rangle = 0$. El ángulo θ entre dos vectores $u, v \in V$ diferentes de cero se define como $\theta = \arccos(\langle u, v \rangle / (\|u\| \|v\|))$ para $\theta \in [0, \pi]$. El conjunto $U^\perp = \{v \in V \mid \langle u, v \rangle = 0 \forall u \in U, U \text{ un subconjunto de } V\}$ se llama *conjunto ortogonal* a U y resulta ser un subespacio de V . Sea $\{v_i\}_{i=1}^n$ un conjunto de vectores de V . $\{v_i\}_{i=1}^n$ es *ortogonal* si $\langle v_i, v_j \rangle = 0$ para $i \neq j$ y *ortonormal* si $\langle v_i, v_j \rangle = \delta_{ij}$. El siguiente teorema es de particular importancia y en su demostración se establece un procedimiento para encontrar una base ortonormal de un espacio vectorial de dimensión finita llamado *procedimiento de Gram-Schmidt*:

TEOREMA. *Sea $\{u_i\}_{i=1}^n$ una base del espacio vectorial de dimensión finita V sobre \mathbb{R} o \mathbb{C} . Entonces existe una base ortonormal $\{v_i\}_{i=1}^n$ de V tal que la matriz de transición es triangular.*

Un resultado útil es el siguiente: si U es un subespacio de V entonces $V \cong U \oplus U^\perp$. Así, podemos hablar de una proyección llamada *ortogonal* de V en U , $p_U: V \rightarrow V$ tal que $\text{im } p_U = U$ y $\text{ker } p_U = U^\perp$.

Un espacio que posee un producto escalar se llama *espacio con producto escalar*. Sea V un espacio vectorial con producto escalar sobre un campo $K = \mathbb{R}$ o \mathbb{C} y $g: V \rightarrow V^*$ dada por $g(v)(u) = g_v(u) = \langle u, v \rangle$. Así claramente, cada vector $v \in V$ nos determina un funcional g_v . Lo inverso también sucede: si V es de dimensión finita y $f: V \rightarrow K$ un funcional, entonces existe un vector único $v \in V$ tal que $f(u) = \langle u, v \rangle$, para toda $u \in V$. Estos resultados nos dicen que cualquier funcional es igual al producto escalar con un vector fijo de V .

Sea u un elemento fijo de un espacio vectorial de dimensión finita V con producto escalar y $\rho \in \text{Hom}_K(V, V)$. Consideremos $f \in V^*$ un funcional dada por $f(v) = \langle \rho(v), u \rangle$. Luego, existe un vector único $u' \in V$ tal que $\langle \rho(v), u \rangle = \langle v, u' \rangle$ para toda $v \in V$. Definimos $\rho^*: V \rightarrow V$ tal que $\rho^*(u) = u'$. Entonces $\langle \rho(v), u \rangle = \langle v, \rho^*(u) \rangle$, ρ^* resulta ser lineal, único, y se le llama *operador adjunto* de ρ .

Si A es la matriz asociada a ρ con respecto a una base ortonormal de V entonces la matriz asociada a ρ^* es $A^* = {}^t\bar{A}$.

Se define un *isomorfismo* $f: V \rightarrow V'$ entre espacios vectoriales con producto escalar como un isomorfismo que preserva productos escalares, es decir, tal que $\langle f(v), f(u) \rangle = \langle v, u \rangle$ para toda $v, u \in V$.

Sea $\phi: \text{Hom}_K(V, V) \rightarrow \text{Hom}_K(V, V)$ el operador dado por $\phi(\rho) = \rho^*$. Son equivalentes las afirmaciones $\rho^* = \rho^{-1}$ y $\langle \rho(v), \rho(u) \rangle = \langle v, u \rangle$ para toda v, u en V . Un *operador unitario (ortogonal)* $\rho: V \rightarrow V$ definido en un espacio vectorial con producto escalar V sobre $K = \mathbb{C}$ ($K = \mathbb{R}$) es un isomorfismo de espacios vectoriales con producto escalar $\rho: V \rightarrow V$. Entonces ρ es unitario (ortogonal) si $K = \mathbb{C}$ ($K = \mathbb{R}$) y $\rho^* = \rho^{-1}$. La matriz asociada a un operador unitario ρ es A (llamada *matriz unitaria*) si, y sólo si, $A^* = A^{-1}$. La matriz asociada a un operador ortogonal ρ es A (llamada *matriz ortogonal*) si, y sólo si, ${}^t A = A^{-1}$. Si A es una matriz ortogonal, $|A| = \pm 1$. El conjunto de matrices ortogonales de $n \times n$ posee una estructura de grupo, llamado *grupo ortogonal* y es denotado con $O(n)$. El conjunto de matrices ortogonales que poseen determinante 1 es denotado con $SO(n)$ y llamado *grupo ortogonal especial*. Finalmente, decimos que un operador $\rho: V \rightarrow V$ es *normal* si conmuta con su adjunto, es decir, si $\rho\rho^* = \rho^*\rho$. Análogamente, una matriz compleja A es *normal* si conmuta con su conjugada traspuesta, i.e. $AA^* = A^*A$. Así, los operadores ortogonales y unitarios son normales.

Un operador $\rho \in \text{Hom}_K(V, V)$ es *autoadjunto* si $\phi(\rho) = \rho^* = \rho$. Si $K = \mathbb{R}$ le llamaremos también *simétrico* y si $K = \mathbb{C}$ le llamaremos *hermitiano*. Los operadores autoadjuntos son importantes por lo siguiente: si ρ es simétrico, su polinomio característico se factoriza en factores lineales, los vectores característicos son ortogonales y posee una matriz asociada diagonal. En términos de matrices, si A es una matriz simétrica real entonces existe una ortogonal N tal que $B = N^{-1}AN = {}^t N A N$ es diagonal. En forma similar, si $\rho \in \text{Hom}_K(V, V)$ es normal entonces posee una matriz asociada diagonal. En términos de matrices, si A es normal entonces existe una matriz unitaria N tal que $B = N^{-1}AN = N^*AN$ es diagonal.

El siguiente resultado establece una descomposición de ρ llamada *espectral* de ρ : sea ρ un operador normal, entonces existen proyecciones ortogonales $p_{V_{\lambda_i}}$ de V en V_{λ_i} tales que $\rho = \sum_{i=1}^s \lambda_i p_{V_{\lambda_i}}$, $\sum p_{V_{\lambda_i}} = I$ y $p_{V_{\lambda_i}} \circ p_{V_{\lambda_j}} = 0$ si $i \neq j$.

Generalizando el concepto de función bilineal, diremos que una función $f: V_1 \times V_2 \times \cdots \times V_m \rightarrow W$ entre espacios vectoriales sobre un campo K es *multilineal* si para cada $i = 1, \dots, m$ se tiene que

$$f(v_1, \dots, v_i + v'_i, \dots, v_m) = f(v_1, \dots, v_i, \dots, v_m) + f(v_1, \dots, v'_i, \dots, v_m)$$

y

$$f(v_1, \dots, \lambda v_i, \dots, v_m) = \lambda f(v_1, \dots, v_i, \dots, v_m)$$

donde $\lambda \in K$ y $v_i, v'_i \in V_i$. Es decir, f es lineal en v_i si las demás variables se mantienen fijas.

Sea $f: V_1 \times \cdots \times V_m \longrightarrow T$ una función multilinear y $h: T \longrightarrow W$ una función lineal entre espacios vectoriales. Entonces $h \circ f$ es una función multilinear. Nos preguntamos, ¿cuándo es posible tener T y W de tal manera que toda función multilinear se obtenga de esta forma? En otras palabras, ¿cuándo podemos encontrar T y f de tal manera que dada cualquier función multilinear $g: V_1 \times \cdots \times V_m \longrightarrow W$ exista una y solamente una función lineal $h: T \longrightarrow W$ tal que $h \circ f = g$? La pregunta anterior suele conocerse con el nombre de *problema universal* para funciones multilineales. Definimos el *producto tensorial* de los espacios $\{V_i\}_{i=1}^m$ como la pareja (T, f) que resuelve dicho problema universal y denotamos T con $V_1 \otimes V_2 \otimes \cdots \otimes V_m$, o con $\otimes_{i=1}^m V_i$. La condición $g = h \circ f$ puede visualizarse en el siguiente diagrama conmutativo

$$\begin{array}{ccc} V_1 \times \cdots \times V_m & \xrightarrow{f} & T \\ & g \searrow & \downarrow h \\ & & W \end{array}$$

Es relativamente fácil comprobar la unicidad y existencia del producto tensorial que además posee las siguientes propiedades:

- (i) $V \otimes_K K \cong V \cong K \otimes_K V$
- (ii) $(U \otimes_K V) \otimes_K W \cong U \otimes_K (V \otimes W) \cong U \otimes_K V \otimes_K W$
- (iii) $U \otimes_K V \cong V \otimes_K U$.

Existen isomorfismos que relacionan el producto tensorial con el conjunto de homomorfismos (véase [LL1]).

Sea $\{V_i\}_{i=1}^m$ una familia de espacios vectoriales de dimensión finita sobre un campo K . Diremos que la sucesión

$$0 \xrightarrow{f_0} V_1 \xrightarrow{f_1} V_2 \xrightarrow{f_2} V_3 \xrightarrow{f_3} \cdots \xrightarrow{f_{m-2}} V_{m-1} \xrightarrow{f_{m-1}} V_m \xrightarrow{f_m} 0$$

es *exacta en V_i* si $\text{im } f_{i-1} = \ker f_i$ y diremos que es *exacta* si es exacta en cada V_i para $i = 1, \dots, m$.

Sucede que si se tiene una sucesión exacta como la anterior, entonces $\dim V_1 - \dim V_2 + \dim V_3 - \cdots + (-1)^{m-1} \dim V_m = 0$ y si $0 \longrightarrow V_1 \longrightarrow V_2 \longrightarrow V_3 \longrightarrow 0$ es una sucesión exacta (llamada *sucesión exacta corta*) entonces $V_2 \cong V_1 \oplus V_3$.

Consideremos una función multilinear $f: \times_{i=1}^k V_i \longrightarrow W$. Diremos que f es *alternante* si $f(v_1, \dots, v_k) = 0$ siempre que $v_i = v_j$ para algunas $i \neq j$.

Sea $V_i = V$ un espacio vectorial de dimensión finita sobre un campo K . La *potencia exterior de grado k* de V es la pareja $(\bigwedge^k V, f)$ donde $\bigwedge^k V$ es un espacio vectorial sobre K y $f: \times_{i=1}^k V_i \rightarrow \bigwedge^k V$ es una función multilineal alternante tal que para todo espacio vectorial W sobre K y para toda función multilineal alternante $g: \times_{i=1}^k V_i \rightarrow W$, existe una función lineal única $h: \bigwedge^k V \rightarrow W$ tal que $g = h \circ f$, es decir, tal que el siguiente diagrama conmuta

$$\begin{array}{ccc} \times_{i=1}^k V_i & \xrightarrow{f} & \bigwedge^k V \\ & g \searrow & \downarrow h \\ & & W \end{array}$$

Si $\{v_i\}_{i=1}^k$ son vectores de V , denotaremos a $f(v_1, \dots, v_k)$ con $v_1 \wedge \dots \wedge v_k$. También denotaremos $\bigwedge^2 V$ como $V \wedge V$. Es fácil comprobar que $u \wedge v = -v \wedge u$. Si definimos $\bigwedge^k V$ como el cociente $\otimes_{i=1}^k V_i / U$ donde U es el subespacio de $\otimes_{i=1}^k V_i$ generado por todos los elementos de la forma $v_1 \otimes \dots \otimes v_k$ con $v_i = v_j$ para algunas $i \neq j$ es claro que la existencia y unicidad de $\bigwedge^k V$ están garantizadas. La dimensión de $\bigwedge^k V$ resulta ser $\binom{n}{k}$.

Sea $A: V \rightarrow V$ un *endomorfismo* (sinónimo de operador lineal) de V . Supongamos que $\dim V = n$. Definamos la función $g = g_A: \times_{i=1}^n V_i \rightarrow \bigwedge^n V$ donde $V_i = V$ dada por $g_A(v_1, \dots, v_n) = A(v_1) \wedge \dots \wedge A(v_n)$. Como g es multilineal alternante, existe una función lineal única $h = h_A: \bigwedge^n V \rightarrow \bigwedge^n V$ tal que $h_A(v_1 \wedge \dots \wedge v_n) = A(v_1) \wedge \dots \wedge A(v_n)$. Como $\dim \bigwedge^n V = \binom{n}{n} = 1$, h_A es simplemente la multiplicación por un escalar denotado $|A|$ o $\det(A)$, i.e.

$$h_A(v_1 \wedge \dots \wedge v_n) = |A|(v_1 \wedge \dots \wedge v_n).$$

El *determinante* de $A: V \rightarrow V$ se define como el escalar $|A|$. Es fácil comprobar que si $\{v_i\}_{i=1}^n$ es una base de V y si escribimos $A(v_i) = \sum_{j=1}^n \alpha_{ij} v_j$ para $i = 1, \dots, n$ donde (α_{ij}) es la matriz de A con respecto a la base $\{v_1, \dots, v_n\}$, el determinante de A es igual a $\sum_{\sigma} \text{sig} \sigma \alpha_{1\sigma(1)} \dots \alpha_{n\sigma(n)}$ donde σ es una permutación del conjunto $\{1, 2, \dots, n\}$. Así, definimos el *determinante de una matriz* (α_{ij}) como el determinante del endomorfismo $A: V \rightarrow V$ dado por $A(v_i) = \sum_{j=1}^n \alpha_{ij} v_j$ para $i = 1, \dots, n$ y lo denotamos con $|\alpha_{ij}|$ o $\det(\alpha_{ij})$.

Podemos definir una multiplicación llamada *producto exterior* y denotarla por conveniencia con $\wedge: \bigwedge^k V \times \bigwedge^\ell V \rightarrow \bigwedge^{k+\ell} V$ mediante la regla

$$\wedge((u_1 \wedge \dots \wedge u_k), (v_1 \wedge \dots \wedge v_\ell)) = u_1 \wedge \dots \wedge u_k \wedge v_1 \wedge \dots \wedge v_\ell.$$

Este producto exterior es asociativo, distributivo y anticonmutativo. Para completar los índices se define $\bigwedge^0 V = K$ y $\bigwedge^1 V = V$. Entonces tenemos un álgebra graduada

$$\bigwedge V = (K, V, \bigwedge^2 V, \bigwedge^3 V, \dots)$$

llamada *álgebra exterior* o de *Grassman* de V .

También, si $T^k(V) = \otimes^k V = V \otimes_K \cdots \otimes_K V$, llamado *espacio tensorial de grado k de V* definimos una multiplicación

$$\begin{aligned} \cdot: T^k V \times T^\ell V &\longrightarrow T^{k+\ell} && \text{mediante} \\ ((u_1 \otimes \cdots \otimes u_k), (v_1 \otimes \cdots \otimes v_\ell)) &\longmapsto u_1 \otimes \cdots \otimes u_k \otimes v_1 \otimes \cdots \otimes v_\ell \end{aligned}$$

Así tenemos un álgebra graduada (donde $T^0 V = K$ y $T^1 V = V$)

$$\mathbf{TV} = (K, V, T^2 V, T^3 V, T^4 V, \dots)$$

llamada *álgebra tensorial* de V .

Sea V^* el espacio dual de V . Consideremos el espacio tensorial $T^k V$ de grado k de V . Consideremos también $T^\ell V^*$ y denotemos con $T_\ell^k(V)$ el producto $(\otimes^k V) \otimes (\otimes^\ell V^*)$. Es decir, $T^k V \otimes T^\ell(V^*) = T_\ell^k(V)$. Con esta notación se tiene que $T_0^k(V) = T^k(V) = \otimes^k V$, $T_\ell^0(V) = \otimes^\ell V^*$ y $T_0^0(V) = K$. Llamaremos a $T_\ell^k V$ *espacio tensorial* de tipo (k, ℓ) y cada uno de sus elementos lo llamaremos *tensor de tipo (k, ℓ)* . Un tensor de tipo $(k, 0)$ se llamará *tensor contravariante de grado k* y uno de tipo $(0, \ell)$ *tensor covariante de grado ℓ* . Un tensor de tipo $(0, 0)$ es simplemente un *escalar*. Un elemento de $T_0^1 V = V$ se llama *vector contravariante* y uno de $T_1^0 V = V^*$ se llama *vector covariante*. Si $k \neq 0$ y $\ell \neq 0$, un *tensor mixto* es un tensor de tipo (k, ℓ) .

Sea $\{v_i\}_{i=1}^n$ una base de V y $\{f^j\}_{j=1}^n$ la base de V^* , (con índices superiores). Es fácil ver que los tensores

$$v_{i_1} \otimes \cdots \otimes v_{i_k} \otimes f^{j_1} \otimes \cdots \otimes f^{j_\ell}$$

con $i_\mu, j_\eta = 1, \dots, n$; $\mu = 1, \dots, k$ y $\eta = 1, \dots, \ell$ forman una base de $T_\ell^k(V)$. Entonces cualquier tensor del tipo (k, ℓ) puede escribirse en forma única como

$$t = \sum \xi_{j_1, \dots, j_\ell}^{i_1, \dots, i_k} v_{i_1} \otimes \cdots \otimes v_{i_k} \otimes f^{j_1} \otimes \cdots \otimes f^{j_\ell}.$$

Los índices i_μ se llaman *índices contravariantes*, los j_η *índices covariantes* y $\xi_{j_1, \dots, j_\ell}^{i_1, \dots, i_k}$ se llaman *componentes* de t con respecto a la base $\{v_i\}$.

La K -Teoría Algebraica Clásica es parte del Algebra Lineal General. Intuitivamente, la K -Teoría Algebraica Clásica es una generalización del teorema que establece la existencia y unicidad de las bases para espacios vectoriales y también de la Teoría de Grupos del grupo lineal general sobre un campo K .

Definiremos un grupo denotado $K_0(X)$ asociado a un monoide conmutativo X mediante la siguiente propiedad universal: sea $g: X \rightarrow G$ un homomorfismo de monoides del monoide X en el grupo conmutativo G . Definimos el grupo $K_0(X)$ como el único grupo que cumple que, si $f: X \rightarrow K_0(X)$ es un homomorfismo de monoides entonces existe un homomorfismo de grupos único $h: K_0(X) \rightarrow G$ tal que $g = h \circ f$

$$\begin{array}{ccc} X & \xrightarrow{f} & K_0(X) \\ & g \searrow & \downarrow h \\ & & G \end{array}$$

$K_0(X)$ se llama *grupo de Grothendieck* del monoide X .

Sea K un campo y consideremos los espacios vectoriales de dimensión finita sobre K . Denotemos con $\langle V \rangle$ la clase de isomorfismo del espacio vectorial de dimensión finita V . Es inmediato verificar que el conjunto $X = \{\langle V \rangle\}$ de clases de isomorfismo es un monoide conmutativo cuya operación binaria está dada por

$$\langle V \rangle + \langle W \rangle = \langle V \oplus W \rangle.$$

Sea $g: X \rightarrow \mathbb{Z}$ dado por $g(\langle V \rangle) = \dim V$ un homomorfismo de monoides. Sea F el grupo abeliano libre con base el conjunto de clases de isomorfismo de los espacios vectoriales. Sea R el subgrupo de F generado por las expresiones de la forma $\langle V \oplus W \rangle - \langle V \rangle - \langle W \rangle$ donde $0 \rightarrow V \rightarrow V \oplus W \rightarrow W \rightarrow 0$ recorre todas las posibles sucesiones cortas para los espacios vectoriales. Sea $K_0(K) = F/R$ el grupo cociente y denotemos con $[V]$ la proyección o imagen de $[V]$ en el cociente. Entonces, siempre que se tenga una sucesión exacta corta de espacios vectoriales

$$0 \rightarrow V \rightarrow V \oplus W \rightarrow W \rightarrow 0$$

tendremos una expresión de la forma $[V \oplus W] = [V] + [W]$ en $K_0(K)$, es decir, $K_0(K)$ está generado por $\{[V] \mid V \text{ es un espacio vectorial}\}$ sujeta a las relaciones de la forma

$$[V] + [W] = [V \oplus W].$$

El homomorfismo $g: X \rightarrow \mathbb{Z}$ da lugar a un homomorfismo $h: K_0(K) \rightarrow \mathbb{Z}$ dado por $h([V]) = \dim V$ el cual es biyectivo. Es decir, para los espacios vectoriales

sobre un campo K , los cuales podemos representar por el símbolo \mathbf{EV}_K se tiene que $K_0(\mathbf{EV}_K) = K_0(K) \cong \mathbb{Z}$.

¿Qué sucede para otras estructuras cuando consideramos anillos que no necesariamente son campos? Si consideramos los \mathbb{Z} -módulos proyectivos finitamente generados \mathbf{Ab} , es decir, los grupos abelianos libres de rango finito se sabe que $K_0(\mathbb{Z}) \cong \mathbb{Z}$. Sin embargo, si consideramos los \mathbb{Z} -módulos finitos \mathbf{Abf} se sabe que $K_0(\mathbf{Abf}) \cong \mathbb{Q}^+$. Pero si consideramos los \mathbb{Z} -módulos finitamente generados \mathbf{Abfg} se tiene que $K_0(\mathbf{Abfg}) \cong \mathbb{Z}$.

Como antes, sea K un campo y denotemos con K^n el producto $K \times \cdots \times K$ n veces. El producto tensorial de dos espacios vectoriales sobre K es un espacio vectorial sobre K . Como $K^n \otimes K^m \cong K^{nm}$ los espacios vectoriales son cerrados bajo el producto tensorial. Entonces podemos dar a $K_0(K)$ una estructura de anillo mediante

$$[V] \cdot [W] = [V \otimes_K W].$$

Consideremos el conjunto de las transformaciones lineales invertibles a través de uno asociado de matrices. Sea V un espacio vectorial de dimensión n sobre un campo K . Denotemos con $GL(V)$ (o con $Aut_K(V)$) el conjunto de todas las funciones lineales de V en V que sean biyectivas (invertibles). Podemos proporcionarle a este conjunto una estructura de grupo definiendo una operación binaria

$$\circ: GL(V) \times GL(V) \longrightarrow GL(V)$$

mediante la composición

$$(f \circ g)(v) = f(g(v)).$$

Claramente $GL(V)$ es un grupo bajo \circ .

Ahora definamos otro conjunto. Denotemos con $GL_n(K)$ el conjunto de las matrices de $n \times n$ con elementos en el campo K que poseen inverso, es decir, el conjunto de todas las matrices invertibles o no singulares de $n \times n$ con elementos en K . Podemos definir en $GL_n(K)$ una operación binaria

$$\begin{aligned} \cdot: GL_n(K) \times GL_n(K) &\longrightarrow GL_n(K) \\ (A, B) &\longmapsto A \cdot B \end{aligned}$$

donde \cdot denota la multiplicación de matrices. Es fácil comprobar que $(GL_n(K), \cdot)$ es un grupo cuyo elemento de identidad es la matriz diagonal I_n . Llamaremos a $GL_n(K)$ el *grupo lineal general* de grado n sobre K .

Existe una estrecha relación entre los grupos $GL(V)$ y $GL_n(K)$, a saber, si escogemos una base fija de V , cada función lineal biyectiva de V en V posee una

matriz asociada de $n \times n$ con elementos en K la cual es no singular o invertible. Esta correspondencia establece un isomorfismo entre los grupos $GL(V)$ y $GL_n(K)$ debido a que cuando se componen dos funciones lineales, ésta composición está representada por la multiplicación de sus matrices asociadas.

Consideremos un tipo especial de matrices de $GL_n(K)$, que llamaremos *elementales* y que son aquellas que difieren de la matriz identidad I_n sólo en un elemento $\lambda \in K$ fuera de la diagonal. Dichas matrices las denotaremos con el símbolo e_{ij}^λ .

Definimos el símbolo $[A, B]$ como el producto de las matrices $ABA^{-1}B^{-1}$ y lo llamaremos *conmutador* de A y B donde $A, B \in GL_n(K)$.

Se tiene la siguiente fórmula para el conmutador de matrices elementales:

$$[e_{i,j}^\lambda, e_{k,l}^\mu] = \begin{cases} 1 & \text{si } j \neq k, i \neq l \\ e_{il}^{\lambda\mu} & \text{si } j = k, i \neq l \\ e_{kj}^{-\lambda\mu} & \text{si } j \neq k, i = l. \end{cases}$$

Denotemos con $E_n(K)$ el subgrupo de $GL_n(K)$ generado por todas las matrices elementales e_{ij}^λ , $\lambda \in K$, $1 \leq i \neq j \leq n$, llamado *grupo elemental lineal* de K .

Si cada matriz $A \in GL_n(K)$ la identificamos con la matriz

$$\begin{pmatrix} A & 0 \\ 0 & 1 \end{pmatrix} \in GL_{n+1}(K)$$

obtendremos inclusiones $GL_1(K) \subset GL_2(K) \subset GL_3(K) \subset \dots$. Sea $GL(K) = \bigcup_{n=1}^{\infty} GL_n(K)$, la unión de los grupos $GL_n(K)$ la cual llamaremos *grupo lineal general infinito* de K . Podemos concebir a $GL(K)$ como el grupo que consta de todas las matrices invertibles infinitas $A = (a_{ij})$ con $a_{ij} \in K$, $1 \leq i < \infty$, $1 \leq j < \infty$ y $a_{ij} = \delta_{ij}$, la delta de Kronecker para toda i, j excepto un número finito de i, j . Entonces $GL_n(K) \subset GL(K)$ y lo vemos como el subgrupo de todas las $(a_{ij}) \in GL(K)$ con $a_{ij} = \delta_{ij}$ para toda $i, j > n$. La inclusión de $GL_n(K)$ en $GL_{n+1}(K)$ se restringe a la inclusión de $E_n(K)$ en $E_{n+1}(K)$ y, en $GL(K)$, el subgrupo

$$E(K) = \bigcup_{n=1}^{\infty} E_n(K)$$

se llama *grupo elemental infinito* de K . Se puede probar un resultado de Whitehead:

$$[GL(K), GL(K)] = E(K).$$

Definimos el grupo cociente $GL(K)/E(K)$ como el K -grupo algebraico de índice uno del campo K denotado con $K_1(K)$. Luego

$$K_1(K) = GL(K)/[GL(K), GL(K)].$$

Obsérvese que si $f: K \rightarrow K'$ es un homomorfismo de campos se tiene un homomorfismo de grupos inducido por f

$$f_*: GL(K) \rightarrow GL(K')$$

que envía a $E(K)$ en $E(K')$, siendo así que f induce un homomorfismo de grupos $K_1(f): K_1(K) \rightarrow K_1(K')$.

Como K es conmutativo podemos considerar el determinante de una matriz como un homomorfismo de grupos $det: GL(K) \rightarrow K^*$ donde K^* denota las *unidades* de K .

Definamos $SL(K) = ker(det)$, o sea, todas las matrices de $GL(K)$ con determinante uno, y lo llamaremos *grupo especial lineal o infinito* de K . det induce un homomorfismo, que por abuso de notación, también lo denotaremos con

$$det: K_1(K) = GL(K)/E(K) \rightarrow K^*$$

el cual posee un inverso $K^* = GL_1(K) \rightarrow GL(K) \rightarrow GL(K)/E(K) = K_1(K)$.

Si definimos $SK_1(K) = SL(K)/E(K) = ker(det: K_1(K) \rightarrow K^*)$ resulta que $K_1(K) \cong SK_1(K) \oplus K^*$. Como K^* puede considerarse conocido, el cálculo de $K_1(K)$ se limita al de $SK_1(K)$. Obsérvese que $SK_1(K)$ es trivial si, y sólo si, para cualquier matriz $A \in SL_n(K)$ podemos transformar la matriz

$$\begin{pmatrix} A & 0 \\ 0 & I_k \end{pmatrix}$$

para k adecuada, en la identidad I_{n+k} mediante operaciones elementales por renglón o columna. Si $SK_1(K)$ es trivial, entonces $K_1(K) \cong K^*$. Este resulta ser el caso para el campo K y en general para cualquier dominio euclidiano. Así $K_1(K[x]) \cong K^*$ y $K_1(\mathbb{Z}) \cong \{-1, +1\}$.

Podemos generalizar todo lo anterior poniendo un anillo conmutativo Λ en lugar de un campo K . Véanse [LL1] y [LL2] para una lectura posterior.

Capítulo I

CONCEPTOS FUNDAMENTALES

I.1 ESPACIOS VECTORIALES Y FUNCIONES LINEALES

Sea K un campo.

1.1 DEFINICION. Un *espacio vectorial sobre un campo K* es un conjunto no vacío V con una operación binaria

$$\begin{aligned} +: V \times V &\longrightarrow V \\ (u, v) &\longmapsto u + v \end{aligned}$$

y una función

$$\begin{aligned} \mu: K \times V &\longrightarrow V \\ (\alpha, v) &\longmapsto \mu(\alpha, v) = \alpha v \end{aligned}$$

que cumplen los siguientes axiomas:

- (i) $u + v = v + u$
- (ii) $(u + v) + w = u + (v + w)$
- (iii) Existe $O \in V$ tal que $v + O = v$
- (iv) Para cada $v \in V$ existe un elemento, denotado con $-v$, tal que $v + (-v) = O$

- (v) $\alpha(u + v) = \alpha u + \alpha v$
- (vi) $(\alpha + \beta)v = \alpha v + \beta v$
- (vii) $(\alpha\beta)v = \alpha(\beta v)$
- (viii) Para $1 \in K$, $1v = v$; $(\alpha, \beta, 1 \in K; u, v, w \in V)$.

Los elementos u, v, w, \dots del espacio vectorial sobre K se llaman *vectores*. Los elementos del campo K se llaman *escalares* y la función μ se llama *multiplicación escalar*.

Los primeros cuatro axiomas hacen de V , junto con la operación binaria $+$, un *grupo conmutativo*, por lo que no se requiere utilizar paréntesis al sumar vectores y el orden de los sumandos carece de importancia. También, el vector O es único y el inverso $-v$ de v es único (problema 1.1). La resta $u - v$ se define como $u + (-v)$; $u, v \in V$.

Los siguientes cuatro axiomas ((v) a (viii)) se refieren a la “acción” del campo K en V . Veamos algunas propiedades que se pueden deducir de ellas.

1.2 PROPOSICION. *Sea V un espacio vectorial sobre un campo K . Entonces*

- (i) $0v = O$; $0 \in K$, $O \in V$.
- (ii) $(-\alpha)v = \alpha(-v) = -\alpha v$; $\alpha \in K$, $v \in V$.

Demostración. (i) Sabemos que $0 + 0 = 0$ en K . Luego, utilizando el axioma (vi) de 1.1 tenemos que

$$0v + 0v = (0 + 0)v = 0v.$$

Sumando $-0v$ a ambos lados obtenemos que $0v = O$.

- (ii) Como $\alpha + (-\alpha) = 0$, $\alpha \in K$ tenemos que

$$O = 0v = (\alpha + (-\alpha))v = \alpha v + (-\alpha)v.$$

Sumando $-\alpha v$ a ambos lados se tiene que $-\alpha v = (-\alpha)v$. Tomando $v + (-v) = O$ (axioma (iv)), obtenemos (problema 1.2(i))

$$O = \alpha O = \alpha(v + (-v)) = \alpha v + \alpha(-v).$$

Sumando $-\alpha v$ a ambos lados obtenemos $-\alpha v = \alpha(-v)$. Luego, $(-\alpha)v = \alpha(-v) = -\alpha v$. ■

Obsérvese que, en la demostración precedente, cuando tomamos $\alpha + (-\alpha)$, el signo $+$ se refiere a la suma del campo K y cuando consideramos $v + (-v)$, el signo $+$ se refiere a la suma del espacio vectorial V .

A continuación veamos varios ejemplos de espacios vectoriales:

1.3 EJEMPLO. Sea K un campo. Definamos en K^n la suma

$$\begin{aligned}
 & +: K^n \times K^n \longrightarrow K^n \\
 & ((\alpha_1, \dots, \alpha_n), (\beta_1, \dots, \beta_n)) \longmapsto (\alpha_1, \dots, \alpha_n) + (\beta_1, \dots, \beta_n) \quad \text{mediante} \\
 & (\alpha_1, \alpha_2, \dots, \alpha_n) + (\beta_1, \beta_2, \dots, \beta_n) = (\alpha_1 + \beta_1, \alpha_2 + \beta_2, \dots, \alpha_n + \beta_n); \quad \alpha_i, \beta_i \in K.
 \end{aligned}$$

Definamos una multiplicación escalar

$$\begin{aligned}
 & \mu: K \times K^n \longrightarrow K^n \\
 & (\alpha, (\alpha_1, \dots, \alpha_n)) \longmapsto \alpha(\alpha_1, \dots, \alpha_n) \\
 & \text{mediante} \quad \alpha(\alpha_1, \dots, \alpha_n) = (\alpha\alpha_1, \dots, \alpha\alpha_n); \quad \alpha, \alpha_i \in K.
 \end{aligned}$$

Es fácil comprobar que K^n junto con las operaciones de suma y multiplicación escalar es un espacio vectorial sobre K . Observe que este ejemplo establece que, en particular, el campo de los números reales $\mathbb{R} = \mathbb{R}^1$, así como \mathbb{R}^n , (n un entero mayor o igual que 1), son espacios vectoriales sobre \mathbb{R} . También \mathbb{C}^n es un espacio vectorial sobre \mathbb{C} y sobre \mathbb{R} . Sin embargo, \mathbb{R}^n no es un espacio vectorial sobre \mathbb{C} .

1.4 EJEMPLO. Sea K un campo. Sea $V = M_{m \times n}K$, el conjunto de matrices de $m \times n$ con elementos en K , con la suma y multiplicación por un escalar usuales. Entonces V es un espacio vectorial sobre K .

1.5 EJEMPLO. Sea V el conjunto de funciones de un conjunto no vacío S en un campo K , i.e., $V = K^S$. Si definimos la suma de dos funciones $f, g \in K^S$ como $(f + g)(s) = f(s) + g(s)$ y la multiplicación escalar mediante $(\alpha f)(s) = \alpha f(s)$, $s \in S$, es inmediato comprobar que V es un espacio vectorial sobre K .

1.6 EJEMPLO. Sea $V = K[x]$ el conjunto de todos los polinomios $\alpha_0 + \alpha_1x + \alpha_2x^2 + \dots + \alpha_nx^n$ con coeficientes en un campo K . Entonces, V es un espacio vectorial sobre K si definimos la suma y la multiplicación por un escalar de la manera usual.

Veamos como relacionar dos espacios vectoriales sobre un campo K mediante una función que preserve la estructura de espacio vectorial.

1.7 DEFINICION. Sean U y V espacios vectoriales sobre un campo K . Una función $f: U \rightarrow V$ se llama *lineal* o también *homomorfismo de espacios vectoriales* si

- (i) $f(u + v) = f(u) + f(v)$ y
- (ii) $f(\alpha v) = \alpha f(v)$; $u, v \in U$; $\alpha \in K$.

Obsérvese que el $+$ de $u + v$ se refiere a la suma de U y que el $+$ de $f(u) + f(v)$ se refiere a la suma de V . Lo mismo que αv denota la multiplicación escalar de U y $\alpha f(v)$ la de V .

Si en (ii) tomamos $\alpha = 0 \in K$, tenemos que $f(0v) = f(O) = 0f(v) = O$, luego $f(O) = O$, i.e., todo homomorfismo de espacios vectoriales (o función lineal) envía el vector cero del dominio en el vector cero del codominio.

Es obvio que las condiciones (i) y (ii) de la definición 1.7 son equivalentes a la siguiente:

$$f(\alpha u + \beta v) = \alpha f(u) + \beta f(v); \quad \alpha, \beta \in K; u, v \in U.$$

También se suele llamar a una función lineal f , *aplicación lineal* o *transformación lineal*. Utilizaremos cualquiera de estas denominaciones.

Nota. Por abuso de notación se acostumbra escribir 0 en lugar de O .

1.8 EJEMPLO. Sea $U = \mathbb{R}^3$ y $V = \mathbb{R}$ con la suma y multiplicación escalar usuales. Definamos $f: U \rightarrow V$ mediante la regla $f(x, y, z) = 3x - 2y + 2z$. Veamos que f es lineal. Como

$$\begin{aligned} f((x_1, y_1, z_1) + (x_2, y_2, z_2)) &= f(x_1 + x_2, y_1 + y_2, z_1 + z_2) \\ &= 3(x_1 + x_2) - (y_1 + y_2) + 2(z_1 + z_2) \end{aligned}$$

$$\text{y} \quad f(x_1, y_1, z_1) + f(x_2, y_2, z_2) = (3x_1 - 2y_1 + 2z_1) + (3x_2 - 2y_2 + 2z_2),$$

claramente se cumple la condición (i) de 1.7. También, $f(\alpha(x, y, z)) = f(\alpha x, \alpha y, \alpha z) = 3\alpha x - 2\alpha y + 2\alpha z = \alpha(3x - 2y + 2z) = \alpha f(x, y, z)$, por lo que se cumple (ii) de 1.7.

1.9 EJEMPLO. Sea $U = V = \mathbb{R}^2$. Definamos $f: U \rightarrow V$ mediante $f(x, y) = (x + 2, y + 3)$. Como $f(0, 0) = (2, 3) \neq (0, 0)$, f no es lineal pues todo homomorfismo de espacios vectoriales envía el vector cero del dominio en el vector cero del codominio.

1.10 PROPOSICION. *La composición de dos homomorfismos de espacios vectoriales sobre un campo K es un homomorfismo de espacios vectoriales sobre K .*

Demostración. Sean $f:U \rightarrow V$ y $g:V \rightarrow W$ funciones lineales. Luego

$$\begin{aligned}(g \circ f)(u + v) &= g(f(u + v)) \\ &= g(f(u) + f(v)) \\ &= g(f(u)) + g(f(v)) \\ &= (g \circ f)(u) + (g \circ f)(v)\end{aligned}$$

Además, $(g \circ f)(\alpha u) = g(f(\alpha u)) = g(\alpha f(u)) = \alpha g(f(u)) = \alpha(g \circ f)(u)$. Por lo tanto $(g \circ f)$ es una función lineal. ■

1.11 DEFINICION. Sea $f:U \rightarrow V$ un homomorfismo (o función lineal o aplicación lineal) de espacios vectoriales sobre un campo K . Diremos que f es un *isomorfismo*, y escribiremos $f:U \xrightarrow{\cong} V$, si existe un homomorfismo $g:V \rightarrow U$ tal que $g \circ f = 1_U$ y $f \circ g = 1_V$.

Es fácil comprobar (problema 1.9) que, si g existe, está determinada en forma única; la denotaremos con f^{-1} y se llama *inverso* de f . Así, $f:U \rightarrow V$ es isomorfismo si, y sólo si, es biyectiva. Diremos que dos espacios U y V sobre un campo K son *isomorfos* si existe un isomorfismo $f:U \xrightarrow{\cong} V$ y escribiremos $U \cong V$.

PROBLEMAS

1.1 Sea V un espacio vectorial sobre un campo K . Pruebe que el vector $O \in V$ es único y que el inverso de $v \in V$ es también único.

1.2 Pruebe que

- (i) $\alpha O = O$; $\alpha \in K$, $O \in V$.
- (ii) Si $\alpha v = O$ entonces, $\alpha = 0$ o $v = O$; $\alpha \in K$, $v \in V$.
- (iii) $(-1)v = -v$; $v \in V$.

1.3 Proporcione con todo detalle el hecho de que V es un espacio vectorial en los ejemplos 1.3, 1.4, 1.5 y 1.6.

1.4 Sea $U = V = K^n$. Pruebe que $f:U \rightarrow V$ dada por $f(u_1, \dots, u_n) = (u_1, u_2, \dots, u_{n-1}, 0)$ es lineal.

1.5 Sea V un espacio vectorial sobre un campo K . Pruebe que la función $1_V:V \rightarrow V$ y la función $O_V:V \rightarrow V$ dadas por $1_V(v) = v$ y $O_V(v) = O$

$\forall v \in V$, son lineales. 1_V se llama *homomorfismo identidad* de V y O_V se llama *homomorfismo trivial*.

1.6 Compruebe cuales aplicaciones son lineales y cuales no lo son:

- (i) $f: K^n \rightarrow K^m$, $f(v) = Av$ donde A es una matriz de $m \times n$ con elementos en el campo K .
- (ii) $f: K^2 \rightarrow K^2$, $f(x, y) = (4y, 0)$
- (iii) $f: K^3 \rightarrow K^3$, $f(x, y, z) = (-z, x, y)$
- (iv) $f: K^2 \rightarrow K^2$, $f(x, y) = (x^2, 2y)$
- (v) $f: K^5 \rightarrow K^4$, $f(u, v, x, y, z) = (2uy, 3xz, 0, 4u)$
- (vi) $f: K^3 \rightarrow K^3$, $f(x, y, z) = (x + 2, y + 2, z + 2)$

1.7 Establezca que, (i) si $V = R[x]$ es el espacio vectorial de los polinomios en x sobre \mathbb{R} , entonces la diferencial $D: V \rightarrow V$ dada por $D(f) = df/dx$ y la integral $I: V \rightarrow \mathbb{R}$ dada por $I(f) = \int_0^1 f(x)dx$ son lineales.

(ii) la *traza* $tr: M_n(K) \rightarrow K$ de una matriz cuadrada (la suma de los elementos de su diagonal) es una función lineal y que el *determinante* $det: M_n(K) \rightarrow K$ no es una función lineal.

1.8 Sea K un campo. Denotemos con $Hom_K(U, V)$ el conjunto de homomorfismos o funciones lineales del espacio vectorial U sobre K en el espacio vectorial V sobre K . Defina $f + g: U \rightarrow V$ mediante $(f + g)(u) = f(u) + g(u)$, $u \in U$ y $\alpha f: U \rightarrow V$ mediante $(\alpha f)(u) = \alpha(f(u))$, $\alpha \in K, u \in U$. Pruebe que $Hom_K(U, V)$ es un espacio vectorial sobre K con las operaciones definidas. A menudo, también se utilizan las notaciones $L(U, V)$ y $A(U, V)$ en lugar de $Hom_K(U, V)$.

1.9 Pruebe que si $f: U \rightarrow V$ es como en 1.11, g está determinada en forma única y que f es isomorfismo si, y sólo si es biyectiva.

1.10 Sea $f: U \rightarrow V$ una aplicación lineal biyectiva de espacios vectoriales sobre un campo K . Pruebe que la función inversa $f^{-1}: V \rightarrow U$ es también lineal.

1.11 Sea K un campo y V un espacio vectorial sobre K . Considere K como un espacio vectorial sobre sí mismo. Pruebe que dado un vector $v \in V$, existe una función lineal única $h: K \rightarrow V$ tal que $h(1) = v$. (Esta función está dada por $h(\alpha) = \alpha v$.)

I.2 SUBESPACIOS VECTORIALES

Siempre que consideramos un objeto matemático nos preguntamos por sus subobjetos. Es natural definir un subespacio vectorial de un espacio vectorial sobre un campo K como un subconjunto que es a su vez un espacio vectorial sobre K bajo las mismas operaciones. Sin embargo para nuestra conveniencia lo definiremos de otra forma equivalente (problema 2.1).

2.1 DEFINICION. Un subconjunto U de un espacio vectorial V sobre un campo K se llama *subespacio vectorial* de V si

- (i) el vector O de V pertenece a U ,
- (ii) si $v, w \in U$ entonces $v + w \in U$ y
- (iii) si $\alpha \in K$ y $v \in U$ entonces $\alpha v \in U$.

2.2 EJEMPLO. El conjunto U de vectores de la forma $(u_1, \dots, u_{n-1}, 0)$ con u_i en el campo K forman un subespacio del espacio vectorial K^n sobre K .

2.3 EJEMPLO. Sea V un espacio vectorial sobre un campo K . Los conjuntos V y $\{O\}$ son subespacios de V , llamado este último *subespacio trivial* y, por abuso de notación, se acostumbra escribirlo simplemente como 0 .

2.4 EJEMPLO. Sea $V = M_n K$ el espacio vectorial de las *matrices* de $n \times n$ o *cuadradas*. Sea U el subconjunto de $M_n K$ que consiste de las matrices que cumplan que $a_{ij} = a_{ji}$, llamadas *simétricas*. Entonces U es un subespacio de $M_n K$.

2.5 DEFINICION. Sea $f: U \rightarrow V$ un homomorfismo (función lineal) de espacios vectoriales sobre un campo K . El *núcleo* de f , denotado $\ker f$, es el conjunto de todos los elementos $u \in U$ tales que $f(u) = 0$. La *imagen* de f , denotada $\text{im } f$, es el conjunto de $f(u)$ con $u \in U$.

2.6 PROPOSICION. Sea $f: U \rightarrow V$ un homomorfismo (función lineal) de espacios vectoriales sobre un campo K . Entonces, si U' es un subespacio de U , $f(U')$ es un subespacio de V y, si V' es un subespacio de V , $f^{-1}(V')$ es un subespacio de U .

Demostración. Veamos que $f(U') = \{f(u) | u \in U'\}$ es un subespacio de V . Sean $v, w \in f(U')$, luego, existen $u, u' \in U'$ tales que $f(u) = v$, $f(u') = w$. Como U' es subespacio de U , $u + u' \in U'$ y $\alpha u \in U'$. Como f es lineal,

$$\begin{aligned} f(O) &= O \in f(U'), \\ v + w &= f(u) + f(u') = f(u + u') \in f(U'), \\ \alpha v &= \alpha f(u) = f(\alpha u) \in f(U'). \end{aligned}$$

Por lo tanto, $f(U')$ es un subespacio de V .

Veamos que $f^{-1}(V') = \{u \in U | f(u) \in V'\}$ es un subespacio de U . Sean $u, u' \in f^{-1}(V')$, entonces $f(u)$ y $f(u')$ están en V' . Como V' es un subespacio de V y f es lineal,

$$\begin{aligned} f(O) &= O \in V' \\ f(u + u') &= f(u) + f(u') \in V' \\ f(\alpha u) &= \alpha f(u) \in V', \quad \alpha \in K. \end{aligned}$$

Luego, $f^{-1}(V')$ es un subespacio de U . ■

2.7 COROLARIO. Sea $f: U \rightarrow V$ lineal. Entonces $\text{im } f$ es un subespacio de V y $\ker f$ es un subespacio de U .

Demostración. Inmediata de 2.6 tomando $U' = U$ y $V' = 0$. ■

2.8 PROPOSICION. Sea $\{V_i\}_{i \in I}$ una familia de subespacios de un espacio vectorial V sobre un campo K indizada por un conjunto I . Entonces $\bigcap_{i \in I} V_i$ es un subespacio de V .

Demostración. Sea $\alpha \in K$; $u, v \in \bigcap_{i \in I} V_i$. Como $\bigcap_{i \in I} V_i \subset V_i$ para cualquier $i \in I$, tenemos que $u, v \in V_i$. Como V_i es subespacio de V , $O \in V_i$, $u + v \in V_i$ y $\alpha u \in V_i$ para toda $i \in I$. Por lo tanto $O \in \bigcap_{i \in I} V_i$, $u + v \in \bigcap_{i \in I} V_i$ y $\alpha u \in \bigcap_{i \in I} V_i$. ■

2.9 DEFINICION. Sean U y V subespacios vectoriales de un espacio vectorial W sobre un campo K . La *suma* de U y V , denotada $U + V$, es el conjunto de todas las sumas $u + v$ donde $u \in U$, $v \in V$.

Es inmediato comprobar que $U + V$ es un subespacio de W , (problema 2.5).

2.10 DEFINICION. Sean U y V subespacios del espacio vectorial W sobre un campo K . Diremos que W es la *suma directa interna* de U y V si cada elemento $w \in W$ puede escribirse de una y solamente una manera como $w = u + v$; $u \in U$, $v \in V$. En tal caso escribiremos $W = U \oplus V$.

2.11 EJEMPLO. Sea $W = \mathbb{R}^4$, $U = \{(x, y, z, 0) \in \mathbb{R}^4 \mid x, y, z, 0 \in \mathbb{R}\}$ y $V = \{(0, 0, z, t) \in \mathbb{R}^4 \mid z, t, 0 \in \mathbb{R}\}$. Entonces $\mathbb{R}^4 = U + V$ pues cualquier vector en \mathbb{R}^4 es suma de vectores de U y de vectores de V . \mathbb{R}^4 no es suma directa de U y V pues no son únicas las expresiones para $w \in W$. Por ejemplo, el vector $(4, 2, 3, 5)$ se puede escribir de varias maneras, como $(4, 2, 1, 0) + (0, 0, 2, 5)$, o como $(4, 2, 2, 0) + (0, 0, 1, 5)$. Sin embargo si tomamos $W = \mathbb{R}^4$, $U = \{(x, y, 0, 0) \in \mathbb{R}^4 \mid x, y, 0 \in \mathbb{R}\}$ y $V = \{(0, 0, z, t) \in \mathbb{R}^4 \mid z, t, 0 \in \mathbb{R}\}$ es claro que cualquier vector (x, y, z, t) puede escribirse como suma de un vector en U y otro en V en una, y solamente una forma:

$$(x, y, z, t) = (x, y, 0, 0) + (0, 0, z, t).$$

Entonces $\mathbb{R}^4 = U \oplus V$.

Podemos definir la *suma directa externa* de espacios vectoriales $\{V_i\}, i = 1, \dots, n$, sobre un campo K y la denotamos con $\bigoplus_{i=1}^n V_i$ como sigue: los elementos de $\bigoplus_{i=1}^n V_i$ son listas de la forma (v_1, \dots, v_n) ; dos elementos de $\bigoplus V_i$ son iguales, si son iguales coordenada a coordenada; su suma y multiplicación escalar están dadas mediante

$$\begin{aligned} (v_1, \dots, v_n) + (v'_1, \dots, v'_n) &= (v_1 + v'_1, \dots, v_n + v'_n) && \text{y} \\ \alpha(v_1, \dots, v_n) &= (\alpha v_1, \dots, \alpha v_n), \quad \alpha \in K. \end{aligned}$$

Es inmediato comprobar que $\bigoplus_{i=1}^n V_i$ con las operaciones de suma y multiplicación escalar es un espacio vectorial sobre un campo K , y que, si V es la suma directa interna de V_1, \dots, V_n , entonces V es isomorfo a la suma directa externa de V_1, \dots, V_n (véase el problema 2.10). En vista de esto hablaremos de *la* suma directa.

2.12 TEOREMA. $W = U \oplus V$ si, y sólo si, $W = U + V$ y $U \cap V = \{0\}$.

Demostración. Supongamos que $W = U \oplus V$, esto es, $w \in W$ se escribe de manera única como $w = u + v$; $u \in U$, $v \in V$. Luego $W = U + V$. Supongamos

que $w \in U \cap V$. Entonces podemos escribir w como $w = w + 0$ con $w \in U$, $0 \in V$ y también $w = 0 + w$ con $0 \in U$, $w \in V$. Como cada expresión para w es única por hipótesis, $w = 0$. Luego $U \cap V = \{0\}$.

Ahora, supongamos que $W = U + V$ con $U \cap V = \{0\}$. Sea $w \in W$. Como $w = u + v$, lo único que debemos probar es que dicha expresión para w es única. Supongamos que existe otra expresión para w de la forma $w = u' + v'$. Entonces $u + v = u' + v'$. Luego $u - u' = v' - v$. Pero $u - u' \in U$ y $v' - v \in V$ y como $U \cap V = \{0\}$, $u - u' = 0$ y $v' - v = 0$. Luego $u = u'$ y $v = v'$. Por lo tanto, w se expresa en forma única y $W = U \oplus V$. ■

2.13 COROLARIO. Sean $f: U \rightarrow V$ y $g: V \rightarrow W$ funciones lineales entre espacios vectoriales sobre un campo K tales que $g \circ f$ es isomorfismo. Entonces $V \cong \text{im } f \oplus \ker g$.

Demostración. Veamos que $\text{im } f + \ker g = V$. Sea $v \in V$ y $g(v) \in W$. Como $gf: U \rightarrow W$ es un isomorfismo, existe $u \in U$ tal que $gf(u) = g(v)$. Sea $v' = f(u) \in \text{im } f$ y $v'' = v - v'$. Entonces $g(v'') = g(v - v') = g(v) - g(v') = gf(u) - g(f(u)) = 0$. Luego $v'' \in \ker g$ y, por lo tanto, $v' + v'' \in \text{im } f + \ker g$ pues v era arbitraria.

Veamos que $\text{im } f \cap \ker g = \{0\}$. Sea $v \in \text{im } f \cap \ker g$. Entonces, como $v \in \text{im } f$, existe $u \in U$ tal que $f(u) = v$. Como $v \in \ker g$, $g(v) = 0$. Luego $gf(u) = g(v) = 0$. Como gf es un isomorfismo, $u = 0$. Luego $f(u) = 0$ y, por lo tanto, $v = 0$. Por 2.12, $V \cong \text{im } f \oplus \ker g$. ■

A continuación estableceremos una propiedad, llamada *universal*, de la suma directa.

2.14 TEOREMA. Sea V un espacio vectorial sobre un campo K , $\varphi_i: V_i \rightarrow V$, $i = 1, 2$ funciones lineales de espacios vectoriales e $\iota_i: V_i \rightarrow V_1 \oplus V_2$, $i = 1, 2$ las inclusiones naturales. Entonces existe una función lineal única $\varphi: V_1 \oplus V_2 \rightarrow V$ tal que $\varphi \circ \iota_i = \varphi_i$, $i = 1, 2$.

Demostración. La afirmación del enunciado puede representarse en el siguiente diagrama:

$$\begin{array}{ccccc}
 & & V & & \\
 & \nearrow \varphi_1 & \uparrow \varphi & \nwarrow \varphi_2 & \\
 V_1 & \xrightarrow{\iota_1} & V_1 \oplus V_2 & \xleftarrow{\iota_2} & V_2
 \end{array}$$

Definamos $\varphi(v_1, v_2) = \varphi_1(v_1) + \varphi_2(v_2)$. Es fácil comprobar que $\varphi: V_1 \oplus V_2 \rightarrow V$ es la única función lineal tal que el diagrama anterior *conmuta*, i.e., $\varphi \circ \iota_i = \varphi_i$, $i = 1, 2$. Problema 2.9 ■

El teorema precedente caracteriza a la suma directa y se puede generalizar fácilmente a n sumandos con solamente considerar $i = 1, 2, \dots, n$. El diagrama correspondiente es

$$\begin{array}{ccc} & & V \\ & \nearrow \varphi_j & \uparrow \varphi \\ V_j & \xrightarrow{\iota_j} & \bigoplus_{i=1}^n V_i \end{array}$$

donde ι_j denota la inclusión natural de V_j en $\bigoplus_{i=1}^n V_i$.

2.15 DEFINICION. Decimos que un vector v de un espacio vectorial V sobre un campo K es una *combinación lineal* de elementos de un subconjunto S de V si existe un número finito de elementos $\{v_i\}_{i=1}^n$ de S tal que $v = \alpha_1 v_1 + \dots + \alpha_n v_n$, $\alpha_i \in K$. Las α_i se llaman *coeficientes*.

Para simplificar la notación, y cuando no haya posibilidad de confusión, quitaremos los límites del conjunto. Por ejemplo escribiremos $\{v_j\}$ en lugar de $\{v_j\}_{j=1}^n$.

2.16 TEOREMA. *El conjunto de todas las combinaciones lineales $\langle S \rangle$ de un subconjunto no vacío S del espacio vectorial V sobre un campo K es un subespacio de V que contiene a S y es el subespacio más pequeño de V que contiene a S .*

Demostración. Sea $v \in S$, como $v = 1v$ entonces $v \in \langle S \rangle$ y es inmediato comprobar que $0 \in \langle S \rangle$. Si $u, v \in \langle S \rangle$ entonces $u = \alpha_1 u_1 + \dots + \alpha_n u_n$ y $v = \beta_1 v_1 + \dots + \beta_m v_m$; $\alpha_i, \beta_j \in K$; $u_i, v_j \in S$. Entonces $u + v = \alpha_1 u_1 + \dots + \alpha_n u_n + \beta_1 v_1 + \dots + \beta_m v_m$ y $\alpha u = \alpha(\alpha_1 u_1 + \dots + \alpha_n u_n) = \alpha \alpha_1 u_1 + \dots + \alpha \alpha_n u_n$. Luego $u + v$ y αu pertenece a $\langle S \rangle$. Así, $\langle S \rangle$ es un subespacio de V .

Supongamos que U es un subespacio de V que contiene a S y supongamos que $u_1, \dots, u_n \in S \subset U$. Entonces $\alpha_1 u_1, \dots, \alpha_n u_n \in U$ con $\alpha_i \in K$. Esto significa que U contiene a todas las combinaciones lineales de S , i.e., U contiene a $\langle S \rangle$. ■

2.17 DEFINICION. El subespacio más pequeño de un espacio vectorial V

sobre un campo K que contiene a un subconjunto S de V se llama *subespacio generado por S* .

Por el teorema 2.16, $\langle S \rangle$ es el subespacio generado por un subconjunto S de V . Además, observe que como es el subespacio más pequeño de V que contiene a S , $\langle S \rangle$ es igual a la intersección de todos los subespacios que contienen a S . Si $\langle S \rangle = V$, todo elemento de V es una combinación lineal de elementos de S . En este caso, diremos que V está *generado* por el subconjunto S de V .

2.18 EJEMPLO. Sea $S = \{(1, 0, 0, 0), (0, 1, 0, 0), (0, 0, 1, 0), (0, 0, 0, 1)\}$ un subconjunto de \mathbb{R}^4 . Considere las combinaciones lineales de elementos de S , i.e., expresiones de la forma

$$\alpha_1(1, 0, 0, 0) + \alpha_2(0, 1, 0, 0) + \alpha_3(0, 0, 1, 0) + \alpha_4(0, 0, 0, 1).$$

Es claro que cualquier vector de \mathbb{R}^4 puede escribirse como combinación lineal de vectores de S ; luego $\langle S \rangle = \mathbb{R}^4$.

2.19 EJEMPLO. Sea $S = \{u_1, u_2\}$ donde $u_1 = (2, 3, 4)$ y $u_2 = (1, 6, 7)$ son vectores en $V = \mathbb{R}^3$. Entonces $\langle S \rangle$ es el plano dado por la ecuación $(x, y, z) = \alpha_1(2, 3, 4) + \alpha_2(1, 6, 7)$. Es decir, cada punto $(x, y, z) \in \langle S \rangle$ es tal que $x = 2\alpha_1 + \alpha_2$, $y = 3\alpha_1 + 6\alpha_2$ y $z = 4\alpha_1 + 7\alpha_2$.

PROBLEMAS

2.1 Pruebe que el subconjunto U del espacio vectorial V es un subespacio de V si, y sólo si, U es un espacio vectorial sobre K con respecto a las mismas operaciones de V .

2.2 Muestre con todo detalle el hecho de que U sea un subespacio de V en los ejemplos 2.2 y 2.4.

2.3 Pruebe, sin utilizar la proposición 2.6, la afirmación de 2.7.

2.4 Pruebe que el *conjunto solución* (i.e., el conjunto de todas las soluciones) X de un sistema de m ecuaciones lineales homogéneo con n incógnitas y coeficientes en un campo K

$$\begin{array}{rccccccc} \alpha_{11}x_1 & + & \cdots & + & \alpha_{1n}x_n & = & 0 \\ \alpha_{21}x_1 & + & \cdots & + & \alpha_{2n}x_n & = & 0 \\ \vdots & & & & \vdots & & \\ \alpha_{m1}x_1 & + & \cdots & + & \alpha_{mn}x_n & = & 0 \end{array}$$

es un subespacio del espacio vectorial K^n sobre K llamado *subespacio solución*. Observe que el conjunto solución de un sistema no homogéneo de ecuaciones lineales con n incógnitas no es un subespacio de K^n .

2.5 Pruebe que $U + V$ es un subespacio de W donde U, V y W son los definidos en 2.9.

2.6 Geométricamente, explique qué significan en \mathbb{R}^3 las combinaciones lineales de uno, dos y tres vectores en todas sus posibles elecciones.

2.7 Verifique si el vector $v = (2, 8, 1) \in \mathbb{R}^3$ es una combinación lineal de los vectores

- (i) $v_1 = (3, 8, 0)$, $v_2 = (2, 1, 4)$ y $v_3 = (0, 2, 8)$ y
- (ii) $v_1 = (0, 4, 6)$, $v_2 = (3, 6, 8)$ y $v_3 = (2, 3, 1)$.

2.8 Demuestre que una función lineal $f: U \rightarrow V$ entre espacios vectoriales sobre un campo K es inyectiva si, y sólo si, $\ker f = \{0\}$.

2.9 Pruebe que φ en 2.14 es lineal y única.

2.10 (a) Compruebe que la suma directa externa $\bigoplus_{i=1}^n V_i$ de los espacios vectoriales V_i , $i = 1, \dots, n$ sobre K es un espacio vectorial sobre un campo K .

(b) Establezca el isomorfismo entre la suma directa interna y externa de la familia $\{V_i\}_{i=1}^n$.

2.11 Sean v_1, \dots, v_n vectores de un espacio vectorial V sobre un campo K . Se dice que estos vectores son *linealmente dependientes* si existen escalares $\alpha_1, \dots, \alpha_n \in K$, no todos iguales a cero, tales que

$$\alpha_1 v_1 + \cdots + \alpha_n v_n = \sum_{i=1}^n \alpha_i v_i = 0.$$

También se dice que un conjunto de vectores es *linealmente independiente* si no es linealmente dependiente.

Pruebe que un conjunto finito de vectores de V es linealmente dependiente si, y sólo si, algún vector del conjunto es una combinación lineal de los restantes.

I.3 ESPACIOS VECTORIALES DE DIMENSION FINITA

Iniciaremos esta sección estudiando, desde otro punto de vista, los conceptos de dependencia e independencia lineal definidos en el problema 2.11. Veremos que ambos puntos de vista coinciden.

Consideremos la suma directa $K^n = \bigoplus_{j=1}^n K_j$ con cada K_j igual a K considerado como espacio vectorial sobre sí mismo. Sea $\iota_i: K_i \rightarrow \bigoplus_{j=1}^n K_j$ la inclusión natural dada por $\iota_i(\alpha) = (0, \dots, \alpha, \dots, 0)$, (α en el lugar i). Por el problema 1.11 y como ι_i es lineal, la inclusión queda determinada por su valor en 1, $\iota_i(1) = (0, \dots, 1, \dots, 0) = e_i$. Observe que cada $u \in \bigoplus_{j=1}^n K_j$ puede escribirse en forma única como $u = \alpha_1 e_1 + \alpha_2 e_2 + \dots + \alpha_n e_n$ con $\alpha_j \in K_j$. Denotaremos con g la función

$$g: \{1, \dots, n\} \rightarrow \bigoplus_{j=1}^n K_j \\ i \mapsto e_i$$

dada por $g(i) = e_i$. (g es simplemente una función.)

3.1 PROPOSICION. *Para todo espacio vectorial V sobre un campo K y para toda función $f: \{1, 2, \dots, n\} \rightarrow V$ existe una función lineal única $\phi: \bigoplus_{j=1}^n K_j \rightarrow V$ tal que $f = \phi \circ g$.*

$$\begin{array}{ccc} & V & \\ & \uparrow \phi & \swarrow f \\ \bigoplus_{j=1}^n K_j & \xleftarrow{g} & \{1, \dots, n\} \end{array}$$

Demostración. Sea $u = \alpha_1 e_1 + \dots + \alpha_n e_n \in \bigoplus_{j=1}^n K_j$ y sean $v_1 = f(1), \dots, v_n = f(n)$. Como la expresión de u es única podemos definir una función ϕ mediante la fórmula $\phi(\alpha_1 e_1 + \dots + \alpha_n e_n) = \alpha_1 v_1 + \dots + \alpha_n v_n$. Es inmediato comprobar que ϕ es lineal y que $\phi(e_i) = f(i)$, es decir, $\phi g(i) = f(i)$, o sea, $\phi \circ g = f$. ■

3.2 DEFINICION. Diremos que el conjunto $\{v_j\}$, $j \in \{1, 2, \dots, n\}$, de elementos de un espacio vectorial V sobre un campo K es

- (i) *linealmente independiente* si ϕ es inyectiva
- (ii) un conjunto de *generadores* de V si ϕ es suprayectiva

(iii) una *base* de V si ϕ es biyectiva.

En otras palabras, el conjunto $\{v_j\}$, $j \in \{1, 2, \dots, n\}$ es linealmente independiente si $\phi(\sum_{j=1}^n \alpha_j e_j) = \sum_{j=1}^n \alpha_j v_j = 0$ implica que $\alpha_j = 0$ para toda j en $\{1, 2, \dots, n\}$, $\alpha_j \in K_j$.

Diremos que el conjunto $\{v_j\}$, $j \in \{1, 2, \dots, n\}$, de elementos de un espacio vectorial V sobre un campo K es *linealmente dependiente* si dicho conjunto no es linealmente independiente. Es decir, $\{v_j\}$ es linealmente dependiente si existen escalares $\alpha_i \in K$ no todos cero tales que

$$\alpha_1 v_1 + \alpha_2 v_2 + \dots + \alpha_n v_n = 0.$$

Esta última expresión es válida para $\alpha_j = 0$, $\forall j \in \{1, 2, \dots, n\}$ y si ésta última expresión es válida *únicamente* para $\alpha_j = 0$, $j \in \{1, 2, \dots, n\}$ entonces el conjunto $\{v_j\}$ es linealmente independiente. En otras palabras, el conjunto $\{v_j\}$ es linealmente independiente si, y sólo si, toda combinación lineal no trivial de vectores del conjunto $\{v_j\}$ es diferente del vector 0.

Decir que ϕ en 3.2 es suprayectiva equivale a decir que todo elemento de V puede escribirse como $\sum_{j=1}^n \alpha_j v_j$, i.e., como una combinación lineal. El que ϕ sea biyectiva quiere decir que todo elemento $v \in V$ puede escribirse de una y solamente una manera en la forma $v = \sum_{j=1}^n \alpha_j v_j$, $\forall j \in \{1, 2, \dots, n\}$.

Es claro que el conjunto $\{e_j\}$, $j \in \{1, \dots, n\}$, es una base de $\oplus_{j=1}^n K_j$ (llamada *canónica*). Frecuentemente se identifica el conjunto $\{1, 2, \dots, n\}$ con el conjunto de los e_j mediante la biyección dada por $j \mapsto e_j$.

3.3 EJEMPLO. Los vectores del espacio vectorial \mathbb{R}^4 sobre \mathbb{R} , $v_1 = (2, 3, 1, 4)$, $v_2 = (3, 2, 1, 0)$ y $v_3 = (17, 18, 7, 16)$, son linealmente dependientes puesto que $4(2, 3, 1, 4) + 3(3, 2, 1, 0) - (17, 18, 7, 16) = (0, 0, 0, 0)$.

3.4 EJEMPLO. Sean $v_1 = (5, 4, 7)$, $v_2 = (0, 3, 1)$ y $v_3 = (0, 0, 2)$ vectores del espacio vectorial \mathbb{R}^3 sobre \mathbb{R} . Sea $\alpha_1 v_1 + \alpha_2 v_2 + \alpha_3 v_3 = 0$ una combinación lineal igual a cero. Entonces tenemos un sistema de ecuaciones lineales

$$\begin{aligned} 5\alpha_1 + 0\alpha_2 + 0\alpha_3 &= 0 \\ 4\alpha_1 + 3\alpha_2 + 0\alpha_3 &= 0 \\ 7\alpha_1 + 1\alpha_2 + 2\alpha_3 &= 0. \end{aligned}$$

De la primera ecuación, tenemos que $\alpha_1 = 0$. De la segunda ecuación con $\alpha_1 = 0$ tenemos que $\alpha_2 = 0$, y de la tercera ecuación con $\alpha_1 = \alpha_2 = 0$ tenemos que $\alpha_3 = 0$. Luego $\alpha_1 = \alpha_2 = \alpha_3 = 0$ y los vectores v_1, v_2 y v_3 son linealmente independientes.

3.5 PROPOSICION. *El conjunto de vectores diferentes de cero $\{v_i\}_{i=1}^n$ es linealmente dependiente si, y sólo si, uno de ellos es combinación lineal de los vectores precedentes.*

Demostración. Supongamos que son linealmente dependientes; entonces $\alpha_1 v_1 + \cdots + \alpha_n v_n = 0$ con alguna $\alpha_i \neq 0$. Sea j el mayor entero tal que $\alpha_j \neq 0$. Entonces $\alpha_1 v_1 + \cdots + \alpha_j v_j + 0v_{j+1} + \cdots + 0v_n = 0$, i.e., $\alpha_1 v_1 + \cdots + \alpha_j v_j = 0$. Si $j = 1$ entonces $\alpha_1 v_1 = 0$ con $\alpha_1 \neq 0$, luego $v_1 = 0$. Si $j > 1$, como los vectores v_j son diferentes de cero y

$$v_j = -\alpha_j^{-1} \alpha_1 v_1 - \cdots - \alpha_j^{-1} \alpha_{j-1} v_{j-1},$$

v_j es combinación lineal de los vectores precedentes.

Supongamos ahora que $v_j = \alpha_1 v_1 + \cdots + \alpha_{j-1} v_{j-1}$. Entonces podemos reescribir esto como

$$\alpha_1 v_1 + \cdots + \alpha_{j-1} v_{j-1} - v_j + 0v_{j+1} + \cdots + 0v_n = 0$$

con $\alpha_j \neq 0$. Luego, $\{v_i\}_{i=1}^n$ es linealmente dependiente. ■

3.6 OBSERVACION. Es inmediato de la definición 3.2 que si V es un espacio vectorial sobre un campo K con base $\{v_1, \dots, v_n\}$ entonces es isomorfo a K^n .

3.7 TEOREMA. *Sea $X = \{u_i\}_{i=1}^n$ un conjunto de generadores de un espacio vectorial V sobre un campo K .*

- (i) *Si u_j es combinación lineal de los vectores $\{u_i\}_{i=1}^{j-1}$ entonces el conjunto $\{u_1, \dots, u_{j-1}, u_{j+1}, \dots, u_n\}$ genera a V .*
- (ii) *Si $Y = \{v_1, \dots, v_r\}$ es linealmente independiente entonces $r \leq n$ y V está generado por un conjunto de la forma $\{v_1, \dots, v_r, u_{i_1}, \dots, u_{i_{n-r}}\}$ con $u_{i_j} \in X$.*
- (iii) *Cualquier base de V posee la misma cardinalidad.*

Demostración. (i) Supongamos que u_j es combinación de $\{u_i\}_{i=1}^{j-1}$, entonces $u_j = \sum_{i=1}^{j-1} \beta_i u_i$. Sea $w \in V$. Como $\{u_i\}_{i=1}^n$ genera a V , $w = \sum_{i=1}^n \alpha_i u_i$, susti-

tuyendo u_j con $\sum_{i=1}^{j-1} \beta_i u_i$ tenemos que

$$w = \sum_{i=1}^{j-1} \alpha_i u_i + \alpha_j \sum_{i=1}^{j-1} \beta_i u_i + \sum_{i=j+1}^n \alpha_i u_i = \sum_{i=1}^{j-1} (\alpha_i + \alpha_j \beta_i) u_i + \sum_{i=j+1}^n \alpha_i u_i.$$

Por lo tanto, como w era arbitrario, $\{u_1, \dots, u_{j-1}, u_{j+1}, \dots, u_n\}$ genera a V .

(ii) Como $\{u_i\}_{i=1}^n$ genera a V , si le agregamos el vector v_1 , entonces $\{v_1, u_1, \dots, u_n\}$ es linealmente dependiente y genera a V (véase el problema 3.3). Por 3.5 uno de los vectores del conjunto $\{v_1, u_1, \dots, u_n\}$ es una combinación lineal de los vectores precedentes. No puede ser v_1 , pues $\{v_1\}$ es linealmente independiente, tiene que ser uno de los de X , digamos u_j . Por (i) podemos omitir a u_j y obtener un conjunto $\{v_1, u_1, \dots, u_{j-1}, u_{j+1}, \dots, u_n\}$ que genera. Repetimos el procedimiento con v_2 . Entonces $\{v_1, v_2, u_1, \dots, u_{j-1}, u_{j+1}, \dots, u_n\}$ es linealmente dependiente y genera a V . Por 3.5 uno de los vectores del conjunto es una combinación lineal de los precedentes. Como $\{v_1, v_2\}$ es linealmente independiente, ese vector debe ser una u_k . Por (i) podemos omitir u_k y obtener un conjunto $\{v_1, v_2, u_1, \dots, u_{j-1}, u_{j+1}, \dots, u_{k-1}, u_{k+1}, \dots, u_n\}$ que genera a V . Si continuamos el proceso obtendremos un conjunto, para $r \leq n$, $\{v_1, v_2, \dots, v_r, u_1, \dots, u_{i_n-r}\}$ que genera a V .

(iii) Sea $\{u_1, \dots, u_n\}$ una base de V y $\{v_1, v_2, \dots\}$ otra base de V . Como $\{u_i\}_{i=1}^n$ genera a V , la base $\{v_j\}$ debe contener n o menos vectores, pues, si no, sería linealmente dependiente (por (ii) o el problema 3.4). Si la base $\{v_j\}$ contiene menos de n vectores, entonces $\{u_i\}_{i=1}^n$ es linealmente dependiente (por (ii) o por el problema 3.4). Luego, la base $\{v_j\}$ contiene n elementos. ■

Obsérvese que los espacios vectoriales K^n y K^m son isomorfos si, y sólo si, $n = m$.

3.8 DEFINICION. La *dimensión* de un espacio vectorial V sobre un campo K , denotada $\dim V$, es el número de elementos de una base de V .

A continuación estableceremos un resultado que relaciona la dimensión de la suma de subespacios, con la de cada uno de ellos.

3.9 TEOREMA. Sean U y V subespacios de un espacio vectorial W sobre un campo K de dimensión finita. Entonces

$$\dim(U + V) = \dim U + \dim V - \dim(U \cap V).$$

Demostración. Sea $n = \dim U$, $m = \dim V$ y $r = \dim(U \cap V)$. Sea $\{u_i\}_{i=1}^r$ una base de $U \cap V$. Por el problema 3.5 (iii) $\{u_i\}_{i=1}^r$ es parte de una base de U y también de una base de V , digamos $A = \{u_1, \dots, u_r, v_1, \dots, v_{n-r}\}$ y $B = \{u_1, \dots, u_r, w_1, \dots, w_{m-r}\}$ respectivamente.

Consideremos el conjunto $C = \{u_1, \dots, u_r, v_1, \dots, v_{n-r}, w_1, \dots, w_{m-r}\}$ y veamos que es una base de $U + V$ con lo que habremos terminado. Como A genera a U y B genera a V , C genera a $U + V$. Nos resta probar que C es linealmente independiente, pero esto lo dejamos como ejercicio al lector (problema 3.8).■

3.10 COROLARIO. $\dim(U \oplus V) = \dim U + \dim V$.■

3.11 PROPOSICION. Sea $f: U \rightarrow V$ una transformación lineal de espacios vectoriales sobre un campo K . Entonces

$$\dim U = \dim(\operatorname{im} f) + \dim(\ker f).$$

Demostración. Sea $n = \dim U$. Como $\ker f$ es un subespacio de U , $\dim(\ker f) \leq \dim U = n$. Sea $r = \dim(\ker f) \leq n$. Veamos que $\dim(\operatorname{im} f) = n - r$. Sea $\{v_1, \dots, v_r\}$ una base de $\ker f$. Podemos extenderla a una base de U de la forma $\{v_1, \dots, v_r, w_1, \dots, w_{n-r}\}$. Consideremos $\{f(w_1), \dots, f(w_{n-r})\}$ y veamos que es una base de $\operatorname{im} f$.

Sea $v \in \operatorname{im} f$. Entonces existe $u \in U$ tal que $f(u) = v$. Como $\{v_1, \dots, v_r, w_1, \dots, w_{n-r}\}$ genera a U , $u = \alpha_1 v_1 + \dots + \alpha_r v_r + \beta_1 w_1 + \dots + \beta_{n-r} w_{n-r}$ con $\alpha_i, \beta_i \in K$. Como $f(v_i) = 0$ para $i = 1, \dots, r$ pues $v_i \in \ker f$, tenemos que $f(u) = v = f(\alpha_1 v_1 + \dots + \alpha_r v_r + \beta_1 w_1 + \dots + \beta_{n-r} w_{n-r}) = \beta_1 f(w_1) + \dots + \beta_{n-r} f(w_{n-r})$. Así, $f(w_i)$ genera a la imagen de f .

Ahora veamos la independencia lineal: sea $\beta_1 f(w_1) + \beta_2 f(w_2) + \dots + \beta_{n-r} f(w_{n-r}) = 0$. Entonces $f(\beta_1 w_1 + \dots + \beta_{n-r} w_{n-r}) = 0$ y por lo tanto $\sum_{i=1}^{n-r} \beta_i w_i \in \ker f$. Como $\{v_i\}$ genera a $\ker f$, existe $\alpha_i \in K$, $i = 1, \dots, r$ tal que

$$\beta_1 w_1 + \beta_2 w_2 + \dots + \beta_{n-r} w_{n-r} = \alpha_1 v_1 + \alpha_2 v_2 + \dots + \alpha_r v_r$$

i.e.,

$$\beta_1 w_1 + \dots + \beta_{n-r} w_{n-r} - \alpha_1 v_1 - \dots - \alpha_r v_r = 0.$$

Como $\{v_1, \dots, v_r, w_1, \dots, w_{n-r}\}$ es una base de U , es linealmente independiente

y por lo tanto $\beta_i = \alpha_i = 0$. En particular $\beta_i = 0$, $i = 1, \dots, n - r$. Luego, los $f(w_i)$ son linealmente independientes. Por lo tanto $\dim(\operatorname{im} f) = n - r$. ■

PROBLEMAS

3.1 Compruebe que

- (i) Si en la expresión $\sum_{j=1}^n \alpha_j v_j = 0$ una de las α_j no es cero, entonces el conjunto $\{v_j\}$ es linealmente dependiente.
- (ii) Si en el conjunto $\{v_j\}$ alguna $v_j = 0$ entonces el conjunto $\{v_j\}$ es linealmente dependiente.
- (iii) Cualquier vector diferente de cero es, por sí mismo, linealmente independiente.
- (iv) Si en $\{v_j\}$, $v_i = v_j$ para alguna $i \neq j$ entonces la familia es linealmente dependiente.
- (v) Dos elementos de un espacio vectorial V sobre un campo K son linealmente dependientes si, y sólo si, uno de ellos es un múltiplo del otro.
- (vi) Un conjunto de vectores que contiene un conjunto linealmente dependiente es linealmente dependiente.
- (vii) Un subconjunto de un conjunto linealmente independiente es linealmente independiente.

3.2 Demuestre que un conjunto de elementos $\{v_j\}$, $j \in \{1, 2, \dots, n\}$ de un espacio vectorial V sobre un campo K es linealmente dependiente si, y sólo si, uno de ellos es combinación lineal de los restantes.

3.3 Sea $X = \{u_1, \dots, u_n\}$ un conjunto generador de un espacio vectorial V sobre un campo K . Pruebe que si v es cualquier vector de V , entonces $\{v, u_1, \dots, u_n\}$ es un conjunto linealmente dependiente y genera V .

3.4 Supongamos que $\{v_1, \dots, v_n\}$ genera el espacio vectorial V sobre un campo K . Pruebe que $n + 1$ vectores o más de V son linealmente dependientes.

3.5 Sea S un subconjunto de un espacio vectorial V . Un subconjunto de elementos $\{u_i\}_{i=1}^r$ de S se llama *subconjunto independiente máximo* de S , si es un

subconjunto linealmente independiente de S y si $v \in V$ es cualquier elemento de S , entonces $\{u_i\}_{i=1}^r \cup \{v\}$ es linealmente dependiente. Pruebe que:

- (i) si $\langle S \rangle = V$ y $\{u_i\}_{i=1}^n$ es un subconjunto independiente máximo de S entonces $\{u_i\}_{i=1}^n$ es una base de V .
- (ii) si $\#S < \infty$ y $\langle S \rangle = V$ entonces $\dim V < \infty$ y un subconjunto de S es una base de V .
- (iii) cualquier conjunto linealmente independiente es parte de una base de V .

3.6 Sea U un subespacio de un espacio vectorial V sobre un campo K de dimensión finita. Pruebe que $\dim U \leq \dim V$ y que si $\dim U = \dim V$ entonces $U \cong V$.

3.7 Caracterice los subespacios de dimensiones 0, 1, 2 y 3 del espacio vectorial \mathbb{R}^3 sobre \mathbb{R} .

3.8 Pruebe que el conjunto C del teorema 3.9 es linealmente independiente.

3.9 Sea V el espacio vectorial de las matrices de $m \times n$ sobre un campo K . Sea $E_{ij}^1 \in V$ la matriz con 1 en el lugar ij y cero en todos los demás. Pruebe que $\{E_{ij}^1\}$ es una base de V y que $\dim V = mn$.

3.10 Sean U_1, \dots, U_s subespacios de un espacio vectorial V sobre un campo K . Pruebe que $V = \bigoplus_{i=1}^s U_i$ si, y sólo si la unión de los elementos de las bases de cada U_i es una base de V .

I.4 APLICACIONES

Considérese el sistema de ecuaciones lineales

$$\begin{array}{ccccccc} a_{11}x_1 & + & \cdots & + & a_{1n}x_n & = & b_1 \\ \vdots & & & & \vdots & & \vdots \\ a_{m1}x_1 & + & \cdots & + & a_{mn}x_n & = & b_m \end{array}$$

con a_{ij} en el campo K . Podemos reescribirlo de la siguiente manera:

$$x_1 \begin{pmatrix} a_{11} \\ \vdots \\ a_{m1} \end{pmatrix} + \cdots + x_n \begin{pmatrix} a_{1n} \\ \vdots \\ a_{mn} \end{pmatrix} = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix} \quad (*)$$

4.1 PROPOSICION. *Considérese el sistema de ecuaciones lineales homogéneo*

$$x_1A^1 + \cdots + x_nA^n = 0$$

donde $A^i = {}^t(a_{1i}, \dots, a_{mi}) \in K^m$ y $n > m$. Entonces existe una solución no trivial del sistema.

Demostración. Como $n > m$, y más de m vectores son linealmente dependientes, existen elementos del campo $s_1, \dots, s_n \in K$ no todos cero tales que $s_1A^1 + \cdots + s_nA^n = 0$. Por lo tanto $S = \{s_1, \dots, s_n\}$ es una solución no trivial del sistema. ■

4.2 PROPOSICION. *Si en (*), $m = n$ y el conjunto $\{A^i\}$ es linealmente independiente entonces el sistema posee una solución y ésta es única.*

Demostración. Como $\{A^i\}$ es linealmente independiente, forma una base de K^n y en consecuencia el vector columna ${}^t(b_1, \dots, b_n) = B$ puede expresarse de manera única como combinación lineal de las A^i , es decir, como

$$x_1A^1 + \cdots + x_nA^n = B.$$

Luego, $X = (x_1, \dots, x_n)$ es la única solución. ■

También podemos escribir el sistema original de ecuaciones lineales en la forma

$$AX = B$$

donde

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix}, \quad X = {}^t(x_1, \dots, x_n) \text{ y } B = {}^t(b_1, \dots, b_m).$$

Observe que cualquier matriz A de $m \times n$ lugares determina una aplicación lineal $f: K^n \rightarrow K^m$ dada por $v \mapsto Av$ donde los vectores de K^n y K^m se ponen como vectores columna. La linealidad es inmediata pues $A(v_1 + v_2) = Av_1 + Av_2$ y $A(\alpha v) = \alpha A(v)$, $\alpha \in K$. Por ejemplo, si $A = \begin{pmatrix} 3 & 1 & 2 \\ 2 & 1 & 3 \end{pmatrix}$, A determina $f: K^3 \rightarrow K^2$ dada por $f(v) = Av$. Así, si $v = \begin{pmatrix} 2 \\ 4 \\ 6 \end{pmatrix}$, $f(v) = Av = \begin{pmatrix} 3 & 1 & 2 \\ 2 & 1 & 3 \end{pmatrix} \begin{pmatrix} 2 \\ 4 \\ 6 \end{pmatrix} = \begin{pmatrix} 22 \\ 26 \end{pmatrix}$.

De esta forma, la solución de la ecuación $AX = 0$ es el núcleo de la aplicación lineal $f = A: K^n \rightarrow K^m$.

Recuerde que el rango de una matriz A es el número máximo de renglones linealmente independientes (que también es el número máximo de columnas linealmente independientes). Recuerde también que el rango de una matriz se obtiene reduciéndola mediante operaciones elementales a una matriz escalonada. Por definición, si $f: U \rightarrow V$ es una transformación lineal, el *rango de f* es la dimensión de la imagen y la *nulidad* de f es la dimensión del núcleo de f , i.e.

$$\text{rango } f = \dim(\text{im } f) \text{ y } \text{nul } f = \dim(\text{ker } f).$$

4.3 PROPOSICION. *La dimensión del espacio solución de un sistema de ecuaciones lineales homogéneo $AX = 0$ es $n - r$ donde n es el número de incógnitas y r es el rango de la matriz de coeficientes.*

Demostración. Por la proposición 3.11,

$$\dim(\text{ker } f) = \dim K^n - \dim(\text{im } f) = n - \text{rango } A.$$

Pero n es el número de incógnitas, luego el resultado es inmediato. ■

Observe que, si vemos a A como transformación lineal, entonces la dimensión de la imagen de A coincide con el rango de A pues $\text{im } A$ corresponde a su espacio columna. (Veáse el problema 4.8.)

4.4 EJEMPLO. Consideremos el siguiente sistema de ecuaciones lineales

$$\begin{aligned}x + 3y - 2z + 4s - t &= 0 \\x + 3y - z + 3s + t &= 0 \\2x + 6y - z + 5s + 4t &= 0.\end{aligned}$$

Obtenemos el rango de la matriz de coeficientes

$$\begin{pmatrix} 1 & 3 & -2 & 4 & -1 \\ 1 & 3 & -1 & 3 & 1 \\ 2 & 6 & -1 & 5 & 4 \end{pmatrix} \sim \begin{pmatrix} 1 & 3 & -2 & 4 & -1 \\ 0 & 0 & -1 & 1 & -2 \\ 0 & 0 & -3 & 3 & -6 \end{pmatrix} \sim \begin{pmatrix} 1 & 3 & -2 & 4 & -1 \\ 0 & 0 & -1 & 1 & -2 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Así tenemos un sistema de dos ecuaciones con cinco incógnitas

$$\begin{aligned}x + 3y - 2z + 4s - t &= 0 \\-z + s - 2t &= 0.\end{aligned}$$

La dimensión del espacio solución es $n - r$, es decir 3. Tenemos tres variables libres (y, s, t) .

En términos de y, s, t obtenemos $z = s - 2t$ y $x = -3y - 2s - 3t$. Luego, las soluciones son de la forma $(-3y - 2s - 3t, y, s - 2t, s, t)$. Si damos valores particulares para (y, s, t) obtenemos valores particulares que son solución del sistema. Por ejemplo, si damos valores $(1, 0, 0), (0, 1, 0)$ y $(0, 0, 1)$ para la terna (y, s, t) obtenemos las soluciones $u_1 = (-3, 1, 0, 0, 0)$, $u_2 = (-2, 0, 1, 1, 0)$ y $u_3 = (-3, 0, -2, 0, 1)$ respectivamente, las cuales forman una base del espacio solución.

Sea V un espacio vectorial sobre un campo K . Sea $f: V \rightarrow V$ una aplicación lineal de V en sí mismo. Tal aplicación lineal se llama *operador lineal*.

4.5 DEFINICION. Diremos que un operador lineal $f: V \rightarrow V$ es *invertible* si posee un inverso f^{-1} , i.e., si existe $f^{-1}: V \rightarrow V$ tal que $f \circ f^{-1} = 1_V = f^{-1} \circ f$.

4.6 DEFINICION. Sea $g: U \rightarrow V$ una transformación lineal de espacios vectoriales sobre un campo K . Diremos que g es *no singular* si $\ker g = \{0\}$. g es *singular* si no es no singular.

4.7 PROPOSICION. Sea V un espacio vectorial de dimensión finita sobre un campo K . El operador $f: V \rightarrow V$ es invertible si, y sólo si, es no singular.

Demostración. Si f es invertible, entonces es biyectiva, en particular es inyectiva y, por el problema 2.8, $\ker f = \{0\}$. Luego, f es no singular.

Si f es no singular, entonces $\ker f = \{0\}$ y por lo tanto f es inyectiva (problema 2.8). Como $\dim V < \infty$, $\dim V = \dim(\operatorname{im} f) + \dim(\ker f) = \dim(\operatorname{im} f)$. Por lo tanto $\operatorname{im} f = V$, i.e., f es suprayectiva. Así, f es biyectiva, luego invertible. ■

Consideremos el sistema de n ecuaciones con n incógnitas $AX = B$. Si suponemos que el sistema homogéneo asociado $AX = 0$ solamente posee la solución trivial, i.e., $\ker A = \{0\}$, entonces, por 4.7, A es biyectiva y por lo tanto $AX = B$ posee una solución única para cualquier B .

Si suponemos que A es singular, i.e., $\ker A \neq \{0\}$, existe una solución no trivial de $AX = 0$. Luego A no es suprayectiva, i.e., existe $B \in K^n$ para el cual $AX = B$ no tiene solución. Aún más, si existe una solución, ésta no es única.

PROBLEMAS

4.1 Pruebe que los renglones diferentes de cero de una matriz escalonada son linealmente independientes.

4.2 Encuentre la dimensión y una base para el espacio solución de los sistemas

$$\begin{aligned} \text{(a)} \quad & 2x + 3y + 2z - 4t = 0 \\ & 2x + 3y + z + 2t = 0 \\ & 6x - 2y + 8z - t = 0 \end{aligned}$$

$$\begin{aligned} \text{(b)} \quad & x + 2y + 2z - s + 3t = 0 \\ & x + 2y + 3z + s + t = 0 \\ & 3x + 6y + 8z + s + 5t = 0 \end{aligned}$$

4.3 Encuentre el sistema homogéneo cuyo espacio solución está generado por el conjunto $\{(1, 3, 2), (4, 5, 8), (3, 8, 6)\}$. Haga lo mismo para el conjunto $\{(1, -2, 0, 3), (1, -1, -1, 4), (1, 0, -2, 5)\}$.

4.4 Sean U y V los subespacios de \mathbb{R}^3 dados por $U = \{(x, y, z) | x + y + z = 0\}$, $V = \{(x, y, z) | x + y = 0\}$. Encuentre una base y la dimensión de U , V y $U \cap V$. Haga lo mismo para $U = \{(a, b, c, d) \in \mathbb{R}^4 | b + c + d = 0\}$ y $W = \{(a, b, c, d) \in \mathbb{R}^4 | a + b = 0, c = 2d\}$.

4.5 Sean U y V los subespacios de \mathbb{R}^4 generados por $\{(1, 4, 3, 2), (2, 4, 6, 8), (3, 6, 4, 2)\}$ y $\{(2, 3, 1, 4), (1, 1, 2, 0), (3, 1, 2, 4)\}$. Encuentre la base y la dimensión de $U + V$ y de $U \cap V$. Haga lo mismo para los subespacios U y V de \mathbb{R}^5 generados por los conjuntos $\{(1, 3, -2, 2, 3), (1, 4, -3, 4, 2), (2, 3, -1, -2, 9)\}$ y $\{(1, 3, 0, 2, 1), (1, 5, -6, 6, 3), (2, 5, 3, 2, 1)\}$.

4.6 Sea A una matriz de $n \times n$. Pruebe que el rango de A es n si, y sólo si A es invertible.

4.7 Encuentre la transformación inversa del operador lineal $f: \mathbb{R}^3 \rightarrow \mathbb{R}^3$ dado por $f(x, y, z) = (y, 2x - z, z)$.

4.8 Pruebe que el rango por columna de una matriz A de $m \times n$ es igual al rango de la función lineal $f: K^n \rightarrow K^m$ determinada por A . (Sugerencia: verifique que para cada $j = 1, \dots, n$, el vector $f(e_j) = Ae_j$ es la j -columna de A .)

I.5 LA MATRIZ ASOCIADA A UNA TRANSFORMACION LINEAL

Sea K un campo. Denotemos con $\text{Hom}_K(U, V)$ el conjunto de transformaciones lineales del espacio vectorial U sobre K en el espacio V sobre K . Sean $f, g: U \rightarrow V$ aplicaciones lineales y definamos $f + g: U \rightarrow V$ mediante $(f + g)(u) = f(u) + g(u)$. También, si $f: U \rightarrow V$ y $\alpha \in K$ definamos una multiplicación escalar $\alpha f: U \rightarrow V$ mediante $(\alpha f)(u) = \alpha(f(u))$. Es inmediato comprobar que $f + g$ y αf son lineales (problema 5.1(a)). Aún más, (véanse los problemas 5.1(b) y 1.8) tenemos el siguiente resultado:

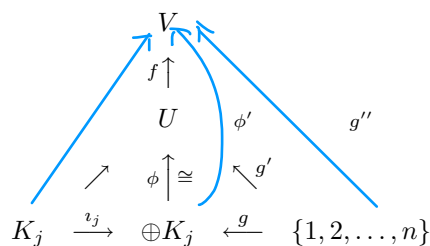
5.1 TEOREMA. *Sean U y V espacios vectoriales sobre un campo K . Entonces $\text{Hom}_K(U, V)$ con las operaciones definidas arriba es un espacio vectorial sobre K . ■*

¿Cuál será la dimensión del espacio $\text{Hom}_K(U, V)$ si U y V son de dimensión finita? Para responder esta pregunta, primero veamos un resultado previo que nos dice que una transformación lineal está totalmente determinada si conocemos la imagen de los elementos de la base de U .

5.2 PROPOSICION. *Sean U y V espacios vectoriales sobre un campo K . Sea $\{u_i\}_{i=1}^n$ una base de U y $\{v_i\}_{i=1}^n$ cualesquiera vectores de V . Entonces existe una función lineal única $f: U \rightarrow V$ tal que $f(u_i) = v_i$, $i = 1, \dots, n$.*

Demostración. Daremos dos demostraciones.

(1) Consideremos el diagrama



Por 3.1 y 3.6, basta tomar $f = \phi' \circ \phi^{-1}$ donde $\phi': \oplus_{j=1}^n K_j \rightarrow V$ es la función lineal única tal que $\phi' \circ g = g''$ pues ϕ es biyectiva.

(2) Definamos $f: U \rightarrow V$ mediante $f(u) = f(\alpha_1 u_1 + \cdots + \alpha_n u_n) = \alpha_1 v_1 + \cdots + \alpha_n v_n$. En particular $f(u_i) = f(0u_1 + \cdots + 1u_i + \cdots + 0u_n) = v_i$. Veamos que f es lineal: sean $u = \sum_{i=1}^n \alpha_i u_i$ y $u' = \sum_{i=1}^n \beta_i u_i$ entonces $f(u+u') = \sum_{i=1}^n (\alpha_i + \beta_i) v_i = \sum_{i=1}^n \alpha_i v_i + \sum_{i=1}^n \beta_i v_i = f(u) + f(u')$ y $f(\alpha u) = \sum_{i=1}^n \alpha \alpha_i v_i = \alpha \sum_{i=1}^n \alpha_i v_i = \alpha f(u)$. Veamos que f es única: sea $f': U \rightarrow V$ otra aplicación lineal tal que $f'(u_i) = v_i$, $i = 1, \dots, n$. Entonces $f'(u) = f'(\sum \alpha_i u_i) = \sum \alpha_i f'(u_i) = \sum \alpha_i v_i = f(u)$. Como u es arbitraria, $f' = f$. ■

5.3 TEOREMA. Si $\dim U = n$ y $\dim V = m$ entonces $\dim \text{Hom}_K(U, V) = nm$.

Demostración. Sea $\{u_i\}_{i=1}^n$ una base de U y $\{v_j\}_{j=1}^m$ una base de V . Encontramos una base para $\text{Hom}_K(U, V)$ y contemos el número de elementos de dicha base. Para ello definimos $f_{ij} \in \text{Hom}_K(U, V)$ mediante

$$f_{ij}(u_k) = \begin{cases} v_j & \text{si } k = i \\ 0 & \text{si } k \neq i. \end{cases}$$

Veamos que $\{f_{ij}\}$ es linealmente independiente: supongamos que $\sum_{i=1}^n \sum_{j=1}^m \alpha_{ij} f_{ij} = 0$; $\alpha_{ij} \in K$. Pero para u_k

$$0 = \sum_{i=1}^n \sum_{j=1}^m \alpha_{ij} f_{ij}(u_k) = \sum_{j=1}^m \alpha_{kj} f_{kj}(u_k) = \sum_{j=1}^m \alpha_{kj} v_j;$$

pero como las v_j son linealmente independientes, para $k = 1, \dots, n$ tenemos que $\alpha_{k1} = \alpha_{k2} = \cdots = \alpha_{km} = 0$. Luego $\alpha_{ij} = 0$ y por lo tanto $\{f_{ij}\}$ es linealmente independiente.

Veamos que $\{f_{ij}\}$ genera a $\text{Hom}_K(U, V)$: sea f cualquier elemento de $\text{Hom}_K(U, V)$. Sea $w_i = f(u_i)$, $i = 1, \dots, n$. Como $w_k \in V$, $w_k = \alpha_{k1} v_1 + \cdots + \alpha_{km} v_m$; $k = 1, \dots, n$; $\alpha_{ij} \in K$. Luego, al evaluar en u_k , $\sum_{i=1}^n \sum_{j=1}^m \alpha_{ij} f_{ij}(u_k) = \sum_{j=1}^m \alpha_{kj} f_{kj}(u_k) = \sum_{j=1}^m \alpha_{kj} v_j = w_k$. Pero $w_k = f(u_k)$. Luego, $f = \sum_{i=1}^n \sum_{j=1}^m \alpha_{ij} f_{ij}$ y por lo tanto $\{f_{ij}\}$ genera a $\text{Hom}_K(U, V)$. Como hay nm elementos en $\{f_{ij}\}$, $\dim \text{Hom}_K(U, V) = nm$. ■

Sea $f: U \rightarrow V$ una aplicación de espacios vectoriales U y V con $\dim U = m$ y $\dim V = n$. Supongamos que $\beta = \{u_1, \dots, u_m\}$ y $\beta' = \{v_1, \dots, v_n\}$ son bases para U y V respectivamente. Como $f(u_i) \in V$, tenemos que

$$\begin{aligned} f(u_1) &= \alpha_{11} v_1 + \cdots + \alpha_{1n} v_n \\ &\vdots \\ f(u_m) &= \alpha_{m1} v_1 + \cdots + \alpha_{mn} v_n \end{aligned}$$

El sistema de ecuaciones anterior lo podemos escribir como

$$\begin{pmatrix} f(u_1) \\ \vdots \\ f(u_m) \end{pmatrix} = \begin{pmatrix} \alpha_{11}v_1 + \cdots + \alpha_{1n}v_n \\ \vdots \\ \alpha_{m1}v_1 + \cdots + \alpha_{mn}v_n \end{pmatrix} = \begin{pmatrix} \alpha_{11} & \cdots & \alpha_{1n} \\ \vdots & & \vdots \\ \alpha_{m1} & \cdots & \alpha_{mn} \end{pmatrix} \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}.$$

A la matriz

$$[f]_{\beta}^{\beta'} = {}^t \begin{pmatrix} \alpha_{11} & \cdots & \alpha_{1n} \\ \vdots & & \vdots \\ \alpha_{m1} & \cdots & \alpha_{mn} \end{pmatrix} = \begin{pmatrix} \alpha_{11} & \cdots & \alpha_{m1} \\ \vdots & & \vdots \\ \alpha_{1n} & \cdots & \alpha_{mn} \end{pmatrix}$$

se le llama *matriz asociada a la transformación lineal f*, y decimos que *representa a f*.

5.4 EJEMPLO. Sea $f: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ dado por $f(x, y) = (2x - y, x + y)$. Calculemos $[f]_{\beta}^{\beta'}$ con respecto a la base $\beta = \beta' = \{(1, 0), (0, 1)\}$. Entonces

$$\begin{aligned} f(1, 0) &= (2, 1) = 2(1, 0) + 1(0, 1) \quad \text{y} \\ f(0, 1) &= (-1, 1) = -1(1, 0) + 1(0, 1). \end{aligned}$$

Luego $[f]_{\beta}^{\beta'} = \begin{pmatrix} 2 & -1 \\ 1 & 1 \end{pmatrix}$.

5.5 EJEMPLO. Sea $f: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ la aplicación lineal dada por $f(x, y) = (4x + y, 2x - 4y)$. Calculemos $[f]_{\gamma}^{\gamma'}$ donde $\gamma = \gamma' = \{(1, 1), (-1, 0)\}$:

$$\begin{aligned} f(1, 1) &= (5, -2) = (-2)(1, 1) + (-7)(-1, 0) \quad \text{y} \\ f(-1, 0) &= (-4, -2) = (-2)(1, 1) + (2)(-1, 0). \quad \text{Luego} \end{aligned}$$

$$[f]_{\gamma}^{\gamma'} = \begin{pmatrix} -2 & -2 \\ -7 & 2 \end{pmatrix}.$$

Observemos que si $u = (3, 5)$ entonces, en términos de la base γ , $u = (3, 5) = 5(1, 1) + 2(-1, 0)$. Luego $f(u) = f(3, 5) = (17, -14) = -14(1, 1) + (-31)(-1, 0)$. Así que, el vector traspuesto de coordenadas de u es $[u]_{\gamma} = \begin{pmatrix} 5 \\ 2 \end{pmatrix}$ y el vector traspuesto de coordenadas de $f(u)$ es $[f(u)]_{\gamma'} = \begin{pmatrix} -14 \\ -31 \end{pmatrix}$. Finalmente

$$[f]_{\gamma}^{\gamma'} [u]_{\gamma} = \begin{pmatrix} -2 & -2 \\ -7 & 2 \end{pmatrix} \begin{pmatrix} 5 \\ 2 \end{pmatrix} = \begin{pmatrix} -14 \\ -31 \end{pmatrix} = [f(u)]_{\gamma'}.$$

Tenemos el siguiente resultado que establece lo que observamos en el ejemplo anterior:

5.6 PROPOSICION. Sean $\beta = \{u_1, \dots, u_m\}$ y $\beta' = \{v_1, \dots, v_n\}$ bases para los espacios vectoriales U y V sobre un campo K respectivamente. Sea $f: U \rightarrow V$ una transformación lineal. Entonces $[f]_{\beta'}^{\beta}[u]_{\beta} = [f(u)]_{\beta'}$.

Demostración. Consideremos $f(u_i) = \alpha_{i1}v_1 + \alpha_{i2}v_2 + \dots + \alpha_{in}v_n = \sum_{j=1}^n \alpha_{ij}v_j$. Entonces $[f]_{\beta'}^{\beta}$ es la matriz cuyo renglón j es $(\alpha_{1j}, \alpha_{2j}, \dots, \alpha_{mj})$.

Supongamos que $u = \gamma_1u_1 + \dots + \gamma_mu_m = \sum_{i=1}^m \gamma_iu_i$. Luego $[u]_{\beta} = {}^t(\gamma_1, \dots, \gamma_m)$. Aplicando la transformación lineal f a u obtenemos $f(u) = f(\sum_{i=1}^m \gamma_iu_i) = \sum_{i=1}^m \gamma_i f(u_i) = \sum_{i=1}^m \gamma_i (\sum_{j=1}^n \alpha_{ij}v_j) = \sum_{j=1}^n (\sum_{i=1}^m \alpha_{ij}\gamma_i)v_j = \sum_{j=1}^n (\alpha_{1j}\gamma_1 + \dots + \alpha_{mj}\gamma_m)v_j$.

Luego $[f(u)]_{\beta'}$ es el vector columna cuyo coeficiente en el nivel j es $\alpha_{1j}\gamma_1 + \dots + \alpha_{mj}\gamma_m$. Calculando

$$[f]_{\beta'}^{\beta}[u]_{\beta} = \begin{pmatrix} \alpha_{11} & \dots & \alpha_{m1} \\ \vdots & & \vdots \\ \alpha_{1n} & \dots & \alpha_{mn} \end{pmatrix} \begin{pmatrix} \gamma_1 \\ \vdots \\ \gamma_m \end{pmatrix} = \begin{pmatrix} \alpha_{11}\gamma_1 + \dots + \alpha_{m1}\gamma_m \\ \vdots \\ \alpha_{1n}\gamma_1 + \dots + \alpha_{mn}\gamma_m \end{pmatrix} = [f(u)]_{\beta'}. \blacksquare$$

La proposición anterior nos dice que, el multiplicar el vector de coordenadas de u con respecto a la base $\beta = \{u_1, \dots, u_m\}$ por la matriz $[f]_{\beta'}^{\beta}$ nos da el vector de coordenadas del vector $f(u)$ con respecto a la base $\beta' = \{v_1, \dots, v_n\}$.

5.7 DEFINICION. Sean $\beta = \{u_1, \dots, u_n\}$ y $\gamma = \{u'_1, \dots, u'_n\}$ bases de U . Considérese

$$\begin{aligned} 1_U(u_1) &= u_1 = \alpha_{11}u'_1 + \dots + \alpha_{1n}u'_n \\ &\vdots \\ 1_U(u_n) &= u_n = \alpha_{n1}u'_1 + \dots + \alpha_{nn}u'_n. \end{aligned}$$

Luego, la matriz cuadrada

$$N_{\beta}^{\gamma} = \begin{pmatrix} \alpha_{11} & \dots & \alpha_{n1} \\ \vdots & & \vdots \\ \alpha_{1n} & \dots & \alpha_{nn} \end{pmatrix}$$

se llama *matriz de transición* de la base β en la base γ . Con frecuencia escribimos simplemente N en lugar de N_{β}^{γ} . Si $\beta = \gamma$, N_{β}^{β} se denota N_{β} y se llama *matriz asociada a f con respecto (o relativa) a β* .

La matriz de transición N_β^γ puede verse como la matriz asociada a la función lineal $1_U: U \rightarrow U$ con respecto a las bases β y γ , es decir $N_\beta^\gamma = [1_U]_\beta^\gamma$.

5.8 EJEMPLO. Considere $U = \mathbb{R}^2$ con bases $\beta = \{(1, 0), (0, 1)\}$ y $\gamma = \{(1, 1), (-1, 0)\}$. Entonces

$$\begin{aligned}(1, 0) &= 0(1, 1) + (-1)(-1, 0) \quad y \\ (0, 1) &= 1(1, 1) + (1)(-1, 0).\end{aligned}$$

Luego $N_\beta^\gamma = \begin{pmatrix} 0 & 1 \\ -1 & 1 \end{pmatrix}$. Por otro lado,

$$\begin{aligned}(1, 1) &= (1)(1, 0) + (1)(0, 1) \quad y \\ (-1, 0) &= (-1)(1, 0) + 0(0, 1).\end{aligned}$$

Luego $N_\gamma^\beta = \begin{pmatrix} 1 & -1 \\ 1 & 0 \end{pmatrix}$.

Observe que $N_\gamma^\beta N_\beta^\gamma = I$.

5.9 LEMA. Sea $N = N_\beta^\gamma$ la matriz de transición de la base $\beta = \{u_i\}_{i=1}^n$ en la base $\gamma = \{u'_i\}_{i=1}^n$ del espacio vectorial U sobre un campo K . Entonces $N[u]_\gamma = [u]_\beta$, y $[u]_\gamma = N^{-1}[u]_\beta$ para toda $u \in U$.

Demostración. Sea $u'_i = \alpha_{i1}u_1 + \alpha_{i2}u_2 + \dots + \alpha_{in}u_n = \sum_{j=1}^n \alpha_{ij}u_j$, para cada $i = 1, \dots, n$. Entonces N es la matriz cuadrada con renglón j igual a $(\alpha_{1j}, \alpha_{2j}, \dots, \alpha_{nj})$.

Si suponemos que $u = \lambda_1 u'_1 + \lambda_2 u'_2 + \dots + \lambda_n u'_n = \sum_{i=1}^n \lambda_i u'_i$ entonces $[u]_\gamma = {}^t(\lambda_1, \dots, \lambda_n)$. Luego $u = \sum_{i=1}^n \lambda_i u'_i = \sum_{i=1}^n \lambda_i (\sum_{j=1}^n \alpha_{ij} u_j) = \sum_{j=1}^n (\sum_{i=1}^n \alpha_{ij} \lambda_i) u_j = \sum_{j=1}^n (\alpha_{1j} \lambda_1 + \alpha_{2j} \lambda_2 + \dots + \alpha_{nj} \lambda_n) u_j$. Así, $[u]_\beta$ es el vector columna con coeficiente j igual a $\alpha_{1j} \lambda_1 + \alpha_{2j} \lambda_2 + \dots + \alpha_{nj} \lambda_n$.

Por otro lado, el coeficiente j de $N[u]_\gamma$ se obtiene multiplicando el renglón j de N por $[u]_\gamma$, i.e., multiplicando $(\alpha_{1j}, \alpha_{2j}, \dots, \alpha_{nj})$ por ${}^t(\lambda_1, \dots, \lambda_n)$. Dicha multiplicación es precisamente $\alpha_{1j} \lambda_1 + \dots + \alpha_{nj} \lambda_n$. Luego $N[u]_\gamma$ y $[u]_\beta$ tienen los mismos coeficientes. Por lo tanto $N[u]_\gamma = [u]_\beta$. Finalmente, si multiplicamos por N^{-1} , (véase el problema 5.6), obtenemos $N^{-1}[u]_\beta = N^{-1}N[u]_\gamma = [u]_\gamma$. ■

5.10 TEOREMA. Sea N la matriz de transición de la base $\beta = \beta' = \{u_i\}$ a la base $\gamma = \gamma' = \{u'_i\}$ del espacio vectorial U . Sea $f: U \rightarrow U$ un operador lineal. Entonces $[f]_\gamma^{\gamma'} = N^{-1}[f]_\beta^{\beta'} N$ donde $N = N_\beta^{\beta'}$.

Demostración. Sea $u \in U$, luego $N^{-1}[f]_{\beta}^{\beta'} N[u]_{\gamma} \stackrel{(5.9)}{=} N^{-1}[f]_{\beta}^{\beta'} [u]_{\beta} \stackrel{(5.6)}{=} N^{-1}[f(u)]_{\beta'} \stackrel{(5.9)}{=} [f(u)]_{\gamma'}$. Como $[f]_{\gamma}^{\gamma'} [u]_{\gamma} = [f(u)]_{\gamma'}$ por 5.6, tenemos que $N^{-1}[f]_{\beta}^{\beta'} N[u]_{\gamma} = [f]_{\gamma}^{\gamma'} [u]_{\gamma}$. Luego $N^{-1}[f]_{\beta}^{\beta'} N = [f]_{\gamma}^{\gamma'}$. ■

5.11 EJEMPLO. Considere el operador $f: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ dado por $f(x, y) = (4x + y, 2x - 4y)$. Sean $\beta = \beta'$ y $\gamma = \gamma'$ las bases del ejemplo 5.8. Luego

$$\begin{aligned} f(1, 0) &= (4, 2) = 4(1, 0) + 2(0, 1) \text{ y} \\ f(0, 1) &= (1, -4) = 1(1, 0) + (-4)(0, 1). \end{aligned}$$

Así que

$$[f]_{\beta}^{\beta'} = \begin{pmatrix} 4 & 1 \\ 2 & -4 \end{pmatrix}.$$

Calculemos $[f]_{\gamma}^{\gamma'}$ utilizando el teorema 5.10 con la $N = N_{\gamma}^{\beta}$ obtenida en 5.8:

$$\begin{aligned} [f]_{\gamma}^{\gamma'} &= N^{-1}[f]_{\beta}^{\beta'} N \\ &= \begin{pmatrix} 0 & 1 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} 4 & 1 \\ 2 & -4 \end{pmatrix} \begin{pmatrix} 1 & -1 \\ 1 & 0 \end{pmatrix} \\ &= \begin{pmatrix} 0 & 1 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} 5 & -4 \\ -2 & -2 \end{pmatrix} = \begin{pmatrix} -2 & -2 \\ -7 & 2 \end{pmatrix} \end{aligned}$$

la cual coincide con la matriz $[f]_{\gamma}^{\gamma'}$ de 5.5.

PROBLEMAS

5.1 a) Sean $f, g: U \rightarrow V$ funciones lineales de espacios vectoriales sobre un campo K y $\alpha \in K$. Demuestre que $f + g$ y αf son funciones lineales.

b) Pruebe el teorema 5.1 con detalle. Si $U = V$ se acostumbra denotar al conjunto de funciones lineales de V en sí mismo con $End_K(V)$ y se llama conjunto de *endomorfismos* de V .

5.2 Sea $f: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ el operador lineal dado por $f(x, y) = (3x - 4y, x - y)$. Calcule la matriz $[f]_{\beta}^{\beta}$ asociada a f con respecto a la base $\beta = \{(1, 1), (-1, 0)\}$.

5.3 Sea f como en el problema anterior. ¿Cuáles son las coordenadas del vector $f(3, 5)$ con respecto a β ?

5.4 a) Describa la matriz de transición de un espacio de dimensión 1.

b) Calcule la matriz de transición de la base $\beta = \{(1, 1), (-1, 0)\}$ a la base $\beta' = \{(2, 6), (4, 10)\}$ del espacio vectorial \mathbb{R}^2 del problema 5.2.

5.5 Calcule la matriz asociada $[f]_{\beta'}^{\beta'}$ en el problema 5.4.

5.6 Pruebe que la matriz de transición N_{β}^{γ} definida en 5.7 es invertible.

5.7 Pruebe que existe un isomorfismo del espacio $Hom_K(U, U)$ con el espacio de las matrices cuadradas $M_n K$ dado por $f \mapsto [f]_{\beta}^{\beta}$.

5.8 Pruebe que si $f, g \in Hom_K(U, U)$ entonces $[f \circ g]_{\beta}^{\beta} = [f]_{\beta}^{\beta} [g]_{\beta}^{\beta}$.

5.9 Pruebe resultados análogos al problema 5.7 para $Hom_K(U, V)$ i.e., $Hom_K(U, V) \cong M_{m \times n} K$ y al problema 5.8 para cuando $f \in Hom_K(U, V)$ y $g \in Hom_K(V, W)$ son transformaciones lineales de espacios vectoriales U, V y W sobre un campo K con bases β, γ y η respectivamente, i.e. pruebe que $[g \circ f]_{\beta}^{\eta} = [g]_{\gamma}^{\eta} [f]_{\beta}^{\gamma}$.

5.10 Sea N la matriz de transición de la base β en la base γ de U . Sea M la matriz de transición de la base β' en la base γ' de V . Pruebe que si $f: U \rightarrow V$ es una aplicación lineal, entonces

$$[f]_{\gamma'}^{\gamma'} = M^{-1} [f]_{\beta}^{\beta'} N.$$



Capítulo II

VECTORES CARACTERISTICOS Y FORMAS CANONICAS

II.1 VALORES Y VECTORES CARACTERISTICOS

1.1 DEFINICION. Sean A y B matrices cuadradas. Si $A = N^{-1}BN$ con N una matriz invertible entonces diremos que A es *similar* o *semejante* a B .

1.2 PROPOSICION. *La relación de similaridad (o semejanza) es una relación de equivalencia.*

Demostración. Claramente $A = I^{-1}AI$ donde I es la matriz identidad, la cual es invertible. Luego, A es similar a A . Si A es similar a B , existe una matriz invertible N tal que $A = N^{-1}BN$. Entonces $NAN^{-1} = N(N^{-1}BN)N^{-1} = B$ o bien $(N^{-1})^{-1}AN^{-1} = B$. Luego, B es similar a A . Finalmente, si A es similar a B y B es similar a C , entonces existen matrices invertibles N y M tales que $A = N^{-1}BN$ y $B = M^{-1}CM$. Sustituyendo obtenemos $A = N^{-1}(M^{-1}CM)N = (MN)^{-1}C(MN)$ y como MN es invertible, A es similar a C . ■

1.3 PROPOSICION. Las matrices A y B representan al mismo operador lineal $f: U \rightarrow U$ si, y sólo si, son similares una con la otra.

Demostración. Sean $\beta = \beta'$ y $\gamma = \gamma'$ bases de U . Supongamos que $A = [f]_{\gamma}^{\gamma'}$ y $B = [f]_{\beta}^{\beta'}$ representan el mismo operador lineal. Por I.5.10, A es similar a B .

Sea $\beta = \beta'$ una base de U y supongamos que A es similar a $B = [f]_{\beta}^{\beta'}$, es decir, que $A = N^{-1}[f]_{\beta}^{\beta'}N$ para alguna matriz invertible N . Sea $\gamma = \gamma'$ la base de U obtenida mediante N a partir de $\beta = \beta'$. Es fácil comprobar (problema 1.5) que A representa a f con respecto a la base γ . ■

Dado un operador lineal f , podemos decir que sus matrices asociadas forman una clase de equivalencia de matrices similares.

1.4 DEFINICION. Diremos que un operador lineal $f: U \rightarrow U$ es *diagonalizable* si para alguna base de U , la matriz asociada es diagonal. Diremos que dicha base *diagonaliza* a f .

El siguiente resultado es consecuencia inmediata de 1.3:

1.5 COROLARIO. Sea $B = [f]_{\beta}^{\beta'}$ la matriz asociada al operador lineal f con respecto a la base $\beta = \beta'$. Entonces f es diagonalizable si, y sólo si, existe una matriz invertible N tal que $N^{-1}BN$ es una matriz diagonal. ■

Hacemos la observación de que no todo operador es diagonalizable. Como veremos más adelante, a todo operador f se le puede asociar una matriz especial llamada *forma canónica*, (véase §4).

1.6 DEFINICION. Sea U un espacio vectorial sobre un campo K . Sea $f: U \rightarrow U$ un operador lineal. Si existe un vector $u \in U$ distinto de cero y un escalar $\lambda \in K$ tal que $f(u) = \lambda u$ entonces llamaremos a λ *valor característico* o *valor propio* de f y a u *vector característico* o *vector propio correspondiente* a λ .

Observe que si en la definición anterior $u = 0$ entonces u sería un vector característico correspondiente a cualquier $\lambda \in K$. Nótese que u es un vector característico correspondiente a λ si $f(u)$ es un múltiplo de u , i.e., $f(u) = \lambda u$, $\lambda \in K$. También los múltiplos αu de un vector característico son vectores característicos pues $f(\alpha u) = \alpha f(u) = \alpha(\lambda u) = \lambda(\alpha u)$. Además, puede haber muchos vectores

característicos correspondientes a un valor característico. Por ejemplo, la función identidad $1_U: U \rightarrow U$ posee a $1 \in K$ como su único valor característico, pero cualquier vector distinto de cero de U es un vector característico correspondiente al 1.

1.7 PROPOSICION. *Sea λ un valor característico (fijo) de $f: U \rightarrow U$. Entonces el conjunto de los vectores $u \in U$ tales que $f(u) = \lambda u$ es un subespacio no trivial de U denotado U_λ .*

Demostración. Si $u = 0 \in U$, $f(u) = f(0) = 0 = \lambda 0 = \lambda u$, para $\lambda \in K$. Luego $0 \in U_\lambda$. Si $u, u' \in U_\lambda$ entonces $f(u) = \lambda u$ y $f(u') = \lambda u'$. Luego, $f(u + u') = f(u) + f(u') = \lambda u + \lambda u' = \lambda(u + u')$. Además $f(\alpha u) = \alpha f(u) = \alpha(\lambda u) = \lambda(\alpha u)$. Así, $u + u' \in U_\lambda$ y $\alpha u \in U_\lambda$. Luego U_λ es un subespacio de U . ■

Llamaremos a U_λ *espacio característico de λ* . Observe que U_λ consiste del cero y de todos los vectores característicos correspondientes a λ .

1.8 EJEMPLO. Si A es una matriz cuadrada, los valores característicos de A son los valores característicos de A vista como operador lineal. Por ejemplo, si $A = \begin{pmatrix} 2 & 4 \\ 3 & 6 \end{pmatrix}$ encontremos los valores característicos y sus vectores característicos no triviales asociados, i.e., si $X = \begin{pmatrix} x \\ y \end{pmatrix}$, busquemos $\lambda \in K$ tal que $AX = \lambda X$. Esto es

$$\begin{pmatrix} 2 & 4 \\ 3 & 6 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \lambda \begin{pmatrix} x \\ y \end{pmatrix}.$$

Entonces se tiene $\left. \begin{matrix} 2x + 4y = \lambda x \\ 3x + 6y = \lambda y \end{matrix} \right\}$, por lo que $\left. \begin{matrix} (2 - \lambda)x + 4y = 0 \\ 3x + (6 - \lambda)y = 0 \end{matrix} \right\}$.

Así, el sistema homogéneo anterior tiene solución no trivial si, y sólo si, el determinante de la matriz de coeficientes es cero:

$$\begin{aligned} \begin{vmatrix} (2 - \lambda) & 4 \\ 3 & (6 - \lambda) \end{vmatrix} &= (2 - \lambda)(6 - \lambda) - 12 = 12 - 2\lambda - 6\lambda + \lambda^2 - 12 \\ &= \lambda^2 - 8\lambda = \lambda(\lambda - 8) = 0. \end{aligned}$$

Así, λ es un valor característico de A correspondiente a un vector característico distinto de cero si, y sólo si, $\lambda = 0$ o $\lambda = 8$. Si $\lambda = 0$, entonces $\left. \begin{matrix} 2x + 4y = 0 \\ 3x + 6y = 0 \end{matrix} \right\}$

o bien, $x + 2y = 0$. Luego $X = \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} -2 \\ 1 \end{pmatrix}$ es un vector característico no

trivial correspondiente al valor característico asociado a $\lambda = 0$. Cualquier otro vector característico asociado a $\lambda = 0$ es un múltiplo de X . Si $\lambda = 8$, entonces $\left. \begin{array}{l} -6x + 4y = 0 \\ 3x - 2y = 0 \end{array} \right\}$ o bien $3x - 2y = 0$. Luego $Y = \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 2 \\ 3 \end{pmatrix}$ es un vector característico no trivial correspondiente al valor característico $\lambda = 8$. Cualquier otro vector característico de $\lambda = 8$ es un múltiplo de Y .

1.9 TEOREMA. *Sea $f: U \rightarrow U$ un operador lineal de un espacio vectorial U sobre un campo K . Entonces $\lambda \in K$ es un valor característico de f si, y sólo si, $\lambda I - f$ es singular.*

Demostración. λ es un valor característico de $f \iff$ existe un vector no trivial u tal que $f(u) = \lambda u \iff (\lambda I)(u) - f(u) = 0 \iff (\lambda I - f)(u) = 0$ i.e., $\lambda I - f$ es singular, (o no invertible).■

1.10 COROLARIO. *El espacio de vectores característicos de λ es igual al núcleo de $\lambda I - f$.■*

1.11 PROPOSICION. *Los vectores característicos correspondientes a valores característicos diferentes son linealmente independientes.*

Demostración. Sea $f: U \rightarrow U$ un operador lineal con u_1, \dots, u_n vectores característicos correspondientes a los valores característicos distintos $\lambda_1, \dots, \lambda_n$. Veamos que $\{u_i\}_{i=1}^n$ es linealmente independiente. Para ello hagamos inducción sobre n . Si $n = 1$, entonces u_1 es linealmente independiente pues $u_1 \neq 0$. Supongamos que, para $n > 1$,

$$\alpha_1 u_1 + \dots + \alpha_n u_n = 0, \quad \alpha_i \in K. \quad (*)$$

Luego, $f(\alpha_1 u_1 + \dots + \alpha_n u_n) = \alpha_1 f(u_1) + \dots + \alpha_n f(u_n) = f(0) = 0$. Por hipótesis, $f(u_i) = \lambda_i u_i$, y tenemos que

$$\alpha_1 \lambda_1 u_1 + \dots + \alpha_n \lambda_n u_n = 0.$$

Si multiplicamos (*) por λ_n obtenemos

$$\alpha_1 \lambda_n u_1 + \dots + \alpha_n \lambda_n u_n = 0.$$

Restando las dos últimas ecuaciones obtenemos

$$\alpha_1 (\lambda_1 - \lambda_n) u_1 + \dots + \alpha_{n-1} (\lambda_{n-1} - \lambda_n) u_{n-1} = 0.$$

Por inducción, $\alpha_i(\lambda_i - \lambda_n) = 0$ para $i = 1, \dots, n - 1$. Como las λ_i son distintas entre sí, $\lambda_i - \lambda_n \neq 0$ para $i \neq n$. Luego, $\alpha_i = 0$ para $i = 1, \dots, n - 1$. Al sustituir en (*) obtenemos que $\alpha_n u_n = 0$ y, por lo tanto, $\alpha_n = 0$. Así, las u_i son linealmente independientes. ■

Nota. Vectores característicos linealmente independientes pueden corresponder a un mismo valor característico.

El siguiente teorema relaciona los conceptos de matriz diagonal y vector característico.

1.12 TEOREMA. *Sea $f: U \rightarrow U$ un operador lineal. Entonces f tiene como matriz asociada a (o puede representarse por) una matriz diagonal si, y sólo si, U posee una base que consta de vectores característicos de f .*

Demostración. f puede representarse por una matriz diagonal

$$B = \begin{pmatrix} \lambda_1 & 0 & 0 & \cdots & 0 \\ 0 & \lambda_2 & 0 & \cdots & 0 \\ \vdots & & & & \vdots \\ 0 & 0 & 0 & \cdots & \lambda_n \end{pmatrix}$$

si, y sólo si, existe una base $\{u_i\}_{i=1}^n$ de U tal que

$$f(u_1) = \lambda_1 u_1 + 0u_2 + \cdots + 0u_n$$

$$f(u_2) = 0u_1 + \lambda_2 u_2 + \cdots + 0u_n$$

$$\vdots$$

$$f(u_n) = 0u_1 + 0u_2 + \cdots + \lambda_n u_n.$$

i.e. si, y sólo si, los vectores $\{u_i\}_{i=1}^n$ son vectores característicos de f correspondientes a los valores característicos $\{\lambda_i\}_{i=1}^n$. ■

Observe que los elementos de la diagonal de B son precisamente los valores característicos correspondientes.

Observe también que, por 1.12, una matriz cuadrada A es similar a una matriz diagonal B si, y sólo si, A posee n vectores característicos linealmente independientes. Aquí también B posee en su diagonal a los valores característicos. Si N es la matriz cuyas columnas son los n vectores característicos de A , entonces $B = N^{-1}AN$.

1.13 EJEMPLO. Sea A la matriz del ejemplo 1.8. Los vectores característicos

son $X = \begin{pmatrix} -2 \\ 1 \end{pmatrix}$ y $Y = \begin{pmatrix} 2 \\ 3 \end{pmatrix}$. Sea $N = \begin{pmatrix} -2 & 2 \\ 1 & 3 \end{pmatrix}$, $N^{-1} = \begin{pmatrix} -3/8 & 2/8 \\ 1/8 & 2/8 \end{pmatrix}$. Luego A es similar a la matriz diagonal

$$B = N^{-1}AN = \begin{pmatrix} -3/8 & 2/8 \\ 1/8 & 2/8 \end{pmatrix} \begin{pmatrix} 2 & 4 \\ 3 & 6 \end{pmatrix} \begin{pmatrix} -2 & 2 \\ 1 & 3 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 8 \end{pmatrix}$$

donde los elementos de la diagonal son los valores característicos correspondientes.

PROBLEMAS

1.1 Recuerde que la *traza* de una matriz cuadrada A (i.e. la función $tr: M_n(K) \rightarrow K$), denotada trA , es la suma de los elementos de su diagonal. Pruebe que:

- (i) $tr(AB) = tr(BA)$ y que
- (ii) $tr(A) = tr(B)$ si A y B son similares. Debido a esto último, decimos que la traza es un *invariante* bajo la relación de *similaridad* o semejanza.

1.2 Calcule los valores característicos y los vectores característicos correspondientes de la matriz $A = \begin{pmatrix} 1 & 5 \\ 4 & 3 \end{pmatrix}$ y encuentre una matriz invertible N tal que $N^{-1}AN$ sea diagonal.

1.3 Calcule los valores característicos y los vectores característicos correspondientes de la matriz $A = \begin{pmatrix} 1 & 4 & 2 \\ 2 & 3 & 2 \\ 1 & 1 & 2 \end{pmatrix}$ y encuentre una matriz invertible N tal que $N^{-1}AN$ sea diagonal.

1.4 Sean $A, B \in M_{m \times n}K$. Se dice que A es *equivalente* a B si existen matrices invertibles M y N tales que $A = MBN$. Pruebe que la equivalencia de matrices es una relación de equivalencia.

1.5 Pruebe con todo detalle que A representa a f con respecto a la base γ en la proposición 1.3.

1.6 (i) Supóngase que A y B son matrices similares de $n \times n$, i.e. $B = N^{-1}AN$ para alguna matriz invertible N . Pruebe que $B^k = N^{-1}A^kN$.

(ii) Suponga que los conejos no se reproducen durante su primer mes de vida, pero que a partir del segundo mes cada pareja de conejos produce un nuevo par. Suponga que ningún conejo muere. Si comenzamos con un par de conejos, ¿cuántas parejas de conejos hay a los n meses? Sugerencia: escriba la fórmula de recursión $u_n = u_{n-1} + u_{n-2}$ como la igualdad de matrices $(u_n, u_{n-1}) = (u_{n-1}, u_{n-2})B$ donde $B = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}$. Encuentre los valores característicos de B y utilice (i). La sucesión así obtenida se llama *sucesión de Fibonacci*, sus términos se llaman *números de Fibonacci* y el valor característico positivo obtenido se llama *sección áurea* o *proporción divina*. La solución de muchos problemas de “aplicación” de la teoría con largos enunciados se reduce a formular ecuaciones adecuadas como en este problema.

II.2 TEOREMA DE CAYLEY-HAMILTON

2.1 DEFINICION. Sea $f(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0$ un polinomio con coeficientes en un campo K . Sea A una matriz cuadrada con coeficientes también en K . Definimos el polinomio $f(A)$ como $a_n A^n + a_{n-1} A^{n-1} + \cdots + a_1 A + a_0 I$, donde I es la matriz identidad.

En forma análoga tenemos la siguiente

2.2 DEFINICION. Sea $\rho: U \rightarrow U$ un operador lineal del espacio vectorial U sobre el campo K . Sea $f(x)$ un polinomio con coeficientes en K . Definimos $f(\rho) = a_n \rho^n + \cdots + a_1 \rho + a_0 I$ donde I es la aplicación de identidad. Aquí, $\rho^n = \rho \circ \cdots \circ \rho$, n veces.

Se dice que una matriz A (o un operador ρ) es una *raíz* del polinomio f si $f(A) = 0$ (o si $f(\rho) = 0$).

2.3 DEFINICION. Sea A la matriz cuadrada

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}.$$

La matriz cuadrada $\lambda I_n - A$ se llama *matriz característica*. Llamamos *polinomio característico* $p_A(\lambda)$ de A al determinante de la matriz característica de A , i.e., $p_A(\lambda) = |\lambda I_n - A|$.

Así, $\lambda I_n - A$ es

$$\begin{pmatrix} \lambda - a_{11} & -a_{12} & \cdots & -a_{1n} \\ -a_{21} & \lambda - a_{22} & \cdots & -a_{2n} \\ \vdots & \vdots & & \vdots \\ -a_{n1} & -a_{n2} & \cdots & \lambda - a_{nn} \end{pmatrix}$$

y $p_A(\lambda)$ es $(\lambda - a_{11})(\lambda - a_{22}) \cdots (\lambda - a_{nn}) - \cdots$; o bien,

$$p_A(\lambda) = \lambda^n - (a_{11} + a_{22} + \cdots + a_{nn})\lambda^{n-1} + \cdots = \lambda^n - (\text{tr} A)\lambda^{n-1} + \cdots + (-1)^n |A|$$

donde $(-1)^n |A| = p_A(0)$ es el término constante.

Recordemos que si A es la matriz cuadrada de $n \times n$

$$\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix}$$

se denota con N_{ij} a la submatriz de $(n-1) \times (n-1)$ obtenida de A suprimiendo su renglón i y su columna j y, el determinante $|N_{ij}|$ se llama *menor* del elemento a_{ij} de A . El *cofactor* de a_{ij} se define como el escalar

$$A_{ij} = (-1)^{i+j} |N_{ij}|.$$

Por ejemplo, si $A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}$ entonces

$$N_{12} = \begin{pmatrix} 4 & 6 \\ 7 & 9 \end{pmatrix} \text{ y } A_{12} = (-1)^{1+2} \begin{vmatrix} 4 & 6 \\ 7 & 9 \end{vmatrix} = -(36 - 42) = 6.$$

Recordemos que $|A| = a_{i1}A_{i1} + a_{i2}A_{i2} + \cdots + a_{in}A_{in} = \sum_{j=1}^n a_{ij}A_{ij}$ y $|A| = a_{1j}A_{1j} + a_{2j}A_{2j} + \cdots + a_{nj}A_{nj} = \sum_{i=1}^n a_{ij}A_{ij}$.

También recordemos que la *matriz adjunta clásica* de A es la matriz traspuesta de cofactores, i.e.,

$$\tilde{A} = \begin{pmatrix} A_{11} & \cdots & A_{n1} \\ \vdots & & \vdots \\ A_{1n} & \cdots & A_{nn} \end{pmatrix}$$

y tiene la siguiente propiedad: $A\tilde{A} = \tilde{A}A = |A|I$; lo cual implica que $A^{-1} = \frac{1}{|A|}\tilde{A}$.

2.4 TEOREMA. (Cayley-Hamilton). *Toda matriz cuadrada es raíz de su polinomio característico.*

Demostración. Sea A una matriz cuadrada y $p_A(\lambda) = |\lambda I - A| = \lambda^n + a_{n-1}\lambda^{n-1} + \cdots + a_1\lambda + a_0$ su polinomio característico. Sea C_λ la adjunta clásica de la matriz $\lambda I - A$. Como los elementos de C_λ son cofactores de $\lambda I - A$, son polinomios en λ de grado menor o igual que $n-1$. Luego

$$C_\lambda = C_{n-1}\lambda^{n-1} + \cdots + C_1\lambda + C_0$$

donde C_i es una matriz cuadrada con elementos en K . Así que, aplicando la propiedad que mencionamos de la matriz adjunta clásica se tiene que $(\lambda I - A)C_\lambda = |\lambda I - A|I$. Esto es

$$(\lambda I - A)(C_{n-1}\lambda^{n-1} + \cdots + C_1\lambda + C_0) = (\lambda^n + a_{n-1}\lambda^{n-1} + \cdots + a_1\lambda + a_0)I.$$

Luego $-AC_0 = a_0I$, $C_0 - AC_1 = a_1I, \dots, C_{n-3} - AC_{n-2} = a_{n-2}I$, $C_{n-2} - AC_{n-1} = a_{n-1}I$ y $C_{n-1} = I$. Si multiplicamos ambos lados de cada igualdad por I, A, A^2, \dots, A^n respectivamente, obtenemos: $-AC_0 = a_0I$, $AC_0 - A^2C_1 = a_1A, \dots, A^{n-2}C_{n-3} - A^{n-1}C_{n-2} = a_{n-2}A^{n-2}$, $A^{n-1}C_{n-2} - A^nC_{n-1} = a_{n-1}A^{n-1}$, $A^nC_{n-1} = A^n$. Al sumar las ecuaciones obtenemos $0 = a_0I + a_1A + a_2A^2 + \dots + a_{n-1}A^{n-1} + A^n$. ■

2.5 EJEMPLO. Sea $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$. Luego

$$p_A(\lambda) = |\lambda I - A| = \begin{vmatrix} \lambda - 1 & -2 \\ -3 & \lambda - 4 \end{vmatrix} = \lambda^2 - \lambda - 4\lambda + 4 - 6 = \lambda^2 - 5\lambda - 2.$$

Por el teorema 2.4

$$\begin{aligned} p_A(A) &= \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} - 5 \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} - 2 \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 7 & 10 \\ 15 & 22 \end{pmatrix} - \begin{pmatrix} 5 & 10 \\ 15 & 20 \end{pmatrix} - \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}. \end{aligned}$$

2.6 PROPOSICION. Sea A una matriz cuadrada con elementos en un campo K . Un elemento $\alpha \in K$ es un valor característico de A correspondiente a un vector característico (diferente de 0), si, y sólo si, α es una raíz del polinomio característico $p_A(\lambda)$.

Demostración. Por 1.9, α es un valor característico de A si, y sólo si, $\alpha I - A$ es singular, es decir, no invertible. Pero $\alpha I - A$ es no invertible si, y sólo si, $|\alpha I - A| = 0$ (recuerde que una matriz es invertible si, y sólo si, su determinante es distinto de cero). i.e., α es raíz de $p_A(\lambda)$. ■

2.7 PROPOSICION. Supóngase que $p_A(\lambda)$ es de la forma

$$(\lambda - a_1)(\lambda - a_2) \cdots (\lambda - a_n)$$

con todas las a_i distintas, entonces A es similar a una matriz diagonal con las a_i en la diagonal.

Demostración. Como las a_i son las distintas raíces diferentes de cero de $p_A(\lambda)$, éstas son, por 2.6, los valores característicos de A y, por 1.11, los vectores característicos correspondientes son linealmente independientes. Luego, los vectores

característicos asociados a dichos valores característicos son una base del espacio vectorial y, por 1.12, A es similar a una matriz diagonal.■

2.8 PROPOSICION. Si A y B son matrices similares, entonces

$$p_A(\lambda) = p_B(\lambda).$$

Demostración. Si A es similar a B entonces $A = N^{-1}BN$ con N invertible. Como $\lambda I = N^{-1}\lambda IN$, tenemos que $|\lambda I - A| = |\lambda I - N^{-1}BN| = |N^{-1}\lambda IN - N^{-1}BN| = |N^{-1}(\lambda I - B)N| = |N^{-1}||\lambda I - B||N| = |\lambda I - B||N^{-1}||N| = |\lambda I - B|$.■

En otras palabras, la función $M_n(K) \rightarrow K[\lambda]$ que asigna a cada matriz su polinomio característico es un *invariante* bajo *similaridad* o semejanza.

Puesto que todo polinomio sobre \mathbb{C} posee una raíz, si A es una matriz con coeficientes en \mathbb{C} entonces A posee al menos un valor característico.

PROBLEMAS

2.1 Pruebe que si f y g son polinomios con coeficientes en K entonces $(f+g)(A) = f(A) + g(A)$, $(fg)(A) = f(A)g(A)$ y $(\lambda f)(A) = \lambda(f(A))$, $\lambda \in K$.

2.2 Sea A la matriz asociada a ρ . Pruebe que $f(A)$ es la matriz asociada a $f(\rho)$.

2.3 Pruebe que si \tilde{A} es la matriz adjunta clásica de A entonces $A\tilde{A} = \tilde{A}A = |A|I$.

2.4 Sea $A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}$. Calcule la adjunta clásica C_λ de la matriz $\lambda I - A$. Compruebe que $C_\lambda = C_2\lambda^2 + C_1\lambda + C_0$ donde C_i es una matriz cuadrada con elementos en K .

2.5 Pruebe que $p_A(\lambda) = p_{tA}(\lambda)$.

2.6 Pruebe que si A es una matriz triangular, su polinomio característico posee como valores característicos a los elementos de la diagonal.

2.7 Sea $\rho: \mathbb{R}^3 \rightarrow \mathbb{R}^3$ un operador lineal dado por $\rho(x, y, z) = (x + y, 2x, 4y)$.

Encuentre los valores característicos α y una base para cada espacio característico U_α .

2.8 Calcule los valores y vectores característicos de

$$A = \begin{pmatrix} 2 & -1 \\ 1 & 3 \end{pmatrix} \text{ y } B = \begin{pmatrix} -1 & 4 \\ -2 & 1 \end{pmatrix} \text{ sobre } \mathbb{R} \text{ y sobre } \mathbb{C}.$$

Propóngase usted mismo ejercicios similares a éstos dos últimos si presiente que no ha dominado la técnica.

II.3 EL POLINOMIO MINIMO

En esta sección presentaremos otro invariante para la relación de similaridad o semejanza.

3.1 DEFINICION. Sea A una matriz cuadrada. El *polinomio mínimo* $m_A(\lambda)$ de A es el polinomio mónico (con coeficiente inicial 1) de menor grado tal que $m_A(A) = 0$.

Por el teorema de Cayley-Hamilton, A es raíz de un polinomio $p_A(\lambda)$ diferente de cero. Supongamos que tiene grado n y que es el más pequeño de los grados tal que $p_A(A) = 0$. Si dividimos $p_A(\lambda)$ entre su coeficiente inicial obtendremos un polinomio mónico $m_A(\lambda)$ de grado n con A como raíz. Si $m'_A(\lambda)$ es otro polinomio mónico de grado n para el cual $m'_A(A) = 0$ entonces $m_A(\lambda) - m'_A(\lambda)$ es un polinomio diferente de cero de grado menor que n con raíz A . Esto contradice el que n sea el más pequeño de los grados tal que $p_A(A) = 0$. Por lo tanto $m_A(\lambda)$ es único.

Aún más, por el algoritmo de la división, si $f(\lambda)$ es un polinomio tal que $f(A) = 0$ entonces $f(\lambda) = m_A(\lambda)g(\lambda) + r(\lambda)$ con $r(\lambda) = 0$ o *gr* $r(\lambda) < \text{gr } m_A(\lambda)$. Como $f(A) = 0$ y $m_A(A) = 0$ tenemos que $r(A) = 0$. Si $r(\lambda) \neq 0$, entonces $r(\lambda)$ es un polinomio de grado menor que el de $m_A(\lambda)$ que posee a A como raíz, lo cual contradice el hecho de que $m_A(\lambda)$ sea mínimo. Luego $r(\lambda) = 0$ y $f(\lambda) = m_A(\lambda)g(\lambda)$. Podemos resumir lo anterior en la siguiente

3.2 PROPOSICION. *El polinomio mínimo de una matriz A existe, es único y divide a cualquier otro polinomio que tenga a A como raíz, en particular, divide al polinomio característico de A .■*

3.3 PROPOSICION. *El polinomio mínimo y el polinomio característico de una matriz A poseen los mismos factores irreducibles.*

Demostración. Sea $m_A(\lambda)$ el polinomio mínimo de A de la forma $m_A(\lambda) = \lambda^t + a_{t-1}\lambda^{t-1} + \dots + a_1\lambda + a_0$. Consideremos las matrices $C_0 = I$, $C_1 = A + a_{t-1}I$, $C_2 = A^2 + a_{t-1}A + a_{t-2}I$, \dots , $C_{t-1} = A^{t-1} + a_{t-1}A^{t-2} + \dots + a_1I$. Luego $C_0 = I$, $C_1 - AC_0 = a_{t-1}I$, $C_2 - AC_1 = a_{t-2}I$, \dots , $C_{t-1} - AC_{t-2} = a_1I$. Multiplicando

C_{t-1} por $-A$ obtenemos $-AC_{t-1} = a_0I - (A^t + a_{t-1}A^{t-1} + \cdots + a_1A + a_0I) = a_0I - m_A(A) = a_0I$. Sea $C(\lambda) = \lambda^{t-1}C_0 + \lambda^{t-2}C_1 + \cdots + \lambda C_{t-2} + C_{t-1}$. Entonces

$$\begin{aligned} (\lambda I - A) \cdot C(\lambda) &= (\lambda^t C_0 + \lambda^{t-1} C_1 + \cdots + \lambda C_{t-1}) \\ &\quad - (\lambda^{t-1} A C_0 + \lambda^{t-2} A C_1 + \cdots + \lambda A C_{t-2} + A C_{t-1}) \\ &= \lambda^t C_0 + \lambda^{t-1} (C_1 - A C_0) + \cdots + \lambda (C_{t-1} - A C_{t-2}) + a_0 I \\ &= \lambda^t I + a_{t-1} \lambda^{t-1} I + \cdots + a_1 \lambda I + a_0 I \\ &= m_A(\lambda) I. \end{aligned}$$

Luego, $|\lambda I - A| |C(\lambda)| = |m_A(\lambda) I| = m_A(\lambda)^n$. Por lo tanto, $|\lambda I - A|$ divide a $m_A(\lambda)^n$. Es decir, el polinomio característico de A divide a $m_A(\lambda)^n$.

Sea $g(\lambda)$ un polinomio irreducible. Si $g(\lambda) | m_A(\lambda)$ entonces $g(\lambda) | p_A(\lambda)$ pues $m_A(\lambda) | p_A(\lambda)$. Por otro lado, si $g(\lambda) | p_A(\lambda)$ entonces $g(\lambda) | m_A(\lambda)^n$. Como $g(\lambda)$ es irreducible, $g(\lambda) | m_A(\lambda)$. Luego $m_A(\lambda)$ y $p_A(\lambda)$ poseen los mismos factores irreducibles. ■

3.4 COROLARIO. *Un elemento $\alpha \in K$ es valor característico de A si, y sólo si, α es raíz del polinomio mínimo de A .*

Demostración. Por 3.3, el polinomio mínimo y el polinomio característico de A poseen los mismos factores lineales irreducibles y, por lo tanto, poseen las mismas raíces. ■

3.5 EJEMPLO. Sea $A = \begin{pmatrix} 3 & 0 & 1 & 0 \\ 0 & 3 & 0 & 0 \\ 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 8 \end{pmatrix}$. El polinomio característico de A es $p_A(\lambda) = |\lambda I - A| = (\lambda - 3)^3(\lambda - 8)$. Por los resultados precedentes, $m_A(\lambda) | p_A(\lambda)$ y además $m_A(\lambda)$ posee los mismos factores irreducibles, entonces $m_A(\lambda)$ puede ser uno de los siguientes:

$$(\lambda - 3)(\lambda - 8), (\lambda - 3)^2(\lambda - 8), \text{ ó } (\lambda - 3)^3(\lambda - 8).$$

Por el teorema (de Cayley-Hamilton) 2.4, $m_A(A) = 0$. Luego, calculando, $m_A(\lambda)$ debe ser $(\lambda - 3)^2(\lambda - 8)$.

PROBLEMAS

3.1 Encuentre el polinomio mínimo de $A = \begin{pmatrix} 2 & 0 & 0 & 4 \\ 0 & 4 & 2 & 0 \\ 0 & 0 & 2 & 4 \\ 0 & 0 & 0 & 9 \end{pmatrix}$.

3.2 Encuentre los valores característicos de $A = \begin{pmatrix} 8 & 0 & 3 & 1 \\ 4 & 2 & 3 & 2 \\ 0 & 0 & 8 & 0 \\ 0 & 0 & 0 & 2 \end{pmatrix}$.

3.3 Sea V un espacio vectorial de dimensión finita sobre un campo K . Pruebe que $\rho \in \text{Hom}_K(V, V)$ es invertible si, y sólo si, el término constante de $m_\rho(\lambda)$ de ρ no es cero.

3.4 Sea V en espacio vectorial de dimensión finita sobre un campo K . Pruebe que:

- (a) si $\rho \in \text{Hom}_K(V, V)$ es invertible, entonces ρ^{-1} es una expresión polinomial en ρ sobre K y
- (b) si ρ es singular, entonces existe $\eta \in \text{Hom}_K(V, V)$, $\eta \neq 0$ tal que $\rho\eta = \eta\rho = 0$.

3.5 Pruebe que una matriz cuadrada A es diagonalizable si, y sólo si, $m_A(\lambda) = (\lambda - \lambda_1)(\lambda - \lambda_2) \cdots (\lambda - \lambda_r)$ donde $\lambda_1, \dots, \lambda_r$ son los valores característicos distintos de A .

II.4 FORMA CANONICA TRIANGULAR

En 1.1 definimos el concepto de similaridad de matrices, (el cual se traduce en uno para operadores), y en 1.2 se vio que era una relación de equivalencia. El problema que consideraremos será el de cómo distinguir las clases de equivalencia, o bien, dicho de otra manera, cómo determinamos si dos operadores lineales son similares. Lo resolveremos definiendo ciertas matrices llamadas formas canónicas, una para cada clase de equivalencia. Formalmente, definimos un conjunto de *formas canónicas* para una relación de equivalencia \sim en un conjunto C como un subconjunto F de C que consiste de exactamente un elemento de cada clase de equivalencia de \sim . Entonces, una vez obtenidas las formas canónicas, bastará comparar si éstas son las mismas o no para cada operador. Comenzaremos estudiando el importante concepto de espacio cociente que utilizaremos en esta sección y en el resto del texto.

4.1 NOTACION. Sea V un espacio vectorial sobre un campo K . Sea U un subespacio de V y $v \in V$. Denotaremos con $v + U$ el conjunto $\{v + u \mid u \in U\}$. Dichos elementos $v + U$ los llamaremos *clases laterales* de U en V .

Como $0 \in U$ y $v = v + 0 \in v + U$, cada $v \in V$ pertenece a una clase lateral. Es inmediato comprobar que cualesquiera dos clases laterales son ajenas o son iguales (problema 4.1).

Sea V/U el conjunto de todas las clases laterales de U en V . Démosle a V/U una estructura de espacio vectorial mediante

$$\begin{aligned} +: V/U \times V/U &\longrightarrow V/U \text{ dada por} \\ ((v + U), (w + U)) &\longmapsto ((v + w) + U) \quad y \end{aligned}$$

$$\begin{aligned} \mu: K \times V/U &\longrightarrow V/U \text{ dada por} \\ (\lambda, v + U) &\longmapsto \lambda v + U. \end{aligned}$$

Es fácil comprobar que las operaciones anteriores están bien definidas (problema 4.2) y que definen una estructura de espacio vectorial en V/U . Llamaremos a V/U , *espacio cociente*.

Nota: Sea U un subespacio del espacio vectorial V sobre un campo K . Si $u \in v + U$ entonces existe $w \in U$ tal que $u = v + w$. Luego $u - v = w \in U$. Si $u - v \in U$ entonces $u - v = w \in U$. Luego $u = v + w \in v + U$. También

$u - v \in U \iff -(u - v) = v - u \in U \iff v \in u + U$. En resumen,
 $u \in v + U \iff u - v \in U \iff v \in u + U$.

Sea $p: V \rightarrow V/U$ dada por $v \mapsto v + U$. Si $v, w \in V$, entonces $p(v + w) = (v + w) + U = (v + U) + (w + U) = p(v) + p(w)$. Si $\lambda \in K$, $p(\lambda v) = \lambda v + U = \lambda(v + U) = \lambda p(v)$. Por lo tanto, p es una aplicación lineal llamada *proyección canónica*.

4.2 DEFINICION. Sea V un espacio vectorial sobre un campo K y U un subespacio de V . Decimos que U es *invariante bajo un operador lineal* $\rho: V \rightarrow V$ si $\rho(U) \subset U$.

Observe que ρ define un operador lineal de U que denotamos con $\rho|_U$.

4.3 PROPOSICION. Sea $\rho: V \rightarrow V$ un operador lineal de V y U un subespacio de V invariante bajo ρ . Entonces ρ induce un operador lineal $\rho_{V/U}: V/U \rightarrow V/U$ dado por $\rho_{V/U}(v + U) = \rho(v) + U$.

Demostración. Veamos que $\rho_{V/U}$ está bien definido: sea $u + U = u' + U$, entonces $u - u' \in U$ y como U es invariante bajo ρ , $\rho(u - u') = \rho(u) - \rho(u') \in U$. Luego

$$\rho_{V/U}(u + U) = \rho(u) + U = \rho(u') + U = \rho_{V/U}(u' + U).$$

Así, si $u + U = u' + U$ entonces $\rho_{V/U}(u + U) = \rho_{V/U}(u' + U)$ y, por lo tanto, $\rho_{V/U}$ está bien definida.

Veamos que $\rho_{V/U}$ es lineal:

$$\begin{aligned} \rho_{V/U}((u + U) + (u' + U)) &= \rho_{V/U}(u + u' + U) \\ &= \rho(u + u') + U = \rho(u) + \rho(u') + U \\ &= \rho(u) + U + \rho(u') + U \\ &= \rho_{V/U}(u + U) + \rho_{V/U}(u' + U). \\ \rho_{V/U}(\lambda(u + U)) &= \rho_{V/U}(\lambda u + U) \\ &= \rho(\lambda u) + U \\ &= \lambda \rho(u) + U \\ &= \lambda(\rho(u) + U) \\ &= \lambda \rho_{V/U}(u + U). \blacksquare \end{aligned}$$

4.4 PROPOSICION. Sea $\rho: V \rightarrow V$ un operador lineal. Si f es un polinomio tal que $f(\rho) = 0$ entonces $f(\rho_{V/U}) = 0$.

Demostración. Sea $v + U$ una clase lateral de V/U . Observe que $(\rho^2)_{V/U} = (\rho_{V/U})^2$ pues $(\rho^2)_{V/U}(v + U) = \rho^2(v) + U = \rho(\rho(v)) + U = \rho_{V/U}(\rho(v) + U) = \rho_{V/U}(\rho_{V/U}(v + U)) = (\rho_{V/U})^2(v + U)$. Análogamente, $(\rho^i)_{V/U} = (\rho_{V/U})^i$. Luego, para el polinomio $f = \sum_{i=0}^n \alpha_i t^i$ tenemos que

$$\begin{aligned} [f(\rho)]_{V/U}(v + U) &= f(\rho)(v) + U \\ &= \sum_{i=0}^n \alpha_i \rho^i(v) + U \\ &= \sum_{i=0}^n \alpha_i (\rho^i(v) + U) \\ &= \sum_{i=0}^n \alpha_i (\rho^i)_{V/U}(v + U) \\ &= \sum_{i=0}^n \alpha_i (\rho_{V/U})^i(v + U) \\ &= \left(\sum_{i=0}^n \alpha_i (\rho_{V/U})^i \right) (v + U) \\ &= f(\rho_{V/U})(v + U). \end{aligned}$$

Así que $[f(\rho)]_{V/U} = f(\rho_{V/U})$. Esto es, si ρ es raíz de f entonces $[f(\rho)]_{V/U} = 0_{V/U} = U = f(\rho_{V/U})$, luego $\rho_{V/U}$ es raíz de f . ■

Observe que, debido a la proposición anterior, el polinomio mínimo de $\rho_{V/U}$ divide al polinomio mínimo de ρ .

4.5 PROPOSICION. Sea U un subespacio de un espacio vectorial V sobre un campo K . Sea $\{u_1, \dots, u_t\}$ una base de U y $\{\bar{v}_1, \dots, \bar{v}_r\}$ una base de V/U . Entonces $\{v_1, \dots, v_r, u_1, \dots, u_t\}$ es una base de V y $\dim V = \dim U + \dim V/U$.

Demostración. Sea $v \in V$. Como $\{\bar{v}_i\}$ es una base de V/U , $\bar{v} = v + U = \alpha_1 \bar{v}_1 + \alpha_2 \bar{v}_2 + \dots + \alpha_r \bar{v}_r$. Luego $v = \alpha_1 v_1 + \dots + \alpha_r v_r + u$ donde $u \in U$. Como $\{u_i\}$ es una base de U , $v = \alpha_1 v_1 + \dots + \alpha_r v_r + \beta_1 u_1 + \dots + \beta_t u_t$. Luego $\{v_1, \dots, v_r, u_1, \dots, u_t\}$ genera a V .

Supongamos que $\gamma_1 v_1 + \dots + \gamma_r v_r + \delta_1 u_1 + \dots + \delta_t u_t = 0$. Luego $\gamma_1 \bar{v}_1 + \dots + \gamma_r \bar{v}_r = \bar{0} = U$ pues cada $\bar{v}_i = v_i + U$. Como $\{\bar{v}_i\}$ es linealmente independiente, $\gamma_i = 0$. Luego $\delta_1 u_1 + \dots + \delta_t u_t = 0$ y como las $\{u_i\}$ son linealmente independientes, las $\delta_i = 0$. Luego $\{v_1, \dots, v_r, u_1, \dots, u_t\}$ es linealmente independiente. ■

Sea $\rho \in \text{Hom}_K(V, V)$ un operador lineal de V . Diremos que ρ puede represen-

tarse por una matriz triangular si su matriz asociada es triangular. Si A es la matriz triangular de la forma

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & a_{22} & \cdots & a_{2n} \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & a_{nn} \end{pmatrix},$$

su polinomio característico es

$$p_A(\lambda) = |\lambda I - A| = (\lambda - a_{11})(\lambda - a_{22}) \cdots (\lambda - a_{nn}).$$

El siguiente teorema establece el sentido inverso:

4.6 TEOREMA. *Sea $\rho \in \text{Hom}_K(V, V)$ tal que su polinomio característico se factoriza como producto de polinomios lineales. Entonces existe una base de V para la cual la matriz asociada a ρ es triangular.*

Demostración. Utilizando inducción sobre la dimensión de V tenemos que, si $\dim V = 1$, la matriz asociada a ρ , de 1×1 , es triangular.

Supongamos que $\dim V = n$ con $n > 1$ y que el teorema es válido para dimensiones menores que n . El polinomio característico se factoriza como producto de polinomios lineales por hipótesis, luego ρ posee al menos un valor característico y un vector característico asociado diferente de cero. Sea v dicho vector característico tal que $\rho(v) = a_{11}v$. Sea U el subespacio de dimensión 1 generado por v . Luego, por 4.5, $\dim V/U = \dim V - \dim U = n - 1$. Como U es invariante bajo ρ , por 4.3 ρ induce un operador $\rho_{V/U}$ cuyo polinomio mínimo divide al polinomio mínimo de ρ . Como el polinomio característico de ρ es producto de polinomios lineales, el polinomio mínimo de ρ también es producto de polinomios lineales y por ende lo son los polinomios característicos y mínimos de $\rho_{V/U}$. Luego V/U y $\rho_{V/U}$ satisfacen las hipótesis del teorema. Por inducción, existe una base $\{\bar{v}_2, \dots, \bar{v}_n\}$ de V/U tal que

$$\begin{aligned} \rho_{V/U}(\bar{v}_2) &= a_{22}\bar{v}_2 \\ \rho_{V/U}(\bar{v}_3) &= a_{32}\bar{v}_2 + a_{33}\bar{v}_3 \\ &\vdots \\ \rho_{V/U}(\bar{v}_n) &= a_{n2}\bar{v}_2 + a_{n3}\bar{v}_3 + \cdots + a_{nn}\bar{v}_n. \end{aligned}$$

Sea $v_i \in \bar{v}_i$. Por 4.5, $\{v, v_2, \dots, v_n\}$ es una base de V . Como $\rho_{V/U}(\bar{v}_2) = a_{22}\bar{v}_2$, $\rho_{V/U}(\bar{v}_2) - a_{22}\bar{v}_2 = 0$. Luego $\rho(v_2) - a_{22}v_2 \in U$. Como U está generado por v ,

Demostración. Sea f cualquier polinomio. Supongamos que el grado n de f es mayor que 1, pues si es de grado cero o uno es válida la afirmación. Si f es de la forma $\alpha_n x^n + \alpha_{n-1} x^{n-1} + \cdots + \alpha_1 x^1 + \alpha_0$ y el resultado vale para grados menores que n , entonces $f(\rho|_U)(u) = (\alpha_n \rho|_U^n + \cdots + \alpha_1 \rho|_U + \alpha_0 I)(u) = (\alpha_n \rho|_U^{n-1})(\rho|_U(u)) + (\alpha_{n-1} \rho|_U^{n-1} + \cdots + \alpha_0 I)(u) = (\alpha_n \rho^{n-1})(\rho(u)) + (\alpha_{n-1} \rho^{n-1} + \cdots + \alpha_0 I)(u) = f(\rho)(u)$. Si $m_\rho(\lambda)$ denota al polinomio mínimo de ρ , entonces $m_\rho(\rho|_U)(u) = m_\rho(\rho)(u) = 0(u) = 0 \forall u \in U$. Luego $\rho|_U$ es raíz del polinomio mínimo de ρ y, por la proposición 3.2, el polinomio mínimo de $\rho|_U$ divide a $m_\rho(\lambda)$. ■

PROBLEMAS

4.1 Pruebe que cualesquiera clases laterales o son ajenas o son iguales.

4.2 Compruebe que las operaciones definidas en V/U están bien definidas y que proporcionan una estructura de espacio vectorial en V/U .

4.3 Sea U un subespacio de un espacio vectorial V sobre un campo K . Pruebe que si el conjunto de clases laterales $\{\bar{v}_1, \dots, \bar{v}_n\}$ de V/U es linealmente independiente entonces $\{v_1, \dots, v_n\}$ es linealmente independiente.

4.4 Describa las clases laterales de U en \mathbb{R}^3 donde U es el espacio solución de la ecuación $8x + 2y + 5z = 0$.

4.5 Pruebe que si $\rho \in \text{Hom}_K(V, V)$ es un operador lineal y f un polinomio cualquiera, entonces el núcleo de $f(\rho)$ es invariante bajo ρ .

4.6 Proporcione con todo detalle la demostración de la Proposición 4.8.

4.7 Compruebe que si A es una matriz cuadrada cuyo polinomio característico se factoriza en polinomios lineales entonces A es similar a una matriz triangular. (Sugerencia: Teorema 4.6)

4.8 Pruebe que si $f: V \rightarrow W$ es una función lineal entre espacios vectoriales sobre un campo K , entonces existe una función lineal única $h: V/\ker f \rightarrow W$ tal que $h \circ p = f$. Además, h es inyectiva y $f(V) = \text{im} f \cong V/\ker f$.

4.9 Sea $f: U \rightarrow V$ una función lineal. Pruebe que existen bases de U y de V tales que la matriz asociada a f es de la forma $\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}$ donde I_r es matriz identidad y r es el rango de f . Este es un ejemplo de *forma canónica* llamada *normal* de f bajo la relación de equivalencia siguiente: una matriz A es *equivalente* a una matriz B si, y sólo si, el rango de A es igual al rango de B .

4.10 Sea $I_r^{m,n}$ la matriz de $m \times n$ cuyos primeros r renglones son la base canónica de K^r y cuyos renglones restantes son cero, i.e., $I_r^{m,n} = \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}$. Pruebe que, si $A, B \in M_{m \times n}K$ entonces A es equivalente a $I_r^{m,n}$ si, y sólo si, el rango de A es r . Concluya que A es equivalente a B si, y sólo si, sus rangos son iguales (véase el problema 1.4) y que las matrices $I_r^{m,n}$ ($r = 1, 2, \dots, \min(m, n)$) son un conjunto de formas canónicas para la relación de equivalencia en $M_{m \times n}K$.

4.11 Sea $\rho \in \text{Hom}_K(V, V)$ tal que $\rho = \rho_1 \oplus \rho_2$ con respecto a la suma $V = U_1 \oplus U_2$ donde U_i ($i = 1, 2$) es invariante bajo ρ . Pruebe que $p_\rho(\lambda) = p_{\rho_1}(\lambda)p_{\rho_2}(\lambda)$ donde p_ρ, p_{ρ_1} y p_{ρ_2} son los polinomios característicos.

II.5 FORMA CANONICA DE JORDAN

5.1 DEFINICION. Sea $\rho \in \text{Hom}_K(V, V)$ y $U_i, i = 1, \dots, s$ subespacios de V . Diremos que ρ (es *descomponible*) se puede descomponer como suma directa de operadores $\rho|_{U_i}$, si $V = \bigoplus_{i=1}^s U_i$ con U_i invariante bajo ρ . En este caso escribiremos $\rho = \bigoplus_{i=1}^s \rho|_{U_i}$.

5.2 LEMA. Sea $\rho \in \text{Hom}_K(V, V)$ tal que $\rho = \rho_1 \oplus \rho_2$ con respecto a la suma $V = U_1 \oplus U_2$ donde $U_i, i = 1, 2$ es invariante bajo ρ . Entonces $m_\rho(\lambda)$ es el m.c.m. de $m_{\rho_1}(\lambda)$ y de $m_{\rho_2}(\lambda)$.

Demostración. Por el problema 4.6, $m_{\rho_1}(\lambda)|m_\rho(\lambda)$ y $m_{\rho_2}(\lambda)|m_\rho(\lambda)$. Sea h un polinomio múltiplo de m_{ρ_1} y m_{ρ_2} . Luego $h(\rho_1) = 0$ y $h(\rho_2) = 0$. Sea $v \in V$ tal que $v = u_1 + u_2, u_1 \in U_1, u_2 \in U_2$. Entonces

$$h(\rho)(v) = h(\rho)(u_1) + h(\rho)(u_2) = 0.$$

Por lo tanto, ρ es raíz de h . Luego, por 3.2, $m_\rho(\lambda)|h(\lambda)$ y, por lo tanto, $m_\rho(\lambda)$ es el mínimo común múltiplo. ■

5.3 PROPOSICION. Sea $\rho \in \text{Hom}_K(V, V)$. Sean f, g y h polinomios tales que $f(\rho) = 0, f = gh$ y $(g, h) = 1$. Entonces $V = \ker g(\rho) \oplus \ker h(\rho)$. Aún más, si f es el polinomio mínimo de ρ entonces g y h son los polinomios mínimos de $\rho|_{\ker g(\rho)}$ y de $\rho|_{\ker h(\rho)}$ respectivamente.

Demostración. Por el problema 4.5, $\ker g(\rho)$ y $\ker h(\rho)$ son invariantes bajo ρ . Como g y h son primos relativos

$$rg + sh = 1$$

para ciertos polinomios r y s . Así que,

$$r(\rho)g(\rho) + s(\rho)h(\rho) = I$$

el cual, calculado en un elemento $v \in V$ nos da

$$r(\rho)g(\rho)(v) + s(\rho)h(\rho)(v) = v.$$

Como $h(\rho)r(\rho)g(\rho)(v) = r(\rho)g(\rho)h(\rho)(v) = r(\rho)f(\rho)(v) = r(\rho)0(v) = 0$, $r(\rho)g(\rho)(v) \in \ker h(\rho)$. Análogamente $g(\rho)s(\rho)h(\rho)(v) = s(\rho)g(\rho)h(\rho)(v) = s(\rho)f(\rho)(v) = s(\rho)0(v) = 0$ y $s(\rho)h(\rho) \in \ker g(\rho)$. Luego $V = \ker g(\rho) + \ker h(\rho)$.

Ahora, necesitamos probar que si $v = v' + v''$ con $v' \in \ker g(\rho)$ y $v'' \in \ker h(\rho)$, v' y v'' están determinadas en forma única por v : aplicamos $r(\rho)g(\rho)$ a $v = v' + v''$ y como $g(\rho)v' = 0$ tenemos que $r(\rho)g(\rho)(v) = r(\rho)g(\rho)(v') + r(\rho)g(\rho)(v'') = r(\rho)g(\rho)(v'')$. También $v'' = I(v'') = r(\rho)g(\rho)(v'') + s(\rho)h(\rho)(v'') = r(\rho)g(\rho)(v'')$ pues $h(\rho)(v'') = 0$. Las fórmulas anteriores nos dan $v'' = r(\rho)g(\rho)(v)$ y así v'' está determinada en forma única por v . De manera semejante, v' está determinada en forma única por v ; luego V es suma directa de $\ker g(\rho)$ y de $\ker h(\rho)$.

Sean m_g y m_h los polinomios mínimos de $\rho|_{\ker g(\rho)}$ y de $\rho|_{\ker h(\rho)}$ respectivamente. Sabemos que $g(\rho|_{\ker g(\rho)}) = 0$ y que $h(\rho|_{\ker h(\rho)}) = 0$. Luego $m_g|g$ y $m_h|h$. Por 5.2, f es el m.c.m. de m_g y m_h . Pero m_g y m_h son primos relativos puesto que g y h lo son. Luego $f = m_g m_h$. Pero $f = gh$. Esto implica que $g = m_g$ y $h = m_h$. ■

5.4 TEOREMA. (Descomposición primaria). *Sea $\rho \in \text{Hom}_K(V, V)$ con polinomio mínimo*

$$m_\rho(\lambda) = f_1(\lambda)^{\eta_1} f_2(\lambda)^{\eta_2} \cdots f_s(\lambda)^{\eta_s}$$

donde los $f_i(\lambda)$ son polinomios mónicos e irreducibles distintos. Entonces $V = \bigoplus_{i=1}^s \ker f_i(\rho)^{\eta_i}$ donde los subespacios $\ker f_i(\rho)^{\eta_i}$ son invariantes bajo ρ y $f_i(\lambda)^{\eta_i}$ es el polinomio mínimo de $\rho|_{\ker f_i(\rho)^{\eta_i}}$.

Demostración. Utilizando inducción sobre s , tenemos que para $s = 1$, el resultado es trivial. Supongamos que el teorema es válido para $s - 1$. Por la proposición 5.3, $V = \ker f_1(\rho)^{\eta_1} \oplus \ker (f_2(\rho)^{\eta_2} \cdots f_s(\rho)^{\eta_s})$ y los polinomios mínimos de $\rho|_{\ker f_1(\rho)^{\eta_1}}$ y de $\rho|_{\ker f_2(\rho)^{\eta_2} \cdots f_s(\rho)^{\eta_s}}$ son $f_1(\lambda)^{\eta_1}$ y $f_2(\lambda)^{\eta_2} \cdots f_s(\lambda)^{\eta_s}$ respectivamente. Sea $\rho_1 = \rho|_{\ker f_2(\rho)^{\eta_2} \cdots f_s(\rho)^{\eta_s}}$. Por hipótesis de inducción, $\ker (f_2(\rho)^{\eta_2} \cdots f_s(\rho)^{\eta_s}) = \bigoplus_{i=2}^s \ker f_i(\rho_1)^{\eta_i}$ tal que $f_i(\lambda)^{\eta_i}$ es el polinomio mínimo de la restricción de ρ_1 a $\ker f_i(\rho_1)^{\eta_i}$. Como $f_i(\lambda)^{\eta_i}$ divide a $f_2(\lambda)^{\eta_2} \cdots f_s(\lambda)^{\eta_s}$, $\ker f_i(\rho)^{\eta_i} \subset \ker (f_2(\rho)^{\eta_2} \cdots f_s(\rho)^{\eta_s})$, $i = 2, \dots, s$. Luego, $\ker f_i(\rho)^{\eta_i} = \ker f_i(\rho_1)^{\eta_i}$.

Como $\rho|_{\ker f_i(\rho_1)^{\eta_i}} = \rho_1|_{\ker f_i(\rho_1)^{\eta_i}}$ para $i = 2, \dots, s$, $f_i(\lambda)^{\eta_i}$ es también el polinomio mínimo de $\rho|_{\ker f_i(\rho_1)^{\eta_i}}$. Luego $V = \ker f_1(\rho)^{\eta_1} \oplus \ker f_2(\rho)^{\eta_2} \oplus \cdots \oplus \ker f_s(\rho)^{\eta_s}$ y la descomposición de ρ es $\rho = \bigoplus_{i=1}^s \rho|_{\ker f_i(\rho)^{\eta_i}}$. ■

5.5 COROLARIO. *Un operador $\rho \in \text{Hom}_K(V, V)$ posee una matriz asociada diagonal si, y sólo si, su polinomio mínimo $m_\rho(\lambda)$ es producto de polinomios lineales distintos.*

Demostración. Sea ρ un operador lineal cuya matriz asociada es diagonal. Entonces V posee una base que consiste de vectores característicos de ρ con valores característicos distintos $\lambda_1, \dots, \lambda_r$. Por consiguiente el operador $f(\rho) = (\rho - \lambda_1 I)(\rho - \lambda_2 I) \cdots (\rho - \lambda_r I)$ envía cualquier vector de la base al cero (véase el problema 5.3). Luego $f(\rho) = 0$ y, por lo tanto, el polinomio mínimo $m(\lambda)$ de ρ divide al polinomio $f(\rho) = (\lambda - \lambda_1)(\lambda - \lambda_2) \cdots (\lambda - \lambda_r)$. Así, $m(\lambda)$ es producto de polinomios lineales distintos.

Ahora, supongamos que $m(\lambda) = (\lambda - \lambda_1)(\lambda - \lambda_2) \cdots (\lambda - \lambda_s)$ con las $\lambda_i \in K$ diferentes. Por el teorema 5.4, $V = \bigoplus_{i=1}^s \ker(\rho - \lambda_i I)$. Sea $v \in \ker(\rho - \lambda_i I)$, luego $(\rho - \lambda_i I)(v) = 0$, i.e., $\rho(v) = \lambda_i v$. Esto significa que cada vector de $\ker(\rho - \lambda_i I)$ es un vector característico perteneciente al valor característico λ_i . Por el problema I.3.10, la unión de las bases de $\ker(\rho - \lambda_i I)$ es una base de V que consiste de vectores característicos. Luego ρ es diagonalizable. ■

A continuación estudiaremos operadores lineales cuyas raíces de su polinomio mínimo son todas cero. Sabemos entonces, por el teorema 4.6, que existe una base del espacio vectorial tal que la matriz asociada a ρ es triangular. Sin embargo, el encontrar formas canónicas para dichos operadores nilpotentes nos permitirán encontrar formas canónicas para cualquier operador que se factorice como producto de polinomios lineales.

5.6 DEFINICION. Un operador lineal $\rho \in \text{Hom}_K(V, V)$ se llama *nilpotente* si $\rho^n = 0$ para alguna $n > 0$. Llamaremos al entero r *índice de nilpotencia* de ρ si $\rho^r = 0$ pero $\rho^{r-1} \neq 0$. También diremos que una matriz cuadrada A es *nilpotente* si $A^n = 0$ y r es el *índice de nilpotencia* de A si $A^r = 0$ pero $A^{r-1} \neq 0$.

Observación. El polinomio mínimo de un operador nilpotente de índice r es $m(\lambda) = \lambda^r$. Su único valor característico es el cero. (Problema 5.7).

5.7 PROPOSICION. *Sea $\rho \in \text{Hom}_K(V, V)$ un operador nilpotente de índice r , y $v \in V$ tal que $\rho^{r-1}(v) \neq 0$. Entonces el conjunto $\{\rho^{r-1}(v), \rho^{r-2}(v), \dots, \rho(v), v\}$ es una base del subespacio que genera, cuya matriz aso-*

ciada posee índice de nilpotencia r y es de la forma

$$\begin{pmatrix} 0 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & 1 \\ 0 & 0 & 0 & \cdots & 0 & 0 \end{pmatrix}.$$

Demostración. Por el problema 5.4, el conjunto $\{\rho^{r-1}(v), \dots, v\}$ es linealmente independiente y el subespacio que genera es invariante bajo ρ . Por el problema 5.5, como ρ es nilpotente de índice r , $(\rho')^r(\rho^k(v)) = \rho^{r+k}(v) = 0$ luego

$$\begin{array}{rcl} \rho'(\rho^{r-1}(v)) & = & \rho^r(v) = 0 \\ \rho'(\rho^{r-2}(v)) & = & \rho^{r-1}(v) \\ \rho'(\rho^{r-3}(v)) & = & \rho^{r-2}(v) \\ \vdots & & \ddots \\ \rho'(\rho(v)) & = & \rho^2(v) \\ \rho'(v) & = & \rho(v) \end{array}$$

Por lo tanto, la matriz asociada es la requerida. Es inmediato comprobar que es de índice r . ■

5.8 LEMA. Sea $\rho \in \text{Hom}_K(V, V)$. Entonces $\ker \rho^k \subset \ker \rho^{k+1}$ y $\rho(\ker \rho^{k+1}) \subset \ker \rho^k$.

Demostración. Sea $v \in \ker \rho^k$, luego $\rho^k(v) = 0$ y $\rho^{k+1}(v) = \rho(\rho^k(v)) = \rho(0) = 0$. Por lo tanto $v \in \ker \rho^{k+1}$ y, como v es arbitraria, $\ker \rho^k \subset \ker \rho^{k+1}$. Si $v \in \ker \rho^{k+1}$ entonces $\rho^{k+1}(v) = 0$. Queremos ver que $\rho(v) \in \ker \rho^k$ i.e. $\rho^k(\rho(v)) = 0$. Pero como $\rho^{k+1}(v) = 0$ y $\rho^k(\rho(v)) = \rho^{k+1}(v) = 0$ hemos terminado. ■

5.9 PROPOSICION. Sea $\rho \in \text{Hom}_K(V, V)$ nilpotente de índice r . Entonces ρ posee una matriz asociada diagonal por bloques que son de la forma

$$\begin{pmatrix} 0 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & 1 \\ 0 & 0 & 0 & \cdots & 0 & 0 \end{pmatrix}.$$

Demostración. Sea $n = \dim V$, $m_i = \dim \ker \rho^i$, $i = 1, \dots, r$. Como $\rho^r = 0$, $V = \ker \rho^r$ y como $\rho^{r-1} \neq 0$, $\ker \rho^{r-1} \neq V$. Luego $m_{r-1} < m_r = n$. Por el lema

5.8, $\ker \rho^1 \subset \ker \rho^2 \subset \dots \subset \ker \rho^r = V$. Por inducción, podemos escoger una base $\{v_1, \dots, v_n\}$ de V tal que $\{v_1, \dots, v_{m_i}\}$ es una base de $\ker \rho^i$.

Escojamos una base nueva de V para la cual ρ posee la forma deseada: pongámosla con cierto orden;

$$\begin{aligned} u_{(1,r)} &= v_{m_{r-1}+1} \\ u_{(2,r)} &= v_{m_{r-1}+2} \\ &\vdots \\ u_{(m_r-m_{r-1},r)} &= v_{m_r} \quad \text{y} \end{aligned}$$

$$\begin{aligned} u_{(1,r-1)} &= \rho u_{(1,r)} \\ u_{(2,r-1)} &= \rho u_{(2,r)} \\ &\vdots \\ u_{(m_r-m_{r-1},r-1)} &= \rho u_{(m_r-m_{r-1},r)}. \end{aligned}$$

Por el problema 5.6, $X_1 = \{v_1, \dots, v_{m_{r-2}}, u_{(1,r-1)}, \dots, u_{(m_r-m_{r-1},r-1)}\}$ es un subconjunto linealmente independiente de $\ker \rho^{r-1}$. Extendemos X_1 a una base de $\ker \rho^{r-1}$ adjuntando nuevos elementos como

$$\begin{aligned} &u_{(m_r-m_{r-1}+1,r-1)} \\ &u_{(m_r-m_{r-1}+2,r-1)} \\ &\vdots \\ &u_{(m_{r-1}-m_{r-2},r-1)}. \end{aligned}$$

Sea

$$\begin{aligned} u_{(1,r-2)} &= \rho u_{(1,r-1)} \\ u_{(2,r-2)} &= \rho u_{(2,r-1)} \\ &\vdots \\ u_{(m_{r-1}-m_{r-2},r-2)} &= \rho u_{(m_{r-1}-m_{r-2},r-1)}. \end{aligned}$$

Por el problema 5.6 se tiene que

$$X_2 = \{v_1, \dots, v_{m_{r-3}}, u_{(1,r-2)}, \dots, u_{(m_{r-1}-m_{r-2},r-2)}\}$$

es un subconjunto linealmente independiente de $\ker \rho^{r-2}$ el cual podemos extender a una base de $\ker \rho^{r-2}$ adjuntándole elementos de la forma

$$\begin{aligned} &u_{(m_{r-1}-m_{r-2}+1,r-2)} \\ &u_{(m_{r-1}-m_{r-2}+2,r-2)} \\ &\vdots \\ &u_{(m_{r-2}-m_{r-3},r-2)}. \end{aligned}$$

Si seguimos con este proceso obtendremos una base para V rearmada como

$$\begin{aligned} &u_{(1,r)}, \dots, u_{(m_r - m_{r-1}, r)} \\ &u_{(1,r-1)}, \dots, u_{(m_r - m_{r-1}, r-1)}, \dots, u_{(m_{r-1} - m_{r-2}, r-1)} \\ &\quad \vdots \\ &u_{(1,2)}, \dots, u_{(m_r - m_{r-1}, 2)}, \dots, u_{(m_{r-1} - m_{r-2}, 2)}, \dots, u_{(m_2 - m_1, 2)} \\ &u_{(1,1)}, \dots, u_{(m_r - m_{r-1}, 1)}, \dots, u_{(m_{r-1} - m_{r-2}, 1)}, \dots, u_{(m_2 - m_1, 1)}, \dots, u_{(m_1, 1)} \end{aligned}$$

El último renglón es una base para $\ker \rho^1$. El último y el penúltimo renglón son una base para $\ker \rho^2$ y así sucesivamente. Obsérvese que cada vector se transforma bajo ρ en el vector inmediato inferior, o en cero si el vector es del último renglón, i.e.,

$$\rho u_{(i,j)} = \begin{cases} u_{(i,j-1)} & \text{si } j > 1 \\ 0 & \text{si } j = 1. \end{cases}$$

Por la proposición 5.7, ρ tendrá la forma requerida si las $u_{(i,j)}$ se ordenan lexicográficamente, comenzando por $u_{(1,1)}$ y subiendo por la primera columna hasta $u_{(1,r)}$; luego brincando a $u_{(2,1)}$ hasta arriba, etcétera. ■

5.10 COROLARIO. *En la proposición 5.9, existe al menos una matriz de orden r , las otras son de órdenes menores o iguales a r . El número de matrices de cada orden posible está determinado en forma única por ρ y el número de matrices de todos los órdenes es igual a la dimensión de $\ker \rho$.*

Demostración. Hay $m_r - m_{r-1}$ elementos de la diagonal de orden r ; $(m_{r-1} - m_{r-2}) - (m_r - m_{r-1}) = 2m_{r-1} - m_r - m_{r-2}$ elementos de la diagonal de orden $r - 1$; ... , $2m_2 - m_1 - m_3$ elementos de la diagonal de orden 2 y $2m_1 - m_2$ elementos de la diagonal de orden 1.

Como los números m_1, \dots, m_r están determinados en forma única por ρ , el número de elementos diagonales de cada orden están determinados en forma única por ρ .

Como $m_1 = (m_r - m_{r-1}) + (2m_{r-1} - m_r - m_{r-2}) + \dots + (2m_2 - m_1 - m_3) + (2m_1 - m_2)$, $m_1 = \dim \ker \rho^1$ es el número de elementos de la diagonal de ρ . ■

Observe que, por 5.7 la matriz de 5.9 es nilpotente de índice igual a su orden.

Las formas canónicas triangular y de Jordan para un operador lineal ρ existen si, y sólo si, el polinomio característico de ρ posee todas sus raíces en un campo base K , (lo cual es cierto si $K = \mathbb{C}$ pero no lo es si $K = \mathbb{R}$ en general). Sin embargo, si esto no sucede, siempre podemos obtener un campo K adjuntando raíces de tal forma que los polinomios mínimo y característico se factoricen en factores lineales y entonces podríamos decir que todo operador posee una forma canónica de Jordan y que toda matriz es similar a una matriz en forma canónica de Jordan. También vimos que una clase de transformaciones que tienen todas sus raíces características en un campo K fueron las nilpotentes, luego siempre se pueden poner en forma triangular.

PROBLEMAS

5.1 Sea $\rho \in \text{Hom}_K(V, V)$ y $V = \bigoplus_{i=1}^s U_i$ con U_i invariante bajo ρ . Pruebe que si la matriz asociada a $\rho|_{U_i}$ es X_i entonces ρ posee una matriz asociada por bloques de la forma

$$X = \begin{pmatrix} X_1 & 0 & \cdots & 0 \\ 0 & X_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & X_s \end{pmatrix}$$

que denotaremos como $X = \bigoplus_{i=1}^s X_i$.

5.2 Sea M una matriz cuadrada diferente de I tal que $M^5 = I$. Compruebe si M es similar a una matriz diagonal con los coeficientes de M reales o complejos.

5.3 Pruebe con detalle que el operador $f(\rho)$ dado en la demostración de 5.5 envía cualquier vector de la base al cero.

5.4 Sea $\rho \in \text{Hom}_K(V, V)$ un operador nilpotente de índice r , y $v \in V$ tal que $\rho^{r-1}(v) \neq 0$. Pruebe que el conjunto $\{v, \rho(v), \rho^2(v), \dots, \rho^{r-1}(v)\}$ es linealmente independiente y genera un subespacio invariante bajo ρ .

5.5 Pruebe que la restricción ρ' de ρ al subespacio generado por el conjunto $\{v, \rho(v), \dots, \rho^{r-1}(v)\}$ del problema 5.4 es nilpotente de índice r .

5.6 Sea $\rho \in \text{Hom}_K(V, V)$, $\{v_1, \dots, v_t\}$ una base de $\ker \rho^{r-2}$, $\{v_1, \dots, v_t,$

u_1, \dots, u_k una base de $\ker \rho^{r-1}$ y $\{v_1, \dots, v_t, u_1, \dots, u_k, w_1, \dots, w_s\}$ una base de $\ker \rho^r$. Pruebe que $\{v_1, \dots, v_t, \rho(w_1), \dots, \rho(w_s)\}$ es un subconjunto linealmente independiente de $\ker \rho^{r-1}$.

5.7 Pruebe que el polinomio mínimo de un operador nilpotente de índice r es $m(\lambda) = \lambda^r$ y que su único valor característico es el cero.

5.8 Pruebe que, en la terminología de 5.11, existe al menos una matriz J_{ij} de orden η_i y que el resto de las J_{ij} poseen órdenes menores o iguales a η_i . También demuestre que la suma de los órdenes de J_{ij} es igual a μ_i .

5.9 Encuentre todas las posibles formas canónicas de Jordan de un operador $\rho \in \text{Hom}_K(V, V)$ cuyo polinomio característico es $p_\rho(\lambda) = (\lambda - 3)^3(\lambda - 7)^2$.

5.10 Si $m_\rho(\lambda) = (\lambda - 7)^2$ es el polinomio mínimo de una matriz de orden 7, encuentre todas las posibles formas canónicas de Jordan.

5.11 Pruebe que dos operadores lineales cualesquiera son similares (proporcione una definición adecuada de similaridad de operadores) si, y sólo si, poseen la misma forma canónica de Jordan.

— |

| —

— |

| —

Capítulo III

FORMAS Y OPERADORES

III.1 FORMAS BILINEALES

1.1 DEFINICION. Sean U , V y W espacios vectoriales sobre un campo K . Una función $f: U \times V \rightarrow W$ se llama *bilineal* si:

- (i) $f(u_1 + u_2, v) = f(u_1, v) + f(u_2, v)$
- (ii) $f(u, v_1 + v_2) = f(u, v_1) + f(u, v_2)$ y
- (iii) $f(\lambda u, v) = \lambda f(u, v) = f(u, \lambda v)$; $u_1, u_2, u \in U$; $v_1, v_2, v \in V$; $\lambda \in K$.

Es decir, $f: U \times V \rightarrow W$ es bilineal si es lineal en cada variable cuando la otra se mantiene fija.

Observación. Lo anterior significa que si $v \in V$ se mantiene fija en $U \times V$, la función $u \mapsto f(u, v)$ es lineal y por lo tanto es un elemento de $\text{Hom}_K(U, W)$. De igual forma, si $u \in U$ se mantiene fija en $U \times V$, la función $v \mapsto f(u, v)$ es lineal y pertenece a $\text{Hom}_K(V, W)$. Esto no significa que f sea lineal como función $f: U \times V \rightarrow W$. Por ejemplo, $f: \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$ dada por $f(u, v) = uv$ es bilineal pero no es lineal. Otro ejemplo, $f: \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$ dada por $f(u, v) = u + v$ es lineal pero no es bilineal.

1.2 EJEMPLO. Sea A una matriz de $m \times n$. Definamos

$$f: K^m \times K^n \longrightarrow K$$

mediante $f(X, Y) = {}^t XAY$. Esto es

$$\begin{aligned} & (x_1, \dots, x_m) \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \\ &= \left(\sum_{i=1}^m x_i a_{i1}, \dots, \sum_{i=1}^m x_i a_{in} \right) \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \\ &= \sum_{j=1}^n \sum_{i=1}^m x_i a_{ij} y_j \\ &= \sum_{j=1}^n \sum_{i=1}^m a_{ij} x_i y_j. \end{aligned}$$

Es inmediato comprobar que $f(X, Y) \in K$ y que es bilineal, ya que las propiedades de las matrices establecen que ${}^t XA(Y + Y') = {}^t XAY + {}^t XAY'$ y ${}^t XA(\lambda Y) = \lambda({}^t XAY)$.

Por ejemplo, si $A = \begin{pmatrix} 2 & 1 & 1 \\ 1 & 3 & 3 \\ 2 & 1 & 1 \end{pmatrix}$, $X = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$ y $Y = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}$ entonces

$$\begin{aligned} f(X, Y) &= (x_1, x_2, x_3) \begin{pmatrix} 2 & 1 & 1 \\ 1 & 3 & 3 \\ 2 & 1 & 1 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} \\ &= (2x_1 + x_2 + 2x_3, x_1 + 3x_2 + x_3, x_1 + 3x_2 + x_3) \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} \\ &= 2x_1y_1 + x_2y_1 + 2x_3y_1 + x_1y_2 + 3x_2y_2 + x_3y_2 + x_1y_3 + 3x_2y_3 + x_3y_3. \end{aligned}$$

Si en 1.1, $W = K$ diremos que f es una *forma bilineal*. Denotamos con $L^2(U, V; K)$ el conjunto de formas bilineales de $U \times V$ en K . Si $U = V$, simplemente denotamos a $L^2(V, V; K)$ con $L^2(V; K)$ o con $Bil(V)$ entendiéndose que se trata de formas bilineales de $V \times V$ en K , que son las que consideraremos en adelante.

A $Bil(V)$ le podemos dar una estructura de espacio vectorial mediante las reglas $(f + g)(u, v) = f(u, v) + g(u, v)$ y $(\lambda f)(u, v) = \lambda f(u, v)$ para $f, g \in Bil(V)$, $\lambda \in K$.

Consideremos el caso en que tengamos el espacio vectorial de homomorfismos $\text{Hom}_K(V, K)$. Sus elementos $f: V \rightarrow K$, que son homomorfismos o aplicaciones lineales, se acostumbra llamarlos *funcionales lineales* o *formas lineales*. También se acostumbra denotar a $\text{Hom}_K(V, K)$ con $L^1(V; K)$ o simplemente V^* y se le llama *espacio dual* de V . Su estructura está dada por

$$\begin{aligned}(f + g)(v) &= f(v) + g(v) \quad y \\ (\lambda f)(v) &= \lambda f(v); \quad v \in V, \lambda \in K.\end{aligned}$$

Aquí vemos a K como espacio vectorial sobre sí mismo.

1.3 EJEMPLO. Sea $V = M_n(K)$ el espacio de las matrices cuadradas de $n \times n$. Sea $f = \text{tr}: M_n(K) \rightarrow K$ dada por $\text{tr}(A) = a_{11} + a_{22} + \cdots + a_{nn}$, i.e., la traza de la matriz $A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix}$. Como $\text{tr}(A + B) = \text{tr}A + \text{tr}B$ y $\text{tr}(\lambda A) = \lambda \text{tr}A$, tr es un funcional.

1.4 EJEMPLO. Sea $V = K^n$. Si $f \in \text{Hom}_K(V, K) = V^*$, f tiene una representación matricial de la forma

$$f(x_1, \dots, x_n) = (a_1, a_2, \dots, a_n) \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = a_1x_1 + a_2x_2 + \cdots + a_nx_n$$

llamada también *forma lineal*.

Sabemos que si $\dim V = n$ entonces $\dim V^* = n$ pues $V^* = \text{Hom}_K(V, K)$ y por el teorema I.5.3 $\dim \text{Hom}_K(V, K) = n \cdot 1 = n$.

Veamos como determinar una base para V^* a partir de una base de V .

1.5 PROPOSICION. Sea $\{v_1, \dots, v_n\}$ una base del espacio vectorial V sobre K . Sean $f_1, \dots, f_n \in \text{Hom}_K(V, K) = V^*$ funcionales dadas por $f_i(v_j) = \delta_{ij}$, donde δ_{ij} es la delta de Kronecker, i.e., $\delta_{ij} = \begin{cases} 1 & \text{si } i = j \\ 0 & \text{si } i \neq j \end{cases}$. Entonces $\{f_i\}_{i=1}^n$ es una base de V^* .

Demostración. Veamos que $\{f_i\}_{i=1}^n$ es linealmente independiente: Supongamos que $\lambda_1 f_1 + \cdots + \lambda_n f_n = 0$. Evaluando en v_1 obtenemos $\lambda_1 f_1(v_1) + \cdots + \lambda_n f_n(v_1) = \lambda_1 \cdot 1 = 0(v_1) = 0$. Luego $\lambda_1 = 0$. Análogamente, evaluando en

v_2, v_3, \dots, v_n obtenemos que $\lambda_2 = \lambda_3 = \dots = \lambda_n = 0$. Luego $\{f_i\}_{i=1}^n$ es linealmente independiente. Como $\dim V^* = n$ y $\{f_i\}_{i=1}^n$ es linealmente independiente, es una base de V^* . Sin embargo veamos directamente que $\{f_i\}_{i=1}^n$ genera a V^* : sea $f \in V^*$ tal que $f(v_i) = \lambda_i$. Sea $\phi = \sum_{i=1}^n \lambda_i f_i$. Luego $\phi(v_1) = \sum_{i=1}^n \lambda_i f_i(v_1) = \lambda_1$, $\phi(v_2) = \lambda_2, \dots, \phi(v_n) = \lambda_n$. Así que $f(v_i) = \phi(v_i)$ para toda $i = 1, \dots, n$. Puesto que f y ϕ son iguales al evaluarlas en los elementos de la base de V , $f = \phi = \sum_{i=1}^n \lambda_i f_i$. Luego $\{f_i\}_{i=1}^n$ genera a V^* . ■

La base de V^* , así obtenida, se llama *base dual*.

1.6 EJEMPLO. Consideremos la base $\{(1, 1), (3, 1)\}$ de \mathbb{R}^2 . Encontramos su base dual para $(\mathbb{R}^2)^* = \text{Hom}_{\mathbb{R}}(\mathbb{R}^2, \mathbb{R})$. Deseamos encontrar funcionales $f_1(x, y) = \alpha x + \beta y$ y $f_2(x, y) = \gamma x + \delta y$ tales que $f_1(1, 1) = 1$, $f_1(3, 1) = 0$, $f_2(1, 1) = 0$, $f_2(3, 1) = 1$. Luego

$$\left. \begin{aligned} f_1(1, 1) &= 1\alpha + 1\beta = 1 \\ f_1(3, 1) &= 3\alpha + 1\beta = 0 \end{aligned} \right\}$$

Resolviendo el sistema obtenemos $\alpha = -\frac{1}{2}$ y $\beta = \frac{3}{2}$.

También

$$\left. \begin{aligned} f_2(1, 1) &= \gamma + \delta = 0 \\ f_2(3, 1) &= 3\gamma + \delta = 1 \end{aligned} \right\}$$

Resolviendo el sistema obtenemos $\gamma = \frac{1}{2}$ y $\delta = -\frac{1}{2}$. Por lo tanto, una base dual es $\left\{ f_1(x, y) = -\frac{1}{2}x + \frac{3}{2}y, f_2(x, y) = \frac{1}{2}x - \frac{1}{2}y \right\}$.

1.7 PROPOSICION. Sea V un espacio vectorial de dimensión finita sobre un campo K . Sea $\{v_i\}_{i=1}^n$ una base de V y $\{f_i\}_{i=1}^n$ su base dual. Entonces

(i) si $v \in V$, v es de la forma

$$v = f_1(v)v_1 + f_2(v)v_2 + \dots + f_n(v)v_n \quad \text{y}$$

(ii) si $f \in V^*$, f es de la forma

$$f = f(v_1)f_1 + f(v_2)f_2 + \dots + f(v_n)f_n.$$

Demostración. (i) Sea $v = a_1v_1 + a_2v_2 + \dots + a_nv_n$. Evaluando $f_i(v) = f_i(a_1v_1 + \dots + a_nv_n) = a_i$ para $i = 1, \dots, n$. Luego $v = f_1(v)v_1 + \dots + f_n(v)v_n$.

(ii) Sea $v = f_1(v)v_1 + \dots + f_n(v)v_n$. Luego $f(v) = f_1(v)f(v_1) + \dots + f_n(v)f(v_n) = f(v_1)f_1(v) + \dots + f(v_n)f_n(v) = (f(v_1)f_1 + \dots + f(v_n)f_n)(v)$ para toda v en V . Así que $f = f(v_1)f_1 + \dots + f(v_n)f_n$. ■

A continuación encontremos una base para $Bil(V)$.

1.8 PROPOSICION. *Sea V un espacio vectorial sobre un campo K de dimensión n . Si $\{f_i\}_{i=1}^n$ es una base para V^* entonces $\{f_{ij}\}_{i,j=1}^n$ dado por $f_{ij}(u, v) = f_i(u)f_j(v)$ es una base para $Bil(V)$.*

Demostración. Sea $\{v_i\}_{i=1}^n$ una base de V , dual de $\{f_i\}_{i=1}^n$. Veamos que $\{f_{ij}\}$ es linealmente independiente: supongamos que $\sum a_{ij}f_{ij} = 0$. Entonces para índices $r, s = 1, \dots, n$ tenemos que $(\sum a_{ij}f_{ij})(v_r, v_s) = \sum a_{ij}f_{ij}(v_r, v_s) = \sum a_{ij}f_i(v_r)f_j(v_s) = \sum a_{ij}\delta_{ir}\delta_{js} = a_{rs} = 0(v_r, v_s) = 0$. Por lo tanto $\{f_{ij}\}$ es linealmente independiente.

Veamos que $\{f_{ij}\}$ genera a $Bil(V)$: sea $f \in Bil(V)$ y $a_{ij} = f(v_i, v_j)$. Basta probar que $f(v_r, v_s) = (\sum a_{ij}f_{ij})(v_r, v_s)$ para $r, s = 1, \dots, n$. Pero como antes, $(\sum a_{ij}f_{ij})(v_r, v_s) = a_{rs} = f(v_r, v_s)$, luego $\{f_{ij}\}$ genera $Bil(V)$. ■

Observe que $\dim Bil(V) = n^2$.

Sea V un espacio vectorial con base $\gamma = \{v_1, \dots, v_n\}$ y sea $f: V \times V \rightarrow K$ una forma bilineal de V . Si $u = \alpha_1v_1 + \dots + \alpha_nv_n$ y $v = \beta_1v_1 + \dots + \beta_nv_n$ son vectores de V ,

$$\begin{aligned} f(u, v) &= f\left(\sum_{i=1}^n \alpha_i v_i, \sum_{j=1}^n \beta_j v_j\right) \\ &= \alpha_1\beta_1 f(v_1, v_1) + \alpha_1\beta_2 f(v_1, v_2) + \dots + \alpha_n\beta_n f(v_n, v_n) \\ &= \sum_{i,j=1}^n \alpha_i\beta_j f(v_i, v_j). \end{aligned}$$

Sea $A = (a_{ij})$ la matriz cuadrada tal que $a_{ij} = f(v_i, v_j)$; luego

$$\begin{aligned} f(u, v) &= \sum_{i,j=1}^n \alpha_i\beta_j a_{ij} \\ &= (\alpha_1, \dots, \alpha_n)A \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_n \end{pmatrix} \\ &= {}^t[u]_\gamma A [v]_\gamma. \end{aligned}$$

Llamaremos a A *matriz asociada a la forma bilineal f con respecto a la base γ* . A menudo denotamos a A como $[f]_\gamma$.

1.9 EJEMPLO. Sea $f: \mathbb{R}^2 \times \mathbb{R}^2 \rightarrow \mathbb{R}$ una forma bilineal dada por $f((\alpha_1, \alpha_2), (\beta_1, \beta_2)) = 4\alpha_2\beta_2$ y $\gamma = \{\gamma_1, \gamma_2\} = \{(1, 1), (3, 1)\}$ una base de \mathbb{R}^2 . Calculemos la matriz asociada a f con respecto a γ , i.e., $A = (a_{ij})$ donde $a_{ij} = f(\gamma_i, \gamma_j)$

$$a_{11} = f(\gamma_1, \gamma_1) = f((1, 1), (1, 1)) = 4 \cdot 1 \cdot 1 = 4$$

$$a_{12} = f(\gamma_1, \gamma_2) = f((1, 1), (3, 1)) = 4 \cdot 1 \cdot 1 = 4$$

$$a_{21} = f(\gamma_2, \gamma_1) = 4$$

$$a_{22} = f(\gamma_2, \gamma_2) = 4$$

Luego $A = \begin{pmatrix} 4 & 4 \\ 4 & 4 \end{pmatrix}$.

1.10 EJEMPLO. Sea f como en el ejemplo anterior. Calculemos la matriz B asociada a f con respecto a la base $\gamma' = \{\gamma'_1, \gamma'_2\} = \{(2, 1), (1, -1)\}$:

$$b_{11} = f(\gamma'_1, \gamma'_1) = f((2, 1), (2, 1)) = 4$$

$$b_{12} = f(\gamma'_1, \gamma'_2) = f((2, 1), (1, -1)) = -4$$

$$b_{21} = f(\gamma'_2, \gamma'_1) = f((1, -1), (2, 1)) = -4$$

$$b_{22} = f(\gamma'_2, \gamma'_2) = f((1, -1), (1, -1)) = 4$$

Luego $B = \begin{pmatrix} 4 & -4 \\ -4 & 4 \end{pmatrix}$.

Ahora, calculemos la matriz de transición N de la base γ a la base γ' del ejemplo 1.9:

$$\gamma'_1 = (2, 1) = \lambda(1, 1) + \mu(3, 1) \implies \lambda = \frac{1}{2} = \mu$$

$$\gamma'_2 = (1, -1) = \eta(1, 1) + \delta(3, 1) \implies \eta = -2, \delta = 1.$$

Luego $N = \begin{pmatrix} 1/2 & -2 \\ 1/2 & 1 \end{pmatrix}$.

Observe que ${}^tNAN = \begin{pmatrix} 1/2 & 1/2 \\ -2 & 1 \end{pmatrix} \begin{pmatrix} 4 & 4 \\ 4 & 4 \end{pmatrix} \begin{pmatrix} 1/2 & -2 \\ 1/2 & 1 \end{pmatrix} = \begin{pmatrix} 4 & -4 \\ -4 & 4 \end{pmatrix} = B$.

Establezcamos la observación del ejemplo 1.10 en el siguiente teorema:

1.11 TEOREMA. Sea $f: V \times V \rightarrow K$ una forma bilineal. Si N es la matriz de transición de una base γ a una base γ' de V entonces la matriz B asociada a f con respecto a la base γ' es

$$B = {}^tNAN$$

i.e.
$$B = \begin{pmatrix} \beta_{11} & \cdots & \beta_{n1} \\ \vdots & & \vdots \\ \beta_{1n} & \cdots & \beta_{nn} \end{pmatrix} = {}^t(\beta_{ij}).$$

Deseamos ver que $B = {}^tN^{-1}$. Para ello, veamos que $B^tN = I$, y así tendremos que $B = {}^tN^{-1}$. Pero

$$\begin{aligned} \delta_{ij} = g_i(v_j) &= (\beta_{i1}f_1 + \cdots + \beta_{in}f_n)(\alpha_{j1}u_1 + \cdots + \alpha_{jn}u_n) \\ &= \sum_k \sum_\ell \beta_{ik}\alpha_{j\ell}f_k(u_\ell) \\ &= \sum_{k=1}^n \beta_{ik}\alpha_{jk} \\ &= \sum_{k=1}^n \beta_{ik}\alpha'_{kj} \quad (\alpha'_{kj} = \alpha_{jk}). \end{aligned}$$

Luego $B^tN = I$. ■

¿Qué sucede cuando consideramos el espacio dual de V^* ? Lo denotaremos con $V^{**} = (V^*)^*$.

¿Qué relación existe entre V y V^{**} ? Veremos que $V \cong V^{**}$, pero antes necesitaremos un resultado previo que vale para espacios de dimensión infinita pero que no probaremos aquí.

1.14 LEMA. *Sea V un espacio vectorial de dimensión finita n y $v \in V$ diferente de cero. Entonces existe un elemento $f \in V^*$ tal que $f(v) \neq 0$.*

Demostración. Como $v \neq 0$, podemos completar una base de V de la forma $\{v_1, v_2, v_3, \dots, v_n\}$ con $v_1 = v$. Sea $\{f_1, \dots, f_n\}$ la base dual. Entonces $f_1(v_1) = f_1(v) = 1 \neq 0$. ■

1.15 TEOREMA. *Sea V un espacio vectorial de dimensión finita sobre un campo K . La aplicación $\psi: V \rightarrow V^{**}$ dada por $\psi(v) = \bar{v}$ donde $\bar{v}(f) = f(v) \forall v \in V$ es un isomorfismo.*

Demostración. Veamos que ψ es lineal:

$$\begin{aligned} \psi(u+v)(f) &= \overline{u+v}(f) = f(u+v) = f(u) + f(v) \\ &= \bar{u}(f) + \bar{v}(f) = \psi(u)(f) + \psi(v)(f). \\ \psi(\lambda u)(f) &= \overline{\lambda u}(f) = f(\lambda u) = \lambda f(u) \\ &= \lambda \bar{u}(f) = \lambda \psi(u)(f). \end{aligned}$$

Luego ψ es lineal.

Sea $v \neq 0$, $v \in V$. Por el lema 1.14, existe un funcional $f \in V^*$ tal que $f(v) \neq 0$. Luego $0 \neq f(v) = \bar{v}(f) = \psi(v)(f)$ para toda $v \neq 0$, por lo que $\psi \neq 0$, es decir, ψ es no singular ($\ker \psi = \{0\}$). Por I.4.7 ψ es invertible y como $\dim V^{**} = \dim V^* = \dim V$, ψ es un isomorfismo. ■

Observe que la función lineal $\psi: V \longrightarrow V^{**}$ se definió sin hacer mención de una base.

A ψ de 1.15 se le llama *aplicación* u *homomorfismo natural* de V en V^{**} . Si V no es de dimensión finita, ψ no es suprayectiva.

PROBLEMAS

1.1 Pruebe que la función $f: \mathbb{R}^m \times \mathbb{R}^m \longrightarrow \mathbb{R}$ dada por $f((x_1, \dots, x_m), (y_1, \dots, y_m)) = x_1y_1 + x_2y_2 + \dots + x_my_m$ es bilineal. A f , así definida, se le llama *producto escalar*.

1.2 Encuentre la base del espacio dual de V para

- (a) $V = \mathbb{R}^2$ con base $\{(2, 1), (-3, 87)\}$.
- (b) $V = \mathbb{R}^3$ con base $\{(0, 1, 1), (0, 1, -1), (2, 4, -3)\}$.

1.3 Sea $f: K^m \times K^n \longrightarrow K$ una aplicación bilineal. Pruebe que existe una matriz única A tal que $f(X, Y) = {}^tXAY$ y que el conjunto de aplicaciones bilineales $L^2(K^m, K^n; K)$ es un espacio vectorial sobre K isomorfo a $M_{m \times n}(K)$.

1.4 Considere $f: \mathbb{R}^2 \times \mathbb{R}^2 \longrightarrow \mathbb{R}$ dado por $f((\alpha_1, \alpha_2), (\beta_1, \beta_2)) = 3\alpha_1\beta_2 - \alpha_2\beta_1$.

- (i) Compruebe que f es bilineal.
- (ii) Encuentre la matriz asociada a f con respecto a la base $\gamma = \{(2, 1), (-3, 87)\}$.
- (iii) Encuentre la matriz asociada a f con respecto a la base $\gamma' = \{(1, -1), (2, 1)\}$.

1.5 Sea $V = \mathbb{R}^3$ con bases $\gamma = \{(0, 1, 1), (0, 1, -1), (2, 4, -3)\}$ y $\gamma' = \{(1, 0, 0), (0, 1, 0), (0, 0, 1)\}$.

- (i) Encuentre la matriz de transición de la base γ en la base γ' .

- (ii) Encuentre las bases duales de γ y γ' en $(\mathbb{R}^3)^*$.
- (iii) Encuentre la matriz de transición de la base dual de γ en la base dual de γ' .

1.6 Sea U un subconjunto de un espacio vectorial V sobre un campo K . Un funcional $f \in V^*$ se llama *aniquilador* o *anulador* de U si $f(u) = 0$ para toda $u \in U$. Sea $U^a = \{f \in V^* | f(u) = 0 \ \forall u \in U\}$. Pruebe que:

- (i) U^a es un subespacio de V^* .
- (ii) Si V es de dimensión finita y U es un subespacio de V entonces $\dim U^a + \dim U = \dim V$ y $(U^a)^a = U$.

1.7 Sea U el subespacio de \mathbb{R}^3 generado por los vectores $u_1 = (4, 3, 1)$ y $u_2 = (1, 4, 3)$. Encuentre una base para U^a .

1.8 El *rango* de una forma bilineal $f: V \times V \rightarrow K$ se define como el rango de la matriz asociada a f y se denota $rg(f)$. Se dice que f es una forma bilineal *degenerada* si $rg(f) < \dim V$. Se dice que f es *no degenerada* si $rg(f) = \dim V$. Pruebe que el producto escalar del problema 1.1 es una forma bilineal no degenerada.

III.2 FORMAS BILINEALES SIMÉTRICAS, ANTISIMÉTRICAS, ALTERNANTES Y HERMITIANAS. FORMAS CUADRÁTICAS.

Sea V un espacio vectorial sobre un campo K .

2.1 DEFINICION. Sea f una forma bilineal de V . Diremos que f es *simétrica* si $f(u, v) = f(v, u)$ para toda $u, v \in V$.

Recuerde que una matriz se llama *simétrica* si es igual a su traspuesta. Veamos como se relaciona el concepto de simetría de una forma bilineal con el de una matriz simétrica. Sea A la matriz asociada a la forma bilineal simétrica f . Entonces, si V es de dimensión finita, podemos escribir f como en el ejemplo 1.2

$$f: K^n \times K^n \longrightarrow K \quad \text{dada por}$$

$$f(X, Y) = {}^t XAY.$$

Pero ${}^t XAY = {}^t({}^t XAY)$, puesto que es un escalar. Luego ${}^t XAY = {}^t Y^t AX$. Como f es simétrica, $f(X, Y) = {}^t XAY = {}^t Y^t AX = {}^t YAX = f(Y, X) \quad \forall (X, Y)$. Luego ${}^t A = A$. Análogamente, si $A = {}^t A$ entonces $f(X, Y) = f(Y, X)$. Por lo tanto, una forma bilineal es simétrica si, y sólo si, su matriz asociada es simétrica. Observemos que si f posee una matriz asociada en forma diagonal entonces f es simétrica pues toda matriz diagonal es simétrica.

2.2 DEFINICION. Sea $f: V \times V \longrightarrow K$ una forma bilineal simétrica. La *forma cuadrática* asociada a f es la función $q: V \longrightarrow K$ dada por $q(v) = f(v, v)$, $v \in V$.

Observe que una forma cuadrática es lineal solamente cuando f es la función trivial.

Sea A la matriz simétrica asociada a la forma bilineal simétrica f . Entonces podemos escribir

$$q(X) = f(X, X) = {}^t XAX$$

En otra forma, esto es

$$\begin{aligned}
& (x_1, \dots, x_n) \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \\
&= \sum_{i,j} a_{ij} x_i x_j \\
&= a_{11} x_1^2 + a_{22} x_2^2 + \cdots + a_{nn} x_n^2 + 2 \sum_{i < j} a_{ij} x_i x_j.
\end{aligned}$$

Si A es diagonal, entonces $q(X) = a_{11}x_1^2 + \cdots + a_{nn}x_n^2$ y como más adelante veremos, éste es siempre el caso.

Considere la expresión y las igualdades siguientes

$$\begin{aligned}
q(u+v) - q(u) - q(v) &= f(u+v, u+v) - f(u, u) - f(v, v) \\
&= f(u, u) + f(u, v) + f(v, u) + f(v, v) - f(u, u) - f(v, v) \\
&= 2f(u, v)
\end{aligned}$$

Si el campo K es tal que $1+1 \neq 0$ podremos dividir entre 2 y tendremos la siguiente expresión que permite obtener f a partir de q :

$$f(u, v) = \frac{1}{2} [q(u+v) - q(u) - q(v)].$$

2.3 EJEMPLO. Sea $X = (x, y)$ y $q(x, y) = 9x^2 - 8xy - 3y^2$ una forma cuadrática. La matriz simétrica asociada a q posee elementos diagonales iguales a los coeficientes de los términos cuadrados y elementos a_{ij} , $i \neq j$, iguales a la mitad de los coeficientes de los términos no cuadrados. Luego

$$\begin{pmatrix} 9 & -4 \\ -4 & -3 \end{pmatrix}$$

es la matriz simétrica asociada a q .

Ahora encontremos la forma bilineal simétrica asociada a q : Consideremos vectores $X_1 = (x_1, y_1)$ y $X_2 = (x_2, y_2)$; luego

$$\begin{aligned}
f(X_1, X_2) &= \frac{1}{2} [q(X_1 + X_2) - q(X_1) - q(X_2)] \\
&= \frac{1}{2} [9(x_1 + x_2)^2 - 8(x_1 + x_2)(y_1 + y_2) - 3(y_1 + y_2)^2 - 9x_1^2 + 8x_1y_1 \\
&\quad + 3y_1^2 - 9x_2^2 + 8x_2y_2 + 3y_2^2] \\
&= \frac{1}{2} [2 \cdot 9x_1x_2 - 8x_2y_1 - 8x_1y_2 - 2 \cdot 3y_1y_2] \\
&= 9x_1x_2 - 4x_2y_1 - 4x_1y_2 - 3y_1y_2.
\end{aligned}$$

Observe que los coeficientes de f son los elementos de la matriz simétrica asociada a q .

2.4 TEOREMA. *Sea f una forma bilineal simétrica $f: V \times V \rightarrow K$ tal que $1+1 \neq 0$ en K (i.e. tal que K sea de característica diferente de 2). Entonces existe una base de V tal que f posee una matriz asociada diagonal.*

Demostración. La afirmación es obvia si $\dim V = 1$ o si $f = 0$. Supongamos que $n = \dim V > 1$ y que $f \neq 0$. Si $f(v, v) = 0$ para toda $v \in V$, utilizando la fórmula que permite obtener f a partir de q (la anterior al ejemplo 2.3), $f = 0$. Por lo tanto podemos suponer que existe un elemento $v_1 \in V$ tal que $f(v_1, v_1) \neq 0$ pues estamos considerando el caso en que $f \neq 0$. Sea $V_1 = \langle v_1 \rangle$ y U el subespacio generado por los elementos $v \in V$ tales que $f(v_1, v) = 0$. Veremos que $V = V_1 \oplus U$ y así, f restringida a U es una forma bilineal simétrica en U . Así tendríamos que $\dim U = n - 1$ y por lo tanto existiría una base $\{v_2, \dots, v_n\}$ de U tal que $f(v_i, v_j) = 0$ para $i \neq j$ y $2 \leq i, j \leq n$. Pero por la definición de U , $f(v_1, v_j) = 0$ para $j = 2, \dots, n$. Luego $\{v_1, \dots, v_n\}$ es la base requerida tal que $f(v_i, v_j) = 0$, $i \neq j$ y f posee una matriz asociada diagonal (véase la definición de matriz asociada a una forma bilineal f anterior al ejemplo 1.9).

Nos resta probar que $V = V_1 \oplus U$: veamos que $V_1 \cap U = \{0\}$. Sea $w \in V_1 \cap U$. Como $w \in V_1$, $w = \lambda v_1$ para $\lambda \in K$. Como $w \in U$, $0 = f(w, w) = f(\lambda v_1, \lambda v_1) = \lambda^2 f(v_1, v_1)$. Pero $f(v_1, v_1) \neq 0$, luego $\lambda = 0$ y por lo tanto $w = \lambda v_1 = 0$ y $V_1 \cap U = \{0\}$. Ahora, veamos que $V = V_1 + U$. Sea $v \in V$ y sea $u = v - \frac{f(v_1, v)}{f(v_1, v_1)} v_1$. Entonces

$$f(v_1, u) = f(v_1, v) - \frac{f(v_1, v)}{f(v_1, v_1)} f(v_1, v_1) = 0.$$

Luego $u \in U$ y, por lo anterior, $v = u + \frac{f(v_1, v)}{f(v_1, v_1)} v_1$ y así, v es la suma de un elemento de V_1 y de un elemento de U . ■

Diremos que una matriz simétrica B es *congruente* a una matriz simétrica A si existe una matriz no singular o invertible N tal que $B = {}^t N A N$.

En términos de congruencia de matrices podemos interpretar el teorema anterior como sigue: si A es una matriz simétrica (con elementos en un campo de característica diferente de 2) entonces A es congruente con una matriz diagonal, i.e., existe una matriz no singular o invertible N tal que ${}^t N A N$ es diagonal.

Observe que 1.11 lo podemos formular para el caso de formas bilineales simétricas

como sigue: las matrices asociadas a una misma forma bilineal simétrica son congruentes entre sí. Nótese que si A es simétrica entonces tNAN es simétrica.

También observe que, por el teorema 2.4, cuando la característica de K es diferente de 2, siempre podemos encontrar una matriz diagonal tal que la forma cuadrática posea una representación de la forma

$$q(X) = a_{11}x_1^2 + \cdots + a_{nn}x_n^2.$$

2.5 DEFINICION. Sea $f: V \times V \rightarrow K$ una forma bilineal. Diremos que f es *antisimétrica* si $f(u, v) = -f(v, u)$ para toda $u, v \in V$.

Es claro que si V es de dimensión finita, f es antisimétrica si, y sólo si, su matriz asociada A es tal que $A = -{}^tA$, es decir si, y sólo si, A es *antisimétrica*.

Observe que si el campo K es tal que $1+1 \neq 0$ (i.e. es de característica diferente de dos) y f antisimétrica entonces $f(v, v) = -f(v, v)$ lo cual implica que $f(v, v) = 0$.

2.6 PROPOSICION. Si $f: V \times V \rightarrow K$ es una forma bilineal, entonces f es la suma de una forma bilineal simétrica y de una forma bilineal antisimétrica.

Demostración. Sea $s(u, v) = \frac{1}{2}(f(u, v) + f(v, u))$ y $a(u, v) = \frac{1}{2}(f(u, v) - f(v, u))$. Luego $s(u, v) = \frac{1}{2}(f(u, v) + f(v, u)) = \frac{1}{2}(f(v, u) + f(u, v)) = s(v, u)$ y $a(u, v) = \frac{1}{2}(f(u, v) - f(v, u)) = -\frac{1}{2}(f(v, u) - f(u, v)) = -a(v, u)$. Así, s es simétrica y a es antisimétrica. Luego $f = s + a$. ■

2.7 DEFINICION. Una forma bilineal f se llama *alternante* si $f(v, v) = 0$ para toda $v \in V$.

Ahora, consideremos formas bilineales $f: V \times V \rightarrow K$ donde $K = \mathbb{R}$.

2.8 TEOREMA. Sea $f: V \times V \rightarrow \mathbb{R}$ una forma bilineal simétrica. Entonces V posee una base tal que la matriz asociada a f es diagonal y cualquier otra matriz diagonal asociada a f posee el mismo número de elementos positivos y el mismo número de elementos negativos.

Demostración. Por el teorema anterior existe una base $\{v_i\}_{i=1}^n$ tal que f posee una matriz asociada diagonal. Sea $\{v'_i\}_{i=1}^n$ otra base tal que la matriz asociada a

f es diagonal. Sean n_+ y n'_+ los números de elementos positivos, y n_- y n'_- los números de elementos negativos de las matrices asociadas a f con respecto a las bases $\{v_i\}$ y $\{v'_i\}$.

Sabemos que $rg(f) = n_+ + n_- = n'_+ + n'_-$. Veamos que $n_+ = n'_+$: sea V_+ el subespacio generado por $\{v_1, \dots, v_{n_+}\}$ y V'_- el subespacio generado por $\{v'_{n'_++1}, \dots, v'_{n'}\}$. Luego $f(v, v) > 0$ si $0 \neq v \in V_+$ y $f(v, v) \leq 0$ si $0 \neq v \in V'_-$. Así que $V_+ \cap V'_- = \{0\}$ y $\dim V_+ = n_+$ y $\dim V'_- = n - n'_+$. Entonces $\dim (V_+ + V'_-) = \dim V_+ + \dim V'_- - \dim (V_+ \cap V'_-) = n_+ + (n - n'_+) - 0 = n_+ - n'_+ + n$. Pero $\dim (V_+ + V'_-) \leq \dim V = n$. Luego $n_+ - n'_+ + n \leq n$, i.e., $n_+ \leq n'_+$. Análogamente $n'_+ \leq n_+$ y por lo tanto tendremos $n_+ = n'_+$. Es claro que esto es suficiente para probar el teorema pues podríamos comenzar la demostración con los elementos negativos. ■

El teorema anterior se llama *ley de la inercia de Sylvester*.

Consideremos el caso en que $K = \mathbb{C}$. Diremos que una forma $f: V \times V \rightarrow \mathbb{C}$ es *hermitiana* si

- (i) $f(\lambda v_1 + \mu v_2, v) = \lambda f(v_1, v) + \mu f(v_2, v)$ y
- (ii) $f(u, v) = \overline{f(v, u)}$ para $v_1, v_2, v \in V$; $\lambda, \mu \in \mathbb{C}$.

A veces se les llama *formas bilineales hermitianas*, a pesar de que no son bilineales, pues el escalar de la segunda variable sale “conjugado” (problema 2.4).

Definimos $q: V \rightarrow \mathbb{R}$ dado por $q(v) = f(v, v)$ como la *forma cuadrática hermitiana* o *compleja* asociada a la forma bilineal hermitiana f . f se obtiene a partir de q mediante la fórmula $f(u, v) = \frac{1}{4}(q(u+v) - q(u-v)) + \frac{i}{4}(q(u+iv) - q(u-iv))$ y también se tiene un resultado análogo a 2.8 para formas bilineales hermitianas.

PROBLEMAS

2.1 Considere las formas cuadráticas siguientes y encuentre su matriz asociada, así como la forma bilineal correspondiente:

- (i) $q(x, y) = 18x^2 - 4xy - 9y^2$
- (ii) $q(x, y) = x^2 - xy$
- (iii) $q(x, y, z) = x^2 + 4yz - xz$

2.2 Para cada matriz A asociada a las formas del problema 2.1 que sea simétrica encuentre una matriz no singular N tal que tNAN sea diagonal.

2.3 En los términos del teorema 2.8 se define el *signo* de una forma bilineal $f: V \times V \rightarrow \mathbb{R}$ como $sgf = n_+ - n_-$. Se dice que f está *definida positivamente* si $f(v, v) > 0 \forall v \neq 0$ y que está *semidefinida no negativamente* si $f(v, v) \geq 0 \forall v \in V$.

- (i) Pruebe que f está definida positivamente si, y sólo si, $\dim V = sg(f)$.
- (ii) Pruebe que f está semidefinida no negativamente si, y sólo si, $sg(f) = rg(f)$
y
- (iii) analice las formas de los problemas 2.1 y 1.1 y diga sus características.

2.4 Sea $f: V \times V \rightarrow \mathbb{C}$ una forma bilineal hermitiana. Pruebe que $f(v, \lambda v_1 + \mu v_2) = \bar{\lambda}f(v, v_1) + \bar{\mu}f(v, v_2)$.

2.5 Sea $\bar{A} = (\bar{a}_{ij})$ la matriz cuadrada donde \bar{a}_{ij} denota al conjugado complejo de a_{ij} . Sea $A^* = {}^t\bar{A} = \bar{{}^tA}$. Una matriz se llama *hermitiana* si $A^* = A$, i.e., si $a_{ij} = \bar{a}_{ji}$. Pruebe que si A es hermitiana, $f(X, Y) = {}^tXA\bar{Y}$ define una forma hermitiana en \mathbb{C}^n .

2.6 Compruebe la fórmula que permite obtener la forma hermitiana a partir de la forma cuadrática hermitiana asociada

$$f(u, v) = \frac{1}{4}(q(u+v) - q(u-v)) + \frac{i}{4}(q(u+iv) - q(u-iv)).$$

2.7 Establezca el resultado análogo al del teorema 2.8 para formas bilineales hermitianas.

2.8 Pruebe que la congruencia de matrices simétricas es una relación de equivalencia.

2.9 Pruebe que cualquier forma cuadrática sobre \mathbb{R} es de la forma $q(X) = x_1^2 + \cdots + x_s^2 - x_{s+1}^2 - \cdots - x_r^2$ donde $X = (x_1, \dots, x_n)$. (Sugerencia: considere la matriz diagonal N con elementos iguales a 1 si $a_{ii} = 0$ y $1/\sqrt{|a_{ii}|}$ si $a_{ii} \neq 0$ donde $A = (a_{ii})$ es la matriz diagonal asociada a la forma cuadrática $q(X) = a_{11}x_1^2 + \cdots + a_{nn}x_n^2$. Luego A es congruente con la matriz diagonal requerida).

III.3 PRODUCTO ESCALAR

3.1 DEFINICION. Sea V un espacio vectorial sobre un campo $K = \mathbb{R}$ o \mathbb{C} . Si $K = \mathbb{R}$, un *producto escalar en V sobre \mathbb{R}* es una forma bilineal simétrica definida positivamente. Si $K = \mathbb{C}$, un *producto escalar en V sobre \mathbb{C}* es una forma hermitiana definida positivamente (i.e. tal que $f(v, v) > 0 \forall v \neq 0$).

Es decir, si V es un espacio vectorial real o complejo, la forma bilineal $\langle \cdot, \cdot \rangle: V \times V \rightarrow K$ tal que

- (i) $\langle \lambda v_1 + \mu v_2, v \rangle = \lambda \langle v_1, v \rangle + \mu \langle v_2, v \rangle$
- (ii) $\langle u, v \rangle = \overline{\langle v, u \rangle}$ y
- (iii) $\langle v, v \rangle > 0$ si $v \neq 0$

se llama *producto escalar*. También se le acostumbra llamar *producto interno*.

Un espacio vectorial que posee un producto escalar se llama *espacio con producto escalar*. Si el producto escalar es sobre \mathbb{R} , al espacio vectorial se le llama *espacio euclidiano* y si el campo es \mathbb{C} , se le llama *espacio unitario*.

Observe que

$$\begin{aligned} \langle v, \lambda v_1 + \mu v_2 \rangle &= \overline{\langle \lambda v_1 + \mu v_2, v \rangle} \\ &= \overline{\lambda \langle v_1, v \rangle + \mu \langle v_2, v \rangle} \\ &= \overline{\lambda} \overline{\langle v_1, v \rangle} + \overline{\mu} \overline{\langle v_2, v \rangle} \\ &= \overline{\lambda} \langle v, v_1 \rangle + \overline{\mu} \langle v, v_2 \rangle. \end{aligned}$$

3.2 EJEMPLOS. (i) Los problemas 1.1 y 2.3 nos dicen que

$$\langle \cdot, \cdot \rangle: \mathbb{R}^m \times \mathbb{R}^m \rightarrow \mathbb{R}$$

es un producto escalar en $V = \mathbb{R}^m$ sobre \mathbb{R} .

(ii) $\langle \cdot, \cdot \rangle: \mathbb{C}^m \times \mathbb{C}^m \rightarrow \mathbb{C}$ dado por

$$\langle X, Y \rangle = x_1 \bar{y}_1 + \cdots + x_m \bar{y}_m$$

donde $X = (x_1, \dots, x_m) \in \mathbb{C}^m$ y $Y = (y_1, \dots, y_m) \in \mathbb{C}^m$ es un producto escalar en $V = \mathbb{C}^m$ sobre \mathbb{C} .

(iii) $\langle \cdot, \cdot \rangle: M_n K \times M_n K \longrightarrow K$ dado por

$$\langle A, B \rangle = \begin{cases} \text{tr}({}^t BA) & \text{si } K = \mathbb{R} \text{ y} \\ \text{tr}(B^* A) & \text{si } K = \mathbb{C} \end{cases}$$

son productos escalares en $M_n K$ sobre $K = \mathbb{R}$ o \mathbb{C} .

A continuación definimos la *norma* o *longitud* de un vector v de un espacio vectorial V sobre $K = \mathbb{R}$ o \mathbb{C} como

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

Observe que si $\langle v, v \rangle = 1$ entonces $\|v\| = 1$ y llamaremos a v *vector unitario* o *normalizado*.

3.3 PROPOSICION. Si $v, w \in V$ entonces $|\langle v, w \rangle| \leq \|v\| \|w\|$. (Desigualdad de Cauchy-Schwarz).

Demostración. Si $w = 0$ entonces $0 \leq 0$ y el resultado es cierto. Si $w \neq 0$, para cualquier $\lambda \in \mathbb{R}$, como $\overline{\langle v, w \rangle} = \langle w, v \rangle$ y $\lambda = \bar{\lambda}$,

$$\begin{aligned} 0 &\leq \|v - \langle v, w \rangle \lambda w\|^2 = \langle v - \langle v, w \rangle \lambda w, v - \langle v, w \rangle \lambda w \rangle \\ &= \langle v, v \rangle - \langle v, w \rangle \lambda \langle w, v \rangle - \overline{\langle v, w \rangle} \lambda \langle v, w \rangle + \langle v, w \rangle \overline{\langle v, w \rangle} \lambda^2 \langle w, w \rangle \\ &= \langle v, v \rangle - \langle v, w \rangle \lambda \overline{\langle v, w \rangle} - \overline{\langle v, w \rangle} \lambda \langle v, w \rangle + \langle v, w \rangle \overline{\langle v, w \rangle} \lambda^2 \langle w, w \rangle \\ &= \|v\|^2 - 2\lambda |\langle v, w \rangle|^2 + |\langle v, w \rangle|^2 \lambda^2 \|w\|^2 \quad \text{pues } |\langle v, w \rangle|^2 = \langle v, w \rangle \overline{\langle v, w \rangle}. \end{aligned}$$

Ahora, si tomamos $\lambda = \frac{1}{\|w\|^2}$, obtenemos $0 \leq \|v\|^2 - \frac{|\langle v, w \rangle|^2}{\|w\|^2}$. Luego $|\langle v, w \rangle|^2 \leq \|v\|^2 \|w\|^2$. Tomando raíz cuadrada obtenemos el resultado requerido. ■

3.4 PROPOSICION. (i) $\|v\| > 0$ si $v \neq 0$

(ii) $\|\lambda v\| = |\lambda| \|v\|$

(iii) $\|v + w\| \leq \|v\| + \|w\|$ (*desigualdad del triángulo*).

Demostración. (i) Por 3.1, $\langle v, v \rangle > 0$ si $v \neq 0$, luego $\|v\| = \sqrt{\langle v, v \rangle} > 0$ si $v \neq 0$.

(ii) Consideremos $\|\lambda v\|^2 = \langle \lambda v, \lambda v \rangle = \lambda \bar{\lambda} \langle v, v \rangle = |\lambda|^2 \|v\|^2$, luego tomemos la raíz cuadrada de ambos lados.

(iii) Por 3.3

$$\begin{aligned} \|v + w\|^2 &= \langle v + w, v + w \rangle = \langle v, v \rangle + \langle v, w \rangle + \overline{\langle v, w \rangle} + \langle w, w \rangle \\ &\leq \|v\|^2 + 2\|v\| \|w\| + \|w\|^2 = (\|v\| + \|w\|)^2. \end{aligned}$$

Luego tómesese la raíz cuadrada de ambos lados. ■

Sea V un espacio vectorial con producto escalar. Diremos que los vectores $v, w \in V$ son *ortogonales* si $\langle v, w \rangle = 0$.

Si V es un espacio euclidiano, definimos el *ángulo* entre dos vectores no nulos $u, v \in V$ como $\theta = \arccos(\langle u, v \rangle / \|u\| \|v\|)$ para $\theta \in [0, \pi]$.

Sea U un subconjunto del espacio V y $U^\perp = \{v \in V \mid \langle u, v \rangle = 0 \ \forall u \in U\}$ el *conjunto ortogonal* a U .

3.5 PROPOSICION. U^\perp es un subespacio de V .

Demostración. Como $\langle u, 0 \rangle = 0$, $0 \in U^\perp$. Sean $v, v' \in U^\perp$, entonces $\langle u, \lambda v + \mu v' \rangle = \lambda \langle u, v \rangle + \mu \langle u, v' \rangle = \lambda 0 + \mu 0 = 0$. Por lo tanto $\lambda v + \mu v' \in U^\perp$. ■

3.6 DEFINICION. Sea $\{v_i\}_{i=1}^n$ un conjunto de vectores de un espacio vectorial V sobre un campo $K = \mathbb{R}$ o \mathbb{C} . Diremos que $\{v_i\}$ es *ortogonal* si $\langle v_i, v_j \rangle = 0$ para $i \neq j$ y *ortonormal* si $\langle v_i, v_j \rangle = \delta_{ij} = \begin{cases} 0 & \text{si } i \neq j \\ 1 & \text{si } i = j \end{cases}$.

A continuación estableceremos un procedimiento para encontrar una base ortonormal de un espacio vectorial V de dimensión finita, llamado *procedimiento de Gram-Schmidt*.

3.7 TEOREMA. Sea $\{u_i\}_{i=1}^n$ una base del espacio vectorial de dimensión finita V sobre un campo $K = \mathbb{R}$ o \mathbb{C} . Entonces existe una base ortonormal $\{v_i\}_{i=1}^n$ de V tal que la matriz de transición es triangular.

Demostración. Utilizaremos inducción sobre $n = \dim V$. Si $n = 1$, consideremos $v_1 = \frac{u_1}{\|u_1\|}$, luego $\{v_1\}$ es ortonormal. Supongamos que el conjunto $\{v_1, v_2, \dots, v_{n-1}\}$ es ortonormal. Veamos que si $v_n = \frac{w_n}{\|w_n\|}$, donde $w_n = u_n - \langle u_n, v_1 \rangle v_1 - \langle u_n, v_2 \rangle v_2 - \langle u_n, v_3 \rangle v_3 - \dots - \langle u_n, v_{n-1} \rangle v_{n-1}$, entonces el conjunto

$\{v_1, \dots, v_n\}$ es ortonormal. Pero un cálculo directo nos muestra que $\langle w_n, v_1 \rangle = \langle w_n, v_2 \rangle = \dots = \langle w_n, v_{n-1} \rangle = 0$. Así, w_n es ortogonal a cada v_1, \dots, v_{n-1} . Luego, el conjunto $\{v_1, \dots, v_n\}$ es ortonormal.

Nos resta probar que el conjunto ortonormal es linealmente independiente: considere $\lambda_1 v_1 + \dots + \lambda_n v_n = 0$ y calculemos $\langle 0, v_i \rangle$:

$$\begin{aligned} 0 &= \langle 0, v_i \rangle = \langle \lambda_1 v_1 + \dots + \lambda_n v_n, v_i \rangle \\ &= \lambda_1 \langle v_1, v_i \rangle + \dots + \lambda_i \langle v_i, v_i \rangle + \dots + \lambda_n \langle v_n, v_i \rangle = \lambda_i. \end{aligned}$$

Luego $\{v_i\}_{i=1}^n$ es una base de V . Resumiendo, la base ortogonal se obtiene como sigue:

$$\begin{aligned} v_1 &= \frac{u_1}{\|u_1\|} \\ v_2 &= \frac{u_2 - \langle u_2, v_1 \rangle v_1}{\|u_2 - \langle u_2, v_1 \rangle v_1\|} \\ v_3 &= \frac{u_3 - \langle u_3, v_1 \rangle v_1 - \langle u_3, v_2 \rangle v_2}{\|u_3 - \langle u_3, v_1 \rangle v_1 - \langle u_3, v_2 \rangle v_2\|} \\ &\vdots \\ v_n &= \frac{u_n - \langle u_n, v_1 \rangle v_1 - \dots - \langle u_n, v_{n-1} \rangle v_{n-1}}{\|u_n - \langle u_n, v_1 \rangle v_1 - \dots - \langle u_n, v_{n-1} \rangle v_{n-1}\|} \end{aligned}$$

Claramente, la matriz de transición de la base $\{u_i\}$ a $\{v_i\}$ es triangular. ■

3.8 EJEMPLO. Sean $u_1 = (1, 0, 0)$, $u_2 = (1, 1, 0)$ y $u_3 = (1, 1, 1)$ los elementos de una base para \mathbb{R}^3 . Utilizando el procedimiento de 3.7 obtengamos una base ortonormal v_1, v_2, v_3 para \mathbb{R}^3 .

$$\text{Sea } v_1 = \frac{u_1}{\|u_1\|} = \frac{(1, 0, 0)}{\sqrt{1}} = (1, 0, 0). \text{ Luego,}$$

$$\begin{aligned} w_2 &= u_2 - \langle u_2, v_1 \rangle v_1 \\ &= (1, 1, 0) - \langle (1, 1, 0), (1, 0, 0) \rangle (1, 0, 0) \\ &= (1, 1, 0) - (1, 0, 0) = (0, 1, 0) \end{aligned}$$

$$\text{Consideremos } v_2 = \frac{w_2}{\|w_2\|} = \frac{(0, 1, 0)}{\sqrt{1}} = (0, 1, 0).$$

$$\begin{aligned} \text{Luego } w_3 &= u_3 - \langle u_3, v_1 \rangle v_1 - \langle u_3, v_2 \rangle v_2 \\ &= (1, 1, 1) - \langle (1, 1, 1), (1, 0, 0) \rangle (1, 0, 0) - \langle (1, 1, 1), (0, 1, 0) \rangle (0, 1, 0) \\ &= (1, 1, 1) - (1, 0, 0) - (0, 1, 0) = (0, 0, 1). \end{aligned}$$

Normalizando w_3 obtenemos $v_3 = \frac{w_3}{\|w_3\|} = \frac{(0, 0, 1)}{\sqrt{1}} = (0, 0, 1)$. Luego $\{v_1, v_2, v_3\}$ es una base ortonormal para \mathbb{R}^3 .

3.9 TEOREMA. Si U es un subespacio de un espacio vectorial V de dimensión finita sobre un campo K , entonces $V \cong U \oplus U^\perp$.

Demostración. Sea $\{u_1, \dots, u_s\}$ una base de U . Extiéndase a una base $\{u_1, \dots, u_n\}$ de V y aplique el procedimiento de Gram-Schmidt para obtener una base ortonormal $\{v_1, \dots, v_n\}$ de V . Por lo tanto $\{v_1, \dots, v_s\}$ es una base ortonormal de U y $\{v_{s+1}, \dots, v_n\} \subset U^\perp$. Veamos que $V = U + U^\perp$: si $v \in V$ entonces $v = \lambda_1 v_1 + \dots + \lambda_n v_n$ donde $\lambda_1 v_1 + \dots + \lambda_s v_s \in U$ y $\lambda_{s+1} v_{s+1} + \dots + \lambda_n v_n \in U^\perp$. Luego $V = U + U^\perp$. Veamos que $U \cap U^\perp = \{0\}$: si $u \in U \cap U^\perp$ entonces $\langle u, u \rangle = 0$ y por el problema 3.7, $u = 0$. Luego $U \cap U^\perp = \{0\}$. ■

Observe que si $V \cong U \oplus U^\perp$ entonces existe una proyección única (I.2.4) $p_U: V \rightarrow V$ tal que $\text{im } p_U = U$ y $\ker p_U = U^\perp$ llamada *proyección ortogonal* de V en U . Por ejemplo, sea $V = \mathbb{R}^3$, $U = \{(0, 0, \eta) \mid \eta \in \mathbb{R}\}$ y $U^\perp = \{(\lambda, \mu, 0) \mid \lambda, \mu \in \mathbb{R}\}$. Entonces $\mathbb{R}^3 = U \oplus U^\perp$ y $p_U(\lambda, \mu, \eta) = \text{im } p_U = (0, 0, \eta)$ es la proyección ortogonal de \mathbb{R}^3 en U .

Otro ejemplo es el siguiente: denotemos con $AX = 0$ al sistema de m ecuaciones lineales homogéneo con n incógnitas con coeficientes reales. El espacio solución U es el núcleo del operador lineal A . Es decir, U es el conjunto de todos los vectores ortogonales a cada renglón de A . Luego, U es el complemento ortogonal del espacio generado por los renglones de A . Además, por I.3.10, el teorema 3.9 nos dice que $\dim U = n - \text{rango } A$.

PROBLEMAS

3.1 Verifique que los ejemplos de 3.2 son efectivamente productos escalares.

3.2 Sea $V = C_{[0,1]}$ el espacio de las funciones continuas sobre el intervalo $[0,1]$ y $K = \mathbb{R}$ o \mathbb{C} . Compruebe que $\langle \cdot, \cdot \rangle: C_{[0,1]} \times C_{[0,1]} \rightarrow K$ dado por $\langle f, g \rangle = \int_0^1 f(x)\overline{g(x)}dx$, $x \in [0, 1]$ es un producto escalar en $V = C_{[0,1]}$ sobre K .

3.3 Compruebe utilizando los resultados de la sección III.2 que $\langle \cdot, \cdot \rangle: \mathbb{R}^2 \times \mathbb{R}^2 \rightarrow \mathbb{R}$ dado por $\langle X, Y \rangle = 2x_1y_1 - 3x_1y_2 - 3x_2y_1 + 2x_2y_2$ donde $X = (x_1, x_2)$ y $Y = (y_1, y_2)$ es un producto escalar.

3.4 Calcule la norma del vector $u = (6, 7) \in \mathbb{R}^2$ con respecto al producto escalar del problema 3.3.

3.5 Sea $f = x^2 + x + 1$ y $g = 3x + 2i$. Calcule $\langle f, g \rangle$, $\|f\|$ y $\|g\|$ con respecto al producto escalar del problema 3.2.

3.6 Encuentre una base ortonormal para el subespacio de $C_{[0,1]}$ generado por $\{1, x, x^2\}$ con el producto escalar definido en el problema 3.2.

3.7 Sea U un espacio vectorial sobre un campo $K = \mathbb{R}$ o \mathbb{C} . Pruebe que

- (i) si $\langle u, u \rangle = 0$ entonces $u = 0$, $u \in U$ y
- (ii) $\|u\| = 0$ si y sólo si $u = 0$.

3.8 Sea $\{v_1, \dots, v_s\}$ un subconjunto ortogonal de un espacio vectorial V de dimensión finita n . Pruebe que para cualquier $v \in V$,

$$\sum_{i=1}^s \frac{|\langle v, v_i \rangle|^2}{\|v_i\|^2} \leq \|v\|^2$$

(*desigualdad de Bessel*).

3.9 Pruebe que, (i) si U es un subespacio de un espacio vectorial V de dimensión finita, entonces $U = U^{\perp\perp}$;

- (ii) si V no es de dimensión finita entonces $U \subset (U^\perp)^\perp$.

3.10 Sea V un espacio vectorial de dimensión finita con producto escalar. Pruebe que $(U + U')^\perp = U^\perp \cap U'^\perp$ y que $(U \cap U')^\perp = U^\perp + U'^\perp$.

3.11 Sea $\{v_1, \dots, v_n\}$ una base ortonormal de un espacio vectorial V sobre un campo K . Pruebe que si $v \in V$ entonces $v = \langle v, v_1 \rangle v_1 + \langle v, v_2 \rangle v_2 + \dots + \langle v, v_n \rangle v_n$, y que si $\rho: V \rightarrow V$ es un operador lineal entonces el elemento ij de la matriz asociada A a ρ con respecto a la base dada es $\langle \rho(v_j), v_i \rangle$.

III.4 OPERADORES ADJUNTOS

En esta sección K denota al campo \mathbb{R} o \mathbb{C} .

Sea V un espacio vectorial con producto escalar sobre el campo K . Sea

$$g: V \longrightarrow V^* \quad \text{tal que}$$

$$v \longmapsto g_v$$

la función dada por $g_v(u) = \langle u, v \rangle$. Efectivamente, como $g_v(\lambda u_1 + \mu u_2) = \langle \lambda u_1 + \mu u_2, v \rangle = \lambda \langle u_1, v \rangle + \mu \langle u_2, v \rangle = \lambda g_v(u_1) + \mu g_v(u_2)$, tenemos que $g_v \in V^*$. Entonces cada elemento $v \in V$ nos proporciona un funcional g_v . Veamos que si V es de dimensión finita sucede lo inverso:

4.1 TEOREMA. *Sea V un espacio vectorial de dimensión finita sobre el campo K con producto escalar. Sea $f: V \longrightarrow K$ un funcional. Entonces existe un único $v \in V$ tal que $f(u) = \langle u, v \rangle$ para toda $u \in V$.*

Demostración. Sea $\{u_1, \dots, u_n\}$ una base ortonormal de V . Sea $v = \overline{f(u_1)}u_1 + \dots + \overline{f(u_n)}u_n$ un elemento de V . Definamos

$$g_v(u_i) = \langle u_i, v \rangle = \langle u_i, \overline{f(u_1)}u_1 + \dots + \overline{f(u_n)}u_n \rangle = f(u_i).$$

Luego $g_v = f$ pues coinciden en los elementos de la base. Veamos que v es único: supongamos que $v' \in V$ es tal que $f(u) = \langle u, v' \rangle$ para toda $u \in V$. Entonces $\langle u, v \rangle = \langle u, v' \rangle$, i.e. $\langle u, v - v' \rangle = 0$. Consideremos $u = v - v'$ luego $\langle v - v', v - v' \rangle = 0$. Por lo tanto $v - v' = 0$ y $v = v'$. ■

4.2 TEOREMA. *Sea V como en 4.1 y $\rho: V \longrightarrow V$ un operador lineal. Entonces existe un operador lineal único $\rho^*: V \longrightarrow V$ tal que*

$$\langle \rho(v), u \rangle = \langle v, \rho^*(u) \rangle$$

para toda $u, v \in V$.

Demostración. Sea $u \in V$ un elemento fijo de V . Consideremos la aplicación $f: V \longrightarrow K$ dada por $f(v) = \langle \rho(v), u \rangle$ la cual es un funcional de V . Por 4.1 existe un elemento único $u' \in V$ tal que $\langle \rho(v), u \rangle = \langle v, u' \rangle$ para toda $v \in V$. Sea $\rho^*: V \longrightarrow V$ tal que $\rho^*(u) = u'$. Entonces

$$\langle \rho(v), u \rangle = \langle v, \rho^*(u) \rangle$$

para toda $u, v \in V$. Veamos que ρ^* es lineal:

$$\begin{aligned}\langle v, \rho^*(\lambda u_1 + \mu u_2) \rangle &= \langle \rho(v), \lambda u_1 + \mu u_2 \rangle \\ &= \bar{\lambda} \langle \rho(v), u_1 \rangle + \bar{\mu} \langle \rho(v), u_2 \rangle \\ &= \bar{\lambda} \langle v, \rho^*(u_1) \rangle + \bar{\mu} \langle v, \rho^*(u_2) \rangle \\ &= \langle v, \lambda \rho^*(u_1) + \mu \rho^*(u_2) \rangle\end{aligned}$$

$u_1, u_2, v \in V$; $\lambda, \mu \in K$. Como v es arbitraria, por la unicidad de 4.1

$$\rho^*(\lambda u_1 + \mu u_2) = \lambda \rho^*(u_1) + \mu \rho^*(u_2),$$

y por lo tanto ρ^* es lineal. Dejamos probar la unicidad de ρ^* al lector (problema 4.1).■

Si la base de V es ortonormal se tiene la siguiente

4.3 PROPOSICION. Si $A = (a_{ij})$ es la matriz asociada a ρ con respecto a una base ortonormal $\{u_1, \dots, u_n\}$ de V entonces la matriz asociada a ρ^* es A^* .

Demostración. Por el problema 3.11 los elementos ij de las matrices asociadas a ρ y ρ^* son $\langle \rho(u_j), u_i \rangle$ y $\langle \rho^*(u_j), u_i \rangle$ respectivamente. Pero $\langle \rho^*(u_j), u_i \rangle = \overline{\langle u_i, \rho^*(u_j) \rangle} = \overline{\langle \rho(u_i), u_j \rangle} = \bar{a}_{ji}$. Luego A^* es la matriz asociada a ρ^* .■

4.4 DEFINICION. Sea V un espacio con producto escalar. Un operador $\rho: V \rightarrow V$ posee un *operador adjunto* ρ^* en V si

$$\langle \rho(v), u \rangle = \langle v, \rho^*(u) \rangle.$$

Observe que si V es un espacio vectorial de dimensión finita, por 4.2 siempre existe un operador adjunto.

4.5 EJEMPLO. Sea $\rho: \mathbb{C}^3 \rightarrow \mathbb{C}^3$ dado por

$$\rho(x, y, z) = (2x + 3iy + 4z, 8x + 9iy + 10z, 6x + 7iy + 8z).$$

La matriz A asociada a ρ con respecto a la base canónica es

$$A = \begin{pmatrix} 2 & 3i & 4 \\ 8 & 9i & 10 \\ 6 & 7i & 8 \end{pmatrix}$$

Entonces la matriz asociada a ρ^* es

$$A^* = \begin{pmatrix} 2 & 8 & 6 \\ -3i & -9i & -7i \\ 4 & 10 & 8 \end{pmatrix}$$

Luego $\rho^*(x, y, z) = (2x + 8y + 6z, -3ix - 9iy - 7iz, 4x + 10y + 8z)$.

4.6 PROPOSICION. Sean $\rho, \eta: V \rightarrow V$ operadores lineales de un espacio vectorial de dimensión finita sobre K . Entonces

- (i) $(\rho + \eta)^* = \rho^* + \eta^*$
- (ii) $(\rho\eta)^* = \eta^*\rho^*$
- (iii) $(\lambda\rho)^* = \bar{\lambda}\rho^*$
- (iv) $(\rho^*)^* = \rho$
- (v) $I^* = I$
- (vi) Si ρ es invertible entonces $(\rho^{-1})^* = (\rho^*)^{-1}$.

Demostración. Tomando en cuenta la proposición 4.3, la demostración puede hacerse utilizando las propiedades de las matrices, pero también se puede hacer directamente como sigue:

- (i) Considere las igualdades siguientes:

$$\begin{aligned} \langle (\rho + \eta)(v), u \rangle &= \langle \rho(v) + \eta(v), u \rangle \\ &= \langle \rho(v), u \rangle + \langle \eta(v), u \rangle \\ &= \langle v, \rho^*(u) \rangle + \langle v, \eta^*(u) \rangle \\ &= \langle v, \rho^*(u) + \eta^*(u) \rangle \\ &= \langle v, (\rho^* + \eta^*)(u) \rangle; \quad u, v \in V. \end{aligned}$$

Por la unicidad de 4.2, $(\rho + \eta)^* = \rho^* + \eta^*$.

- (ii) $\langle (\rho\eta)(v), u \rangle = \langle \rho(\eta(v)), u \rangle = \langle \eta(v), \rho^*(u) \rangle = \langle v, \eta^*(\rho^*(u)) \rangle = \langle v, (\eta^*\rho^*)(u) \rangle$;
 $u, v \in V$. Por la unicidad de 4.2, $(\rho\eta)^* = \eta^*\rho^*$.
- (iii) $\langle (\lambda\rho)(v), u \rangle = \langle \lambda\rho(v), u \rangle = \lambda\langle \rho(v), u \rangle = \lambda\langle v, \rho^*(u) \rangle = \langle v, \bar{\lambda}\rho^*(u) \rangle$;
 $= \langle v, (\bar{\lambda}\rho^*)(u) \rangle$; $u, v \in V$; $\lambda \in K$. Por la unicidad de 4.2, $(\lambda\rho)^* = \bar{\lambda}\rho^*$.
- (iv) $\langle \rho^*(v), u \rangle = \overline{\langle u, \rho^*(v) \rangle} = \overline{\langle \rho(u), v \rangle} = \langle v, \rho(u) \rangle$; $u, v \in V$. Por 4.2, $(\rho^*)^* = \rho$.
- (v) Lo dejamos como ejercicio al lector.
- (vi) Como $I = I^*$ por (v), $I^* = (\rho\rho^{-1})^* = (\rho^{-1})^*\rho^*$. Así que $(\rho^{-1})^* = (\rho^*)^{-1}$. ■

Considere el conjunto $A_V = \text{Hom}_K(V, V)$ de todos los operadores lineales $\rho: V \rightarrow V$ donde V es un espacio vectorial de dimensión finita con producto escalar. Sabemos que A_V posee una estructura de espacio vectorial. Considere el operador $\phi: A_V \rightarrow A_V$ dado por $\phi(\rho) = \rho^*$.

A continuación estudiaremos el caso específico en que $\phi(\rho) = \rho^*$ es ρ^{-1} .

4.7 DEFINICION. Sean V y V' espacios vectoriales con producto escalar $\langle \cdot, \cdot \rangle_V$ y $\langle \cdot, \cdot \rangle_{V'}$ respectivamente. Un *isomorfismo entre espacios vectoriales con producto escalar* es un isomorfismo $f: V \rightarrow V'$ que preserva productos escalares, es decir, $\langle f(v), f(u) \rangle_{V'} = \langle v, u \rangle_V$ para toda $v, u \in V$.

4.8 DEFINICION. Un *operador unitario (ortogonal)* $\rho: V \rightarrow V$ definido en un espacio vectorial V con producto escalar sobre $K = \mathbb{C}$ ($K = \mathbb{R}$) es un isomorfismo $\rho: V \rightarrow V$ de espacios vectoriales con producto escalar.

Por el problema 4.4, $\rho: V \rightarrow V$ es unitario (ortogonal) si $K = \mathbb{C}$ ($K = \mathbb{R}$) y $\rho^* = \rho^{-1}$. Observe que, por 4.3, si $K = \mathbb{C}$, la matriz asociada a un operador unitario ρ es A (llamada *matriz unitaria*) si, y sólo si $A^* = A^{-1}$, i.e. $AA^* = I = A^*A$. También, si $K = \mathbb{R}$, la matriz asociada a un operador ortogonal ρ es A (llamada *matriz ortogonal*) si, y sólo si ${}^tA = A^{-1}$ (pues $A^* = {}^tA$), i.e. $A({}^tA) = I = {}^tAA$.

Si A es una matriz ortogonal, $A({}^tA) = I$ y como $|A| = |{}^tA|$, $|A{}^tA| = |I| = 1$. Pero $|A{}^tA| = |A||{}^tA| = |A|^2$. Luego $|A| = \pm 1$.

Observe que el conjunto de las matrices ortogonales de $n \times n$ cumplen las propiedades (ii), (iii) y (iv) de I.1.1, es decir, forman un *grupo* que llamaremos *grupo ortogonal* denotado con $O(n)$. El conjunto de matrices ortogonales que poseen determinante 1 lo denotaremos con $SO(n)$ y lo llamaremos *grupo ortogonal especial*.

4.9 DEFINICION. Diremos que un operador $\rho: V \rightarrow V$ es *normal* si conmuta con su adjunto, es decir, si $\rho\rho^* = \rho^*\rho$. Diremos que una matriz compleja A es *normal* si conmuta con su conjugada traspuesta, es decir, si $AA^* = A^*A$.

Observe que los operadores ortogonales y unitarios son normales.

PROBLEMAS

4.1 Compruebe la unicidad de ρ^* del teorema 4.2.

4.2 Pruebe que, si $\rho: V \rightarrow V$ es un operador de V y U un subespacio invariante bajo ρ entonces U^\perp es un subespacio invariante bajo ρ^* .

4.3 Sea λ un valor característico de un operador $\rho: V \rightarrow V$ donde V es un espacio vectorial complejo. Pruebe que

- (i) si $\rho^* = \rho$ entonces $\lambda \in \mathbb{R}$.
- (ii) si $\rho^* = \rho^{-1}$ entonces $|\lambda| = 1$.

4.4 Pruebe que las siguientes afirmaciones son equivalentes:

- (i) $\rho^* = \rho^{-1}$
- (ii) $\langle \rho(v), \rho(u) \rangle = \langle v, u \rangle$ para toda $v, u \in V$
- (iii) $\|\rho(u)\| = \|u\|$ para toda $u \in V$.

Un operador ρ que cumple la afirmación (ii) se dice que *preserva productos escalares* y si cumple (iii) se dice que *preserva longitudes*.

4.5 Pruebe que si A es una matriz ortogonal de 2×2 con $|A| = 1$ entonces A es de la forma

$$\begin{pmatrix} \cos t & -\operatorname{sen} t \\ \operatorname{sen} t & \cos t \end{pmatrix} \quad \text{para alguna } t \in \mathbb{R}.$$

4.6 Pruebe que si A es una matriz ortogonal (o unitaria) entonces los renglones y las columnas forman conjuntos ortonormales y viceversa.

III.5 EL TEOREMA ESPECTRAL

Considere el operador $\phi: A_V \rightarrow A_V$ dado por $\phi(\rho) = \rho^*$ de la sección anterior, donde V es un espacio vectorial de dimensión finita con producto escalar sobre $K = \mathbb{R}$ ó \mathbb{C} .

5.1 DEFINICION. Diremos que un operador $\rho: V \rightarrow V$ es *autoadjunto* si $\phi(\rho) = \rho^* = \rho$. Si $K = \mathbb{R}$ lo llamaremos también *simétrico*; si $K = \mathbb{C}$ lo llamaremos *hermitiano*.

Veamos algunas propiedades de los operadores simétricos:

5.2 PROPOSICION. Sea $\rho: V \rightarrow V$ un operador simétrico. Entonces el polinomio característico $p_\rho(\lambda)$ de ρ se factoriza en factores lineales sobre \mathbb{R} y los vectores característicos de ρ que corresponden a valores característicos distintos son ortogonales.

Demostración. Sea A la matriz asociada a ρ con respecto a una base ortonormal de V . Entonces ${}^t A = A$. Sea $p_A(\lambda)$ el polinomio característico de A . Por el problema 4.3(i) el operador autoadjunto considerado como operador autoadjunto complejo ρ posee únicamente valores característicos reales. Entonces

$$p_A(\lambda) = (\lambda - \lambda_1)(\lambda - \lambda_2) \cdots (\lambda - \lambda_n)$$

donde $\lambda_i \in \mathbb{R}$.

Si $\rho(u) = \lambda u$ y $\rho(v) = \lambda' v$ con $\lambda \neq \lambda'$ entonces $\lambda \langle u, v \rangle = \langle \lambda u, v \rangle = \langle \rho(u), v \rangle = \langle u, \rho(v) \rangle = \langle u, \lambda' v \rangle = \lambda' \langle u, v \rangle$. Luego $(\lambda - \lambda') \langle u, v \rangle = 0$. Como $\lambda \neq \lambda'$, $\langle u, v \rangle = 0$. ■

5.3 TEOREMA. Sea $\rho: V \rightarrow V$ un operador simétrico (V es un espacio vectorial sobre \mathbb{R}). Entonces ρ posee una matriz asociada diagonal.

Demostración. Utilizaremos inducción sobre la dimensión de V . Si $\dim V = 1$ entonces el teorema se cumple. Sea $\dim V = n$ con $n > 1$. Por 5.2 existe al menos un vector característico v_1 diferente de cero de ρ . Sea U el subespacio generado por v_1 y sea $w_1 = \frac{v_1}{\|v_1\|}$. Como v_1 es un vector característico el subespacio U es invariante bajo ρ . Por el problema 4.2 U^\perp es invariante bajo

$\rho^* = \rho$. Así, $\rho|_{U^\perp}$ es un operador simétrico. Por 3.9, $V \cong U \oplus U^\perp$. Como $\dim U = 1$, $\dim U^\perp = n - 1$. Por inducción, existe una base ortonormal $\{w_2, \dots, w_n\}$ de U^\perp que consta de vectores característicos de $\rho|_{U^\perp}$ y por lo tanto de ρ . Como $w_i \in U^\perp$, $\langle w_1, w_i \rangle = 0$ para $i = 2, \dots, n$ y por lo tanto $\{w_i\}_{i=1}^n$ es un conjunto ortonormal de vectores característicos de ρ y por II.1.12 posee una matriz asociada diagonal. ■

5.4 COROLARIO. *Sea A una matriz simétrica real. Entonces existe una matriz ortogonal N tal que la matriz $B = N^{-1}AN = {}^tNAN$ es diagonal. ■*

Observe que por el problema 5.1 y el teorema anterior tenemos que un espacio vectorial sobre \mathbb{R} posee una base ortonormal que consiste de vectores característicos de ρ si, y sólo si, ρ es autoadjunto.

A continuación veamos un resultado semejante al teorema 5.3 para operadores normales en espacios complejos.

5.5 TEOREMA. *Sea $\rho: V \rightarrow V$ un operador normal donde V es un espacio vectorial complejo con producto escalar. Entonces ρ posee una matriz asociada diagonal.*

Demostración. Utilizaremos el proceso de inducción sobre la dimensión de V . Si $\dim V = 1$ el teorema es válido. Supongamos que $\dim V = n$ para $n > 1$. Como el campo es \mathbb{C} , ρ posee al menos un valor característico y por ende un vector característico $w \neq 0$. Sea U el subespacio generado por w y sea v_1 un vector unitario de U . Como w es un vector característico de ρ , U es invariante bajo ρ . Por el problema 5.4(i), w es también un vector característico de ρ^* y por lo tanto, U es invariante bajo ρ^* . Por el problema 4.2, U^\perp es invariante bajo $\rho = \rho^{**}$. Ahora, continúe la demostración en forma análoga a 5.3. ■

5.6 COROLARIO. *Sea A una matriz normal. Entonces existe una matriz unitaria N tal que la matriz $B = N^{-1}AN = N^*AN$ es diagonal. ■*

5.7 EJEMPLO. Consideremos la matriz simétrica $A = \begin{pmatrix} 7 & -2 \\ -2 & 4 \end{pmatrix}$. Encontramos una matriz ortogonal N tal que $B = {}^tNAN$ sea diagonal. El polinomio característico $p_A(\lambda)$ de A es

$$p_A(\lambda) = |\lambda I - A| = \begin{vmatrix} \lambda - 7 & 2 \\ 2 & \lambda - 4 \end{vmatrix} = (\lambda - 3)(\lambda - 8).$$

Entonces $\lambda = 3$ y $\lambda = 8$ son los valores característicos. Para $\lambda = 3$, obtenemos

$$\begin{pmatrix} 7 & -2 \\ -2 & 4 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = 3 \begin{pmatrix} x \\ y \end{pmatrix}$$

lo que nos da el sistema de ecuaciones

$$\left. \begin{array}{l} 4x - 2y = 0 \\ 2x - y = 0 \end{array} \right\}.$$

El vector $(-2, -4)$ es una solución del sistema y por lo tanto es un vector característico correspondiente al valor característico $\lambda = 3$. Para $\lambda = 8$, obtenemos

$$\begin{pmatrix} 7 & -2 \\ -2 & 4 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = 8 \begin{pmatrix} x \\ y \end{pmatrix}$$

lo que nos da el sistema de ecuaciones

$$\left. \begin{array}{l} x + 2y = 0 \\ 2x + 4y = 0 \end{array} \right\}.$$

El vector $(-4, 2)$ es una solución del sistema y por lo tanto es un vector característico correspondiente al valor característico $\lambda = 8$.

Por la proposición 5.2 sabemos que dichos vectores característicos son ortogonales. Normalizamos los vectores característicos y obtenemos una base ortonormal

$$\left\{ \left(-2/\sqrt{20}, -4/\sqrt{20} \right), \left(-4/\sqrt{20}, 2/\sqrt{20} \right) \right\}.$$

Definamos N como la matriz

$$\begin{pmatrix} -2/\sqrt{20} & -4/\sqrt{20} \\ -4/\sqrt{20} & 2/\sqrt{20} \end{pmatrix}.$$

Luego ${}^tNAN = \begin{pmatrix} 3 & 0 \\ 0 & 8 \end{pmatrix} = B$. Observe que la diagonal de B consiste de los valores característicos correspondientes a los vectores característicos.

A continuación estableceremos el *teorema espectral*.

5.8 TEOREMA. Sea $\rho: V \rightarrow V$ un operador simétrico (normal) donde V es un espacio vectorial de dimensión finita con producto escalar sobre $K = \mathbb{R}$ ($K = \mathbb{C}$). Entonces existen proyecciones ortogonales $p_{V_{\lambda_i}}$ de V en V_{λ_i} tales que

- (i) $\rho = \lambda_1 p_{V_{\lambda_1}} + \lambda_2 p_{V_{\lambda_2}} + \cdots + \lambda_s p_{V_{\lambda_s}}$
- (ii) $p_{V_{\lambda_1}} + \cdots + p_{V_{\lambda_s}} = I$
- (iii) $p_{V_{\lambda_i}} \circ p_{V_{\lambda_j}} = 0$ si $j \neq i$.

Demostración. (i) Como ρ es simétrico (normal), por 5.3 (5.5) existe una base ortonormal que consiste de vectores característicos asociados a valores característicos $\lambda_1, \dots, \lambda_s$ de ρ . Luego $V \cong V_{\lambda_1} \oplus \cdots \oplus V_{\lambda_s}$. Sea $v = v_1 + \cdots + v_s$ donde $v_i \in V_{\lambda_i}$. Entonces $\rho(v) = \rho(v_1) + \cdots + \rho(v_s) = \lambda_1 v_1 + \cdots + \lambda_s v_s$. Denotemos con $W = V_{\lambda_1} \oplus \cdots \oplus \widehat{V_{\lambda_i}} \oplus \cdots \oplus V_{\lambda_s}$. Sea $u \in V_{\lambda_i}$, $w = \sum_{j \neq i} w_j$ donde $w_j \in V_{\lambda_j}$. Entonces $\langle u, w \rangle = \langle u, \sum_{j \neq i} w_j \rangle = \sum_{j \neq i} \langle u, w_j \rangle$. Por el problema 5.4, esta última igualdad es 0 y por lo tanto $W = V_{\lambda_i}^\perp$. Así, por definición, el operador $p_{V_{\lambda_i}}: V \rightarrow V$ dado por $p_{V_{\lambda_i}}(v) = v_i$ es la proyección ortogonal sobre V_{λ_i} . Luego $\rho(v) = \lambda_1 p_{V_{\lambda_1}}(v) + \cdots + \lambda_s p_{V_{\lambda_s}}(v)$ y $\rho = \lambda_1 p_{V_{\lambda_1}} + \cdots + \lambda_s p_{V_{\lambda_s}}$.

- (ii) Sea $v = v_1 + \cdots + v_s$. Por (i), $v = p_{V_{\lambda_1}}(v) + \cdots + p_{V_{\lambda_s}}(v) = (p_{V_{\lambda_1}} + \cdots + p_{V_{\lambda_s}})(v)$. Luego $I = p_{V_{\lambda_1}} + \cdots + p_{V_{\lambda_s}}$.
- (iii) $(p_{V_{\lambda_i}} \circ p_{V_{\lambda_j}})(v) = p_{V_{\lambda_i}}(p_{V_{\lambda_j}}(v)) = p_{V_{\lambda_i}}(v_j)$ con $v_j \in V_{\lambda_j}$. Pero $p_{V_{\lambda_i}}(v_j) = 0$ pues $p_{V_{\lambda_i}}$ es la proyección ortogonal. Luego $p_{V_{\lambda_i}} \circ p_{V_{\lambda_j}} = 0$. ■

La descomposición de ρ del teorema anterior se llama *descomposición o resolución espectral* de ρ .

5.9 EJEMPLO. Considere el operador $\rho: \mathbb{R}^3 \rightarrow \mathbb{R}^3$ dado por $\rho(x, y, z) = (4x + 5y, 5x + 4y, 3z)$. Su matriz asociada es

$$A = \begin{pmatrix} 4 & 5 & 0 \\ 5 & 4 & 0 \\ 0 & 0 & 3 \end{pmatrix}$$

la cual es simétrica (y por ende ρ lo es) y posee una descomposición espectral en proyecciones ortogonales. Veamos como se obtienen. Los valores característicos se obtienen de

$$\begin{aligned} \begin{vmatrix} \lambda - 4 & -5 & 0 \\ -5 & \lambda - 4 & 0 \\ 0 & 0 & \lambda - 3 \end{vmatrix} &= (\lambda - 4)^2(\lambda - 3) - 25(\lambda - 3) \\ &= ((\lambda - 4)^2 - 25)(\lambda - 3) \\ &= (\lambda - 9)(\lambda + 1)(\lambda - 3). \end{aligned}$$

Sean $\lambda_1 = 9$, $\lambda_2 = -1$ y $\lambda_3 = 3$ los valores característicos. Ahora encontremos los espacios característicos generados por los vectores característicos asociados a λ_1 , λ_2 y λ_3 :

(i) Para $\lambda_1 = 9$, $\lambda_1 I - A = \begin{pmatrix} 5 & -5 & 0 \\ -5 & 5 & 0 \\ 0 & 0 & 6 \end{pmatrix}$, luego

$$\begin{pmatrix} 5 & -5 & 0 \\ -5 & 5 & 0 \\ 0 & 0 & 6 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

y obtenemos el sistema de ecuaciones $\left. \begin{array}{l} 5x - 5y = 0 \\ -5x + 5y = 0 \\ 6z = 0 \end{array} \right\}$ cuya solución es $x = y = a$ y $z = 0$.

(ii) Para $\lambda_2 = -1$, $\lambda_2 I - A = \begin{pmatrix} -5 & -5 & 0 \\ -5 & -5 & 0 \\ 0 & 0 & -4 \end{pmatrix}$, luego

$$\begin{pmatrix} -5 & -5 & 0 \\ -5 & -5 & 0 \\ 0 & 0 & -4 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

y obtenemos el sistema de ecuaciones $\left. \begin{array}{l} -5x - 5y = 0 \\ -5x - 5y = 0 \\ -4z = 0 \end{array} \right\}$ cuya solución es $x = c$, $y = -c$ y $z = 0$.

(iii) Análogamente, para $\lambda_3 = 3$ obtenemos $x = y = 0$ y $z = b$. Así, los espacios característicos son

$$V_{\lambda_1} = \{(a, a, 0) | a \in \mathbb{R}\}, \quad \dim V_{\lambda_1} = 1$$

$$V_{\lambda_2} = \{(c, -c, 0) | c \in \mathbb{R}\}, \quad \dim V_{\lambda_2} = 1 \text{ y}$$

$$V_{\lambda_3} = \{(0, 0, b) | b \in \mathbb{R}\}, \quad \dim V_{\lambda_3} = 1.$$

Luego $\mathbb{R}^3 \cong V_{\lambda_1} \oplus V_{\lambda_2} \oplus V_{\lambda_3}$. Si $(x, y, z) \in \mathbb{R}^3$ entonces

$$\begin{aligned} (x, y, z) &= (a, a, 0) + (c, -c, 0) + (0, 0, b) \\ &= (a + c, a - c, b). \end{aligned}$$

Como $x = a + c$, $y = a - c$ y $z = b$,

$$x + y = 2a \quad \text{y} \quad -x + y = -2c$$

de donde $a = \frac{x+y}{2}$, $c = \frac{-x+y}{-2} = \frac{x-y}{2}$. Las proyecciones ortogonales son

$$p_{V_{\lambda_1}} = \left(\frac{x+y}{2}, \frac{x+y}{2}, 0 \right)$$

$$p_{V_{\lambda_2}} = \left(\frac{x-y}{2}, \frac{y-x}{2}, 0 \right)$$

$$p_{V_{\lambda_3}} = (0, 0, z)$$

y podemos expresar ρ como

$$\rho(x, y, z) = 9 \left(\frac{x+y}{2}, \frac{x+y}{2}, 0 \right) - \left(\frac{x-y}{2}, \frac{y-x}{2}, 0 \right) + 3(0, 0, z).$$

Es fácil comprobar que $\rho(x, y, z) = (4x + 5y, 5x + 4y, 3z)$.

PROBLEMAS

5.1 Pruebe que si V es un espacio vectorial sobre \mathbb{R} de dimensión finita con producto escalar y $\rho: V \rightarrow V$ es un operador lineal para el cual existe una base ortonormal que consiste de vectores característicos de ρ entonces ρ es simétrico, i.e. es autoadjunto sobre \mathbb{R} .

5.2 (i) Proporcione una demostración del corolario 5.4.

(ii) Pruebe que si A es una matriz hermitiana entonces existe una matriz unitaria N tal que N^*AN es diagonal.

5.3 Pruebe que si ρ es un operador normal entonces

- (i) $\rho - \lambda I$ es normal
- (ii) $\rho^*(v) = 0$ si, y sólo si, $\rho(v) = 0$

5.4 Sea ρ un operador normal. Pruebe que

- (i) si $\rho(v) = \lambda v$ entonces $\rho^*(v) = \bar{\lambda}v$
- (ii) si $\rho(v) = \lambda v$ y $\rho(v') = \lambda'v'$ y $\lambda \neq \lambda'$ entonces $\langle v, v' \rangle = 0$.

5.5 Escriba con detalle el final de la demostración de 5.5.

5.6 (i) Pruebe que si p es una proyección entonces $p^2 = p$.

(ii) Verifique y pruebe en caso de ser válida la doble implicación del siguiente enunciado: Un operador $\rho: V \rightarrow V$ es normal si, y sólo si ρ es autoadjunto.

5.7 Pruebe que si ρ es un operador normal y f un polinomio sobre \mathbb{C} , entonces $f(\rho)$ es un operador normal.

5.8 Sea $A = \begin{pmatrix} 9 & -2 \\ -2 & 6 \end{pmatrix}$ una matriz simétrica. Encuentre una matriz ortogonal N tal que $B = {}^tNAN$ sea diagonal.

5.9 Encuentre la descomposición espectral de $\rho: \mathbb{R}^3 \rightarrow \mathbb{R}^3$ dado por $\rho(x, y, z) = (5x - 2y, -2x + 2y, 4z)$.

Capítulo IV

ALGEBRA MULTILINEAL Y

K-TEORIA ALGEBRAICA CLASICA

IV.1 PRODUCTO TENSORIAL

A continuación definiremos un espacio vectorial en el cual solamente se tienen relaciones bilineales.

1.1 DEFINICION. Sean U y V espacios vectoriales de dimensión finita sobre un campo K . El *producto tensorial* de U y V , es la pareja (T, f) donde T es un espacio vectorial de dimensión finita y $f: U \times V \rightarrow T$ es una función bilineal, tal que si W es un espacio vectorial de dimensión finita y $g: U \times V \rightarrow W$ es bilineal, entonces existe una función lineal única $h: T \rightarrow W$ tal que $g = h \circ f$.

La condición $g = h \circ f$ se puede representar mediante el diagrama

$$\begin{array}{ccc} U \times V & \xrightarrow{f} & T \\ & g \searrow & \downarrow h \\ & & W \end{array}$$

Veamos a continuación que, si existe, el producto tensorial de dos espacios vecto-

riales de dimensión finita es único. Es decir, dados dos productos tensoriales (T, f) y (T', f') de U y V existe un isomorfismo entre T y T' . Esto es inmediato, pues, por ser T un producto tensorial, existe $h: T \rightarrow T'$ tal que $f' = h \circ f$. Análogamente, como T' es un producto tensorial, existe $h': T' \rightarrow T$ tal que $f = h' \circ f'$. Consideremos los siguientes diagramas

$$\begin{array}{ccc}
 & T & \\
 f \nearrow & & \downarrow h \\
 U \times V & \xrightarrow{f'} & T' \\
 f \searrow & & \downarrow h' \\
 & T & \\
 & \xrightarrow{1_T} &
 \end{array}
 \qquad
 \begin{array}{ccc}
 & T' & \\
 f' \nearrow & & \downarrow h' \\
 U \times V & \xrightarrow{f} & T \\
 f' \searrow & & \downarrow h \\
 & T' & \\
 & \xrightarrow{1_{T'}} &
 \end{array}$$

Por ser T un producto tensorial, y como $1_T: T \rightarrow T$ es tal que $1_T \circ f = f$ y también $h' \circ h \circ f = f$, por la unicidad tenemos que $h' \circ h = 1_T$. De manera semejante, por ser T' un producto tensorial, y como $1_{T'}: T' \rightarrow T'$ es tal que $1_{T'} \circ f' = f'$ y también $h \circ h' \circ f' = f'$, se tiene, por unicidad, que $h \circ h' = 1_{T'}$. Por lo tanto, h es un isomorfismo.

Entonces podemos hablar de *el producto tensorial* T de U y V , denotado con $T = U \otimes_K V$ o simplemente $U \otimes V$.

En otras palabras, 1.1 nos dice que cualquier función bilineal $g: U \times V \rightarrow W$ puede expresarse en términos de $f: U \times V \rightarrow T = U \otimes_K V$ como $g(u, v) = h(f(u, v))$ para una función *lineal* única $h: U \otimes_K V \rightarrow W$.

Ahora veamos que, dados dos espacios vectoriales de dimensión finita sobre un campo K , siempre existe su producto tensorial.

1.2 PROPOSICION. *Sean U y V espacios vectoriales de dimensión finita sobre un campo K . Entonces existe un espacio vectorial T de dimensión finita sobre K que cumple la definición 1.1.*

Demostración. Sea $\{u_1, \dots, u_m\}$ una base de U y $\{v_1, \dots, v_n\}$ una base de V . Sea $T = K^{mn}$ el espacio vectorial de dimensión mn sobre K y $\{e_{ij}\}$ con $1 \leq i \leq m$ y $1 \leq j \leq n$ la base canónica. Los elementos de T se pueden expresar en forma única como

$$\sum_{i=1}^m \sum_{j=1}^n \lambda_{ij} e_{ij} \quad \text{con } \lambda_{ij} \in K.$$

Sea $f: U \times V \rightarrow T$ la función dada por

$$f(u, v) = f(\alpha_1 u_1 + \dots + \alpha_m u_m, \beta_1 v_1 + \dots + \beta_n v_n) = \sum_{i=1}^m \sum_{j=1}^n \alpha_i \beta_j e_{ij}.$$

En particular, $f(u_i, v_j) = e_{ij}$.

Veamos que se cumple 1.1. Comprobemos que f es bilinear: sea $u' = \alpha'_1 u_1 + \dots + \alpha'_m u_m$.

$$\begin{aligned} f(u + u', v) &= \sum_{i=1}^m \sum_{j=1}^n (\alpha_i + \alpha'_i) \beta_j e_{ij} \\ &= \sum_{i=1}^m \sum_{j=1}^n (\alpha_i \beta_j + \alpha'_i \beta_j) e_{ij} \\ &= \sum_{i=1}^m \sum_{j=1}^n (\alpha_i \beta_j e_{ij} + \alpha'_i \beta_j e_{ij}) \\ &= \sum_{i=1}^m \sum_{j=1}^n \alpha_i \beta_j e_{ij} + \sum_{i=1}^m \sum_{j=1}^n \alpha'_i \beta_j e_{ij} \\ &= f(u, v) + f(u', v) \end{aligned}$$

Análogamente se tiene que $f(u, v + v') = f(u, v) + f(u, v')$ y que $f(\lambda u, v) = \lambda f(u, v) = f(u, \lambda v)$.

Finalmente, sea $g: U \times V \rightarrow W$ una función bilinear. Por la proposición I.5.2 existe una función lineal única $h: T \rightarrow W$ tal que $h(e_{ij}) = g(u_i, v_j)$. Así, $g(u, v) = g\left(\sum_{i=1}^m \alpha_i u_i, \sum_{j=1}^n \beta_j v_j\right) = \sum_{i=1}^m \sum_{j=1}^n \alpha_i \beta_j g(u_i, v_j) = h\left(\sum_{i=1}^m \sum_{j=1}^n \alpha_i \beta_j e_{ij}\right) = h(f(u, v))$. ■

La función bilinear f se llama *función bilinear universal* (cualquier otra función bilinear $g: U \times V \rightarrow W$ se obtiene de f). Decimos que debido a la propiedad universal, el espacio vectorial $U \otimes_K V$ está determinado en forma única salvo isomorfismo. Es decir, que si en la construcción del espacio vectorial $U \otimes_K V$ hubiéramos tomado bases diferentes para U y V obtendríamos un espacio vectorial isomorfo a $U \otimes_K V$. Aún más, podríamos haber construido el producto tensorial $U \otimes_K V$ sin utilizar bases de U y V tal que 1.1 se cumpla, véase [LL1, p.60].

Para cada $u \in U$ y $v \in V$, el elemento $f(u, v)$ lo escribiremos en la forma $u \otimes v$. Es fácil comprobar (problema 1.9) que $f(U \times V)$ genera el producto tensorial T , el cual denotamos $U \otimes_K V$. De manera que cada elemento de $U \otimes_K V$ se puede escribir en la forma $\sum_{i=1}^r \lambda_i (u_i \otimes v_i)$ con $\lambda_i \in K$, $u_i \in U$, $v_i \in V$. Esta expresión no es única pues de la bilinealidad de f se tiene que

$$\begin{aligned} (\lambda_1 u_1 + \lambda_2 u_2) \otimes v &= \lambda_1 (u_1 \otimes v) + \lambda_2 (u_2 \otimes v) \quad \text{y} \\ u \otimes (\mu_1 v_1 + \mu_2 v_2) &= \mu_1 (u \otimes v_1) + \mu_2 (u \otimes v_2), \end{aligned}$$

donde $\lambda_1, \lambda_2, \mu_1, \mu_2, \in K$; $u_1, u_2, u \in U$ y $v_1, v_2, v \in V$.

Como caso particular se tiene que

$$(\lambda u) \otimes v = \lambda(u \otimes v) = u \otimes (\lambda v).$$

Si $\lambda = -1$ se tiene que $(-u) \otimes v = -(u \otimes v) = u \otimes (-v)$ y si $\lambda = 0$ se tiene que $0 \otimes v = 0 = u \otimes 0$.

Por lo tanto, cualquier elemento de $U \otimes_K V$ puede escribirse en la forma

$$\sum_{i=1}^r (u_i \otimes v_i)$$

donde $u_i \in U$, $v_i \in V$.

A continuación estableceremos algunas propiedades del producto tensorial.

1.3 PROPOSICION. *Sea V un espacio vectorial de dimensión finita n sobre un campo K . Entonces*

$$V \otimes_K K \cong V \cong K \otimes_K V.$$

Demostración. Sea $g: V \times K \rightarrow V$ la función bilineal dada por $g(v, \lambda) = \lambda v$, $\lambda \in K$, $v \in V$. Entonces por 1.1 existe una función lineal única $h: V \otimes_K K \rightarrow V$ tal que $h \circ f = g$, es decir, el siguiente diagrama conmuta:

$$\begin{array}{ccc} V \times K & \xrightarrow{f} & V \otimes_K K \\ & g \searrow & \downarrow h \\ & & V \end{array}$$

La función bilineal g es suprayectiva pues $g(v, 1) = 1 \cdot v = v$. Como $h \circ f = g$ entonces h es suprayectiva.

Veamos que h es inyectiva: sea $x \in V \otimes_K K$. Sea $\{v_i\}_{i=1}^n$ una base de V , x es de la forma $\sum_{i=1}^n (v_i \otimes \lambda_i)$ para $v_i \in V$, $\lambda_i \in K$ y tal que $x = \sum_{i=1}^n (v_i \otimes \lambda_i) = \sum_{i=1}^n (\lambda_i v_i \otimes 1) = (\sum_{i=1}^n \lambda_i v_i) \otimes 1 = v \otimes 1$. Luego $h(x) = h(v \otimes 1) = h(f(v, 1)) = g(v, 1) = 1 \cdot v = v$. Si $h(v \otimes 1) = 0$ entonces $v = 0$ y por lo tanto $x = v \otimes 1 = 0$. Así, h es inyectivo. Dejamos al lector probar que $V \cong K \otimes_K V$ (Problema 1.1). ■

1.4 DEFINICION. Sean V_1, V_2, \dots, V_m, W una colección de espacios vectoriales sobre un campo K . Diremos que una función

$$f: V_1 \times V_2 \times \dots \times V_m \rightarrow W$$

es *multilineal* si para cada $i = 1, \dots, m$

$$\begin{aligned} f(v_1, \dots, v_i + v'_i, \dots, v_m) &= f(v_1, \dots, v_i, \dots, v_m) + f(v_1, \dots, v'_i, \dots, v_m) \quad \text{y} \\ f(v_1, \dots, \lambda v_i, \dots, v_m) &= \lambda f(v_1, \dots, v_i, \dots, v_m) \end{aligned}$$

donde $\lambda \in K$, $v_i, v'_i \in V_i$.

Es decir, f es lineal en cada variable cuando las otras se mantienen fijas. También llamaremos *forma multilineal* a una función multilineal con codominio K . Podemos enunciar la definición equivalente a 1.1 para productos multitenoriales.

1.5 DEFINICION. Sean $\{V_i\}_{i=1}^m$ espacios vectoriales sobre un campo K . El *producto tensorial* de $\{V_i\}$ es una pareja (T, f) donde T es un espacio vectorial sobre K y f es una función multilineal $f: V_1 \times \dots \times V_m \rightarrow T$ tal que si W es un espacio vectorial sobre un campo K y $g: V_1 \times \dots \times V_m \rightarrow W$ es multilineal, entonces existe una función lineal única $h: T \rightarrow W$ tal que $g = h \circ f$, es decir, tal que el siguiente diagrama conmuta:

$$\begin{array}{ccc} V_1 \times \dots \times V_m & \xrightarrow{f} & T \\ & g \searrow & \downarrow h \\ & & W \end{array}$$

Denotaremos a T con $V_1 \otimes \dots \otimes V_m$ o con $\otimes_{i=1}^m V_i$ y es fácil comprobar la unicidad y existencia de T .

1.6 PROPOSICION. Sean U, V, W espacios vectoriales sobre un campo K . Entonces $(U \otimes V) \otimes W \cong U \otimes (V \otimes W) \cong U \otimes V \otimes W$.

Demostración. Consideremos la función bilineal $g'': U \times V \rightarrow U \otimes V \otimes W$ dada por $g''(u, v) = u \otimes v \otimes w$ para $w \in W$ fija, la cual induce una función lineal $h_w: U \otimes V \rightarrow U \otimes V \otimes W$ tal que $h_w(u \otimes v) = u \otimes v \otimes w$. Sea $g: (U \otimes V) \times W \rightarrow U \otimes V \otimes W$ dada por $g(t, w) = h_w(t)$. g es bilineal y por lo tanto induce una función lineal $h: (U \otimes V) \otimes W \rightarrow U \otimes V \otimes W$ tal que $h((u \otimes v) \otimes w) = u \otimes v \otimes w$.

Construyamos ahora una función $h': U \otimes V \otimes W \rightarrow (U \otimes V) \otimes W$ tal que $h' \circ h = 1_{(U \otimes V) \otimes W}$ y $h \circ h' = 1_{U \otimes V \otimes W}$. Para construir h' considere la función $g': U \times V \times W \rightarrow (U \otimes V) \otimes W$ dada por $g'(u, v, w) = (u \otimes v) \otimes w$. g' es lineal en cada variable, luego induce una función lineal $h': U \otimes V \otimes W \rightarrow (U \otimes V) \otimes W$ tal que $h(u \otimes v \otimes w) = (u \otimes v) \otimes w$. Es inmediato comprobar que $h' \circ h = 1_{(U \otimes V) \otimes W}$

y que $h \circ h' = 1_{U \otimes V \otimes W}$ y, por lo tanto, h y h' son isomorfismos. La demostración de que $U \otimes (V \otimes W) \cong U \otimes V \otimes W$ es análoga. ■

1.7 DEFINICION. Sean V_1, \dots, V_m espacios vectoriales de dimensión finita sobre un campo K . Diremos que la sucesión

$$0 \xrightarrow{f_0} V_1 \xrightarrow{f_1} V_2 \xrightarrow{f_2} \cdots \longrightarrow V_{m-1} \xrightarrow{f_{m-1}} V_m \xrightarrow{f_m} 0$$

es *exacta en V_i* si $\text{im } f_{i-1} = \ker f_i$. Diremos que la sucesión es *exacta* si es exacta en cada V_i , para toda $i = 1, \dots, m$.

1.8 PROPOSICION. Sean V_1, \dots, V_m espacios vectoriales de dimensión finita sobre un campo K y

$$0 \xrightarrow{f_0} V_1 \xrightarrow{f_1} V_2 \xrightarrow{f_2} V_3 \xrightarrow{f_3} \cdots \xrightarrow{f_{m-2}} V_{m-1} \xrightarrow{f_{m-1}} V_m \xrightarrow{f_m} 0$$

una sucesión exacta. Entonces

$$\dim V_1 - \dim V_2 + \dim V_3 - \cdots + (-1)^{m-1} \dim V_m = 0.$$

Demostración. Utilizaremos el proceso de inducción sobre m . Para $m = 1$ tenemos la sucesión $0 \xrightarrow{f_0} V_1 \xrightarrow{f_1} 0$. Por exactitud, $\text{im } f_0 = 0 = \ker f_1 = V_1$. Luego $V_1 = 0$ y $\dim V_1 = 0$. Supongamos que la proposición es válida para $m - 1$. Como $f_2: V_2 \rightarrow V_3$ posee núcleo $\ker f_2 = \text{im } f_1$, induce una función lineal $V_2/\text{im } f_1 \rightarrow V_3$ (problema II.4.8) que es inyectiva. Luego, se tiene una sucesión exacta

$$0 \longrightarrow V_2/\text{im } f_1 \longrightarrow V_3 \longrightarrow \cdots \longrightarrow V_{m-1} \longrightarrow V_m \longrightarrow 0$$

y, por hipótesis de inducción $\dim V_2/\text{im } f_1 - \dim V_3 + \cdots = 0$, es decir, por II.4.5 $\dim V_2 - \dim (\text{im } f_1) - \dim V_3 + \cdots = 0$ y por I.3.11 $\dim V_1 = \dim (\text{im } f_1) + \dim (\ker f_1) = \dim (\text{im } f_1) + 0 = \dim (\text{im } f_1)$ pues $\text{im } f_0 = 0 = \ker f_1$. Luego $\dim V_2 - \dim V_1 - \dim V_3 + \cdots = 0$. ■

1.9 COROLARIO. Si $0 \rightarrow V_1 \rightarrow V_2 \rightarrow V_3 \rightarrow 0$ es una sucesión exacta corta de espacios vectoriales de dimensión finita sobre un campo K entonces $V_2 \cong V_1 \oplus V_3$.

Demostración. Por 1.8, $\dim V_1 - \dim V_2 + \dim V_3 = 0$. Luego $\dim V_2 = \dim V_1 + \dim V_3$ y por lo tanto $V_2 \cong V_1 \oplus V_3$. ■

PROBLEMAS

1.1 Pruebe que $K^n \otimes K^m = K^{nm}$ y que $K \otimes_K V \cong V$ para V un espacio vectorial como en 1.3.

1.2 Proporcione con todo detalle la demostración de 1.6

1.3 Pruebe que $U \otimes V \cong V \otimes U$.

1.4 Pruebe que si

$$T \xrightarrow{f} U \xrightarrow{g} V \xrightarrow{h} W$$

es una sucesión exacta de espacios vectoriales sobre un campo K , entonces

- (i) h es inyectiva si, y sólo si, g es *trivial* (i.e. $g(U) = 0$)
- (ii) g es trivial si, y sólo si, f es suprayectivo.

1.5 Pruebe que, en una sucesión exacta de espacios vectoriales sobre un campo K

$$S \xrightarrow{f} T \xrightarrow{g} U \xrightarrow{h} V \xrightarrow{k} W$$

f es suprayectiva y k es inyectiva si, y sólo si, $U = 0$.

1.6 Pruebe que, si $0 \rightarrow V \rightarrow 0$ es una sucesión exacta de espacios vectoriales sobre un campo K , entonces $V = 0$.

1.7 Sea $R \xrightarrow{f} S \xrightarrow{g} T \xrightarrow{h} U \xrightarrow{k} V \xrightarrow{q} W$ una sucesión exacta de espacios vectoriales sobre un campo K . Pruebe que g, k son funciones lineales triviales si, y sólo si, h es isomorfismo, y que h es isomorfismo si, y sólo si, f es suprayectiva y q es inyectiva.

1.8 Pruebe que, si $0 \rightarrow U \xrightarrow{h} V \rightarrow 0$ es una sucesión exacta de espacios vectoriales sobre un campo K , entonces h es un isomorfismo.

1.9 Verifique que $f(U \times V)$ genera a $U \otimes_K V$. (Sugerencia: defina una función lineal $i: U \times V \rightarrow U \otimes_K V$ y utilice la unicidad de 1.1 para mostrar que i es suprayectiva.)

IV.2 PRODUCTO EXTERIOR

2.1 DEFINICION. Sean V y W espacios vectoriales de dimensión finita sobre un campo K . Diremos que la función bilinear $f: V \times V \rightarrow W$ es *alternante* si $f(v, v) = 0$ para toda $v \in V$.

2.2 DEFINICION. La *potencia exterior (de grado 2)* de un espacio vectorial V de dimensión finita sobre un campo K , es una pareja $(V \wedge V, f)$ en donde $V \wedge V$ es un espacio vectorial sobre K y $f: V \times V \rightarrow V \wedge V$ es una función bilinear alternante, tal que para todo espacio vectorial W y toda $g: V \times V \rightarrow W$ función bilinear alternante, existe una función lineal única $h: V \wedge V \rightarrow W$ tal que $g = h \circ f$, es decir, tal que el siguiente diagrama conmuta:

$$\begin{array}{ccc} V \times V & \xrightarrow{f} & V \wedge V \\ & g \searrow & \downarrow h \\ & & W \end{array}$$

Análogamente al caso de suma directa y al de producto tensorial se demuestra (problema 2.1) que, si existe, la potencia exterior es única.

La existencia se demuestra fácilmente utilizando los conceptos de producto tensorial y espacio vectorial cociente:

2.3 PROPOSICION. Sea V un espacio vectorial de dimensión finita sobre K y $(V \otimes V, f')$ el producto tensorial. Sea U el subespacio de $V \otimes V$ generado por los elementos $v \otimes v$ con $v \in V$. Sea $p: V \otimes V \rightarrow (V \otimes V)/U$ la proyección canónica. Entonces la pareja $((V \otimes V)/U, f)$, con $f = pf'$ es la potencia exterior $V \wedge V$.

Demostración. Veamos que f es bilinear alternante. En efecto, es bilinear porque f' lo es y p es lineal; es alternante porque $f(v, v) = pf'(v, v) = p(v \otimes v) = (v \otimes v) + U = U$ que es el cero de $(V \otimes V)/U$.

Demostraremos que satisface (la propiedad universal de) la definición de potencia exterior. Dada una función bilineal alternante arbitraria $g: V \times V \rightarrow W$ consideremos el diagrama

$$\begin{array}{ccccc} V \times V & \xrightarrow{f'} & V \otimes V & \xrightarrow{p} & V \otimes V/U = V \wedge V \\ & g \searrow & \downarrow h' & \swarrow h & (f = pf') \\ & & W & & \end{array}$$

Por ser $V \otimes V$ el producto tensorial existe h' , función lineal tal que $g = h'f'$. Así, $h'(v \otimes v) = h'f'(v, v) = g(v, v) = 0$. Luego, $h'(v \otimes v) = h'f'(v, v) = 0$ para toda $v \in V$, lo cual implica que $h'(U) = 0$. Entonces existe una función lineal única h tal que $h' = hp$. Por lo tanto $g = hp f' = hf$. ■

Observe que como $im f'$ genera a $V \otimes V$ y p es suprayectiva, resulta que $im f$ genera a $V \wedge V$.

Si $f: V \times V \rightarrow V \wedge V$ es la función bilineal alternante que da la potencia exterior $V \wedge V$ usaremos la notación

$$f(u, v) = u \wedge v.$$

2.4 EJEMPLO. Con la notación anterior, también resulta que $u \wedge v = -v \wedge u$. En efecto, ya que $f(u+v, u+v) = 0$, se tiene que $f(u, u) + f(u, v) + f(v, u) + f(v, v) = 0$ de donde $f(u, v) = -f(v, u)$.

2.5 EJEMPLO. Sea V un espacio vectorial de dimensión 3 y $(V \wedge V, f)$ la potencia exterior. Sea $\{u_i\}_{i=1}^3$ una base de V . Si $u = \sum_{i=1}^3 \alpha_i u_i$, $v = \sum_{i=1}^3 \beta_i u_i$ entonces

$$\begin{aligned} f(u, v) &= f\left(\sum \alpha_i u_i, \sum \beta_j u_j\right) \\ &= \alpha_1 \beta_1 f(u_1, u_1) + \alpha_1 \beta_2 f(u_1, u_2) + \alpha_1 \beta_3 f(u_1, u_3) + \alpha_2 \beta_1 f(u_2, u_1) \\ &\quad + \alpha_2 \beta_2 f(u_2, u_2) + \alpha_2 \beta_3 f(u_2, u_3) + \alpha_3 \beta_1 f(u_3, u_1) + \alpha_3 \beta_2 f(u_3, u_2) \\ &\quad + \alpha_3 \beta_3 f(u_3, u_3) \\ &= \sum_{i=1}^3 \sum_{j=1}^3 \alpha_i \beta_j f(u_i, u_j). \end{aligned}$$

Por ser f alternante $f(u_i, u_i) = 0$, $f(u_2, u_1) = -f(u_1, u_2)$, $f(u_3, u_1) = -f(u_1, u_3)$ y $f(u_3, u_2) = -f(u_2, u_3)$. Así obtenemos

$$f(u, v) = (\alpha_1\beta_2 - \alpha_2\beta_1)f(u_1, u_2) + (\alpha_1\beta_3 - \alpha_3\beta_1)f(u_1, u_3) + (\alpha_2\beta_3 - \alpha_3\beta_2)f(u_2, u_3),$$

es decir,

$$u \wedge v = (\alpha_1\beta_2 - \alpha_2\beta_1)u_1 \wedge u_2 + (\alpha_1\beta_3 - \alpha_3\beta_1)u_1 \wedge u_3 + (\alpha_2\beta_3 - \alpha_3\beta_2)u_2 \wedge u_3.$$

Mostraremos a continuación que $\{u_1 \wedge u_2, u_1 \wedge u_3, u_2 \wedge u_3\}$ es una base de $V \wedge V$ y, por lo tanto la dimensión de la potencia exterior, en este caso, es 3.

Acabamos de demostrar que todo elemento de $\text{im } f$ es combinación lineal de estos tres vectores y como la $\text{im } f$ genera a $V \wedge V$, el conjunto $\{u_1 \wedge u_2, u_1 \wedge u_3, u_2 \wedge u_3\}$ genera a $V \wedge V$.

Para demostrar que son linealmente independientes podemos utilizar la propiedad universal de la manera siguiente. Sea W un espacio vectorial de dimensión 3 y $\{t_{12}, t_{13}, t_{23}\}$ una base de W . Sea $g: V \times V \rightarrow W$ la función dada por

$$g(u, v) = \sum_{1 \leq i, j \leq 3} (\alpha_i\beta_j - \alpha_j\beta_i)t_{ij}$$

en donde $u = \sum_{i=1}^3 \alpha_i u_i$, $v = \sum_{j=1}^3 \beta_j u_j$. En particular se tiene que $g(u_i, u_j) = t_{ij}$. Esta función (problema 2.3) es bilineal y alternante. Entonces, por la propiedad universal, existe una función lineal única $h: V \wedge V \rightarrow W$ tal que $g = hf$. Obtenemos así $g(u, v) = hf(u, v) = h(u \wedge v)$ y, en particular, $h(u_i \wedge u_j) = g(u_i, u_j) = t_{ij}$. Supongamos ahora que $\sum \lambda_{ij} u_i \wedge u_j = 0$ para $(1 \leq i < j \leq 3)$. Entonces como h es lineal, $h(\sum \lambda_{ij} u_i \wedge u_j) = \sum \lambda_{ij} h(u_i \wedge u_j) = \sum \lambda_{ij} g(u_i, u_j) = \sum \lambda_{ij} t_{ij} = 0$. Como $\{t_{12}, t_{13}, t_{23}\}$ es base de W , entonces $\lambda_{12} = \lambda_{13} = \lambda_{23} = 0$.

2.6 EJEMPLO. Si $V = \mathbb{R}^3$, recuérdese que el producto vectorial o producto cruz de dos vectores $u = (\alpha_1, \alpha_2, \alpha_3)$ y $v = (\beta_1, \beta_2, \beta_3)$ se define como $u \times v = (\alpha_2\beta_3 - \alpha_3\beta_2, \alpha_3\beta_1 - \alpha_1\beta_3, \alpha_1\beta_2 - \alpha_2\beta_1)$, es decir,

$$u \times v = (\alpha_2\beta_3 - \alpha_3\beta_2)e_1 + (\alpha_3\beta_1 - \alpha_1\beta_3)e_2 + (\alpha_1\beta_2 - \alpha_2\beta_1)e_3$$

lo cual puede expresarse como

$$u \times v = \begin{vmatrix} e_1 & e_2 & e_3 \\ \alpha_1 & \alpha_2 & \alpha_3 \\ \beta_1 & \beta_2 & \beta_3 \end{vmatrix}$$

En particular $e_1 \times e_2 = e_3$, $e_2 \times e_3 = e_1$, $e_3 \times e_1 = e_2$, por lo que podemos escribir

$$u \times v = (\alpha_1\beta_2 - \alpha_2\beta_1)e_1 \times e_2 + (\alpha_1\beta_3 - \alpha_3\beta_1)e_1 \times e_3 + (\alpha_2\beta_3 - \alpha_3\beta_2)e_2 \times e_3.$$

Si comparamos esta expresión de $u \times v$ con la expresión para $u \wedge v$ del ejemplo anterior (con $\{u_1, u_2, u_3\} = \{e_1, e_2, e_3\}$) vemos que $\mathbb{R}^3 \wedge \mathbb{R}^3 = \mathbb{R}^3$ y $u \wedge v = u \times v$; la potencia exterior, en este caso, es el producto vectorial.

2.7 PROPOSICION. *Si $\dim V = n$ entonces $\dim V \wedge V = \binom{n}{2}$.*

Demostración. Sea $\{v_i\}_{i=1}^n$ una base de V . Sabemos que $\{v_i \otimes v_j\}$ ($i, j = 1, \dots, n$) es una base de $V \otimes V$ y como $p: V \otimes V \rightarrow V \wedge V$ es suprayectiva, $\{v_i \wedge v_j\}$ ($i, j = 1, \dots, n$) genera a $V \wedge V$. Como $v_i \wedge v_j = 0$ ($i = j$) y $v_i \wedge v_j = -v_j \wedge v_i$ ($i \neq j$) de este conjunto podemos eliminar los vectores innecesarios y quedarnos con $\{v_i \wedge v_j\}$ ($1 \leq i < j \leq n$) y sigue generando $V \wedge V$. Además (problema 2.4) este conjunto es linealmente independiente, por lo que es base de $V \wedge V$. Como el número de vectores es $\frac{n(n-1)}{2} = \binom{n}{2}$, la dimensión de $V \wedge V$ es $\binom{n}{2}$. ■

El concepto de potencia exterior puede extenderse a más de dos factores. Diremos que una función multilinear $f: \times_{i=1}^k V_i \rightarrow W$ con $V_i = V$ es *alternante* si $f(v_1, \dots, v_k) = 0$ siempre que $v_i = v_j$ para algunas $i \neq j$.

2.8 DEFINICION. *La potencia exterior de grado k* de un espacio vectorial V de dimensión finita sobre un campo K es una pareja $(\wedge^k V, f)$ en donde $\wedge^k V$ es un espacio vectorial y $f: \times_{i=1}^k V_i \rightarrow \wedge^k V$ ($V_i = V$) es una función multilinear alternante tal que para todo espacio vectorial W y toda $g: \times_{i=1}^k V_i \rightarrow W$ ($V_i = V$) función multilinear alternante, existe una única función lineal $h: \wedge^k V \rightarrow W$ tal que $g = h \circ f$, es decir, tal que el siguiente diagrama conmuta:

$$\begin{array}{ccc} \times_{i=1}^k V_i & \xrightarrow{f} & \wedge^k V \quad (V_i = V) \\ & g \searrow & \downarrow h \\ & & W \end{array}$$

La demostración de la unicidad y existencia de la potencia exterior sigue exactamente los mismos pasos que para el caso de $\wedge^2 V = V \wedge V$.

Dejamos al lector imitar el caso $k = 2$ para probar (problema 2.6) que si $\{v_1, \dots, v_n\}$ es una base de V entonces

$$\{v_{j_1} \wedge \dots \wedge v_{j_k} \mid 1 \leq j_1 < \dots < j_k \leq n\}$$

es base de $\bigwedge^k V$. Puesto que este conjunto tiene $\binom{n}{k}$ elementos, la dimensión de $\bigwedge^k V$ es $\binom{n}{k}$.

Si denotamos $J = \{j_1, \dots, j_k\}$ con $1 \leq j_1 < \dots < j_k \leq n$ entonces los elementos $v \in \bigwedge^k V$ son de la forma $v = \sum_J \alpha_J v_J$ en donde $v_J = v_{j_1} \wedge \dots \wedge v_{j_k}$.

Para completar la definición de $\bigwedge^k V$ definimos $\bigwedge^0 V = K$ y $\bigwedge^1 V = V$. Obsérvese que para índices $k > n = \dim V$, claramente $\bigwedge^k V = 0$.

En lo que sigue analizaremos la relación que existe entre los conceptos de potencia exterior y determinante.

Sea V un espacio vectorial de dimensión n y $A: V \rightarrow V$ un *endomorfismo* (sinónimo de operador lineal) de V , es decir una función lineal de V en V . Defínase una función $g = g_A: \times_{i=1}^n V_i \rightarrow \bigwedge^n V$ ($V_i = V$) dada por

$$g_A(v_1, \dots, v_n) = A(v_1) \wedge \dots \wedge A(v_n)$$

para $\{v_1, \dots, v_n\}$ una base de V . Como g es multilineal alternante existe una función lineal única $h = h_A: \bigwedge^n V \rightarrow \bigwedge^n V$ tal que

$$h_A(v_1 \wedge \dots \wedge v_n) = g_A(v_1, \dots, v_n) = A(v_1) \wedge \dots \wedge A(v_n)$$

$$\begin{array}{ccc} \times_{i=1}^n V_i & \xrightarrow{f} & \bigwedge^n V & & (v_1, \dots, v_n) & \xrightarrow{f} & v_1 \wedge \dots \wedge v_n \\ & & \downarrow h & & \downarrow h & & \\ & & \bigwedge^n V & & A(v_1) \wedge \dots \wedge A(v_n) & & \end{array}$$

Pero $\dim \bigwedge^n V = 1$, es decir $\bigwedge^n V = K$ por lo que $h_A: K \rightarrow K$ es una función lineal y por lo tanto h_A es la multiplicación por un escalar. Entonces, si denotamos con $|A|$ a este escalar, tenemos

$$h_A(v_1 \wedge \dots \wedge v_n) = |A|(v_1 \wedge \dots \wedge v_n) = A(v_1) \wedge \dots \wedge A(v_n).$$

Entonces definimos el *determinante* de $A: V \rightarrow V$ como el escalar $|A|$.

Si (α_{ij}) es la matriz de la función lineal $A: V \rightarrow V$ con respecto a la base $\{v_1, \dots, v_n\}$, es decir, si $A(v_i) = \sum_{j=1}^n \alpha_{ij} v_j$ ($1 \leq i \leq n$) entonces obtenemos

$$A(v_1) \wedge \dots \wedge A(v_n) = \sum \alpha_{1j_1} \dots \alpha_{nj_n} v_{j_1} \wedge \dots \wedge v_{j_n}; \quad 1 \leq j_i \leq n, i = 1, \dots, n$$

Ahora bien, en esta suma los productos $v_{j_1} \wedge \dots \wedge v_{j_n}$ con dos (o más) índices iguales valen 0. Es decir, solamente quedan aquellos productos tales que los índices

j_1, \dots, j_n son todos distintos entre sí. Dicho de otra manera, los índices son las *permutaciones* de $1, \dots, n$, es decir $\sigma(i) = j_i$.

Además, como $v_{j_1} \wedge \dots \wedge v_{j_n} = \text{sig}\sigma v_1 \wedge \dots \wedge v_n$ obtenemos

$$A(v_1) \wedge \dots \wedge A(v_n) = \sum_{\sigma} \text{sig}\sigma \alpha_{1\sigma(1)} \dots \alpha_{n\sigma(n)} v_1 \wedge \dots \wedge v_n$$

y como $|A|v_1 \wedge \dots \wedge v_n = A(v_1) \wedge \dots \wedge A(v_n)$ obtenemos $|A| = \text{sig}\sigma \alpha_{1\sigma(1)} \dots \alpha_{n\sigma(n)}$. Entonces definimos el *determinante de una matriz* (α_{ij}) como el determinante de la función lineal $A: V \rightarrow V$ dada por $A(v_i) = \sum_{j=1}^n \alpha_{ij} v_j$ ($i = 1, \dots, n$). Se denotará con $|\alpha_{ij}|$.

A continuación definiremos una multiplicación la cual denotaremos también con el símbolo \wedge por conveniencia. Sea

$$\wedge: \bigwedge^k V \times \bigwedge^\ell V \longrightarrow \bigwedge^{k+\ell} V$$

dado por

$$\wedge(u_1 \wedge \dots \wedge u_k, v_1 \wedge \dots \wedge v_\ell) = u_1 \wedge \dots \wedge u_k \wedge v_1 \wedge \dots \wedge v_\ell.$$

Esta multiplicación suele llamarse *producto exterior* y posee las siguientes propiedades:

- (i) \wedge es asociativo
- (ii) \wedge es distributivo con respecto a la suma directa.
- (iii) $u_1 \wedge \dots \wedge u_k \wedge v_1 \wedge \dots \wedge v_\ell = (-1)^{k\ell} v_1 \wedge \dots \wedge v_\ell \wedge u_1 \wedge \dots \wedge u_k$.

PROBLEMAS

2.1 Pruebe la unicidad de la potencia exterior.

2.2 Sea V un espacio vectorial sobre un campo K y $u, v \in V$. Pruebe que si u y v son linealmente dependientes entonces $u \wedge v = 0$.

2.3 Pruebe que la función g definida en el ejemplo 2.5 es bilineal y alternante.

2.4 Pruebe que el conjunto $\{v_i \wedge v_j\}$ ($1 \leq i < j \leq n$) de la demostración de 2.7 es linealmente independiente. (Siga el método usado en el ejemplo 2.5).

2.5 Pruebe que si σ es una permutación de $\{1, 2, \dots, k\}$ entonces $v_{\sigma(1)} \wedge \dots \wedge v_{\sigma(k)} = \text{sig}(\sigma)v_1 \wedge \dots \wedge v_k$, en donde $\text{sig}(\sigma)$ denota el signo de la permutación.

2.6 Pruebe que si $\{v_1, \dots, v_n\}$ es una base de V entonces $\{v_{j_1} \wedge \dots \wedge v_{j_k} \mid 1 \leq j_1 < \dots < j_k \leq n\}$ es base de $\bigwedge^k V$ y que el número de tales vectores es $\binom{n}{k}$.

2.7 Pruebe que si $A, B: V \rightarrow V$ son endomorfismos de V entonces

(i) $|AB| = |A||B|$

(ii) $|1_V| = 1$

(iii) $|A^{-1}| = |A|^{-1}$ para A biyectivo (i.e. A *automorfismo* de V)

2.8 Pruebe que si una columna de la matriz (α_{ij}) es combinación de las demás, entonces $|\alpha_{ij}| = 0$.

2.9 Verifique que el producto exterior posee las tres propiedades mencionadas.

IV.3 ESTRUCTURAS ALGEBRAICAS

En esta sección definiremos varias estructuras algebraicas algunas de las cuales ya han sido implícitamente estudiadas. Tiene como finalidad la de (i) presentar un breve panorama de las estructuras algebraicas; (ii) definir el álgebra tensorial y exterior, (iii) situar al lector en la posibilidad de entender los objetos de estudio del álgebra lineal y (iv) proporcionar los objetos necesarios para el estudio de la K -Teoría Algebraica Clásica que haremos en la próxima sección.

Sea $(V, +, \mu)$ un espacio vectorial sobre un campo K tal como se definió en I.1.1. Si quitamos la multiplicación escalar μ nos quedaremos con un conjunto con una operación binaria $+$ que cumple (i) a (iv) de I.1.1. Entonces diremos que $(V, +)$ es un grupo conmutativo bajo $+$. Formalmente tenemos la siguiente:

3.1 DEFINICION. Un *grupo* es una pareja $(G, +)$ donde G es un conjunto y $+: G \times G \rightarrow G$ es una operación binaria $(u, v) \mapsto +(u, v)$ donde, por abuso de notación se escribe $+(u, v) = u + v$, tal que

- (i) $(u + v) + w = u + (v + w)$
- (ii) existe un elemento $0 \in G$ tal que $v + 0 = v$
- (iii) para cada $v \in G$ existe un elemento, denotado con $-v$, tal que $v + (-v) = 0$.

Diremos que el grupo es *conmutativo* si además satisface

- (iv) $u + v = v + u$.

Si en la definición anterior consideramos un conjunto S con una operación binaria $+$ que cumpla (i) diremos que $(S, +)$ es un *semigrupo*.

También, si en la definición 3.1 consideramos un conjunto M con una operación binaria $+$ que cumpla (i) y (ii) diremos que $(M, +)$ es un *monoide*.

3.2 EJEMPLOS. El lector podrá comprobar que $(\mathbb{Z}, +)$, $(\mathbb{Q}, +)$, $(\mathbb{Q}^* = \mathbb{Q} - \{0\}, \cdot)$, $(\mathbb{R}, +)$, $(\mathbb{R}^* = \mathbb{R} - \{0\}, \cdot)$, $(\mathbb{C}, +)$, $(\mathbb{C}^* = \mathbb{C} - \{0\}, \cdot)$, $(\mathbb{Z}_n, +)$, $(V, +)$, $(M_n K, +)$ son grupos.

Para relacionar dos grupos es necesario definir una función que preserve la estructura de grupo.

3.3 DEFINICION. Sean (G, \diamond) y (G', \star) dos grupos. Un *homomorfismo* de grupos es una función $f: G \rightarrow G'$ tal que $f(u \diamond v) = f(u) \star f(v)$.

3.4 DEFINICION. Un *anillo* es una terna $(\Lambda, +, \cdot)$ donde Λ es un conjunto, $+$ y \cdot son operaciones binarias tales que

- (i) $(\Lambda, +)$ es un grupo conmutativo
- (ii) (Λ, \cdot) es un semigrupo
- (iii) $u(v + w) = uv + uw$ y $(u + v)w = uw + vw$

3.5 EJEMPLOS. El lector podrá comprobar que $(\mathbb{Z}, +, \cdot)$, $(\mathbb{Z}_n, +, \cdot)$, $(\mathbb{Q}, +, \cdot)$, $(\mathbb{R}, +, \cdot)$, $(M_n K, +, \cdot)$, $(K, +, \cdot)$, $(K[x], +, \cdot)$, $(\mathbb{C}, +, \cdot)$ son anillos.

Si un anillo $(\Lambda, +, \cdot)$ satisface

- (iv) (Λ, \cdot) es un semigrupo conmutativo, entonces $(\Lambda, +, \cdot)$ se llamará *anillo conmutativo*.

Si en 3.4 (Λ, \cdot) es un monoide, diremos que $(\Lambda, +, \cdot)$ es un *anillo con identidad* o *con uno*.

Si el anillo $(\Lambda, +, \cdot)$ no posee divisores de cero, se llamará *dominio entero*. Si un dominio entero posee un inverso multiplicativo para cada elemento no nulo, se dice que es un *anillo con división*. Un *dominio euclidiano* es un dominio entero Λ junto con una función $d: \Lambda \rightarrow \mathbb{Z}^+$ tal que (i) para todo $x, y \in \Lambda$, distintos de cero, $d(x) \leq d(xy)$ y (ii) para todo $x, y \in \Lambda$, distintos de cero, existen $q, r \in \Lambda$ tales que $x = qy + r$ donde $r = 0$ o $d(r) < d(y)$ (*algoritmo de la división*). Finalmente, un *campo* es un anillo conmutativo con división.

¿Cómo se relacionan dos anillos? Mediante funciones que preserven la estructura de anillos. Si $(\Lambda, \diamond, \star)$ y $(\Lambda', +, \cdot)$ son anillos, un *homomorfismo de anillos* es una función que es un homomorfismo del grupo conmutativo de Λ en el grupo conmutativo de Λ' y que también es un homomorfismo del semigrupo de Λ en el semigrupo de Λ' , es decir,

$$\begin{aligned} f(u \diamond v) &= f(u) + f(v) & \text{y} \\ f(u \star v) &= f(u) \cdot f(v). \end{aligned}$$

Si en la definición I.1.1 consideramos un anillo $(\Lambda, +, \cdot)$ conmutativo con 1 en lugar de un campo K , obtendremos una estructura algebraica llamada Λ -*módulo*

(izquierdo). Entonces, como caso particular de los Λ -módulos están los K -módulos, i.e. los espacios vectoriales sobre un campo K .

Todos los puntos tratados en las secciones I.1 y I.2 son válidos para los Λ -módulos, basta tomar $K = \Lambda$ un anillo conmutativo con 1. En particular, relacionamos dos Λ -módulos mediante un *homomorfismo de Λ -módulos* (definición I.1.7). Los Λ -módulos son generalizaciones de los conceptos de grupo conmutativo y de espacio vectorial, y son los objetos de estudio del Algebra Homológica. Imitando a los espacios vectoriales, si un Λ -módulo posee una *base*, le llamaremos *Λ -módulo libre*. No todo Λ -módulo posee base, es decir, no todo Λ -módulo es libre, pero todo espacio vectorial o K -módulo es libre, es decir, sí posee una base. Diremos que un Λ -módulo es *proyectivo* si es sumando directo de un libre y que es *finitamente generado* si posee un conjunto finito de generadores.

Un *álgebra sobre Λ* es un conjunto A que simultáneamente es un anillo y un Λ -módulo. Es decir, un algebra $(A, +, \mu, \cdot)$ es un Λ -módulo con otra operación binaria, llamada *multiplicación* con una condición extra que hace compatibles las operaciones binarias y multiplicación escalar, la cual es la siguiente:

$$(\lambda u + \lambda' v)w = \lambda(uw) + \lambda'(vw) \quad \text{y}$$

$$w(\lambda u + \lambda' v) = \lambda(wu) + \lambda'(wv) \quad \text{para } \lambda, \lambda' \in \Lambda; u, v, w \in A$$

En particular se tiene que $(\lambda u)v = \lambda(uv) = u(\lambda v)$ y por lo tanto λuv es un elemento bien definido de A .

Dejamos al lector proporcionar la definición de homomorfismo de álgebras así como percatarse de varios ejemplos de álgebras ya introducidos implícitamente.

Si se imponen condiciones en la multiplicación de un álgebra se obtienen *álgebras conmutativas*, *álgebras asociativas*, *álgebras con uno*. Un álgebra asociativa con uno tal que todo elemento diferente de cero sea invertible se llama *álgebra con división*.

Definimos un *álgebra graduada* como una sucesión

$$A = (A_0, A_1, A_2, \dots)$$

de álgebras A_i , una para cada índice $i \in \mathbf{N}$.

3.6 EJEMPLO. Sea $T^k(V) = \otimes^k V = V \otimes_K \cdots \otimes_K V$ el producto tensorial de un espacio vectorial V sobre un campo K , k veces. Llamaremos a $T^k(V)$ *espacio tensorial de grado k de V* . Si definimos una multiplicación

$$\therefore T^k V \times T^\ell V \longrightarrow T^{k+\ell} V \quad \text{mediante}$$

$$(u_1 \otimes \cdots \otimes u_k) \cdot (v_1 \otimes \cdots \otimes v_\ell) = u_1 \otimes \cdots \otimes u_k \otimes v_1 \otimes \cdots \otimes v_\ell$$

tenemos un álgebra graduada (donde definimos $T^0V = K$ y $T^1V = V$) $TV = (K, V, T^2V, T^3V, T^4V, \dots)$ llamada *álgebra tensorial* de V .

3.7 EJEMPLO. Sea $\bigwedge^k V = V \wedge \dots \wedge V$ el producto exterior de un espacio vectorial V sobre un campo K , k veces. Consideremos la multiplicación exterior definida en la sección 2, a saber

$$\wedge: \bigwedge^k V \times \bigwedge^\ell V \longrightarrow \bigwedge^{k+\ell} V.$$

Entonces tenemos un álgebra graduada

$$\bigwedge V = (K, V, \bigwedge^2 V, \bigwedge^3 V, \dots)$$

llamada *álgebra exterior* o *álgebra de Grassmann* de V .

Sea V^* el espacio dual de V . Consideremos el espacio tensorial de grado k de V , T^kV , del ejemplo 3.6. Consideremos también $T^\ell V^*$ y denotemos con $T_\ell^k(V)$ el producto $(\otimes^k V) \otimes (\otimes^\ell V^*)$. Es decir, $T^kV \otimes T^\ell(V^*) = T_\ell^k(V)$. Con esta notación se tiene que $T_0^k(V) = T^k(V) = \otimes^k V$, $T_\ell^0(V) = \otimes^\ell V^*$ y $T_0^0(V) = K$. Llamaremos a T_ℓ^kV *espacio tensorial de tipo (k, ℓ)* y cada uno de sus elementos lo llamaremos *tensor de tipo (k, ℓ)* . Un tensor de tipo $(k, 0)$ se llamará *tensor contravariante de grado k* y uno de tipo $(0, \ell)$ *tensor covariante de grado ℓ* . Un tensor de tipo $(0, 0)$ es simplemente un *escalar*. Un elemento de $T_0^1V = V$ se llama *vector contravariante* y uno de $T_1^0V = V^*$ se llama *vector covariante*. Si $k \neq 0$ y $\ell \neq 0$, un *tensor mixto* es un tensor de tipo (k, ℓ) .

PROBLEMAS

3.1 Verifique que los conjuntos con operación binaria del ejemplo 3.2 son grupos.

3.2 Verifique que los conjuntos con operaciones binarias del ejemplo 3.5 son anillos y que \mathbb{Z} , K y $K[x]$ son dominios euclidianos.

3.3 Considere la “parte aditiva” de un espacio vectorial $(V, +, \mu)$ sobre un campo K . Entonces $(V, +)$ es un grupo conmutativo bajo la suma. Si U es un subespacio de V , $(U, +)$ es un *subgrupo* conmutativo de $(V, +)$. Proporcione una definición

adecuada de *subgrupo* de un grupo. Considere el espacio cociente V/U definido en II.4. Recuerde que $(V/U, +, \mu)$ posee una estructura de espacio vectorial sobre K . Si nos fijamos solamente en la parte aditiva $(V/U, +)$ recibe el nombre *grupo cociente* y consiste también de las clases laterales de U en V . Considere el subgrupo $n\mathbb{Z}$ de \mathbb{Z} y describa las clases laterales de $\mathbb{Z}/n\mathbb{Z}$. Compruebe que existe una correspondencia biyectiva entre los conjuntos $\mathbb{Z}/n\mathbb{Z}$ y \mathbb{Z}_n . Defina adecuadamente el concepto de *isomorfismo* de grupos y compruebe que $\mathbb{Z}/n\mathbb{Z}$ es isomorfo a \mathbb{Z}_n .

3.4 Establezca un resultado análogo al del problema II.4.8 para grupos.

3.5 Sea $(G, +)$ un grupo conmutativo. Una *base* de G es un conjunto $\{v_i\}_{i=1}^n$, $n \geq 1$ de elementos de G tal que cualquier elemento $v \in G$ puede escribirse de manera única como

$$v = \lambda_1 v_1 + \cdots + \lambda_n v_n$$

con $\lambda_1, \dots, \lambda_n \in \mathbb{Z}$. Demuestre que, si G es un grupo conmutativo con base $\{v_i\}_{i=1}^n$, G' es otro grupo conmutativo y $\{v'_i\}_{i=1}^n$ es cualquier conjunto de elementos de G' entonces existe un homomorfismo de grupos único $h: G \rightarrow G'$ tal que $h(v_i) = v'_i$ para toda $i = 1, \dots, n$. Diremos que $(G, +)$ es un *grupo abeliano libre* si G es conmutativo y posee una base. Observe que el concepto de base para un grupo abeliano se define de una manera similar a la de base para un espacio vectorial tomando en cuenta que los coeficientes λ_i son enteros.

3.6 Sea U un subespacio de un espacio vectorial V sobre un campo K . Considere el espacio cociente V/U . Sea $i: U \rightarrow V$ la inclusión y $p: V \rightarrow V/U$ la proyección al cociente. Pruebe que

(i) $0 \rightarrow U \xrightarrow{i} V \xrightarrow{p} V/U \rightarrow 0$ es una sucesión exacta y

(ii) cualquier sucesión de la forma

$$0 \rightarrow M' \rightarrow M \rightarrow M'' \rightarrow 0$$

llamada *sucesión exacta corta* de espacios vectoriales es esencialmente una sucesión del tipo del inciso (i), es decir, una sucesión con un subespacio y un espacio cociente en los extremos.

3.7 Sea $\{v_i\}_{i=1}^n$ una base de V y $\{f^j\}_{j=1}^n$ la base dual de $\{v_i\}$ en V^* (con índices superiores). Compruebe que los tensores

$$v_{i_1} \otimes \cdots \otimes v_{i_k} \otimes f^{j_1} \otimes \cdots \otimes f^{j_\ell}$$

con $i_\mu, j_\eta = 1, \dots, n$, $\mu = 1, \dots, k$ y $\eta = 1, \dots, \ell$ forman una base de $T_\ell^k(V)$. Entonces cualquier tensor del tipo (k, ℓ) puede escribirse en forma única como

$$t = \sum \xi_{j_1, \dots, j_\ell}^{i_1, \dots, i_k} v_{i_1} \otimes \dots \otimes v_{i_k} \otimes f^{j_1} \otimes \dots \otimes f^{j_\ell}.$$

Los índices i_μ se llaman *índices contravariantes*, los j_η *índices covariantes* y $\xi_{j_1, \dots, j_\ell}^{i_1, \dots, i_k}$ se llaman *componentes* de t con respecto a la base $\{v_i\}$.

3.8 Sean ρ_1 y ρ_2 elementos de $\text{Hom}_K(V, V)$. Defina una operación binaria en $\text{Hom}_K(V, V)$ mediante $\rho_2 \rho_1(v) = \rho_2(\rho_1(v))$ para toda $v \in V$ y compruebe que $\text{Hom}_K(V, V)$ es un álgebra asociativa sobre K también denotada con $\mathcal{A}_K(V)$.

IV.4 K_0 Y K_1

En esta sección estudiaremos algunos conceptos de la K -Teoría Algebraica Clásica la cual es parte del Algebra Lineal. Intuitivamente, la K -Teoría Algebraica Clásica es una generalización del teorema que establece la existencia y unicidad de las bases para espacios vectoriales y también de la Teoría de Grupos del grupo lineal general sobre un campo K .

Recordemos la construcción de \mathbb{Z} a partir de \mathbb{N} : en el conjunto $\mathbb{N} \times \mathbb{N}$ definimos la relación

$$(a, b) \sim (c, d) \iff a + d = b + c.$$

Es muy fácil (problema 4.1(i)) ver que es de equivalencia. Entonces se tiene el conjunto de clases de equivalencia $\mathbb{N} \times \mathbb{N} / \sim$. Si denotamos con $\overline{(a, b)}$ a la clase de equivalencia de (a, b) es fácil demostrar que la fórmula

$$\overline{(a, b)} + \overline{(c, d)} = \overline{(a + c, b + d)}$$

define una operación binaria en $\mathbb{N} \times \mathbb{N} / \sim$ y que con esta operación el conjunto $\mathbb{N} \times \mathbb{N} / \sim$ constituye un grupo conmutativo que se denotará $K_0(\mathbb{N})$ o con \mathbb{Z} . (Se debe demostrar primero que la adición dada por la fórmula anterior “está bien definida”, es decir, es independiente de los representantes de las clases de equivalencia. Después debe comprobarse que satisface los axiomas de grupo conmutativo, (problema 4.1(ii)).

Además se tiene un homomorfismo (de monoides) $f: \mathbb{N} \rightarrow K_0(\mathbb{N})$ dado por $a \mapsto \overline{(a, 0)}$.

Esta misma construcción sirve para cualquier monoide conmutativo (X, \oplus) si alteramos la relación de equivalencia de la siguiente manera: en $X \times X$,

$$(a, b) \sim (c, d) \iff a \oplus d \oplus h = b \oplus c \oplus h$$

para alguna $h \in X$. Es fácil ver que \sim es de equivalencia (problema 4.2). (En el caso de los naturales se puede tomar la misma $h = 0$ siempre. En el caso en que no se cuente con la ley de cancelación se necesitan estos “sumandos de compensación”.)

Se define entonces la operación \oplus en $X \times X / \sim$ y se obtiene un grupo conmutativo que se denota $K_0(X)$ el cual se llama *grupo de Grothendieck* del monoide conmutativo X . También se tiene el homomorfismo de monoides $f: X \rightarrow K_0(X)$ dado por $f(x) = \overline{(x, 0)}$.

4.1 PROPOSICION. $f: X \rightarrow K_0(X)$ *satisface la siguiente propiedad universal: para cada homomorfismo (de monoïdes) $g: X \rightarrow G$ de X en un grupo abeliano G existe un único homomorfismo (de grupos) $h: K_0(X) \rightarrow G$ tal que $g = hf$, es decir, tal que el siguiente diagrama conmuta:*

$$\begin{array}{ccc} X & \xrightarrow{f} & K_0(X) \\ g \searrow & & \downarrow h \\ & & G \end{array}$$

La demostración es fácil (problema 4.3). También, siguiendo los mismos pasos que en el producto tensorial \otimes y que en la potencia exterior \wedge , es fácil demostrar que $K_0(X)$ es único (problema 4.3).

Así, para el conjunto de los números naturales $\mathbf{N} = \{0, 1, 2, \dots\}$ con la operación de sumar (el cual es un monoïde conmutativo en donde vale la ley de cancelación) el grupo de Grothendieck $K_0(\mathbf{N})$ de este monoïde es el grupo aditivo \mathbf{Z} , es decir, $K_0(\mathbf{N}) = \mathbf{Z}$.

Ahora, sea K un campo y consideremos los espacios vectoriales de dimensión finita sobre K , es decir, los K -módulos proyectivos finitamente generados. Denotemos con $\langle V \rangle$ la clase de isomorfismo del espacio vectorial V de dimensión finita, es decir, la clase de todos los espacios vectoriales que son isomorfos a V . Es inmediato verificar que el conjunto $X = \{\langle V \rangle\}$ de clases de isomorfismo es un monoïde conmutativo cuya operación binaria está dada por

$$\langle V \rangle + \langle W \rangle = \langle V \oplus W \rangle.$$

Sea $g: X \rightarrow \mathbf{Z}$ dado por $g(\langle V \rangle) = \dim V$ un homomorfismo de monoïdes, i.e. $g(\langle V \rangle + \langle W \rangle) = g(\langle V \oplus W \rangle) = \dim(V \oplus W) = \dim V + \dim W = g(\langle V \rangle) + g(\langle W \rangle)$.

Sea F el grupo abeliano libre con base el conjunto de clases de isomorfismo de los espacios vectoriales, es decir, con base $\{\langle V \rangle \mid V \text{ es un espacio vectorial}\}$. Sea R el subgrupo de F generado por las expresiones de la forma $\langle V \oplus W \rangle - \langle V \rangle - \langle W \rangle$ donde $0 \rightarrow V \rightarrow V \oplus W \rightarrow W \rightarrow 0$ recorre todas la posibles sucesiones exactas cortas para los espacios vectoriales. Sea $K_0(K) = F/R$ el grupo cociente y denotemos con $[V]$ la proyección o imagen de $\langle V \rangle$ en el cociente. Entonces, siempre que se tenga una sucesión exacta corta de espacios vectoriales

$$0 \rightarrow V \rightarrow V \oplus W \rightarrow W \rightarrow 0$$

tendremos una expresión de la forma $[V \oplus W] = [V] + [W]$ en $K_0(K)$, es decir, $K_0(K)$ está generado por $\{[V] \mid V \text{ es un espacio vectorial}\}$ sujeta a las relaciones de la forma

$$[V] + [W] = [V \oplus W]$$

El homomorfismo $g: X \rightarrow \mathbb{Z}$ da lugar a un homomorfismo $h: K_0(K) \rightarrow \mathbb{Z}$ dado por $h([V]) = \dim V$. Pero h es inyectivo, pues si $h([V]) = h([W])$ entonces $h([V] - [W]) = 0$, lo cual implica que $\dim V - \dim W = 0$ y que $\dim V = \dim W$ y por lo tanto $V \cong W$. También, h es suprayectivo pues $h([K]) = 1$. En resumen, para los espacios vectoriales sobre un campo K , los cuales podemos representar por el símbolo \mathbf{EV}_K se tiene que $K_0(\mathbf{EV}_K) = K_0(K) \cong \mathbb{Z}$. Aquí hemos denotado, por abuso de notación, $K_0(K)$ en lugar de $K_0(\mathbf{EV}_K)$.

¿Qué sucede para otras estructuras cuando consideramos anillos que no necesariamente son campos? Para los \mathbb{Z} -módulos proyectivos finitamente generados \mathbf{Ab} , es decir, los grupos abelianos libres de rango finito (i.e. con base finita) se sabe que $K_0(\mathbb{Z}) \cong \mathbb{Z}$.

Sin embargo, para los \mathbb{Z} -módulos finitos \mathbf{Abf} se sabe que $K_0(\mathbf{Abf}) \cong \mathbb{Q}^+$. Pero para los \mathbb{Z} -módulos finitamente generados \mathbf{Abfg} se tiene que $K_0(\mathbf{Abfg}) \cong \mathbb{Z}$. Véase [LL1] para una demostración de éstos resultados.

Como antes, sea K un campo y denotemos con K^n el producto $K \times \cdots \times K$ n veces. El producto tensorial de dos espacios vectoriales sobre K es un espacio vectorial sobre K . Como $K^n \otimes K^m \cong K^{nm}$ los espacios vectoriales son cerrados bajo el producto tensorial. Entonces (véase el problema 4.4) podemos darle a $K_0(K)$ una estructura de anillo mediante

$$[V] \cdot [W] = [V \otimes_K W].$$

A continuación estudiaremos el conjunto de las transformaciones lineales invertibles a través de uno asociado de matrices. Sea V un espacio vectorial de dimensión n sobre un campo K . Denotemos con $GL(V)$ (o con $Aut_K(V)$) el conjunto de todas las funciones lineales de V en V que sean biyectivas (invertibles). Podemos proporcionarle a este conjunto una estructura de grupo definiendo una operación binaria

$$\circ: GL(V) \times GL(V) \rightarrow GL(V)$$

mediante la composición

$$(f \circ g)(v) = f(g(v)).$$

Claramente (problema 4.5) $GL(V)$ es un grupo bajo \circ .

Ahora definamos otro conjunto. Denotemos con $GL_n(K)$ el conjunto de las matrices de $n \times n$ con elementos en el campo K que poseen inverso, es decir, el conjunto de todas las matrices invertibles o no singulares de $n \times n$ con elementos en K . Podemos definir en $GL_n(K)$ una operación binaria

$$\cdot : GL_n(K) \times GL_n(K) \longrightarrow GL_n(K)$$

mediante

$$(A, B) \longmapsto A \cdot B$$

donde \cdot denota la multiplicación de matrices. Es fácil comprobar (problema 4.6) que $(GL_n(K), \cdot)$ es un grupo cuyo elemento de identidad es la matriz diagonal 1_n . Llamaremos a $GL_n(K)$ *grupo lineal general* de grado n sobre K .

Existe una estrecha relación entre los grupos $GL(V)$ y $GL_n(K)$, a saber, si escogemos una base fija de V , cada función lineal biyectiva de V en V posee una matriz asociada de $n \times n$ con elementos en K la cual es no singular o invertible. Esta correspondencia establece un isomorfismo entre los grupos $GL(V)$ y $GL_n(K)$ debido a que cuando se componen dos funciones lineales, ésta composición está representada por la multiplicación de sus matrices asociadas.

Consideremos un tipo especial de matrices de $GL_n(K)$, que llamaremos *elementales** y que son aquellas que difieren de la matriz identidad 1_n en un sólo elemento $\lambda \in K$ fuera de la diagonal. Dichas matrices las denotaremos con el símbolo e_{ij}^λ , el cual nos indica que tenemos una matriz con unos en la diagonal y λ en el lugar i, j con $i \neq j$. Por ejemplo e_{13}^λ representa dentro de $GL_3(K)$ a la matriz

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

* También se acostumbra llamar matriz elemental (en otros textos) a aquella matriz obtenida de la matriz 1_n mediante una sola operación elemental y dicho concepto no coincide con el aquí definido. También, en algunos textos se le llama *transvecciones* a las matrices que arriba definimos como elementales. El nombre de *elementales* es el que se utiliza en la K -Teoría Algebraica desde la década de los sesenta.

Como $e_{ij}^\lambda \in GL_n(K)$ entonces es fácil comprobar que $(e_{ij}^\lambda)^{-1} = e_{ij}^{-\lambda}$ (problema 4.8). Por ejemplo, en $GL_3(K)$,

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

es decir, $e_{13}^3 \cdot e_{13}^{-3} = 1_3$.

Es fácil ver (problema 4.9) que si e_{ij}^λ es una matriz elemental y $A \in GL_n(K)$, multiplicar A por e_{ij}^λ por la izquierda equivale a sumar λ veces el renglón j al renglón i , y que multiplicar por la derecha equivale a sumar λ veces la columna i a la columna j . Es decir, que la multiplicación $e_{ij}^\lambda A$ o Ae_{ij}^λ equivale a hacer operaciones elementales por renglón o columna en matrices. Por ejemplo, en $GL_3(K)$, $e_{23}^\lambda A$ es

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & \lambda \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} + \lambda a_{31} & a_{22} + \lambda a_{32} & a_{23} + \lambda a_{33} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}.$$

Definimos el símbolo $[A, B]$ como el producto de las matrices $ABA^{-1}B^{-1}$ y lo llamaremos *conmutador* de A y B donde $A, B \in GL_n(K)$.

Dejamos al lector probar la siguiente fórmula para el conmutador de matrices elementales:

$$[e_{ij}^\lambda, e_{k\ell}^\mu] = \begin{cases} 1 & \text{si } j \neq k, i \neq \ell \\ e_{i\ell}^{\lambda\mu} & \text{si } j = k, i \neq \ell \\ e_{kj}^{-\lambda\mu} & \text{si } j \neq k, i = \ell \end{cases}$$

Denotemos con $E_n(K)$ el subgrupo de $GL_n(K)$ generado por todas las matrices elementales e_{ij}^λ , $\lambda \in K$, $1 \leq i \neq j \leq n$, llamado *grupo elemental lineal* de K .

Si cada matriz $A \in GL_n(K)$ la identificamos con la matriz

$$\begin{pmatrix} A & 0 \\ 0 & 1 \end{pmatrix} \in GL_{n+1}(K)$$

obtenemos inclusiones $GL_1(K) \subset GL_2(K) \subset GL_3(K) \subset \dots$. Sea $GL(K) = \cup_{n=1}^{\infty} GL_n(K)$, la unión de los grupos $GL_n(K)$ que llamaremos *grupo lineal general infinito* de K . Podemos concebir a $GL(K)$ como el grupo que consta de todas las matrices invertibles infinitas $A = (a_{ij})$ con $a_{ij} \in K$, $1 \leq i < \infty$, $1 \leq j < \infty$ y $a_{ij} = \delta_{ij}$, la delta de Kronecker para toda i, j excepto un número finito de i, j . Entonces $GL_n(K) \subset GL(K)$ y lo vemos como el subgrupo de todas las $(a_{ij}) \in GL(K)$

con $a_{ij} = \delta_{ij}$ para toda $i, j > n$. La inclusión de $GL_n(K)$ en $GL_{n+1}(K)$ se restringe a la inclusión de $E_n(K)$ en $E_{n+1}(K)$ y, en $GL(K)$, el subgrupo

$$E(K) = \cup_{n=1}^{\infty} E_n(K)$$

se llama *grupo elemental infinito* de K .

4.2 LEMA. (Whitehead) $[GL(K), GL(K)] = E(K)$.

Demostración. Primero veamos que cada matriz elemental puede expresarse como el conmutador de otras dos matrices elementales para $n \geq 3$. Sea $e_{ij}^{\lambda} \in E_n(K)$ una matriz elemental. Consideremos las matrices e_{ik}^1 y e_{kj}^{λ} con $k \neq i, j$. Como $n \geq 3$ siempre podemos encontrar dicha k . Entonces por la fórmula para el conmutador (problema 4.10) $[e_{ik}^1, e_{kj}^{\lambda}] = e_{ij}^{\lambda}$, $0 \leq k \leq n$ y $k \neq i, j$. Pero esto es válido para cualquier $e_{ij}^{\lambda} \in E_n(K)$, por lo tanto

$$E_n(K) \subset [E_n(K), E_n(K)], \quad n \geq 3.$$

También, el producto de matrices elementales es una matriz elemental, i.e. el grupo conmutador siempre es un subgrupo, luego

$$E_n(K) \supset [E_n(K), E_n(K)].$$

Así, $E_n(K) = [E_n(K), E_n(K)]$ y, por lo tanto, $E(K) = [E(K), E(K)]$.

Por otro lado, $E(K) \subset GL(K)$, luego $[E(K), E(K)] \subset [GL(K), GL(K)]$. Así que $E(K) \subset [GL(K), GL(K)]$.

Ahora veamos que $E(K) \supset [GL(K), GL(K)]$: sean $A, B \in GL_n(K)$. Veamos que $ABA^{-1}B^{-1} \in E(K)$, i.e. que $ABA^{-1}B^{-1}$ es una matriz elemental, i.e. que se puede expresar como producto de matrices elementales. Consideremos la siguiente matriz en $GL_{2n}(K)$

$$\begin{pmatrix} ABA^{-1}B^{-1} & 0 \\ 0 & I \end{pmatrix}$$

Pero

$$\begin{pmatrix} ABA^{-1}B^{-1} & 0 \\ 0 & I \end{pmatrix} = \begin{pmatrix} A & 0 \\ 0 & A^{-1} \end{pmatrix} \begin{pmatrix} B & 0 \\ 0 & B^{-1} \end{pmatrix} \begin{pmatrix} (BA)^{-1} & 0 \\ 0 & BA \end{pmatrix}.$$

Esto significa que podemos descomponerla como el producto de tres matrices de la forma

$$\begin{pmatrix} X & 0 \\ 0 & X^{-1} \end{pmatrix} \quad \text{donde } X \in GL_n(K).$$

Veamos que cada matriz de la forma $\begin{pmatrix} X & 0 \\ 0 & X^{-1} \end{pmatrix}$ es producto de matrices elementales (o lo que es lo mismo, invertibles) y así tendremos que $\begin{pmatrix} X & 0 \\ 0 & X^{-1} \end{pmatrix} \in E_{2n}(K)$.

Como

$$\begin{pmatrix} X & 0 \\ 0 & X^{-1} \end{pmatrix} = \begin{pmatrix} I & X \\ 0 & I \end{pmatrix} \begin{pmatrix} I & 0 \\ -X^{-1} & I \end{pmatrix} \begin{pmatrix} I & X \\ 0 & I \end{pmatrix} \begin{pmatrix} 0 & -I \\ I & 0 \end{pmatrix}$$

las tres primeras elementales (i.e. pueden reducirse a I_{2n} mediante operaciones elementales por renglón) y como

$$\begin{pmatrix} 0 & -I \\ I & 0 \end{pmatrix} = \begin{pmatrix} I & -I \\ 0 & I \end{pmatrix} \begin{pmatrix} I & 0 \\ I & I \end{pmatrix} \begin{pmatrix} I & -I \\ 0 & I \end{pmatrix}$$

$\begin{pmatrix} X & 0 \\ 0 & X^{-1} \end{pmatrix}$ es producto de matrices elementales. Luego $ABA^{-1}B^{-1} \in E(K)$ y por lo tanto $[GL(K), GL(K)] \subset E(K)$. ■

4.3 DEFINICION. El grupo cociente $GL(K)/E(K)$ es el K -grupo algebraico de índice uno del campo K denotado con $K_1(K)$.

Así, $K_1(K) = GL(K)/[GL(K), GL(K)]$.

Obsérvese que si $f: K \rightarrow K'$ es un homomorfismo de campos se tiene un homomorfismo de grupos inducido por f

$$f_*: GL(K) \rightarrow GL(K')$$

que envía $E(K)$ en $E(K')$, siendo así que f induce un homomorfismo de grupos $K_1(f): K_1(K) \rightarrow K_1(K')$.

Como K es conmutativo podemos considerar el determinante de una matriz como un homomorfismo $det: GL(K) \rightarrow K^*$ donde K^* denota las *unidades* de K , i.e. los elementos $\lambda \in K$ tal que λ divide al 1, o sea, elementos λ tales que $\lambda\lambda' = 1$ para alguna $\lambda' \in K$. En un campo K , todo elemento diferente de cero es una unidad.

Definamos $SL(K) = \ker(\det)$, o sea, todas las matrices de $GL(K)$ con determinante uno, y lo llamaremos *grupo especial lineal o infinito* de K . Nótese que

$$SL(K) = \cup_{n=1}^{\infty} SL_n(K)$$

donde $SL_n(K) = \ker(\det: GL_n(K) \rightarrow K^*)$ y que las matrices elementales siempre tienen determinante 1, así que, para toda n

$$E_n(K) \subset SL_n(K) \quad \text{y} \quad E(K) \subset SL(K).$$

Nótese que $\det: GL_n(K) \rightarrow K^*$ es invariante bajo la inclusión $GL_n(K) \rightarrow GL_{n+1}(K)$ y, por lo tanto, $\det: GL(K) \rightarrow K^*$ está bien definido. Luego, \det induce un homomorfismo, que por abuso de notación, también denotaremos con

$$\det: K_1(K) = GL(K)/E(K) \rightarrow K^*$$

el cual posee un inverso

$$K^* = GL_1(K) \rightarrow GL(K) \rightarrow \frac{GL(K)}{E(K)} = K_1(K).$$

Si definimos

$$SK_1(K) = SL(K)/E(K) = \ker(\det: K_1(K) \rightarrow K^*)$$

resulta que (problema 4.11) $K_1(K) \cong SK_1(K) \oplus K^*$, i.e., se tiene la sucesión exacta corta que se escinde

$$1 \rightarrow SL(K)/E(K) \rightarrow GL(K)/E(K) \xrightarrow{\det} K^* \rightarrow 1.$$

Como K^* puede considerarse conocido, el cálculo de $K_1(K)$ se limita al de $SK_1(K)$. Obsérvese que $SK_1(K)$ es trivial si, y sólo si, para cualquier matriz $A \in SL_n(K)$ podemos transformar la matriz

$$\begin{pmatrix} A & 0 \\ 0 & I_k \end{pmatrix}$$

para k adecuada, en la identidad I_{n+k} mediante operaciones elementales por renglón o columna. Si $SK_1(K)$ es trivial, entonces $K_1(K) \cong K^*$. Este resulta ser el caso para el campo K . (Véase el problema 4.12).

Podemos generalizar todo lo anterior poniendo, en lugar de un campo K , un anillo conmutativo Λ . Véanse [LL1], [LL2], [V] y [L] para una lectura posterior.

PROBLEMAS

4.1 (i) Pruebe que en $\mathbf{N} \times \mathbf{N}$, $(a, b) \sim (c, d) \iff a + d = b + c$ es una relación de equivalencia.

(ii) Pruebe que la adición en $\mathbf{N} \times \mathbf{N} / \sim$ dada por $\overline{(a, b)} + \overline{(c, d)} = \overline{(a + c, b + d)}$ está bien definida y $\mathbf{N} \times \mathbf{N} / \sim$ es un grupo conmutativo.

4.2 Pruebe que si (X, \oplus) es un monoide conmutativo, la relación $(a, b) \sim (c, d) \iff a \oplus b \oplus h = b \oplus c \oplus h$ para alguna $h \in X$ es una relación de equivalencia.

4.3 Pruebe que $f: X \longrightarrow K_0(X)$ satisface la propiedad universal de 4.1.

4.4 Un *ideal* I de un anillo conmutativo Λ es un subconjunto de Λ que es un subgrupo aditivo y tal que $\Lambda I \subset I$ (es decir, si $\alpha \in I$ y $\lambda \in \Lambda$ entonces $\alpha\lambda \in I$). El grupo cociente Λ/I hereda de Λ una multiplicación bien definida que lo hace un anillo, llamado *anillo cociente*. Defina en $K_0(K) = F/R$ una estructura de anillo mediante $\langle V \rangle \cdot \langle W \rangle = \langle V \otimes_K W \rangle$ en F . Compruebe que R es un ideal de F .

4.5 Sea V un espacio vectorial de dimensión n sobre un campo K . Pruebe que $(GL(V), \circ)$ o $(Aut_K(V), \circ)$ es un grupo. Obsérvese que $Aut_K(V)$ es el grupo de las unidades de $End_K(V)$.

4.6 Pruebe que $(GL_n(K); \cdot)$ es un grupo.

4.7 Sea V un espacio vectorial sobre un campo K . Establezca detalladamente el isomorfismo de grupos

$$GL(V) \cong GL_n(K).$$

4.8 Pruebe que $(e_{ij}^\lambda)^{-1} = e_{ij}^{-\lambda}$.

4.9 Demuestre que si e_{ij}^λ es una matriz elemental y $A \in GL_n(K)$, multiplicar A por e_{ij}^λ por la izquierda equivale a sumar λ veces el rengón j al rengón i , y que multiplicar por la derecha equivale a sumar λ veces la columna i a la columna j .

4.10 Pruebe la fórmula dada por el conmutador $[e_{ij}^\lambda, e_{k\ell}^\mu]$ y observe que no existe una fórmula sencilla para $[e_{ij}^\lambda, e_{ji}^\mu]$, o sea, para el caso $j = k, i = \ell$.

4.11 Demuestre que $K_1(K) \cong SK_1(K) \oplus K^*$.

4.12 Pruebe que si Λ es un dominio euclidiano entonces $K_1(\Lambda) \cong \Lambda^*$. En particular $K_1(K) \cong K^*$, $K_1(K[x]) \cong K^*$ y $K_1(\mathbb{Z}) \cong \{-1, +1\}$. (Sugerencia: Considere el isomorfismo $K_1(\Lambda) \cong SK_1(\Lambda) \oplus \Lambda^*$. Pruebe que $SK_1(\Lambda) = 0$, es decir, que $SL(\Lambda) = E(\Lambda)$. Para probar que si $A = (a_{ij}) \in SL_n(\Lambda)$ entonces $A \in E_n(\Lambda)$, sea $\varepsilon = \min_j\{|a_{1j}| : a_{1j} \neq 0\}$. Supongamos que $\varepsilon = |a_{1k}|$. Después de restar los múltiplos adecuados de la columna k de las otras columnas, pasamos a una nueva matriz $B = (b_{ij})$, con $\min_j\{|b_{1j}| : b_{1j} \neq 0\} < \varepsilon$. Continuando de esta manera obtenemos una matriz $C = (c_{ij})$ con $\min_j\{|c_{1j}| : c_{1j} \neq 0\} = 1$, es decir, alguna c_{1k} es una unidad. De hecho, podemos suponer que c_{11} es una unidad (¿por qué?). Después de transformaciones por columna y rengón, podemos suponer que el primer rengón, y la primera columna de C son cero, excepto c_{11} . Una vez logrado esto, repita el proceso al menor de c_{11} en C . Finalmente se llegará a una matriz $\delta = \text{diag}(d_1, \dots, d_n)$ donde necesariamente $d_1 \cdots d_n = 1$. Luego, $\delta \in E_n(\Lambda)$.)

APENDICE

NOTAS HISTORICAS

*El Algebra Lineal una de las ramas
más antiguas de la Matemática
y a la vez una de las más nuevas.[B]*

AI.1, A.I.2 y A.I.3 Espacios Vectoriales, Funciones Lineales, Subespacios Vectoriales y Espacios Vectoriales de Dimensión Finita

Hermann Gunther Grassmann (1809 -1877) fue el primero en presentar una teoría detallada de espacios vectoriales de dimensión mayor que tres [K]. En sus dos versiones del *Cálculo de Extensión* (1844 y 1862), expresa simbólicamente las ideas geométricas sobre espacios vectoriales y distingue el Algebra Lineal como una teoría formal, donde la geometría solo es una aplicación particular. El trabajo de Grassmann no fue entendido por sus contemporáneos debido a su forma abstracta y filosófica. [K]

La definición de *espacio vectorial* antes llamado *espacio lineal* llegó a ser ampliamente conocida alrededor del año 1920, cuando Hermann Weyl (1885-1955) y otros publicaron la definición formal. De hecho, tal definición (para dimensión finita e infinita) había sido dada treinta años antes por Peano (1858-1932), quien fue uno de los primeros matemáticos que apreciaron en todo su valor la obra de Grassmann, y además con una notación completamente moderna dio la definición de función lineal. Grassmann no dio la definición formal de función lineal, ya que el lenguaje no estaba disponible, sin embargo no hay duda de que conocía el concepto.

Grassmann comienza su trabajo en 1862 con el concepto de un vector como

un segmento de línea recta con longitud y dirección fija. Dos vectores pueden ser sumados uniendo el punto inicial del segundo vector con el punto final del primero. La resta es simplemente la suma del negativo, esto es, el vector con la misma longitud y dirección contraria. [K]

A partir de estas ideas geométricas, Grassmann define el *espacio vectorial* generado por las “unidades” e_1, e_2, e_3, \dots , considerando *combinaciones lineales* $\sum \alpha_i e_i$ donde las α_i son números reales. Establece la suma y la multiplicación por un número real de la siguiente manera:

$$\begin{aligned}\sum \alpha_i e_i + \sum \beta_i e_i &= \sum (\alpha_i + \beta_i) e_i \\ \alpha (\sum \alpha_i e_i) &= \sum (\alpha \alpha_i) e_i\end{aligned}$$

y demuestra formalmente las propiedades de espacio vectorial para esas operaciones (no es claro si el conjunto de unidades puede ser infinito, pero la finitud es implícitamente asumida en algunas de sus demostraciones). Desarrolla la teoría de independencia lineal similarmente a la presentación que uno encuentra en los textos actuales sobre Álgebra Lineal. [F-S]

Grassmann denota a la función lineal que envía los elementos de la base e_1, e_2, \dots, e_n a los de la base b_1, b_2, \dots, b_n respectivamente como

$$Q = \frac{b_1, b_2, \dots, b_n}{e_1, e_2, \dots, e_n}$$

y considera a Q como un cociente generalizado. Mientras que hay problemas obvios con esta notación, ésta tiene cierta elegancia; por ejemplo, si b_1, b_2, \dots, b_n son linealmente independientes, entonces el inverso de Q es

$$\frac{e_1, e_2, \dots, e_n}{b_1, b_2, \dots, b_n}.$$

Da la representación matricial de Q como $Q = \sum \alpha_{r,s} E_{r,s}$ donde

$$E_{r,s} = \frac{0, \dots, 0, e_s, 0, \dots, 0}{e_1, \dots, e_r, \dots, e_n}$$

El *determinante* de Q lo define como el número

$$\frac{b_1 b_2 \cdots b_n}{e_1 e_2 \cdots e_n}. \quad [F - S]$$

Define los conceptos de *subespacio vectorial*, *independencia lineal*, *espacio generado*, *dimensión* de un espacio vectorial, intersección de subespacios y se da

cuenta de la necesidad de demostrar que cualquier base de un espacio vectorial tiene el mismo número de elementos. Entre otras cosas prueba que todo espacio vectorial de dimensión finita tiene un subconjunto linealmente independiente que genera el mismo espacio y que cualquier subconjunto linealmente independiente se extiende a una base. Demuestra la identidad importante

$$\dim(U + V) = \dim U + \dim V - \dim(U \cap V).$$

William Rowan Hamilton (1805-1865), presenta paralelamente a Grassmann una teoría de espacios vectoriales, sólo que lo hace en el caso de dimensión cuatro. Llama *cuaterniones* a los vectores y demuestra que forman un espacio vectorial sobre el campo de los números reales, definiendo su suma y la multiplicación por un escalar.

AI.5 La Matriz Asociada a una Función Lineal

El Algebra Matricial se obtuvo como una consecuencia del tratamiento de la teoría aritmética de formas cuadráticas binarias $aX^2 + 2bXY + cY^2$ contenidas en las *Disquisiciones Aritméticas* de Gauss (1777-1855). Durante su estudio, Gauss introdujo nuevas convenciones notacionales. Luego, para *sustituciones (funciones) lineales* en dos variables

$$\begin{aligned} x &= \alpha x' + \beta y' \\ y &= \gamma x' + \delta y' \end{aligned} \tag{1}$$

decidió, “por brevedad”, referirse a ellas por sus coeficientes:

$$\begin{array}{cc} \alpha & \beta \\ \gamma & \delta \end{array}$$

Cuando estaba tratando con ejemplos numéricos, modificó la notación anterior, agregándole llaves:

$$\left\{ \begin{array}{cc} 1 & 0 \\ 0 & -2 \end{array} \right\}$$

Cada sustitución lineal la denotó con una sola letra mayúscula cuando no era necesario referirse a los coeficientes.

La composición de sustituciones lineales también fue una parte importante en la teoría aritmética de formas. Como en el caso binario, Gauss observó que si la sustitución definida por (1) transforma una forma $F = ax^2 + 2bxy + cy^2$ en F' y si F' es transformada en F'' por una segunda sustitución lineal

$$\begin{aligned} x' &= \varepsilon x'' + \zeta y'' \\ y' &= \eta x'' + \theta y'' \end{aligned}$$

entonces existe una nueva sustitución lineal que transforma F directamente en F'' :

$$\begin{aligned}x &= (\alpha\varepsilon + \beta\eta)x'' + (\alpha\zeta + \beta\theta)y'' \\y &= (\gamma\varepsilon + \delta\eta)x'' + (\gamma\eta + \delta\theta)y''\end{aligned}$$

Gauss escribió la *matriz* de coeficientes de esta nueva sustitución lineal, la cual es el producto de las dos matrices de coeficientes de las dos sustituciones lineales originales. También realizó un cálculo análogo en su estudio de formas cuadráticas ternarias $Ax^2 + 2Bxy + Cy^2 + 2Dxz + 2Eyz + Fz^2$, obteniendo la regla para multiplicar matrices de 3×3 . Sin embargo no designó explícitamente este proceso como un tipo de multiplicación de objetos que no son números. Esto no significa que tales ideas fueran muy abstractas para él, ya que fue el primero en introducirlas, pero en un contexto diferente. [K]

Cauchy (1789-1857) escribió en 1815 una memoria sobre la teoría de determinantes, su trabajo fue inspirado tanto en forma como en contenido por las *Disquisiciones* de Gauss. Siguiendo la dirección de Gauss, introdujo un “sistema simétrico”

$$\begin{array}{cccccc}a_{1,1}, & a_{1,2}, & a_{1,3}, & \cdots, & a_{1,n} \\a_{2,1}, & a_{2,2}, & a_{2,3}, & \cdots, & a_{2,n} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\a_{n,1}, & a_{n,2}, & a_{n,3}, & \cdots, & a_{n,n},\end{array}$$

que representó en forma abreviada (a_{ij}) , al cual se le asocia un determinante. Aún más, en su formulación del teorema del producto para determinantes, Cauchy reconoció explícitamente la idea de Gauss de componer dos sistemas (a_{ij}) y (b_{ij}) para formar un nuevo sistema (m_{ij}) donde $m_{ij} = \sum_{k=1}^n a_{i,k}b_{k,j}$. Entonces demostró que el determinante del nuevo sistema (m_{ij}) es el producto de los determinantes de (a_{ij}) y (b_{ij}) .

James Joseph Sylvester (1814-1897) en 1850, utiliza el término *matriz* para denotar un arreglo rectangular de números.

Ferdinand Gotthold Eisenstein (1823-1852) hizo un estudio cuidadoso de las *Disquisiciones Aritméticas*, que inspiraron gran parte de su trabajo. Introdujo la notación $S \times T$ para “la sustitución compuesta de S y T ” y \bar{S} para “el sistema inverso del sistema S ” en un artículo sobre formas cuadráticas ternarias publicado en 1844.

Eisenstein consideró a las sustituciones lineales como entes que pueden ser sumados y multiplicados como si fueran números ordinarios, solo que la multiplicación no es conmutativa. Consideraba lo anterior como algo generalmente conocido ya que es sugerido por el teorema del producto para determinantes formulado por Cauchy.

Gauss había empleado un simbolismo para la composición de formas binarias que reflejaban su analogía con la aritmética ordinaria, luego Eisenstein lo empleó para sustituciones lineales. Eisenstein nunca desarrolló plenamente su idea de un álgebra de sustituciones.

El simbolismo de Eisenstein fue adoptado por su contemporáneo, Charles Hermite. Tanto Hermite como Eisenstein emplearon el simbolismo como un medio de resumir resultados y no como un modo de razonamiento. Su estudio del álgebra simbólica no lo desarrollaron más allá de lo que necesitaban en su trabajo sobre la Teoría de Números y Funciones Elípticas.

El problema de Cayley-Hermite [H] es muy importante para el entendimiento del desarrollo histórico del álgebra simbólica de matrices. Hermite fue el primero en formular dicho problema y le interesó por la relevancia en su trabajo sobre la teoría aritmética de formas cuadráticas ternarias. Tres matemáticos, quienes llevaron la investigación más allá que Eisenstein y Hermite (Cayley, Laguerre y Frobenius), estuvieron preocupados por este problema. Al menos en el caso de Cayley y Frobenius, fue lo primero que motivó su estudio del álgebra simbólica de matrices.

En 1855 Cayley introdujo la notación

$$\begin{pmatrix} a, & b, & c, & \dots \\ a', & b', & c', & \dots \\ a'', & b'', & c'', & \dots \\ \vdots & & & \end{pmatrix}$$

para representar lo que hoy llamamos una *matriz* y notó que el uso de matrices es muy conveniente para la teoría de ecuaciones lineales. Luego escribió

$$(u, v, w, \dots) = \begin{pmatrix} a, & b, & c, & \dots \\ a', & b', & c', & \dots \\ a'', & b'', & c'', & \dots \\ \vdots & & & \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

para representar el sistema de ecuaciones

$$\begin{aligned} u &= ax + by + cz + \dots, \\ v &= a'x + b'y + c'z + \dots, \\ w &= a''x + b''y + c''z + \dots, \\ &\vdots \end{aligned}$$

También determinó la solución del sistema usando lo que llamó la *inversa de la matriz*:

$$(x, y, z) = \begin{pmatrix} a, & b, & c, & \dots \\ a', & b', & c', & \dots \\ a'', & b'', & c'', & \dots \\ \vdots & & & \end{pmatrix}^{-1} \begin{pmatrix} u \\ v \\ w \end{pmatrix}$$

En 1858, Cayley introdujo la notación de una sola letra para matrices y definió no solamente cómo multiplicarlas sino también cómo sumarlas y restarlas.

De esas innovaciones notacionales surge la idea de que las matrices se comportan como entes que pueden ser sumadas y multiplicadas: la ley de adición de matrices es precisamente similar a la adición de cantidades algebraicas ordinarias, y como se puede ver en su multiplicación, en general no es conmutativa.

Así, Cayley desarrolló el álgebra matricial, haciendo uso constante de su analogía con las manipulaciones algebraicas ordinarias, pero notando cuidadosamente dónde fallan estas analogías. Luego, usando la fórmula para la inversa de una matriz, escribió que “la noción de inversa de una matriz falla cuando el determinante es cero; la matriz en este caso se dice que es indeterminada; y que el producto de dos matrices puede ser cero sin que alguno de los factores sea cero”.

Las innovaciones notacionales sugeridas por el problema de Cayley-Hermite le permitieron a Cayley obtener el teorema de Hamilton-Cayley.

Laguerre en 1867 continúa el trabajo de Eisenstein y Hermite en la aplicación del álgebra simbólica de las funciones lineales a problemas en la Teoría Aritmética de Formas y a la Teoría de Funciones Elípticas. Las memorias de Laguerre constan de dos partes: la primera está dedicada al desarrollo del álgebra simbólica de funciones lineales en la línea de la memoria de Cayley de 1858. La segunda parte contiene el problema de Cayley-Hermite. El trabajo de Laguerre es importante históricamente porque muestra que el desarrollo del álgebra simbólica de matrices fue una consecuencia del trabajo de Eisenstein y Hermite ya que Laguerre no conocía las notas de Cayley de 1855 ni sus memorias de 1858.

Sin conocer las memorias de Cayley y Laguerre, Georg Frobenius también investigó el álgebra simbólica de matrices en su trabajo de formas bilineales de 1878. Frobenius hizo más contribuciones a la teoría de matrices, en particular estudió algunas propiedades de varias clases especiales de matrices, como las simétricas, hermitianas, ortogonales y unitarias.

Aún en sus últimas memorias de 1858 y 1866, Cayley nunca usó su álgebra simbólica de matrices como un modo de razonamiento, mientras que Frobenius sí

lo hizo. En 1878, Frobenius demostró convincentemente la potencia del método simbólico de razonamiento cuando lo combina con los resultados de la teoría espectral de formas bilineales.

Fue Emmy Noether (1882-1935) quien liberó al Álgebra Lineal de las matrices y determinantes.

AII.1 Valores y Vectores Característicos

El concepto de *matriz similar*, como muchos otros conceptos de la teoría de matrices provienen del trabajo de Cauchy. En 1826, en su trabajo sobre formas cuadráticas demostró que, si dos formas cuadráticas están relacionadas mediante un cambio de variable, esto es, si sus matrices son similares, entonces sus polinomios característicos son iguales. Sin embargo Georg Frobenius (1849-1917) fue el primero que analizó y definió formalmente el concepto de similaridad, bajo el nombre de transformación contragrediente, en un artículo de 1878. Comenzó analizando el caso general: estableció que dos matrices A y D son equivalentes, si existen matrices invertibles P y Q tales que $D = PAQ$. Trató los casos especiales de $P = {}^tQ$ (matrices congruentes) y de $P = Q^{-1}$, y demostró varios resultados sobre matrices similares incluyendo el teorema que dice: si A es similar a D , entonces $f(A)$ es similar a $f(D)$, donde f es cualquier función matricial polinomial.

El concepto de *valor característico* se originó independientemente de la teoría de matrices. El contexto en el cual se obtuvieron los primeros problemas de valores característicos durante el siglo XVIII fue en la solución de sistemas de ecuaciones diferenciales lineales con coeficientes constantes. En 1743, Leonhard Euler (1707-1783) introdujo por primera vez el método estándar de resolución de una ecuación diferencial de orden n con coeficientes constantes

$$y^n + a_{n-1}y^{(n-1)} + \cdots + a_1y' + a_0y = 0$$

usando funciones de la forma $y = e^{\lambda t}$ donde λ es una raíz de la ecuación característica

$$z^n + a_{n-1}z^{n-1} + \cdots + a_1z + a_0 = 0.$$

Esta es la misma ecuación que se obtiene al sustituir

$$y_1 = y, \quad y_2 = y', \quad y_3 = y'', \quad \dots, \quad y_n = y^{(n-1)}$$

reemplazando la única ecuación de orden n por un sistema de n ecuaciones de

primer orden

$$\begin{aligned} y_1' &= & y_2 \\ y_2' &= & y_3 \\ &\vdots \\ y_n' &= -a_0 y_1 - a_1 y_2 - \cdots - a_{n-1} y_n \end{aligned}$$

y, por último, calculando el polinomio característico de la matriz de coeficientes de este sistema. [F-B]

Veinte años más tarde, Lagrange (1736 -1813) dio una versión más explícita de esta misma idea al hallar la solución de un sistema de ecuaciones diferenciales encontrando las raíces del polinomio característico de la matriz de coeficientes. El sistema particular de ecuaciones diferenciales surgió del examen de los “movimientos infinitesimales” de un sistema mecánico en la vecindad de su posición de equilibrio. En 1774, Lagrange resolvió un problema similar de Mecánica Celeste usando la misma técnica.

Por otro lado, D'Alembert en trabajos que datan de 1743 a 1758, y motivado por la consideración del movimiento de una cuerda cargada con un número finito de masas (restringidas aquí por simplicidad a tres), había considerado el sistema

$$\frac{d^2 y_i}{dt^2} + \sum_{k=1}^3 a_{ik} y_k = 0 \quad i = 1, 2, 3.$$

Para resolver este sistema decidió multiplicar la i -ésima ecuación por una constante v_i para cada i y sumar las ecuaciones para obtener

$$\sum_{i=1}^3 v_i \frac{d^2 y_i}{dt^2} + \sum_{i,k=1}^3 v_i a_{ik} y_k = 0.$$

Si las v_i son elegidas tal que $\sum_{i=1}^3 v_i a_{ik} + \lambda v_k = 0$ para $k = 1, 2, 3$, esto es, si (v_1, v_2, v_3) es un vector característico correspondiente al valor característico $-\lambda$ para la matriz $A = (a_{ik})$, la sustitución $u = v_1 y_1 + v_2 y_2 + v_3 y_3$ reduce el sistema original a la única ecuación diferencial

$$\frac{d^2 u}{dt^2} + \lambda u = 0.$$

Tal ecuación, después del trabajo de Euler sobre ecuaciones diferenciales, fue resuelta fácilmente y permitió encontrar las soluciones para las tres y_i . Un estudio de

las tres ecuaciones en las cuales esto aparece demuestra que λ está determinada por una ecuación cúbica la cual tiene tres raíces. D'Alembert y Lagrange, un poco más tarde, se dieron cuenta de que para que las soluciones tuvieran sentido físico tenían que ser acotadas cuando $t \rightarrow \infty$. Esto debería ser verdadero si los tres valores de λ fueran distintos, reales y positivos. Cuando Lagrange consideró sistemas similares derivados de sus estudios de Mecánica Celeste, también había determinado los valores para λ , los cuales en muchos casos satisfacían ecuaciones de grado mayor que tres. No podía determinar la naturaleza de esos valores matemáticamente, luego argumentaba que de la estabilidad del sistema solar se obtenía que los valores fueran tales que las correspondientes soluciones a las ecuaciones diferenciales permanecían acotadas.

Cauchy fue el primero que resolvió el problema de determinar en un caso especial los valores característicos de la matriz (a_{ik}) . Se considera que no estuvo influenciado por el trabajo de D'Alembert ni de Lagrange sobre ecuaciones diferenciales, sino que lo hizo como resultado de su estudio sobre superficies cuadráticas (el cual formaba parte del curso de Geometría Analítica que enseñaba en el Ecole Polytechnique desde 1815). Una superficie cuadrática (centrada en el origen) está dada por una ecuación $f(x, y, z) = k$, donde f es una forma cuadrática ternaria $Ax^2 + 2Bxy + Cy^2 + 2Dxz + 2Eyz + Fz^2$. Para clasificar tales superficies, Cauchy, como Euler más tempranamente, necesitaban encontrar una transformación de coordenadas bajo las cuales f se convertía en una suma o diferencia de cuadrados. En términos geométricos, este problema significa encontrar un nuevo conjunto de ejes ortogonales en un espacio de tres dimensiones en el cual expresar la superficie. Pero entonces, Cauchy generalizó el problema a formas cuadráticas con n variables, cuyos coeficientes pueden ser escritos como una matriz simétrica. Por ejemplo, la forma cuadrática binaria $ax^2 + 2bxy + cy^2$ determina la matriz simétrica de 2×2

$$\begin{pmatrix} a & b \\ b & c \end{pmatrix}.$$

El propósito de Cauchy fue encontrar una función lineal tal que la matriz resultante de ésta fuera diagonal, lo cual logró en un artículo de 1829. Debido a que la esencia de la demostración de Cauchy reside en el caso de dos variables, veremos este caso: para encontrar una transformación lineal la cual mande a la forma cuadrática binaria $f(x, y) = ax^2 + 2bxy + cy^2$ en una suma de cuadrados, es necesario encontrar el máximo y mínimo de $f(x, y)$ sujeto a la condición $x^2 + y^2 = 1$. El punto en el cual tal valor extremo de f ocurre es un punto sobre el círculo unitario, el cual también está al final de un eje de uno de la familia de elipses (o hipérbolas) descritas por la forma cuadrática. Si se toma la línea del origen a ese punto como uno de los ejes y la

perpendicular a esa línea como el otro, la ecuación en relación a esos ejes contendrá únicamente los cuadrados de las variables. Por el principio de los multiplicadores de Lagrange, los valores extremos ocurren cuando los cocientes $f_x/2x$ y $f_y/2y$ son iguales. Dándole a cada uno de ellos el valor de λ , obtenemos las ecuaciones

$$\frac{ax + by}{x} = \lambda \quad \text{y} \quad \frac{bx + cy}{y} = \lambda$$

las cuales pueden reescribirse como el sistema

$$\begin{aligned}(a - \lambda)x + by &= 0 \\ bx + (c - \lambda)y &= 0.\end{aligned}$$

Cauchy sabía que su sistema tiene soluciones no triviales si, y sólo si, su determinante es igual a cero, esto es, si $(a - \lambda)(c - \lambda) - b^2 = 0$. En términos de matrices, ésta es la ecuación característica $\det(A - \lambda I) = 0$, la cual Cayley trabajó unos treinta años más tarde.

Para ver cómo las raíces de esta ecuación nos permiten diagonalizar la matriz, sean λ_1 y λ_2 las raíces de la ecuación característica y (x_1, y_1) , (x_2, y_2) las soluciones correspondientes para x, y . Luego

$$\begin{aligned}(a - \lambda_1)x_1 + by_1 &= 0 \\ (a - \lambda_2)x_2 + by_2 &= 0.\end{aligned}$$

Si se multiplica la primera de esas ecuaciones por x_2 , la segunda por x_1 , y se restan, el resultado es la ecuación

$$(\lambda_2 - \lambda_1)x_1x_2 + b(y_1x_2 - x_1y_2) = 0.$$

Similarmente, comenzando con las dos ecuaciones y utilizando $-\lambda_i$, se obtiene la ecuación

$$b(y_2x_1 - y_1x_2) + (\lambda_2 - \lambda_1)y_1y_2 = 0.$$

Sumando estas dos ecuaciones, tenemos que $(\lambda_2 - \lambda_1)(x_1x_2 + y_1y_2) = 0$. Luego, si $\lambda_1 \neq \lambda_2$ (y esto es verdadero en el caso considerado, a menos que la forma original ya sea diagonal), entonces $x_1x_2 + y_1y_2 = 0$. Las condiciones del problema también implican que $x_1^2 + y_1^2 = 1$ y $x_2^2 + y_2^2 = 1$. En términos modernos, la función lineal

$$\begin{aligned}x &= x_1u + x_2v \\ y &= y_1u + y_2v\end{aligned}$$

es ortogonal. Se puede calcular fácilmente que la nueva forma cuadrática que se obtiene de esta transformación es $\lambda_1 u^2 + \lambda_2 v^2$, como se quería. Que λ_1 y λ_2 son reales, se sigue de suponer lo contrario, que son complejos y conjugados uno del otro. En ese caso, x_1 debe ser el conjugado de x_2 y y_1 el de y_2 , además $x_1 x_2 + y_1 y_2$ no puede ser cero. Por consiguiente, Cauchy había demostrado que todos los valores característicos de una matriz simétrica son reales y al menos en el caso en el que son todos distintos, la matriz puede ser diagonalizada utilizando transformaciones ortogonales. [K]

Por otro lado, veamos cómo se demostró que los vectores característicos no triviales correspondientes a valores característicos diferentes son linealmente independientes: supongamos que

$$f(x) = \lambda x$$

donde $x = \sum \xi_i e_i \neq 0$. Escribiendo $C_i = (\lambda I - f)(e_i)$, se tiene que $\sum \xi_i C_i = 0$; luego C_1, C_2, \dots, C_n son linealmente dependientes, de aquí que

$$[(\lambda I - f)(e_1)][(\lambda I - f)(e_2)] \cdots [(\lambda I - f)(e_n)] = 0.$$

Como $e_1 e_2 \cdots e_n \neq 0$, esto es equivalente a que el determinante de la función lineal $\lambda I - f$ sea cero, con lo cual queda demostrado el teorema. Este teorema fue demostrado por Cauchy en 1829 y posteriormente por Weierstrass en 1858.

AII.2 Teorema de Cayley-Hamilton

Cayley menciona en su artículo de 1858, el que ahora es llamado *Teorema de Cayley-Hamilton* para matrices cuadradas de cualquier orden. Demuestra el teorema para el caso de matrices de 3×3 y afirma que la demostración del caso general no es necesaria.

El uso de la notación de una única letra para representar matrices fue lo que probablemente le sugirió a Cayley el teorema que establece que

$$\det \begin{pmatrix} a - M & b \\ c & d - M \end{pmatrix} = 0$$

donde M representa la matriz $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$.

Cayley le comunicó a Sylvester este notable teorema en una carta el 19 de noviembre de 1857. En ella y en el artículo de 1858, Cayley hizo cálculos con M como si fuera una cantidad algebraica ordinaria y demostró que $(a - M)(d - M) - bc =$

$(ad - bc)M^0 - (a + d)M^1 + M^2 = 0$ donde $M^0 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$. Lo que motivó a Cayley a establecer este resultado fue demostrar que “cualquier matriz satisface una ecuación algebraica de su mismo orden” y por lo tanto que “toda función racional y entera (de la matriz) puede ser expresada como una función racional y entera de un orden a lo más igual al de la matriz menos la unidad”. Después demostró que este resultado es cierto para funciones irracionales. En particular demuestra como calcular $L^2 = M$, donde M es una matriz de 2×2 . Supongamos que

$$M = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \quad \text{y} \quad L = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}$$

Como L satisface $L^2 - (\alpha + \delta)L + (\alpha\delta - \beta\gamma) = 0$ y $L^2 = M$, sustituyendo obtenemos

$$L = \frac{1}{\alpha + \delta} [M + (\alpha\delta - \beta\gamma)].$$

El propósito era expresar $X = \alpha + \delta$ y $Y = \alpha\delta - \beta\gamma$ en términos de los coeficientes de la matriz M . Cayley resuelve esto haciendo notar que la expresión anterior para L implica que

$$L = \begin{pmatrix} \frac{a+Y}{X} & \frac{b}{X} \\ \frac{c}{X} & \frac{d+Y}{X} \end{pmatrix}.$$

Como la suma de las diagonales de esta matriz es X y el determinante es Y , se obtienen dos ecuaciones para las variables X y Y que involucran únicamente los coeficientes de la matriz original.

Hamilton enunció el teorema en términos de funciones lineales y no en términos de matrices. En particular, escribió el resultado para dimensión 3 en términos de una función lineal de vectores y para dimensión 4 en términos de una función lineal de cuaterniones. Cabe señalar que no dio una demostración general del teorema. Su principal interés estuvo en derivar una expresión para la inversa de una función lineal. Por ejemplo, si la ecuación $M^2 + aM + b = 0$ entonces $bM^{-1} = -(a + M)$ y M^{-1} puede ser fácilmente encontrada. Luego, se puede explotar esta idea tratando con potencias negativas de una matriz dada. [K]

Frobenius en un artículo de 1878 dio la primera demostración general del Teorema de Cayley-Hamilton.

AII.3 El Polinomio Mínimo

Frobenius define en 1878 el *polinomio mínimo*, como el polinomio de menor grado el cual satisface la matriz. Establece que está formado por los factores del polinomio

característico y que es único. Sin embargo no fue hasta 1904 que Kurt Hensel (1861-1941) demostró la afirmación de unicidad de Frobenius. [KL]

AII.5 Forma Canónica de Jordan

Grassmann demuestra para un operador lineal ρ (construyendo una base apropiada), que el espacio completo se puede descomponer como la suma directa de subespacios invariantes $\ker(\rho - \lambda I)^k$ donde λ recorre las raíces características de ρ y k es la multiplicidad algebraica de λ . Grassmann fue el primero en demostrar este resultado el cual es llamado *teorema de descomposición primaria*.

En uno de sus artículos, Peirce (1809-1880) introduce la idea de *elemento nilpotente*. [K]

El trabajo de Cauchy proviene de una teoría que trata de los valores característicos de varios tipos de matrices. Sin embargo, estos resultados fueron escritos en términos de formas y no en términos de matrices a mediados del siglo XIX. Las formas cuadráticas nos conducen a matrices simétricas. El caso más general de formas bilineales, funciones de $2n$ variables de la forma

$$\sum_{i,j=1}^n a_{ij}x_iy_j$$

conducen a matrices cuadradas. Estas fueron estudiadas primero en detalle por Jacobi alrededor del año 1850 y un poco más tarde por Weierstrass y Jordan, entre otros. Lo principal de este estudio fue encontrar maneras para clasificar tales formas, usualmente determinando las funciones lineales las cuales transforman unas formas en otras de un tipo particularmente simple. La idea central de determinar dichas funciones fue, como en el trabajo de Cauchy, la consideración del polinomio característico $\det(A - \lambda I)$, donde A es la matriz que representa la forma bilineal. En 1851 Sylvester demostró que un cierto conjunto de factores del polinomio característico son preservados bajo cualquier sustitución lineal. Estos factores, los divisores elementales, resultan fundamentales en el proceso de clasificación. Luego, Weierstrass en 1868, tratando con valores complejos y teniendo todos los polinomios irreducibles lineales, demostró el converso del resultado de Sylvester el cual dice que “una forma bilineal dada A puede ser transformada en otra forma B , si y sólo si, ambas formas tienen el mismo conjunto de divisores elementales”. En notación moderna, una forma bilineal puede ser expresada como XAY donde X es una matriz de $1 \times n$, Y una matriz de $n \times 1$, y A una matriz de $n \times n$. Una función lineal sobre X se puede expresar como $X = X'P$ donde P es una matriz invertible de $n \times n$. Similarmente, una función lineal sobre Y puede escribirse como $Y = QY'$.

Se sigue que la forma bilineal transformada se puede expresar como $X'PAQY'$. Weierstrass demostró que se pueden encontrar P y Q tales que la matriz PAQ se puede escribir como

$$\begin{pmatrix} D_1 & 0 & 0 & \cdots & 0 \\ 0 & D_2 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & D_k \end{pmatrix}$$

donde cada submatriz D_i es de la forma

$$\begin{pmatrix} 0 & \cdots & 0 & 0 & 1 & u \\ 0 & \cdots & 0 & 1 & u & 0 \\ \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\ u & \cdots & 0 & 0 & 0 & 0 \end{pmatrix}$$

con u una raíz del polinomio característico.

Dos años más tarde, Jordan desarrolló la misma forma canónica en su *Tra-tado de Sustituciones*. Jordan no llegó al problema de clasificación a través del estudio de formas bilineales, sino a través del estudio de las funciones lineales mismas. Había hecho un estudio detallado del trabajo de Galois sobre soluciones de ecuaciones algebraicas y especialmente sobre la resolución de ecuaciones de grado la potencia de un primo. Estas soluciones involucran el estudio de funciones lineales, funciones cuyos coeficientes podían ser considerados como elementos de un campo finito de orden p . Tal función lineal, con las raíces x_1, x_2, \dots, x_n puede ser expresada en términos de una matriz A . En otras palabras, si X representa la matriz de $n \times 1$ de las raíces x_i , entonces la función puede ser escrita como $X' \equiv AX \pmod{p}$. El propósito de Jordan fue encontrar lo que llamó una “transformación de índices” tal que la función pudiera ser expresada en los términos más simples posibles. En notación matricial, esto significa que quería encontrar una matriz invertible P de $n \times n$, tal que $PA \equiv DP$ donde D es la matriz más “simple” posible.

Luego, si $Y \equiv PX$, entonces $PAP^{-1}Y \equiv PAX \equiv DPX \equiv DY$ y la sustitución sobre Y es “simple”. Usando otra vez el polinomio característico para A , Jordan notó que si todas las raíces de $\det(A - \lambda I) = 0$ son distintas, entonces D puede ser diagonal, donde los elementos de la diagonal son los valores característicos. Por otro lado, si hay raíces múltiples, Jordan demostró que puede encontrarse una sustitución tal que la D resultante tenga forma de bloque, como en la forma anterior de Weierstrass, donde cada bloque D_i es una matriz de la forma

$$\begin{pmatrix} k & 1 & 0 & \cdots & 0 & 0 \\ 0 & k & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & k & 1 \\ 0 & 0 & 0 & \cdots & 0 & k \end{pmatrix}$$

y $k \neq 0 \pmod{p}$ es una raíz del polinomio característico. La forma canónica, conocida hoy como *forma canónica de Jordan*, donde los valores k de la diagonal mayor de la matriz son todos sustituidos por el valor 1, fue introducida por Jordan en 1871 cuando se dio cuenta que su método podía ser aplicado a la solución de sistemas de ecuaciones diferenciales lineales, cuyos coeficientes, en lugar de ser tomados en un campo de p elementos, fueran números reales o complejos. [K]

AIII.1 Formas Bilineales

Mientras los matemáticos mostraban una tendencia a desdeñar las ecuaciones de primer grado, la resolución de ecuaciones diferenciales fue un problema muy importante. Las ecuaciones lineales se distinguieron desde el principio y su estudio contribuyó a poner de manifiesto la linealidad correspondiente. D'Alembert, Lagrange y Euler estudiaron esto, pero el primero es el único que considera útil indicar que la solución general de la ecuación no homogénea es suma de una solución particular y de la solución general de la ecuación homogénea correspondiente; además, cuando estos autores enuncian que la solución general de la ecuación lineal homogénea de orden n es combinación lineal de n soluciones particulares, no mencionan que éstas deben ser linealmente independientes. Este punto, como tantos otros, no se aclararán hasta la enseñanza de Cauchy en la Escuela Politécnica.

Lagrange introdujo (aunque solamente para el Cálculo y sin darle nombre) la ecuación adjunta $L^*(y) = 0$ de una ecuación diferencial lineal $L(y) = 0$, la cual es un ejemplo típico de dualidad en virtud de la relación

$$\int zL(y) dx = \int L^*(z)y dx$$

válida para y y z anulándose en los extremos del intervalo de integración. Con más precisión, y treinta años antes de que Gauss definiera explícitamente la traspuesta de una sustitución lineal de 3 variables, vemos aquí el primer ejemplo de un operador funcional L^* traspuesto de un operador L dado mediante una *función bilineal* (en este caso la integral) $\int yz dx$.

Posteriormente, el estudio de las formas cuadráticas y bilineales, y de sus matrices y sus invariantes conduce al descubrimiento de principios generales sobre la resolución de sistemas de ecuaciones lineales; principios que Jacobi no había alcanzado por carecer de la noción de rango. [B]

AIII.3 Producto Escalar

A principios de 1840, Grassmann había desarrollado el concepto de *producto interior* (*escalar*) de dos vectores como el “producto algebraico de un vector mul-

tiplicado por la proyección perpendicular del segundo sobre él” y estableció algebraicamente el producto interior $\alpha|\beta$ de α y β (cuando $n = 3$) como

$$\alpha|\beta = \alpha_1\beta_1 + \alpha_2\beta_2 + \alpha_3\beta_3$$

donde $\alpha = \alpha_1e_1 + \alpha_2e_2 + \alpha_3e_3$ y $\beta = \beta_1e_1 + \beta_2e_2 + \beta_3e_3$ son hipernúmeros, las α_i y β_i son números reales y e_1, e_2 y e_3 son unidades primarias representadas geoméricamente por vectores de longitud unitaria, dibujados desde un origen común tal como lo determina un sistema de ejes ortogonales de la mano derecha. Las α_ie_i son múltiplos de las unidades primarias y son representadas geoméricamente por las longitudes α_i a lo largo de los ejes respectivos, mientras que α esta representada por un vector en el espacio cuyas proyecciones sobre los ejes son las longitudes α_i . Se cumple lo mismo para β_i y β . Para el producto interior Grassmann postula que

$$e_i|e_i = 1 \quad \text{y que} \quad e_i|e_j = 0 \quad \text{para } i \neq j.$$

El valor numérico o magnitud de un hipernúmero α lo define como

$$a = \sqrt{\alpha|\alpha} = \sqrt{\alpha_1^2 + \alpha_2^2 + \alpha_3^2}.$$

Luego, la magnitud de α es numéricamente igual a la longitud del vector, el cual se representa geoméricamente. También desarrolla la fórmula para el cambio de coordenadas bajo cambio de base e introduce algunas de las ideas básicas relacionadas con las funciones lineales. Finalmente, aplicó algunas de estas ideas en el desarrollo de la geometría simbólica.

Hamilton también define el producto escalar de dos vectores, pero en dimensión a lo más cuatro y demuestra su bilinealidad.

AIV.1 Producto Tensorial

El Cálculo Tensorial se inicia a principios del siglo XIX con Grassmann y Cayley, pero prospera hasta fines del mismo. Riemann multiplicó el poder del cálculo tensorial adoptando las coordenadas curvilíneas de Gauss. Un nuevo progreso fue realizado por Christoffel en 1869 al organizar sistemáticamente el nuevo algoritmo, descubriendo las derivadas que después se han llamado invariante y covariante. Finalmente Ricci y su discípulo Levi Civita le dieron la forma definitiva en 1888.

La física relativista de Einstein sacó del olvido este poderoso instrumento, hoy de uso frecuente en las más diversas teorías.

El ejemplo más importante es la curvatura de un espacio y precisamente este tensor de Riemann fue el que le permitió a Einstein expresar su ley de gravitación y su teoría general de relatividad. [R]

AIV.2 Producto Exterior

Grassmann introduce el producto exterior denotado con $[\]$, para el cual postula que para las unidades primarias e_i

$$[e_i e_j] = -[e_j e_i] \quad i \neq j, \quad [e_i e_i] = 0.$$

Estos corchetes son llamados unidades del segundo orden y no son reducidos por Grassmann (mientras que Hamilton sí lo hace) a unidades del primer orden. Sin embargo eran equivalentes a unidades del primer orden $[e_1 e_2] = e_3$, y así sucesivamente. Con la ayuda del producto exterior de las unidades primarias, *el producto exterior* P de los hipernúmeros $\alpha = \alpha_1 e_1 + \alpha_2 e_2 + \alpha_3 e_3$ y $\beta = \beta_1 e_1 + \beta_2 e_2 + \beta_3 e_3$ lo expresó como sigue:

$$P = [\alpha\beta] = (\alpha_2\beta_3 - \alpha_3\beta_2)[e_2e_3] + (\alpha_3\beta_1 - \alpha_1\beta_3)[e_3e_1] + (\alpha_1\beta_2 - \alpha_2\beta_1)[e_1e_2].$$

Este producto es un hipernúmero del segundo orden y está expresado en términos de unidades independientes del segundo orden. Su magnitud $|P|$ se obtiene a partir de la definición del producto interior, $P|P$, de dos hipernúmeros del segundo orden y es

$$\begin{aligned} |P| &= \sqrt{P|P} = \sqrt{(\alpha_2\beta_3 - \alpha_3\beta_2)^2 + (\alpha_3\beta_1 - \alpha_1\beta_3)^2 + (\alpha_1\beta_2 - \alpha_2\beta_1)^2} \\ &= |\alpha||\beta| \sqrt{1 - \left(\frac{\alpha_1\beta_1}{|\alpha||\beta|} + \frac{\alpha_2\beta_2}{|\alpha||\beta|} + \frac{\alpha_3\beta_3}{|\alpha||\beta|}\right)^2} \\ &= |\alpha||\beta|\text{sen}\theta \end{aligned}$$

donde θ es el ángulo entre los vectores α y β .

Luego, la magnitud $|P|$ del producto exterior $[\alpha\beta]$, se representa geométricamente por el área del paralelogramo construido sobre los vectores, los cuales son las representaciones geométricas de α y β .

Hamilton estableció el producto exterior de dos vectores, como el que ahora es dado para el sistema de los números complejos. Durante 15 años estuvo tratando de generalizarlo a un espacio de tres dimensiones. Un día mientras caminaba por el canal Royal se le ocurrió la fórmula $i^2 = j^2 = k^2 = ijk$ que le permitió definir el producto exterior de dos vectores, pero en un espacio de dimensión cuatro (cuaterniones).

AIV.4 K_0 y K_1

La K -Teoría Algebraica comienza por el hecho de que no todos los módulos proyectivos son libres.

La K -Teoría Algebraica comenzó a finales de la década de los cincuentas con la analogía de Serre entre haces vectoriales y módulos proyectivos: existe una correspondencia biyectiva entre clases de isomorfismo de haces vectoriales reales sobre X y clases de isomorfismo de módulos finitamente generados sobre el anillo de funciones reales continuas sobre X . Por medio de esta correspondencia, el conocimiento de los haces vectoriales se utiliza para sugerir resultados, demostraciones y construcciones para los módulos proyectivos.

En geometría algebraica, los haces vectoriales sobre variedades algebraicas se definen como gavillas localmente libres. De hecho, existe una correspondencia biyectiva entre haces vectoriales y gavillas localmente libres de rango finito. El espacio afín A_k^n sobre un campo k es el espectro primo $\text{Spec } k[t_1, \dots, t_n]$. Este es un esquema afín sobre $k[t_1, \dots, t_n]$ y las gavillas localmente libres corresponden a módulos proyectivos finitamente generados sobre $k[t_1, \dots, t_n]$. Así pues, corresponden a haces vectoriales sobre A_k^n los módulos proyectivos finitamente generados.

La conjetura de Serre era la siguiente: ¿Es todo haz vectorial sobre A_k^n un haz trivial? o equivalentemente, ¿son libres todos los módulos proyectivos finitamente generados sobre $k[t_1, \dots, t_n]$?

Una de las metas de la K -Teoría Algebraica fue en un principio, la de proveer nuevas técnicas para atacar el problema de Serre. Sin embargo, ninguna de las soluciones independientes de Suslin y Quillen en 1976 se basó en la K -Teoría Algebraica.

La K -Teoría Algebraica Clásica es una parte del Algebra Lineal General. Es, intuitivamente, una generalización del teorema que establece la existencia y unicidad de las bases para espacios vectoriales y también de la Teoría de Grupos del grupo lineal general sobre un campo.

La K -Teoría Algebraica tiene sus raíces en el trabajo de Grothendieck sobre el teorema de Riemann-Roch, donde introdujo el funtor K_0 . Sea Λ un anillo con unidad. Definimos $K_0(\Lambda)$ como el grupo abeliano con un generador $[P]$ para cada Λ -módulo proyectivo finitamente generado P , y una relación $[P] = [P'] + [P'']$ para cada sucesión exacta

$$0 \longrightarrow P' \longrightarrow P \longrightarrow P'' \longrightarrow 0.$$

Llamamos a $K_0(\Lambda)$ el *grupo de Grothendieck* de Λ , i.e., el grupo abeliano de clases de isomorfismo de Λ -módulos proyectivos finitamente generados.

El cálculo de K_0 de un campo mide, hasta cierto punto, cuánto les falta a los Λ -módulos proyectivos finitamente generados para poseer una teoría de dimensión como la de los espacios vectoriales. Así, podemos considerar aquella parte de la

K -Teoría Algebraica correspondiente a K_0 como un intento de generalizar ciertas propiedades elementales del Algebra Lineal a módulos sobre un anillo cualquiera.

Observamos que la operación aditiva en $K_0(\Lambda)$ proviene de la suma directa de módulos, i.e., $[P] + [Q] = [P \oplus Q]$ y que, si Λ es conmutativo, entonces el producto tensorial de módulos hace de $K_0(\Lambda)$ un anillo conmutativo, i.e., $[P] \cdot [Q] = [P \otimes Q]$.

Sea $GL_n \Lambda$ el grupo lineal general de matrices invertibles de $n \times n$ con coeficientes en Λ . Si identificamos cada matriz $A \in GL_n \Lambda$ con la matriz $\begin{pmatrix} A & 0 \\ 0 & 1 \end{pmatrix} \in GL_{n+1}(\Lambda)$ obtenemos inclusiones $GL_1(\Lambda) \subset GL_2(\Lambda) \subset GL_3(\Lambda) \dots$. Denotaremos su límite directo con $GL(\Lambda)$. En 1939, Whitehead demostró que el subgrupo $E(\Lambda) \subset GL(\Lambda)$ generado por todas las matrices elementales era igual al subgrupo conmutador $[GL(\Lambda), GL(\Lambda)]$.

LEMA (Whitehead) $[GL(\Lambda), GL(\Lambda)] = E(\Lambda)$.

Por lo tanto, $E(\Lambda)$ es un subgrupo normal de $GL(\Lambda)$ y el cociente $GL(\Lambda)/E(\Lambda)$ es un grupo abeliano bien definido. Este grupo se define como $K_1 \Lambda$ y se llama *grupo de Whitehead de Λ* . Es inmediato comprobar que K_1 es un funtor de la categoría de anillos en la categoría de grupos abelianos.

Durante los últimos años de la década de los sesenta, uno de los problemas mayores de la K -Teoría Algebraica era el de definir funtores $K_n \Lambda$ para toda $n \in \mathbb{Z}$. Esto fue sugerido por analogía con la K -Teoría Topológica.

En 1969, Milnor propuso una definición de $K_2(\Lambda)$ que poseía propiedades análogas a K_0 y K_1 . El observó que, en el grupo $E_n(\Lambda)$, las matrices elementales e_{ij}^λ satisfacían ciertas relaciones obvias. Siguiendo a Steinberg, Milnor introdujo un grupo abstracto $St_n(\Lambda)$ definido por generadores x_{ij}^λ y relaciones que imitaban el comportamiento de esas matrices elementales. Definiendo el homomorfismo canónico

$$\phi_n: St_n(\Lambda) \longrightarrow E_n(\Lambda) \subset GL_n(\Lambda)$$

dado por $\phi_n(x_{ij}^\lambda) = e_{ij}^\lambda$, y pasando al límite directo obtenemos $\phi: St(\Lambda) \longrightarrow GL(\Lambda)$ tal que $\phi(St(\Lambda)) = E(\Lambda)$. Entonces Milnor definió $K_2(\Lambda) = \ker \phi$.

La K -Teoría Algebraica se convirtió en un tema importante porque relaciona dos áreas de las matemáticas. El desarrollo de la K -Teoría Algebraica superior de Quillen relaciona la Topología con el Algebra de una manera nueva y fundamental. La K -Teoría Algebraica utiliza métodos topológicos para encontrar invariantes algebraicos, y también proporciona una manera de traducir conceptos algebraicos en

conceptos topológicos. La K -Teoría Algebraica estudia las propiedades de ciertos grupos $K_i(\Lambda)$, construidos a partir de un anillo Λ .

Quillen en los setentas definió, para $i \geq 1$, el i -ésimo K -grupo algebraico de Λ como $K_i\Lambda = \pi_i(BGL\Lambda^+)$. Como en los casos $i = 1, 2$, K_i es un funtor covariante de la categoría de anillos a la categoría de grupos. [LL1] y [LL2]

Bibliografía de las Notas Históricas

- [B] Bourbaki, N. *Elementos de Historia de las Matemáticas*. Madrid. Alianza. (1976).
- Boyer, C. B. *A History of Mathematics*. New York. Wiley. (1989).
- Cajori, F. *A History of Mathematics*. New York. The Macmillan. (1929).
- Crowe, M. J. *A History of Vector Analysis*. University of Notre Dame Press. (1967).
- Dieudonné, J. *A History of Algebraic and Differential Topology 1900-1960*. Birkhauser. (1989).
- [F-S] Fearnley, D., Sander. Hermann Grassmann and the Creation of Linear Algebra. *Mathematical Monthly* 86. p.809-817. (1879).
- [F-B] Fraleigh, J. B., Bearegard R. A. *Linear Algebra*. Addison-Wesley. (1989).
- Gauss, C. F. *Recherches Arithmétiques*. Jaques Gabay. (1807).
- Grassmann, H. G. *Teoría de la Extensión*. Espasa Calpe. (1947).
- [H] Hawkins, T. Another Look at Cayley and the Theory of Matrices. *Archives Internationales D'histoire des Sciences*. Vol. 27 No.100. Boston (1977).
- Hawkins, T. Cauchy and the Spectral Theory of Matrices. *Historia Matematica*. Vol. 2. p.1-29. (1975).
- [K] Katz, V. J. *Historical Ideas in Teaching Linear Algebra and about some concepts of Linear Algebra*. Comunicación personal no publicada. (1993).
- [KL] Kline, M. *Mathematical Thought from Ancient to Modern Times*. Oxford University Press. (1972).

- [LL1] Lluís-Puebla, E. *Algebra Homológica, Cohomología de Grupos y K-Teoría Algebraica Clásica*. Addison Wesley Ib. (1990).
- [LL2] Lluís-Puebla, E., Loday J.-L., Gillet H., Soulé C., Snaith V. *Higher Algebraic K-Theory: an overview*. Lecture Notes in Mathematics. Vol. **1491**. Springer Verlag. (1992).
- [R] Rey, P., Babini, J. *Historia de la Matemática*. Espasa Calpe. (1951).
- van der Waerden, B. L. *A History of Algebra*. Springer Verlag. (1985).
- van der Waerden, B. L. Hamilton's Discovery of Quaternions. *Mathematics Magazine*. Vol. **49**. No.5. p.227-234. (1976).

— |

| —

— |

| —

BIBLIOGRAFIA

- Birkhoff, G., Mac Lane S. *Algebra*. Macmillan. (1968).
- Birkhoff, G., Mac Lane S. *A survey of Modern Algebra*. Macmillan. (1977).
- Bourbaki, N. *Elements of Mathematics. Algebra I*. Addison Wesley. (1974).
- Fraleigh, J.B., Bearegard R.A. *Algebra Lineal*. Addison Wesley Ib. (1989).
- Guelfand, I.M. *Lecciones de Algebra Lineal*. U. del País Vasco. (Traducción de la versión rusa de 1971).
- Herstein, I.N. *Algebra Moderna*. Trillas. (1970).
- Hoffman, K., Kunze R. *Linear Algebra*. Prentice Hall. (1961).
- Hungerford, T.W. *Algebra*. Springer Verlag. (1980).
- [L] Lam, T.Y., Siu, M.K. K_0 and K_1 -An Introduction to Algebraic K -Theory. *Amer. Math. Monthly*. April, p.329-364. (1975).
- Lang, S. *Algebra*. Addison Wesley. (1971).
- Lang, S. *Linear Algebra*. Addison Wesley. (1971).
- Lipschutz, S. *Linear Algebra*. McGraw-Hill. (1974).
- [LL1] Lluís-Puebla, E. *Algebra Homológica, Cohomología de Grupos y K -Teoría Algebraica Clásica*. Addison Wesley Ib. (1990).
- [LL2] Lluís-Puebla, E., Loday J.-L., Gillet H., Soulé C., Snaith V. *Higher*

Algebraic K-Theory: an overview. Lecture Notes in Mathematics **1491**. Springer Verlag. (1992).

Nomizu, K. *Fundamentals of Linear Algebra.* McGraw-Hill. (1966).

Northcott, D. *Multilinear Algebra.* Cambridge U. Press. (1984).

[V] Vaserstein, L.N. Linear Algebra and Algebraic *K*-Theory. *Contemporary Mathematics* vol. **82**. p.191-198. (1989).

LISTA DE SIMBOLOS

- K , 21
 $M_{m \times n}K$, 23
 K^S , 23 $K[x]$, 23
 $U \cong V$, 25
 1_V , 26
 O_V , 26
 tr , 26
 det , 26, 147
 $Hom_K(U, V)$, 26
 $L(U, V)$, 26
 $A(U, V)$, 26
 M_nK , 27
 $ker f$, 27
 $im f$, 27
 $U + V$, 28
 $U \oplus V$, 29
 $\bigoplus_{i=1}^n V_i$, 31
 $\langle S \rangle$, 31
 $\dim V$, 38
rango f , 43
nul f , 43
 $[f]_{\beta}^{\beta'}$, 49
 $[u]_{\beta}$, 49
 $[f(u)]_{\beta'}$, 49
 N_{β}^{γ} , 50
 $End_K(V)$, 52
 U_{α} , 57
 $p_A(\lambda)$, 62
 $|N_{ij}|$, 63
 A_{ij} , 63
 \tilde{A} , 63
 $m_A(\lambda)$, 67
 V/U , 70
 ρ_U , 71
 J_{ij} , 83

$L^2(V, V; K)$, 88	$T^k(V)$, 137
$L^2(V; K)$, 88	TV , 138
$Bil(V)$, 88	$\bigwedge^k V$, 138
$L^1(V; K)$, 89	$\bigwedge V$, 138
V^* , 89	$T_\ell^k(V)$, 138
$[f]_\gamma$, 91	$\mathcal{A}_K(V)$, 140
V^{**} , 94	$K_0(X)$, 141
U^a , 96	$\langle V \rangle$, 142
$rg(f)$, 96	$[V]$, 142
n_+ , 100	EV_K , 143
n_- , 100	Ab , 143
sgf , 101	Abf , 143
A^* , 102	$Abfg$, 143
\langle , \rangle , 103	$GL(V)$, 143
$\ \ $, 104	$Aut_K(V)$, 143
U^\perp , 105	$GL_n(K)$, 144
ρ^* , 109	e_{ij}^λ , 144
A^* , 112	$[A, B]$, 145
$O(n)$, 112	$GL(K)$, 145
$SO(n)$, 112	$E_n(K)$, 145
$U \otimes_k V$, 121	$E(K)$, 145
$u_i \otimes v_j$, 122	$K_1(K)$, 147
$V \wedge V$, 128	$SL(K)$, 147
$u \wedge v$, 129	$SL_n(K)$, 147
$u \times v$, 130	$SK_1(K)$, 148
$\bigwedge^k V$, 131	
$ A $, 132	

INDICE ANALITICO

A

álgebra
 de Grassmann, 138
 exterior, 138
 graduada, 137
 tensorial, 138
álgebra sobre un anillo, 137
 asociativa, 137
 con uno, 137
 conmutativa, 137
algoritmo de la división, 136
anillo, 136
 cociente, 149
 con división, 136
 con identidad, 136
 con uno, 136
 conmutativo, 136
aniquilador, 96
anulador, 96
aplicación
 lineal, 24
 natural, 95
automorfismo, 134

B

base, 36
 canónica, 36
 de un módulo, 137
 dual, 90
Bessel
 desigualdad de, 108

C

campo, 136
Cauchy-Schwarz
 desigualdad de, 104
Cayley-Hamilton
 teorema de, 63
clase lateral, 70
coeficientes, 31
cofactor, 63
combinación lineal, 31
componente de un tensor, 140
conjunto
 base, 36
 de generadores, 35
 linealmente dependiente, 33, 36
 linealmente independiente, 33, 35
 ortogonal, 105
 ortogonal a un subconjunto, 105
 ortonormal, 105
 solución, 32
conmutador, 145

D

descomposición espectral, 117
desigualdad
 de Bessel, 108
 de Cauchy-Schwarz, 104
 del triángulo, 104
determinante, 132
 de una matriz, 133

- diagrama
 - conmutativo, 31
- dimensión, 38
- dominio
 - entero, 136
 - euclidiano, 136
- E**
- endomorfismo, 52, 132
- escalar, 22, 138
- espacio
 - característico, 57
 - cociente, 70
 - con producto escalar, 103
 - dual, 89
 - euclidiano, 103
 - unitario, 103
- espacio vectorial
 - dimensión de, 38
 - invariante bajo un operador lineal, 71
 - sobre un campo, 21
 - tensorial de grado k , 138
 - tensorial de tipo (k, ℓ) , 138
- espacios isomorfos, 25
- F**
- forma bilineal, 88
 - alternante, 100
 - antisimétrica, 100
 - definida positivamente, 101
 - degenerada, 96
 - no degenerada, 96
 - rango de una, 96
 - semidefinida no negativamente, 101
 - signo de una, 101
 - simétrica, 97
- forma canónica, 56, 70
 - de Jordan, 83
 - normal, 76
- forma cuadrática, 97
 - compleja, 101
 - hermitiana, 101
- forma hermitiana, 101
- forma lineal, 89
- forma multilineal, 125
- función
 - bilineal, 87
 - bilineal alternante, 128
 - multilineal, 125
- función lineal, 24
 - matriz que representa a una, 49
 - nulidad de una, 43
 - rango de una, 43
 - trivial, 26, 127
- funcional lineal, 89
- G**
- Gram-Schmidt
 - procedimiento de, 105
- generadores, 35
- grupo, 135
 - abeliano libre, 139
 - base de un, 139
 - cociente, 139
 - conmutativo, 22, 135
 - de Grothendieck, 141
 - elemental infinito, 145
 - elemental lineal, 145
 - especial infinito, 147
 - especial lineal, 147
 - lineal general, 144
 - lineal general infinito, 145
 - ortogonal, 112
 - ortogonal especial, 112
- H**
- homomorfismo
 - de anillos, 136
 - de espacios vectoriales, 24
 - de grupos, 136
 - de identidad, 26
 - de Λ -módulos, 137
 - natural, 95
 - trivial, 26
- I**
- ideal, 149
- imagen, 27
- índice
 - contravariante, 140
 - covariante, 140
 - de nilpotencia, 79
- invariante bajo similaridad, 60, 65
- isomorfismo
 - de espacios vectoriales, 25
 - de grupos, 139
 - de espacios con producto escalar, 112
- K**
- K -grupo algebraico de índice uno, 147
- L**
- lema de Whitehead, 145
- ley de la inercia de Sylvester, 101
- longitud de un vector, 104

M

matriz

- adjunta clásica, 63
 - antisimétrica, 100
 - asociada a una función lineal, 49
 - asociada a una forma bilineal, 91
 - asociada con respecto a una base, 50
 - asociada relativa a una base, 50
 - característica, 62
 - congruente a, 99
 - cuadrada, 27
 - de transición, 50
 - determinante de una, 133
 - elemental, 144
 - equivalente a, 60, 76
 - hermitiana, 102
 - nilpotente, 79
 - normal, 112
 - ortogonal, 112
 - semejante a, 55
 - simétrica, 27
 - similar a, 55
 - traza de una, 26, 60
 - unitaria, 112
- menor del elemento, 63
- módulo
- finitamente generado, 137
 - izquierdo, 137
 - libre, 137
 - proyectivo, 137
 - sobre un anillo, 137
- monoide, 135
- multiplicación, 137
- multiplicación escalar, 22

N

- norma de un vector, 104
- núcleo, 27
- nulidad, 43
- números de Fibonacci, 61

O

operador

- adjunto, 110
 - autoadjunto, 114
 - hermitiano, 114
 - normal, 112
 - ortogonal, 112
 - que preserva longitudes, 113
 - que preserva productos escalares, 113
 - simétrico, 114
 - unitario, 112
- operador lineal, 44

- descomponible, 77
- determinante de un, 132
- diagonalizable, 56
- invertible, 44
- nilpotente, 79
- no singular, 44
- representado por matriz triangular, 73
- singular, 44

P

- permutaciones, 133
- polinomio
- característico, 62
 - mínimo, 67
- potencia
- exterior de grado dos, 128
 - exterior de grado k , 131
- procedimiento de Gram-Schmidt, 105
- producto
- escalar, 95, 103
 - exterior, 133
 - interno, 103
 - tensorial, 121
- propiedad universal
- de la potencia exterior, 128, 131
 - de la suma directa, 30
 - del producto tensorial, 121
- proporción divina, 61
- proyección ortogonal, 107

R

- rango, 43
- resolución espectral, 117

S

- sección áurea, 61
- semigrupo, 135
- subconjunto independiente máximo, 41
- subespacio
- generado por, 31
 - invariante bajo un operador, 71
 - ortogonal, 105
 - trivial, 27
 - vectorial, 27
- subgrupo, 139
- sucesión
- de Fibonacci, 61
 - exacta, 126
 - exacta corta, 139
- suma
- de subespacios vectoriales, 28
 - directa, 29
 - directa externa, 29

directa interna, 29
Sylvester, ley de la inercia, 101

T

tensor
 componente de un, 140
 contravariante de grado k , 138
 covariante de grado ℓ , 138
 de tipo (k, ℓ) , 138
 mixto, 138
teorema
 de Cayley-Hamilton, 63
 de descomposición primaria, 78
 espectral, 116
transformación lineal, 24
transvección, 144
traza, 26, 60

V

valor
 característico, 56
 propio, 56

vector

 característico, 56
 contravariante, 138
 covariante, 138
 longitud de un, 104
 norma de un, 104
 normalizado, 104
 propio, 56
 unitario, 104

vectores, 22

 linealmente dependientes, 33
 linealmente independientes, 33
 ortogonales, 105

W

Whitehead
 lema de, 145

— |

| —

— |

| —

La intención de la presente obra es la de proveer a los estudiantes de las carreras científicas de un enfoque serio, fundamentado y moderno de los conceptos básicos del Álgebra Lineal. Se ha incluido el Álgebra Multilineal, tema de fundamental importancia así como algunos conceptos de la K-Teoría Algebraica Clásica, una de las ramas más recientes de la Matemática. Se incluyen diversos ejercicios de cálculo explícito, así como una gran cantidad de problemas interesantes que le brindan al estudiante la oportunidad de crear y redactar matemática.

En la introducción se presenta un panorama del libro y se incluye un apéndice que contiene notas históricas sobre los conceptos definidos. El texto está escrito en un lenguaje claro, conciso y elegante. Está diseñado para un curso de un año o dos semestres al nivel licenciatura o bien de un semestre para algunos posgrados.

Este libro cumple ya más de diez años de ser utilizado exitosamente como texto sobre la materia en diversas universidades del Continente Americano, incluyendo algunas universidades de Estados Unidos de Norteamérica y, desde luego, en México.

Emilio Lluis nació en la Ciudad de México en septiembre de 1952. Realizó sus Estudios Profesionales y de Maestría en Matemática en México. En 1980 obtuvo su Doctorado (Ph.D.) en Matemática en Canadá. Es catedrático de la Universidad Nacional Autónoma de México en sus Divisiones de Estudios Profesionales y de Posgrado desde hace treinta años. Ha formado varios profesores e investigadores que laboran tanto en México como en el extranjero.

Es autor de varios libros sobre K-Teoría Algebraica, Álgebra Homológica, Álgebra Lineal y Teoría Matemática de la Música publicados en las editoriales con distribución mundial Addison Wesley, Birkhäuser y Springer Verlag entre otras.

Su trabajo matemático ha quedado establecido en sus artículos de investigación y divulgación que ha publicado sobre la K-Teoría Algebraica, la Cohomología de Grupos y la Teoría Matemática de la Música en las más prestigiadas revistas nacionales e internacionales. Ha sido Profesor Visitante en Canadá.

Recibió varias distinciones académicas, entre otras, la medalla Gabino Barreda al más alto promedio en la Maestría, Investigador Nacional (1984-1990) y Cátedra Patrimonial de Excelencia del Conacyt (1992-1993).

Es miembro de varias asociaciones científicas como la Real Sociedad Matemática Española y la American Mathematical Society. Es presidente de la Academia de Ciencias del Instituto Mexicano de Ciencias y Humanidades, presidente de la Academia de Matemática de la Sociedad Mexicana de Geografía y Estadística y presidente 2000-2002 de la Sociedad Matemática Mexicana.