(Preprint for)

J. Rojo, C. Núñez, A.I. Alonso y J. Álvarez

Curso básico de Álgebra lineal.

Proyecto de investigación financiado por la
Consejería de la Junta de Castilla y León
referencia VA35/99
solicitado el 09/10/98, concedido el 09/04/99
para los años 1999 y 2000

Curso básico de Álgebra lineal

Jesús ROJO
Carmen NÚÑEZ
Ana I. ALONSO

Doctores/as en Matemáticas

Jorge ÁLVAREZ

Licenciado en Matemáticas

Departamento de Matemática Aplicada
E.T.S. de Ingenieros Industriales de Valladolid

Jesús ROJO
Carmen NÚÑEZ
Ana I. ALONSO
Jorge ÁLVAREZ

Curso básico de Álgebra lineal

Proyecto de la Junta de C.-L.

AMS Subject Classification: 15-01 Clasificación Decimal: 512.5

Jesús ROJO
Carmen NÚÑEZ
Ana I. ALONSO
Jorge ÁLVAREZ

Departamento de Matemática Aplicada
E.T.S. de Ingenieros Industriales
Paseo del Cauce, s/n
47011 VALLADOLID , España

Contenido

Prólogo.

El presente libro ha sido realizado merced al apoyo de la Consejería de Educación y Cultura de la Junta de Castilla y León y la Unión Europea (Fondo Social Europeo) al proyecto VA19/00B 'Elaboración de material de apoyo para el Álgebra lineal de primer curso de enseñanzas técnicas'. Al amparo de este proyecto se ha realizado una selección de los temas que más interesan a los alumnos de la asignatura de Álgebra lineal del primer curso de la E.T.S. de Ingenieros Industriales de Valladolid y a los alumnos de asignaturas parecidas en las restantes carreras técnicas.

Hemos tratado de que el libro sea breve y pueda estudiarse en el poco tiempo que los actuales programas conceden a esta materia. El estudiante interesado completará, sin duda, los conocimientos que aquí impartimos con estudios siguiendo alguno de los libros que recomendamos al final de éste.

Los autores agradecerán cuantas sugerencias les sean enviadas con objeto de mejorar el texto. Toda correspondencia con los autores puede dirigirse directamente a la E.T.S. de Ingenieros Industriales de Valladolid.

Valladolid, 15 de noviembre de 2000 — Los Autores

Capítulo 1

Espacios vectoriales.

1.1 Espacios vectoriales, aplicaciones lineales.

1.1.1 DEFINICIÓN. Sea $\mathbb{K}$ un cuerpo conmutativo de operaciones '+' (suma) y '·' (producto) y elementos neutros 0 y 1 .

Un *espacio vectorial sobre $\mathbb{K}$* es un trío $(E, +, \cdot)$ formado por un conjunto E, una operación interna sobre E

$$\begin{aligned} + : E \times E &\to E \\ (x, y) &\to x + y \end{aligned}$$

y una operación externa sobre E

$$\begin{aligned} \cdot : \mathbb{K} \times E &\to E \\ (\lambda, x) &\to \lambda x \end{aligned}$$

que verifican:

(a) $(E, +)$ es un grupo conmutativo, o sea,

(ev1) $(\forall x, y, z \in E) \quad (x + y) + z = x + (y + z)$,
(ev2) $(\exists 0_E \in E)(\forall x \in E) \quad 0_E + x = x + 0_E = x$,
(ev3) $(\forall x \in E)(\exists x' \in E) \quad x + x' = x' + x = 0_E$
(tal x' será designado por $-x$) ,
(ev4) $(\forall x, y \in E) \quad x + y = y + x$;

(b) y además la operación externa cumple

(ev5) $(\forall x \in E)(\forall \lambda, \mu \in \mathbb{K}) \quad \lambda(\mu x) = (\lambda\mu)x$,
(ev6) $(\forall x \in E) \quad 1x = x$,
(ev7) $(\forall x \in E)(\forall \lambda, \mu \in \mathbb{K}) \quad (\lambda + \mu)x = \lambda x + \mu x$,
(ev8) $(\forall x, y \in E)(\forall \lambda \in \mathbb{K}) \quad \lambda(x + y) = \lambda x + \lambda y$.

No es estrictamente necesario poner en esta lista **(ev4)**, pues se la puede deducir de las restantes; no obstante, la mantenemos por respeto al enunciado tradicional.

Los elementos de $\mathbb{K}$ se representan por minúsculas griegas y reciben el nombre de *escalares*.

Para los elementos de E emplearemos generalmente minúsculas ordinarias. Les daremos el nombre de *vectores*.

Las operaciones de E, representadas generalmente por los mismos símbolos que las de $\mathbb{K}$, suelen llamarse 'suma' y 'producto externo'.

$x - y$ representa $x + (-y)$ y, si $\lambda \neq 0$, x/λ representa $\lambda^{-1}x$.

Cuando hablemos de 'el espacio vectorial E' nos referiremos a un trío $(E, +, \cdot)$ cuyas operaciones supondremos conocidas.

En lo que sigue pondremos con frecuencia 'e.v.' para abreviar 'espacio vectorial'.

1.1.2 Ejemplo. Hay una forma natural de dar a $\mathbb{K}$ una estructura de espacio vectorial sobre sí mismo. Si $(\mathbb{K}, +, \cdot)$ es un cuerpo conmutativo, entonces el trío $(\mathbb{K}, +, \cdot)$ es también un espacio vectorial sobre $\mathbb{K}$. Cuando hablamos del espacio vectorial $\mathbb{K}$ (por ejemplo, el espacio vectorial $\mathbb{R}$, o el espacio vectorial $\mathbb{C}$) nos referimos, siempre que no indiquemos lo contrario, a $\mathbb{K}$ dotado de esta estructura.

1.1.3 Ejemplo. Sea X un conjunto y $\mathbb{K}$ un cuerpo conmutativo. Representemos por $\mathcal{F}(X, \mathbb{K})$ el conjunto de las aplicaciones de X en $\mathbb{K}$. Si $f, g \in \mathcal{F}(X, \mathbb{K})$ y $\lambda \in \mathbb{K}$, definimos las aplicaciones $f + g$ y λf poniendo para cada $x \in X$

$$\begin{aligned} (f+g)(x) &= f(x) + g(x)\,, \\ (\lambda f)(x) &= \lambda f(x)\,. \end{aligned}$$

Con las operaciones así definidas, $\mathcal{F}(X, \mathbb{K})$ es un espacio vectorial sobre $\mathbb{K}$.

De manera análoga, si E es un espacio vectorial sobre $\mathbb{K}$ y $\mathcal{F}(X, E)$ el conjunto de las aplicaciones de X en E, las mismas fórmulas de arriba definen dos operaciones con las que $\mathcal{F}(X, E)$ es un espacio vectorial sobre $\mathbb{K}$.

1.1.4 Ejemplo. El conjunto $\mathbb{K}[X]$ de los polinomios en una indeterminada X con coeficientes en $\mathbb{K}$, con la suma ordinaria de polinomios y el producto de un polinomio por un elemento de $\mathbb{K}$, es un espacio vectorial sobre $\mathbb{K}$.

Lo mismo se puede decir del conjunto $\mathbb{K}_n[X]$ de los polinomios de grado igual o menor que n (incluído el polinomio 0) y las operaciones ordinarias.

1.1.5 Ejemplo. Representemos por $S_{\mathbb{K}}$ el conjunto de todas las sucesiones de elementos de $\mathbb{K}$. Si

$$x = (x_n)_{n\in\mathbb{N}} = (x_1, x_2, x_3, \ldots) \quad \text{e} \quad y = (y_n)_{n\in\mathbb{N}} = (y_1, y_2, y_3, \ldots)$$

son elementos de $S_{\mathbb{K}}$ y $\lambda \in \mathbb{K}$, se definen las sucesiones $x+y$ y λx mediante las fórmulas

$$\begin{aligned} x+y &= (x_n+y_n)_{n\in\mathbb{N}} = (x_1+y_1, x_2+y_2, x_3+y_3, \ldots) \\ \lambda x &= \quad (\lambda x_n)_{n\in\mathbb{N}} \quad = (\lambda x_1, \lambda x_2, \lambda x_3, \ldots) \,. \end{aligned}$$

Con las operaciones dadas por estas fórmulas, $S_{\mathbb{K}}$ es un espacio vectorial sobre $\mathbb{K}$.

De hecho, $S_{\mathbb{K}}$ es $\mathcal{F}(\mathbb{N}, \mathbb{K})$ y las fórmulas anteriores coinciden con las del ejemplo (1.1.3), como se puede observar fácilmente si se escribe la sucesión $x = (x_n)_{n\in\mathbb{N}}$ con la notación habitual para las funciones, esto es, representando por $x(n)$ la imagen de n. Así pues, este ejemplo es un caso particular, para $X = \mathbb{N}$, del que vimos en (1.1.3).

1.1.6 Sea E un espacio vectorial. A causa del axioma **(ev2)**, E posee al menos un elemento: el elemento 0_E.

Por otra parte, sea $E = \{x\}$ un conjunto que posee un solo elemento y $\mathbb{K}$ un cuerpo conmutativo. La fórmula

$$x + x = x$$

define la única operación interna sobre E, y la fórmula

$$(\forall \lambda \in \mathbb{K}) \quad \lambda \cdot x = x$$

define la única operación externa posible sobre E. Con estas operaciones, E es un espacio vectorial sobe $\mathbb{K}$ y $x = 0_E$. Representamos este espacio vectorial por $\{0\}$.

Cuando ponemos

$$E \neq \{0\}$$

queremos decir que el espacio vectorial E posee al menos dos elementos, o sea, que existe un vector $x \in E$ tal que $x \neq 0_E$.

Todo cuerpo conmutativo $\mathbb{K}$ posee al menos dos elementos 0 y 1 , luego para el espacio vectorial $\mathbb{K}$ (v. 1.1.2) tenemos que

$$\mathbb{K} \neq \{0\} \,.$$

1.1.7 Consideraciones sobre las sumas de una cantidad finita de elementos.

Los razonamientos que haremos a continuación son válidos tanto para la suma de vectores de un e.v. E (sobre $\mathbb{K}$) como para la suma de escalares de $\mathbb{K}$. En general sirven para la suma de elementos de un conjunto X sobre el que se ha definido una suma '+' asociativa y conmutativa; el lector que encuentre dificultades en razonar con un conjunto X cualquiera, puede imaginarse que X es E o $\mathbb{K}$.

Si $x_1, x_2 \in X$, x_1+x_2 representa la suma de dichos elementos. Si $x_1, x_2, x_3 \in X$, $(x_1+x_2)+x_3$ es la suma de x_1+x_2 y x_3, y $x_1+(x_2+x_3)$ es la suma de x_1 y x_2+x_3; como la suma es asociativa, los elementos $(x_1+x_2)+x_3$ y $x_1+(x_2+x_3)$ coinciden; representamos ambos por $x_1+x_2+x_3$. Teniendo en cuenta que la suma es conmutativa, el elemento $x_1+x_2+x_3$ es también $(x_2+x_1)+x_3$, $x_3+(x_1+x_2)$,..., y así hasta un total de 12 formas distintas. Análogamente, si $x_1, x_2, x_3, x_4 \in X$, $x_1 + x_2 + x_3 + x_4$ se puede definir de varias formas diferentes, pero el resultado es siempre el mismo a causa de la asociatividad y conmutatividad de la suma.

De manera general, poniendo

$$x_1 + x_2 + \cdots + x_n = (x_1 + x_2 + \cdots + x_{n-1}) + x_n$$

resulta, aplicando el principio de inducción, que hemos definido así todas las expresiones del tipo

$$x_1 + x_2 + \cdots + x_n \, .$$

Además, como la suma es asociativa y conmutativa, para calcular $x_1+x_2+\cdots+x_n$ podemos agrupar los elementos por parejas de manera arbitraria.

La suma $x_1 + x_2 + \cdots + x_n$ se representa a veces por

$$\sum_{i\in\{1,\ldots,n\}} x_i \quad \text{o} \quad \sum_{i=1}^{n} x_i \, ,$$

con objeto de acortar las expresiones. En este libro lo haremos únicamente cuando resulte inevitable (lo que será frecuente en el capítulo 3, por ejemplo).

Obsérvese que en la expresión

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

el conjunto $\{1,\ldots,n\}$ juega un papel importante, pero el símbolo i carece de importancia y puede ser substituido por otro. Es decir,

$$\sum_{i=1}^{n} x_i = \sum_{j=1}^{n} x_j \, ,$$

puesto que ambos miembros representan

$$x_1 + x_2 + \cdots + x_n \, .$$

Si I es un conjunto finito, aunque no sea del tipo $\{1,2,\ldots,n\}$, y, para todo $i \in I$, $x_i \in X$, pondremos

$$\sum_{i\in I} x_i$$

para representar la suma de los elementos x_i en cualquier orden. De manera más precisa: podemos ordenar los elementos de I mediante una aplicación biyectiva

$$\begin{aligned} \{1,2,\ldots,n\} &\to I \\ j &\to i_j \, , \end{aligned}$$

poniendo entonces

$$\sum_{i \in I} x_i = \sum_{j=1}^{n} x_{i_j} = x_{i_1} + \cdots + x_{i_n} .$$

En general podremos ordenar I de muchas maneras, aunque el resultado es el mismo para todas ellas.

Si J es un subconjunto de I y $J' = \mathcal{C}_I J$ su complementario, entonces

$$\sum_{i \in I} x_i = \sum_{i \in J} x_i + \sum_{i \in J'} x_i .$$

Supongamos que I y J son dos conjuntos finitos y que

$$\begin{aligned} J &\to I \\ j &\to i_j , \end{aligned}$$

es una aplicación biyectiva; nótese que necesariamente I y J poseen el mismo número de elementos. Entonces

$$\sum_{j \in J} x_{i_j} = \sum_{i \in I} x_i$$

puesto que ambos símbolos representan la suma de los mismos elementos.

Vamos a suponer por fin que en X tenemos además definido un producto distributivo respecto de la suma; es el caso cuando $X = \mathbb{K}$. Si I y J son dos conjuntos finitos, se tiene entonces que

$$\left(\sum_{i \in I} x_i \right) \left(\sum_{j \in J} x_j \right) = \sum_{i \in I} \left(\sum_{j \in J} x_i x_j \right)$$

o también

$$= \sum_{j \in J} \left(\sum_{i \in I} x_i x_j \right)$$

o también

$$= \sum_{(i,j) \in I \times J} x_i x_j ;$$

cualquiera de estas expresiones (que representan el mismo elemento) se denota por

$$= \sum_{i \in I, j \in J} x_i x_j .$$

No terminan aquí todas las consideraciones que se pueden hacer sobre el empleo del símbolo $\sum$, si bien he procurado referirme a todas las operaciones que efectuaremos con él en este libro. Convendrá sin embargo que, llegado el caso de utilizarlo, el lector compruebe que se hace con el suficiente rigor, ayudándose de las consideraciones precedentes.

1.1.8 PROPOSICIÓN

Si E es un espacio vectorial, se verifican las siguientes propiedades:

(a) $(\forall\lambda \in \mathbb{K})(\forall x, y \in E) \quad \lambda(x-y) = \lambda x - \lambda y$,
(b) $(\forall\lambda \in \mathbb{K}) \quad \lambda 0_E = 0_E$,
(c) $(\forall\lambda \in \mathbb{K})(\forall x \in E) \quad \lambda(-x) = -\lambda x$,
(d) $(\forall\lambda, \mu \in \mathbb{K})(\forall x \in E) \quad (\lambda - \mu)x = \lambda x - \mu x$,
(e) $(\forall x \in E) \quad 0x = 0_E$,
(f) $(\forall\lambda \in \mathbb{K})(\forall x \in E) \quad (-\lambda)x = -\lambda x$,
(g) $\lambda x = 0_E \Rightarrow \lambda = 0 \quad \text{o} \quad x = 0_E$.

Se tiene

$$\lambda(x-y) + \lambda y = \lambda[(x-y)+y] = \lambda x \, ,$$

$$(\lambda - \mu)x + \mu x = [(\lambda - \mu) + \mu]x = \lambda x \, ,$$

lo que demuestra (a) y (d).

(b) y (e) se demuestran haciendo $x = y$ en (a) y $\lambda = \mu$ en (d). (c) y (f) se demuestran poniendo $-x = 0_E - x$ y $-\lambda = 0 - \lambda$. Finalmente, para (g), si se tiene $\lambda x = 0_E$ y suponemos que $\lambda \neq 0$, entonces

$$\lambda^{-1}(\lambda x) = \lambda^{-1} 0_E = 0_E$$

y

$$\lambda^{-1}(\lambda x) = (\lambda^{-1}\lambda)x = 1x = x \, ,$$

luego $x = 0_E$.

1.1.9 Siempre que no exista posibilidad de confusión, pondremos 0 para referirnos tanto a 0 como a 0_E .

1.1.10 DEFINICIÓN. Sean E, E' dos espacios vectoriales sobre el mismo cuerpo conmutativo $\mathbb{K}$ y sea $f : E \to E'$ una aplicación. Decimos que f es una *aplicación lineal* (o también un *homomorfismo de espacios vectoriales* cuando:

(ln1) $(\forall x, y \in E) \quad f(x+y) = f(x) + f(y)$,
(ln2) $(\forall\lambda \in \mathbb{K})(\forall x \in E) \quad f(\lambda x) = \lambda f(x)$,

o, lo que es equivalente, cuando

(ln) $(\forall\lambda, \mu \in \mathbb{K})(\forall x, y \in E) \quad f(\lambda x + \mu y) = \lambda f(x) + \mu f(y)$.

(Nótese que en las fórmulas precedentes intervienen dos sumas distintas, la de E y la de E', aun cuando empleemos el mismo símbolo para ambas. Lo mismo ocurre con los dos productos externos.)

Si f es una aplicación lineal, f es en particular un homomorfismo entre los grupos $(E, +)$ y $(E', +)$.

La aplicación compuesta de dos aplicaciones lineales es lineal.

De una aplicación lineal que además sea biyectiva decimos que es un *isomorfismo.*

Si f es un isomorfismo, la aplicación inversa f^{-1} es también un isomorfismo. La aplicación compuesta de dos isomorfismos es un isomorfismo.

Si entre dos espacios vectoriales E y E' existe un isomorfismo $f : E \to E'$, decimos que E y E' son *isomorfos* y ponemos

$$E \simeq E' .$$

De una aplicación lineal de un espacio vectorial E en sí mismo decimos que es un *endomorfismo* de E. De un isomorfismo de E en sí mismo decimos que es un *automorfismo* de E.

1.1.11 Ejemplo. Si E es un espacio vectorial, la aplicación idéntica de E, id_E, dada por

$$(\forall x \in E) \quad id_E(x) = x ,$$

es un automorfismo de E.

1.1.12 PROPOSICIÓN

Si $f : E \to E'$ es una aplicación lineal, [1]

(a) $f(-x) = -f(x)$,
(b) $f(x - y) = f(x) - f(y)$,
(c) $f(0_E) = 0_{E'}$,
(d) $f(\lambda^1 x_1 + \cdots + \lambda^n x_n) = \lambda^1 f(x_1) + \cdots + \lambda^n f(x_n)$.

Para probar las dos primeras basta aplicar la definición de aplicación lineal. La tercera se obtiene haciendo $x = y$ en la segunda. La cuarta se prueba utilizando el principio de inducción) y la propiedad **(ln)**.

1.2 Producto de espacios; subespacios.

1.2.1 DEFINICIÓN. Sean E_1 y E_2 dos espacios vectoriales sobre el mismo cuerpo $\mathbb{K}$. Sobre el producto cartesiano $E_1 \times E_2$ definimos las operaciones

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

y

$$\lambda(x_1, x_2) = (\lambda x_1, \lambda x_2) ,$$

[1] En λ^i de la propiedad (d), i es un índice y no un exponente. Rara vez utilizaremos exponentes, salvo al final del libro; cuando lo hagamos, trataremos siempre de evitar que se produzcan confusiones por esa razón. Por ejemplo, si en una situación que pueda resultar ambigua queremos representar $\lambda \cdot \lambda$, pondremos $(\lambda)^2$ en lugar de λ^2.

donde $x_1, y_1 \in E_1$, $x_2, y_2 \in E_2$, $\lambda \in \mathbb{K}$. Con estas operaciones, $E_1 \times E_2$ es un e.v. sobre $\mathbb{K}$ que denominamos *espacio producto* de E_1 por E_2 y denotamos por $E_1 \times E_2$.

De forma análoga se define el producto $E_1 \times E_2 \times \cdots \times E_n$ de n espacios vectoriales sobre $\mathbb{K}$.

E^n representa el espacio producto $E_1 \times E_2 \times \cdots \times E_n$ cuando $E_1 = E_2 = \ldots = E_n = E$.

1.2.2 Ejemplo. Sea $\mathbb{K}$ un cuerpo conmutativo. $\mathbb{K}^n$ es el espacio producto del e.v. $\mathbb{K}$ (v. 1.1.2) por sí mismo n veces. Recuérdese que, como conjunto, $\mathbb{K}^n$ está formado por las n-uplas $(\lambda^1, \ldots, \lambda^n)$ de elementos de $\mathbb{K}$.

Así es como obtenemos en particular los espacios $\mathbb{R}^n$ y $\mathbb{C}^n$ que utilizaremos con frecuencia.

1.2.3 DEFINICIÓN. Sea E un e.v. sobre $\mathbb{K}$ y $F \subset E$. Decimos que F es un *subespacio vectorial* de E cuando

$$\begin{array}{ll} \textbf{(sev1)} & F \neq \emptyset\,, \\ \textbf{(sev2)} & x, y \in F \Rightarrow x + y \in F\,, \\ \textbf{(sev3)} & \lambda \in \mathbb{K},\ x \in F \Rightarrow \lambda x \in F\,, \end{array}$$

es decir, cuando F es no vacío y cerrado para las operaciones de E. Estas tres propiedades se pueden substituir por

$$\begin{array}{ll} \textbf{(sev1)} & F \neq \emptyset\,, \\ \textbf{(sev4)} & \lambda, \mu \in \mathbb{K},\ x, y \in F \Rightarrow \lambda x + \mu y \in F\,. \end{array}$$

Si F es un subespacio de E, se cumplen las propiedades siguientes

$$\begin{array}{ll} \text{(a)} & 0_E \in F\,, \\ \text{(b)} & x \in F \Rightarrow -x \in F\,. \end{array}$$

Si F es un subespacio de E, entonces F, con las operaciones inducidas por las de E, es también un espacio vectorial sobre $\mathbb{K}$. Su elemento neutro es el de E. Si $x \in F$, su opuesto en F es el mismo que tenía en E.

Si F es un subespacio de E, entonces F es un subgrupo de $(E, +)$.

1.2.4 Si F es un subespacio de E, aplicando la propiedad **(sev4)** y el principio de inducción resulta que

$$\lambda^1, \ldots, \lambda^n \in \mathbb{K},\ x_1, \ldots, x_n \in F \Rightarrow \lambda^1 x_1 + \cdots + \lambda^n x_n \in F\,.$$

1.2.5 Ejemplo. Consideremos el espacio vectorial $\mathcal{F}(X, \mathbb{K})$ de las aplicaciones del conjunto X en $\mathbb{K}$ (v. 1.1.3). Si $x_0 \in X$, el subconjunto formado por las aplicaciones f que verifican $f(x_0) = 0$ es un subespacio vectorial.

1.2.6 Ejemplo. Si $\mathbb{K}[X]$ es el e.v. de los polinomios con coeficientes en $\mathbb{K}$ (v. 1.1.4), el subconjunto $\mathbb{K}_n[X]$ formado por el polinomio 0 y los polinimios de grado $\leq n$, es un subespacio vectorial.

1.2.7 Ejemplo. Sea $S_{\mathbb{K}}$ el e.v. de las sucesiones de elementos de $\mathbb{K}$ (v. 1.1.5). Una sucesión

$$x = (x_n)_{n \in \mathbb{N}} \in S_{\mathbb{K}}$$

es *finita* cuando existe $n_0 \in \mathbb{N}$ tal que

$$(\forall n > n_0) \quad x_n = 0 \, .$$

El subconjunto c_{00} formado por las sucesiones finitas es un subespacio de $S_{\mathbb{K}}$.

En tanto que espacio vectorial, c_{00} es isomorfo al espacio $\mathbb{K}[X]$ de los polinomios, puesto que la aplicación que envía cada sucesión finita $\mu = (\mu_1, \mu_2, \mu_3, \ldots)$ al polinomio $\mu_1 + \mu_2 X + \mu_3 X^2 + \cdots$ es un isomorfismo.

Cuando $\mathbb{K} = \mathbb{R}$ o $\mathbb{C}$, consideraremos también el subconjunto c_0 de $S_{\mathbb{K}}$ formado por las sucesiones que convergen a 0; es un subespacio de $S_{\mathbb{K}}$. Además $c_{00} \subset c_0$.

1.2.8 Sea E un espacio vectorial. $\{0\}$ y E son siempre subespacios de E; se les llama subespacios *triviales*. Los otros subespacios de E se llaman *propios*.

Para la relación de orden dada en $\mathcal{P}(E)$ por la inclusión, $\{0\}$ es el menor subespacio de E, y el propio E es el mayor.

Señalemos que $\emptyset$ no es un subespacio de E, pues no verifica **(sev1)**.

1.2.9 Sea E un e.v. y $x \in E$. El subconjunto

$$\{\lambda x | \lambda \in \mathbb{K}\}$$

es un subespacio de E, que se suele denotar por $\mathbb{K}x$. Para $x = 0$, tenemos $\mathbb{K}x = \{0\}$.

1.2.10 Sea F un subespacio del e.v. E. La aplicación

$$\begin{aligned} i : F &\to E \\ x &\to x \, , \end{aligned}$$

que envía cada vector $x \in F$ a sí mismo, es lineal e inyectiva. Esta aplicación recibe el nombre de *inyección canónica* del subespacio F en el espacio E.

1.2.11 Sea A un conjunto y

$$(a_1, a_2, \ldots, a_n)$$

una n-upla ordenada de A, es decir, un elemento del producto cartesiano A^n. Cuando A es un espacio vectorial o un subconjunto de un espacio vectorial, se suele decir que $(a_1, a_2, \ldots, a_n)$ es un *sistema* de A de n elementos.

Cada uno de los a_i se dice que es un 'elemento del sistema'. El conjunto de estos elementos es un subconjunto de A que recibe el nombre de 'conjunto de los elementos del sistema'.

Un sistema $(a_1, \ldots, a_n)$ de n elementos de A se puede considerar como una aplicación

$$I \to A$$

de un conjunto I de n elementos (generalmente el conjunto $\{1, 2, \ldots, n\}$) en A, justamente la aplicación que a cada $i \in I$ hace corresponder el elemento a_i de A.

Cuando esta aplicación es inyectiva, o sea, cuando

$$i \neq j \Rightarrow a_i \neq a_j \ ,$$

se suele decir que los elementos de $(a_1, a_2, \ldots, a_n)$ son 'distintos dos a dos'.

1.2.12 Ejemplo. $((1,0),(0,0),(1,0),(0,2))$ es un sistema de 4 elementos de $\mathbb{R}^2$. El conjunto de los elementos del sistema es

$$\{(1,0),(0,0),(0,2)\} \ ,$$

o sea, es un conjunto de 3 elementos.

1.2.13 DEFINICIÓN. Sea E un e.v. y $(x_1, \ldots, x_p)$ un sistema de E. Decimos que todo elemento de E de la forma

$$\lambda^1 x_1 + \cdots + \lambda^p x_p = \sum_{i=1}^{p} \lambda^i x_i$$

es una *combinación lineal* de $(x_1, \ldots, x_p)$.

En particular, los vectores $x_1, \ldots, x_p$ son todos de esa forma, pues

$$x_i = 0x_1 + \cdots + 1x_i + \cdots + 0x_p \ ,$$

o sea, son combinación lineal de $(x_1, \ldots, x_p)$.

1.2.14 Ejemplo. En el e.v. $\mathbb{R}[X]$ de los polinomios con coeficientes reales, el polinomio $3 - X + X^2$ es combinación lineal del sistema

$$(1 + X^2, X + 2X^2)$$

puesto que

$$3 - X + X^2 = 3(1 + X^2) - (X + 2X^2) \ .$$

1.2.15 Ejemplo. El vector $(3,-2)$ de $\mathbb{R}^2$ es combinación lineal del sistema $((1,0),(0,0),(1,0),(0,2))$ de (1.2.12), pues

$$(3,-2) = (1,0) + 5(0,0) + 2(1,0) - (0,2) \ .$$

1.2.16 DEFINICIÓN. Sea E un e.v. y $(x_1, \ldots, x_p)$ un sistema de E. El conjunto de los vectores de E que son combinación lineal de $(x_1, \ldots, x_p)$ es un subespacio vectorial de E; decimos que es el *subespacio generado* por el sistema $(x_1, \ldots, x_p)$ y lo representamos por

$$\langle x_1, \ldots, x_p \rangle \,.$$

Ya hemos visto en (1.2.13) que

$$x_1, \ldots, x_p \in \langle x_1, \ldots, x_p \rangle \,.$$

Si F es un subespacio de E y

$$F = \langle x_1, \ldots, x_p \rangle \,,$$

decimos que $(x_1, \ldots, x_p)$ es un *sistema generador* de F, o también que '$(x_1, \ldots, x_p)$ genera F'. Nótese que esto implica que $x_1, \ldots, x_p \in F$, o sea, que $(x_1, \ldots, x_p)$ es también un sistema de F.

Si en un e.v. E existe un sistema (finito) generador de E, $(x_1, \ldots, x_p)$, es decir, tal que

$$E = \langle x_1, \ldots, x_p \rangle \,,$$

decimos que E es un espacio vectorial *de dimensión finita*.

En particular, si $(x_1, \ldots, x_p)$ es un sistema del e.v. E, el subespacio $\langle x_1, \ldots, x_p \rangle$ es un e.v. de dimensión finita.

1.2.17 Ejemplo. El sistema de 4 elementos de $\mathbb{R}^2$ del ejemplo (1.2.12) es un sistema generador de $\mathbb{R}^2$. El e.v. $\mathbb{R}^2$ es, pues, de dimensión finita.

1.2.18 Si $x \in E$, el subespacio $\langle x \rangle$ generado por el sistema (x) es $\mathbb{K}x$ (v. 1.2.9).

Si $\lambda \neq 0$, entonces $\langle x \rangle = \langle \lambda x \rangle$, o sea, $\mathbb{K}x = \mathbb{K}\lambda x$. Más generalmente, si $\lambda^1 \neq 0, \ldots, \lambda^p \neq 0$, entonces

$$\langle x_1, \ldots, x_p \rangle = \langle \lambda^1 x_1, \ldots, \lambda^p x_p \rangle \,.$$

1.2.19 El espacio vectorial $\{0\}$ (v. 1.1.6) es de dimensión finita, pues el sistema (0) genera el espacio.

1.2.20 PROPOSICIÓN

Si uno de los elementos x_i de un sistema $(x_1, \ldots, x_p)$ de E es combinación lineal del sistema $(x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_p)$ de $p-1$ vectores formado por los restantes, entonces

$$\langle x_1, \ldots, x_p \rangle = \langle x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_p \rangle \,.$$

La demostración de esta propiedad no presentará problemas al lector.

1.2.21 PROPOSICIÓN

La intersección $\bigcap_{i\in I} F_i$ de una familia cualquiera $(F_i)_{i\in I}$ de subespacios de un e.v. E es un subespacio de E.

No haremos la demostración, que es formalmente muy sencilla.

1.2.22 DEFINICIÓN. Sea E un e.v. y $A \subset E$. La intersección de todos los subespacios de E que contienen a A es un subespacio; es precisamente el menor subespacio que contiene a A (el menor para el orden dado en $\mathcal{P}(E)$ por la inclusión. Se llama *subespacio generado* por el conjunto A y se lo representa por $\langle A\rangle$.

Si F es un subespacio de E, es evidente que

$$\langle F\rangle = F\,.$$

Si F es un subespacio de E, A un subconjunto de E y

$$F = \langle A\rangle\,,$$

decimos que A es un *conjunto generador* de F, o también que 'A genera F'. Nótese que esto implica que $A \subset F$.

1.2.23 Ejemplo. Si E es un espacio vectorial,

$$\langle\emptyset\rangle = \langle\{0\}\rangle = \{0\} \quad \text{y} \quad \langle E\rangle = E\,.$$

1.2.24 Ejemplo. Si E es un e.v. y $x \in E$, entonces

$$\mathbb{K}x = \langle\{x\}\rangle\,.$$

1.2.25 Ejemplo. El subespacio del espacio vectorial $\mathbb{R}^2$ que está generado por el conjunto $\{(1,0),(0,0),(0,2)\}$ es $\mathbb{R}^2$.

1.2.26 Se demuestra sin dificultad que, si S y S' son subconjuntos de un e.v. E tales que $S \subset S'$, entonces $\langle S\rangle \subset \langle S'\rangle$.

1.2.27 Veamos las relaciones existentes entre los dos conceptos de subespacio generado, el de (1.2.16) y el de (1.2.22). Sea E un espacio vectorial:

a) Si $(x_1,\ldots,x_p)$ es un sistema de E y S es el conjunto de los elementos del sistema (v. 1.2.11), entonces

$$S = \langle x_1,\ldots,x_p\rangle\,.$$

b) Si S es un subconjunto de E (posiblemente infinito) y $S \neq \emptyset$, el subconjunto de E formado por todas las combinaciones lineales de los distintos sistemas (finitos) de S es un subespacio de E. Este subespacio coincide con $\langle S\rangle$.

1.3 Espacio cociente; suma de subespacios.

1.3.1 Sea E un e.v. sobre $\mathbb{K}$ y F un subespacio de E. Si $x, y \in E$, ponemos

$$x \mathrm{R}_F y \quad \text{si} \quad y - x \in F\,.$$

R_F es una relación de equivalencia sobre E.

La clase de equivalencia $\dot{x}$ de un elemento $x \in E$ por la relación R_F es

$$\dot{x} = \{x + z | z \in F\}$$

y este conjunto se representa por $x + F$ (es el conjunto formado por los vectores que son suma de x y de un vector de F). En particular, la clase $\dot{0}$, la del cero de E, es F.

Representamos el conjunto cociente de E por la relación R_F mediante el símbolo

$$E/F\,.$$

Se cumplen las propiedades

$$\begin{array}{lccc} \text{(a)} & x' \in \dot{x} \text{ y } y' \in \dot{y} & \Rightarrow & x' + y' \in (x \dot{+} y) \\ \text{(b)} & \lambda \in \mathbb{K} \text{ y } x' \in \dot{x} & \Rightarrow & \lambda x' \in (\dot{\lambda x})\,. \end{array}$$

1.3.2 Ejemplo. Si $E = \mathbb{R}^2$ y $F = \{(x, y) \in \mathbb{R}^2 | x = 0\}$, entonces los vectores (x, y) y (x', y') estan relacionados por $\mathrm{R_F}$ si y sólo si $x = x'$. La clase de equivalencia de un vector (x, y) es

$$\begin{aligned} \{(x, y)\} &= \{(x', y') \in \mathbb{R}^2 | x' = x\} \\ &= \{(x, y') \in \mathbb{R} | y' \in \mathbb{R}\}\,, \end{aligned}$$

o sea, la 'recta vertical' que pasa por (x, y). (En la figura 1.1 se ha representado esta clase mediante una línea de trazo grueso.)

1.3.3 DEFINICIÓN. Sea E un e.v. sobre $\mathbb{K}$, F un subespacio de E y E/F el conjunto cociente. Las fórmulas

$$\dot{x} + \dot{y} = (x \dot{+} y) \quad \text{y} \quad \lambda\dot{x} = (\dot{\lambda x})$$

definen una 'suma' de clases y un 'producto' de una clase por un escalar. Obsérvese que el cálculo de $\dot{x} + \dot{y}$ y $\lambda\dot{x}$ se efectúa utilizando elementos particulares, x e y, de las clases; sin embargo, las propiedades (a) y (b) de (1.3.1) nos garantizan que el resultado obtenido no depende del representante que hayamos elegido en cada clase.

Con estas operaciones, E/F es un espacio vectorial sobre $\mathbb{K}$, al que llamamos *espacio vectorial cociente* de E por F.

La aplicación

$$\begin{aligned} j : E &\to E/F \\ x &\to \dot{x}\,, \end{aligned}$$

que envía cada elemento de E a su clase de equivalencia, es lineal (v. 1.1.10) y sobre. Recibe el nombre de *sobreyección canónica* de E en el cociente E/F.

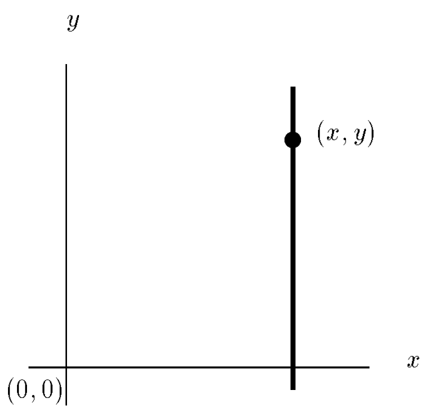

Figura 1.1: Clase de equivalencia de (x, y) por el subespacio $x = 0$.

1.3.4 Al contrario de lo que ocurre con la intersección de subespacios (v. 1.2.21), la reunión de una familia de subespacios de un e.v. E no es, en general, un subespacio de E.

Los párrafos que siguen describen cuál es el menor subespacio que contiene a la reunión de un número finito de subespacios. Por su especial importancia, hacemos hincapié en el caso en que los subespacios son dos.

1.3.5 DEFINICIÓN. Sea E un e.v. y F_1, F_2 dos subespacios de E. El subconjunto de E

$$\{x_1 + x_2 | x_1 \in F_1, x_2 \in F_2\}$$

es un subespacio de E; le llamamos *subespacio suma* de F_1 y F_2 y lo representamos por

$$F_1 + F_2 .$$

Se cumple que

$$F_1 \subset F_1 + F_2 \quad \text{y} \quad F_2 \subset F_1 + F_2 ,$$

o sea,

$$F_1 \cup F_2 \subset F_1 + F_2 .$$

Además,

$$\langle F_1 \cup F_2 \rangle = F_1 + F_2 ,$$

o sea, $F_1 + F_2$ es el menor subespacio que contiene al conjunto $F_1 \cup F_2$.

1.3.6 Cuando $F_1 \cup F_2$ es un subespacio, entonces $F_1 + F_2 = F_1 \cup F_2$. En particular,

$$F_1 + E = E \quad \text{y} \quad F_1 + \{0\} = F_1 .$$

1.3.7 Ejemplo. Si en $\mathbb{R}^2$ consideramos los subespacios

$$F_1 = \{(x, y) \in \mathbb{R}^2 | x = y\}\,, \quad F_2 = \{(x, y) \in \mathbb{R}^2 | y = 0\}$$

(que en la figura 1.2 representamos por sendas líneas gruesas), tenemos que

$$\mathbb{R}^2 = F_1 + F_2\,.$$

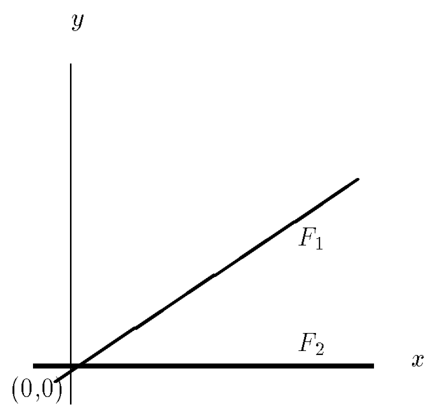

Figura 1.2: $\mathbb{R}^2$ como suma de dos subespacios.

1.3.8 Ejemplo. Si en $\mathbb{R}^3$ consideramos los subespacios

$$F_1 = \{(x, y, z) \in \mathbb{R}^3 | x = 0\}\,, \quad F_2 = \{(x, y, z) \in \mathbb{R}^3 | z = 0\}$$

tenemos que

$$\mathbb{R}^3 = F_1 + F_2\,.$$

1.3.9 DEFINICIÓN. Sea E un e.v., F_1 y F_2 dos subespacios y $F = F_1 + F_2$. Si $x \in F$, x se puede expresar como

$$x = x_1 + x_2\,,$$

con $x_1 \in F_1$ y $x_2 \in F_2$ (v. 1.3.5). Por lo general, esta forma de descomponer x no es única (v. 1.3.11).

Cuando para cada $x \in F = F_1 + F_2$ tenemos

$$x = x_1 + x_2$$

con $x_1 \in F_1$ y $x_2 \in F_2$, en forma única, decimos que F es *suma directa* de F_1 y F_2 y escribimos

$$F = F_1 \oplus F_2\,.$$

1.3.10 Ejemplo. En el caso del ejemplo (1.3.7), cada vector (x, y) de $\mathbb{R}^2$ es la suma

$$(x, y) = (y, y) + (x - y, 0)$$

con $(y, y) \in F_1$ y $(x - y, 0) \in F_2$, y ésta es la única forma de hacer tal descomposición. En consecuencia

$$\mathbb{R}^2 = F_1 \oplus F_2 \,.$$

1.3.11 Ejemplo. En el ejemplo (1.3.8), $\mathbb{R}^3$ no es suma directa de F_1 y F_2 pues, por ejemplo,

$$(0, 0, 0) = (0, 0, 0) + (0, 0, 0) \,,$$

pero también

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

u otras formas, todas ellas suma de un vector de F_1 y otro de F_2.

1.3.12 PROPOSICIÓN

Sea E un e.v. y F, F_1 y F_2 subespacios de E. Son equivalentes las proposiciones siguientes:

(i) $F = F_1 \oplus F_2 \,,$

(ii) $F = F_1 + F_2 \quad \text{y} \quad F_1 \cap F_2 = \{0\} \,.$

(i)$\Rightarrow$(ii). Si (i) es cierta, $F = F_1 + F_2$. Además, si $x \in F_1 \cap F_2$,

$$x = 0 + x \quad \text{y} \quad x = x + 0$$

son dos formas (distintas si $x \neq 0$) de descomponer x en suma de un vector de F_1 y otro de F_2. Como por hipótesis la descomposición es única, necesariamente $x = 0$, lo que demuestra que $F_1 \cap F_2 = \{0\}$.

(ii)$\Rightarrow$(i). Supongamos que (ii) es cierta. En particular, $F = F_1 + F_2$. Además, si $x \in F$ y

$$x = x_1 + x_2 \quad \text{y} \quad x = x_1' + x_2'$$

con $x_1, x_1' \in F_1$ y $x_2, x_2' \in F_2$, resulta que

$$x_1 - x_1' = x_2 - x_2' \,,$$

luego $x_1 = x_1'$ y $x_2 = x_2'$, lo que demuestra que la forma de descompones x es única. Esto significa entonces que $F = F_1 \oplus F_2$.

1.3.13 Si E es un e.v. y F, F_1 y F_2 subespacios tales que

$$F = F_1 \oplus F_2 \,,$$

entonces

$$F \simeq F_1 \times F_2$$

puesto que la aplicación

$$\begin{aligned} F_1 \times F_2 &\to F \\ (x_1, x_2) &\to x_1 + x_2 \end{aligned}$$

es un isomorfismo.

1.3.14 DEFINICIÓN. Sea E un e.v. y F_1 y F_2 subespacios tales que

$$E = F_1 \oplus F_2 \,.$$

Decimos entonces que F_1 y F_2 son subespacios *suplementarios* (uno del otro) de E.

Si F_1 es un subespacio de E, F_1 admite siempre un suplementario, aunque nosotros sólo demostraremos este hecho para ciertos espacios vectoriales (v. 1.5.7). Por otra parte, un subespacio puede tener varios suplementarios (v. 1.3.16).

1.3.15 Ejemplo. Si E es un e.v., $\{0\}$ es el único suplementario de E y E es el único suplementario de $\{0\}$.

1.3.16 Ejemplo. En el ejemplo (1.3.7), F_1 y F_2 son suplementarios, pero también el subespacio

$$F_3 = \{(x, y) \in \mathbb{R}^2 | x = 0\}$$

es suplementario de F_1.

1.3.17 Ejemplo. No se deben confundir los conceptos de subespacio suplementario y de subconjunto complementario. Así, en el ejemplo precedente, el complementario de F_1 es

$$\{(x, y) \in \mathbb{R}^2 | x \neq y\} \,,$$

que ni siquiera es un subespacio.

1.3.18 Si E es un e.v. y F_1 y F_2 subespacios tales que

$$E = F_1 \oplus F_2 \,.$$

entonces

$$F_2 \simeq E/F_1$$

puesto que la aplicación

$$\begin{aligned} F_2 &\to E/F_1 \\ x &\to \dot{x} \,, \end{aligned}$$

que envía cada vector x de F_2 (luego de E) a su clase de equivalencia $\dot{x}$ por la relación R_{F_1} (v. 1.3.1), es un isomorfismo. Nótese que esta aplicación es la restricción a F_2 de la sobreyección canónica de E en E/F_1 (v. 1.3.3).

1.3.19 PROPOSICIÓN

En un espacio vectorial E, todos los subespacios suplementarios a uno dado son isomorfos entre sí.

Si F_1 es un subespacio de E y

$$E = F_1 \oplus F_2 \quad \text{y} \quad E = F_1 \oplus F_2' ,$$

aplicando el resultado precedente

$$F_2 \simeq E/F_1 \quad \text{y} \quad F_2' \simeq E/F_1$$

luego

$$F_2 \simeq F_2' .$$

(Véase también la argumentación de 1.5.8 para un caso menos general.)

1.3.20 DEFINICIÓN. Sea E un e.v. y $F_1, F_2, \ldots, F_n$ subespacios de E. El subconjunto de E

$$\{x_1 + x_2 + \cdots + x_n | x_1 \in F_1, x_2 \in F_2, \ldots, x_n \in F_n\}$$

es un subespacio de E; le llamamos *subespacio suma* de $F_1, F_2, \ldots, F_n$ y lo representamos por

$$F_1 + F_2 + \cdots + F_n .$$

Como en (1.3.5), resulta que

$$F_1 + F_2 + \cdots + F_n = \langle F_1 \cup F_2 \cup \cdots \cup F_n \rangle .$$

1.3.21 Ejemplo. Si $(x_1, \ldots, x_p)$ es un sistema de E, entonces

$$\langle x_1, \ldots, x_p \rangle = \mathbb{K}x_1 + \cdots + \mathbb{K}x_p ,$$

donde $\mathbb{K}x_i$ es el subespacio descrito en (1.2.9).

1.3.22 DEFINICIÓN. Sea E un e.v., $F_1, F_2, \ldots, F_n$ subespacios y

$$F = F_1 + F_2 + \cdots + F_n .$$

Cada $x \in F$ es una suma

$$x = x_1 + x_2 + \cdots + x_n ,$$

con $x_1 \in F_1, x_2 \in F_2, \ldots, x_n \in F_n$. Cuando cada $x \in F$ posee una única descomposición de esta forma, decimos que F es *suma directa* de $F_1, F_2, \ldots, F_n$ y escribimos

$$F = F_1 \oplus F_2 \oplus \cdots \oplus F_n .$$

1.3.23 Ejemplo. Como veremos más adelante (v. 1.4.21), si en las condiciones del ejemplo (1.3.21) tenemos además

$$\langle x_1, \ldots, x_p \rangle = \mathbb{K}x_1 \oplus \cdots \oplus \mathbb{K}x_p \,,$$

entonces $(x_1, \ldots, x_p)$ es una 'base' de $\langle x_1, \ldots, x_p \rangle$.

1.3.24 Ejemplo. $\mathbb{R}^3$ es suma directa de los subespacios

$$\begin{aligned} F_1 &= \{(x,y,z) \in \mathbb{R}^3 | y = z = 0\} \,, \\ F_2 &= \{(x,y,z) \in \mathbb{R}^3 | x = z = 0\} \,, \\ F_3 &= \{(x,y,z) \in \mathbb{R}^3 | x = y = 0\} \,. \end{aligned}$$

1.3.25 PROPOSICIÓN

Sea E un e.v. y $F, F_1, F_2, \ldots, F_n$ subespacios de E. Las siguientes propiedades son equivalentes:

(i) $F = F_1 \oplus F_2 \oplus \cdots \oplus F_n$,

(ii) $F = F_1 + F_2 + \cdots + F_n$ y, además,

$$(F_1 + \cdots + F_i) \cap F_{i+1} = \{0\}$$

para $i = 1, \ldots, n-1$.

La demostración es similar a la de (1.3.12).

1.3.26 Si $F, F_1, F_2, \ldots, F_n$ son subespacios de un e.v. E tales que

$$F = F_1 \oplus F_2 \oplus \cdots \oplus F_n \,,$$

entonces

$$F \simeq F_1 \times F_2 \times \cdots \times F_n$$

puesto que la aplicación

$$\begin{aligned} F_1 \times F_2 \times \cdots \times F_n &\to F \\ (x_1, x_2, \ldots, x_n) &\to x_1 + x_2 + \cdots + x_n \end{aligned}$$

es un isomorfismo.

1.4 Bases de un espacio vectorial.

1.4.1 Sea E un e.v. sobre $\mathbb{K}$ y $(x_1, \ldots, x_p)$ un sistema de E (v. 1.2.11). El vector 0 es siempre combinación lineal de $(x_1, \ldots, x_p)$, puesto que

$$0 = 0x_1 + \cdots + 0x_p \,.$$

Pero es posible que 0 sea combinación lineal de $(x_1, \ldots, x_p)$ en alguna otra forma diferente de ésta (que es la llamada forma 'trivial'), es decir, es posible que

$$0 = \lambda^1 x_1 + \cdots + \lambda^p x_p$$

con escalares $\lambda^1, \ldots \lambda^p$ no todos nulos.

1.4.2 Ejemplo. En el espacio $\mathbb{R}^2$, 0 es combinación lineal del sistema de vectores $((1,0),(0,2))$ únicamente en la forma

$$0 = (0,0) = 0(1,0) + 0(0,2)$$

o sea, en la forma trivial.

0 es también combinación lineal del sistema $((1,0),(0,2),(2,0))$ en la forma trivial:

$$0 = (0,0) = 0(1,0) + 0(0,2) + 0(2,0)\,,$$

pero en este caso lo es también en otras formas; por ejemplo,

$$0 = (0,0) = -4(1,0) + 0(0,2) + 2(2,0)\,.$$

1.4.3 DEFINICIÓN. Sea E un e.v. sobre $\mathbb{K}$ y $(x_1, \ldots, x_p)$ un sistema de E. Decimos que el sistema es *libre*[2] cuando la única forma en que 0 es combinación lineal del sistema es la trivial, o sea, cuando

$$\textbf{(lb1)} \quad \lambda^1 x_1 + \cdots + \lambda^p x_p = 0 \quad \Rightarrow \quad \lambda^1 = \cdots = \lambda^p = 0\,,$$

y decimos que el sistema es *ligado* cuando 0 es combinación lineal del sistema en alguna forma distinta de la trivial, o sea, cuando

(lg1) Existen $\lambda^1, \ldots, \lambda^p$ no todos nulos, tales que

$$\lambda^1 x_1 + \cdots + \lambda^p x_p = 0\,.$$

De las definiciones se desprende que cualquier sistema finito o bien es libre o bien es ligado.

1.4.4 Ejemplo. De los dos sistemas del ejemplo (1.4.2), el primero es libre y el segundo es ligado.

[2] En algunos textos se dice también 'sistema linealmente independiente' o 'sistema de vectores independientes' además de sistema libre. Hemos preferido no utilizar esta terminología.

1.4.5 La propiedad **(lb1)** es equivalente a la siguiente:

> **(lb2)** Cada vector x de $\langle x_1, \ldots, x_p \rangle$ es, de forma única, una combinación lineal
> $$x = \lambda^1 x_1 + \cdots + \lambda^p x_p$$
> de $(x_1, \ldots, x_p)$.

Obsérvese que, por definición (v. 1.2.16), todo vector x de $\langle x_1, \ldots, x_p \rangle$ es combinación lineal de $(x_1, \ldots, x_p)$; lo que afirma **(lb2)** es que la igualdad $x = \lambda^1 x_1 + \cdots + \lambda^p x_p$ es cierta para un único grupo de escalares $\lambda^1, \ldots, \lambda^p$.

La propiedad **(lb1)** es también equivalente a

> **(lb3)** $\langle x_1, \ldots, x_p \rangle = \mathbb{K}x_1 \oplus \cdots \oplus \mathbb{K}x_p$,

que afirma que la suma de (1.3.21) es directa.

1.4.6 La propiedad **(lg1)** es equivalente a la siguiente:

> **(lg2)** Uno de los vectores x_i del sistema $(x_1, \ldots, x_p)$ es combinación lineal de los restantes, o sea,
> $$x_i = \lambda^1 x_1 + \cdots + \lambda^{i-1} x_{i-1} + \lambda^{i+1} x_{i+1} + \cdots + \lambda^p x_p \,.$$

La propiedad **(lg2)** no significa que, en un sistema ligado, cada vector del sistema sea combinación lineal de los restantes; así lo muestra el ejemplo que sigue. En general no podremos decir para qué vector o vectores del sistema tal cosa es cierta. No obstante, la proposición (1.4.11) nos muestra un caso en el que podemos precisar más esta cuestión.

1.4.7 El sistema de $\mathbb{R}^2$

$$((0,2),(1,0),(0,0),(2,0))$$

es ligado y los tres últimos vectores son cada uno combinación lineal de los restantes; sin embargo, $(0,2)$ no es combinación lineal de los tres restantes.

1.4.8 Si $(x_1, \ldots, x_p)$ es libre (o si es ligado), cualquier reordenación del sistema posee el mismo carácter, o sea, si q es una permutación de $\{1, 2, \ldots, p\}$, los sistemas

$$(x_1, \ldots, x_p) \quad \text{y} \quad (x_{q(1)}, \ldots, x_{q(p)})$$

son ambos libres o ambos ligados.

1.4.9 Si $(x_1, \ldots, x_p)$ es libre, entonces todos los x_i son no nulos. En efecto, si $x_i = 0$, tendríamos

$$0x_1 + \cdots + 1x_i + \cdots + 0x_p = 0$$

y el sistema sería ligado.

Todo sistema que contenga a 0 es, pues, ligado.

1.4.10 Si $(x_1, \ldots, x_p)$ es libre, entonces el sistema es inyectivo (v. 1.2.11), o sea, los x_i son todos distintos entre sí. En efecto, si $i \neq j$ y $x_i = x_j$, tendríamos

$$0x_1 + \cdots + 1x_i + \cdots + (-1)x_j + \cdots + 0x_p = 0$$

y el sistema sería ligado.

1.4.11 PROPOSICIÓN

Si $(x_1, \ldots, x_p)$ es libre y $(x_1, \ldots, x_p, x_{p+1})$ es ligado, entonces

$$x_{p+1} \in \langle x_1, \ldots, x_p \rangle \, .$$

Existen $\lambda^1, \ldots \lambda^p, \lambda^{p+1}$ no todos nulos y tales que

$$\lambda^1 x_1 + \cdots + \lambda^p x_p + \lambda^{p+1} x_{p+1} = 0 \, .$$

Si $\lambda^{p+1} = 0$, entonces $(x_1, \ldots, x_p)$ sería ligado, luego $\lambda^{p+1} \neq 0$ y resulta que

$$x_{p+1} = -\frac{\lambda^1}{\lambda^{p+1}} x_1 - \cdots - \frac{\lambda^p}{\lambda^{p+1}} x_p \, .$$

1.4.12 PROPOSICIÓN

---

a) Si $(x_1, \ldots, x_p)$ es libre y $1 \leq i \leq p$, entonces $(x_1, \ldots, x_i)$ es libre.

b) Si $(x_1, \ldots, x_p)$ es ligado, entonces $(x_1, \ldots, x_p, x_{p+1}, \ldots, x_n)$ es ligado, cualesquiera que sean los vectores $x_{p+1}, \ldots, x_n$.

---

a) Como $(x_1, \ldots, x_p)$ es libre,

$$\lambda^1 x_1 + \cdots \lambda^p x_p = 0 \quad \Rightarrow \quad \lambda^1 = \cdots = \lambda^p = 0 \, ;$$

entonces, si

$$\lambda^1 x_1 + \cdots \lambda^i x_i = 0$$

resulta que

$$\lambda^1 x_1 + \cdots \lambda^i x_i + 0x_{i+1} + \cdots + 0x_p = 0$$

y por lo tanto

$$\lambda^1 = \cdots = \lambda^i = 0 \, ,$$

lo que demuestra que $(x_1, \ldots, x_i)$ es libre.

b) Si $(x_1, \ldots, x_p)$ es ligado, entonces $(x_1, \ldots, x_p, x_{p+1}, \ldots, x_n)$ ha de ser ligado pues, en caso contrario, aplicando el apartado a), resultaría que $(x_1, \ldots, x_p)$ es libre.

1.4.13 La proposición precedente y el resultado sobre reordenación de sistemas que hemos visto en (1.4.8) nos permiten afirmar que si suprimimos vectores de un sistema libre obtenemos un sistema libre, y que si añadimos vectores a un sistema ligado obtenemos un sistema ligado.

1.4.14 Un sistema (x) con un solo vector es libre si $x \neq 0$, y es ligado si $x = 0$.

1.4.15 Si un sistema (x_1, x_2) de dos vectores es ligado, decimos que x_1 y x_2 son *colineales*. En ese caso tenemos que, o bien $x_1 = \lambda x_2$, o bien $x_2 = \lambda x_1$ (pero no necesariamente ambas cosas).

1.4.16 PROPOSICIÓN

Si $(y_1, \ldots, y_{p+1})$ es un sistema cuyos $p+1$ vectores son combinación lineal de un sistema $(x_1, \ldots, x_p)$ de p vectores, entonces $(y_1, \ldots, y_{p+1})$ es ligado.

Lo demostraremos por inducción sobre p.

Para $p = 1$ es cierto; en efecto, si $y_1, y_2 \in \langle x_1 \rangle$, se tiene $y_1 = \lambda x_1, y_2 = \mu x_1$, luego $\mu y_1 - \lambda y_2 = 0$. Si $\mu \neq 0$ o $\lambda \neq 0$, esto prueba que (y_1, y_2) es ligado; si $\mu = \lambda = 0$, entonces, como $y_1 = y_2 = 0$, también es ligado.

Supondremos cierta la proposición para $p - 1$. Sean entonces

$$y_1, y_2, \ldots, y_{p+1} \in \langle x_1, \ldots, x_p \rangle \,,$$

es decir,

$$\begin{aligned} y_1 &= \lambda_1^1 x_1 + \cdots + \lambda_1^p x_p \\ \ldots & \quad \ldots \quad \ldots \quad \ldots \quad \ldots \\ y_{p+1} &= \lambda_{p+1}^1 x_1 + \cdots + \lambda_{p+1}^p x_p \,. \end{aligned}$$

Si $y_1 = 0$, el sistema $(y_1, \ldots, y_p, y_{p+1})$ es ligado; si $y_1 \neq 0$, alguno de los λ_1^i es no nulo; supondremos, con el fin de simplificar la notación, que $\lambda_1^1 \neq 0$.

Consideremos entonces los vectores $y'_2, \ldots, y'_{p+1}$ definidos por

$$y'_2 = y_2 - \frac{\lambda_2^1}{\lambda_1^1} y_1 \,, \ldots, y'_{p+1} = y_{p+1} - \frac{\lambda_{p+1}^1}{\lambda_1^1} y_1 \,.$$

Se tiene

$$y'_2, \ldots, y'_{p+1} \in \langle x_2, \ldots, x_p \rangle \,.$$

Entonces, como hemos supuesto cierta la proposición para $p - 1$, resulta que $(y'_2, \ldots, y'_{p+1})$ es ligado. Es decir, existen $\mu^2, \ldots, \mu^{p+1}$ no todos nulos tales que

$$\mu^2 y'_2 + \cdots + \mu^{p+1} y'_{p+1} = 0 \,.$$

Haciendo ahora la substitución de $y'_2, \dots, y'_{p+1}$ obtenemos

$$\theta y_1 + \mu^2 y_2 + \cdots + \mu^{p+1} y_{p+1} = 0\,,$$

donde

$$\theta = -\frac{1}{\lambda_1^1}\left(\mu^2\lambda_2^1 + \mu^3\lambda_3^1 + \cdots + \mu^{p+1}\lambda_{p+1}^1\right)\,.$$

Como alguno de los $\mu^2, \dots, \mu^{p+1}$ es no nulo, el sistema $(y_1, \dots, y_{p+1})$ es ligado. Esto completa la recurrencia.

1.4.17 COROLARIO

Sea $(x_1, \dots, x_p)$ un sistema generador de E. Entonces cualquier sistema de E que tenga más de p vectores es ligado.

1.4.18 Sea E un e.v. y supongamos que el sistema $(x_1, \dots, x_n)$ de E es libre y que además genera E (v. 1.2.16), o sea, que $E = \langle x_1, \dots, x_n\rangle$. Utilizando la propiedad **(lb2)** de (1.4.5) resulta entonces que cada vector x de E es, de forma única, una combinación lineal $x = \lambda^1 x_1 + \cdots \lambda^n x_n$ del sistema. Es decir, a cada vector de E podemos asociarle unívocamente un sistema $(\lambda^1, \dots, \lambda^n)$ de escalares, o sea, un elemento de $\mathbb{K}^n$.

Un sistema libre y generador de E es lo que en (1.4.21) llamaremos 'base' de E. Con un sistema de este tipo los vectores de E se representan como elementos de $\mathbb{K}^n$.

Evidentemente, si en E existe una base, E es de dimensión finita (v. 1.2.16) y además $E \neq \{0\}$, pues no existe (a causa de 1.4.9) ningún sistema libre en $\{0\}$.

Veremos en (1.4.30) que en todo espacio de dimensión finita distinto de $\{0\}$ hay al menos una base, y en (1.4.27) que, cuando $\mathbb{K}$ es infinito, hay infinitas bases.

1.4.19 Existe una noción de base más general que la nuestra, con la cual todo espacio vectorial, aunque no sea de dimensión finita, admite una base. Sin embargo, no realizaremos el estudio de esta cuestión.

1.4.20 TEOREMA

Sea E un e.v. sobre $\mathbb{K}$ y $(x_1, \dots, x_n)$ un sistema de E. Son equivalentes las siguientes proposiciones:

(b1) $(x_1, \dots, x_n)$ es libre y es un sistema generador de E.

(b2) Cada vector $x \in E$ es, de forma única, una combinación lineal $x = \lambda^1 x_1 + \cdots + \lambda^n x_n$ de $(x_1, \dots, x_n)$.

(b3) $E = \mathbb{K}x_1 \oplus \mathbb{K}x_2 \oplus \cdots \oplus \mathbb{K}x_n$.

(b4) $(x_1, \dots, x_n)$ es un sistema generador de E y, además, ninguno de los sistemas de $n-1$ vectores que resultan de suprimir algún vector de $(x_1, \ldots, x_n)$ genera E.

(b5) $(x_1, \ldots, x_n)$ es un sistema libre y, además, el sistema $(x_1, \ldots, x_n, x_{n+1})$ de $n+1$ vectores es ligado, cualquiera que sea el vector $x_{n+1} \in E$.

(b1)⇔(b2) y **(b1)⇔(b3)** son consecuencias inmediatas de (1.4.5).

(b1)⇒(b4) se prueba fácilmente por reducción al absurdo; en efecto, si suponemos por ejemplo que $(x_1, \ldots, x_{n-1})$ genera E, resulta que x_n es combinación lineal de $(x_1, \ldots, x_{n-1})$ y que $(x_1, \ldots, x_n)$ es ligado (v. 1.4.6).

(b4)⇒(b1) se prueba también por reducción al absurdo; en efecto, si suponemos que $(x_1, \ldots, x_n)$ es ligado, entonces se verifica la propiedad **(lg2)** de (1.4.6) y el sistema $(x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n)$ genera E (v. 1.2.20).

(b1)⇒(b5), puesto que, si $(x_1, \ldots, x_n)$ genera E y $x_{n+1} \in E$, entonces el vector x_{n+1} es combinación lineal de $(x_1, \ldots, x_n)$ y, aplicando (1.4.6), resulta que $(x_1, \ldots, x_n, x_{n+1})$ es ligado.

(b5)⇒(b1) resulta inmediatamente de aplicar (1.4.11).

1.4.21 DEFINICIÓN. Sea E un espacio vectorial. De un sistema $(x_1, \ldots, x_n)$ de E que verifica las propiedades del teorema precedente decimos que es una *base* de E.

Si $(x_1, \ldots, x_n)$ es una base de E y $x \in E$, entonces

$$x = \lambda^1 x_1 + \cdots + \lambda^n x_n$$

de forma única. Decimos que los escalares $\lambda^1, \ldots, \lambda^n$ son las *coordenadas* del vector x en la base $(x_1, \ldots, x_n)$ y representamos este hecho poniendo

$$x = (\lambda^1, \ldots, \lambda^n)_{(x_1, \ldots, x_n)} ,$$

o también

$$x = (\lambda^1, \ldots, \lambda^n)_{(x_i)}$$

cuando poner (x_i) baste para saber a qué base $(x_1, \ldots, x_n)$ nos referimos.

Será usual representar una base de un espacio vectorial por $(a_1, \ldots, a_n)$ o $(b_1, \ldots, b_n)$. Si $(a_1, \ldots, a_n)$ es una base de E y $x \in E$, es también frecuente poner

$$x = (x^1, \ldots, x^n)_{(a_i)}$$

denotando los escalares $x_1, \ldots, x_n$ por minúsculas latinas, en contra de lo dicho en (1.1.1). La comodidad de estas notaciones resultará pronto evidente.

1.4.22 Ejemplo. Si $(a_1, \ldots, a_n)$ es una base de E, entonces

$$0 = (0, \ldots, 0)_{(a_i)}$$

y

$$a_i = (0, \ldots, 1, \ldots, 0)_{(a_i)} ,$$

con 1 en el lugar número i, para $i = 1, 1, \ldots, n$.

1.4.23 Si $(a_1, \ldots, a_n)$ es una base de E, la aplicación

$$\begin{aligned} \mathbb{K}^n &\to E \\ (\lambda^1, \ldots, \lambda^n) &\to \lambda^1 a_1 + \cdots + \lambda^n a_n \end{aligned}$$

es un isomorfismo entre los espacios vectoriales $\mathbb{K}^n$ y E. La aplicación inversa de este isomorfismo (que es tambián un isomorfismo) envía cada vector de E a la n-upla formada por sus coordenadas en la base $(a_1, \ldots, a_n)$.

Estas afirmaciones encierran, a pesar de su sencillez, hechos importantes que citamos a continuación:

a) Si $x = (x^1, \ldots, x^n)_{(a_i)}$ y $y = (y^1, \ldots, y^n)_{(a_i)}$, entonces

$$x + y = (x^1 + y^1, \ldots, x^n + y^n)_{(a_i)} .$$

b) Si $x = (x^1, \ldots, x^n)_{(a_i)}$ y $\lambda \in \mathbb{K}$, entonces

$$\lambda x = (\lambda x^1, \ldots, \lambda x^n)_{(a_i)} .$$

c) Si $x = (x^1, \ldots, x^n)_{(a_i)}$ y $y = (y^1, \ldots, y^n)_{(a_i)}$, y si $x^i = y^i$ para $i = 1, \ldots, n$, entonces $x = y$.

d) Si $(\lambda^1, \ldots, \lambda^n) \in \mathbb{K}^n$, existe un vector de E cuyas coordenadas en $(a_1, \ldots, a_n)$ son $\lambda^1, \ldots, \lambda^n$, y es el vector $\lambda^1 a_1 + \cdots + \lambda^n a_n$.

1.4.24 Ejemplo. El sistema $(e_1, \ldots, e_n)$ de $\mathbb{K}^n$ formado por los vectores de $\mathbb{K}^n$ definidos por

$$e_i = (0, \ldots, 1, \ldots, 0)$$

(con 1 en el lugar número i), es una base de $\mathbb{K}^n$ que recibe el nombre de *base canónica* de $\mathbb{K}^n$.

Si $(\lambda^1, \ldots, \lambda^n)$ es un vector de K^n, las coordenadas de este vector en la base canónica son justamente $\lambda^1, \ldots, \lambda^n$, o sea,

$$(\lambda^1, \ldots, \lambda^n) = (\lambda^1, \ldots, \lambda^n)_{(e_1, \ldots, e_n)} .$$

Cuando un vector x de $\mathbb{K}^n$ tenga por coordenadas en la base canónica $\lambda^1, \ldots, \lambda^n$, pondremos simplemente

$$x = (\lambda^1, \ldots, \lambda^n)$$

omitiendo escribir (e_i). Por lo que acabamos de ver, tal omisión no da lugar, en este caso, a confusión alguna.

El isomorfismo

$$\begin{aligned} \mathbb{K}^n &\to \mathbb{K}^n \\ (\lambda^1, \ldots, \lambda^n) &\to \lambda^1 e_1 + \cdots + \lambda^n e_n \end{aligned}$$

del párrafo precedente es, en este caso, la aplicación idéntica de $\mathbb{K}^n$ (v. 1.1.11).

1.4.25 TEOREMA

Dos bases cualesquiera de un e.v. E tienen el mismo número de elementos.

O sea, si $(a_1, \ldots, a_n)$ y $(b_1, \ldots, b_m)$ son dos bases de E, entonces $n = m$. En efecto, si suponemos por ejemplo que $n < m$, entonces el sistema $(b_1, \ldots, b_m)$ sería ligado (v. 1.4.17), lo que no puede ser cierto.

1.4.26 Para la noción de base a la que aludíamos en (1.4.19), y que comprende el caso de bases que son conjuntos infinitos, el teorema precedente es aún cierto en el sentido siguiente: dos bases cualesquiera de un espacio vectorial son equipotentes.

1.4.27 Si $(a_1, \ldots, a_n)$ es una base de E y $\lambda^1 \neq 0, \ldots, \lambda^n \neq 0$, entonces $(\lambda^1 a_1, \ldots$ $\ldots, \lambda^n a_n)$ es una base de E, como se puede ver utilizando (1.2.18) y caracterizando una base mediante la propiedad **(b3)**.

1.4.28 Si $(a_1, \ldots, a_n)$ es una base de E y $p \in G_n$ es una permutación de $\{1, 2, \ldots, n\}$, entonces $(a_{p(1)}, \ldots, a_{p(n)})$ es una base de E (véase 1.4.8).

1.4.29 PROPOSICIÓN

Sea E un e.v. de dimensión finita (v. 1.2.16) y $(x_1, \ldots, x_p)$ un sistema libre de E. Entonces existe algún sistema $(x_1, \ldots, x_n)$ de E, con $n \geq p$, cuyos primeros p vectores son $x_1, \ldots, x_p$, que es una base de E.

Si $(x_1, \ldots, x_p)$ es ya una base de E, la proposición es cierta. Supondremos, pues, que $(x_1, \ldots, x_p)$ no es una base, lo que significa (v. 1.4.20, propiedad **(b5)**) que existe un vector $x_{p+1} \in E$ tal que $(x_1, \ldots, x_p, x_{p+1})$ es libre.

Como E es de dimensión finita, existe un sistema $(y_1, \ldots, y_m)$ que genera E; sabemos entonces (v. 1.4.17) que todo sistema con más de m vectores es ligado. En particular resulta que $p + 1 \leq m$.

Consideremos todos los sistemas de la forma $(x_1, \ldots, x_p, x_{p+1}, \ldots, x_q)$ que sean libres. El conjunto de los números naturales q para los que existe un sistema de este tipo está contenido en $\{p + 1, \ldots, m\}$ y es, pues, finito. Sea n el máximo de tales números q .

El número n verifica entonces que:

— existen $x_{p+1}, \ldots, x_n \in E$ tales que $(x_1, \ldots, x_p, x_{p+1}, \ldots, x_n)$ es libre;

— cualquiera que sea $x_{n+1} \in E$, el sistema $(x_1, \ldots, x_p, x_{p+1}, \ldots, x_n,$ $x_{n+1})$ es ligado.

Resulta así que el sistema $(x_1, \ldots, x_p, x_{p+1}, \ldots, x_n)$ verifica la propiedad **(b5)** de (1.4.20) y es una base de E.

1.4.30 TEOREMA

Si E es un e.v. de dimensión finita y $E \neq \{0\}$, existe una base de E.

Como $E \neq \{0\}$, existe $x \in E, x \neq 0$; el sistema de un solo vector (x) es libre (v. 1.4.14) y, aplicando la proposición anterior, o bien es una base, o bien hay una base que lo prolonga.

1.4.31 DEFINICIÓN. Sea E un e.v. de dimensión finita y $E \neq \{0\}$. E admite una base (v. 1.4.30), y además todas las bases de E tienen igual número de vectores (v. 1.4.25). Del número de elementos de una base de E decimos que es la *dimensión* de E; lo denotamos por $\dim E$.

También pondremos $\dim\{0\} = 0$.

Todo e.v. de dimensión $n \neq 0$ es isomorfo a $\mathbb{K}^n$, como vimos en (1.4.23).

1.4.32 Ejemplo. Si $\mathbb{K}$ es un cuerpo conmutativo, el espacio $\mathbb{K}^n$ es de dimensión n (v. 1.4.24); en particular, $\mathbb{K}$ es de dimensión 1.

1.4.33 Modificando la demostración de (1.4.29) se puede probar que, si E es de dimensión n, si $(x_1, \ldots, x_p)$ es un sistema libre de E con $p < n$ y si $\langle S \rangle = E$ (v. 1.2.22), entonces existen $x_{p+1}, \ldots, x_n \in S$ tales que $(x_1, \ldots, x_p, x_{p+1}, \ldots, x_n)$ es una base de E. Este enunciado suele recibir el nombre de 'teorema de la base incompleta'.

1.4.34 PROPOSICIÓN

Sea E un e.v. de dimensión n; entonces:

a) todo sistema que tenga $n + 1$ elementos o más es ligado,

b) todo sistema libre tiene a lo sumo n elementos,

c) todo sistema generador de E tiene por lo menos n elementos.

a) y c) son consecuencias sencillas de (1.4.17); b) no es más que otra manera de enunciar a).

1.4.35 COROLARIO

Si $(x_1, \ldots, x_p)$ es un sistema generador de un espacio vectorial E, entonces $\dim E \leq p$.

1.4.36 COROLARIO

Sea E un e.v. de dimensión $n \neq 0$ y $(x_1, \ldots, x_n)$ un sistema de E con n elementos (nótese que se trata de la misma cantidad que la dimensión del espacio); entonces son equivalentes las siguientes proposiciones:

$$\begin{aligned} \text{(i)} &\ (x_1, \ldots, x_n) \text{ es una base de } E\,, \\ \text{(ii)} &\ (x_1, \ldots, x_n) \text{ es libre}\,, \\ \text{(iii)} &\ (x_1, \ldots, x_n) \text{ genera } E\,. \end{aligned}$$

(i)⇒(ii) y (i)⇒(iii) resultan inmediatamente de la definición de base.

(ii)⇒(i): si $(x_1, \ldots, x_n)$ es libre, aplicando el apartado a) de la proposición precedente resulta que $(x_1, \ldots, x_n)$ verifica **(b5)** y, por tanto, que es una base.

(iii)⇒(i): si $(x_1, \ldots, x_n)$ genera E, aplicando el apartado c) de la proposición precedente resulta que $(x_1, \ldots, x_n)$ verifica **(b4)** y, por tanto, que es una base.

1.4.37 PROPOSICIÓN

Sean E, E' dos e.v. sobre $\mathbb{K}$ y $f : E \to E'$ un isomorfismo. Si $(a_1, \ldots, a_n)$ es una base de E, entonces $(f(a_1), \ldots, f(a_n))$ es una base de E'.

Si
$$\lambda^1 f(a_1) + \cdots + \lambda^n f(a_n) = 0\,,$$
resulta que (v. 1.1.12)
$$f(\lambda^1 a_1 + \cdots + \lambda^n a_n) = 0 = f(0)\,;$$
como f es inyectiva, entonces
$$\lambda^1 a_1 + \cdots + \lambda^n a_n = 0$$
y, como $(a_1, \ldots, a_n)$ es libre,
$$\lambda^1 = \cdots = \lambda^n = 0\,,$$
lo que demuestra que $(f(a_1), \ldots, f(a_n))$ es libre.

Si $y \in E'$, como f es sobre, resulta que existe $x \in E$ con $f(x) = y$. Pero
$$x = \lambda^1 a_1 + \cdots + \lambda^n a_n\,,$$
luego
$$y = f(x) = \lambda^1 f(a_1) + \cdots + \lambda^n f(a_n)\,.$$
Esto demuestra que $(f(a_1), \ldots, f(a_n))$ genera E' y, como ya hemos visto que es libre, resulta que es una base de E'.

1.4.38 PROPOSICIÓN

a) Si E y E' son isomorfos y $\dim E = n$, entonces $\dim E' = n$.

b) Dos espacios vectoriales de dimensión finita son isomorfos si y sólo si tienen la misma dimensión.

a) Para $n = 0$ (o sea $E = \{0\}$) es evidente; para $n \neq 0$ es la proposición anterior.

b) Sean E y E' de dimensión finita. Si suponemos que $E \simeq E'$, entonces, aplicando el apartado a), resulta que $\dim E = \dim E'$. Recíprocamente, si suponemos que $\dim E = \dim E'$, entonces, o bien ambos son $\{0\}$ y por tanto isomorfos, o bien su dimensión común es $n \neq 0$ y, como ambos son isomorfos a $\mathbb{K}^n$, lo son entre sí.

1.5 Dimensión de un subespacio.

1.5.1 Sea E un espacio vectorial, F un subespacio y $(x_1, \ldots, x_p)$ un sistema de F. El sistema $(x_1, \ldots, x_p)$ es también un sistema de E. Las propiedades **(lb1)** y **(lg1)** de (1.4.3) referidas al sistema $(x_1, \ldots, x_p)$ tienen el mismo significado tanto si consideramos que $(x_1, \ldots, x_p)$ es un sistema de E como si consideramos que lo es de F, puesto que ambos, E y F, son espacios vectoriales sobre el mismo cuerpo, poseen las mismas operaciones y $0_E = 0_F$. Decir que $(x_1, \ldots, x_p)$ es libre (o ligado) representa lo mismo en ambos casos, por lo que nunca será necesario precisar en qué aspecto lo estamos considerando.

1.5.2 PROPOSICIÓN

Sea E un e.v. y $(x_1, \ldots, x_p)$ un sistema de E. El subespacio $\langle x_1, \ldots, x_p\rangle$ es siempre de dimensión finita y

$$\dim\langle x_1, \ldots, x_p\rangle \leq p\,.$$

La dimensión es igual a p si y sólo si $(x_1, \ldots, x_p)$ es libre, en cuyo caso es una base de $\langle x_1, \ldots, x_p\rangle$.

$(x_1, \ldots, x_p)$ es un sistema generador de $\langle x_1, \ldots, x_p\rangle$, de lo cual resulta la primera afirmación. Además, si $(x_1, \ldots, x_p)$ es libre, es un sistema libre y generador de $\langle x_1, \ldots, x_p\rangle$, luego es una base y $\dim\langle x_1, \ldots, x_p\rangle = p$. Recíprocamente, si la dimensión es p, entonces $(x_1, \ldots, x_p)$ es libre (v. 1.4.36).

1.5.3 Sea E un e.v. de dimensión $n \neq 0$ y $(a_1, \ldots, a_n)$ una base de E; sea F un subespacio de E. Si bien todo elemento de F es combinación lineal de $(a_1, \ldots, a_n)$, no hay que pensar que $F = \langle a_1, \ldots, a_n\rangle$, pues, en general, los vectores a_i no estarán en F. Es cierto, sin embargo, que F está generado por un sistema (finito) de vectores, aunque la argumentación sea más complicada, como vemos a continuación.

1.5.4 TEOREMA

Sea E un e.v. de dimensión finita y F un subespacio de E.

a) F es también de dimensión finita y

$$\dim F \leq \dim E \,.$$

b) Si $\dim F = \dim E$, entonces $F = E$.

Ambas propiedades resultan evidentes cuando $F = \{0\}$; supondremos entonces que $F \neq \{0\}$, luego, en particular, que $\dim E = n \neq 0$.

a) Existe un vector x de F no nulo, luego en F existe al menos un sistema libre.

Consideremos todos los sistemas $(x_1, \ldots, x_q)$ de F que son libres; como son al mismo tiempo sistemas libres de E, resulta que $q \leq n$. El conjunto de los números naturales q para los que existe un sistema de ese tipo está contenido en $\{1, \ldots, n\}$ y es, pues, finito. Sea p el máximo de tales números q .

El número p verifica entonces que:

— existen $x_1, \ldots, x_p \in F$ tales que $(x_1, \ldots, x_p)$ es libre;

— cualquiera que sea $x_{p+1} \in F$, el sistema $(x_1, \ldots, x_p, x_{p+1})$ es ligado;

— $p \leq n$.

Resulta así que el sistema $(x_1, \ldots, x_p)$ verifica la propiedad **(b5)** de (1.4.20) y es una base de F. Luego F es de dimensión finita, con $\dim F \leq \dim E$.

b) Si $\dim F = \dim E = n \neq 0$, sea $(x_1, \ldots, x_n)$ una base de F; esta base es un sistema libre, luego es un sistema libre de E y por tanto (v. 1.4.36) una base de E. Resulta entonces que

$$F = \langle x_1, \ldots, x_n \rangle = E \,.$$

1.5.5 Sea E un e.v. sobre $\mathbb{K}$. Los subespacios de dimensión 1 se denominan *rectas vectoriales*; son los subespacios de la forma

$$\mathbb{K}x = \langle x \rangle = \{\lambda x | \lambda \in \mathbb{K}\} \,,$$

con $x \neq 0$ (v. 1.2.9).

Si E es de dimensión $n \neq 0$, los subespacios de dimensión $n-1$ se denominan *hiperplanos vectoriales*.

1.5.6 PROPOSICIÓN

Sea E un e.v. de dimensión $n \neq 0$, F un subespacio de E de dimensión p con $0 < p < n$ y $(a_1, \ldots, a_p)$ una base de F. Existen vectores $a_{p+1}, \ldots, a_n$ de E tales que $(a_1, \ldots, a_p, a_{p+1}, \ldots, a_n)$ es una base de E.

Basta utilizar la proposición (1.4.29).

1.5.7 PROPOSICIÓN

Sea E un e.v. de dimensión finita.

a) Sea F un subespacio de E; F admite un suplementario (v. 1.3.14).

b) Si $E = F_1 \oplus F_2 \oplus \cdots \oplus F_p$ y si, para todo $i = 1, \ldots, p$, $(a^i_1, \ldots, a^i_{n_i})$ es una base de F_i, entonces $(a^1_1, \ldots, a^1_{n_1}, a^2_1, \ldots, a^2_{n_2}, \ldots, a^p_{n_p})$ es una base de E.

c) Si $E = F_1 \oplus F_2$, entonces $\dim E = \dim F_1 + \dim F_2$.

---

a) Si F es un subespacio trivial (v. 1.2.8), según hemos visto en (1.3.15) F admite un suplementario. Si F es propio, sea $(a_1, \ldots, a_p)$ una base de F y sean $a_{p+1}, \ldots, a_n \in E$ tales que $(a_1, \ldots, a_p, a_{p+1}, \ldots, a_n)$ es una base de E. Entonces $\langle a_{p+1}, \ldots, a_n \rangle$ es un suplementario de F.

b) es sencilla; c) es una consecuencia de b).

1.5.8 Sea E un e.v. de dimensión finita y F un subespacio de E. Todos los suplementarios de F tienen la misma dimensión, $\dim E - \dim F$, a causa del apartado c) de la proposición precedente, luego son isomorfos (v. 1.4.38). Llegamos así al mismo resultado de (1.3.19), pero sólo para el caso de dimensión finita.

1.5.9 DEFINICIÓN. Sea E un e.v. y $(x_1, \ldots, x_p)$ un sistema de E. Denominamos *rango* de $(x_1, \ldots, x_p)$ y denotamos por

$$\mathrm{rg}(x_1, \ldots, x_p)$$

la dimensión del subespacio $\langle x_1, \ldots, x_p \rangle$. Hemos visto en (1.5.2) que

$$\mathrm{rg}(x_1, \ldots, x_p) \leq p$$

y que $\mathrm{rg}(x_1, \ldots, x_p) = p$ si y sólo si $(x_1, \ldots, x_p)$ es libre.

Si además $\dim E = n$, entonces

$$\mathrm{rg}(x_1, \ldots, x_p) \leq n\,.$$

1.5.10 Sea $(a_1, \ldots, a_n)$ una base de E y sea $(x_1, \ldots, x_p)$ un sistema de E, con

$$x_i = (x^1_i, \ldots, x^n_i)_{(a_i)} = \sum_{j=1}^{n} x^j_i a_j\,, \quad i = 1, \ldots, p\,.$$

Si $p \leq n$ y además

$$\begin{array}{lll} \text{(a)} & (\forall i) & x^i_i \neq 0\,, \\ \text{(b)} & j < i \quad \Rightarrow & x^j_i = 0\,, \end{array}$$

entonces el sistema $(x_1, \ldots, x_p)$ es libre. La demostración de este hecho es sencilla.

1.5.11 Ejemplo. El sistema (x_1, x_2, x_3) de $\mathbb{R}^4$, formado por

$$\begin{aligned} x_1 &= (2, 1, 3, \quad 0) \\ x_2 &= (0, 1, 3, \quad 4) \\ x_3 &= (0, 0, 4, -6) \end{aligned}$$

cumple las condiciones de (1.5.10) y es libre.

Obsérvese que, formando una columna con los vectores y sus coordenadas, la condición a) significa que los escalares de la 'diagonal' que pasa por x_1^1 son no nulos, y la condición b) significa que los escalares por debajo de esta 'diagonal' son todos nulos.

1.5.12 Sea E un e.v. y $(x_1, \ldots, x_p)$ un sistema de E. Cuando substituimos un vector x_i por un vector

$$x_i' = \sum_{j=1}^{p} \lambda^j x_j \,,$$

que es una combinación lineal de $(x_1, \ldots, x_p)$ con la condición de que $\lambda^i \neq 0$, los sistemas $(x_1, \ldots, x_i, \ldots, x_p)$ y $(x_1, \ldots, x_i', \ldots, x_p)$ generan el mismo subespacio. En particular esto sucede cuando substituimos x_i por $x_i' = \lambda x_i + \mu x_j$ con $i \neq j$ y $\lambda \neq 0$.

1.5.13 Los resultados (1.5.10) y (1.5.12) nos permiten calcular el rango de un sistema de vectores, como podemos ver en el ejemplo siguiente. Los determinantes, que estudiaremos en el capítulo 3, nos aportarán otro método diferente para la realización de este cálculo.

1.5.14 Ejemplo. Consideremos el sistema (x_1, x_2, x_3, x_4) de $\mathbb{R}^4$, formado por

$$\begin{aligned} x_1 &= (\ \ \ 2,\ 1,\ \ \ 3,\ 0\ ) \\ x_2 &= (\ -1,\ 0,\ \ \ 0,\ 2\ ) \\ x_3 &= (\ \ \ 3,\ 1,\ \ \ 1,\ 1\ ) \\ x_4 &= (\ \ \ 4,\ 1,\ -1,\ 2\ )\,, \end{aligned}$$

y sea $F = \langle x_1, x_2, x_3, x_4 \rangle$.

Haciendo $x_2' = x_1 + 2x_2$ tenemos (v. 1.5.12)

$$F = \langle x_1, x_2', x_3, x_4 \rangle \,.$$

Si ahora ponemos $x_3' = 3x_1 - 2x_3$,

$$F = \langle x_1, x_2', x_3', x_4 \rangle \,.$$

Finalmente, si hacemos $x_4' = -2x_1 + x_4$,

$$F = \langle x_1, x_2', x_3', x_4' \rangle \,.$$

con

$$\begin{aligned} x_1 &= (\ 2, \quad 1, \quad 3, \quad 0\ ) \\ x'_2 &= (\ 0, \quad 1, \quad 3, \quad 4\ ) \\ x'_3 &= (\ 0, \quad 1, \quad 7, -2\ ) \\ x'_4 &= (\ 0, -1, -7, \quad 2\ )\,. \end{aligned}$$

Es decir, hemos conseguido un sistema generador de F con ceros en las primeras coordenadas situadas debajo de la 'diagonal'.

Repetimos el proceso para la segunda columna de coordenadas haciendo ahora $x''_3 = -x'_2 + x'_3$ y $x''_4 = x'_2 + x'_4$ y obtenemos

$$F = \langle x_1, x'_2, x''_3, x''_4 \rangle$$

con

$$\begin{aligned} x_1 &= (\ 2, 1, \quad 3, \quad 0\ ) \\ x'_2 &= (\ 0, 1, \quad 3, \quad 4\ ) \\ x''_3 &= (\ 0, 0, \quad 4, -6\ ) \\ x''_4 &= (\ 0, 0, -4, \quad 6\ )\,. \end{aligned}$$

Por fin, una nueva fase, haciendo $x'''_4 = x''_3 + x''_4$, nos lleva a que

$$F = \langle x_1, x'_2, x''_3, x'''_4 \rangle$$

con

$$\begin{aligned} x_1 &= (\ 2, 1, 3, \quad 0\ ) \\ x'_2 &= (\ 0, 1, 3, \quad 4\ ) \\ x''_3 &= (\ 0, 0, 4, -6\ ) \\ x'''_4 &= (\ 0, 0, 0, \quad 0\ )\,, \end{aligned}$$

o sea,

$$F = \langle x_1, x'_2, x''_3 \rangle\,,$$

y como (x_1, x'_2, x''_3) es libre (v. 1.5.10 y 1.5.11), a causa de (1.5.2) es una base de F y resulta que

$$\mathrm{rg}(x_1, x_2, x_3, x_4) = \mathrm{rg}(x_1, x'_2, x''_3) = 3\,.$$

1.5.15 El método del ejemplo precedente permite además obtener una base para todo subespacio F del que se conozca un sistema generador (y las coordenadas de los vectores de dicho sistema en alguna base).

Capítulo 2

Aplicaciones lineales y matrices.

2.1 Propiedades de las aplicaciones lineales.

2.1.1 En el capítulo precedente definíamos lo que entendemos por aplicación lineal (véase 1.1.10) entre dos espacios vectoriales E y E' sobre el mismo cuerpo conmutativo $\mathbb{K}$ y resumíamos en una proposición (v. 1.1.12) algunas propiedades elementales de dichas aplicaciones. En la presente sección realizaremos un estudio más profundo de las aplicaciones lineales.

2.1.2 Si E y E' son dos e.v. sobre $\mathbb{K}$, representamos por $\mathcal{L}(E, E')$ el conjunto de las aplicaciones lineales de E en E'.

Por $\mathcal{L}(E)$ representamos el conjunto de los endomorfismos del e.v. E, o sea, $\mathcal{L}(E, E)$.

Estos conjuntos son no vacíos. Para el caso particular de $\mathcal{L}(E)$, la aplicación id_E (v. 1.1.11) es un elemento de $\mathcal{L}(E)$; en el ejemplo siguiente construiremos una aplicación que es un elemento de $\mathcal{L}(E, E')$.

2.1.3 **Ejemplo.** Si E y E' son dos e.v. sobre $\mathbb{K}$, la aplicación

$$\begin{aligned} E &\to E' \\ x &\to 0\,, \end{aligned}$$

que envía cada vector de E al cero de E', es una aplicación lineal; la representaremos por 0, o a veces por $0_{\mathcal{L}(E,E')}$. Tenemos que

$$(\forall x \in E) \quad 0(x) = 0\,.$$

2.1.4 $\mathcal{L}(E, E')$ posee, pues, al menos un elemento; en ciertos casos, éste es su único elemento.

2.1.5 Sean E y E' dos e.v. y supongamos que E es de dimensión $n \neq 0$; sea $(a_1, \ldots, a_n)$ una base de E. Si $f \in \mathcal{L}(E, E')$ y x es un vector de E, con

$$x = (x^1, \ldots, x^n)_{(a_i)}\,,$$

entonces

$$f(x) = f(x^1 a_1 + \cdots + x^n a_n) = x^1 f(a_1) + \cdots + x^n f(a_n)\,;$$

es decir, si conocemos las imágenes por f de los vectores de una base de E, la aplicación lineal f queda completamente determinada.

La proposición siguiente afirma que la elección de los vectores $f(a_1), \ldots, f(a_n)$ es arbitraria.

2.1.6 PROPOSICIÓN

Sea E un e.v. de dimensión $n \neq 0$ sobre $\mathbb{K}$, $(a_1, \ldots, a_n)$ una base de E y E' un e.v. sobre $\mathbb{K}$.

Si $y_1, \ldots, y_n$ son vectores cualesquiera de E', existe una aplicación lineal $f \in \mathcal{L}(E, E')$, única, que verifica

$$(\forall i = 1, \ldots, n) \quad f(a_i) = y_i\,.$$

Además, si E' es también de dimensión n y $(y_1, \ldots, y_n)$ es una base de E', la aplicación f es un isomorfismo.

Si para cada $x \in E$, con $x = (x^1, \ldots, x^n)_{(a_i)}$, ponemos

$$f(x) = x^1 y_1 + \cdots + x^n y_n\,,$$

f es una aplicación lineal de E en E'; se prueba sin dificultad que $f(a_i) = y_i$ para $i = 1, \ldots, n$.

Tal aplicación lineal es la única que cumple dicha propiedad, pues, si $g \in \mathcal{L}(E, E')$ y

$$(\forall i) \quad g(a_i) = y_i\,,$$

resulta que, cualquiera que sea $x \in E$ con $x = (x^1, \ldots, x^n)_{(a_i)}$,

$$g(x) = g(x^1 a_1 + \cdots + x^n a_n) = x^1 y_1 + \cdots + x^n y_n = f(x)\,,$$

o sea, que $g = f$.

Finalmente, si $(y_1, \ldots, y_n)$ es una base de E', entonces f es una aplicación compuesta

$$\begin{array}{ccccc} E & \to & \mathbb{K}^n & \to & E' \\ (x^1, \ldots, x^n)_{(a_i)} & \to & (x^1, \ldots, x^n) & \to & (x^1, \ldots, x^n)_{(y_i)} = x^1 y_1 + \cdots + x^n y_n \end{array}$$

de dos isomorfismos (v. 1.4.23), luego es un isomorfismo.

2.1.7 La proposición precedente nos permite construir así todas las aplicaciones lineales de E en E' cuando E es de dimensión finita.

2.1.8 Ejemplo. Sea (e_1, e_2) la base canónica de $\mathbb{R}^2$ y pongamos $v_1 = (3,0)$ y $v_2 = (1,1)$. La única aplicación $f \in \mathcal{L}(\mathbb{R}^2)$ que cumple $f(e_1) = v_1$ y $f(e_2) = v_2$ viene dada por

$$f(x,y) = f(xe_1 + ye_2) = xf(e_1) + yf(e_2) = x(3,0) + y(1,1) = (3x + y, y)\,.$$

Como (v_1, v_2) es una base de $\mathbb{R}^2$, f es un automorfismo.

2.1.9 PROPOSICIÓN

Sea $f \in \mathcal{L}(E, E')$.

a) Si F es un subespacio de E, entonces $f(F)$ es un subespacio de E'.

b) Si F' es un subespacio de E', entonces $f^{-1}(F')$ es un subespacio de E.

La demostración es sencilla.

2.1.10 DEFINICIÓN. Sea $f \in \mathcal{L}(E, E')$. La imagen de f, $\operatorname{Im} f$ (o también $f(E)$), es un subespacio de E'.

La imagen recíproca del subespacio $\{0_{E'}\}$ de E', $f^{-1}(\{0_{E'}\})$, que es el subconjunto de E formado por los vectores cuya imagen es $0_{E'}$, es un subespacio de E. Lo representamos por $\operatorname{Ker} f$ y decimos que es el *núcleo* de f. Se tiene, pues,

$$\operatorname{Ker} f = \{x \in E | f(x) = 0\}\,.$$

La aplicación f es sobre cuando $\operatorname{Im} f = E'$. La siguiente proposición caracteriza el caso en que f es inyectiva.

2.1.11 PROPOSICIÓN

Sea $f \in \mathcal{L}(E, E')$. La aplicación f es inyectiva si y sólo si

$$\operatorname{Ker} f = \{0_E\}\,.$$

Si f es inyectiva, la condición se verifica de forma evidente. Recíprocamente, supongamos que $\operatorname{Ker} f = \{0\}$; entonces, si $f(x) = f(y)$, resulta que $f(x - y) = 0$, o sea, $x - y \in \operatorname{Ker} f$ y $x - y = 0$, luego $x = y$; esto demuestra que f es inyectiva.

2.1.12 Ejemplo. Sea F un subespacio de un e.v. E. Para la inyección canónica

$$i : F \to E$$

(v. 1.2.10) tenemos

$$\operatorname{Im} i = F\,, \quad \operatorname{Ker} i = \{0\}\,.$$

Para la sobreyección canónica

$$j : E \to E/F$$

(v. 1.3.3) tenemos

$$\operatorname{Im} j = E/F\,, \quad \operatorname{Ker} j = F\,.$$

2.1.13 PROPOSICIÓN

Sea $f \in \mathcal{L}(E, E')$ y $(x_1, \ldots, x_p)$ un sistema de E. Tenemos que:

a) si $(x_1, \ldots, x_p)$ genera E, entonces $(f(x_1), \ldots, f(x_p))$ genera $\operatorname{Im} f$;

b) si $(x_1, \ldots, x_p)$ es ligado, entonces $(f(x_1), \ldots, f(x_p))$ es ligado;

c) si $(f(x_1), \ldots, f(x_p))$ es libre, entonces $(x_1, \ldots, x_p)$ es libre.

Cuando f es inyectiva, tenemos además que:

d) si $(x_1, \ldots, x_p)$ es libre, entonces $(f(x_1), \ldots, f(x_p))$ es libre;

e) si $(x_1, \ldots, x_p)$ es una base de E, entonces $(f(x_1), \ldots, f(x_p))$ es una base de $\operatorname{Im} f$.

Cuando f es sobre, tenemos además que:

f) si $(x_1, \ldots, x_p)$ genera E, entonces $(f(x_1), \ldots, f(x_p))$ genera E'.

Finalmente, cuando f es un isomorfismo, resulta que:

g) $(x_1, \ldots, x_p)$ es una base de E si y sólo si $(f(x_1), \ldots, f(x_p))$ es una base de E'.

Las demostraciones de estas propiedades son sencillas; para llevarlas a cabo se utilizan las técnicas empleadas en la demostración de (1.4.37); por lo tanto, las dejaremos como ejercicio para el lector.

2.1.14 PROPOSICIÓN

Sea $f \in \mathcal{L}(E, E')$ y supongamos que E es de dimensión finita; entonces $\operatorname{Im} f$ es de dimensión finita y además

$$\dim \operatorname{Im} f \leq \dim E\,.$$

Si $E = \{0\}$, entonces $\operatorname{Im} f = \{0\}$ y la proposición es evidente. Si $E \neq \{0\}$, sea $(a_1, \ldots, a_n)$ una base de E; como consecuencia del apartado a) de la proposición precedente, resulta que $(f(a_1), \ldots, f(a_n))$ genera $\operatorname{Im} f$, luego $\operatorname{Im} f$ es de dimensión finita y $\dim \operatorname{Im} f \leq n$, lo que demuestra la proposición en este caso.

2.1.15 Sea $f : E \to E'$ una aplicación lineal e inyectiva y F un subespacio de E. Podemos considerar la aplicación

$$\begin{aligned} F &\to f(F) \\ x &\to f(x)\,, \end{aligned}$$

que envía cada vector x de F al vector $f(x)$ del subespacio $f(F)$ de E'; esta aplicación suele denotarse también por f y es un isomorfismo entre los espacios vectoriales F y $f(F)$.

2.1.16 Si E es un e.v. de dimensión $n \neq 0$ y fijamos una base $(a_1, \ldots, a_n)$ de E, el isomorfismo

$$\begin{aligned} \mathbb{K}^n &\to E \\ (\lambda^1, \ldots, \lambda^n) &\to \lambda^1 a_1 + \cdots + \lambda^n a_n \end{aligned}$$

que vimos en (1.4.23) y su inverso,

$$\begin{aligned} E &\to \mathbb{K}^n \\ (\lambda^1, \ldots, \lambda^n)_{(a_i)} &\to (\lambda^1, \ldots, \lambda^n)\,, \end{aligned}$$

nos permiten estudiar algunos aspectos del espacio E trabajando con vectores de $\mathbb{K}^n$. Así, por ejemplo, si $(x_1, \ldots, x_p)$ es un sistema de E con

$$\begin{aligned} x_1 &= (x_1^1, \ldots, x_1^n)_{(a_i)} \\ &\cdots\cdots\cdots\cdots\cdots \\ x_p &= (x_p^1, \ldots, x_p^n)_{(a_i)} \end{aligned}$$

y consideramos los correspondientes vectores de $\mathbb{K}^n$,

$$\begin{aligned} v_1 &= (x_1^1, \ldots, x_1^n) \\ &\cdots\cdots\cdots\cdots\cdots \\ v_p &= (x_p^1, \ldots, x_p^n)\,, \end{aligned}$$

los sistemas $(x_1, \ldots, x_p)$ y $(v_1, \ldots, v_p)$ son ambos libres o ambos ligados (v. 2.1.13). Además, los vectores correspondientes (por el isomorfismo $\mathbb{K}^n \to E$) a una base de $\langle v_1, \ldots, v_p \rangle$ forman una base de $\langle x_1, \ldots, x_p \rangle$ (v. 2.1.15 y 2.1.13); en particular, ambos sistemas poseen el mismo rango.

2.1.17 Sea $f \in \mathcal{L}(E, E')$. Como $\operatorname{Ker} f$ es un subespacio de E, podemos considerar el espacio vectorial cociente $E/\operatorname{Ker} f$ (v. 1.3.3). Sean $x, y \in E$, con $\dot{x} = \dot{y}$; entonces $y - x \in \operatorname{Ker} f$ y $f(y - x) = 0$, es decir, $f(y) - f(x) = 0$ y $f(x) = f(y)$. Poniendo

$$\overline{f}(\dot{x}) = f(x)$$

para cada $\dot{x} \in E/\operatorname{Ker} f$, definimos sin ambigüedad una aplicación

$$\overline{f} : E/\operatorname{Ker} f \to \operatorname{Im} f\,.$$

Se prueba fácilmente que $\overline{f}$ es un isomorfismo de espacios vectoriales.

Además, considerando la inyección canónica (v. 1.2.10)

$$i\ : \operatorname{Im} f \to E'$$

y la sobreyección canónica (v. 1.3.3)

$$j : E \to E/\operatorname{Ker} f\,,$$

resulta que f es la aplicación compuesta de $i, \overline{f}$ y j :

$$E \xrightarrow{j} E/\operatorname{Ker} f \xrightarrow{\overline{f}} \operatorname{Im} f \xrightarrow{i} E'$$

es decir,

$$\boxed{f = i \circ \overline{f} \circ j}$$

Esta expresión recibe el nombre de *descomposición canónica* de la aplicación lineal f.

2.1.18 Se deduce de lo anterior una consecuencia importante. Si F es un suplementario de $\operatorname{Ker} f$, o sea, si

$$E = \operatorname{Ker} f \oplus F\ ,$$

entonces la aplicación

$$\begin{aligned} F &\to E/\operatorname{Ker} f \\ x &\to \dot{x} \end{aligned}$$

es un isomorfismo (v. 1.3.18). La aplicación compuesta de este isomorfismo y del isomorfismo

$$\begin{aligned} \overline{f} : E/\operatorname{Ker} f &\to \operatorname{Im} f \\ \dot{x} &\to f(x) \end{aligned}$$

es un isomorfismo y envía cada vector x de F a su imagen por f, $f(x)$. Resulta así que $\operatorname{Im} f$ es isomorfo a todo suplementario de $\operatorname{Ker} f$.

2.1.19 Cuando $f \in \mathcal{L}(E, E')$. y E es de dimensión finita, entonces $\operatorname{Ker} f$ admite un suplementario F (v. 1.5.7) y además

$$\dim E = \dim \operatorname{Ker} f + \dim F\ .$$

Como, según hemos visto, $\operatorname{Im} f \simeq F$, aplicando (1.4.38b) resulta lo siguiente:

2.1.20 PROPOSICIÓN

Sea $f \in \mathcal{L}(E, E')$ y E de dimensión finita. Entonces

$$\dim E = \dim \operatorname{Ker} f + \dim \operatorname{Im} f\ .$$

2.1.21 DEFINICIÓN. Sea $f \in \mathcal{L}(E, E')$ y E de dimensión finita. El número $\dim \operatorname{Im} f$ se denomina *rango* de f y se representa por $\operatorname{rg} f$.

Como vimos en (2.1.14), $\operatorname{rg} f \leq \dim E$ y, más exactamente,

$$\operatorname{rg} f = \dim E - \dim \operatorname{Ker} f\,.$$

Recordando la definición de rango de un sistema de vectores (v. 1.5.9) y el apartado a) de la proposición (2.1.13) resulta que, si $(x_1, \ldots, x_p)$ genera E, entonces

$$\operatorname{rg} f = \operatorname{rg}(f(x_1), \ldots, f(x_p))\,.$$

Cuando E' es de dimensión finita, como $\operatorname{Im} f$ es un subespacio de E', tenemos también que

$$\operatorname{rg} f \leq \dim E'\,.$$

2.1.22 Ejemplo. La aplicación $f \in \mathcal{L}(\mathbb{R}^3)$ dada por

$$f(x, y, z) = (-x + z, y + 2z, 2x + y)$$

tiene rango 2, puesto que, si (e_1, e_2, e_3) es la base canónica de $\mathbb{R}^3$,

$$\begin{aligned} f(e_1) &= (\ -1,\ 0,\ 2) \\ f(e_2) &= (\ \ \ 0,\ 1,\ 1) \\ f(e_3) &= (\ \ \ 1,\ 2,\ 0) \end{aligned}$$

y $\operatorname{rg}(f(e_1), f(e_2), f(e_3)) = 2$.

2.1.23 PROPOSICIÓN

Sea $f \in \mathcal{L}(E, E')$, con E de dimensión n. Entonces f es inyectiva si y sólo si $\operatorname{rg} f = n$. Si, además, E' es de dimensión m, entonces f es sobre si y sólo si $\operatorname{rg} f = m$.

f es inyectiva cuando $\operatorname{Ker} f = \{0\}$, o sea, cuando $\dim \operatorname{Ker} f = 0$, y esto significa que $\operatorname{rg} f = \dim E - \dim \operatorname{Ker} f = n$. Por otra parte, f es sobre cuando $\operatorname{Im} f = E'$, o sea, cuando $\dim \operatorname{Im} f = \dim E' = m$.

2.1.24 COROLARIO

Sea $f \in \mathcal{L}(E, E')$ con $\dim E = \dim E' = n$. Entonces son equivalentes las siguientes propiedades:

(i) f es un isomorfismo,
(ii) f es sobre,
(iii) f es inyectiva,
(iv) $\operatorname{rg} f = n$.

Basta observar que, como consecuencia de la proposición precedente, cada una de las tres primeras propiedades equivale a la cuarta.

2.1.25 Sea $f \in \mathcal{L}(E, E')$ con $\dim E = n \neq 0$ y sea $(a_1, \ldots, a_n)$ una base de E. El sistema $(f(a_1), \ldots, f(a_n))$ de $\operatorname{Im} f$ genera $\operatorname{Im} f$ (v. 2.1.13), pero no es necesariamente una base.

Si $\operatorname{rg} f = n$, entonces $(f(a_1), \ldots, f(a_n))$ sí es una base de $\operatorname{Im} f$.

Si $\operatorname{rg} f = 0$, entonces $\operatorname{Im} f = \{0\}$ (es decir, $f = 0$).

La proposición siguiente proporciona un método de obtención de una base de $\operatorname{Im} f$ para el caso de $0 < \operatorname{rg} f < n$.

2.1.26 PROPOSICIÓN

Sea $f \in \mathcal{L}(E, E')$ con $\dim E = n \neq 0$ y $\operatorname{rg} f = r$ tal que

$$0 < r < n\,.$$

La dimensión de $\operatorname{Ker} f$ es $n - r$ (v. 2.1.21); sea $(a_{r+1}, \ldots, a_n)$ una base de $\operatorname{Ker} f$ y $(a_1, \ldots, a_r, a_{r+1}, \ldots, a_n)$ una base de E que prolonga la de $\operatorname{Ker} f$ (v. 1.5.6). Entonces $(f(a_1), \ldots, f(a_r))$ es una base de $\operatorname{Im} f$.

Como $\dim \operatorname{Im} f = r$, bastará probar que $(f(a_1), \ldots, f(a_r))$ es un sistema libre (v. 1.4.36). Si $\lambda^1, \ldots, \lambda^r$ son escalares tales que

$$\lambda^1 f(a_1) + \cdots + \lambda^r f(a_r) = 0\,,$$

entonces

$$\lambda^1 a_1 + \cdots + \lambda^r a_r \in \operatorname{Ker} f\,,$$

luego existen escalares $\lambda^{r+1}, \ldots, \lambda^n$ tales que

$$\lambda^1 a_1 + \cdots + \lambda^r a_r = \lambda^{r+1} a_{r+1} + \cdots + \lambda^n a_n\,,$$

o sea

$$\lambda^1 a_1 + \cdots + \lambda^r a_r - \lambda^{r+1} a_{r+1} - \cdots - \lambda^n a_n = 0\,,$$

y, como $(a_1, \ldots, a_r, a_{r+1}, \ldots, a_n)$ es libre, esto implica que

$$\lambda^1 = \cdots = \lambda^r = \lambda^{r+1} = \cdots = \lambda^n = 0$$

y resulta así que $\lambda^1 = \cdots = \lambda^r = 0$ y que $(f(a_1), \ldots, f(a_r))$ es libre.

2.1.27 Ejemplo. Para la aplicación $f \in \mathcal{L}(\mathbb{R}^3)$ descrita en el ejemplo (2.1.22), $\dim \operatorname{Ker} f = 1$. Como $(1, -2, 1) \in \operatorname{Ker} f$, este vector constituye una base de $\operatorname{Ker} f$. Además

$$((1,0,0), (0,1,0), (1,-2,1))$$

es una base de $\mathbb{R}^3$ que la prolonga. Como

$$f(1,0,0) = (-1,0,2) \quad \text{y} \quad f(0,1,0) = (0,1,1)\,,$$

resulta que

$$((-1,0,2), (0,1,1))$$

es una base de $\operatorname{Im} f$.

2.2 Matrices. Matriz de una aplicación lineal.

2.2.1 Las estructuras rectangulares de números de la forma[1]

$$\begin{bmatrix} -1. & 2.3 \\ 1.5 & 0 \end{bmatrix} \qquad \begin{bmatrix} 0 & -1.2 \\ 1. & 5. \\ -0.3 & -1.7 \end{bmatrix},$$

reciben el nombre de 'matrices'. Para nosotros estas estructuras serán de gran interés, como veremos en lo que sigue. Las matrices son una herramienta esencial en el dominio de las aplicaciones prácticas, tanto del álgebra lineal como de otras muchas ramas de la matemática.

Comencemos por establecer la definición formal de matriz.

2.2.2 DEFINICIÓN. Sea X un conjunto y $n, m \in \mathbb{N}$, $n \neq 0, m \neq 0$. Una *matriz* $n \times m$, o bien una *matriz de* n *columnas y* m *filas* (con elementos en X), es una aplicación

$$A : \{1, \ldots, n\} \times \{1, \ldots, m\} \to X$$

que a cada pareja de números naturales $(i\,, j)$, con $1 \leq i \leq n$ y $1 \leq j \leq m$, hace corresponder un elemento de X, que se suele representar por a_i^j . (Son también usuales otras representaciones similares, pero ésta es la más acorde con nuestros propósitos.) La matriz A se representa entonces por $[a_i^j]_{1\leq i\leq n,\, 1\leq j\leq m}$, o simplemente por $[a_i^j]$. También se representa mediante un rectángulo en el que se disponen los elementos a_i^j como sigue:

$$\begin{bmatrix} a_1^1 & a_2^1 & \cdots & a_i^1 & \cdots & a_n^1 \\ a_1^2 & a_2^2 & \cdots & a_i^2 & \cdots & a_n^2 \\ \vdots & \vdots & & \vdots & & \vdots \\ a_1^j & a_2^j & \cdots & a_i^j & \cdots & a_n^j \\ \vdots & \vdots & & \vdots & & \vdots \\ a_1^m & a_2^m & \cdots & a_i^m & \cdots & a_n^m \end{bmatrix}$$

De acuerdo con esta representación, decimos que a_i^j es el elemento de la columna i y la fila j; por la misma razón, llamamos índice de columna de a_i^j al número i,

[1] El símbolo 2.3 representa el número real cuya expresión decimal es $2'3$, o sea, $2 + (3 \times 10^{-1})$. Hemos optado por utilizar el 'punto decimal' en lugar de la coma porque se presta menos a confusión en expresiones del tipo

$$(2.3, 1.1, 1.5)$$

o similares. No resultará extraño al lector, acostumbrado sin duda a leer expresiones con punto decimal en la pantalla de las pequeñas máquinas calculadoras.

Para los distintos ejemplos con números reales, hubiéramos podido limitarnos a elegir casi siempre números enteros, teniendo en cuenta que, en estos capítulos, lo importante es la exposición de los conceptos y no los métodos de cálculo. Hemos preferido, sin embargo, utilizar algunos números reales no enteros (pero siempre con expresión decimal finita y corta) para proporcionar una mejor ilusión óptica de número real.

La excepción es siempre el 0 (y a veces el 1), porque en general se necesita que 0 sea el cero de $\mathbb{R}$ y no algún número 'suficientemente próximo' a 0. También hemos omitido el punto decimal cuando, en caso de incluirlo, las expresiones hubieran podido perder claridad.

que figura en posición inferior, e índice de fila al número j, que figura en posición superior.

2.2.3 Ejemplo. La matriz

$$A = \begin{bmatrix} 0 & -1.2 \\ 1. & 5. \\ -0.3 & -1.7 \end{bmatrix}$$

es la aplicación

$$A : \{1, 2\} \times \{1, 2, 3\} \to \mathbb{R}$$

dada por

$$a_1^1 = 0,\ a_1^2 = 1.,\ a_1^3 = -0.3,\ a_2^1 = -1.2,\ a_2^2 = 5.,\ a_2^3 = -1.7\,.$$

2.2.4 El conjunto de las matrices $n \times m$ con elementos en X se representa por $M_X(n, m)$, o por $M(n, m)$ si no hay posibilidad de confusión.

El conjunto $M(n, n)$ se representa por $M(n)$; sus elementos se llaman *matrices cuadradas* (de orden n).

Las matrices de $M(1, m)$ son las *matrices columna*; las de $M(n, 1)$ son las *matrices fila*.

Si $A = [a_i^j] \in M(n, m)$, entonces decimos que la matriz columna

$$\begin{bmatrix} a_i^1 \\ \vdots \\ a_i^m \end{bmatrix}$$

es la columna número i de A, y que la matriz fila

$$\begin{bmatrix} a_1^j & \dots & a_n^j \end{bmatrix}$$

es la fila número j de A.

2.2.5 Si $A = [a_i^j] \in M(n)$ es una matriz cuadrada, los n elementos a_i^i, $i = 1, \dots, n$, reciben el nombre de *elementos diagonales* de A, y el sistema $(a_1^1, \dots, a_n^n)$ de X se llama *diagonal* de A.

2.2.6 DEFINICIÓN. Si $A = [a_i^j] \in M(n, m)$, la matriz de $M(m, n)$ que representamos por $A^t = [{}^t a_i^j]$ y cuyo elemento ${}^t a_i^j$ (de la columna i y la fila j) es

$${}^t a_i^j = a_j^i$$

recibe el nombre de *traspuesta* de A. Se suele decir que A^t se obtiene cambiando en A las filas por columnas y viceversa.

La aplicación

$$\begin{aligned} M(n, m) &\to M(m, n) \\ A &\to A^t \end{aligned}$$

es una aplicación biyectiva. La inversa de esta aplicación es

$$\begin{aligned} M(m,n) &\to M(n,m) \\ A &\to A^t \end{aligned}$$

puesto que

$$(A^t)^t = A\,.$$

Si $A \in M(n)$ es una matriz cuadrada, su traspuesta, A^t, es también de $M(n)$. En este caso, A y A^t tienen la misma diagonal. Cuando

$$A^t = A$$

decimos que A es *simétrica*. Obsérvese que esto significa que

$$(\forall i\,, j) \quad a_i^j = a_j^i\,.$$

2.2.7 Ejemplo. Si

$$A = \begin{bmatrix} 0 & -1.2 \\ 1. & 5. \\ -0.3 & -1.7 \end{bmatrix},$$

entonces

$$A^t = \begin{bmatrix} 0 & 1. & -0.3 \\ -1.2 & 5. & -1.7 \end{bmatrix}.$$

2.2.8 En adelante, supondremos que las matrices tienen sus elementos en un cuerpo conmutativo $\mathbb{K}$ (o sea, que $X = \mathbb{K}$), ya que este es el caso en que nos serán más útiles. Es decir, trabajaremos prácticamente siempre con matrices formadas por escalares.

Decimos que una matriz cuadrada $A = [a_i^j] \in M(n)$ es *diagonal* cuando

$$i \neq j \quad \Rightarrow \quad a_i^j = 0 \quad (\text{el } 0 \text{ de } \mathbb{K})\,,$$

es decir, cuando son nulos todos los elementos no diagonales de A.

Una matriz diagonal es siempre simétrica.

2.2.9 Ejemplo. La matriz

$$A = \begin{bmatrix} 0.2 & 0 & 0 \\ 0 & -1. & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

de $M_{\mathbb{R}}(3)$ es diagonal.

2.2.10 Decimos que una matriz $A = [a_i^j] \in M(n)$ es *triangular inferior* cuando

$$i > j \quad \Rightarrow \quad a_i^j = 0\,,$$

o sea, cuando son nulos todos los elementos situados 'encima' de la diagonal de A, y decimos que es *triangular superior* cuando

$$j > i \quad \Rightarrow \quad a_i^j = 0\,,$$

o sea, cuando son nulos todos los elementos situados 'debajo' de la diagonal de A. Decimos que A es *triangular* cuando es triangular inferior o triangular superior.

Si A es triangular inferior (respectivamente, superior) entonces A^t es triangular superior (resp., inferior).

2.2.11 Ejemplo. La matriz

$$A = \begin{bmatrix} 0.2 & 0 & 0 \\ 0.5 & -1. & 0 \\ -7.2 & 1.3 & 0 \end{bmatrix}$$

es triangular inferior; su traspuesta

$$A^t = \begin{bmatrix} 0.2 & 0.5 & -7.2 \\ 0 & -1. & 1.3 \\ 0 & 0 & 0 \end{bmatrix}$$

es triangular superior.

En el caso de matrices triangulares o diagonales, es frecuente no escribir los ceros de la matriz, o al menos parte de ellos, siempre que esto no suponga crear una situación ambigua o mal definida. Las dos matrices de este ejemplo y la de (2.2.9) se pueden representar así por

$$\begin{bmatrix} 0.2 & & \\ 0.5 & -1. & \\ -7.2 & 1.3 & 0 \end{bmatrix}, \begin{bmatrix} 0.2 & 0.5 & -7.2 \\ & -1. & 1.3 \\ & & 0 \end{bmatrix}, \begin{bmatrix} 0.2 & & \\ & -1. & \\ & & 0 \end{bmatrix}.$$

2.2.12 Si $A = [a_i^j] \in M(n, m)$, su columna

$$\begin{bmatrix} a_i^1 \\ \vdots \\ a_i^m \end{bmatrix}$$

puede ser identificada con el vector

$$(a_i^1, \ldots, a_i^m)$$

de $\mathbb{K}^m$, y lo mismo podemos hacer con una fila

$$\begin{bmatrix} a_1^j & \ldots & a_n^j \end{bmatrix}$$

de A y el correspondiente vector de $\mathbb{K}^n$. Hablaremos entonces de los vectores columna de A (vectores de $\mathbb{K}^m$) y los vectores fila de A (vectores de $\mathbb{K}^n$).

Los vectores columna de A (resp., vectores fila) son los vectores fila de A^t (resp., vectores columna).

2.2.13 Si $x = (x^1, \ldots, x^n)$ es un vector de $\mathbb{K}^n$, identificamos x con la matriz columna

$$\begin{bmatrix} x^1 \\ \vdots \\ x^n \end{bmatrix}$$

de $M(1, n)$.

2.2.14 También identificamos cada elemento λ de $\mathbb{K}$ con la matriz $[\lambda] \in M(1, 1)$.

2.2.15 Sean $p, q, r, s \in \mathbb{N}$ y $n = p + q$, $m = r + s$. Consideremos las matrices $A \in M(p, r)$, $B \in M(q, r)$, $C \in M(p, s)$ y $D \in M(q, s)$ siguientes:

$$A = \begin{bmatrix} a_1^1 & \cdots & a_p^1 \\ \vdots & & \vdots \\ a_1^r & \cdots & a_p^r \end{bmatrix} \quad B = \begin{bmatrix} b_1^1 & \cdots & b_q^1 \\ \vdots & & \vdots \\ b_1^r & \cdots & b_q^r \end{bmatrix}$$

$$C = \begin{bmatrix} c_1^1 & \cdots & c_p^1 \\ \vdots & & \vdots \\ c_1^s & \cdots & c_p^s \end{bmatrix} \quad D = \begin{bmatrix} d_1^1 & \cdots & d_q^1 \\ \vdots & & \vdots \\ d_1^s & \cdots & d_q^s \end{bmatrix}.$$

La matriz $M \in M(n, m)$ dada por

$$M = \begin{bmatrix} a_1^1 & \cdots & a_p^1 & b_1^1 & \cdots & b_q^1 \\ \vdots & & \vdots & \vdots & & \vdots \\ a_1^r & \cdots & a_p^r & b_1^r & \cdots & b_q^r \\ c_1^1 & \cdots & c_p^1 & d_1^1 & \cdots & d_q^1 \\ \vdots & & \vdots & \vdots & & \vdots \\ c_1^s & \cdots & c_p^s & d_1^s & \cdots & d_q^s \end{bmatrix}$$

se representa por

$$M = \left[\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right]$$

y de ella se suele decir que está 'definida por bloques'.

2.2.16 Ejemplo. Si

$$A = \begin{bmatrix} 3. \end{bmatrix} \quad B = \begin{bmatrix} 0 & -1. \end{bmatrix}$$

$$C = \begin{bmatrix} 1. \\ 1.2 \end{bmatrix} \quad D = \begin{bmatrix} 0 & 5. \\ -1. & 0 \end{bmatrix},$$

la matriz

$$M = \left[\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right]$$

es

$$M = \begin{bmatrix} 3. & 0 & -1. \\ 1. & 0 & 5. \\ 1.2 & -1. & 0 \end{bmatrix}.$$

2.2.17 DEFINICIÓN. Sean E y E' espacios vectoriales de dimensión finita, $(a_1, \ldots, a_n)$ una base de E, $(b_1, \ldots, b_m)$ una base de E' y $f \in \mathcal{L}(E, E')$. Llamamos *matriz asociada a f y a las bases $(a_1, \ldots, a_n)$ y $(b_1, \ldots, b_m)$* a la matriz

$$[\alpha_i^j] \in M(n, m)$$

cuyo elemento general α_i^j es la coordenada número j del vector $f(a_i)$ en la base $(b_1, \ldots, b_m)$. Representamos esta matriz por

$$[f, (a_1, \ldots, a_n), (b_1, \ldots, b_m)]$$

o abreviadamente, por

$$[f, (a_i), (b_j)]\,.$$

Es decir, si

$$f(a_i) = (\alpha_i^1, \ldots, \alpha_i^m)_{(b_1, \ldots, b_m)}\,,$$

entonces

$$\begin{bmatrix} \alpha_i^1 \\ \vdots \\ \alpha_i^m \end{bmatrix}$$

es la columna número i de $[f, (a_i), (b_j)]$.

Si E es un e.v. de dimensión finita, $(a_1, \ldots, a_n)$ y $(b_1, \ldots, b_n)$ bases de E y $f \in \mathcal{L}(E)$, la correspondiente matriz $[f, (a_i), (b_j)]$ no es apenas utilizada. En cambio, tiene gran importancia la matriz $[f, (a_i), (a_i)]$ asociada a f cuando las bases (a_i) y (b_j) son la misma; representamos esta matriz por

$$[f, (a_i)]\,.$$

Observemos que si

$$\begin{bmatrix} \alpha_i^1 \\ \vdots \\ \alpha_i^n \end{bmatrix}$$

es la columna número i de esta matriz, entonces

$$f(a_i) = (\alpha_i^1, \ldots, \alpha_i^n)_{(a_1, \ldots, a_n)}\,.$$

2.2.18 Ejemplo. Para la aplicación

$$f : \mathbb{R}^3 \to \mathbb{R}^2 ,$$

dada por

$$f(x,y,z) = (2x+z, y+z) ,$$

y las bases canónicas (e_1, e_2, e_3) y $(\bar{e}_1, \bar{e}_2)$ de dichos espacios, tenemos

$$[f, (e_i), (\bar{e}_j)] = \begin{bmatrix} 2. & 0 & 1. \\ 0 & 1. & 1. \end{bmatrix}$$

puesto que

$$\begin{aligned} f(e_1) &= (2,0) = 2\bar{e}_1 , \\ f(e_2) &= (0,1) = \bar{e}_2 , \\ f(e_3) &= (1,1) = \bar{e}_1 + \bar{e}_2 . \end{aligned}$$

2.2.19 Ejemplo. Para la aplicación

$$f : \mathbb{R}^3 \to \mathbb{R}^3$$

del ejemplo (2.1.22) y la base canónica (e_1, e_2, e_3), tenemos

$$[f, (e_i)] = \begin{bmatrix} -1. & 0 & 1. \\ 0 & 1. & 2. \\ 2. & 1. & 0 \end{bmatrix} .$$

2.2.20 La asociación de matrices a las aplicaciones lineales resulta de gran importancia práctica, como tendremos ocasión de observar en éste y en sucesivos capítulos. Antes, sin embargo, veremos en la sección siguiente que la correspondencia entre aplicaciones lineales y matrices posee un 'buen comportamiento' respecto de las estructuras algebraicas.

2.3 Los espacios vectoriales $\mathcal{L}(E,E')$ y $M(n,m)$.

2.3.1 Sean E y E' dos espacios vectoriales sobre un cuerpo conmutativo $\mathbb{K}$. Si en el conjunto $\mathcal{F}(E,E')$ de todas las aplicaciones de E en E' consideramos las operaciones $f+g$ (suma) y λf (producto por un elemento de $\mathbb{K}$), definidas para $f, g \in \mathcal{F}(E,E')$ y $\lambda \in \mathbb{K}$ por

$$(f+g)(x) = f(x) + g(x) ,$$

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

tenemos en dicho conjunto una estructura de espacio vectorial sobre $\mathbb{K}$ (v. 1.1.3). El subconjunto $\mathcal{L}(E,E')$ de $\mathcal{F}(E,E')$, formado por las aplicaciones de $\mathcal{F}(E,E')$ que son lineales, es un subespacio de $\mathcal{F}(E,E')$, como se prueba de forma sencilla (recuérdese que en 2.1.3 vimos que $\mathcal{L}(E,E')$ es no vacío).

Es decir, con las operaciones mencionadas, $\mathcal{L}(E, E')$ es un espacio vectorial sobre $\mathbb{K}$. El cero de este e.v. es la aplicación 0 de (2.1.3); si $f \in \mathcal{L}(E, E')$, su opuesto para la suma es la aplicación $-f$, definida para todo $x \in E$ por

$$(-f)(x) = -f(x) .$$

2.3.2 Consideremos el conjunto $M(n, m)$ de las matrices $n \times m$ con elementos en $\mathbb{K}$. Observemos que

$$M(n, m) = \mathcal{F}(\{1, \ldots, n\} \times \{1, \ldots, m\}, \mathbb{K}) ,$$

o sea, que $M(n, m)$ es el conjunto de todas las aplicaciones de $\{1, \ldots, n\} \times \{1, \ldots, m\}$ en $\mathbb{K}$. Sabemos que este conjunto posee una estructura de espacio vectorial sobre $\mathbb{K}$ con las operaciones que ya vimos en (1.1.3). Para la notación habitual de las matrices resunta que, si $A, B \in M(n, m)$ y $A = [a_i^j]$ y $B = [b_i^j]$, entonces $A + B$ es la matrix de $M(n, m)$ dada por

$$A + B = [a_i^j + b_i^j]$$

y, si $\lambda \in \mathbb{K}$, entonces λA es la matrix de $M(n, m)$ dada por

$$\lambda A = [\lambda a_i^j] .$$

El cero del e.v. $M(n, m)$ es la matriz, representada por 0 (o por $0_{M(n,m)}$), cuyos elementos son todos el 0 de $\mathbb{K}$. Si $A = [a_i^j] \in M(n, m)$, la matriz $-A$, opuesta de A, es

$$-A = [-a_i^j] .$$

2.3.3 Ejemplo.

$$\begin{bmatrix} 2.1 & -1. & 1.1 \\ 0 & 0.5 & 3. \end{bmatrix} + \begin{bmatrix} 0 & 2.7 & 0 \\ -1.2 & 0.5 & 0 \end{bmatrix} = \begin{bmatrix} 2.1 & 1.7 & 1.1 \\ -1.2 & 1. & 3. \end{bmatrix} ,$$

$$-2. \begin{bmatrix} 2.1 & -1. & 1.1 \\ 0 & 0.5 & 3. \end{bmatrix} = \begin{bmatrix} -4.2 & 2. & -2.2 \\ 0 & -1. & -6. \end{bmatrix} .$$

2.3.4 La aplicación biyectiva

$$\begin{aligned} M(n, m) &\to M(m, m) \\ A &\to A^t \end{aligned}$$

de (2.2.6) es un isomorfismo, puesto que

$$\begin{aligned} (A + B)^t &= A^t + B^t \\ (\lambda A)^t &= \lambda A^t , \end{aligned}$$

como se comprueba con facilidad.

2.3.5 Vamos a describir una base del e.v. $M(n, m)$. Siendo $1 \leq i \leq n$ y $1 \leq j \leq m$, representamos por E_i^j la matriz $n \times m$ cuyos elementos son todos nulos salvo el de la columna número i y la fila j, que vale 1. Si consideramos elementos α_i^j de $\mathbb{K}$, para $1 \leq i \leq n$ y $1 \leq j \leq m$, resulta que la combinación lineal

$$\sum_{i,j} \alpha_i^j E_i^j$$

es la matriz $A = [\alpha_i^j]$ de $M(n, m)$. Teniendo en cuenta este hecho, se prueba sin dificultad que el sistema de $n \times m$ matrices constituido por las matrices E_i^j es una base de $M(n, m)$; decimos que esta base es la 'base canónica' de $M(n, m)$. Resulta entonces que

$$\dim M(n, m) = n \times m \,.$$

2.3.6 Ejemplo. La base canónica de $M(2, 3)$ es el sistema de 6 elementos formado por las matrices

$$E_1^1 = \begin{bmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \qquad E_2^1 = \begin{bmatrix} 0 & 1 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}$$

$$E_1^2 = \begin{bmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 0 \end{bmatrix} \qquad E_2^2 = \begin{bmatrix} 0 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}$$

$$E_1^3 = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 1 & 0 \end{bmatrix} \qquad E_2^3 = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 1 \end{bmatrix} \,.$$

2.3.7 Sean E y E' dos espacios vectoriales de dimensiones $n \neq 0$ y $m \neq 0$, y sean $(a_1, \ldots, a_n)$ y $(b_1, \ldots, b_m)$ bases de E y E'.

Sea $f \in \mathcal{L}(E, E')$ y consideremos la matriz $[f, (a_i), (b_j)]$ de $M(n, m)$, definida en (2.2.17). Tendremos así una aplicación

$$\begin{aligned} \mathcal{L}(E, E') &\to M(n, m) \\ f &\to [f, (a_i), (b_j)] \,, \end{aligned}$$

de la que hacemos notar, en primer lugar, que depende de las bases que hemos fijado en E y E'.

La aplicación es inyectiva, pues, si

$$[f, (a_i), (b_j)] = [g, (a_i), (b_j)] \,,$$

aplicando la definición (2.2.17) resulta que

$$(\forall i = 1, \ldots, n) \quad f(a_i) = g(a_i)$$

y, utilizando (2.1.6), tenemos que $f = g$.

La aplicación es también sobre; en efecto, si $A = [\alpha_i^j] \in M(n, m)$, existe una única aplicación lineal $f \in \mathcal{L}(E, E')$ que verifica

$$(\forall i) \quad f(a_i) = (\alpha_i^1, \ldots, \alpha_i^m)_{(b_j)}$$

(v. 2.1.6), y para esta aplicación

$$[f, (a_i), (b_j)] = A\,.$$

A causa de la inyectividad, f es la única aplicación lineal cuya matriz asociada (para las bases fijadas) es A. En la sección 2.5 veremos una formulación más explícita de f.

La aplicación $f \to [f, (a_i), (b_j)]$ es además una aplicación lineal, o sea,

$$[f + g, (a_i), (b_j)] = [f, (a_i), (b_j)] + [g, (a_i), (b_j)]$$

$$[\lambda f, (a_i), (b_j)] = \lambda[f, (a_i), (b_j)]\,.$$

En efecto; como

$$(f + g)(a_i) = f(a_i) + g(a_i)\,,$$

poniendo $[f, (a_i), (b_j)] = [\alpha_i^j]$ y $[g, (a_i), (b_j)] = [\beta_i^j]$ resulta que

$$f(a_i) = (\alpha_i^1, \ldots, \alpha_i^m)_{(b_j)} \quad \text{y} \quad g(a_i) = (\beta_i^1, \ldots, \beta_i^m)_{(b_j)}\,,$$

luego

$$(f + g)(a_i) = (\alpha_i^1 + \beta_i^1, \ldots, \alpha_i^m + \beta_i^m)_{(b_j)}$$

y el elemento de la columna i y la fila j de $[f + g, (a_i), (b_j)]$ es

$$\alpha_i^j + \beta_i^j\,,$$

lo que demuestra la primera propiedad. La segunda se prueba de forma análoga.

Estos hechos se resumen en la proposición siguiente.

2.3.8 PROPOSICIÓN

Sean E y E' dos e.v. de dimensiones $n \neq 0$ y $m \neq 0$, y sean $(a_1, \ldots, a_n)$ y $(b_1, \ldots, b_m)$ bases de E y E'. La aplicación

$$\begin{aligned} \mathcal{L}(E, E') &\to M(n, m) \\ f &\to [f, (a_i), (b_j)] \end{aligned}$$

es un isomorfismo de espacios vectoriales.

2.3.9 De esta proposición y de (1.1.12) resulta que

$$[0_{\mathcal{L}(E,E')}, (a_i), (b_j)] = 0_{M(n,m)}$$

y que

$$[-f, (a_i), (b_j)] = -[f, (a_i), (b_j)] .$$

2.3.10 También resulta de esta proposición y de (1.4.38) que, si E y E' son e.v. de dimensión finita, entonces

$$\dim \mathcal{L}(E, E') = \dim E \times \dim E' .$$

2.4 Los anillos $\mathcal{L}(E)$ y $M(n)$. Matrices inversibles.

2.4.1 Si $f \in \mathcal{L}(E, E')$ y $g \in \mathcal{L}(E', E'')$, entonces $g \circ f \in \mathcal{L}(E, E'')$ (v. 1.1.10).
La composición de aplicaciones es asociativa; en particular resulta que:
a) si $f \in \mathcal{L}(E, E')$, $g \in \mathcal{L}(E', E'')$ y $h \in \mathcal{L}(E'', E''')$, entonces

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

Se demuestra fácilmente que la composición de aplicaciones lineales es distributiva (a ambos lados) respecto de la suma, o sea, que:
b) si $f_1, f_2 \in \mathcal{L}(E, E')$, y $g \in \mathcal{L}(E', E'')$, entonces

$$g \circ (f_1 + f_2) = (g \circ f_1) + (g \circ f_2) ;$$

c) si $f \in \mathcal{L}(E, E')$, y $g_1, g_2 \in \mathcal{L}(E', E'')$, entonces

$$(g_1 + g_2) \circ f = (g_1 \circ f) + (g_2 \circ f) .$$

2.4.2 Sea E un e.v. El conjunto $\mathcal{L}(E)$ de los endomorfismos de E es un e.v. con la suma y el producto por un escalar considerados en (2.3.1). En particular, $(\mathcal{L}(E), +)$ es un grupo conmutativo (v. 1.1.1).

Si $f, g \in \mathcal{L}(E)$, entonces $g \circ f \in \mathcal{L}(E)$; es decir, la composición de aplicaciones es una operación interna de $\mathcal{L}(E)$. Utilizando en esta situación particular las propiedades a), b) y c) del párrafo precedente, resulta que $(\mathcal{L}(E), +, \circ)$ es un anillo.

Por otra parte, el endomorfismo id_E de (1.1.11) verifica que

$$(\forall f \in \mathcal{L}(E)) \quad f \circ id_E = id_E \circ f = f$$

y es la unidad del anillo $\mathcal{L}(E)$; se trata, pues, de un anillo unitario.

2.4.3 El anillo $\mathcal{L}(E)$ no es, en general, conmutativo, como muestra el ejemplo siguiente.

2.4.4 Ejemplo. Consideremos los endomorfismos f y g de $\mathbb{R}^2$ dados por

$$f(x,y) = (x,0) \quad \text{y} \quad g(x,y) = (0,x)\,.$$

Se tiene

$$(g \circ f)(x,y) = g(x,0) = (0,x)\,,$$
$$(f \circ g)(x,y) = f(0,x) = (0,0)\,,$$

y basta con hacer $(x,y) = (1,0)$ para demostrar que

$$g \circ f \neq f \circ g\,.$$

Además, $f \neq 0$, $g \neq 0$ y $f \circ g = 0$.

2.4.5 En el ejemplo precedente hemos comprobado también que en general $\mathcal{L}(E)$ no es un anillo íntegro, es decir, que en $\mathcal{L}(E)$ puede haber divisores de 0.

2.4.6 Tratemos de ver cuáles son los elementos inversibles del anillo $\mathcal{L}(E)$. Un elemento $f \in \mathcal{L}(E)$ es inversible cuando

$$f \circ g = g \circ f = id_E$$

para algún $g \in \mathcal{L}(E)$ (necesariamente único).

Si f es inversible, resulta que f es una aplicación biyectiva, o sea, un automorfismo de E. Recíprocamente, todo automorfismo f de E es inversible, pues la aplicación inversa f^{-1} es un automorfismo de E y

$$f \circ f^{-1} = f^{-1} \circ f = id_E\,.$$

Los elementos inversibles de $\mathcal{L}(E)$ son entonces los automorfismos de E; el conjunto de estos automorfismos se representa por $GL(E)$. Es fácil ver que $(GL(E), \circ)$ es un grupo; se le suele llamar 'grupo lineal de E'.

Si $f, g \in GL(E)$, entonces

$$(g \circ f)^{-1} = f^{-1} \circ g^{-1}\,.$$

2.4.7 Enunciemos otra propiedad, de demostración sencilla, que relaciona el producto por escalares con la composición: si $f, g \in \mathcal{L}(E)$ y $\lambda \in \mathbb{K}$, entonces

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

2.4.8 Si $f \in \mathcal{L}(E)$, las potencias f^p de f se definen poniendo

$$f^0 = id_E\,, \quad f^1 = f \quad \text{y} \quad f^{p+1} = f^p \circ f\,.$$

Estas relaciones definen f^p para cualquier $p \in \mathbb{N}$, como se prueba utilizando el principio de inducción. También por inducción se demuestra que, si $p, q \in \mathbb{N}$, entonces

$$f^p \circ f^q = f^q \circ f^p = f^{p+q}\,,$$

lo que significa en particular que dos potencias de un mismo endomorfismo siempre conmutan.

Si $p(X) \in \mathbb{K}[X]$ es el polinomio

$$p(X) = \mu_0 + \mu_1 X + \mu_2 X^2 + \cdots + \mu_r X^r$$

con coeficientes en $\mathbb{K}$, ponemos por definición

$$p(f) = \mu_0\, id_E + \mu_1 f + \mu_2 f^2 + \cdots + \mu_r f^r \,.$$

Se demuestra utilizando el principio de inducción que, si $p(X), q(X) \in \mathbb{K}[X]$, entonces

$$p(f) + q(f) = (p+q)(f)\,, \quad p(f) \circ q(f) = (p\,q)(f) \quad \text{y} \quad \lambda p(f) = (\lambda p)(f)\,,$$

donde $(p+q)(X)$ y $(p\,q)(X)$ son los polinomios suma y producto de p y q y $(\lambda p)(X)$ es el producto de λ por el polinomio p.

En particular, como el producto de polinomios es conmutativo, resulta que

$$p(f) \circ q(f) = q(f) \circ p(f)\,,$$

es decir, que dos polinomios de un mismo endomorfismo siempre conmutan.

2.4.9 Consideremos ahora el conjunto $M(n)$ de las matrices cuadradas de orden n; es un espacio vectorial (en la forma que vimos en 2.3.2).

Si E es un e.v. de dimensión $n \neq 0$ y $(a_1, \ldots, a_n)$ es una base de E, la aplicación

$$\begin{aligned} \mathcal{L}(E) &\to M(n) \\ f &\to [f, (a_i)] \end{aligned}$$

es un isomorfismo de espacios vectoriales (v. 2.3.8).

Como $\mathcal{L}(E)$ es un anillo, nos interesa introducir una segunda operación interna en $M(n)$ de manera que

- $M(n)$, provisto de la suma y de esta nueva operación, sea un anillo;

- la aplicación $f \to [f, (a_i)]$ sea un isomorfismo de anillos.

Como ocurría en el caso de la composición de endomorfismos, la segunda operación interna de $M(n)$ se puede definir en un marco más general.

2.4.10 DEFINICIÓN. Sean $A = [a_i^j] \in M(n, m)$ y $B = [b_i^j] \in M(m, p)$. (Obsérvese que hacemos coincidir el número de filas de A con el de columnas de B.) La matriz $[c_i^j]$ de $M(n, p)$, donde, para $1 \leq i \leq n$ y $1 \leq j \leq p$, el elemento c_i^j viene dado por la fórmula

$$\textbf{(pm)} \quad c_i^j = \sum_{k=1}^{m} b_k^j a_i^k = b_1^j a_i^1 + \cdots + b_m^j a_i^m \,,$$

se denomina *producto* de A y B y se representa por $B\,A$. (El orden aparentemente invertido en el símbolo $B\,A$, utilizado para representar el producto de A y B, tiene por objeto acordar este producto con la composición de aplicaciones, donde se presentaba la misma anomalía.) La matriz $B\,A$ tiene el mismo número de columnas que A y el mismo número de filas que B.

Los elementos que intervienen en la fórmula para el cálculo de c_i^j son los de la fila número j de B y los de la columna número i de A. Figuradamente se dice que, para obtener el término c_i^j de la columna i y la fila j de $B\,A$, multiplicamos la fila j de B por la columna i de A.

2.4.11 Ejemplo.

$$\begin{bmatrix} 1. & -1. & 0 & 3.1 \\ 0 & 2.2 & -1.2 & 0 \\ -1. & -1. & 0 & 2. \end{bmatrix} \begin{bmatrix} 2.1 & -1. \\ 0 & 0.5 \\ 1.2 & 0 \\ 1.1 & 1. \end{bmatrix} = \begin{bmatrix} -5.51 & 1.6 \\ \boxed{-1.44} & 1.1 \\ 0.1 & 2.5 \end{bmatrix}.$$

El elemento -1.44 se ha obtenido como

$$-1.44 = 0 \times 2.1 + 2.2 \times 0 + (-1.2) \times 1.2 + 0 \times 1.1\,.$$

2.4.12 PROPOSICIÓN

Si $A \in M(n,m)$ y $B \in M(m,p)$, entonces $A^t B^t$ también existe y

$$(B\,A)^t = A^t B^t\,.$$

La posibilidad del producto citado es inmediata. Además, utilizando las notaciones $A = [a_i^j]$, $B = [b_i^j]$, $B\,A = C = [c_i^j]$ y $A^t B^t = D = [d_i^j]$, tenemos

$$d_i^j = \sum_{k=1}^{m} {}^t a_k^j \, {}^t b_i^k = \sum_{k=1}^{m} a_j^k b_k^i = c_j^i\,,$$

luego $D = C^t$, es decir, $A^t B^t = (B\,A)^t$.

2.4.13 Se prueban sin dificultad las siguientes propiedades del producto de matrices:

a) si $A \in M(n,m)$, $B \in M(m,p)$ y $C \in M(p,q)$, entonces

$$C(B\,A) = (C\,B)A\,;$$

b) si $A, A' \in M(n,m)$ y $B \in M(m,p)$, entonces

$$B(A + A') = B\,A + B\,A'\,;$$

c) si $A \in M(n,m)$ y $B, B' \in M(m,p)$, entonces

$$(B + B')A = B\,A + B'\,A\,.$$

2.4.14 Si $A, B \in M(n)$, entonces $B\,A \in M(n)$, o sea, el producto de matrices es una operación interna en $M(n)$. Consideremos el trío $(M(n), +, \times)$, donde '+' es la suma de matrices y '$\times$' el producto de matrices que acabamos de definir. $(M(n), +)$ es un grupo conmutativo, y el producto es asociativo y distributivo respecto de la suma, como se desprende del párrafo precedente. Resulta así que $(M(n), +, \times)$ es un anillo.

2.4.15 Representamos por I_n la matriz de $M(n)$ cuyo término general a_i^j vale

$$a_i^j = \begin{cases} 1 & \text{si } i = j \\ 0 & \text{si } i \neq j\,, \end{cases}$$

es decir, I_n es la matriz diagonal cuyos elementos diagonales son todos 1 (el 1 de $\mathbb{K}$); le llamamos *matriz unidad* de orden n.

Se verifica que

$$(\forall A \in M(n)\,) \quad A\,I_n = I_n\,A = A\,,$$

o sea, que I_n es la unidad del anillo $M(n)$; éste es, pues, un anillo unitario.

2.4.16 Se suele utilizar el símbolo δ_i^j, con $i, j \in \mathbb{N}$, para representar el escalar 0 cuando $i \neq j$ y el 1 cuando $i = j$. El símbolo δ_i^j recibe el nombre de 'delta de Kronecker'[2].

Utilizando este símbolo resulta que

$$I_n = [\delta_i^j]\,.$$

2.4.17 Ejemplo.

$$I_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad I_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

2.4.18 Si $A, B \in M(n)$ y $\lambda \in \mathbb{K}$, se tiene que

$$\lambda(B\,A) = (\lambda B)A = B(\lambda A)\,.$$

En particular, si $A \in M(n)$ y $\lambda \in \mathbb{K}$,

$$\lambda A = (\lambda I_n)A = A(\lambda I_n)\,.$$

2.4.19 Una matriz A de $M(n)$ es un elemento inversible del anillo cuando existe una matriz de $M(n)$, representada por A^{-1}, que verifica

$$AA^{-1} = A^{-1}A = I_n\,.$$

[2]En honor al matemático alemán Leopold Kronecker (1823-1891).

Decimos que un elemento inversible de $M(n)$ es una *matriz inversible* y que A^{-1} es la *inversa* de A. Cuando existe esta inversa, es única.

El subconjunto de $M(n)$ formado por las matrices inversibles se representa por $GL_{\mathbb{K}}(n)$.

$(GL_{\mathbb{K}}(n), \times)$ es un grupo; se le suele llamar 'grupo lineal de orden n'.

Fácilmente se demuestran las siguientes propiedades:

a) si $A \in M(n)$ es inversible, A^{-1} es inversible y

$$(A^{-1})^{-1} = A\,;$$

b) si $A \in M(n)$ es inversible, A^t es inversible y

$$(A^t)^{-1} = (A^{-1})^t\,;$$

c) si $A \in M(n)$ es inversible y $\lambda \neq 0$, λA es inversible y

$$(\lambda A)^{-1} = \lambda^{-1} A^{-1}\,;$$

a) si $A, B \in M(n)$ son inversibles, BA es inversible y

$$(BA)^{-1} = A^{-1} B^{-1}\,.$$

2.4.20 Ejemplo. I_n es inversible y su inversa es la propia I_n.

2.4.21 Ejemplo. Si $\alpha_1^1 \alpha_2^2 - \alpha_1^2 \alpha_2^1 \neq 0$, entonces la matriz

$$A = \begin{bmatrix} \alpha_1^1 & \alpha_2^1 \\ \alpha_1^2 & \alpha_2^2 \end{bmatrix}$$

es inversible y su inversa es

$$A^{-1} = \frac{1}{\alpha_1^1 \alpha_2^2 - \alpha_1^2 \alpha_2^1} \begin{bmatrix} \alpha_2^2 & -\alpha_2^1 \\ -\alpha_1^2 & \alpha_1^1 \end{bmatrix}\,.$$

2.4.22 TEOREMA

Sean E, E' y E'' espacios vectoriales sobre $\mathbb{K}$, de dimensiones n, m y p; sean $(a_1, \ldots, a_n)$, $(b_1, \ldots, b_m)$ y $(c_1, \ldots, c_p)$ bases de E, E' y E''. Sean $f \in \mathcal{L}(E, E')$ y $g \in \mathcal{L}(E', E'')$, y sea $g \circ f \in \mathcal{L}(E, E'')$ la compuesta de f y g. Entonces

$$[g \circ f, (a_i), (c_k)] = [g, (b_j), (c_k)]\,[f, (a_i), (b_j)]\,.$$

Hagamos

$$\begin{aligned} [f, (a_i), (b_j)] &= A = [\alpha_i^j] \in M(n, m) \\ [g, (b_j), (c_k)] &= B = [\beta_i^j] \in M(m, p) \\ [g \circ f, (a_i), (c_k)] &= C = [\Delta_i^j] \in M(n, p) \\ B\,A &= D = [\gamma_i^j] \in M(n, p)\,. \end{aligned}$$

Con estas notaciones, se trata de probar que $C = D$.

Aplicando la definición (2.2.17) tenemos

$$f(a_i) = (\alpha_i^1, \ldots, \alpha_i^m)_{(b_j)} = \alpha_i^1 b_1 + \cdots + \alpha_i^m b_m \,,$$

luego

$$g \circ f(a_i) = \alpha_i^1 g(b_1) + \cdots + \alpha_i^m g(b_m)\,;$$

pero, por otra parte, aplicando la misma definición,

$$g(b_j) = (\beta_j^1, \ldots, \beta_j^p)_{(c_k)} = \beta_j^1 c_1 + \cdots + \beta_j^p c_p\,;$$

por lo tanto

$$\begin{aligned} g \circ f(a_i) &= \alpha_i^1(\beta_1^1 c_1 + \cdots + \beta_1^p c_p) + \cdots + \alpha_i^m(\beta_m^1 c_1 + \cdots + \beta_m^p c_p) \\ &= (\alpha_i^1 \beta_1^1 + \cdots + \alpha_i^m \beta_m^1)\, c_1 + \cdots + (\alpha_i^1 \beta_1^p + \cdots + \alpha_i^m \beta_m^p)\, c_p \end{aligned}$$

y, aplicando de nuevo la misma definición,

$$\Delta_i^j = \alpha_i^1 \beta_1^j + \cdots \alpha_i^m \beta_m^j = \sum_{l=1}^{m} \alpha_i^l \beta_l^j$$

luego (v. 2.4.10) $\Delta_i^j = \gamma_i^j$, lo que demuestra que $C = D$.

2.4.23 Sea E un e.v. de dimensión n y $(a_1, \ldots, a_n)$ una base de E. La aplicación biyectiva

$$\begin{aligned} \mathcal{L}(E) &\to M(n) \\ f &\to [f, (a_i)]\,, \end{aligned}$$

que es un isomorfismo de espacios vectoriales, verifica en particular que

$$[f + g, (a_i)] = [f, (a_i)] + [g, (a_i)]\,;$$

pero además, utilizando el teorema precedente,

$$[g \circ f, (a_i)] = [g, (a_i)]\,[f, (a_i)]\,.$$

Hemos demostrado así la siguiente proposición.

2.4.24 PROPOSICIÓN

Sea E un e.v. de dimensión $n \neq 0$ y $(a_1, \ldots, a_n)$ una base de E. La aplicación

$$\begin{aligned} \mathcal{L}(E) &\to M(n) \\ f &\to [f, (a_i)] \end{aligned}$$

es un isomorfismo de anillos.

2.4.25 Resulta entonces que

$$\begin{aligned}[][0_{\mathcal{L}(E)}, (a_i)] &= 0_{M(n)}\,, \\ [\,-f, (a_i)] &= -[f, (a_i)]\,, \\ [\,id_E, (a_i)] &= I_n\,.\end{aligned}$$

2.4.26 Otras consecuencias de demostración sencilla son las siguientes:

a) Si f es un automorfismo de E, entonces $[f, (a_i)]$ es una matriz inversible y

$$[f^{-1}, (a_i)] = [f, (a_i)]^{-1}.$$

Recíprocamente, si $A \in M(n)$ y es inversible, y si $A = [f, (a_i)]$, entonces f es un automorfismo de E.

b) La aplicación

$$\begin{aligned} GL(E) &\to GL_{\mathbb{K}}(n) \\ f &\to [f, (a_i)] \end{aligned}$$

es un isomorfismo de grupos.

c) El anillo $M(n)$ no es, en general, ni conmutativo ni íntegro.

2.4.27 Ejemplo. Para encontrar dos matrices $A, B \in M(2)$ tales que $AB \neq BA$, basta tomar

$$A = [f, (e_1, e_2)] \quad \text{y} \quad B = [g, (e_1, e_2)]\,,$$

donde f y g son los endomorfismos de $\mathbb{R}^2$ del ejemplo (2.4.4), o sea,

$$A = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}.$$

Además, en este caso, $A \neq 0$ y $B \neq 0$, pero $AB = 0$, luego A y B son divisores de 0.

2.4.28 Nos interesamos ahora por las matrices A de $M(n)$ tales que

$$(\forall B \in M(n)) \quad AB = BA\,,$$

o sea, aquellas que conmutan con cualquier otra matriz de $M(n)$. Por ejemplo, I_n es una matriz de este tipo, pues

$$(\forall B \in M(n)) \quad I_n\, B = B\, I_n \quad (= B)\,.$$

Vamos a ver que las matrices que poseen esta propiedad son justamente las de la forma λI_n con $\lambda \in \mathbb{K}$, es decir, las matrices diagonales cuyos elementos diagonales son todos iguales.

Hemos visto en (2.4.18) que λI_n conmuta con cualquier otra matriz. Recíprocamente, vamos a ver qué ocurre si $A \in M(n)$ es tal que $AB = BA$ para toda matriz $B \in M(n)$. Pongamos $A = [a_i^j]$ y consideremos las matrices E_i^j de la base canónica de $M(n)$ (v. 2.3.5); en particular

$$(\forall\, i, j\,) \quad AE_i^j = E_i^j A\,.$$

El elemento de la columna i y la fila j es a_j^j para AE_i^j y a_i^i para $E_i^j A$, luego

$$(\forall\, i, j\,) \quad a_i^i = a_j^j\,.$$

Por otra parte, si $i \neq j$, el elemento diagonal de la columna i y la fila i es a_j^i para AE_i^j y 0 para $E_i^j A$, luego

$$i \neq j \quad \Rightarrow \quad a_j^i = 0\,;$$

o sea, A es diagonal y $a_1^1 = \cdots = a_n^n$, es decir, $A = \lambda I_n$, donde $\lambda = a_1^1 = \cdots = a_n^n$; esto completa la demostración.

Llamaremos *matrices escalares* a las matrices de la forma λI_n; el sentido de esta denominación proviene de que

$$(\lambda I_n)A = A(\lambda I_n) = \lambda A\,,$$

como vimos en (2.4.18).

2.4.29 Si $A \in M(n)$, las potencias A^p de A se definen poniendo

$$A^0 = I_n\,, \quad A^1 = A \quad \text{y} \quad A^{p+1} = A^p A\,.$$

Si $p, q \in \mathbb{N}$,

$$A^p A^q = A^q A^p = A^{p+q},$$

lo que significa en particular que dos potencias de una misma matriz siempre conmutan.

Si $p(X) \in \mathbb{K}[X]$ es el polinomio

$$p(X) = \mu_0 + \mu_1 X + \mu_2 X^2 + \cdots + \mu_r X^r$$

con coeficientes en $\mathbb{K}$, ponemos por definición

$$p(A) = \mu_0 I + \mu_1 A + \mu_2 A^2 + \cdots + \mu_r A^r\,.$$

Se demuestra, utilizando el principio de inducción, que si $p(X), q(X) \in \mathbb{K}[X]$ y $\lambda \in \mathbb{K}$, entonces

$$p(A) + q(A) = (p+q)(A), \quad p(A)\,q(A) = (p\,q)(A) \quad \text{y} \quad \lambda p(A) = (\lambda p)(A)\,.$$

En particular, como el producto de polinomios es conmutativo, resulta que

$$p(A)\,q(A) = q(A)\,p(A)\,,$$

es decir, que dos polinomios de una misma matriz siempre conmutan.

Sea E un e.v. de dimensión $n \neq 0$, $(a_1, \dots, a_n)$ una base de E y $f \in \mathcal{L}(E)$. Teniendo en cuenta las fórmulas que proporcionan la matriz de una suma de aplicaciones lineales, la de la compuesta de aplicaciones lineales y la del producto de un escalar por una aplicación lineal, se tiene para toda potencia de f que

$$[f^p, (a_i)] = [f, (a_i)]^p$$

y, para todo polinomio $p(X) \in \mathbb{K}[X]$,

$$[p(f), (a_i)] = p\left([f, (a_i)]\right).$$

2.4.30 Finalizaremos esta sección proporcionando un método para calcular la inversa de una matriz. No es el único metodo que veremos, pero se trata de un método elemental y eficaz.

2.4.31 Las siguientes operaciones sobre las matrices $n \times n$ se denominan (sobre las filas de la matriz):

(oe1) multiplicación de la fila número j por un escalar $\lambda \neq 0$; denotamos esta operación por f_j^λ ;

(oe2) intercambio de las filas número j y número k, operación que representamos por $f_{j,k}$;

(oe3) substitución de la fila número j por la suma de dicha fila con la fila número k $(k \neq j)$ multiplicada por el escalar λ, operación que representamos por $f_{j,k}^\lambda$.

Si $A \in M(n)$ es una matriz $n \times n$, entonces $f_j^\lambda(A)$, $f_{j,k}(A)$ o $f_{j,k}^\lambda(A)$ representan las matrices que se obtienen al aplicar a la matriz A la correspondiente operación elemental.

En particular, si I es la matriz unidad $n \times n$, pondremos

$$F_j^\lambda = f_j^\lambda(I), \quad F_{j,k} = f_{j,k}(I) \quad \text{y} \quad F_{j,k}^\lambda = f_{j,k}^\lambda(I).$$

Las matrices F_j^λ , $F_{j,k}$ y $F_{j,k}^\lambda$ reciben el nombre de *matrices elementales* de tamaño $n \times n$.

2.4.32 Ejemplo. Para la matriz

$$A = \begin{bmatrix} 2. & 2. & 1. \\ 4. & 4. & 1. \\ 1. & 0 & 0 \end{bmatrix}$$

y la operación elemental $f_{2,1}^{-2.}$, se tiene

$$f_{2,1}^{-2.}(A) = \begin{bmatrix} 2. & 2. & 1. \\ 0 & 0 & -1. \\ 1. & 0 & 0 \end{bmatrix}.$$

Por otro lado

$$F_{2,1}^{-2.} = f_{2,1}^{-2.}(I) = \begin{bmatrix} 1. & 0 & 0 \\ -2. & 1. & 0 \\ 0 & 0 & 1. \end{bmatrix}.$$

Se puede comprobar que

$$f_{2,1}^{-2.}(A) = F_{2,1}^{-2.}\,A\,,$$

que es lo que pasa con carácter general, como veremos a continuación.

2.4.33 Si f es una operación elemental y $F = f(I)$ es la correspondiente matriz elemental, entonces

$$(\forall A \in M(n)\,) \quad f(A) = F\,A\,.$$

La demostración de este hecho es sencilla; la realizaremos para una operación elemental del tipo f_j^λ, dejando al cuidado del lector los dos tipos restantes.

Sea $A = [\alpha_i^j] \in M(n)$, f_k^λ la operación elemental consistente en la multiplicación de la fila número k por el escalar $\lambda \neq 0$ y $F_k^\lambda = [\beta_i^j]$ la correspondiente matriz elemental. Se tiene que

$$\beta_i^j = \begin{cases} 1 & \text{si } i = j \neq k \\ \lambda & \text{si } i = j = k \\ 0 & \text{si } i \neq j\,. \end{cases}$$

Pongamos $f_k^\lambda(A) = [\gamma_i^j]$ y $F_k^\lambda\,A = [\epsilon_i^j]$. Para $j \neq k$ tenemos que $\gamma_i^j = \alpha_i^j$ y que $\epsilon_i^j = \beta_1^j\alpha_i^1 + \beta_2^j\alpha_i^2 + \cdots + \beta_n^j\alpha_i^n = \beta_j^j\alpha_i^j = \alpha_i^j$, o sea, $\gamma_i^j = \epsilon_i^j$. Para $j = k$ tenemos que $\gamma_i^k = \lambda\alpha_i^k$ y que $\epsilon_i^k = \beta_1^k\alpha_i^1 + \beta_2^k\alpha_i^2 + \cdots + \beta_n^k\alpha_i^n = \beta_k^k\alpha_i^k = \lambda\alpha_i^k$, o sea, $\gamma_i^k = \epsilon_i^k$. Así pues, $f_k^\lambda(A) = F_k^\lambda\,A$.

2.4.34 Las matrices elementales son inversibles y sus inversas son también matrices elementales. En efecto, observando que

$$f_j^{1/\lambda}(f_j^\lambda(I)\,) = I, \quad f_{j,k}(f_{j,k}(I)\,) = I \quad \text{y} \quad f_{j,k}^{-\lambda}(f_{j,k}^\lambda(I)\,) = I\,,$$

y teniendo en cuenta el apartado precedente, resulta que

$$F_j^{1/\lambda}\,F_j^\lambda = I, \quad F_{j,k}\,F_{j,k} = I \quad \text{y} \quad F_{j,k}^{-\lambda}\,F_{j,k}^\lambda = I\,;$$

con el mismo razonamiento se prueba que

$$F_j^\lambda\,F_j^{1/\lambda} = I \quad \text{y} \quad F_{j,k}^\lambda\,F_{j,k}^{-\lambda} = I\,.$$

En consecuencia, las matrices elementales son inversibles y

$$(F_j^\lambda)^{-1} = F_j^{1/\lambda}\,, \quad (F_{j,k})^{-1} = F_{j,k} \quad \text{y} \quad (F_{j,k}^\lambda)^{-1} = F_{j,k}^{-\lambda}\,.$$

2.4.35 Cálculo de la inversa de una matriz mediante el empleo de las operaciones elementales.

Sea $A \in M(n)$ una matriz. Supongamos que $f_1, f_2, \ldots, f_p$ son operaciones elementales para las que

$$f_p(f_{p-1} \ldots (f_2(f_1(A))\ldots) = I\,,$$

es decir, que aplicadas sucesivamente a la matriz A proporcionan la matriz unidad. Entonces, para las correspondientes matrices elementales, se cumple

$$F_p\, F_{p-1} \cdots F_2\, F_1\, A = I\,.$$

La matriz $F = F_p\, F_{p-1} \cdots F_2\, F_1$ es producto de matrices inversibles, luego es inversible (v. 2.4.19)). Como $F\,A = I$, resulta que $A = F^{-1}$, luego A es inversible. Además

$$A^{-1} = F = F_p\, F_{p-1} \cdots F_2\, F_1\,,$$

o sea,

$$A^{-1} = f_p(f_{p-1} \ldots (f_2(f_1(I))\ldots)\,.$$

Así pues, se trata de transformar A en la matriz unidad mediante operaciones elementales y de transformar I mediante las mismas operaciones, con lo que se obtiene A^{-1}.

La transformación de una matriz $A \in M(n)$ en la matriz unidad a base de operaciones elementales es posible si y sólo si A es inversible, como veremos en (2.5.32).

2.4.36 Ejemplo. Consideremos la matriz

$$A = \begin{bmatrix} 2. & 2. & 1. \\ 4. & 4. & 1. \\ 1. & 0 & 0 \end{bmatrix}$$

Vamos a utilizar sobre A la sucesión de operaciones elementales que indicamos y, simultáneamente, haremos lo mismo con I_3. Al conseguir transformar A en I_3, la matriz transformada de I_3 será A^{-1}.

$$\begin{bmatrix} 2. & 2. & 1. \\ 4. & 4. & 1. \\ 1. & 0 & 0 \end{bmatrix} \qquad \begin{bmatrix} 1. & 0 & 0 \\ 0 & 1. & 0 \\ 0 & 0 & 1. \end{bmatrix}$$

$$\text{con } f_{2,1}^{-2.} \qquad \begin{bmatrix} 2. & 2. & 1. \\ 0 & 0 & -1. \\ 1. & 0 & 0 \end{bmatrix} \qquad \begin{bmatrix} 1. & 0 & 0 \\ -2. & 1. & 0 \\ 0 & 0 & 1. \end{bmatrix}$$

$$\text{con } f_{3,1}^{-0.5} \qquad \begin{bmatrix} 2. & 2. & 1. \\ 0 & 0 & -1. \\ 0 & -1. & -0.5 \end{bmatrix} \qquad \begin{bmatrix} 1. & 0 & 0 \\ -2. & 1. & 0 \\ -0.5 & 0 & 1. \end{bmatrix}$$

$$\text{con } f_{2,3} \qquad \begin{bmatrix} 2. & 2. & 1. \\ 0 & -1. & -0.5 \\ 0 & 0 & -1. \end{bmatrix} \qquad \begin{bmatrix} 1. & 0 & 0 \\ -0.5 & 0 & 1. \\ -2. & 1. & 0 \end{bmatrix}$$

$$\text{con } f_3^{-1.} \quad \begin{bmatrix} 2. & 2. & 1. \\ 0 & -1. & -0.5 \\ 0 & 0 & 1. \end{bmatrix} \quad \begin{bmatrix} 1. & 0 & 0 \\ -0.5 & 0 & 1. \\ 2. & -1. & 0 \end{bmatrix}$$

$$\text{con } f_{2,3}^{0.5} \quad \begin{bmatrix} 2. & 2. & 1. \\ 0 & -1. & 0 \\ 0 & 0 & 1. \end{bmatrix} \quad \begin{bmatrix} 1. & 0 & 0 \\ 0.5 & -0.5 & 1. \\ 2. & -1. & 0 \end{bmatrix}$$

$$\text{con } f_2^{-1.} \quad \begin{bmatrix} 2. & 2. & 1. \\ 0 & 1. & 0 \\ 0 & 0 & 1. \end{bmatrix} \quad \begin{bmatrix} 1. & 0 & 0 \\ -0.5 & 0.5 & -1. \\ 2. & -1. & 0 \end{bmatrix}$$

$$\text{con } f_{1,3}^{-1.} \quad \begin{bmatrix} 2. & 2. & 0 \\ 0 & 1. & 0 \\ 0 & 0 & 1. \end{bmatrix} \quad \begin{bmatrix} -1. & 1. & 0 \\ -0.5 & 0.5 & -1. \\ 2. & -1. & 0 \end{bmatrix}$$

$$\text{con } f_{1,2}^{-2.} \quad \begin{bmatrix} 2. & 0 & 0 \\ 0 & 1. & 0 \\ 0 & 0 & 1. \end{bmatrix} \quad \begin{bmatrix} 0 & 0 & 2. \\ -0.5 & 0.5 & -1. \\ 2. & -1. & 0 \end{bmatrix}$$

$$\text{con } f_1^{0.5} \quad \begin{bmatrix} 1. & 0 & 0 \\ 0 & 1. & 0 \\ 0 & 0 & 1. \end{bmatrix} \quad \begin{bmatrix} 0 & 0 & 1. \\ -0.5 & 0.5 & -1. \\ 2. & -1. & 0 \end{bmatrix}$$

Resulta así que

$$A^{-1} = \begin{bmatrix} 0 & 0 & 1. \\ -0.5 & 0.5 & -1. \\ 2. & -1. & 0 \end{bmatrix}.$$

2.4.37 Ejemplo. Cuando la matriz posee elementos que no están determinados y cuando la matriz es de tamaño general $n \times n$, sin que debamos concretar este tamaño, se puede usar también este método, si bien todo suele ser más complicado. En general, la complejidad de las expresiones que resultan y el problema de un tamaño no concreto hacen que la dificultad del sencillo método que hemos presentado aumente mucho. No obstante, una buena elección del orden de aplicación de las operaciones elementales puede facilitar las cosas.

Por ejemplo, vamos a llevar adelante el método con la matriz

$$A = \begin{bmatrix} 1+a & 1 & 1 & \cdots & 1 \\ 1 & 1+a & 1 & \cdots & 1 \\ 1 & 1 & 1+a & \cdots & 1 \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & 1 & 1 & \cdots & 1+a \end{bmatrix}.$$

El lector comprobará que otras alternativas diferentes de la que empleamos dificultan las cosas. Esta es una sucesión adecuada de operaciones elementales en este caso. Por brevedad, indicamos varias de ellas de manera conjunta. Cuando concentramos varias de estas operaciones, el orden en que se aplican no es importante.

$$\begin{bmatrix} 1+a & 1 & 1 & \cdots & 1 \\ 1 & 1+a & 1 & \cdots & 1 \\ 1 & 1 & 1+a & \cdots & 1 \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & 1 & 1 & \cdots & 1+a \end{bmatrix} \qquad \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{bmatrix}$$

con $f^1_{1,2}, f^1_{1,3}, \ldots, f^1_{1,n}$

$$\begin{bmatrix} n+a & n+a & n+a & \cdots & n+a \\ 1 & 1+a & 1 & \cdots & 1 \\ 1 & 1 & 1+a & \cdots & 1 \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & 1 & 1 & \cdots & 1+a \end{bmatrix} \qquad \begin{bmatrix} 1 & 1 & 1 & \cdots & 1 \\ 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{bmatrix}$$

con $f_1^{1/n+a}$

$$\begin{bmatrix} 1 & 1 & 1 & \cdots & 1 \\ 1 & 1+a & 1 & \cdots & 1 \\ 1 & 1 & 1+a & \cdots & 1 \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & 1 & 1 & \cdots & 1+a \end{bmatrix} \qquad \begin{bmatrix} \frac{1}{n+a} & \frac{1}{n+a} & \frac{1}{n+a} & \cdots & \frac{1}{n+a} \\ 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{bmatrix}$$

con $f^{-1}_{2,1}, f^{-1}_{3,1}, \ldots, f^{-1}_{n,1}$

$$\begin{bmatrix} 1 & 1 & 1 & \cdots & 1 \\ 0 & a & 0 & \cdots & 0 \\ 0 & 0 & a & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & a \end{bmatrix} \qquad \begin{bmatrix} \frac{1}{n+a} & \frac{1}{n+a} & \frac{1}{n+a} & \cdots & \frac{1}{n+a} \\ -\frac{1}{n+a} & 1-\frac{1}{n+a} & -\frac{1}{n+a} & \cdots & -\frac{1}{n+a} \\ -\frac{1}{n+a} & -\frac{1}{n+a} & 1-\frac{1}{n+a} & \cdots & -\frac{1}{n+a} \\ \vdots & \vdots & \vdots & & \vdots \\ -\frac{1}{n+a} & -\frac{1}{n+a} & -\frac{1}{n+a} & \cdots & 1-\frac{1}{n+a} \end{bmatrix}$$

con $f_2^{1/a}, f_3^{1/a}, \ldots, f_n^{1/a}$

$$\begin{bmatrix} 1 & 1 & 1 & \cdots & 1 \\ 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{bmatrix}$$

$$\frac{1}{a(n+a)}\begin{bmatrix} a & a & a & \cdots & a \\ -1 & n+a-1 & -1 & \cdots & -1 \\ -1 & -1 & n+a-1 & \cdots & -1 \\ \vdots & \vdots & \vdots & & \vdots \\ -1 & -1 & -1 & \cdots & n+a-1 \end{bmatrix}$$

con $f_{1,2}^{-1}, f_{1,3}^{-1}, \ldots, f_{1,n}^{-1}$

$$\begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{bmatrix}$$

$$\frac{1}{a(n+a)}\begin{bmatrix} n+a-1 & -1 & -1 & \cdots & -1 \\ -1 & n+a-1 & -1 & \cdots & -1 \\ -1 & -1 & n+a-1 & \cdots & -1 \\ \vdots & \vdots & \vdots & & \vdots \\ -1 & -1 & -1 & \cdots & n+a-1 \end{bmatrix}$$

Resulta así que

$$A^{-1} = \frac{1}{a(n+a)}\begin{bmatrix} n+a-1 & -1 & -1 & \cdots & -1 \\ -1 & n+a-1 & -1 & \cdots & -1 \\ -1 & -1 & n+a-1 & \cdots & -1 \\ \vdots & \vdots & \vdots & & \vdots \\ -1 & -1 & -1 & \cdots & n+a-1 \end{bmatrix}.$$

2.5 Matrices y coordenadas.

2.5.1 Consideremos una matriz $A = [\alpha_i^j] \in M(n,m)$. Si fijamos dos espacios vectoriales E y E' de dimensiones n y m, y fijamos dos bases $(a_1, \ldots, a_n)$ y $(b_1, \ldots, b_m)$ de E y E', entonces

$$A = [f, (a_i), (b_j)]$$

para una única aplicación lineal $f \in \mathcal{L}(E, E')$, como vimos en (2.3.7). Así, pues, a la matriz A le corresponden diversas aplicaciones lineales, una por cada elección de espacios y de bases.

La elección más sencilla que se puede hacer es la correspondiente a $\mathbb{K}^n, \mathbb{K}^m$ y a sus bases canónicas; la aplicación $f \in \mathcal{L}(\mathbb{K}^n, \mathbb{K}^m)$ que, en las bases canónicas, tiene A como matriz asociada suele recibir el nombre de 'aplicación lineal dada por A'. Si

$$x = (x^1, \ldots, x^n) \in \mathbb{K}^n$$

y

$$f(x) = (y^1, \dots, y^m)\,,$$

resulta que

$$f(x) = x^1 f(e_1) + \cdots + x^n f(e_n)\,,$$

y, como los $f(e_i)$ son los vectores columna de A, tenemos

$$\begin{aligned} f(x) &= x^1\,(\alpha_1^1, \dots, \alpha_1^m) + \cdots + x^n\,(\alpha_n^1, \dots, \alpha_n^m) \\ &= (x^1\alpha_1^1 + \cdots + x^n\alpha_n^1\,, \dots, x^1\alpha_1^m + \cdots + x^n\alpha_n^m)\,, \end{aligned}$$

o sea,

$$\begin{aligned} y^1 &= x^1\alpha_1^1 + \cdots + x^n\alpha_n^1 \\ &\cdots\cdots\cdots\cdots\cdots \\ y^m &= x^1\alpha_1^m + \cdots + x^n\alpha_n^m\,; \end{aligned}$$

esto significa que (v. 2.4.10)

$$\begin{bmatrix} y^1 \\ \vdots \\ y^m \end{bmatrix} = \begin{bmatrix} \alpha_1^1 & \cdots & \alpha_n^1 \\ \vdots & & \vdots \\ \alpha_1^m & \cdots & \alpha_n^m \end{bmatrix} \begin{bmatrix} x^1 \\ \vdots \\ x^n \end{bmatrix},$$

o bien, con la identificación habitual (v. 2.2.13) entre vectores y matrices columna,

$$\textbf{(1)} \qquad \boxed{f(x) = A\,x}\,.$$

Para la aplicación lineal dada por A, la imagen de un vector x de $\mathbb{K}^n$ se obtiene así mediante un sencillo producto de matrices.

2.5.2 Ejemplo. Si

$$A = \begin{bmatrix} 1.2 & 0 \\ -1. & 2. \\ 0 & 3.1 \end{bmatrix},$$

la aplicación $f \in \mathcal{L}(\mathbb{R}^2, \mathbb{R}^3)$ dada por A envía un vector (x, y) de $\mathbb{R}^2$ al vector

$$(1.2x, -x + 2y, 3.1y)$$

de $\mathbb{R}^3$, puesto que

$$\begin{bmatrix} 1.2 & 0 \\ -1. & 2. \\ 0 & 3.1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 1.2x \\ -x + 2y \\ 3.1y \end{bmatrix}.$$

2.5.3 En el caso más general en que

$$A = [f, (a_i), (b_j)]\,,$$

con $f \in \mathcal{L}(E, E')$, se obtiene una expresión matricial análoga para el cálculo de la imagen por f de un vector de E. En primer lugar, expondremos un método para asociar una matriz columna a cada vector de un e.v. de dimensión finita.

2.5.4 DEFINICIÓN. Sea E un e.v. de dimensión $n \neq 0$ y $(a_1, \ldots, a_n)$ una base de E; sea $x \in E$. Si
$$x = (x^1, \ldots, x^n)_{(a_i)},$$
decimos que la matriz columna $1 \times n$
$$\begin{bmatrix} x^1 \\ \vdots \\ x^n \end{bmatrix}$$
es la *matriz asociada a x y a la base* $(a_1, \ldots, a_n)$, y la representamos por $[x, (a_i)]$ (o por $[x, (a_1, \ldots, a_n)]$).

Cuando $E = \mathbb{K}^n$, la matriz asociada a un vector x y a la base canónica es la matriz que identificamos habitualmente con x (v. 2.2.13).

2.5.5 Expresión matricial de una aplicación lineal. Sean E, E' dos e.v. de dimensiones n y m, y $(a_1, \ldots, a_n)$ y $(b_1, \ldots, b_n)$ bases de E y E'; sea $f \in \mathcal{L}(E, E')$ y $A = [\alpha_i^j] = [f, (a_i), (b_j)]$. Si x es un vector de E de coordenadas
$$x = (x^1, \ldots, x^n)_{(a_i)}$$
y su imagen $f(x)$ es el vector de E' de coordenadas
$$f(x) = (y^1, \ldots, y^m)_{(b_j)},$$
entonces
$$\begin{bmatrix} y^1 \\ \vdots \\ y^m \end{bmatrix} = \begin{bmatrix} \alpha_1^1 & \cdots & \alpha_n^1 \\ \vdots & & \vdots \\ \alpha_1^m & \cdots & \alpha_n^m \end{bmatrix} \begin{bmatrix} x^1 \\ \vdots \\ x^n \end{bmatrix},$$
es decir,
$$\textbf{(2)} \qquad \boxed{[f(x), (b_j)] = [f, (a_i), (b_j)]\,[x, (a_i)]}\,.$$
La demostración de esta fórmula es muy parecida a la que hicimos en (2.5.1).

La expresión **(2)** es prácticamente igual a la **(1)** de (2.5.1), con la dificultad suplementaria que impone la descripción de las bases en las que trabajamos.

2.5.6 La matriz $[x, (a_i)]$ se generaliza al caso de un número finito cualquiera de vectores de la manera siguiente. Sea E un e.v. de dimensión $n \neq 0$ y $(a_1, \ldots, a_n)$ una base de E; sea $(x_1, \ldots, x_p)$ un sistema de p vectores de E, con
$$\begin{aligned} x_1 &= (x_1^1, \ldots, x_1^n)_{(a_i)} \\ &\ldots\ldots\ldots\ldots\ldots\ldots \\ x_p &= (x_p^1, \ldots, x_p^n)_{(a_i)}\,. \end{aligned}$$

Representamos por $[(x_1, \ldots, x_p), (a_i)]$ la matriz $p \times n$

$$[(x_1, \ldots, x_p), (a_i)] = \begin{bmatrix} x_1^1 & \cdots & x_p^1 \\ \vdots & & \vdots \\ x_1^n & \cdots & x_p^n \end{bmatrix},$$

cuya columna número j es $[x_j, (a_i)]$, o sea, está formada por las coordenadas del vector x_j en la base $(a_1, \ldots, a_n)$.

Cuando el sistema es una base de E, la matriz (cuadrada en este caso) que resulta es un instrumento importante en los cálculos prácticos que se derivan de un cambio de base.

2.5.7 DEFINICIÓN. Sea E un e.v. de dimensión $n \neq 0$ y $(a_1, \ldots, a_n)$ y $(a'_1, \ldots, a'_n)$ dos bases de E. La matriz $[(a'_1, \ldots, a'_n), (a_i)]$, de tamaño $n \times n$, cuya columna número i está formada por las coordenadas de a'_i en la base $(a_1, \ldots, a_n)$, se denomina *matriz de paso* de la base $(a_1, \ldots, a_n)$ a la base $(a'_1, \ldots, a'_n)$. Abreviaremos frecuentemente la notación poniendo simplemente $[(a'_i), (a_i)]$; también se la representa a veces por $[(a_i) \to (a'_i)]$ o por $P[(a_i) \to (a'_i)]$.

Si

$$a'_i = (\alpha_i^1, \ldots, \alpha_i^n)_{(a_1, \ldots, a_n)},$$

entonces

$$\begin{bmatrix} \alpha_i^1 \\ \vdots \\ \alpha_i^n \end{bmatrix}$$

es la columna número i de $[(a'_i), (a_i)]$. El elemento α_i^j de la columna i y la fila j de $[(a'_i), (a_i)]$ es la coordenada número j en la base $(a_1, \ldots, a_n)$ del vector a'_i de la otra base.

Observando que para la aplicación idéntica de E tenemos

$$id_E(a'_i) = a'_i$$

y recordando la definición (2.2.17), resulta que

$$[(a'_i), (a_i)] = [id_E, (a'_i), (a_i)].$$

2.5.8 Ejemplo. Si $(a_1, \ldots, a_n)$ es una base del e.v. E, se tiene

$$[(a_i), (a_i)] = I_n.$$

2.5.9 Ejemplo. Si en $\mathbb{R}^3$ consideramos la base canónica (e_1, e_2, e_3) y la base (a_1, a_2, a_3) formada por $a_1 = (0.2, 0, 1.)$, $a_2 = (1., 1., 0)$ y $a_3 = (0, 0, -1.1)$, entonces

$$[(a_i), (e_i)] = \begin{bmatrix} 0.2 & 1. & 0 \\ 0 & 1. & 0 \\ 1. & 0 & -1.1 \end{bmatrix}.$$

2.5.10 Ejemplo. Si (e_1, e_2, e_3) es la base canónica de $\mathbb{K}^3$, entonces

$$[(e_2, e_3, e_1), (e_1, e_2, e_3)] = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} .$$

2.5.11 PROPOSICIÓN

Sea E un e.v. de dimensión $n \neq 0$ y $(a_1, \ldots, a_n)$ y $(a'_1, \ldots, a'_n)$ dos bases de E. La matriz de paso

$$[(a'_i), (a_i)]$$

es inversible y su inversa es

$$[(a'_i), (a_i)]^{-1} = [(a_i), (a'_i)]$$

o sea, la matriz de paso de la base (a'_i) a la base (a_i).

Como $[(a'_i), (a_i)] = [id_E, (a'_i), (a_i)]$ y $[(a_i), (a'_i)] = [id_E, (a_i), (a'_i)]$, resulta que, aplicando (2.4.22),

$$\begin{aligned} [(a'_i), (a_i)]\,[(a_i), (a'_i)] &= [id_E, (a'_i), (a_i)]\,[id_E, (a_i), (a'_i)] \\ &= [id_E \circ id_E, (a_i), (a_i)] \\ &= [id_E, (a_i)] \\ &= I_n \end{aligned}$$

y, análogamente,

$$[(a_i), (a'_i)]\,[(a'_i), (a_i)] = I_n ,$$

lo que concluye la demostración.

2.5.12 PROPOSICIÓN

Sea E un e.v. de dimensión $n \neq 0$ y $(a_1, \ldots, a_n)$ una base de E; sea $P \in M(n)$ una matriz inversible. Existe una base $(a'_1, \ldots, a'_n)$ de E tal que

$$P = [(a'_i), (a_i)] .$$

Como consecuencia de (2.3.8), existe un endomorfismo f de E tal que

$$P = [f, (a_i)]$$

y, puesto que P es inversible, f es un automorfismo (v. 2.4.26). Poniendo entonces

$$a'_1 = f(a_1), \ldots, a'_n = f(a_n) ,$$

resulta de (1.4.37) que $(a'_1, \ldots, a'_n)$ es una base de E. Finalmente, como

$$[f, (a_i)] = [(a'_i), (a_i)] ,$$

hemos demostrado la proposición.

2.5.13 Ejemplo. Sea E un e.v. de dimensión 2 y (a_1, a_2) una base de E. La matriz

$$P = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}$$

es inversible (v. 2.4.21) y los vectores

$$a'_1 = (1,1)_{(a_1,a_2)}\,, \quad a'_2 = (1,0)_{(a_1,a_2)}\,,$$

o sea,

$$a'_1 = a_1 + a_2\,, \quad a'_2 = a_1\,,$$

forman una base de E tal que

$$P = [(a'_1, a'_2), (a_1, a_2)]\,.$$

2.5.14 Cambio de base. Sea E un e.v. de dimensión $n \neq 0$ y $(a_1, \ldots, a_n)$ y $(a'_1, \ldots, a'_n)$ dos bases de E; sea $P = [(a'_i), (a_i)]$, la correspondiente matriz de paso. Si $x \in E$, se tiene

$$\textbf{(3)} \qquad \boxed{[x, (a_i)] = [(a'_i), (a_i)]\,[x, (a'_i)]}$$

o, en forma más explícita, si

$$x = (x^1, \ldots, x^n)_{(a_i)} \quad \text{y} \quad x = (x'^1, \ldots, x'^n)_{(a'_i)}\,,$$

entonces

$$\begin{bmatrix} x^1 \\ \vdots \\ x^n \end{bmatrix} = P \begin{bmatrix} x'^1 \\ \vdots \\ x'^n \end{bmatrix},$$

lo que, por otra parte, equivale a

$$\begin{bmatrix} x'^1 \\ \vdots \\ x'^n \end{bmatrix} = P^{-1} \begin{bmatrix} x^1 \\ \vdots \\ x^n \end{bmatrix}.$$

La demostración de la fórmula **(3)** es inmediata si observamos que coincide con

$$[id_E(x), (a_i)] = [id_E, (a'_i), (a_i)]\,[x, (a'_i)]$$

y que ésta es, a su vez, un caso particular de la fórmula **(2)** de (2.5.5).

2.5.15 Ejemplo. Consideremos el e.v. y las bases del ejemplo (2.5.13) con $E = \mathbb{R}^2$. El vector

$$x = (1.2, -1.)_{(a'_1,a'_2)}$$

de $\mathbb{R}^2$ expresado en la base (a_1, a_2) es

$$x = (0.2, 1.2)_{(a_1,a_2)}$$

puesto que

$$\begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 1.2 \\ -1. \end{bmatrix} = \begin{bmatrix} 0.2 \\ 1.2 \end{bmatrix}.$$

La inversa de

$$P = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}$$

es (v. 2.4.21)

$$P^{-1} = \begin{bmatrix} 0 & 1 \\ 1 & -1 \end{bmatrix},$$

luego el vector

$$x = (1.1, 3.)_{(a_1,a_2)}$$

de $\mathbb{R}^2$ expresado en la base (a'_1, a'_2) es

$$x = (3., -1.9)_{(a'_1,a'_2)}$$

puesto que

$$\begin{bmatrix} 0 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} 1.1 \\ 3. \end{bmatrix} = \begin{bmatrix} 3. \\ -1.9 \end{bmatrix}.$$

2.5.16 Sean E y E' dos e.v. de dimensiones n y m, $(a_1, \ldots, a_n)$ y $(a'_1, \ldots, a'_n)$ dos bases de E y $(b_1, \ldots, b_m)$ y $(b'_1, \ldots, b'_m)$ dos bases de E'. Sea $f \in \mathcal{L}(E, E')$ y pongamos

$$A = [f, (a_i), (b_j)] \quad \text{y} \quad A' = [f, (a'_i), (b'_j)].$$

La relación entre estas dos matrices asociadas a f es

$$\textbf{(4)} \qquad \boxed{A' = Q^{-1}\, A\, P},$$

donde hemos puesto

$$P = [(a'_i), (a_i)] \quad \text{y} \quad Q = [f, (b'_j), (b_j)].$$

Probemos la fórmula **(4)**; como

$$f = id_{E'} \circ f \circ id_E\,,$$

aplicando el teorema (2.4.22) tenemos

$$\begin{aligned} A' &= [f, (a'_i), (b'_j)] \\ &= [id_{E'} \circ f \circ id_E, (a'_i), (b'_j)] \\ &= [id_{E'}\,, (b_j), (b'_j)]\,[f, (a_i), (b_j)]\,[id_E, (a'_i), (a_i)] \\ &= [(b_j), (b'_j)]\,[f, (a_i), (b_j)]\,[(a'_i), (a_i)] \\ &= Q^{-1}\, A\, P\,, \end{aligned}$$

como queríamos demostrar.

2.5.17 Cuando E es un e.v. de dimensión n, $(a_1, \ldots, a_n)$ y $(a'_1, \ldots, a'_n)$ bases de E, $f \in \mathcal{L}(E)$ y ponemos

$$A = [f, (a_i)] \quad \text{y} \quad A' = [f, (a'_i)]\,,$$

la relación entre A y A' es

$$\textbf{(5)} \qquad \boxed{A' = P^{-1}\,A\,P}\,,$$

donde $P = [(a'_i), (a_i)]$, pues no se trata más que de un caso particular de la fórmula **(4)** del párrafo precedente.

2.5.18 Sean $A, A' \in M(n,m)$; decimos que A y A' son matrices *equivalentes* cuando

> **(eq)** existen matrices cuadradas, $P \in M(n)$ y $Q \in M(m)$, inversibles y tales que
> $$A' = Q\,A\,P\,.$$

Fácilmente se prueba que la equivalencia así definida en una relación de equivalencia en $M(n,m)$.

2.5.19 Ejemplo. Dos matrices asociadas a una misma aplicación lineal en diferentes bases son equivalentes, como hemos visto en (2.5.16).

2.5.20 DEFINICIÓN. Sea $A \in M(n,m)$; sus vectores columna (v. 2.2.12) forman un sistema de n vectores de $\mathbb{K}^m$; decimos que el rango de este sistema de vectores (v. 1.5.9) es el *rango* de la matriz A, y lo representamos por $\operatorname{rg} A$. Evidentemente,

$$\operatorname{rg} A \leq n \quad \text{y} \quad \operatorname{rg} A \leq m\,,$$

puesto que, por una parte, se trata de un sistema de n vectores y, por otra, estos vectores pertenecen a un espacio de dimensión m.

Más adelante (v. 2.6.38) tendremos ocasión de probar que $\operatorname{rg} A$ es también el rango del sistema formado por los m vectores fila de A (que son de $\mathbb{K}^n$).

2.5.21 Ejemplo. El rango de la matriz

$$A = \begin{bmatrix} 2 & -1 & 3 & 4 \\ 1 & 0 & 1 & 1 \\ 3 & 0 & 1 & -1 \\ 0 & 2 & 1 & 2 \end{bmatrix}$$

de $M_{\mathbb{R}}(4)$ es 3, pues vimos en (1.5.14) que el rango del sistema que forman sus vectores columna es 3.

2.5.22 PROPOSICIÓN

---

a)Sean E y E' dos e.v. de dimensiones n y m, $(a_1, \ldots, a_n)$ y $(b_1, \ldots, b_m)$ bases de E y E', $f \in \mathcal{L}(E, E')$ y $A = [f, (a_i), (b_j)]$. Entonces

$$\operatorname{rg} A = \operatorname{rg} f \,.$$

b)Sea E un e.v. de dimensión n, $(a_1, \ldots, a_n)$ una base de E, $(x_1, \ldots, x_p)$ un sistema de E y $A = [(x_1, \ldots, x_p), (a_i)]$ (v. 2.5.6). Entonces

$$\operatorname{rg} A = \operatorname{rg}(x_1, \ldots, x_p) \,.$$

---

a) Si $A = [\alpha_i^j]$, los vectores columna de A son

$$\begin{aligned} c_1 &= (\alpha_1^1, \ldots, \alpha_1^m) \\ &\cdots\cdots\cdots\cdots \\ c_n &= (\alpha_n^1, \ldots, \alpha_n^m) \,; \end{aligned}$$

por otra parte,

$$\begin{aligned} f(a_1) &= (\alpha_1^1, \ldots, \alpha_1^m)_{(b_j)} \\ &\cdots\cdots\cdots\cdots \\ f(a_n) &= (\alpha_n^1, \ldots, \alpha_n^m)_{(b_j)} \end{aligned}$$

y, como vimos en (2.1.16), tenemos entonces

$$\operatorname{rg}(c_1, \ldots, c_n) = \operatorname{rg}(f(a_1), \ldots, f(a_n)) \,;$$

pero el primer miembro es $\operatorname{rg} A$ y el segundo es $\operatorname{rg} f$ (v. 2.1.21).

b) La demostración es idéntica a la anterior.

2.5.23 LEMA

Toda matriz $A \in M(n, m)$ de rango r es equivalente a la matriz $n \times m$

$$\begin{array}{cl} \left[\begin{array}{c|c} I_r & 0 \\ \hline 0 & 0 \end{array}\right] & \begin{array}{l} r \\ m-r \end{array} \\ \begin{array}{cc} r & n-r \end{array} & \end{array}$$

definida por los bloques cuyo 'tamaño' se indica.

Sea $f \in \mathcal{L}(\mathbb{K}^n, \mathbb{K}^m)$ la aplicación lineal dada por A (v. 2.5.1), que será de rango r. Consideremos entonces una base

$$(a_1, \ldots, a_r, a_{r+1}, \ldots a_n)$$

de $\mathbb{K}^n$ tal que $(a_{r+1}, \ldots, a_n)$ sea una base de $\operatorname{Ker} f$. Sabemos que $(f(a_1), \ldots, f(a_r))$ es una base de $\operatorname{Im} f$ (v. 2.1.26). Ampliemos esta base hasta formar una base

$$(f(a_1), \ldots, f(a_r), b_{r+1}, \ldots, b_m)$$

de $\mathbb{K}^m$. La matriz asociada a f en las bases de $\mathbb{K}^n$ y $\mathbb{K}^m$ que hemos formado así es justamente

$$\left[\begin{array}{c|c} I_r & 0 \\ \hline 0 & 0 \end{array}\right],$$

como se comprueba con facilidad. Por fin, como dos matrices asociadas a una misma aplicación lineal son equivalentes (v. 2.5.19), obtenemos la conclusión.

2.5.24 PROPOSICIÓN

Sean $A, A' \in M(n, m)$; las propiedades que siguen son equivalentes:

(i) $\operatorname{rg} A = \operatorname{rg} A'$;

(ii) A y A' son equivalentes;

(iii) A y A' están asociadas (en distintas bases) a una misma aplicación lineal.

(i)$\Rightarrow$(ii). Si $\operatorname{rg} A = \operatorname{rg} A' = r$, entonces A y A' son equivalentes a una misma matriz, como resulta del lema precedente; en consecuencia, A y A' son equivalentes.

(ii)$\Rightarrow$(iii). Como A y A' son equivalentes, entonces $A' = QAP$, con $P \in M(n)$ y $Q \in M(m)$ inversibles. Sea $f \in \mathcal{L}(\mathbb{K}^n, \mathbb{K}^m)$ la aplicación lineal dada por A. Existen bases $(a_1, \ldots, a_n)$ de $\mathbb{K}^n$ y $(b_1, \ldots, b_m)$ de $\mathbb{K}^m$ tales que

$$P = [(a_i), (e_i)] \quad \text{y} \quad Q^{-1} = [(b_j), (e_j)],$$

como vimos en (2.5.12). Entonces

$$[f, (a_i), (b_j)] = (Q^{-1})^{-1} A\, P = Q\, A\, P = A',$$

o sea, A' está tambiés asociada a f.

(iii)$\Rightarrow$(i). Si f es una aplicación lineal a la que están asociadas tanto A como A', resulta que $\operatorname{rg} A = \operatorname{rg} f = \operatorname{rg} A'$.

2.5.25 DEFINICIÓN. Sean $A, A' \in M(n)$ dos matrices cuadradas; decimos que A y A' son matrices *semejantes* cuando

(sm) existe una matriz cuadrada $P \in M(n)$, inversible y tal que

$$A' = P^{-1} A\, P.$$

La semejanza es una relación de equivalencia en $M(n)$.

2.5.26 Ejemplo. Si $f \in \mathcal{L}(E)$ con $\dim E = n \neq 0$, dos matrices $A = [f, (a_i)]$ y $A' = [f, (a'_i)]$ asociadas a f (en una base distinta) son semejantes (v. 2.5.17).

2.5.27 Ejemplo. La única matriz semejante a una matriz escalar λI_n (v. 2.4.28) es ella misma. En particular, la única matriz semejante a I_n es ella misma.

2.5.28 Ejemplo. Dos matrices semejantes son equivalentes y tienen por lo tanto el mismo rango. Sin embargo, las matrices de $M_{\mathbb{R}}(2)$

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad \text{y} \quad \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$$

son ambas de rango 2, pero no son semejantes, como acabamos de ver en el ejemplo precedente.

2.5.29 PROPOSICIÓN

Sean $A, A' \in M(n)$; las propiedades que siguen son equivalentes:

(i) A y A' son semejantes ;
(ii) A y A' están asociadas (en distinta base) a un mismo endomorfismo .

El ejemplo (2.5.26) afirma que (ii)$\Rightarrow$(i). La demostración de (i)$\Rightarrow$(ii) sigue la técnica empleada en (2.5.24) y no la repetiremos aquí.

2.5.30 La proposición que sigue nos proporciona un primer criterio práctico para reconocer las matrices inversibles. Sin embargo, obtendremos otro aún mejor en el capítulo siguiente (v. 3.2.18).

2.5.31 PROPOSICIÓN

Sea $A \in M(n)$; las siguientes afirmaciones son equivalentes:

(i) A es inversible ;
(ii) el endomorfismo $f \in \mathcal{L}(\mathbb{K}^n)$ dado por A es un automorfismo;
(iii) $\operatorname{rg} A = n$.

La equivalencia de (i) y (ii) es un caso particular de (2.4.26); la de (ii) y (iii) es consecuencia de (2.1.24) y de que $\operatorname{rg} A = \operatorname{rg} f$.

2.5.32 Sea $A \in M(n)$ una matriz cuadrada. Estamos ahora en condiciones de probar la afirmación que hicimos en (2.4.35) sobre la posibilidad de transformar A en I_n utilizando operaciones elementales.

Ya vimos que, si es posible transformar A en I_n, entonces A es inversible. Probaremos ahora que, si A es inversible, entonces siempre es posible transformar A en la matriz unidad mediante operaciones elementales.

Supongamos que A es inversible y llamemos $a_1, \dots, a_n$ a los vectores fila de A. La matriz A^t es también inversible y sus vectores columna son $a_1, \dots, a_n$; como $\mathrm{rg}\, A^t = n$, entonces $(a_1, \dots, a_n)$ es una base de $\mathbb{K}^n$. Representemos por $(e_1, \dots, e_n)$ la base canónica de $\mathbb{K}^n$. Tendremos que

$$e_1 = \alpha^1 a_1 + \alpha^2 a_2 + \cdots + \alpha^n a_n$$

para escalares $\alpha^1, \alpha^2, \dots, \alpha^n$ que no pueden ser todos nulos. Supongamos por comodidad de la exposición que $\alpha^1 \neq 0$ (en todo caso esto siempre se puede lograr cambiando la primera fila por otra de las filas). Entonces la secuencia de operaciones elementales

$$f_1^{\alpha^1}, f_{1,2}^{\alpha^2}, \dots f_{1,n}^{\alpha^n}$$

transforma a_1 en e_1. Se obtiene así una matriz en la que la primera fila ha sido substituída por el vector e_1. Como a_1 se puede despejar de la igualdad que proporcionaba e_1 , resulta que a_1 es combinación lineal de $(e_1, a_2, \dots, a_n)$; por lo tanto, $(e_1, a_2, \dots, a_n)$ genera $\mathbb{K}^n$ y es una base del espacio. A continuación,

$$e_2 = \alpha^1 e_1 + \alpha^2 a_2 + \cdots + \alpha^n a_n$$

(los coeficientes α^i no tiene nada que ver con los de antes). Necesariamente, alguno de los $\alpha^2, \dots, \alpha^n$ es no nulo, pues $e_2 \notin \langle e_1 \rangle$; supondremos que $\alpha^2 \neq 0$. Entonces la secuencia de operaciones elementales

$$f_2^{\alpha^2}, f_{2,1}^{\alpha^1}, f_{2,3}^{\alpha^3}, \dots f_{2,n}^{\alpha^n}$$

transforma a_2 en e_2 . Se obtiene así una matriz donde las dos primeras filas han sido substituídas por e_1 y e_2 . El proceso continúa de la misma manera y nos lleva a reemplazar los vectores fila de A por $e_1, e_2, \dots, e_n$ mediante operaciones elementales, es decir, a obtener la matriz I_n mediante dichas operaciones.

2.6 Dual de un espacio vectorial.

2.6.1 DEFINICIÓN. Sea E un e.v. sobre $\mathbb{K}$. Ya sabemos que $\mathbb{K}$ es un espacio vectorial sobre sí mismo (v. 1.1.2). Las aplicaciones lineales de E en $\mathbb{K}$ reciben el nombre de *formas lineales* sobre E. El conjunto $\mathcal{L}(E, \mathbb{K})$ de estas formas lineales, con las operaciones definidas en (2.3.1), es un espacio vectorial sobre $\mathbb{K}$; recordemos que, si f y g son formas lineales sobre E y $\lambda \in \mathbb{K}$, entonces $f + g$ y λf son las formas lineales definidas para cada $x \in E$ por

$$(f + g)(x) = f(x) + g(x) ,$$

$$(\lambda f)(x) = \lambda f(x)\,.$$

El espacio vectorial $\mathcal{L}(E, \mathrm{I\!K})$ se representa habitualmente por E^* y recibe el nombre de *espacio dual* de E.

Cuando se aborda este tema por vez primera, puede resultar difícil acostumbrarse al hecho de que una forma lineal sobre E es, a la vez, una aplicación (lineal) de E en $\mathrm{I\!K}$ y un elemento del espacio vectorial E^*. Este último aspecto será el que consideremos ahora con más frecuencia. Generalmente vamos a utilizar para los elementos de E^* notaciones como x^*, y^*, etc., aunque a veces los representemos también por f, g, etc., particularmente cuando estudiemos una forma lineal aisladamente. Nótese que, si $x \in E$ y $x^* \in E^*$, x y x^* no tendrán en principio nada que ver uno con el otro, aunque la letra x forme todo o parte de ambos símbolos.

El espacio vectorial $E^* = \mathcal{L}(E, \mathrm{I\!K})$ admite a su vez un espacio dual $\mathcal{L}(E^*, \mathrm{I\!K})$ que representaremos por E^{**} y al que llamaremos *bidual* de E. Sus elementos, para los que emplearemos a veces símbolos como x^{**}, y^{**}, etc., son formas lineales sobre E^*, o sea, aplicaciones lineales de E^* en $\mathrm{I\!K}$.

2.6.2 El cero de E^* es la forma lineal

$$\begin{aligned} 0 : E &\to \mathrm{I\!K} \\ x &\to 0 \end{aligned}$$

que envía todo vector de E al cero de $\mathrm{I\!K}$.

Si $E = \{0\}$, entonces $E^* = \{0\}$ y $E^{**} = \{0\}$.

2.6.3 Ejemplo. Si $\alpha, \beta \in \mathrm{I\!K}$, la aplicación $f : \mathrm{I\!K}^2 \to \mathrm{I\!K}$ dada por

$$f(x, y) = \alpha x + \beta y$$

es una forma lineal sobre $\mathrm{I\!K}^2$, o sea, un elemento de $(\mathrm{I\!K}^2)^*$.

De manera más general, si E es un e.v. de dimensión $n \neq 0$ y $(a_1, \ldots, a_n)$ una base de E, cada elemento $f \in E^*$ tiene asociada una matriz fila $n \times 1$ en la base $(a_1, \ldots, a_n)$ de E y la base canónica $(\,1\,)$ de $\mathrm{I\!K}$ (el sistema de $\mathrm{I\!K}$ formado por el vector 1). Si

$$[f, (a_i), (\,1\,)] = [\alpha_1 \ldots \alpha_n]\,,$$

entonces, para todo $x \in E$ de coordenadas

$$x = (x^1, \ldots, x^n)_{(a_i)}\,,$$

se tiene

$$\begin{aligned} f(x) &= \begin{bmatrix} \alpha_1 & \cdots & \alpha_n \end{bmatrix} \begin{bmatrix} x^1 \\ \vdots \\ x^n \end{bmatrix} \\ &= \alpha_1 x^1 + \cdots + \alpha_n x^n\,. \end{aligned}$$

2.6.4 Ejemplo. Si E es un e.v. y $x \in E$, la aplicación

$$\begin{aligned} E^* &\to \mathbb{K} \\ f &\to f(x)\,, \end{aligned}$$

que hace corresponder a cada forma lineal f sobre E el escalar $f(x)$, es una forma lineal sobre E^*, o sea, un elemento de E^{**}.

2.6.5 Sea E un e.v. y $f \in E^*$. $\operatorname{Im} f$ es un subespacio de $\mathbb{K}$, luego puede ser de dimensión 0 (cuando $\operatorname{Im} f = \{0\}$) o 1 (cuando $\operatorname{Im} f = \mathbb{K}$). Cuando $f = 0_{E^*}$, entonces $\operatorname{Im} f = \{0\}$; cuando $f \neq 0_{E^*}$, entonces $\operatorname{Im} f = \mathbb{K}$ y f es sobre.

Supongamos que E es de dimensión $n \neq 0$. Si $f \neq 0_{E^*}$, el subespacio de E

$$\operatorname{Ker} f = \{x \in E \mid f(x) = 0\}$$

es de dimensión $n-1$, o sea, lo que habíamos llamado un hiperplano de E (v. 1.5.5); se dice que $\operatorname{Ker} f$ es el hiplerplano *de ecuación* $f(x) = 0$. Por otra parte, todo hiperplano H de E admite una ecuación de la forma $f(x) = 0$ con $f \in E^*$ y $f \neq 0$; para demostrar este hecho basta considerar una base $(a_1, \ldots, a_{n-1})$ de H y completarla hasta obtener una base $(a_1, \ldots, a_{n-1}, a_n)$ de E; como existe una forma $f \in E^*$ única tal que

$$f(a_1) = \cdots = f(a_{n-1}) = 0 \quad \text{y} \quad f(a_n) = 1$$

(v. 2.1.6), entonces $H = \operatorname{Ker} f$ y H tiene por ecuación $f(x) = 0$; además $f \neq 0$, puesto que $f(a_n) \neq 0$.

Sea H el hiperplano de E de ecuación $f(x) = 0$. Si $(a_1, \ldots, a_n)$ es una base de E y

$$[f, (a_i), (\,1\,)] = [\alpha_1 \ldots \alpha_n]$$

(v. 2.6.3), entonces H es el conjunto de los vectores

$$x = (x^1, \ldots, x^n)_{(a_i)}$$

de E tales que

$$\alpha_1 x^1 + \cdots + \alpha_n x^n = 0\,.$$

Nótese que, puesto que $f \neq 0$, los α_i no son todos nulos.

2.6.6 Sea E un e.v. y E^* su dual. Si $x \in E$ y $x^* \in E^*$, ponemos

$$\boxed{\langle x, x^* \rangle = x^*(x)}\,,$$

es decir, representamos mediante el símbolo $\langle x, x^* \rangle$ el valor que la forma lineal x^* toma en el vector x, valor que es un escalar (un elemento de $\mathbb{K}$). El símbolo $\langle\ ,\ \rangle$ recibe el nombre de 'corchete de dualidad'. No hemos introducido ningún concepto nuevo, sino una representación distinta de algo que ya conocíamos; el corchete de dualidad permite hacer más manejables ciertas fórmulas complicadas, de las que encontraremos alguna al final de esta sección.

2.6.7 Ejemplo. Siendo $x \in E$ y $f \in E^*$ los del ejemplo (2.6.3), tenemos que

$$\langle x, f \rangle = \alpha_1 x^1 + \cdots + \alpha_n x^n .$$

2.6.8 La aplicación

$$\begin{aligned} E \times E^* &\to \mathbb{K} \\ (x, x^*) &\to \langle x, x^* \rangle \end{aligned}$$

recibe el nombre de *forma bilineal canónica* sobre $E \times E^*$. Esta aplicación verifica las propiedades siguientes:

(dl1) $(\forall x, y \in E)(\forall x^* \in E^*) \quad \langle x + y, x^* \rangle = \langle x, x^* \rangle + \langle y, x^* \rangle$,
(dl2) $(\forall x \in E)(\forall x^*, y^* \in E^*) \quad \langle x, x^* + y^* \rangle = \langle x, x^* \rangle + \langle x, y^* \rangle$,
(dl3) $(\forall \lambda \in \mathbb{K})(\forall x \in E)(\forall x^* \in E^*) \quad \langle \lambda x, x^* \rangle = \lambda \langle x, x^* \rangle$,
(dl4) $(\forall \lambda \in \mathbb{K})(\forall x \in E)(\forall x^* \in E^*) \quad \langle x, \lambda x^* \rangle = \lambda \langle x, x^* \rangle$.

En efecto, **(dl1)** y **(dl3)** no dicen más que los elementos x^* de E^* son formas lineales, y **(dl2)** y **(dl4)** son las definiciones de la suma y del producto externo (producto por un escalar) de E^*. Veremos en el capítulo siguiente que las aplicaciones que cumplen cuatro propiedades como las anteriores se llaman bilineales.

2.6.9 Si en la forma bilineal canónica consideramos fijo un vector x de E, la aplicación que resulta

$$\begin{aligned} E^* &\to \mathbb{K} \\ x^* &\to \langle x, x^* \rangle \end{aligned}$$

es lineal, como consecuencia de **(dl2)** y **(dl4)** o sea, es una forma lineal sobre E^* y por tanto un elemento de E^{**}; vamos a denotarla por $\tilde{x}$. Se tiene, pues, para todo $x^* \in E^*$ que

$$\tilde{x}(x^*) = \langle x, x^* \rangle = x^*(x) .$$

Obsérvese que $\tilde{x}$ es el elemento de E^{**} que obteníamos en el ejemplo (2.6.4).

Podemos considerar ahora la aplicación de E en E^{**} que envía cada $x \in E$ al correspondiente elemento $\tilde{x}$ de E^{**}, obteniendo el resultado que sigue.

2.6.10 PROPOSICIÓN

Sea E un e.v.; la aplicación

$$\begin{aligned} E &\to E^{**} \\ x &\to \tilde{x} \end{aligned}$$

que envía cada $x \in E$ al elemento $\tilde{x} \in E^{**}$ definido por

$$(\forall x^* \in E^*) \quad \tilde{x}(x^*) = \langle x, x^* \rangle$$

es una aplicación lineal. Esta aplicación recibe el nombre de *homomorfismo canónico* de E en E^{**}.

En efecto, las propiedades **(ln1)** y **(ln2)** de (1.1.10) son justamente **(dl1)** y **(dl3)**.

2.6.11 Supongamos ahora que E es un espacio vectorial de dimensión $n \neq 0$ y que $(a_1, \ldots, a_n)$ es una base de E. La proposición (2.1.6) nos asegura que, fijados n escalares $\alpha_1, \ldots, \alpha_n$, existe una forma lineal $x^* \in E^*$ única que verifica

$$\langle a_i, x^* \rangle = \alpha_i \,, \quad i = 1, 2, \ldots, n \,.$$

Esta forma tiene entonces asociada la matriz

$$[x^*, (a_i), (\,1\,)] = [\alpha_1 \ldots \alpha_n]$$

y envía cada vector

$$x = (x^1, \ldots, x^n)_{(a_i)}$$

de E al escalar

$$\begin{aligned} \langle x, x^* \rangle &= \begin{bmatrix} \alpha_1 & \cdots & \alpha_n \end{bmatrix} \begin{bmatrix} x^1 \\ \vdots \\ x^n \end{bmatrix} \\ &= \alpha_1 x^1 + \cdots + \alpha_n x^n \,. \end{aligned}$$

2.6.12 Si fijamos los escalares

$$\alpha_1 = 0 \ \ldots, \alpha_i = 1 \ \ldots, \alpha_n = 0$$

y designamos por a^{*i} la forma que verifica

$$\langle a_1, a^{*i} \rangle = 0 \,, \ldots, \langle a_i, a^{*i} \rangle = 1 \,, \ldots, \langle a_n, a^{*i} \rangle = 0 \,,$$

entonces, para cada vector

$$x = (x^1, \ldots, x^n)_{(a_i)}$$

de E, se tiene

$$\langle x, a^{*i} \rangle = x^i \,,$$

es decir, a^{*i} envía el vector x a su coordenada número i en la base $(a_1, \ldots, a_n)$. Decimos que a^{*i} es la *forma coordenada* número i (respecto de la base $(a_1, \ldots, a_n)$).

Las n formas coordenadas que se definen de esta manera se caracterizan diciendo que $\langle a_j, a^{*i} \rangle$ es 1 cuando $i = j$ y es 0 cuando $i \neq j$, o mejor aún

$$\boxed{\langle a_j, a^{*i} \rangle = \delta_i^j} \,,$$

donde δ_i^j es la delta de Kronecker cuyo significado explicamos en (2.4.16).

La matriz de a^{*i} en las bases $(a_1, \ldots, a_n)$ de E y la base canónica de $\mathbb{K}$ es $[\,0\ \ldots\ 1\ \ldots\ 0]$, o sea, el correspondiente elemento de la base canónica de $M(n, 1)$.

2.6.13 TEOREMA

Sea E un e.v. de dimensión $n \neq 0$ y $(a_1, \ldots, a_n)$ una base de E. Las formas coordenadas $a^{*i} \in E^*$ definidas por

$$\langle a_j, a^{*i} \rangle = \delta_i^j$$

forman una base de E^*. La base $(a^{*1}, \ldots, a^{*n})$ recibe el nombre de *base dual* de la base $(a_1, \ldots, a_n)$ de E. Por lo tanto

$$\dim E^* = \dim E\,.$$

Si $x^* \in E^*$ y $x^* = (\alpha_1, \ldots, \alpha_n)_{(a^{*i})}$, la matriz de x^* en la base $(a_1, \ldots, a_n)$ de E y la base canónica de $\mathbb{K}$ es

$$[x^*, (a_i), (\,1\,)] = [\alpha_1 \ldots \alpha_n]\,.$$

Además, si $x \in E$ y $x^* \in E^*$, con

$$x = (x^1, \ldots, x^n)_{(a_i)}\,, \quad x^* = (\alpha_1, \ldots, \alpha_n)_{(a^{*i})}\,,$$

entonces

$$\langle x, x^* \rangle = x^1\alpha_1 + \cdots + x^n\alpha_n\,.$$

Por el isomorfismo

$$\begin{array}{lcl} E^* = \mathcal{L}(E, \mathbb{K}) & \to & M(n,1) \\ x^* & \to & [x^*, (a_i), (\,1\,)]\,, \end{array}$$

las formas coordenadas a^{*i} se corresponden con los elementos E_i^1 de la base canónica de $M(n,1)$, luego, aplicando (1.4.37), resulta que forman una base de E^*.

Las restantes afirmaciones se deducen fácilmente.

2.6.14 COROLARIO

Sea E un e.v. de dimensión finita y $x \in E$. Si todo elemento de E^* se anula en x, es decir, si

$$(\forall x^* \in E^*) \quad \langle x, x^* \rangle = 0\,,$$

entonces $x = 0$.

Si $E = \{0\}$, el resultado es claro; si ese no es el caso, sea $(a_1, \ldots, a_n)$ una base de E y $(a^{*1}, \ldots, a^{*n})$ la base dual. Si

$$x = (x^1, \ldots, x^n)_{(a_i)}\,,$$

entonces, como

$$(\forall i) \quad \langle x, a^{*i} \rangle = 0\,,$$

resulta que

$$(\forall i) \quad x^i = 0\,,$$

o sea, que $x = 0$.

2.6.15 El resultado precedente es cierto aunque E no tenga dimensión finita. Sin embargo, nosotros carecemos de argumentos para demostrarlo en ese caso.

2.6.16 Sea E un e.v. de dimensión $n \neq 0$, $(a_1, \ldots, a_n)$ una base de E y $(a^{*1}, \ldots, a^{*n})$ la base dual. Si $x^* \in E^*$ y

$$x^* = (\alpha_1, \ldots, \alpha_n)_{(a^{*i})} ,$$

entonces para cada $i = 1, 2, \ldots, n$

$$\langle a_i, x^* \rangle = \alpha_i ,$$

es decir, a_i envía toda forma $x^* \in E^*$ a su coordenada número i en la base $(a^{*1}, \ldots, a^{*n})$.

2.6.17 TEOREMA

Si E es un e.v. de dimensión finita, el homomorfismo canónico

$$\begin{aligned} E &\to E^{**} \\ x &\to \tilde{x} \end{aligned}$$

(v. 2.6.10) es un isomorfismo.

Veamos, en primer lugar, que es una aplicación inyectiva. Si $\tilde{x} = 0_{E^{**}}$, entonces

$$(\forall x^* \in E^*) \quad \tilde{x}(x^*) = 0 ,$$

es decir

$$(\forall x^* \in E^*) \quad \langle x, x^* \rangle = 0$$

y entonces $x = 0$ (v. 2.6.14). El núcleo de la aplicación se reduce a $\{0\}$ y esto significa (v. 2.1.11) que la aplicación es inyectiva.

Por otro lado

$$\dim E^{**} = \dim E^* = \dim E$$

luego (v. 2.1.24) la aplicación es un isomorfismo.

2.6.18 COROLARIO

Sea E un e.v. de dimensión $n \neq 0$, E^* su dual y $(a^{*1}, \ldots, a^{*n})$ una base de E^*. Existe una base $(a_1, \ldots, a_n)$ de E tal que su base dual es $(a^{*1}, \ldots, a^{*n})$.

Consideremos en E^{**} la base dual de $(a^{*1}, \ldots, a^{*n})$, o sea, el sistema $(u_1, \ldots, u_n)$ de E^{**} tal que

$$(\forall i, j) \quad u_j(a^{*i}) = \delta_i^j .$$

Por el teorema anterior, existen $a_1, \ldots, a_n \in E$ (únicos) tales que

$$\tilde{a}_j = u_j \, , \quad j = 1, 2, \ldots, n \, .$$

Entonces $(a_1, \ldots, a_n)$ es una base de E y

$$(\forall i, j) \quad \langle a_j \, , a^{*i} \rangle = u_j(a^{*i}) = \delta_i^j \, ,$$

luego $(a^{*1}, \ldots, a^{*n})$ es la base dual de $(a_1, \ldots, a_n)$.

2.6.19 Ejemplo. Consideremos el espacio $\mathbb{K}^n$ y la base canónica $(e_1, \ldots, e_n)$; la base dual de esta, $(e^{*1}, \ldots, e^{*n})$, la constituyen las formas sobre $\mathbb{K}^n$ dadas por

$$\begin{aligned} e^{*i} : \qquad \mathbb{K}^n &\to \mathbb{K} \\ (x^1, \ldots, x^n) &\to x^i \, . \end{aligned}$$

También es frecuente representar e^{*i} por π_i y llamarle 'proyección número i de $\mathbb{K}^n$'. La proyección π_i envía, pues, cada elemento $(x^1, \ldots, x^n)$ de $\mathbb{K}^n$ a su componente número i, x^i.

Vamos a representar por $(\alpha_1, \ldots, \alpha_n)^*$ el elemento de $(\mathbb{K}^n)^*$

$$(\alpha_1, \ldots, \alpha_n)_{(\pi_i)} \, .$$

Es una manera de identificar $(\mathbb{K}^n)^*$ con $\mathbb{K}^n$. Con esta notación,

$$\pi_i = (0, \ldots, 1, \ldots, 0)^* \, ,$$

donde 1 ocupa el lugar número i.

Como hemos visto en (2.6.13), se verifica

$$\langle (x^1, \ldots, x^n), (\alpha_1, \ldots, \alpha_n)^* \rangle = x^1 \alpha_1 + \cdots + x^n \alpha_n \, .$$

Esta fórmula recordará a quienes lo conozcan un producto escalar, lo que no es extraño puesto que, en un espacio de dimensión finita con producto escalar, cada vector se puede identificar, mediante una fórmula como la precedente, con una forma lineal.

2.6.20 Sea E un e.v., E' un e.v. de dimensión $m \neq 0$ y $(b_1, \ldots, b_m)$ una base de E'. Si $f_1, f_2, \ldots, f_m \in E^*$ son formas lineales sobre E, entonces la aplicación

$$\begin{aligned} E &\to E' \\ x &\to f_1(x) b_1 + \cdots + f_m(x) b_m \end{aligned}$$

es una aplicación lineal.

Recíprocamente, cualquier aplicación lineal

$$f : E \to E'$$

se puede construir de esa manera. En efecto, siendo $(b^{*1},\ldots,b^{*m})$ la base dual de la $(b_1,\ldots,b_m)$, las aplicaciones compuestas $f_j = b^{*j}\circ f$

$$E \xrightarrow{f} E' \xrightarrow{b^{*j}} \mathbb{K}$$
$$\underrightarrow{\quad b^{*j}\circ f \quad}$$

con $j = 1,2,\ldots,n$, son formas lineales sobre E, y como

$$f(x) = (b^{*1}\circ f(x),\ldots,b^{*m}\circ f(x)\,)_{(b_j)}$$

resulta que

$$f(x) = f_1(x)b_1 + \cdots + f_m(x)b_m\,.$$

2.6.21 Ejemplo. Sea E un e.v. de dimensión $n \neq 0$, $(a_1,\ldots,a_n)$ una base de E y $(a^{*1},\ldots,a^{*n})$ la base dual. Escogiendo r formas coordenadas $a^{*i_1},\ldots,a^{*i_r}$ (con $i_1,\ldots,i_r \in \{1,2,\ldots,n\}$), la aplicación

$$\begin{aligned} f : \qquad\qquad E &\to \mathbb{K}^r \\ (x^1,\ldots,x^n)_{(a_i)} &\to (x^{i_1},\ldots,x^{i_r}) \end{aligned}$$

es una aplicación lineal, pues envía a cada $x \in E$ a

$$f(x) = a^{*i_1}(x)e_1 + \cdots + a^{*i_r}(x)e_r\,,$$

donde $(e_1,\ldots,e_r)$ es la base canónica de $\mathbb{K}^r$.

2.6.22 DEFINICIÓN. Sea E un e.v. y E^* su dual. Sean $x \in E$ y $x^* \in E^*$; cuando $\langle x, x^*\rangle = 0$, decimos que x y x^* son *ortogonales.*

Si $A \subset E$ (A no es necesariamente un subespacio), el subconjunto de E^*

$$A^{\perp} = \{x^* \in E^* | (\forall x \in A) \quad \langle x, x^*\rangle = 0\}\,,$$

formado por los elementos x^* de E^* que son ortogonales a todos los elementos de A, es un subespacio de E^*; decimos que es el *subespacio ortogonal* a A.

Análogamente, si $A^* \subset E^*$, el subconjunto de E

$$A^{*\perp} = \{x \in E | (\forall x^* \in A^*) \quad \langle x, x^*\rangle = 0\}$$

es un subespacio de E; decimos que es el *subespacio ortogonal* a A^*.

2.6.23 Ejemplo. Si $x^* \in E^*$, entonces

$$\{x^*\}^{\perp} = \operatorname{Ker} x^*\,.$$

2.6.24 Ejemplo. Es fácil probar que

$$E^{\perp} = \{0_{E^*}\}\,, \quad \{0_E\}^{\perp} = E^* \quad \text{y} \quad \{0_{E^*}\}^{\perp} = E\,.$$

También es cierto que

$$E^{*\perp} = \{0_E\}\,,$$

aunque nosotros no tenemos argumentos para demostrarlo en el caso general. Cuando E es de dimensión finita, este resultado es una consecuencia sencilla de (2.6.14).

2.6.25 Ejemplo. Con las notaciones del ejemplo (2.6.19), si consideramos el subconjunto de $\mathbb{R}^3$

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

resulta que

$$A^{\perp} = \{(\alpha_1, \alpha_2, \alpha_3)^* \in (\mathbb{R}^3)^* \mid \alpha_1 + \alpha_2 = 0\}\,.$$

2.6.26 PROPOSICIÓN

Sea E un e.v. de dimensión finita.

a) Si F es un subespacio de E, entonces

$$\dim F^{\perp} = \dim E - \dim F\,.$$

b) Si F^* es un subespacio de E^*, entonces

$$\dim F^{*\perp} = \dim E^* - \dim F^*\,.$$

---

a) Si $F = \{0\}$ o $F = E$, el resultado es obvio aplicando (2.6.24). Supongamos entonces que F es un subespacio propio de E; llamemos p a la dimensión de F y n a la de E y E^*. Sea $(a_1, \ldots, a_p)$ una base de F y $(a_1, \ldots, a_p, a_{p+1}, \ldots, a_n)$ una base de E que prolonga dicha base de F; consideremos además la base dual $(a^{*1}, \ldots, a^{*p}, a^{*p+1}, \ldots, a^{*n})$. Veamos que

$$F^{\perp} = \langle a^{*p+1}, \ldots, a^{*n} \rangle\,,$$

lo que terminará la demostración, puesto que entonces

$$\dim F^{\perp} = n - p = \dim E - \dim F\,.$$

En primer lugar

$$a^{*p+1}, \ldots, a^{*n} \in F^{\perp}$$

ya que, si $x \in F$, entonces

$$x = x^1 a_1 + \cdots + x^p a_p\,,$$

luego

$$\langle x, a^{*p+1} \rangle = x^1 \langle a_1, a^{*p+1} \rangle + \cdots + x^p \langle a_p, a^{*p+1} \rangle = 0$$

y lo mismo ocurre para $a^{*p+2}, \ldots, a^{*n}$.

Además, si $x^* \in F^{\perp}$ y

$$x^* = \alpha_1 a^{*1} + \cdots \alpha_p a^{*p} + \alpha_{p+1} a^{*p+1} + \cdots \alpha_n a^{*n} ,$$

como x^* verifica en particular que

$$\langle a_1, x^* \rangle = \langle a_2, x^* \rangle = \cdots = \langle a_p, x^* \rangle = 0 ,$$

resulta que

$$\alpha_1 = \alpha_2 = \cdots = \alpha_p = 0 ,$$

o sea,

$$x^* = \alpha_{p+1} a^{*p+1} + \cdots + \alpha_n a^{*n}$$

y esto significa entonces que

$$F^{\perp} = \langle a^{*p+1}, \ldots, a^{*n} \rangle .$$

b) También esta parte es evidente si $F^* = \{0\}$ o $F^* = E^*$. Supongamos entonces que F^* es un subespacio propio de E^* y pongamos $p = \dim F^*$ y $n = \dim E = \dim E^*$. Consideremos entonces una base $(a^{*1}, \ldots, a^{*p})$ de F^* y una base $(a^{*1}, \ldots, a^{*p}, a^{*p+1}, \ldots, a^{*n})$ de E^* que prolonga aquella. Esta base de E^* es la base dual de una base $(a_1, \ldots, a_p, a_{p+1}, \ldots, a_n)$ de E (v. 2.6.18). Se prueba como en el apartado anterior que

$$F^{*\perp} = \langle a_{p+1}, \ldots, a_n \rangle$$

y que esto termina la demostración.

2.6.27 DEFINICIÓN. Sea E y E' dos e.v. sobre $\mathbb{K}$ y E^* y E'^* sus duales; sea $f \in \mathcal{L}(E, E')$. Vamos a definir entonces una aplicación de E'^* en E^*.

Para cada una de las formas $y^* \in E'^*$, podemos considerar la aplicación compuesta $y^* \circ f$

$$\underbrace{E \xrightarrow{f} E' \xrightarrow{y^*}}_{y^* \circ f} \mathbb{K}$$

que será una forma lineal sobre E, por ser compuesta de dos aplicaciones lineales; es decir, $y^* \circ f \in E^*$.

La aplicación

$$\begin{aligned} E'^* &\to E^* \\ y^* &\to y^* \circ f , \end{aligned}$$

que envía cada forma $y^* \in E'^*$ a la forma $y^* \circ f$ de E^*, recibe el nombre de *traspuesta* de f y se representa por f^t. Se tiene, pues, que

$$(\forall y^* \in E'^*) \quad f^t(y^*) = y^* \circ f .$$

2.6.28 Ejemplo. Para la aplicación idéntica, id_E , de un e.v. E, la traspuesta

$$id_E^t : E^* \to E^*$$

envía cada forma $y^* \in E^*$ a la forma

$$id_E^t(y^*) = y^* \circ id_E = y^* \,.$$

Entonces

$$id_E^t = id_{E^*} \,.$$

2.6.29 La igualdad

$$f^t(y^*) = y^* \circ f$$

significa que, para todo vector x de E,

$$\langle x, f^t(y^*) \rangle = y^* \circ f(x) = y^*(\, f(x)\,) = \langle f(x), y^* \rangle \,.$$

La expresión

$$\boxed{(\forall x \in E)(\forall y^* \in E'^*) \quad \langle x, f^t(y^*) \rangle = \langle f(x), y^* \rangle}$$

es la que permite manejar f^t con mayor sencillez. Obsérvese que el corchete de dualidad que aparece en el primer miembro se refiere a E y E^*, mientras que el del segundo miembro se refiere a E' y E'^*.

2.6.30 PROPOSICIÓN

Sea $f \in \mathcal{L}(E, E')$, donde E y E' son e.v. sobre $\mathbb{K}$, y

$$f^t : E'^* \to E^*$$

su aplicación traspuesta. Entonces f^t es una aplicación lineal, es decir, $f^t \in \mathcal{L}(E'^*, E^*)$.

Tenemos que probar que, si $y^*, z^* \in E'^*$ y $\lambda \in \mathbb{K}$, se cumplen las dos igualdades

$$f^t(y^* + z^*) = f^t(y^*) + f^t(z^*) \,,$$
$$f^t(\lambda y^*) = \lambda f^t(y^*) \,.$$

En el primer caso, lo que figura en ambos miembros son elementos de E^*, es decir, aplicaciones de E en $\mathbb{K}$. Para probar que coinciden habrá que comprobar que

$$(\forall x \in E) \quad \langle x, f^t(y^* + z^*) \rangle = \langle x, f^t(y^*) + f^t(z^*) \rangle \,.$$

Para todo $x \in E$,

$$\begin{aligned} \langle x, f^t(y^* + z^*) \rangle &= \langle f(x), y^* + z^* \rangle \\ &= \langle f(x), y^* \rangle + \langle f(x), z^* \rangle \\ &= \langle x, f^t(y^*) \rangle + \langle x, f^t(z^*) \rangle \\ &= \langle x, f^t(y^*) + f^t(z^*) \rangle \,, \end{aligned}$$

lo que demuestra la igualdad.

La segunda igualdad se prueba haciendo el mismo razonamiento.

2.6.31 Ejemplo. Volvemos a utilizar las notaciones del ejemplo (2.6.19). Sea $f : \mathbb{K}^3 \to \mathbb{K}^2$ una aplicación lineal; la aplicación traspuesta

$$f^t : (\mathbb{K}^2)^* \to (\mathbb{K}^3)^*$$

envía cada elemento $(\alpha, \beta)^* \in (\mathbb{K}^2)^*$ al elemento $f^t(\,(\alpha,\beta)^*\,)$ de $(\mathbb{K}^3)^*$ que verifica

$$(\forall (x,y,z) \in \mathbb{K}^3\,) \quad \langle (x,y,z), f^t(\,(\alpha,\beta)^*\,)\rangle = \langle f(x,y,z), (\alpha,\beta)^*\rangle \,.$$

Si, por ejemplo, f viene dada por

$$f(x,y,z) = (x+z, y+z)\,,$$

entonces, para todo $(x,y,z) \in \mathbb{K}^3$,

$$\begin{aligned}
\langle (x,y,z), f^t(\,(\alpha,\beta)^*\,)\rangle &= \langle f(x,y,z), (\alpha,\beta)^*\rangle \\
&= \langle (x+y, y+z), (\alpha,\beta)^*\rangle \\
&= (x+z)\alpha + (y+z)\beta \\
&= x\alpha + y\beta + z(\alpha+\beta) \\
&= \langle (x,y,z), (\alpha,\beta,\alpha+\beta)^*\rangle\,,
\end{aligned}$$

es decir

$$f^t(\,(\alpha,\beta)^*\,) = (\alpha,\beta,\alpha+\beta)^*\,.$$

2.6.32 PROPOSICIÓN

Sea $f \in \mathcal{L}(E, E')$, donde E y E' son e.v. sobre $\mathbb{K}$; sea $f^t \in \mathcal{L}(E'^*, E^*)$ su traspuesta. $\operatorname{Ker} f^t$ es un subespacio de E'^*; también lo es $(\operatorname{Im} f)^\perp$, puesto que $\operatorname{Im} f$ es un subespacio de E'. Se tiene que

$$(\operatorname{Im} f)^\perp = \operatorname{Ker} f^t\,.$$

En efecto,

$$\begin{aligned}
(\operatorname{Im} f)^\perp &= \{y^* \in E'^* | (\forall y \in \operatorname{Im} f) \quad \langle y, y^*\rangle = 0\} \\
&= \{y^* \in E'^* | (\forall x \in E) \quad \langle f(x), y^*\rangle = 0\} \\
&= \{y^* \in E'^* | (\forall x \in E) \quad \langle x, f^t(y^*)\rangle = 0\} \\
&= \{y^* \in E'^* | f^t(y^*)\rangle = 0\} \\
&= \operatorname{Ker} f^t\,.
\end{aligned}$$

2.6.33 COROLARIO

Sean E y E' dos e.v. sobre $\mathbb{K}$, ambos de dimensión finita; sea $f \in \mathcal{L}(E, E')$ y f^t su traspuesta. Entonces

$$\operatorname{rg} f^t = \operatorname{rg} f\,.$$

Utilizando los resultados que se indican al margen, tenemos las sucesivas igualdades

$$\begin{aligned}
\operatorname{rg} f^t &= \dim E'^* - \dim \operatorname{Ker} f^t && \text{(v. 2.1.20)}\\
&= \dim E' - \dim \operatorname{Ker} f^t && \text{(v. 2.6.13)}\\
&= \dim E' - \dim(\operatorname{Im} f)^\perp && \text{(v. 2.6.32)}\\
&= \dim E' - (\dim E' - \dim \operatorname{Im} f) && \text{(v. 2.6.26)}\\
&= \dim \operatorname{Im} f\\
&= \operatorname{rg} f\,.
\end{aligned}$$

2.6.34 PROPOSICIÓN

Sean E y E' dos e.v. de dimensiones $n \neq 0$ y $m \neq 0$, $(a_1, \ldots, a_n)$ una base de E y $(a^{*1}, \ldots, a^{*n})$ la correspondiente base dual de E^*, $(b_1, \ldots, b_m)$ una base de E' y $(b^{*1}, \ldots, b^{*m})$ la correspondiente base dual de E'^*. Sea $f \in \mathcal{L}(E, E')$ y $f^t \in \mathcal{L}(E'^*, E^*)$ su traspuesta. Se cumple la siguiente relación entre las matrices asociadas (v. 2.2.17) a f y f^t :

$$[f^t, (b^{*j}), (a^{*i})] = [f, (a_i), (b_j)]^t\,.$$

Pongamos $[f, (a_i), (b_j)] = [\alpha_i^j]$ y $[f^t, (b^{*j}), (a^{*i})] = [\beta_i^j]$. Si $1 \leq i \leq n$ y $1 \leq j \leq m$, aplicando la fórmula de (2.6.29) tenemos que

$$\langle a_i, f^t(b^{*j})\rangle = \langle f(a_i), b^{*j}\rangle\,.$$

El escalar $\langle f(a_i), b^{*j}\rangle$ es la coordenada número j en la base $(b_1, \ldots, b_m)$ del vector $f(a_i)$ (v. 2.6.12), o sea, α_i^j. El escalar $\langle a_i, f^t(b^{*j})\rangle$ es la coordenada número i en la base $(a^{*1} \ldots, a^{*n})$ de la forma $f^t(b^{*j})$ (v. 2.6.16), o sea, β_j^i. Resulta pues que

$$(\forall i, j) \quad \beta_j^i = \alpha_i^j\,,$$

y por lo tanto la conclusión.

2.6.35 Ejemplo. En el ejemplo (2.6.31), la matriz asociada a f^t en las bases duales de las canónicas es

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 1 \end{bmatrix},$$

justamente la traspuesta de la matriz

$$\begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix},$$

asociada a f y a las bases canónicas de $\mathbb{K}^3$ y $\mathbb{K}^2$.

2.6.36 COROLARIO

Sea E un e.v. de dimensión $n \neq 0$, $(a_1, \ldots, a_n)$ y $(a'_1, \ldots, a'_n)$ dos bases de E y $(a^{*1}, \ldots, a^{*n})$ y $(a'^{*1}, \ldots, a'^{*n})$ las correspondientes bases duales de E^*. Se cumple la siguiente relación entre las matrices de cambio de bases de E y E^*:

$$[(a'^{*i}), (a^{*i})] = ([(a'_i), (a_i)]^t)^{-1}.$$

En efecto, si tenemos en cuenta que $id_{E^*} = id_E^t$ (v. 2.6.28), entonces, aplicando (2.5.11) y la proposición precedente,

$$\begin{aligned} [(a'^{*i}), (a^{*i})] &= [(a^{*i}), (a'^{*i})]^{-1} \\ &= [id_{E^*}, (a^{*i}), (a'^{*i})]^{-1} \\ &= ([id_E, (a'_i), (a_i)]^t)^{-1} \\ &= ([(a'_i), (a_i)]^t)^{-1}. \end{aligned}$$

2.6.37 PROPOSICIÓN

Sea $A \in M(n, m)$ y $A^t \in M(m, n)$ su traspuesta; entonces

$$\operatorname{rg} A^t = \operatorname{rg} A.$$

Designemos por f la aplicación lineal $f : \mathbb{K}^n \to \mathbb{K}^m$ dada por A (v. 2.5.1), para la cual

$$A = [f, (e_i), (\overline{e_j})],$$

siendo $(e_1, \ldots, e_n)$ y $(\overline{e_1}, \ldots, \overline{e_m})$ las bases canónicas de $\mathbb{K}^n$ y $\mathbb{K}^m$. Para su traspuesta, $f^t : (\mathbb{K}^m)^* \to (\mathbb{K}^n)^*$, tenemos

$$A^t = [f^t, (\overline{e}^{*j}), (e^{*i}].$$

Utilizando (2.6.33) y (2.5.22), resulta que

$$\operatorname{rg} A^t = \operatorname{rg} f^t = \operatorname{rg} f = \operatorname{rg} A.$$

2.6.38 Como los vectores columna de A^t son los vectores fila de A, el resultado precedente significa que el rango de A, es decir, el rango del sistema formado por sus n vectores columna, es también el rango del sistema formado por sus m vectores fila.

2.6.39 **Nota.** El lector se puede sentir abrumado por la diversidad de notaciones empleadas en este capítulo, sobre todo en lo que se refiere a la representación de las distintas matrices que aparecen en el texto. Lo esencial no es sin embargo el conocimiento de estas notaciones, sino el manejo de los conceptos que encierran; éste se adquiere cuando se resuelven bastantes problemas, de éste o de otros libros. Evidentemente, cualquier notación es arbitraria y se puede substituir por otra, siempre que quede bien claro cuál es el objeto que se representa. Las bases son sin duda un instrumento eficaz en el tratamiento de los espacios vectoriales de dimensión finita, pero también llevan a notaciones más complejas al tener que figurar sistemáticamente en ellas; ¡ no hay beneficio sin servidumbre!

Capítulo 3

Determinantes.

Solamente resultará novedosa para el lector la introducción del determinante como forma n-lineal alternada, basada en el teorema 3.1.25, que le exigirá sin duda un cierto esfuerzo. Se emplea con profusión el símbolo $\sum$, cuyo manejo hemos explicado en 1.1.7 . Llegado el caso, razónese, empleando los argumentos que allí expusimos, que se utiliza correctamente.

3.1 Formas n-lineales alternadas.

3.1.1 Si

$$A = \begin{bmatrix} \alpha_1^1 & \alpha_2^1 \\ \alpha_1^2 & \alpha_2^2 \end{bmatrix} \quad \text{y} \quad B = \begin{bmatrix} \alpha_1^1 & \alpha_2^1 & \alpha_3^1 \\ \alpha_1^2 & \alpha_2^2 & \alpha_3^2 \\ \alpha_1^3 & \alpha_2^3 & \alpha_3^3 \end{bmatrix}$$

son matrices $A \in M(2)$ y $B \in M(3)$, se suele escribir

$$\det A = \alpha_1^1\alpha_2^2 - \alpha_1^2\alpha_2^1\,,$$

$$\det B = \alpha_1^1\alpha_2^2\alpha_3^3 + \alpha_1^2\alpha_2^3\alpha_3^1 + \alpha_1^3\alpha_2^1\alpha_3^2 - \alpha_1^1\alpha_2^3\alpha_3^2 - \alpha_1^2\alpha_2^1\alpha_3^3 - \alpha_1^3\alpha_2^2\alpha_3^1\,.$$

$\det A$ y $\det B$ son elementos del cuerpo $\mathbb{K}$; se dice que son los 'determinantes' de las matrices A y B. Esto será, sin duda, familiar para el lector.

En la expresión que nos permite calcular $\det A$ figuran dos términos, cada uno de ellos producto de dos elementos de A; si observamos los índices superiores (de fila) de los elementos de estos términos, vemos que constituyen las dos permutaciones $(1,2)$ y $(2,1)$ del conjunto $\{1,2\}$ y que el término precedido del signo '$-$' corresponde a la permutación impar.

Obsérvese que sucede lo mismo en la expresión de $\det B$, donde aparecen las 6 permutaciones del conjunto $\{1,2,3\}$, con el mismo criterio de la paridad a la hora de anteponer el signo '$-$' a los tres últimos términos.

Vamos a mantener el mismo criterio para definir el determinante de una matriz $A = [\alpha_i^j] \in M(n)$; este determinante será una suma de $n!$ términos (recuérdese que $n!$ es el numero de elementos de G_n), todos ellos de la forma

$$\epsilon(p)\, \alpha_1^{p(1)}\alpha_2^{p(2)}\cdots\alpha_n^{p(n)}\,,$$

donde $p \in G_n$ y $\epsilon(p)$ es la paridad de p.

3.1.2 Si $x_1 = (x_1^1, x_1^2)$ y $x_2 = (x_2^1, x_2^2)$ son dos vectores de $\mathbb{K}^2$, podemos poner

$$\det(x_1, x_2) = \det \begin{bmatrix} x_1^1 & x_2^1 \\ x_1^2 & x_2^2 \end{bmatrix} .$$

Se verifican entonces las siguientes propiedades que el lector probablemente conocerá:

$$\begin{array}{ll} \text{(a)} & \det(x_1 + x_1', x_2) = \det(x_1, x_2) + \det(x_1', x_2)\,, \\ \text{(b)} & \det(x_1, x_2 + x_2') = \det(x_1, x_2) + \det(x_1, x_2')\,, \\ \text{(c)} & \det(\lambda x_1, x_2) = \lambda \det(x_1, x_2)\,, \\ \text{(d)} & \det(x_1, \lambda x_2) = \lambda \det(x_1, x_2)\,. \end{array}$$

(Se suele decir que la aplicación

$$\begin{aligned} \det : \mathbb{K}^2 \times \mathbb{K}^2 &\to \mathbb{K} \\ (x_1, x_2) &\to \det(x_1, x_2) \end{aligned}$$

es una forma bilineal.) Además, si $x \in \mathbb{K}^2$, se tiene que

$$\text{(e)} \qquad \det(x, x) = 0\,.$$

(Se dice entonces que la forma bilineal 'det' es alternada.) Finalmente, si (e_1, e_2) es la base canónica de $\mathbb{K}^2$,

$$\text{(f)} \qquad \det(e_1, e_2) = \det \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = 1\,.$$

Lo mismo podemos decir de la aplicación

$$\begin{aligned} \det : \mathbb{K}^3 \times \mathbb{K}^3 \times \mathbb{K}^3 &\to \mathbb{K} \\ (x_1, x_2, x_3) &\to \det(x_1, x_2, x_3) \end{aligned}$$

definida, con $x_1 = (x_1^1, x_1^2, x_1^3)$, $x_2 = (x_2^1, x_2^2, x_2^3)$ y $x_3 = (x_3^1, x_3^2, x_3^3)$, por

$$\det(x_1, x_2, x_3) = \det \begin{bmatrix} x_1^1 & x_2^1 & x_3^1 \\ x_1^2 & x_2^2 & x_3^2 \\ x_1^3 & x_2^3 & x_3^3 \end{bmatrix} .$$

Las cuatro primeras propiedades se convierten en este caso en seis. La propiedad (e) se convierte en

$$x_1 = x_2 \text{ o } x_1 = x_3 \text{ o } x_2 = x_3 \quad \Rightarrow \quad \det(x_1, x_2, x_3) = 0\,.$$

También vamos a generalizar esta noción de determinante. Si E es un e.v. de dimensión $n \neq 0$ y $(a_1, \ldots, a_n)$ una base de E, el determinante relativo a dicha base será una aplicación

$$\begin{aligned} E \times \cdots \times E &\to \mathbb{K} \\ (x_1, \ldots, x_n) &\to \det_{(a_i)}(x_1, \ldots, x_n) \end{aligned}$$

que verificará propiedades análogas a las que acabamos de ver. Demostraremos que existe una aplicación de estas características y que además es única.

3.1.3 DEFINICIÓN. Sea E un e.v. sobre un cuerpo conmutativo $\mathbb{K}$; consideremos el producto cartesiano E^n de E por sí mismo n veces. Los elementos de E^n son los sistemas $(x_1, \ldots, x_n)$ de n vectores de E (v. 1.2.11). No nos interesamos aquí por la estructura de e.v. producto de E^n (v. 1.2.1), sino que haremos intervenir únicamente la estructura de e.v. de E.

Sea

$$\begin{aligned} f: \quad E^n &\to \mathbb{K} \\ (x_1, \ldots, x_n) &\to f(x_1, \ldots, x_n) \end{aligned}$$

una aplicación que hace corresponder a cada sistema $(x_1, \ldots, x_n)$ de n vectores de E un elemento de $\mathbb{K}$. Decimos que f es una *forma n-lineal* sobre E si, para cada índice $i = 1, \ldots, n$, se verifica

(n-ln1) $f(x_1, \ldots, x_i + x_i', \ldots, x_n) =$

$$= f(x_1, \ldots, x_i, \ldots, x_n) + f(x_1, \ldots, x_i', \ldots, x_n)\,,$$

(n-ln2) $f(x_1, \ldots, \lambda x_i, \ldots, x_n) = \lambda f(x_1, \ldots, x_i, \ldots, x_n)\,,$

cualesquiera que sean $\lambda \in \mathbb{K}$ y $x_1, \ldots, x_i, x_i', \ldots, x_n \in E$.

La definición precedente exige, pues, el cumplimiento de dos propiedades para cada uno de los índices $i = 1, \ldots, n$, es decir, de un total de $2n$ propiedades. Compárese esta definición con la de aplicación lineal (v. 1.1.10); las relaciones entre ambas se explican en (3.1.11) y (3.1.5).

3.1.4 Ejemplo. Las dos aplicaciones representadas por det en (3.1.2) son, la primera, una forma 2-lineal, y la segunda, una forma 3-lineal.

3.1.5 Ejemplo. Las formas lineales sobre E, o sea, las aplicaciones lineales de E en $\mathbb{K}$, son justamente las formas 1-lineales sobre E.

3.1.6 Ejemplo. Una aplicación

$$\begin{aligned} f: \quad E^2 &\to \mathbb{K} \\ (x_1, x_2) &\to f(x_1, x_2) \end{aligned}$$

es una forma 2-lineal, o mejor una *forma bilineal*, cuando verifica las 4 $(= 2 \times 2)$ propiedades siguientes:

(bln1) $(\forall x_1, x_1', x_2 \in E) \quad f(x_1 + x_1', x_2) = f(x_1, x_2) + f(x_1', x_2)\,,$
(bln2) $(\forall x_1, x_2, x_2' \in E) \quad f(x_1, x_2 + x_2') = f(x_1, x_2) + f(x_1, x_2')\,,$
(bln3) $(\forall \lambda \in \mathbb{K})(\forall x_1, x_2 \in E) \quad f(\lambda x_1, x_2) = \lambda f(x_1, x_2)\,,$
(bln4) $(\forall \lambda \in \mathbb{K})(\forall x_1, x_2 \in E) \quad f(x_1, \lambda x_2) = \lambda f(x_1, x_2)\,.$

3.1.7 Ejemplo. La forma bilineal canónica de (2.6.8) cumple cuatro propiedades como las precedentes (de ahí su nombre), aunque el conjunto de partida no sea $E^2 = E \times E$, sino $E \times E^*$. La definición (3.1.3) se extiende sin cambios al caso de aplicaciones

$$f : E_1 \times E_2 \times \cdots \times E_n \to E' ,$$

donde $E_1, E_2, \ldots, E_n, E'$ son espacios vectoriales cualesquiera (sobre el mismo cuerpo), dando lugar a lo que se llama 'aplicaciones n-lineales'; prescindiremos del estudio general de este tipo de aplicaciones, que no necesitaremos utilizar en el presente libro, excepción hecha de la forma bilineal canónica.

3.1.8 Ejemplo. Las aplicaciones

$$f, g, h : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$$

definidas por

$$\begin{aligned} f(\,(x,y),(x',y')\,) &= xx' + yy' , \\ g(\,(x,y),(x',y')\,) &= xy' - yx' , \\ h(\,(x,y),(x',y')\,) &= 2xy' + yx' , \end{aligned}$$

son formas bilineales sobre $\mathbb{R}^2$, como se comprueba de forma conceptualmente sencilla (aunque laboriosa). La segunda, g, es la que en (3.1.2) representábamos por det.

3.1.9 Ejemplo. Una forma n-lineal

$$\begin{aligned} f : \qquad E^n &\to \mathbb{K} \\ (x_1, \ldots, x_n) &\to f(x_1, \ldots, x_n) \end{aligned}$$

es una aplicación de n 'variables'. Si $E = \mathbb{K}^m$, cada una de las variables x_i posee a su vez m componentes; de manera mas general, si E es un e.v. de dimensión m y $(a_1, \ldots, a_m)$ una base de E, cada variable x_i posee m coordenadas en la base $(a_1, \ldots, a_m)$. En el ejemplo precedente, $n = m = 2$; de hecho, las formas que nos interesan en el presente capítulo son las formas n-lineales sobre un espacio de dimensión n, puesto que las aplicaciones det son, como vimos en (3.1.2) de este tipo.

3.1.10 Ejemplo. Sea

$$\begin{aligned} f : \qquad E^2 &\to \mathbb{K} \\ (x_1, x_2) &\to f(x_1, x_2) \end{aligned}$$

una forma bilineal sobre E. Entonces

$$\begin{aligned} f(x_1 + x_1', x_2 + x_2') &= f(x_1, x_2 + x_2') + f(x_1', x_2 + x_2') \\ &= f(x_1, x_2) + f(x_1, x_2') + f(x_1', x_2) + f(x_1', x_2') \, ; \end{aligned}$$

es decir, si las dos variables consisten en la suma de dos vectores, la imagen podemos ponerla como suma de 4 $(= 2 \times 2)$ escalares. De manera más general, si

$$\begin{aligned} f: \quad E^n &\to \mathbb{K} \\ (x_1, \ldots, x_n) &\to f(x_1, \ldots, x_n) \end{aligned}$$

es una forma n-lineal y cada uno de los vectores x_i es suma de k_i vectores, entonces $f(x_1, \ldots, x_n)$ es una suma de $k_1 \times k_2 \times \cdots \times k_n$ escalares de la forma $f(x'_1, \ldots, x'_n)$, donde cada x'_i es uno de los vectores cuya suma constituye x_i.

3.1.11 Si

$$\begin{aligned} f: \quad E^n &\to \mathbb{K} \\ (x_1, \ldots, x_n) &\to f(x_1, \ldots, x_n) \end{aligned}$$

es una forma n-lineal sobre E, las propiedades **(n-ln1)** y **(n-ln2)** de (3.1.3) significan que, cuando fijamos $n-1$ vectores $x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n$ de E, la aplicación separada de una variable

$$\begin{aligned} E &\to \mathbb{K} \\ x &\to f(x_1, \ldots, x_{i-1}, x, x_{i+1}, \ldots, x_n) \end{aligned}$$

que resulta, es una forma lineal sobre E, o sea, un elemento de E^*.

3.1.12 Ejemplo. Si en la aplicación

$$f: \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$$

de (3.1.8) hacemos $(x', y') = (2, -1)$, la aplicación separada que queda,

$$\begin{aligned} \mathbb{R}^2 &\to \mathbb{R} \\ (x, y) &\to 2x - y\,, \end{aligned}$$

es una forma lineal sobre $\mathbb{R}^2$.

3.1.13 Representamos por $\mathcal{L}_n(E, \mathbb{K})$ el conjunto de las formas n-lineales sobre E. La aplicación

$$\begin{aligned} 0: \quad E^n &\to \mathbb{K} \\ (x_1, \ldots, x_n) &\to 0\,, \end{aligned}$$

que envía todo sistema de E al 0 de $\mathbb{K}$, es claramente n-lineal, o sea, un elemento de $\mathcal{L}_n(E, \mathbb{K})$.

$\mathcal{L}_n(E, \mathbb{K})$ es un subconjunto del e.v. $\mathcal{F}(E^n, \mathbb{K})$ de todas las aplicaciones del conjunto E^n en $\mathbb{K}$ (v. 1.1.3); además $\mathcal{L}_n(E, \mathbb{K})$ es un subespacio de dicho espacio. Entonces $\mathcal{L}_n(E, \mathbb{K})$ es un espacio vectorial cuyas operaciones son las inducidas por las de $\mathcal{F}(E^n, \mathbb{K})$. Si $f, g \in \mathcal{L}_n(E, \mathbb{K})$ y $\lambda \in \mathbb{K}$, recordemos que $f+g$ y λf vienen dadas para cada sistema $(x_1, \ldots, x_n)$ de E por

$$\begin{aligned}(f+g)(x_1, \ldots, x_n) &= f(x_1, \ldots, x_n) + g(x_1, \ldots, x_n)\,, \\ (\lambda f)(x_1, \ldots, x_n) &= \lambda f(x_1, \ldots, x_n)\,.\end{aligned}$$

El cero del e.v. $\mathcal{L}_n(E, \mathbb{K})$ es la forma n-lineal 0 que hemos visto antes. Si $f \in \mathcal{L}_n(E, \mathbb{K})$, su opuesto para la suma es la forma n-lineal $-f$, que en cada sistema $(x_1, \ldots, x_n)$ de E vale

$$(-f)(x_1, \ldots, x_n) = -f(x_1, \ldots, x_n)\,.$$

3.1.14 Ejemplo. Para $n = 1$, el e.v. $\mathcal{L}_1(E, \mathbb{K})$ es justamente el dual de E, $E^* = \mathcal{L}(E, \mathbb{K})$, puesto que las formas 1-lineales sobre E son las formas lineales (v. 3.1.5).

3.1.15 Sea

$$\begin{aligned} f: \qquad E^n &\to \mathbb{K} \\ (x_1, \ldots, x_n) &\to f(x_1, \ldots, x_n)\end{aligned}$$

una forma n-lineal sobre E. Vamos a estudiar la relación que existe entre la imagen $f(x_1, \ldots, x_n)$ de un sistema $(x_1, \ldots, x_n)$ de E y la imagen $f(x_{p(1)}, \ldots, x_{p(n)})$ de una permutación $(x_{p(1)}, \ldots, x_{p(n)})$ del mismo sistema (donde $p \in G_n$) es una permutación de $\{1, 2, \ldots, n\}$).

Para las formas bilineales de (3.1.8) tenemos, por ejemplo,

$$f(\,(1,1),(0,1)\,) = f(\,(0,1),(1,1)\,)\,,$$

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

y lo mismo ocurre tomando dos vectores cualesquiera de $\mathbb{R}^2$; sin embargo, no existen relaciones de este tipo para h. Además

$$g(\,(x,y),(x,y)\,) = 0$$

cualquiera que sea el vector $(x, y) \in \mathbb{R}^2$.

3.1.16 DEFINICIÓN. Sea $f \in \mathcal{L}_n(E, \mathbb{K})$ una forma n-lineal; recordemos que G_n representa el conjunto de las permutaciones del conjunto $\{1, 2, \ldots, n\}$.

Decimos que f es *simétrica* cuando

(sim) $(\forall x_1, \ldots, x_n \in E)(\forall p \in G_n)$
$$f(x_{p(1)}, \ldots, x_{p(n)}) = f(x_1, \ldots, x_n)\,.$$

Decimos que f es *antisimétrica* cuando

(ant) $(\forall x_1, \ldots, x_n \in E)(\forall p \in G_n)$
$$f(x_{p(1)}, \ldots, x_{p(n)}) = \epsilon(p) f(x_1, \ldots, x_n)\,,$$

donde $\epsilon(p)$ representa la paridad de la permutación p.

Decimos que f es *alternada* cuando

(alt) $i \neq j$ y $x_i = x_j \quad \Rightarrow \quad f(x_1, \ldots, x_i, \ldots, x_j, \ldots, x_n) = 0$,

es decir cuando f vale 0 en los sistemas de E que tienen dos vectores iguales.

3.1.17 Las definiciones precedentes toman una expresión particularmente sencilla en el caso $n = 2$ (o sea, para las formas bilineales). Teniendo en cuenta que G_2 sólo posee dos elementos, $(1, 2)$ y $(2, 1)$, que el primero de ellos es la identidad y que el segundo es una permutación impar, resulta que, si $f \in \mathcal{L}_2(E, \mathbb{K})$ es una forma bilineal sobre E, entonces f es simétrica cuando

(sim) $(\forall x_1, x_2 \in E) \quad f(x_2, x_1) = f(x_1, x_2)$,

y que f es antisimétrica cuando

(ant) $(\forall x_1, x_2 \in E) \quad f(x_2, x_1) = -f(x_1, x_2)$.

Por otra parte, f es alternada cuando

(alt) $(\forall x \in E) \quad f(x, x) = 0$.

3.1.18 Ejemplo. De las formas bilineales del ejemplo (3.1.8), f es simétrica y g es antisimétrica y alternada. h no es ni simétrica, ni antisimétrica, ni alternada.

3.1.19 Ejemplo. Las formas det de (3.1.2) son antisimétricas y alternadas.

3.1.20 Ejemplo. La forma 0 (v. 3.1.13) de $\mathcal{L}_n(E, \mathbb{K})$ es simétrica, antisimétrica y alternada.

3.1.21 PROPOSICIÓN

Sea E un e.v. sobre $\mathbb{K}$.

a) Toda forma f n-lineal alternada sobre E es antisimétrica.

b) Si $\mathbb{K}$ no es de característica 2 , toda forma f n-lineal antisimétrica sobre E es alternada.

---

a) Supongamos que $f : E^n \to \mathbb{K}$ es una forma alternada. Sea $\tau \in G_n$ una trasposición y sean i y j los elementos que intercambia; sea $(x_1, \ldots, x_n)$ un sistema de E. Poniendo el vector $x_i + x_j$ en los lugares i y j, por ser f alternada se tiene

$$\begin{aligned}
0 &= f(x_1, \ldots, x_i + x_j, \ldots, x_i + x_j, \ldots, x_n) \\
&= f(x_1, \ldots, x_i, \ldots, x_i, \ldots, x_n) + f(x_1, \ldots, x_i, \ldots, x_j, \ldots, x_n) \\
&\quad + f(x_1, \ldots, x_j, \ldots, x_i, \ldots, x_n) + f(x_1, \ldots, x_j, \ldots, x_j, \ldots, x_n) \\
&= f(x_1, \ldots, x_i, \ldots, x_j, \ldots, x_n) + f(x_1, \ldots, x_j, \ldots, x_i, \ldots, x_n) \\
&= f(x_1, \ldots, x_i, \ldots, x_j, \ldots, x_n) + f(x_{\tau(1)}, \ldots, x_{\tau(i)}, \ldots, x_{\tau(j)}, \ldots, x_{\tau(n)})
\end{aligned}$$

y por lo tanto, para la trasposición τ,

$$f(x_{\tau(1)}, \ldots, x_{\tau(n)}) = -f(x_1, \ldots, x_n) = \epsilon(\tau) f(x_1, \ldots, x_n), .$$

Sea ahora $p \in G_n$ una permutación cualquiera; sabemos que

$$p = \tau_l \circ \cdots \circ \tau_2 \circ \tau_1 ,$$

donde $\tau_1, \tau_2, \ldots, \tau_l$ son l trasposiciones de G_n; también sabemos que

$$\epsilon(p) = \epsilon(\tau_l) \cdots \epsilon(\tau_2)\, \epsilon(\tau_1) .$$

Aplicando entonces lo que acabamos de demostrar,

$$\begin{aligned}
f(x_{p(1)}, \ldots, x_{p(n)}) &= f(x_{\tau_l \circ \cdots \circ \tau_2 \circ \tau_1(1)}, \ldots, x_{\tau_l \circ \cdots \circ \tau_2 \circ \tau_1(n)}) \\
&= \epsilon(\tau_l)\, f(x_{\tau_{l-1} \circ \cdots \circ \tau_2 \circ \tau_1(1)}, \ldots, x_{\tau_{l-1} \circ \cdots \circ \tau_2 \circ \tau_1(n)}) \\
\cdots & \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\
&= \epsilon(\tau_l) \cdots \epsilon(\tau_2) f(x_{\tau_1(1)}, \ldots, x_{\tau_1(n)}) \\
&= \epsilon(\tau_l) \cdots \epsilon(\tau_2) \epsilon(\tau_1) f(x_1, \ldots, x_n) \\
&= \epsilon(p) f(x_1, \ldots, x_n) ,
\end{aligned}$$

lo que demuestra que f es antisimétrica.

b) Supongamos que $f : E^n \to \mathbb{K}$ es una forma antisimétrica. Sea $(x_1, \ldots, x_n)$ un sistema de E tal que para cierta pareja de índices, i y j, con $i \neq j$, se tiene $x_i = x_j$; si representamos por τ la trasposición que intercambia i con j,

$$x_1 = x_{\tau(1)}, \ldots, x_i = x_{\tau(i)}, \ldots, x_j = x_{\tau(j)}, \ldots, x_n = x_{\tau(n)}$$

(ya que $x_i = x_j$) luego

$$f(x_{\tau(1)}, \ldots, x_{\tau(n)}) = f(x_1, \ldots, x_n) .$$

Pero, como f es antisimétrica y $\epsilon(\tau) = -1$, tenemos por otro lado

$$f(x_{\tau(1)}, \ldots, x_{\tau(n)}) = -f(x_1, \ldots, x_n)$$

y resulta entonces que

$$f(x_1, \ldots, x_n) + f(x_1, \ldots, x_n) = 0$$

y que

$$(1 + 1)\, f(x_1, \ldots, x_n) = 0 .$$

Si $\mathbb{K}$ no es de característica 2, entonces $1 + 1 \neq 0$, luego necesariamente

$$f(x_1, \ldots, x_n) = 0 ,$$

y esto demuestra que f es alternada.

3.1.22 Representamos por $\mathcal{S}_n(E,\mathbb{K})$, $\mathcal{U}_n(E,\mathbb{K})$ y $\mathcal{A}_n(E,\mathbb{K})$ los subconjuntos de $\mathcal{L}_n(E,\mathbb{K})$ constituidos por las formas que son respectivamente simétricas, antisimétricas y alternadas. Los tres son subespacios vectoriales de $\mathcal{L}_n(E,\mathbb{K})$, como se demuestra fácilmente.

En la proposición precedente hemos visto que

$$\mathcal{A}_n(E,\mathbb{K}) \subset \mathcal{U}_n(E,\mathbb{K})$$

y que, si $\mathbb{K}$ no es de característica 2,

$$\mathcal{A}_n(E,\mathbb{K}) = \mathcal{U}_n(E,\mathbb{K})\,.$$

Nos interesa en este capítulo el subespacio $\mathcal{A}_n(E,\mathbb{K})$ de las formas alternadas.

3.1.23 PROPOSICIÓN

Sea $f \in \mathcal{A}_n(E,\mathbb{K})$ una forma n-lineal alternada sobre E y $(x_1,\ldots,x_n)$ un sistema de E. Si $(x_1,\ldots,x_n)$ es ligado, entonces

$$f(x_1,\ldots,x_n) = 0\,.$$

Como $(x_1,\ldots,x_n)$ es ligado, uno de los vectores del sistema es combinación lineal del sistema formado por los restantes (v. 1.4.6). Con objeto de no complicar mucho la notación, vamos a suponer que

$$x_1 = \lambda^2 x_2 + \cdots + \lambda^n x_n$$

con $\lambda^2,\ldots,\lambda^n \in \mathbb{K}$. Entonces

$$\begin{aligned} f(x_1,x_2,\ldots,x_n) &= f(\lambda^2 x_2 + \cdots + \lambda^n x_n, x_2,\ldots,x_n) \\ &= \lambda^2 f(x_2,x_2,\ldots,x_n) + \cdots + \lambda^n f(x_n,x_2,\ldots,x_n) \\ &= 0 \end{aligned}$$

puesto que f es alternada.

3.1.24 Sea $f \in \mathcal{A}_n(E,\mathbb{K})$ una forma n-lineal alternada y supongamos que E es de dimensión m con $m < n$. Si $(x_1,\ldots,x_n)$ es un sistema de n vectores de E, entonces es un sistema ligado (v. 1.4.34), luego $f(x_1,\ldots,x_n) = 0$, es decir, $f = 0$. Resulta así que, si $\dim E < n$,

$$\mathcal{A}_n(E,\mathbb{K}) = \{0\}\,.$$

Vamos a ver que, si $\dim E = n$, entonces $\mathcal{A}_n(E,\mathbb{K})$ es de dimensión 1.

3.1.25 TEOREMA

Sea E un e.v. de dimensión $n \neq 0$ y $(a_1, \ldots, a_n)$ una base de E. Existe una forma n-lineal alternada $d \in \mathcal{A}_n(E, \mathbb{K})$, única, que verifica

$$d(a_1, \ldots, a_n) = 1 \, .$$

Además, si $f \in \mathcal{A}_n(E, \mathbb{K})$, entonces

$$f = \lambda \, d$$

con $\lambda = f(a_1, \ldots, a_n)$. En consecuencia, $\mathcal{A}_n(E, \mathbb{K})$ es de dimensión 1.

La demostración del teorema, aunque utiliza argumentos sencillos, presenta un aspecto complicado debido a las dificultades en la notación que se originan al trabajar con las permutaciones de $\{1, 2, \ldots, n\}$. Es aconsejable, cuando se estudia por primera vez, simultanear su lectura con la del ejemplo posterior, que aborda el caso particular en que $n = 2$.

Sea $f \in \mathcal{A}_n(E, \mathbb{K})$ una forma n-lineal alternada sobre E. Si $(x_1, \ldots, x_n)$ es un sistema de E y

$$\begin{aligned} x_1 &= (x_1^1, \ldots, x_1^n)_{(a_i)} = x_1^1 a_1 + \cdots + x_1^n a_n \\ &\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ x_n &= (x_n^1, \ldots, x_n^n)_{(a_i)} = x_n^1 a_1 + \cdots + x_n^n a_n \, , \end{aligned}$$

entonces (véase el comentario del ejemplo 3.1.10) $f(x_1, \ldots, x_n)$ es la suma de los $n^n = n \times \cdots \times n$ términos posibles de la forma

$$x_1^{j_1} x_2^{j_2} \cdots x_n^{j_n} \, f(a_{j_1}, a_{j_2}, \ldots, a_{j_n}) \, ,$$

donde $j_1, \ldots, j_n$ son elementos de $\{1, \ldots, n\}$ (iguales o distintos), o sea,

$$f(x_1, \ldots, x_n) = \sum_{j_1, \ldots, j_n \in \{1, \ldots, n\}} x_1^{j_1} x_2^{j_2} \cdots x_n^{j_n} \, f(a_{j_1}, a_{j_2}, \ldots, a_{j_n}) \, .$$

Como f es alternada, $f(a_{j_1}, a_{j_2}, \ldots, a_{j_n}) = 0$ siempre que dos de los j_k son iguales; los términos en los que se da esta situación valen cero. Los otros términos son aquellos en los que los números $j_1, j_2, \ldots, j_n$ son todos distintos; poniendo en este caso

$$p(1) = j_1, p(2) = j_2, \ldots, p(n) = j_n \, ,$$

resulta que p es una permutación del conjunto $\{1, 2, \ldots, n\}$. Eliminando de la suma los términos que sabemos que son nulos, tenemos

$$f(x_1, \ldots, x_n) = \sum_{p \in G_n} x_1^{p(1)} x_2^{p(2)} \cdots x_n^{p(n)} \, f(a_{p(1)}, a_{p(2)}, \ldots, a_{p(n)}) \, ,$$

donde la suma posee $n!$ sumandos (el número de elementos de G_n). Como f es también antisimétrica (v. 3.1.21), tenemos que

$$f(a_{p(1)}, a_{p(2)}, \ldots, a_{p(n)}) = \epsilon(p)\, f(a_1, a_2, \ldots, a_n)$$

y entonces, apartando el factor común $f(a_1, a_2, \ldots, a_n)$,

$$f(x_1, \ldots, x_n) = f(a_1, \ldots, a_n) \sum_{p \in G_n} \epsilon(p)\, x_1^{p(1)} x_2^{p(2)} \cdots x_n^{p(n)} .$$

Vamos a definir ahora una aplicación

$$\begin{aligned} d : \qquad E^n &\to \mathbb{K} \\ (x_1, \ldots, x_n) &\to d(x_1, \ldots, x_n) \end{aligned}$$

poniendo

$$d(x_1, \ldots, x_n) = \sum_{p \in G_n} \epsilon(p)\, x_1^{p(1)} x_2^{p(2)} \cdots x_n^{p(n)}$$

cuando

$$\begin{aligned} x_1 &= (x_1^1, \ldots, x_1^n)_{(a_i)} \\ &\ldots\ldots\ldots\ldots\ldots\ldots \\ x_n &= (x_n^1, \ldots, x_n^n)_{(a_i)} . \end{aligned}$$

d es una forma n-lineal; la demostración de este hecho es sencilla si se utilizan los razonzmientos sobre las coordenadas de una suma de vectores y del producto de un escalar por un vector, que vimos en (1.4.23). También es fácil ver que

$$d(a_1, \ldots, a_n) = 1 .$$

Demostremos que d es alternada; sea $(x_1, \ldots, x_n)$ un sistema de E con dos vectores iguales, $x_i = x_j$, correspondientes a índices $i \neq j$; llamemos τ a la trasposición que intercambia i y j. La aplicación

$$\begin{aligned} P &\to P' \\ p &\to p \circ \tau \end{aligned}$$

del conjunto P de las permutaciones de G_n que son pares en el conjunto P' de las que son impares es biyectiva. Entonces

$$\begin{aligned} d(x_1, \ldots, x_n) &= \sum_{p \in G_n} \epsilon(p)\, x_1^{p(1)} x_2^{p(2)} \cdots x_n^{p(n)} \\ &= \sum_{p \in P} x_1^{p(1)} x_2^{p(2)} \cdots x_n^{p(n)} + \sum_{p \in P'} (-1)\, x_1^{p(1)} x_2^{p(2)} \cdots x_n^{p(n)} \\ &= \sum_{p \in P} x_1^{p(1)} x_2^{p(2)} \cdots x_n^{p(n)} + \sum_{p \in P} (-1)\, x_1^{p\circ\tau(1)} x_2^{p\circ\tau(2)} \cdots x_n^{p\circ\tau(n)} \\ &= \sum_{p \in P} \left(x_1^{p(1)} x_2^{p(2)} \cdots x_n^{p(n)} - x_1^{p\circ\tau(1)} x_2^{p\circ\tau(2)} \cdots x_n^{p\circ\tau(n)} \right) . \end{aligned}$$

Consideremos un término

$$x_1^{p(1)} x_2^{p(2)} \cdots x_n^{p(n)} - x_1^{p\circ\tau(1)} x_2^{p\circ\tau(2)} \cdots x_n^{p\circ\tau(n)}$$

de esta suma; si $k \neq i$ y $k \neq j$, entonces

$$x_k^{p\circ\tau(k)} = x_k^{p(k)}$$

puesto que $\tau(k) = k$; por otra parte, $\tau(i) = j$ y $\tau(j) = i$, luego

$$x_i^{p\circ\tau(i)} = x_i^{p(j)} = x_j^{p(j)} \quad \text{y} \quad x_j^{p\circ\tau(j)} = x_j^{p(i)} = x_i^{p(i)}$$

puesto que $x_i = x_j$; entonces $x_1^{p(1)} x_2^{p(2)} \cdots x_n^{p(n)}$ y $x_1^{p\circ\tau(1)} x_2^{p\circ\tau(2)} \cdots x_n^{p\circ\tau(n)}$ son producto de los mismos escalares, aunque dos de ellos estén en distinto orden, luego son iguales a causa de la conmutatividad del producto en $\mathbb{K}$. Cada término de los que forman la suma es, pues, 0 y resulta que $d(x_1, \ldots, x_n) = 0$; esto prueba que d es alternada. d cumple así los requisitos del teorema.

Además, resulta que, si $f \in \mathcal{A}_n(E, \mathbb{K})$,

$$\begin{aligned} f(x_1, \ldots, x_n) &= f(a_1, \ldots, a_n) \sum_{p \in G_n} \epsilon(p)\, x_1^{p(1)} \cdots x_n^{p(n)} \\ &= f(a_1, \ldots, a_n) d(x_1, \ldots, x_n) \end{aligned}$$

para cada sistema $(x_1, \ldots, x_n)$ de E, luego

$$f = f(a_1, \ldots, a_n)\, d\,,$$

como queríamos demostrar.

Esto nos permite además comprobar que d es única, pues si $d' \in \mathcal{A}_n(E, \mathbb{K})$ y $d'(a_1, \ldots, a_n) = 1$, entonces

$$d' = d'(a_1, \ldots, a_n)\, d = 1\, d = d\,.$$

Por fin, el sistema $(\, d\,)$ de $\mathcal{A}_n(E, \mathbb{K})$, formado por un solo elemento, es libre, pues, como $d(a_1, \ldots, a_n) = 1$, entonces $d \neq 0$. Hemos visto antes que $(\, d\,)$ genera $\mathcal{A}_n(E, \mathbb{K})$, luego es una base de $\mathcal{A}_n(E, \mathbb{K})$ y este espacio es de dimensión 1.

3.1.26 Ejemplo. Si E es un e.v. de dimensión 2, (a_1, a_2) una base de E y $f \in \mathcal{A}_2(E, \mathbb{K})$, entonces, para todo sistema (x_1, x_2) de dos vectores de E, con coordenadas

$$x_1 = (x_1^1, x_1^2)_{(a_1, a_1)}\,, \quad x_2 = (x_2^1, x_2^2)_{(a_1, a_1)}\,,$$

tenemos

$$\begin{aligned} f(x_1, x_2) &= f(x_1^1 a_1 + x_1^2 a_2, x_2^1 a_1 + x_2^2 a_2) \\ &= x_1^1 x_2^1\, f(a_1, a_1) + x_1^1 x_2^2\, f(a_1, a_2) + x_1^2 x_2^1\, f(a_2, a_1) + x_1^2 x_2^2\, f(a_2, a_2) \\ &= x_1^1 x_2^2\, f(a_1, a_2) + x_1^2 x_2^1\, f(a_2, a_1) \\ &= x_1^1 x_2^2\, f(a_1, a_2) - x_1^2 x_2^1\, f(a_1, a_2) \\ &= f(a_1, a_2)\, \left(x_1^1 x_2^2 - x_1^2 x_2^1\right)\,. \end{aligned}$$

La aplicación

$$\begin{aligned} d: \quad E^2 &\to \mathbb{K} \\ (x_1, x_2) &\to d(x_1, x_2) \end{aligned}$$

definida por

$$d(x_1, x_2) = x_1^1 x_2^2 - x_1^2 x_2^1$$

es una forma bilineal alternada sobre E, cumple $d(a_1, a_2) = 1$ y además

$$(\forall x_1, x_2 \in E) \quad f(x_1, x_2) = f(a_1, a_2)\, d(x_1, x_2)\,,$$

luego

$$f = f(a_1, a_2)\, d\,.$$

Obsérvese que

$$d(x_1, x_2) = \det \begin{bmatrix} x_1^1 & x_2^1 \\ x_1^2 & x_2^2 \end{bmatrix}.$$

Es justamente la forma d lo que por definición denominaremos determinante.

3.1.27 Ejemplo. Para $E = \mathbb{R}^2$ y la base canónica (e_1, e_2) de $\mathbb{R}^2$, la correspondiente forma bilineal alternada d viene dada por

$$d(\,(x, y), (x', y')\,) = xy' - yx'\,.$$

Toda forma bilineal alternada f sobre $\mathbb{R}^2$ es de la forma

$$f = \lambda\, d\,,$$

o sea,

$$f(\,(x, y), (x', y')\,) = \lambda xy' - \lambda yx'\,,$$

y tenemos que

$$f(e_1, e_2) = f(\,(1, 0), (0, 1)\,) = \lambda\, 1\, 1 - \lambda\, 0\, 0 = \lambda\,.$$

3.1.28 La forma $d \in \mathcal{A}_n(E, \mathbb{K})$ que hemos definido en el teorema precedente y que en cada sistema $(x_1, \ldots, x_n)$ de E con

$$\begin{aligned} x_1 &= (x_1^1, \ldots, x_1^n)_{(a_i)} \\ &\cdots\cdots\cdots\cdots \\ x_n &= (x_n^1, \ldots, x_n^n)_{(a_i)} \end{aligned}$$

vale

$$\textbf{(1)} \qquad \boxed{d(x_1, \ldots, x_n) = \sum_{p \in G_n} \epsilon(p)\, x_1^{p(1)} x_2^{p(2)} \cdots x_n^{p(n)}}$$

depende efectivamente de la base $(a_1, \ldots, a_n)$ considerada en E, puesto que para calcular $d(x_1, \ldots, x_n)$ empleamos las coordenadas de $x_1, \ldots, x_n$ en dicha base.

3.1.29 PROPOSICIÓN

Sea E un e.v. de dimensión n y $f \in \mathcal{A}_n(E, \mathbb{K})$ una forma n-lineal alternada sobre E tal que $f \neq 0$; sea $(x_1, \ldots, x_n)$ un sistema de E. Entonces $f(x_1, \ldots, x_n) = 0$ si y sólo si $(x_1, \ldots, x_n)$ es ligado (o, lo que es igual, $f(x_1, \ldots, x_n) \neq 0$ si y sólo si $(x_1, \ldots, x_n)$ es una base de E).

Si $(x_1, \ldots, x_n)$ es ligado, entonces $f(x_1, \ldots, x_n) = 0$, como hemos visto en (3.1.23). Recíprocamente, si $(x_1, \ldots, x_n)$ es libre, entonces es una base de E; representemos por $d \in \mathcal{A}_n(E, \mathbb{K})$ la única forma n-lineal alternada sobre E cuyo valor en la base $(x_1, \ldots, x_n)$ es

$$d(x_1, \ldots, x_n) = 1\,.$$

El teorema (3.1.25), que garantiza la existencia de d, afirma además que

$$f = f(x_1, \ldots, x_n)\, d$$

(nótese que $(x_1, \ldots, x_n)$ es ahora la base considerada) y, como $f \neq 0$, necesariamente $f(x_1, \ldots, x_n) \neq 0$.

3.2 Determinantes.

3.2.1 Los razonamientos de la sección precedente nos sitúan en condiciones de definir qué se entiende por 'determinante'. Vamos a dar una triple definición; diremos qué es el determinante de una matriz cuadrada (generalizando las nociones de 3.1.1), qué es el determinante de un sistema de vectores relativo a una base (generalizando 3.1.2) y qué es el determinante de un endomorfismo. Estudiaremos también la relación existente entre estas nociones.

3.2.2 DEFINICIÓN. Sea E un e.v. de dimensión $n \neq 0$ y $(a_1, \ldots, a_n)$ una base de E. Hemos visto en (3.1.25) que existe una única forma n-lineal alternada, d, sobre E que verifica $d(a_1, \ldots, a_n) = 1$. Representamos esta forma por

$$\det_{(a_1, \ldots, a_n)} \quad \text{o} \quad \det_{(a_i)}$$

y le llamamos *determinante relativo a la base* $(a_1, \ldots, a_n)$. Así, pues,

$$\det_{(a_i)} \in \mathcal{A}_n(E, \mathbb{K})\,.$$

Como vimos en (3.1.25), si $f \in \mathcal{A}_n(E, \mathbb{K})$, entonces

$$f = f(a_1, \ldots, a_n) \det_{(a_i)}\,.$$

El escalar

$$\det_{(a_i)}(x_1, \ldots, x_n)\,,$$

o sea, el valor que toma $\det_{(a_i)}$ en un sistema $(x_1, \ldots, x_n)$ de n vectores de E, se llama *determinante de* $(x_1, \ldots, x_n)$ *en la base* $(a_1, \ldots, a_n)$. Sabemos que, en particular,

$$\det_{(a_i)}(a_1, \ldots, a_n) = 1\,.$$

Si

$$\begin{aligned} x_1 &= (x_1^1, \ldots, x_1^n)_{(a_i)} \\ &\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ x_n &= (x_n^1, \ldots, x_n^n)_{(a_i)}\,, \end{aligned}$$

entonces

$$\textbf{(2)} \qquad \boxed{\det_{(a_i)}(x_1, \ldots, x_n) = \sum_{p \in G_n} \epsilon(p)\, x_1^{p(1)} x_2^{p(2)} \cdots x_n^{p(n)}\,,}$$

como vimos en (3.1.25).

3.2.3 Ejemplo. Si E es un e.v. de dimensión 2, (a_1, a_2) una base de E y $x_1, x_2 \in E$, con

$$x_1 = (x_1^1, x_1^2)_{(a_1,a_2)}\,, \quad x_2 = (x_2^1, x_2^2)_{(a_1,a_2)}\,,$$

entonces

$$\det_{(a_1,a_2)}(x_1, x_2) = x_1^1 x_2^2 - x_1^2 x_2^1\,.$$

Si E es un e.v. de dimensión 3, (a_1, a_2, a_3) una base de E y $x_1, x_2, x_3 \in E$, con

$$\begin{aligned} x_1 &= (x_1^1, x_1^2, x_1^3)_{(a_1,a_2,a_3)} \\ x_2 &= (x_2^1, x_2^2, x_2^3)_{(a_1,a_2,a_3)} \\ x_3 &= (x_3^1, x_3^2, x_3^3)_{(a_1,a_2,a_3)}\,, \end{aligned}$$

entonces

$$\det_{(a_1,a_2,a_3)}(x_1, x_2, x_3) = x_1^1 x_2^2 x_3^3 + x_1^2 x_2^3 x_3^1 + x_1^3 x_2^1 x_3^2 - x_1^1 x_2^3 x_3^2 - x_1^2 x_2^1 x_3^3 - x_1^3 x_2^2 x_3^1\,.$$

3.2.4 DEFINICIÓN. Sea $A = [\alpha_i^j] \in M_{\mathbb{K}}(n)$ una matriz cuadrada. Sus n vectores columna

$$\begin{aligned} c_1 &= (\alpha_1^1, \alpha_1^2, \ldots, \alpha_1^n) \\ &\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ c_n &= (\alpha_n^1, \alpha_n^2, \ldots, \alpha_n^n) \end{aligned}$$

son vectores de $\mathbb{K}^n$. El determinante

$$\det_{(e_1,\ldots,e_n)}(c_1, \ldots, c_n)$$

del sistema $(c_1, \ldots, c_n)$ en la base canónica $(e_1, \ldots, e_n)$ de $\mathbb{K}^n$ se llama *determinante de la matriz* A y se denota por

$$\det A\,, \quad \det \begin{bmatrix} \alpha_1^1 & \cdots & \alpha_n^1 \\ \vdots & & \vdots \\ \alpha_1^n & \cdots & \alpha_n^n \end{bmatrix} \quad \text{o} \quad \begin{vmatrix} \alpha_1^1 & \cdots & \alpha_n^1 \\ \vdots & & \vdots \\ \alpha_1^n & \cdots & \alpha_n^n \end{vmatrix}\,.$$

Resulta que

$$\textbf{(3)} \qquad \boxed{\det A = \sum_{p \in G_n} \epsilon(p)\, \alpha_1^{p(1)} \alpha_2^{p(2)} \cdots \alpha_n^{p(n)}\,.}$$

3.2.5 Ejemplo. Para la matriz unidad se tiene

$$\det I_n = 1\,,$$

puesto que los vectores columna de I_n son los de la base canónica de $\mathbb{K}^n$.

3.2.6 Ejemplo. Para las matrices 1×1 se tiene

$$\det \left[\ \alpha\ \right] = \alpha\,;$$

para las matrices 2×2,

$$\begin{vmatrix} \alpha_1^1 & \alpha_2^1 \\ \alpha_1^2 & \alpha_2^2 \end{vmatrix} = \alpha_1^1 \alpha_2^2 - \alpha_1^2 \alpha_2^1\,;$$

y para las matrices 3×3,

$$\begin{vmatrix} \alpha_1^1 & \alpha_2^1 & \alpha_3^1 \\ \alpha_1^2 & \alpha_2^2 & \alpha_3^2 \\ \alpha_1^3 & \alpha_2^3 & \alpha_3^3 \end{vmatrix} = \alpha_1^1\alpha_2^2\alpha_3^3 + \alpha_1^2\alpha_2^3\alpha_3^1 + \alpha_1^3\alpha_2^1\alpha_3^2 - \alpha_1^1\alpha_2^3\alpha_3^2 - \alpha_1^2\alpha_2^1\alpha_3^3 - \alpha_1^3\alpha_2^2\alpha_3^1\,.$$

3.2.7 Sea E un espacio vectorial de dimensión $n \neq 0$, $(a_1, \ldots, a_n)$ una base de E y $(x_1, \ldots, x_n)$ un sistema de E con

$$\begin{aligned} x_1 &= (x_1^1, \ldots, x_1^n)_{(a_i)} \\ &\ldots\ldots\ldots\ldots\ldots\ldots \\ x_n &= (x_n^1, \ldots, x_n^n)_{(a_i)}\,. \end{aligned}$$

Basta recordar las fórmulas **(2)** y **(3)** de (3.2.2) y (3.2.4) para ver que

$$\begin{aligned} \det_{(a_i)}(x_1, \ldots, x_n) &= \det\left[(x_1, \ldots, x_n), (a_i)\right] \\ &= \begin{vmatrix} x_1^1 & \cdots & x_n^1 \\ \vdots & & \vdots \\ x_1^n & \cdots & x_n^n \end{vmatrix}\,. \end{aligned}$$

3.2.8 PROPOSICIÓN

Sea E un e.v. de dimensión $n \neq 0$ y $(a_1, \ldots, a_n)$ una base de E. Sea $(x_1, \ldots, x_n)$ un sistema de E, donde cada x_i tiene por coordenadas

$$x_i = (x_i^1, \ldots, x_i^n)_{(a_i)}\,.$$

Entonces, las propiedades siguientes son equivalentes:

(i) $(x_1, \ldots, x_n)$ es una base de E ;

(ii) $\det_{(a_i)}(x_1, \ldots, x_n) \neq 0$;

(iii) $\begin{vmatrix} x_1^1 & \cdots & x_n^1 \\ \vdots & & \vdots \\ x_1^n & \cdots & x_n^n \end{vmatrix} \neq 0$.

La equivalencia de las dos primeras propiedades es otra forma de enunciar la proposición (3.1.29), y la de las dos últimas es una consecuencia del párrafo precedente.

3.2.9 Si $(a_1, \ldots, a_n)$ y $(a'_1, \ldots, a'_n)$ son dos bases de un e.v. E, entonces

$$\det_{(a_i)}(a'_1, \ldots, a'_n) = \det\,[(a'_i), (a_i)]\,,$$

como se sigue de (3.2.7). Resulta entonces de la proposición precedente que el determinante de una matriz de cambio de base es siempre distinto de 0; veremos que ocurre lo mismo con todas las matrices inversibles.

Por otro lado, $\det_{(a_i)}$ y $\det_{(a'_i)}$ son dos formas antisimétricas; entonces (v. 3.2.2)

$$\det_{(a'_i)} = \det_{(a'_i)}(a_1, \ldots, a_n)\ \det_{(a_i)}\,,$$

o sea,

$$\det_{(a'_i)} = \det\,[(a_i), (a'_i)]\ \det_{(a_i)}\ .$$

3.2.10 Sea E un e.v. de dimensión $n \neq 0$ y $(a_1, \ldots, a_n)$ una base de E; sea $f \in \mathcal{L}(E)$ un endomorfismo de E y $[f, (a_i)]$ la matriz asociada a f. Entonces

$$\det\,[f, (a_i)] = \det_{(a_i)}(f(a_1), \ldots, f(a_n))$$

puesto que $[f, (a_i)]$ es la matriz cuyas columnas están formadas por las coordenadas en la base $(a_1, \ldots, a_n)$ de los vectores $f(a_1), \ldots, f(a_n)$.

3.2.11 PROPOSICIÓN

Sea $A = [\alpha_i^j] \in M(n)$; se tiene

$$\det A = \det A^t\,.$$

Recordemos que, si $p \in G_n$ y p^{-1} es su inversa,

$$\epsilon(p) = \epsilon(p^{-1})\,.$$

Además, como el producto de $\mathbb{K}$ es conmutativo,

$$\alpha_1^{p(1)} \cdots \alpha_n^{p(n)} = \alpha_{p^{-1}(1)}^{p\circ p^{-1}(1)} \cdots \alpha_{p^{-1}(n)}^{p\circ p^{-1}(n)} = \alpha_{p^{-1}(1)}^{1} \cdots \alpha_{p^{-1}(n)}^{n} ,$$

luego

$$\det A = \sum_{p\in G_n} \epsilon(p)\, \alpha_1^{p(1)} \cdots \alpha_n^{p(n)} = \sum_{p\in G_n} \epsilon(p^{-1})\, \alpha_{p^{-1}(1)}^{1} \cdots \alpha_{p^{-1}(n)}^{n}$$

y, como la aplicación

$$\begin{aligned} G_n &\to G_n \\ p &\to p^{-1} \end{aligned}$$

es biyectiva, resulta

$$\det A = \sum_{p\in G_n} \epsilon(p)\, \alpha_{p(1)}^{1} \cdots \alpha_{p(n)}^{n} ;$$

por fin, llamando ${}^t\alpha_i^j$ al término general de A^t,

$$\det A = \sum_{p\in G_n} \epsilon(p)\, {}^t\alpha_1^{p(1)} \cdots\ {}^t\alpha_n^{p(n)} = \det A^t .$$

3.2.12 PROPOSICIÓN

Si $A = [\alpha_i^j] \in M(n)$ es una matriz triangular (v. 2.2.10), entonces

$$\det A = \alpha_1^1 \alpha_2^2 \cdots \alpha_n^n ,$$

o sea, su determinante es el producto de sus elementos diagonales.

Supongamos que A es triangular superior (para el caso de triangular inferior se opera de forma análoga); entonces

$$j > i \quad \Rightarrow \quad \alpha_i^j = 0 .$$

Sea $p \in G_n$ una permutación diferente de la identidad; si llamamos i al menor de los elementos j de $\{1,\dots,n\}$ para los que

$$p(j) \neq j ,$$

entonces

$$p(1) = 1, \dots, p(i-1) = i-1 ,$$

pero sin embargo $p(i) \neq i$ y $p(i) \geq i$, o sea, $p(i) > i$. Resulta así que $\alpha_i^{p(i)} = 0$ luego que

$$\alpha_1^{p(1)} \alpha_2^{p(2)} \cdots \alpha_n^{p(n)} = 0 .$$

Entonces

$$\det A = \sum_{p\in G_n} \epsilon(p)\, \alpha_1^{p(1)} \alpha_2^{p(2)} \cdots \alpha_n^{p(n)} = \alpha_1^1 \alpha_2^2 \cdots \alpha_n^n$$

puesto que los otros términos son todos nulos.

3.2.13 Ejemplo. Para una matriz diagonal

$$\det \begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{bmatrix} = \lambda_1 \lambda_2 \cdots \lambda_n \,.$$

En particular

$$\det(\lambda I_n) = \lambda \cdot \lambda \cdots \lambda = (\lambda)^n \,.$$

3.2.14 PROPOSICIÓN

Sean $A \in M(r)$ y $D \in M(s)$ dos matrices cuadradas y $B \in M(s, r)$; consideremos la matriz cuadrada $M \in M(r+s)$

$$M = \left[\begin{array}{c|c} A & B \\ \hline 0 & D \end{array}\right]$$

definida por bloques (v. 2.2.15). (Se suele decir que M es triangular por bloques.) Se tiene entonces que

$$\det M = \det A \cdot \det D \,.$$

Pongamos $M = [m_i^j]$, $A = [\alpha_i^j]$, $B = [\beta_i^j]$ y $D = [\gamma_i^j]$. Por definición

$$\begin{aligned} \det M &= \textstyle\sum_{p \in G_{r+s}} \epsilon(p)\, m_1^{p(1)} \cdots m_r^{p(r)} m_{r+1}^{p(r+1)} \cdots m_{r+s}^{p(r+s)} \,, \\ \det A &= \textstyle\sum_{\tau \in G_r} \epsilon(\tau)\, \alpha_1^{\tau(1)} \cdots \alpha_r^{\tau(r)} \,, \\ \det D &= \textstyle\sum_{\mu \in G_s} \epsilon(\mu)\, \gamma_1^{\mu(1)} \cdots \gamma_s^{\mu(s)} \,. \end{aligned}$$

Entonces

$$\begin{aligned} \det A \cdot \det D &= \left(\sum_{\tau \in G_r} \epsilon(\tau)\, \alpha_1^{\tau(1)} \cdots \alpha_r^{\tau(r)} \right) \left(\sum_{\mu \in G_s} \epsilon(\mu)\, \gamma_1^{\mu(1)} \cdots \gamma_s^{\mu(s)} \right) \\ &= \sum_{\substack{\tau \in G_r \\ \mu \in G_s}} \epsilon(\tau)\epsilon(\mu)\, \alpha_1^{\tau(1)} \cdots \alpha_r^{\tau(r)} \gamma_1^{\mu(1)} \cdots \gamma_s^{\mu(s)} \\ &= \sum_{\substack{\tau \in G_r \\ \mu \in G_s}} \epsilon(\tau)\epsilon(\mu)\, m_1^{\tau(1)} \cdots m_r^{\tau(r)} m_{r+1}^{r+\mu(1)} \cdots m_{r+s}^{r+\mu(s)} \,. \end{aligned}$$

Si $\tau \in G_r$ y $\mu \in G_s$, la aplicación

$$p : \{1, 2, \ldots, r+s\} \to \{1, 2, \ldots, r+s\} \,,$$

definida por

$$p(1) = \tau(1), \ldots, p(r) = \tau(r), p(r+1) = r + \mu(1), \ldots, p(r+s) = r + \mu(s)\,,$$

es una permutación de G_{r+s} ; vamos a representarla por $\tau + \mu$. Es fácil ver que el número de inversiones es

$$I(\tau + \mu) = I(\tau) + I(\mu)$$

y que, por lo tanto,

$$\epsilon(\tau + \mu) = \epsilon(\tau)\epsilon(\mu)\,.$$

Resulta así que

$$\det A \cdot \det D = \sum_{\substack{\tau \in G_r \\ \mu \in G_s}} \epsilon(\tau + \mu)\, m_1^{\tau+\mu(1)} \cdots m_r^{\tau+\mu(r)}\, m_{r+1}^{\tau+\mu(r+1)} \cdots m_{r+s}^{\tau+\mu(r+s)}\,.$$

Si representamos por G'_{r+s} el subconjunto de G_{r+s} formado por las permutaciones $p \in G_{r+s}$ que verifican

$$p(1) \leq r, \ldots, p(r) \leq r\,,$$

entonces la aplicación

$$\begin{aligned} G_r \times G_s &\to G'_{r+s} \\ (\tau, \mu) &\to \tau + \mu \end{aligned}$$

es biyectiva. Tenemos así que

$$\det A \det D = \sum_{p \in G'_{r+s}} \epsilon(p)\, m_1^{p(1)} \cdots m_r^{p(r)} m_{r+1}^{p(r+1)} \cdots m_{r+s}^{p(r+s)}\,.$$

Volvamos ahora a la expresión de $\det M$; podemos poner

$$\begin{aligned} \det M \;=\; & \sum_{p \in G'_{r+s}} \epsilon(p)\, m_1^{p(1)} \cdots m_r^{p(r)} m_{r+1}^{p(r+1)} \cdots m_{r+s}^{p(r+s)} \\ & + \sum_{p \notin G'_{r+s}} \epsilon(p)\, m_1^{p(1)} \cdots m_r^{p(r)} m_{r+1}^{p(r+1)} \cdots m_{r+s}^{p(r+s)}\,. \end{aligned}$$

Si $p \notin G'_{r+s}$, existe $i \in \{1, \ldots, r\}$ tal que

$$p(i) > r\,,$$

y esto significa, a causa de la estructura de M, que

$$m_i^{p(i)} = 0\,,$$

luego que el término correspondiente a p es nulo. Entonces

$$\det M = \sum_{p \in G'_{r+s}} \epsilon(p)\, m_1^{p(1)} \cdots m_r^{p(r)} m_{r+1}^{p(r+1)} \cdots m_{r+s}^{p(r+s)} = \det A \cdot \det D\,.$$

3.2.15 Para una matriz, definida por bloques, de la forma

$$M = \left[\begin{array}{c|c} A & 0 \\ \hline C & D \end{array}\right],$$

donde A y D son matrices cuadradas, también es cierto que

$$\det M = \det A \cdot \det D$$

puesto que

$$M^t = \left[\begin{array}{c|c} A^t & C^t \\ \hline 0 & D^t \end{array}\right]$$

y, aplicando (3.2.11) y la proposición precedente,

$$\det M = \det M^t = \det A^t \cdot \det D^t = \det A \cdot \det D\,.$$

3.2.16 TEOREMA

Si $A, B \in M(n)$, se tiene

$$\det(B\,A) = \det B \cdot \det A\,.$$

Tomemos cualquier espacio vectorial E de dimensión n (por ejemplo $\mathbb{K}^n$) y una base $(a_1, \ldots, a_n)$ cualquiera de E; existen endomorfismos $f, g \in \mathcal{L}(E)$ tales que (v. 2.4.9)

$$A = [f, (a_i)]\,, \quad B = [g, (a_i)]\,.$$

Aplicando (2.4.23), se tiene

$$B\,A = [g \circ f, (a_i)]\,.$$

Hemos visto en (3.2.10) que

$$\det A = \det_{(a_i)}(f(a_1), \ldots, f(a_n)\,)\,,$$

$$\det B = \det_{(a_i)}(g(a_1), \ldots, g(a_n)\,)\,,$$

$$\det(B\,A) = \det_{(a_i)}(g \circ f(a_1), \ldots, g \circ f(a_n)\,)\,.$$

Definimos la aplicación

$$\begin{aligned} h: \quad & E^n \to \mathbb{K} \\ & (x_1, \ldots, x_n) \to \det_{(a_i)}(g(x_1), \ldots, g(x_n)\,)\,; \end{aligned}$$

se comprueba sin dificultad que h es una forma n-lineal alternada sobre E. Entonces

$$h = h(a_1, \ldots, a_n)\det_{(a_i)}\,,$$

pero
$$h(a_1,\ldots,a_n) = \det_{(a_i)}(g(a_1),\ldots,g(a_n)\,) = \det B$$
luego
$$h = \det B \cdot \det_{(a_i)}\ .$$
En particular, para el sistema $(f(a_1),\ldots,f(a_n)\,)$ de E, tenemos
$$h(f(a_1),\ldots,f(a_n)\,) = \det B \cdot \det_{(a_i)}(f(a_1),\ldots,f(a_n)\,) = \det B \cdot \det A$$
y como, por definición de h,
$$h(f(a_1),\ldots,f(a_n)\,) = \det_{(a_i)}(g\circ f(a_1),\ldots,g\circ f(a_n)\,) = \det(B\,A)\,,$$
resulta la igualdad que teníamos que probar.

3.2.17 COROLARIO

Si $A \in M(n)$ y $\lambda \in \mathbb{K}$, entonces
$$\det(\lambda A) = (\lambda)^n\ \det A\,.$$

En efecto,
$$\lambda A = \lambda(I_n A) = (\lambda I_n)A$$
y como vimos en (3.2.13), $\det(\lambda I_n) = (\lambda)^n$.

3.2.18 TEOREMA

Sea $A \in M(n)$; entonces son equivalentes

(i) A es inversible;
(ii) $\det A \neq 0$.

Además, si A es inversible (o sea, si $\det A \neq 0$) y A^{-1} es su inversa, se tiene
$$\det(A^{-1}) = (\det A)^{-1}\,.$$

(i)$\Rightarrow$(ii). Supongamos que A es inversible y representemos por A^{-1} su inversa; como
$$AA^{-1} = I_n$$
resulta que
$$\det A \cdot \det(A^{-1}) = \det I_n = 1$$
y necesariamente $\det A \neq 0$. Además
$$\det(A^{-1}) = (\det A)^{-1}\,.$$

(ii)⇒(i). Supongamos que $\det A \neq 0$. Llamando f al endomorfismo de $\mathbb{K}^n$ dado por A,

$$A = [f, (e_i)]$$

(donde $(e_1, \ldots, e_n)$ es la base canónica de $\mathbb{K}^n$). Resulta entonces que

$$\det_{(e_i)}(f(e_1), \ldots, f(e_n)) = \det[f, (e_i)] = \det A \neq 0$$

y, aplicando (3.2.8), $(f(e_1), \ldots, f(e_n))$ es una base de $\mathbb{K}^n$; esto significa que f es un automorfismo o, lo que es igual, que $\operatorname{rg} f = n$; por lo tanto, A es inversible (v. 2.5.31).

3.2.19 La condición para que una matriz $A \in M(n)$ sea inversible (v. 2.4.19), que exigía la existencia de una matriz cuyo producto 'a ambos lados' por A fuese la matriz unidad, se puede ahora substituir por otra menos exigente. En efecto, si existe una matriz $B \in M(n)$ tal que

$$B\,A = I_n\,,$$

entonces

$$\det B \cdot \det A = \det I_n = 1\,,$$

luego $\det A \neq 0$ y A es inversible; además, B es la inversa de A puesto que

$$B = B\,I_n = B(A\,A^{-1}) = (B\,A)A^{-1} = I_n A^{-1} = A^{-1}\,.$$

Lo mismo podemos decir cuando existe una matriz $B \in M(n)$ tal que

$$A\,B = I_n\,.$$

3.2.20 COROLARIO

Si $A, A' \in M(n)$ son dos matrices semejantes (v. 2.5.25), entonces

$$\det A = \det A'\,.$$

En efecto, puesto que existe una matriz $P \in M(n)$ inversible, tal que

$$A' = P^{-1}AP\,,$$

entonces

$$\begin{aligned}
\det A' &= \det(P^{-1}) \cdot \det A \cdot \det P \\
&= (\det P)^{-1} \cdot \det A \cdot \det P \\
&= (\det P)^{-1} \cdot \det P \cdot \det A \\
&= \det A\,.
\end{aligned}$$

(Recuérdese que el producto de $\mathbb{K}$ es conmutativo.)

3.2.21 DEFINICIÓN. Sea E un e.v. de dimensión $n \neq 0$ y $f \in \mathcal{L}(E)$ un endomorfismo de E. El corolario precedente garantiza que, cualquiera que sea la base $(a_1, \ldots, a_n)$ de E, el determinante de la matriz $[f, (a_i)]$ es el mismo. Decimos que este determinante es el *determinante del endomorfismo* f y lo representamos por $\det f$. Hemos puesto, pues,

$$\det f = \det [f, (a_i)] ,$$

cualquiera que sea la base $(a_1, \ldots, a_n)$ de E. Como vimos en (3.2.10), se tiene también

$$\det f = \det_{(a_i)}(f(a_1), \ldots, f(a_n)) .$$

3.2.22 Si E un e.v. de dimensión $n \neq 0$ y $f \in \mathcal{L}(E)$ un endomorfismo de E, las siguientes proposiciones son equivalentes

(i) f es un automorfismo;
(ii) $\det f \neq 0$.

Además, si f es un automorfiamo y f^{-1} es el automorfismo inverso, tenemos que

$$\det f^{-1} = (\det f)^{-1} .$$

3.2.23 Si E un e.v. de dimensión $n \neq 0$ y $f, g \in \mathcal{L}(E)$, entonces

$$\det(g \circ f) = \det g \cdot \det f .$$

3.3 Cálculo de un determinante. Determinantes e inversión de matrices.

3.3.1 Sea $A = [\alpha_i^j] \in M_{\mathbb{K}}(n)$ una matriz cuadrada. Es posible calcular $\det A$ utilizando la fórmula

$$\det A = \sum_{p \in G_n} \epsilon(p)\, \alpha_1^{p(1)} \alpha_2^{p(2)} \cdots \alpha_n^{p(n)} ,$$

pero este método es lento para valores grandes de n, por la gran cantidad de productos que exige realizar. De hecho, no se lo utiliza para $n \geq 4$; para $n = 1, 2$ y 3 la fórmula precedente se encuentra desarrollada en (3.2.6).

En esta sección abordamos el estudio de ciertos métodos que facilitan el cálculo de un determinante.

3.3.2 Si llamamos $c_1, c_2, \ldots, c_n$ a los vectores columna de A (vectores de $\mathbb{K}^n$), tenemos que (v. 3.2.4)

$$\det A = \det_{(e_i)}(c_1, c_2, \ldots, c_n) ,$$

donde (e_i) representa la base canónica de $\mathbb{K}^n$. Utilizando las propiedades de la aplicación $\det_{(e_i)}$, vamos a enunciar algunos resultados prácticos en orden al cálculo de $\det A$.

Como $\det_{(e_i)}$ es una forma n-lineal, se tiene que

$$\det_{(e_i)}(c_1, \ldots, \lambda c_i, \ldots, c_n) = \lambda \det_{(e_i)}(c_1, \ldots, c_i, \ldots, c_n)\,,$$

lo que puede expresarse de la manera siguiente:

> **Regla 1**: Si multiplicamos todos los elementos de una columna de A por un escalar λ, el determinante de la matriz que resulta es $\lambda \cdot \det A$.

$\det_{(e_i)}$ es una forma alternada; aplicando (3.1.23) se obtiene:

> **Regla 2**: Si los vectores columna de A forman un sistema ligado (de $\mathbb{K}^n$), entonces $\det A = 0$.

En particular esta regla tiene como consecuencias inmediatas las siguientes:

> **Regla 2a**: Si una columna de A es nula, entonces $\det A = 0$.

> **Regla 2b**: Si dos columnas de A son proporcionales, resulta $\det A = 0$.

Consideremos el vector

$$c'_i = c_i + \sum_{j \neq i} \lambda^j c_j\,,$$

que es una combinación lineal de todos los vectores columna de A en la que el coeficiente de c_i es 1; substituyendo c_i por c'_i, tenemos

$$\begin{aligned} \det_{(e_i)}(c_1, \ldots, c'_i, \ldots, c_n) &= \det_{(e_i)}(c_1, \ldots, c_i + \sum_{j \neq i} \lambda^j c_j, \ldots, c_n) \\ &= \det_{(e_i)}(c_1, \ldots, c_i, \ldots, c_n) + \sum_{j \neq i} \lambda^j \det_{(e_i)}(c_1, \ldots, c_j, \ldots, c_n)\,. \end{aligned}$$

Para $j \neq i$, el sistema $(c_1, \ldots, c_j, \ldots, c_n)$ tiene el vector c_j en los lugares i y j; su determinante es, pues, nulo y resulta que

$$\det_{(e_i)}(c_1, \ldots, c'_i, \ldots, c_n) = \det_{(e_i)}(c_1, \ldots, c_i, \ldots, c_n)\,,$$

o sea,

> **Regla 3**: El valor de $\det A$ no varía cuando se substituye una columna de A por una combinación lineal de todas las columnas de A en la que el coeficiente que afecta a la columna que se substituye es 1.

En particular tenemos también

> **Regla 3a**: El valor de $\det A$ no varía cuando se suma a una columna de A otra columna multiplicada por un escalar.

Recordemos además que toda forma alternada es antisimétrica; por esta razón, si $p \in G_n$, entonces

$$\det_{(e_i)}(c_{p(1)}, \ldots, c_{p(n)}) = \epsilon(p)\ \det_{(e_i)}(c_1, \ldots, c_n)\,,$$

o sea,

> **Regla 4**: Toda permutación p de las columnas de A transforma $\det A$ en $\epsilon(p) \det A$,

y. en particular, como toda trasposición es una permutación impar, resulta

> **Regla 4a**: El intercambio de dos columnas de A entre sí transforma $\det A$ en $-\det A$.

3.3.3 Sabemos que $\det A^t = \det A$ (v. 3.2.11) y que los vectores columna de A^t son los vectores fila de A. Enunciando las reglas precedentes para las columnas de A^t, obtenemos las mismas reglas para las filas de A.

3.3.4 La utilización repetida de estas reglas permite obtener el determinante de A mediante el cálculo del determinante de una matriz triangular (de la misma dimensión que A). Como vimos en (3.2.12), el determinante de una matriz triangular tiene una expresión sencilla. El ejemplo siguiente aclara el procedimiento que se puede emplear.

3.3.5 Ejemplo. Consideremos la siguiente matriz 4×4 de números reales:

$$A = \begin{bmatrix} 2. & 1.5 & 0 & -1. \\ -3. & -2.25 & 1.5 & -1. \\ 0 & 2. & 2. & 2. \\ 1. & -1. & -1. & 0 \end{bmatrix}.$$

Utilizando la regla 1 para la segunda fila,

$$\det A = \frac{1}{2.} \begin{vmatrix} 2. & 1.5 & 0 & -1. \\ -6. & -4.5 & 3. & -2. \\ 0 & 2. & 2. & 2. \\ 1. & -1. & -1. & 0 \end{vmatrix};$$

sumando a la segunda fila la primera multiplicada por 3. (regla **3a**),

$$\det A = \frac{1}{2.} \begin{vmatrix} 2. & 1.5 & 0 & -1. \\ 0 & 0 & 3. & -5. \\ 0 & 2. & 2. & 2. \\ 1. & -1. & -1. & 0 \end{vmatrix}.$$

Repitiendo ahora el proceso con la cuarta fila se obtiene

$$\det A = \frac{1}{2.\cdot 2.}\begin{vmatrix} 2. & 1.5 & 0 & -1. \\ 0 & 0 & 3. & -5. \\ 0 & 2. & 2. & 2. \\ 2. & -2. & -2. & 0 \end{vmatrix}$$

$$= \frac{1}{2.\cdot 2.}\begin{vmatrix} 2. & 1.5 & 0 & -1. \\ 0 & 0 & 3. & -5. \\ 0 & 2. & 2. & 2. \\ 0 & -3.5 & -2. & 1. \end{vmatrix}.$$

Se suele decir que se ha utilizado el elemento $\alpha_1^1 = 2.$ como 'pivote' para introducir ceros en los lugares de la primera columna situados bajo la diagonal.

De igual manera se procede con las tres últimas filas para introducir ceros en la segunda columna (sin modificar los que hemos introducido en la primera). Es evidente que no se puede utilizar $\alpha_2^2 = 0$ como 'pivote' y resulta necesario hacer el cambio siguiente (regla **4a**)

$$\det A = \frac{-1.}{2.\cdot 2.}\begin{vmatrix} 2. & 1.5 & 0 & -1. \\ 0 & 2. & 2. & 2. \\ 0 & 0 & 3. & -5. \\ 0 & -3.5 & -2. & 1. \end{vmatrix},$$

prosiguiendo ahora como en la primera parte;

$$\det A = \frac{-1.}{2.\cdot 2.\cdot 2.}\begin{vmatrix} 2. & 1.5 & 0 & -1. \\ 0 & 2. & 2. & 2. \\ 0 & 0 & 3. & -5. \\ 0 & -7. & -4. & 2. \end{vmatrix}$$

$$= \frac{-1.}{2.\cdot 2.\cdot 2.}\begin{vmatrix} 2. & 1.5 & 0 & -1. \\ 0 & 2. & 2. & 2. \\ 0 & 0 & 3. & -5. \\ 0 & 0 & 3. & 9. \end{vmatrix}.$$

Se repite ahora el procedimiento para la tercera columna;

$$\det A = \frac{-1.}{2.\cdot 2.\cdot 2.}\begin{vmatrix} 2. & 1.5 & 0 & -1. \\ 0 & 2. & 2. & 2. \\ 0 & 0 & 3. & -5. \\ 0 & 0 & 0 & 14. \end{vmatrix};$$

finalmente

$$\det A = \frac{(-1.)\cdot 2.\cdot 2.\cdot 3.\cdot 14.}{2.\cdot 2.\cdot 2.} = -21.$$

3.3.6 DEFINICIÓN. Sea $A = [\alpha_i^j] \in M(n)$ una matriz cuadrada, con $n \geq 2$. Si $i, j \in \{1, \ldots, n\}$, representamos por A_i^j la matriz $(n-1) \times (n-1)$ que resulta de suprimir en A la columna i y la fila j (la columna y la fila del elemento α_i^j). El determinante, $\det A_i^j$, de esta matriz recibe el nombre de *menor* del elemento α_i^j. La matriz $n \times n$

$$[\det A_i^j]_{i,j \in \{1,\ldots,n\}} ,$$

cuyo elemento general es $\det A_i^j$, se suele denominar 'matriz de los menores' de A.

El escalar

$$\Delta_i^j = (-1)^{i+j} \det A_i^j$$

recibe el nombre de *adjunto* del elemento α_i^j. A la matriz $n \times n$

$$[\Delta_i^j]_{i,j \in \{1,\ldots,n\}} ,$$

formada por los adjuntos de los elementos de A, se le suele llamar 'matriz de los adjuntos' de A.

La traspuesta de la matriz de los adjuntos de A se representa por A^c y se le llama *matriz complementaria* de A. Si $A^c = [\beta_i^j]$, se tiene entonces

$$\beta_i^j = \Delta_j^i = (-1)^{i+j} \det A_j^i .$$

3.3.7 Ejemplo. Para la matriz 3×3 de números reales

$$A = \begin{bmatrix} 6. & 4. & 1. \\ 2. & 3. & 1. \\ 2. & 0 & 0 \end{bmatrix}$$

los menores de los diferentes elementos son

$$\det A_1^1 = \begin{vmatrix} 3. & 1. \\ 0 & 0 \end{vmatrix} = 0 , \quad \det A_2^1 = \begin{vmatrix} 2. & 1. \\ 2. & 0 \end{vmatrix} = -2. , \quad \det A_3^1 = \begin{vmatrix} 2. & 3. \\ 2. & 0 \end{vmatrix} = -6. ,$$

$$\det A_1^2 = \begin{vmatrix} 4. & 1. \\ 0 & 0 \end{vmatrix} = 0 , \quad \det A_2^2 = \begin{vmatrix} 6. & 1. \\ 2. & 0 \end{vmatrix} = -2. , \quad \det A_3^2 = \begin{vmatrix} 6. & 4. \\ 2. & 0 \end{vmatrix} = -8. ,$$

$$\det A_1^3 = \begin{vmatrix} 4. & 1. \\ 3. & 1. \end{vmatrix} = 1. , \quad \det A_2^3 = \begin{vmatrix} 6. & 1. \\ 2. & 1. \end{vmatrix} = 4. , \quad \det A_3^3 = \begin{vmatrix} 6. & 4. \\ 2. & 3. \end{vmatrix} = 10. ;$$

la matriz de los menores es

$$\begin{bmatrix} 0 & -2. & -6. \\ 0 & -2. & -8. \\ 1. & 4. & 10. \end{bmatrix} ;$$

la matriz de los adjuntos es

$$\begin{bmatrix} 0 & 2. & -6. \\ 0 & -2. & 8. \\ 1. & -4. & 10. \end{bmatrix} ,$$

y la matriz complementaria es

$$A^c = \begin{bmatrix} 0 & 0 & 1. \\ 2. & -2. & -4. \\ -6. & 8. & 10. \end{bmatrix}.$$

3.3.8 TEOREMA (Desarrollo de un determinante por los elementos de una línea)

Sea $A = [\alpha_i^j] \in M(n)$ una matriz cuadrada, con $n \geq 2$. Entonces, cualquiera que sea el índice i,

$$\textbf{(4)} \qquad \boxed{\det A = \sum_{j=1}^{n} \alpha_i^j \Delta_i^j}$$

(desarrollo de $\det A$ por los elementos de la columna número i) y, cualquiera que sea el índice j,

$$\textbf{(5)} \qquad \boxed{\det A = \sum_{i=1}^{n} \alpha_i^j \Delta_i^j}$$

(desarrollo de $\det A$ por los elementos de la fila número j).

Demostraremos en primer lugar la segunda fórmula. Representemos por $(e_1, \ldots, e_n)$ la base canónica de $\mathbb{K}^n$ y por $(\overline{e_1}, \ldots, \overline{e_{n-1}})$ la de $\mathbb{K}^{n-1}$. Si $x_1, \ldots, x_n \in \mathbb{K}^n$, con

$$\begin{aligned} x_1 &= (x_1^1, \ldots, x_1^n) \\ &\cdots\cdots\cdots\cdots\cdots\cdots \\ x_n &= (x_n^1, \ldots, x_n^n), \end{aligned}$$

fijemos un índice j y pongamos

$$f(x_1, \ldots, x_n) = \sum_{i=1}^{n} (-1)^{i+j} \, x_i^j \det_{(\overline{e_k})}(x_1', \ldots, x_{i-1}', x_{i+1}', \ldots, x_n'),$$

donde $x_1', \ldots, x_n'$ son los vectores de $\mathbb{K}^{n-1}$ que resultan de suprimir en cada $x_1, \ldots, x_n$ la componente número j, o sea,

$$x_k' = (x_k^1, \ldots, x_k^{j-1}, x_k^{j+1}, \ldots, x_k^n).$$

Es sencillo demostrar que f es una forma n-lineal sobre $\mathbb{K}^n$ (no obstante, escribir la demostración resulta tedioso). Veamos que f es alternada, o sea, que $f(x_1, \ldots, x_n) = 0$ si hay dos vectores iguales en el sistema $(x_1, \ldots, x_n)$. Con el fin de simplificar la escritura supondremos que $x_1 = x_2$; entonces, para $i = 3, \ldots, n$, los términos de la suma que define f son evidentemente nulos; por otro lado

$$\begin{aligned}
&(-1)^{1+j}\, x_1^j \det_{(\overline{e_k})}(x_2', x_3', \ldots, x_n') + (-1)^{2+j}\, x_2^j \det_{(\overline{e_k})}(x_1', x_3', \ldots, x_n') \\
&= (-1)^{1+j} \Big(x_1^j \det_{(\overline{e_k})}(x_2', x_3', \ldots, x_n') - x_2^j \det_{(\overline{e_k})}(x_1', x_3', \ldots, x_n') \Big) \\
&= 0
\end{aligned}$$

puesto que $x_1^j = x_2^j$ y $x_1' = x_2'$. Hemos visto, pues, que f es una forma n-lineal alternada sobre $\mathbb{K}^n$. Además, $f(e_1, \ldots, e_n) = 1$; en efecto, como $e_j' = 0$ (recuérdese que, con las notaciones antes explicadas, e_j' es el vector de $\mathbb{K}^{n-1}$ que resulta de suprimir en e_j la componente número j, que es su única componente no nula), la suma que define $f(e_1, \ldots, e_n)$ se reduce al término con $i = j$, y este término vale 1 puesto que $(-1)^{j+j} = 1$, la componente número j de e_j es 1 y $e_1' = \overline{e_1}$, …, $e_{j-1}' = \overline{e_{j-1}}$, $e_{j+1}' = \overline{e_j}$, …, $e_n' = \overline{e_{n-1}}$.

La definición de $\det_{(e_i)}$ (v. 3.2.2) nos dice entonces que

$$f = \det_{(e_i)} ,$$

o sea, que

$$\det_{(e_i)}(x_1, \ldots, x_n) = \sum_{i=1}^{n} (-1)^{i+j}\, x_i^j \det_{(\overline{e_k})}(x_1', \ldots, x_{i-1}', x_{i+1}', \ldots, x_n') .$$

Escribiendo la igualdad precedente para los vectores columna $c_1, \ldots, c_n$ de A, lo que se obtiene es justamente

$$\det A = \sum_{i=1}^{n} (-1)^{i+j} \alpha_i^j \det A_i^j = \sum_{i=1}^{n} \alpha_i^j \Delta_i^j ,$$

que es la igualdad que nos proponíamos demostrar.

La primera fórmula resulta de aplicar la segunda a la matriz A^t, si se tiene en cuenta el resultado (3.2.11).

3.3.9 El teorema precedente permite reducir el cálculo de un determinante 'de orden n' (o sea, de una matriz $n \times n$) al cálculo de n determinantes de orden $n-1$.

3.3.10 Ejemplo. Para la matriz

$$A = \begin{bmatrix} 6. & 4. & 1. \\ 2. & 3. & 1. \\ 2. & 0 & 0 \end{bmatrix}$$

del ejemplo (3.3.7) tenemos, desarrollando su determinante por los elementos de la primera columna, que

$$\det A = \alpha_1^1 \Delta_1^1 + \alpha_1^2 \Delta_1^2 + \alpha_1^3 \Delta_1^3 = 6. \cdot 0 + 2. \cdot 0 + 2. \cdot 1. = 2.$$

Ha sido necesario calcular 3 determinantes de orden 2. Parte de este trabajo se puede ahorrar eligiendo para el desarrollo una línea con muchos elementos nulos. En el ejemplo que acabamos de ver resulta conveniente calcular $\det A$ desarrollándolo por los elementos de la tercera fila.

3.3.11 Ejemplo. Para la matriz

$$A = \begin{bmatrix} 2. & 1.5 & 0 & -1. \\ -3. & -2.25 & 1.5 & -1. \\ 0 & 2. & 2. & 2. \\ 1. & -1. & -1. & 0 \end{bmatrix}$$

del ejemplo (3.3.5), es menos costoso modificar primero la matriz para conseguir una línea con un sólo elemento distinto de 0. Como ya vimos en dicho ejemplo,

$$\det A = \frac{1}{2.\cdot 2.} \begin{vmatrix} 2. & 1.5 & 0 & -1. \\ 0 & 0 & 3. & -5. \\ 0 & 2. & 2. & 2. \\ 0 & -3.5 & -2. & 1. \end{vmatrix}$$

y, desarrollando este último determinante por los elementos de la primera columna, resulta que

$$\begin{aligned} \det A &= \frac{1}{2.\cdot 2.} 2.\,(-1)^{1+1} \begin{vmatrix} 0 & 3. & -5. \\ 2. & 2. & 2. \\ -3.5 & -2. & 1. \end{vmatrix} \\ &= \frac{1}{2.}\,(-21. + 20. - 35. - 6.) \\ &= \frac{1}{2.}(-42.) = -21. \end{aligned}$$

con lo que sólo ha sido necesario realizar el cálculo de un determinante de orden 3.

3.3.12 PROPOSICIÓN

Sea $A = [\alpha_i^j] \in M(n)$ una matriz cuadrada, con $n \geq 2$. Si representamos por A^c la matriz complementaria de A (v. 3.3.6), entonces

$$A^c\, A = (\det A) I_n\,.$$

Si ponemos $A^c = [\beta_i^j]$ y $A^c A = [\lambda_i^j]$, entonces

$$\lambda_i^j = \sum_{k=1}^{n} \beta_k^j \alpha_i^k = \sum_{k=1}^{n} \alpha_i^k \Delta_j^k\,.$$

Cuando $i = j$, esta expresión es la del desarrollo de $\det A$ por los elementos de la columna número i, luego

$$(\forall i) \quad \lambda_i^i = \det A\,.$$

Cuando $i \neq j$, la expresión

$$\sum_{k=1}^{n} \alpha_i^k \Delta_j^k$$

es el desarrollo por la columna número j del determinante de la matriz

$$\begin{bmatrix} \alpha_1^1 & \cdots & \alpha_i^1 & \cdots & \alpha_{j-1}^1 & \alpha_i^1 & \alpha_{j+1}^1 & \cdots & \alpha_n^1 \\ \vdots & & \vdots & & \vdots & \vdots & \vdots & & \vdots \\ \vdots & & \vdots & & \vdots & \vdots & \vdots & & \vdots \\ \alpha_1^n & \cdots & \alpha_i^n & \cdots & \alpha_{j-1}^n & \alpha_i^n & \alpha_{j+1}^n & \cdots & \alpha_n^n \end{bmatrix}$$

que se obtiene de A cuando en la columna número j se repite la columna número i; este determinante es cero, luego

$$i \neq j \quad \Rightarrow \quad \lambda_i^j = 0\,,$$

lo que termina la demostración.

3.3.13 Se demuestra de la misma manera que

$$A\,A^c = (\det A) I_n\,.$$

3.3.14 COROLARIO

Sea $A \in M(n)$ una matriz cuadrada inversible, con $n \geq 2$. Entonces

$$A^{-1} = (\det A)^{-1} A^c\,,$$

es decir, la inversa de A se obtiene dividiendo la matriz complementaria de A por el escalar $\det A$.

En efecto, como A es inversible, $\det A \neq 0$ y

$$\big((\det A)^{-1} A^c\big)\, A = (\det A)^{-1} (A^c A) = (\det A)^{-1} \det A\, I_n = I_n\,.$$

3.3.15 Ejemplo. Para la matriz

$$A = \begin{bmatrix} 6. & 4. & 1. \\ 2. & 3. & 1. \\ 2. & 0 & 0 \end{bmatrix}$$

del ejemplo (3.3.7), cuyo determinante es 2. (v. 3.3.10) y cuya matriz complementaria es

$$A^c = \begin{bmatrix} 0 & 0 & 1. \\ 2. & -2. & -4. \\ -6. & 8. & 10. \end{bmatrix}\,,$$

se tiene

$$A^{-1} = \frac{1}{2.} \begin{bmatrix} 0 & 0 & 1. \\ 2. & -2. & -4. \\ -6. & 8. & 10. \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0.5 \\ 1. & -1. & -2. \\ -3. & 4. & 5. \end{bmatrix}\,.$$

3.4 Determinantes y rango.

3.4.1 DEFINICIÓN. Sea $A = [\alpha_i^j] \in M(n,m)$ una matriz (no necesariamente cuadrada); sea r un número natural tal que $r \leq n$ y $r \leq m$. Cuando se suprimen $n - r$ columnas y $m - r$ filas de A, los elementos restantes forman una matriz cuadrada $r \times r$; decimos que es una 'matriz $r \times r$ extraída de A'. Del determinante de esta matriz extraída de A decimos que es un *menor de orden* r de A. De las r columnas que no hemos suprimido, $i_1, i_2, \ldots, i_r$, y que numeramos siempre en orden creciente

$$i_1 < i_2 < \cdots < i_r \ ,$$

decimos que son las columnas que 'forman parte de la matriz extraída' y también que 'forman parte del menor'. (Esto último es un claro abuso de lenguaje puesto que el menor no es más que un escalar; sin embargo, está claro que nos referimos no al menor en sí, sino a la matriz extraída de la que se obtiene el menor.) Lo mismo decimos de las r filas

$$j_1 < j_2 < \cdots < j_r \ ,$$

que no hemos suprimido.

3.4.2 Ejemplo. Consideremos la matriz de $M_{\mathbb{R}}(5,3)$

$$A = \begin{bmatrix} 1. & 0 & 0 & -1. & 0.5 \\ 0 & -1. & 2. & 0 & 3. \\ 2. & 1. & 0 & 1.2 & -2. \end{bmatrix} .$$

Suprimiendo las columnas 1, 4 y 5 y la fila 3, obtenemos la matriz extraída, 2×2,

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

y el correspondiente menor de orden 2

$$\begin{vmatrix} 0 & 0 \\ -1. & 2. \end{vmatrix} = 0 \, ;$$

forman parte de la matriz y del menor las columnas 2 y 3 y las filas 1 y 2.

Si suprimimos las columnas 1, 3 y 5 y la fila 2, obtenemos la matriz extraída, 2×2,

$$\begin{bmatrix} 0 & -1. \\ 1. & 1.2 \end{bmatrix}$$

y el menor de orden 2

$$\begin{vmatrix} 0 & -1. \\ 1. & 1.2 \end{vmatrix} = 1. \, ;$$

de los que forman parte las columnas 2 y 4 y las filas 1 y 3.

3.4.3 Ejemplo. Si $A = [\alpha_i^j] \in M(n)$ es una matriz cuadrada, con $n \geq 2$, la matriz A_i^j (v. 3.3.6) es una matriz $(n-1) \times (n-1)$ extraída de A, de la que forman parte las columnas

$$1, \ldots, i-1, i+1, \ldots, n$$

y las filas

$$1, \ldots, j-1, j+1, \ldots, n\,.$$

El correspondiente menor de orden $n-1$, $\det A_i^j$, es justamente el menor del elemento α_i^j.

3.4.4 PROPOSICIÓN

Sea E un e.v. de dimensión $n \neq 0$ y $(a_1, \ldots, a_n)$ una base de E.

a) Supondremos que $x_1, \ldots, x_r$, con $r \leq n$, son vectores de E, con coordenadas

$$\begin{aligned} x_1 &= (x_1^1, \ldots, x_1^n)_{(a_i)} \\ &\ldots\ldots\ldots\ldots\ldots\ldots \\ x_r &= (x_r^1, \ldots, x_r^n)_{(a_i)}\,, \end{aligned}$$

y consideremos la matriz $r \times n$,

$$[(x_1, \ldots, x_r), (a_i)] = \begin{bmatrix} x_1^1 & \cdots & x_r^1 \\ \vdots & & \vdots \\ \vdots & & \vdots \\ x_1^n & \cdots & x_r^n \end{bmatrix}.$$

Si en esta matriz existe un menor no nulo de orden r, entonces el sistema $(x_1, \ldots, x_r)$ es libre.

b) Suponemos ahora que $r \leq n-1$ y que se cumplen las hipótesis del apartado anterior; en particular, existe un menor no nulo de orden r de $[(x_1, \ldots, x_r), (a_i)]$. Suponemos además que

$$x_{r+1} = (x_{r+1}^1, \ldots, x_{r+1}^n)_{(a_i)}$$

es un vector de E tal que, en la matriz

$$[(x_1, \ldots, x_r, x_{r+1}), (a_i)] = \begin{bmatrix} x_1^1 & \cdots & x_r^1 & x_{r+1}^1 \\ \vdots & & \vdots & \vdots \\ \vdots & & \vdots & \vdots \\ x_1^n & \cdots & x_r^n & x_{r+1}^n \end{bmatrix},$$

todos los menores de orden $r+1$ que amplían el menor no nulo son nulos (nos referimos a los menores que resultan de añadir al menor no nulo de orden r la columna $r+1$ y una de las $n-r$ filas posibles). Entonces x_{r+1} es combinación lineal de $(x_1, \ldots, x_r)$.

Para simplificar la notación, supondremos que el menor de orden r en el que intervienen las r primeras filas de $[(x_1, \ldots, x_r), (a_i)]$ es no nulo, o sea, que

$$\begin{vmatrix} x_1^1 & \cdots & x_r^1 \\ \vdots & & \vdots \\ x_1^r & \cdots & x_r^r \end{vmatrix} \neq 0 \,.$$

a) Como vimos en (2.6.21), la aplicación

$$\begin{aligned} f: \qquad E &\to \mathbb{K}^r \\ (x^1, \ldots, x^n)_{(a_i)} &\to (x^{i_1}, \ldots, x^{i_r}) \end{aligned}$$

es lineal; para esta aplicación

$$\begin{aligned} f(x_1) &= (x_1^1, \ldots, x_1^r) \\ &\cdots\cdots\cdots\cdots \\ f(x_r) &= (x_r^1, \ldots, x_r^r) \,. \end{aligned}$$

Si utilizamos la proposición (3.2.8) para el espacio $\mathbb{K}^r$, su base canónica y el sistema $(f(x_1), \ldots, f(x_r)\,)$, tenemos que este sistema es una base de $\mathbb{K}^r$; en particular, es un sistema libre, y entonces $(x_1, \ldots, x_r)$ es un sistema libre (v. 2.1.13c).

b) En primer lugar, como $(f(x_1), \ldots, f(x_r)\,)$ es una base de $\mathbb{K}^r$, entonces

$$f(x_{r+1}) = \lambda^1 f(x_1) + \cdots + \lambda^r f(x_r)$$

y, además, los escalares $\lambda^1, \ldots, \lambda^r$ que verifican tal igualdad son únicos. Se tiene, pues, que

$$\begin{aligned} x_{r+1}^1 &= \lambda^1 x_1^1 + \cdots + \lambda^r x_r^1 \\ &\cdots\cdots\cdots\cdots\cdots \\ x_{r+1}^r &= \lambda^1 x_1^r + \cdots + \lambda^r x_r^r \,. \end{aligned}$$

Es suficiente probar ahora que, cualquiera que sea el índice $k = r+1, \ldots, n$, también se tiene

$$x_{r+1}^k = \lambda^1 x_1^k + \cdots + \lambda^r x_r^k$$

pues, probado esto, se tiene que

$$x_{r+1} = \lambda^1 x_1 + \cdots + \lambda^r x_r$$

y esto es justamente lo que pretendemos demostrar.

Sea, pues, $k \in \{r{+}1, \ldots, n\}$; conviene considerar las dos aplicaciones siguientes:

$$\begin{aligned} g: \qquad E &\to \mathbb{K}^{r+1} \\ (x^1, \ldots, x^n)_{(a_i)} &\to (x^1, \ldots, x^r, x^k) \end{aligned}$$

y

$$\begin{aligned} \pi: \qquad\qquad \mathbb{K}^{r+1} &\to \mathbb{K}^r \\ (\mu^1,\dots,\mu^r,\mu^{r+1}) &\to (\mu^1,\dots,\mu^r)\,. \end{aligned}$$

Ambas son lineales (v. 2.6.21) y además

$$f = \pi \circ g\,.$$

El sistema $(f(x_1),\dots,f(x_r))$ de $\mathbb{K}^r$ es libre, es decir, $(\pi\circ g(x_1),\dots,\pi\circ g(x_r))$ es libre, luego el sistema $(g(x_1),\dots,g(x_r))$ de $\mathbb{K}^{r+1}$ es libre. Por otro lado, el sistema $(g(x_1),\dots,g(x_r),g(x_{r+1}))$ es ligado, puesto que, por hipótesis,

$$\begin{vmatrix} x_1^1 & \cdots & x_r^1 & x_{r+1}^1 \\ \vdots & & \vdots & \vdots \\ x_1^r & \cdots & x_r^r & x_{r+1}^r \\ x_1^k & \cdots & x_r^k & x_{r+1}^k \end{vmatrix} = 0\,.$$

Resulta entonces (v. 1.4.11) que

$$g(x_{r+1}) = \mu^1 g(x_1) + \cdots + \mu^r g(x_r)$$

y, como π es lineal,

$$\pi\circ g(x_{r+1}) = \mu^1 \pi\circ g(x_1) + \cdots + \mu^r \pi\circ g(x_r)\,,$$

es decir,

$$f(x_{r+1}) = \mu^1 f(x_1) + \cdots + \mu^r f(x_r)\,.$$

La unicidad de las coordenadas de $f(x_{r+1})$ en la base $(f(x_1),\dots,f(x_r))$ hace que

$$\lambda^1 = \mu^1,\dots,\lambda^r = \mu^r$$

y que, por lo tanto,

$$g(x_{r+1}) = \lambda^1 g(x_1) + \cdots + \lambda^r g(x_r)\,;$$

en particular, esto implica que

$$x_{r+1}^k = \lambda^1 x_1^k + \cdots + \lambda^r x_r^k\,,$$

que es la igualdad que queríamos probar.

3.4.5 Ejemplo. Los vectores de $\mathbb{R}^4$

$$x_1 = (1.,0,-1.,2.)\,,\quad x_2 = (1.,0,1.,2.)\,,$$

forman un sistema libre puesto que, en la matriz

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

el menor (del que forman parte las filas 1 y 3)

$$\begin{vmatrix} 1. & 1. \\ -1. & 1. \end{vmatrix} = 2.$$

es no nulo.

El vector

$$x_3 = (3., 0, -1., 6.)$$

es combinación lineal de (x_1, x_2) puesto que los dos menores de orden 3 de la matriz

$$\begin{bmatrix} 1. & 1. & 3. \\ 0 & 0 & 0 \\ -1. & 1. & -1. \\ 2. & 2. & 6. \end{bmatrix}$$

que amplían

$$\begin{vmatrix} 1. & 1. \\ -1. & 1. \end{vmatrix}$$

son nulos, es decir, son nulos los menores

$$\begin{vmatrix} 1. & 1. & 3. \\ 0 & 0 & 0 \\ -1. & 1. & -1. \end{vmatrix} \quad \text{y} \quad \begin{vmatrix} 1. & 1. & 3. \\ -1. & 1. & -1. \\ 2. & 2. & 6. \end{vmatrix}.$$

3.4.6 PROPOSICIÓN

Sea $A \in M(n,m)$, $A \neq 0$, una matriz (no necesariamente cuadrada). El rango de A coincide con el orden máximo de un menor no nulo extraído de A.

Sea r el orden máximo de un menor no nulo extraído de A; como $A \neq 0$, necesariamente $r \geq 1$. El número natural r verifica que

— existe un menor no nulo de orden r;

— cualquier menor de orden $r+1$ es nulo.

Representamos por $c_1, \ldots, c_n$ los vectores columna de A; si las columnas $c_{i_1}, \ldots, c_{i_r}$ son las que forman parte del menor no nulo de orden r, entonces $(c_{i_1}, \ldots, c_{i_r})$ es un sistema libre (por el primer apartado de la proposición anterior); además, si c_i es algún vector columna distinto de los anteriores, c_i es combinación lineal de $(c_{i_1}, \ldots, c_{i_r})$ (por el segundo apartado de la proposición anterior). Resulta así que

$$\langle c_{i_1}, \ldots, c_{i_r} \rangle = \langle c_1, \ldots, c_n \rangle$$

y que $(c_{i_1}, \ldots, c_{i_r})$ es una base de $\langle c_1, \ldots, c_n \rangle$. Esto demuestra que el rango de A es r.

3.4.7 COROLARIO

Sea E un e.v. de dimensión $n \neq 0$ y $(a_1, \ldots, a_n)$ una base de E. Suponemos que $x_1, \ldots, x_r$, con $r \leq n$, son vectores de E, con coordenadas

$$\begin{aligned} x_1 &= (x_1^1, \ldots, x_1^n)_{(a_i)} \\ &\cdots\cdots\cdots\cdots\cdots\cdots \\ x_r &= (x_r^1, \ldots, x_r^n)_{(a_i)} , \end{aligned}$$

y consideramos la matriz $r \times n$,

$$[(x_1, \ldots, x_r), (a_i)] = \begin{bmatrix} x_1^1 & \cdots & x_r^1 \\ \vdots & & \vdots \\ \vdots & & \vdots \\ x_1^n & \cdots & x_r^n \end{bmatrix} .$$

Si todos los menores de orden r de esta matriz son nulos, entonces el sistema $(x_1, \ldots, x_r)$ es ligado. (Compárese este enunciado con la primera parte de 3.4.4 .)

El rango de $x_1, \ldots, x_r$ coincide con el de la matriz $[(x_1, \ldots, x_r), (a_i)]$ (v. 2.5.22), que es menor que r por la proposición precedente. Luego $(x_1, \ldots, x_r)$ es un sistema ligado (v. 1.5.9).

3.4.8 COROLARIO

a) Sea E un e.v. de dimensión $n \neq 0$ y $(a_1, \ldots, a_n)$ una base de E; sea $(x_1, \ldots, x_p)$ un sistema de E. Se considera la matriz

$$A = [(x_1, \ldots, x_p), (a_i)]$$

formada por las coordenadas de los vectores del sistema. Si $\operatorname{rg} A = r \neq 0$ y las columnas $i_1, \ldots, i_r$ son las que forman parte de un menor no nulo de orden r de A, entonces $(x_{i_1}, \ldots, x_{i_r})$ es una base de $\langle x_1, \ldots, x_p \rangle$.

b) Sean E y E' dos e.v. de dimensiones n y m; sean $(a_1, \ldots, a_n)$ y $(b_1, \ldots, b_m)$ bases de E y E'; sea $f \in \mathcal{L}(E, E')$ y

$$A = [f, (a_i), (b_j)] .$$

Supongamos que $\operatorname{rg} A = r \neq 0$. Si las columnas $i_1, \ldots, i_r$ son las que forman parte de un menor no nulo de orden r de A, entonces $(f(a_{i_1}), \ldots, f(a_{i_r}))$ es una base de $\operatorname{Im} f$.

La demostración de ambos apartados es justamente el razonamiento que hemos hecho en (3.4.6).

3.4.9 La proposición (3.4.6) proporciona un método para calcular el rango de una matriz A. En principio el método exige calcular todos los menores de A hasta encontrar r tal que exista un menor no nulo de orden r y que todos los menores de orden $r+1$ sean nulos.

De hecho, si encontramos un menor no nulo de orden r y son nulos todos los menores de orden $r+1$ que amplían dicho menor, podemos afirmar que el rango de A es r, como se observa repasando la demostración de (3.4.6). (Además, en consecuencia, serán nulos todos los menores de orden $r+1$, incluídos los que no amplían el citado menor.)

Este razonamiento permite abreviar considerablemente el trabajo necesario para calcular el rango.

3.4.10 Ejemplo. La matriz

$$A = \begin{bmatrix} 1. & 0 & -1. & 2. \\ 1. & 0 & 1. & 2. \\ 3. & 0 & -1. & 6. \end{bmatrix}$$

de $M_{\mathbb{R}}(4,3)$ tiene rango 2, puesto que

$$\begin{vmatrix} 1. & -1. \\ 1. & 1. \end{vmatrix} = 2. \neq 0$$

y que

$$\begin{vmatrix} 1. & 0 & -1. \\ 1. & 0 & 1. \\ 3. & 0 & -1. \end{vmatrix} = 0 \quad \text{y} \quad \begin{vmatrix} 1. & -1. & 2. \\ 1. & 1. & 2. \\ 3. & -1. & 6. \end{vmatrix} = 0\,.$$

Obsérvese que sólo ha sido necesario calcular dos de los cuatro menores de orden 3. Compruébese que los dos restantes son también nulos.

Capítulo 4

Sistemas de ecuaciones lineales.

En este capítulo ponemos especial énfasis en el estudio de la existencia de soluciones de los sistemas lineales, aunque también proporcionamos métodos para la resolución de los mismos. No se hace, sin embargo, un estudio de otros algoritmos usuales en la resolución numérica de sistemas lineales y tampoco se abordan métodos iterativos de aproximación de las soluciones.

4.1 Estudio general de un sistema.

4.1.1 Consideremos las dos igualdades siguientes

$$\begin{aligned} x + 2z &= 0 \\ y + 3z &= -2\,, \end{aligned}$$

que constituyen lo que se suele llamar un sistema lineal de dos ecuaciones con tres incógnitas. Lo que se pretende es encontrar números (reales en este caso) x, y, z, que cumplan simultáneamente ambas igualdades. Un simple 'tanteo' o ensayo nos permite encontrar números como

$$x = 2.\,, \quad y = 1. \quad \text{y} \quad z = -1.$$

que cumplen las condiciones requeridas, o también

$$x = -3.\,, \quad y = -6.5 \quad \text{y} \quad z = 1.5\,;$$

se suele decir que estos dos grupos, cada uno de tres números, son dos soluciones del sistema lineal.

Si pretendemos encontrar todas las soluciones y no sólo algunas, el problema se vuelve un poco más difícil. También aumenta la dificultad cuando el número de ecuaciones y de incógnitas es elevado. Para resolver estas dificultades conviene tener métodos más seguros que el simple ensayo de números; son los métodos que tratamos de explicar en este capítulo.

Volvamos de nuevo al sistema que hemos planteado; decir que tres números x, y y z forman una solución del sistema equivale a decir que dichos números verifican

$$x(1,0) + y(0,1) + z(2,3) = (0,-2)\,,$$

o sea, que tales números son los coeficientes de una combinación lineal de los vectores $(1,0)$, $(0,1)$ y $(2,3)$ de $\mathbb{R}^2$, cuya suma es el vector $(0,-2)$. Y aún otro punto de vista; si consideramos la aplicación lineal $f : \mathbb{R}^3 \to \mathbb{R}^2$ dada por

$$f(x,y,z) = (x+2z, y+3z)\,,$$

las soluciones del sistema son justamente los vectores de $\mathbb{R}^3$ cuya imagen por f es el vector $(0,-2)$ de $\mathbb{R}^2$. Tales formas de enfocar el problema son más convenientes cuando se pretende un estudio general del mismo, y son los que nos permitirán llegar a métodos eficaces de resolución.

4.1.2 Empezaremos por el tratamiento más abstracto posible. Si A y B son dos conjuntos, $f : A \to B$ una aplicación y $b \in B$, decimos que la expresión

$$\textbf{(ec)} \quad f(x) = b$$

es una *ecuación*. De todo elemento $a \in A$ tal que

$$f(a) = b$$

decimos que es una *solución* de la ecuación **(ec)**.

Obsérvese que el símbolo x no tiene ninguna importancia, como no sea la de señalarnos el lugar en que debemos situar un elemento $a \in A$ para comprobar si tal elemento es una solución de la ecuación. La ecuación podría escribirse

$$f(\ ) = b\,,$$

pero es clásico incluir el símbolo x (u otro similar); este símbolo recibe el nombre de 'incógnita'.

El sistema lineal de (4.1.1) es una ecuación de este tipo, en la que $A = \mathbb{R}^3$, $B = \mathbb{R}^2$, $b = (0,-2)$ y f es la aplicación

$$f(x,y,z) = (x+2z, y+3z)\,;$$

la incógnita es el vector (x,y,z).

4.1.3 Supongamos que f es una aplicación sobre; entonces, sin que importe cuál es el elemento $b \in B$, sabemos que la ecuación **(ec)** admite por lo menos una solución. Sin embargo, si f no es sobre, la ecuación **(ec)** admite solución si $b \in \operatorname{Im} f$, pero no tiene ninguna solución cuando $b \notin \operatorname{Im} f$.

Si f es inyectiva, sabemos que la ecuación **(ec)** admite a lo más una solución.

Si f es biyectiva, resulta entonces que la ecuación **(ec)** admite una solución y además una sola.

4.1.4 DEFINICIÓN. Las ecuaciones que constituyen el objeto de este capítulo son aquellas en las cuales f es una aplicación lineal

$$f : \mathbb{K}^n \to \mathbb{K}^m$$

y $b = (b^1, \ldots, b^m)$ es un vector de $\mathbb{K}^m$. De una ecuación

(sl) $f(x) = b$

de este tipo, decimos que es un *sistema lineal de m ecuaciones con n incógnitas* (con coeficientes en el cuerpo conmutativo $\mathbb{K}$); la razón de esta terminología se pondrá enseguida de manifiesto.

Las soluciones del sistema son aquellos vectores

$$x = (x^1, \dots, x^n) \in \mathbb{K}^n$$

para los que $f(x) = b$, o sea,

(sl1) $f(x^1, \dots, x^n) = b\,;$

cada uno de los n símbolos $x^1, \dots, x^n$ es una *incógnita* del sistema.

Representemos por $A = [\alpha_i^j]$ la matriz $n \times m$ asociada a f en las bases canónicas de $\mathbb{K}^n$ y $\mathbb{K}^m$. La igualdad **(sl)** equivale entonces a (v. 2.5.1)

$$\textbf{(sl2)} \quad A \begin{bmatrix} x^1 \\ \vdots \\ x^n \end{bmatrix} = \begin{bmatrix} b^1 \\ \vdots \\ b^m \end{bmatrix}.$$

Decimos que A es la *matriz del sistema* **(sl)**.

Obsérvese que la igualdad **(sl2)** equivale a su vez al siguiente conjunto de igualdades:

$$\textbf{(sl3)} \quad \begin{array}{l} \alpha_1^1 x^1 \;+\cdots+\; \alpha_i^1 x^i \;+\cdots+\; \alpha_n^1 x^n \;=\; b^1 \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\ \alpha_1^j x^1 \;+\cdots+\; \alpha_i^j x^i \;+\cdots+\; \alpha_n^j x^n \;=\; b^j \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\ \alpha_1^m x^1 \;+\cdots+\; \alpha_i^m x^i \;+\cdots+\; \alpha_n^m x^n \;=\; b^m\,. \end{array}$$

De cada una de las m igualdades precedentes, decimos que es una *ecuación del sistema.* Para todo $j = 1, \dots, m$, la ecuación

$$\alpha_1^j x^1 + \cdots + \alpha_i^j x^i + \cdots + \alpha_n^j x^n = b^j$$

se puede escribir

$$f_j(x) = b^j\,,$$

representando por f_j la forma lineal, $f_j \in (\mathbb{K}^n)^*$, dada por la fila número j de A. O sea, el sistema **(sl)** se escribe también

$$\textbf{(sl4)} \quad \begin{array}{l} f_1(x) \;=\; b^1 \\ \dots\dots\dots\dots \\ f_j(x) \;=\; b^j \\ \dots\dots\dots\dots \\ f_m(x) \;=\; b^m\,. \end{array}$$

Decimos que $f_1, \dots, f_m$ son las *formas fila* del sistema.

Vamos a representar por $c_1, \dots, c_n$ los vectores columna de la matriz A (vectores de $\mathbb{K}^m$), a los que llamaremos *columnas del sistema*. Un vector $x = (x^1, \dots, x^n)$ de K^n es una solución del sistema **(sl)** si y sólo si

$$\textbf{(sl5)} \quad x^1 c_1 + \cdots x^i c_i + \cdots + x^n c_n = b\,;$$

se obtiene así una nueva forma de representar el sistema **(sl)**.

4.1.5 Para los sistemas

$$\textbf{(sl)} \quad f(x) = b$$

que pretendemos estudiar son válidas las conclusiones de (4.1.3). Además, el hecho de que f sea una aplicación lineal entre los espacios vectoriales $\mathbb{K}^n$ y $\mathbb{K}^m$ nos permite obtener más información aún sobre las soluciones. Veamos un primer resultado.

Un sistema lineal del tipo

$$\textbf{(sh)} \quad f(x) = 0\,,$$

en el que el 'término independiente' es el vector 0 de $\mathbb{K}^m$, recibe el nombre de *sistema homogéneo*. Decimos también que el sistema **(sh)** es el 'sistema homogéneo asociado al sistema **(sl)**'.

Las soluciones del sistema homogéneo **(sh)** son los elementos de $\operatorname{Ker} f$ y, en particular, forman un subespacio de $\mathbb{K}^n$. Esto significa además que el vector 0 es siempre una solución del sistema homogéneo; se suele decir que es la solución 'trivial'.

4.1.6 DEFINICIÓN. Consideremos un sistema lineal de m ecuaciones y n incógnitas

$$\textbf{(sl)} \quad f(x) = b\,,$$

donde $f \in \mathcal{L}(\mathbb{K}^n, \mathbb{K}^m)$ y $b = (b^1, \dots, b^m) \in \mathbb{K}^m$. Sea $A = [\alpha_i^j] \in M(n, m)$ la matriz del sistema. Decimos que el número

$$r = \operatorname{rg} f = \operatorname{rg} A$$

es el *rango del sistema*; ya sabemos que $r \leq n$ y $r \leq m$. En lo que sigue supondremos que $r \neq 0$.

Vimos en (3.4.6) que existe en A un menor no nulo de orden r; llamemos $i_1, \dots, i_r$ a los índices de las columnas de A que forman parte de dicho menor y $j_1, \dots, j_r$ a los índices de las filas que forman parte del menor. Las incógnitas $x^{i_1}, \dots, x^{i_r}$ son las que denominaremos *incógnitas principales* del sistema; también llamaremos *columnas principales* a las columnas $c_{i_1}, \dots, c_{i_r}$. Análogamente, llamaremos *ecuaciones principales* y *formas fila principales* del sistema a las correspondientes a los índices $j_1, \dots, j_r$.

Además, vamos a considerar el sistema (de r ecuaciones y n incógnitas) formado por las r ecuaciones principales (completas). Decimos que este sistema es el *subsistema principal*. Nótese que el subsistema principal es también de rango r.

4.1.7 En general existirán varios menores no nulos de orden r de A y, por lo tanto, varias formas (todas válidas) de escoger los elementos principales del sistema y de escoger el subsistema principal.

4.1.8 Ejemplo. Consideremos el sistema lineal, con coeficientes en $\mathbb{R}$, de 4 ecuaciones y 3 incógnitas

$$\begin{array}{rcrcrcr} & & y & + & z & = & 0 \\ & & 2y & + & 2z & = & 2 \\ -x & + & y & & & = & -1 \\ x & & & + & z & = & 2 \end{array} ,$$

cuya matriz

$$A = \begin{bmatrix} 0 & 1. & 1. \\ 0 & 2. & 2. \\ -1. & 1. & 0 \\ 1. & 0 & 1. \end{bmatrix}$$

es de rango 2. De entre los menores no nulos de orden 2 podemos elegir

$$\begin{vmatrix} 1. & 0 \\ 0 & 1. \end{vmatrix} = 1. \neq 0 ,$$

menor formado por las columnas 2 y 3 y las filas 3 y 4. Con esta elección convertimos y y z en incógnitas principales; el subsistema principal es

$$\begin{array}{rcrcrcr} -x & + & y & & & = & -1 \\ x & & & + & z & = & 2 \end{array} .$$

4.1.9 Vamos a ocuparnos en primer lugar de averiguar cuándo un sistema admite alguna solución. De un sistema que admite solución se dice que es *compatible* (o posible); en caso contrario se suele decir que es *incompatible* (o imposible).

4.1.10 PROPOSICIÓN

Todo sistema homogéneo

(sh) $f(x) = 0$

es siempre compatible.

Es justamente lo que hemos visto en (4.1.5).

4.1.11 PROPOSICIÓN

Todo sistema

(sl) $f(x) = b$

cuyo rango coincida con el número de ecuaciones ($r = m$) es compatible.

Como $\operatorname{rg} f = r = m = \dim \mathbb{K}^m$, resulta que la aplicación f es sobre (v. 2.1.23) y entonces, como hemos dicho en (4.1.3), existe siempre una solución.

4.1.12 Hemos visto que el sistema **(sl)** admite una solución $x = (x^1, \ldots, x^n) \in K^n$ si y sólo si

$$x^1 c_1 + \cdots + x^i c_i + \cdots + x^n c_n = b\,,$$

donde $c_1, \ldots, c_n$ son las columnas del sistema. Así, pues, el sistema es compatible si y sólo si

$$b \in \langle c_1, \ldots, c_n \rangle\,;$$

a su vez, esto equivale a que

$$\operatorname{rg}(c_1, \ldots, c_n, b) = \operatorname{rg}(c_1, \ldots, c_n)\,.$$

Es decir, podemos averiguar si el sistema **(sl)** es compatible calculando los rangos de dos sistemas de vectores de $\mathbb{K}^m$. Este método se suele enunciar en la forma matricial que describimos a continuación.

4.1.13 DEFINICIÓN. Sea

(sl) $f(x) = b$

un sistema lineal de m ecuaciones y n incógnitas, A la matriz del sistema y $c_1, \ldots, c_n$ sus columnas. La matriz A', de $n+1$ columnas y m filas, cuyos vectores columna son $c_1, \ldots, c_n$ y b, recibe el nombre de *matriz ampliada* del sistema. La matriz A' se forma, pues, añadiendo a A una columna más, formada por las componentes de b.

4.1.14 TEOREMA

Sea

(sl) $f(x) = b$

un sistema lineal, A la matriz del sistema y A' la matriz ampliada. El sistema es compatible si y sólo si

$$\operatorname{rg} A = \operatorname{rg} A'\,.$$

Es el resultado que hemos visto en (4.1.12), teniendo en cuenta que el rango de una matriz es el rango del sistema formado por sus vectores columna.

4.1.15 Ejemplo. Para el sistema del ejemplo (4.1.8), la matriz ampliada es

$$A' = \begin{bmatrix} 0 & 1. & 1. & 0 \\ 0 & 2. & 2. & 2. \\ -1. & 1. & 0 & -1. \\ 1. & 0 & 1. & 2. \end{bmatrix} .$$

Como

$$\begin{vmatrix} 1. & 1. & 0 \\ 1. & 0 & -1. \\ 0 & 1. & 2. \end{vmatrix} = -1. \neq 0 ,$$

resulta que $\operatorname{rg} A' = 3$ y que el sistema es incompatible.

4.1.16 Para los sistemas compatibles, existe una importante relación entre sus soluciones y las soluciones del sistema homogéneo asociado; es la relación que explicamos en el teorema que sigue.

4.1.17 TEOREMA

Sea

(sl) $f(x) = b$

un sistema lineal compatible y

(sh) $f(x) = 0$

el sistema homogéneo asociado. Supongamos que x_0 es una solución de **(sl)**. Entonces las soluciones de **(sl)** son los vectores del conjunto

$$x_0 + \operatorname{Ker} f ,$$

o sea, los vectores de la forma

$$x_0 + z ,$$

donde $z \in \operatorname{Ker} f$.

Con otras palabras, las soluciones de **(sl)** se obtienen sumando a una solución x_0 de **(sl)** cada una de las soluciones del sistema homogéneo **(sh)**.

Recordemos en primer lugar que las soluciones de **(sh)** son justamente los vectores de $\operatorname{Ker} f$ (v. 4.1.5).

Si $z \in \operatorname{Ker} f$, entonces

$$f(x_0 + z) = f(x_0) + f(z) = f(x_0) = b ,$$

luego $x_0 + z$ es una solución de **(sl)**.

Recíprocamente, si x es una solución de **(sl)**, entonces

$$f(x - x_0) = f(x) - f(x_0) = b - b = 0 ,$$

o sea, el vector $z = x - x_0$ pertenece a $\operatorname{Ker} f$ y $x = x_0 + z$.

4.1.18 COROLARIO

Sea

(sl) $f(x) = b$

un sistema lineal de m ecuaciones y n incógnitas; supongamos que el sistema es compatible. Entonces las soluciones de **(sl)** son las del subsistema principal.

Llamemos r al rango del sistema y representemos por

(sp) $\tilde{f}(x) = \tilde{b}$

el subsistema principal. Recuérdese (v. 4.1.6) que **(sp)** es un sistema con r ecuaciones, n incógnitas y rango r. Para facilitar las notaciones supondremos que el subsistema principal está formado por las r primeras ecuaciones.

Si

$$f(x) = (f_1(x), \ldots, f_r(x), \ldots, f_m(x)) \quad \text{y} \quad b = (b^1, \ldots, b^r, \ldots, b^m)\,,$$

donde $f_1, \ldots, f_m$ son las formas fila del sistema, entonces

$$\tilde{f}(x) = (f_1(x), \ldots, f_r(x)) \quad \text{y} \quad \tilde{b} = (b^1, \ldots, b^r)\,.$$

En primer lugar, está claro que

$$\operatorname{Ker} f \subset \operatorname{Ker} \tilde{f}\,;$$

como además

$$\dim \operatorname{Ker} f = n - r = \dim \operatorname{Ker} \tilde{f}\,,$$

resulta que

$$\operatorname{Ker} f = \operatorname{Ker} \tilde{f}\,.$$

Hemos supuesto que el sistema es compatible; sea x_0 una solución de **(sl)**. El vector x_0 es también una solución de **(sp)**. Pero, como hemos visto en el teorema precedente, las soluciones de **(sl)** son los vectores del conjunto

$$x_0 + \operatorname{Ker} f\,,$$

y las de **(sp)** son los vectores del conjunto

$$x_0 + \operatorname{Ker} \tilde{f}\,.$$

Ambos conjuntos coinciden, lo que termina la demostración.

4.1.19 En el corolario que acabamos de ver es fundamental suponer que el sistema es compatible. Cuando el sistema es incompatible, no admite soluciones, mientras que el subsistema principal admite siempre alguna solución puesto que su rango coincide con el número de ecuaciones.

4.1.20 Ejemplo. El sistema del ejemplo (4.1.8) no posee ninguna solución, como vimos en (4.1.15). El subsistema principal,

$$\begin{array}{rcrcrcr} -x & + & y & & & = & -1 \\ x & & & + & z & = & 2 \end{array} ,$$

sí admite soluciones; por ejemplo,

$$x = 1 \,, \quad y = 0 \,, \quad z = 1$$

es una solución del subsistema principal.

4.2 Obtención de las soluciones de un sistema.

4.2.1 Consideremos un sistema lineal de m ecuaciones y n incógnitas

(sl) $f(x) = b \,,$

donde $f \in \mathcal{L}(\mathbb{K}^n, \mathbb{K}^m)$ y $b = (b^1, \ldots, b^m) \in \mathbb{K}^m$. Representemos por $A = [\alpha_i^j]$ la matriz del sistema y por r ($\neq 0$) el rango del sistema. En esta sección expondremos, de forma práctica, la manera de obtener todas las soluciones del sistema, o sea, todos los vectores $x = (x^1, \ldots, x^n)$ tales que $f(x) = b$.

4.2.2 Caso $n = m = r$ (el número de incógnitas, el número de ecuaciones y el rango coinciden). En este caso la matriz A es cuadrada y $\det A \neq 0$; de un sistema de este tipo decimos que es un *sistema de Cramer*[1]. La aplicación f es biyectiva (v. 2.1.24) y, por lo tanto, el sistema admite una única solución (v. 4.1.3).

Cuando el sistema es además homogéneo, la única solución es la trivial (v. 4.1.5), es decir,

$$x^1 = x^2 = \cdots = x^n = 0 \,.$$

Hay muchas formas prácticas de calcular la solución de un sistema de Cramer. Vamos a exponer en primer lugar el llamado *método de Cramer*, que se basa en las propiedades de los determinantes, Sea $x = (x^1, \ldots, x^n)$ la solución del sistema; sabemos que

$$x^1 c_1 + \cdots + x^n c_n = b \,,$$

donde $c_1, \ldots, c_n$ son las columnas del sistema. Aplicando las reglas **1** y **3** del cálculo de determinantes (v. 3.3.2), se tiene para $i = 1, \ldots, n$

$$\begin{aligned} x^i \det A &= x^i \det_{(e_k)}(c_1, \ldots, c_i, \ldots, c_n) \\ &= \det_{(e_k)}(c_1, \ldots, x^i c_i, \ldots, c_n) \\ &= \det_{(e_k)}(c_1, \ldots, x^1 c_1 + \cdots + x^i c_i + \cdots + x^n c_n, \ldots, c_n) \\ &= \det_{(e_k)}(c_1, \ldots, b, \ldots, c_n) \,, \end{aligned}$$

[1]de Gabriel Cramer (1704-1752), matemático nacido en Ginebra.

o sea,

$$x^i = \frac{\det B_i}{\det A} ,$$

donde B_i representa la matriz que resulta de substituir en A la columna número i por la columna

$$b = \begin{bmatrix} b^1 \\ \vdots \\ b^n \end{bmatrix} ,$$

es decir,

$$B_i = \begin{bmatrix} \alpha_1^1 & \cdots & b^1 & \cdots & \alpha_n^1 \\ \vdots & & \vdots & & \vdots \\ \vdots & & \vdots & & \vdots \\ \alpha_1^n & \cdots & b^n & \cdots & \alpha_n^n \end{bmatrix} .$$

Estas fórmulas permiten entonces calcular la solución.

El método de Cramer es eficaz para pequeños sistemas (para $n = 3$ o menor), y a veces también es recomendable cuando el sistema posee coeficientes que son indeterminados; sim embargo, requiere un número muy grande de operaciones cuando la dimensión del sistema es grande.

4.2.3 Ejemplo. Para el sistema con coeficientes en $\mathbb{R}$

$$\begin{aligned} x + 2z &= 0 \\ y + 3z &= -2 \\ 2x + 2z &= 2 \end{aligned} , \quad A = \begin{bmatrix} 1. & 0 & 2. \\ 0 & 1. & 3. \\ 2. & 0 & 2. \end{bmatrix} ,$$

se tiene $n = 3$, $m = 3$ y, como $\det A = -2. \neq 0$, $r = 3$. Se trata, pues, de un sistema de Cramer, cuya única solución viene dada por

$$x = \frac{\begin{vmatrix} 0 & 0 & 2. \\ -2. & 1. & 3. \\ 2. & 0 & 2. \end{vmatrix}}{\det A} = \frac{-4.}{-2.} = 2. ,$$

$$y = \frac{\begin{vmatrix} 1. & 0 & 2. \\ 0 & -2. & 3. \\ 2. & 2. & 2. \end{vmatrix}}{\det A} = \frac{-2.}{-2.} = 1. ,$$

$$z = \frac{\begin{vmatrix} 1. & 0 & 0 \\ 0 & 1. & -2. \\ 2. & 0 & 2. \end{vmatrix}}{\det A} = \frac{2.}{-2.} = -1.$$

La solución es entonces

$$x = 2. , \quad y = 1. , \quad z = -1.$$

4.2.4 Vamos a exponer otro método que requiere menor número de operaciones que el de Cramer cuando la dimensión es grande. Se trata del *método de Gauss*[2] y utiliza las operaciones elementales que vimos en (2.4.31).

Es fácil comprobar que si
$$A\,x = b\,,$$
con $A \in M(n,m)$, es un sistema con matriz ampliada
$$[\,A\,|\,b\,]$$
y transformamos esta matriz ampliada mediante una operación elemental cualquiera en otra matriz
$$[\,A'\,|\,b'\,]\ ,$$
el sistema
$$A'\,x = b'$$
que resulta posee las mismas soluciones que el primitivo. Se dice que ambos sistemas son *equivalentes*.

Si ahora
$$A\,x = b$$
es un sistema de Cramer, es decir, A es $n \times n$ y de rango n, la matriz A es inversible y se puede transformar en la matriz unidad mediante operaciones elementales (v. 2.5.32). Luego la matriz ampliada
$$[\,A\,|\,b\,]$$
se puede transformar en una matriz de la forma
$$[\,I_n\,|\,b'\,]$$
que proporciona el sistema equivalente
$$I_n\,x = b'$$
cuya solución es inmediata,
$$x = b'\,.$$
Se obtiene así la solución del sistema primitivo.

4.2.5 **Ejemplo.** Para el sistema de Cramer
$$\begin{array}{rcrcr} x & & +\ 2z & = & 0 \\ & y & +\ 3z & = & -2 \\ 2x & & +\ 2z & = & 2 \end{array}$$

[2]Carl Friedrich Gauss (1777-1855), matemático alemán nacido en Brunswich, es para muchos el matemático más importante de todos los tiempos.

del ejemplo (4.2.3), la matriz ampliada es

$$\left[\begin{array}{ccc|c} 1. & 0 & 2. & 0 \\ 0 & 1. & 3. & -2. \\ 2. & 0 & 2. & 2. \end{array}\right] .$$

Mediante operaciones elementales se transforma en

$$\left[\begin{array}{ccc|c} 1. & 0 & 0 & 2. \\ 0 & 1. & 0 & 1. \\ 0 & 0 & 1. & -1. \end{array}\right] ,$$

luego la solución es

$$x = 2. , \quad y = 1. , \quad z = -1.$$

A veces se sigue el procedimiento sólo hasta el punto de obtener una matriz triangular. Si en nuestro ejemplo nos hubiésemos detenido en el punto en que se obtiene

$$\left[\begin{array}{ccc|c} 1. & 0 & 2. & 0 \\ 0 & 1. & 3. & -2. \\ 0 & 0 & -2. & 2. \end{array}\right] ,$$

el sistema equivalente que resulta

$$\begin{array}{rcrcr} x & & + \, 2z & = & 0 \\ & y & + \, 3z & = & -2 \\ & & - \, 2z & = & 2 \end{array} ,$$

proporciona la solución despejando sucesivamente las incógnitas.

4.2.6 Caso $m = r < n$ (el rango coincide con el número de ecuaciones, pero no con el de incógnitas). El sistema es compatible, como vimos en (4.1.11). Puesto que

$$\dim \mathrm{Ker}\, f = n - r \neq 0 ,$$

resulta que $\mathrm{Ker}\, f \neq \{0\}$ y que por lo tanto $\mathrm{Ker}\, f$ posee más de un vector (infinitos vectores si $\mathbb{K}$ es infinito, lo que sucede por ejemplo si $\mathbb{K} = \mathbb{R}$ o $\mathbb{K} = \mathbb{C}$). Se deduce entonces de (4.1.17) que el sistema admite más de una solución.

En este caso hay $r = m$ incógnitas principales, que supondremos son las $x^1, \ldots, x^m$, por hacer más sencilla la notación. Demos valores arbitrarios,

$$x^{m+1} = \lambda^{m+1}, \ldots, x^n = \lambda^n ,$$

a las $n - m$ incógnitas restantes y consideremos el sistema

$$\begin{array}{rcl} \alpha_1^1 x^1 + \cdots + \alpha_m^1 x^m & = & b^1 - \alpha_{m+1}^1 \lambda^{m+1} - \cdots - \alpha_n^1 \lambda^n \\ \multicolumn{3}{c}{\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots} \\ \alpha_1^m x^1 + \cdots + \alpha_m^m x^m & = & b^m - \alpha_{m+1}^m \lambda^{m+1} - \cdots - \alpha_n^m \lambda^n \end{array} .$$

Este sistema es ahora de Cramer, puesto que posee m incógnitas, m ecuaciones y rango m; admite, pues, una solución única

$$\begin{aligned} x^1 &= x^1(\lambda^{m+1}, \ldots, \lambda^n) \\ &\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ x^m &= x^m(\lambda^{m+1}, \ldots, \lambda^n) \end{aligned}$$

dependiente de los valores dados a las incógnitas no principales. Resulta sencillo ver que, cuando $\lambda^{m+1}, \ldots, \lambda^n$ recorren todos los valores posibles de $\mathbb{K}$, entonces

$$\begin{aligned} x^1 &= x^1(\lambda^{m+1}, \ldots, \lambda^n) \\ &\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ x^m &= x^m(\lambda^{m+1}, \ldots, \lambda^n) \\ x^{m+1} &= \lambda^{m+1} \\ x^n &= \lambda^n \end{aligned}$$

recorren todas las soluciones del sistema **(sl)**. Se suele decir que la solución general de **(sl)** viene dada en función de los $n - m$ parámetros $\lambda^{m+1}, \ldots, \lambda^n$.

4.2.7 Ejemplo. Para el sistema con coeficientes en $\mathbb{R}$

$$\begin{array}{rrrrl} 2x & + \; y & & + \;\; t & = 2 \\ & & 2z & + \; 3t & = 0 \\ x \; + & & 3z & & = 1 \end{array} , \quad A = \begin{bmatrix} 2. & 1. & 0 & 1. \\ 0 & 0 & 2. & 3. \\ 1. & 0 & 3. & 0 \end{bmatrix} ,$$

se tiene $n = 4$, $m = 3$ y, como

$$\begin{vmatrix} 2. & 1. & 0 \\ 0 & 0 & 2. \\ 1. & 0 & 3. \end{vmatrix} = 2. \neq 0 \, ,$$

entonces $r = 3$. Con la elección que hemos hecho de un menor no nulo de orden 3, las incógnitas principales son x, y y z. Haciendo $t = \lambda$, obtenemos el sistema de Cramer

$$\begin{array}{rrrl} 2x & + \; y & & = 2 - \lambda \\ & & 2z & = -3\lambda \\ x \; + & & 3z & = 1 \end{array} ,$$

cuya única solución (única para cada valor de λ) es

$$x = 1. + 4.5\lambda \, , \quad y = -10.\lambda \, , \quad z = -1.5\lambda \, .$$

La solución general del sistema es entonces

$$x = 1. + 4.5\lambda \, , \quad y = -10.\lambda \, , \quad z = -1.5\lambda \, , \quad t = \lambda \, .$$

Se obtienen soluciones particulares dando a λ valores reales; así, para $\lambda = 2.$, obtenemos la solución

$$x = 10.\,, \quad y = -20.\,, \quad z = -3.\,, \quad t = 2.\,.$$

y, para $\lambda = -1.$,

$$x = -3.5\,, \quad y = 10.\,, \quad z = 1.5\,, \quad t = -1.$$

4.2.8 Ejemplo. El sistema homogéneo asociado al sistema del ejemplo precedente es

$$\begin{array}{rcrcrcrcl} 2x & + & y & & & + & t & = & 0 \\ & & & & 2z & + & 3t & = & 0 \\ x & + & & & 3z & & & = & 0\ . \end{array}$$

Se resuelve como el anterior. La solución general resulta ser

$$x = 4.5\lambda\,, \quad y = -10.\lambda\,, \quad z = -1.5\lambda\,, \quad t = \lambda\,.$$

Como

$$x = 10.\,, \quad y = -20.\,, \quad z = -3.\,, \quad t = 2.$$

es una solución particular del sistema completo, el teorema (4.1.17) nos permite afirmar que

$$x = 10. + 4.5\lambda\,, \quad y = -20. - 10.\lambda\,, \quad z = -3. - 1.5\lambda\,, \quad t = 2. + \lambda$$

es la solución general del sistema completo. Esta solución difiere en apariencia de la que hemos obtenido en el ejemplo precedente; basta, sin embargo, con hacer $\lambda = -2. + \mu$ para obtener la solución en la misma forma que en el ejemplo precedente, con lo que ambas soluciones son de hecho la misma.

4.2.9 Caso $r < m$ (el rango es menor que el número de ecuaciones). En primer lugar, conviene estudiar si el sistema es comptible o no lo es; se aplica para ello el teorema (4.1.14).

Cuando el sistema es compatible, sus soluciones son las del subsistema principal (v. 4.1.18). Dado que el subsistema principal posee el mismo rango que el número de ecuaciones, se resuelve como hemos visto en los dos casos precedentes.

4.2.10 Ejemplo. Para el sistema con coeficientes en $\mathbb{R}$

$$\begin{array}{rcrcrcl} x & + & y & + & 2z & = & 2 \\ x & & & + & 3z & = & 4\,, \\ & - & y & + & z & = & 2 \end{array} \quad A = \begin{bmatrix} 1. & 1. & 2. \\ 1. & 0 & 3. \\ 0 & -1. & 1. \end{bmatrix},$$

se tiene $n = 3$ y $m = 3$. Por otra parte, $\det A = 0$ y

$$\begin{vmatrix} 1. & 1. \\ 1. & 0 \end{vmatrix} = -1. \neq 0\,,$$

luego $r = 2$. Podemos considerar como subsistema principal el formado por las dos primeras ecuaciones.

La matriz ampliada

$$A' = \left[\begin{array}{ccc|c} 1. & 1. & 2. & 2. \\ 1. & 0 & 3. & 4. \\ 0 & -1. & 1. & 2. \end{array}\right]$$

tiene rango 2 ya que

$$\left|\begin{array}{cc|c} 1. & 1. & 2. \\ 1. & 0 & 4. \\ \hline 0 & -1. & 2. \end{array}\right| = 0\,.$$

El sistema es, pues, compatible y sus soluciones son las del subsistema principal

$$\begin{array}{rcrcrcl} x & + & y & + & 2z & = & 2 \\ x & & & + & 3z & = & 4 \end{array} \;,$$

que son, como se ve fácilmente,

$$x = 4. - 3.\lambda\,, \quad y = \lambda - 2.\,, \quad z = \lambda\,.$$

4.2.11 Ejemplo. Para el sistema con coeficientes en $\mathbb{R}$

$$\begin{array}{rcrcrcl} x & + & 2y & - & z & = & 1 \\ 2x & + & 2y & + & z & = & 3 \\ x & & & + & 2z & = & 1 \end{array} \;, \quad A = \begin{bmatrix} 1. & 2. & -1. \\ 2. & 2. & 1. \\ 1. & 0 & 2. \end{bmatrix} \,,$$

se tiene $n = 3$ y $m = 3$. Como $\det A = 0$ y

$$\begin{vmatrix} 1. & 2. \\ 2. & 2. \end{vmatrix} = -2. \neq 0\,,$$

resulta que $r = 2$. Podemos considerar como principal el subsistema formado por las dos primeras ecuaciones.

La matriz ampliada

$$A' = \left[\begin{array}{ccc|c} 1. & 2. & -1. & 1. \\ 2. & 2. & 1. & 3. \\ 1. & 0 & 2. & 1. \end{array}\right]$$

tiene rango 3 ya que

$$\left|\begin{array}{cc|c} 1. & 2. & 1. \\ 2. & 2. & 3. \\ \hline 1. & 0 & 1. \end{array}\right| = 2. \neq 0\,.$$

El sistema es incompatible, es decir, no admite ninguna solución. El sistema homogéneo asociado al anterior

$$\begin{array}{rcrcrcl} x & + & 2y & - & z & = & 0 \\ 2x & + & 2y & + & z & = & 0 \\ x & & & + & 2z & = & 0 \end{array} \;,$$

sí es compatible, como lo es todo sistema homogéneo (v. 4.1.10); su solución general es

$$x = -2.\lambda\,, \quad y = 1.5\lambda\,, \quad z = \lambda\,.$$

Capítulo 5

Diagonalización de endomorfismos y matrices.

5.1 Subespacios invariantes. Vectores y valores propios.

5.1.1 Sea E un espacio vectorial sobre el cuerpo $\mathbb{K}$ y $f : E \to E$ un endomorfismo de E. Un *subespacio invariante* para f es un subespacio F de E tal que

$$f(F) \subset F \, ,$$

o sea, tal que

$$x \in F \quad \Rightarrow \quad f(x) \in F \, .$$

Ejemplos sencillos de tales subespacios son $\{0\}$, E, $\operatorname{Im} f$ y $\operatorname{Ker} f$, que son invariantes independientemente de la naturaleza del endomorfismo f.

Si F es un subespacio invariante para f, podemos considerar la aplicación

$$\begin{aligned} F &\to F \\ x &\to f(x) \end{aligned}$$

que envía cada vector x de F a su imagen por f; esta aplicación es un endomorfismo de F, el 'endomorfismo inducido' por f en F. Es habitual representar también por f el endomorfismo inducido.

5.1.2 Sea E un e.v. y $f \in \mathcal{L}(E)$. Supongamos que

$$E = F_1 \oplus \cdots \oplus F_p \, ,$$

donde $F_1, \ldots, F_p$ son subespacios invariantes para f, con dimensiones $n_1, \ldots, n_p$ diferentes de 0. Si $(a_1, \ldots, a_{n_1})$ es una base de F_1, $(a_{n_1+1}, \ldots, a_{n_1+n_2})$ una base de F_2, etc., y consideramos la base de E

$$(a_1, \ldots, a_{n_1}, a_{n_1+1}, \ldots, a_{n_1+n_2}, \ldots)$$

formada por la reunión de dichas bases (v. 1.5.7b), entonces la matriz de f en esta base es de la forma

$$[f,(a_i)] = \begin{bmatrix} A_1 & & & \\ & A_2 & & \\ & & \ddots & \\ & & & A_p \end{bmatrix}$$

o sea, lo que se suele llamar una matriz diagonal por bloques. La matriz A_1 es la matriz del endomorfismo inducido por f en F_1 representado en la base $(a_1, \ldots, a_{n_1})$, la matriz A_2 es la matriz del endomorfismo inducido por f en F_2 representado en la base $(a_{n_1+1}, \ldots, a_{n_1+n_2})$, etc.

5.1.3 Ejemplo. Consideremos el endomorfismo de $\mathbb{R}^3$ dado por

$$f(x,y,z) = (4x+2y+z, -2x-z, -x-y+z),$$

es decir, por la matriz

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

Consideremos los vectores

$$a_1 = (1,-1,0), \quad a_2 = (1,0,-1) \quad \text{y} \quad a_3 = (-1,1,1).$$

Es fácil comprobar que

$$\begin{aligned} f(a_1) &= (\;\;2,\; -2,\;\;\; 0) = 2a_1 \\ f(a_2) &= (\;\;3,\; -1,\; -2) = a_1 + 2a_2 \\ f(a_3) &= (-1,\;\;\; 1,\;\;\; 1) = a_3\,. \end{aligned}$$

Los subespacios

$$F_1 = \langle a_1, a_2 \rangle \quad \text{y} \quad F_2 = \langle a_3 \rangle$$

son invariantes; (a_1, a_2) y (a_3) son bases de F_1 y F_2 y (a_1, a_2, a_3) es una base de $\mathbb{R}^3$ (luego $\mathbb{R}^3 = F_1 \oplus F_2$). La matriz de f en la base (a_1, a_2, a_3) es

$$A' = \left[\begin{array}{cc|c} 2. & 1. & 0 \\ 0 & 2. & 0 \\ \hline 0 & 0 & 1. \end{array}\right].$$

5.1.4 El objeto de este capítulo es representar endomorfismos mediante matrices sencillas, entendiendo por tales, matrices con gran número de ceros. Las matrices diagonales por bloques son las más interesantes; lo son aún más cuando se trata de matrices pura y simplemente diagonales. Vamos a estudiar el caso en que un endomorfismo se puede representar por una matriz diagonal, dejando para más tarde las matrices diagonales por bloques, que constituyen un caso más general. El ejemplo siguiente nos muestra una situación en la que es posible representar un endomorfismo mediante una matriz diagonal.

5.1.5 Ejemplo. Consideremos el endomorfismo de $\mathbb{R}^3$ dado por

$$\begin{aligned} f(e_1) &= -\ e_1 - 6e_2 + 2e_3 \\ f(e_2) &= 3e_1 + 8e_2 - 2e_3 \\ f(e_3) &= 6e_1 + 16e_2 - 4e_3 \ , \end{aligned}$$

cuya matriz en la base canónica es

$$A = \begin{bmatrix} -1. & 3. & 6. \\ -6. & 8. & 16. \\ 2. & -2. & -4. \end{bmatrix} .$$

Para la base (a_1, a_2, a_3) de $\mathbb{R}^3$ formada por los vectores

$$a_1 = (0, -2, 1)\,, \quad a_2 = (3, -2, 2) \quad \text{y} \quad a_3 = (1, 1, 0)\,,$$

la matriz de f es

$$A' = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1. & 0 \\ 0 & 0 & 2. \end{bmatrix}$$

puesto que

$$\begin{aligned} f(a_1) &= 0 \\ f(a_2) &= a_2 \\ f(a_3) &= 2a_3\,. \end{aligned}$$

En esta base, f se representa por una matriz diagonal. La matriz A' pone en evidencia algunos aspectos de f que no eran fáciles de ver con la matriz A. Así, ahora es evidente que

$$\operatorname{rg} f = 2\,, \quad \operatorname{Im} f = \langle a_2, a_3 \rangle \quad \text{y} \quad \operatorname{Ker} f = \langle a_1 \rangle\,.$$

5.1.6 A la vista del ejemplo precedente, el problema se centra en dos aspectos:

— averiguar si existe alguna base en la que el endomorfismo se represente por una matriz diagonal, y

— encontrar, cuando exista, una de tales bases.

Nótese que los tres vectores de la base del ejemplo precedente verifican que

$$f(a_i) = \lambda_i\, a_i$$

para escalares λ_i que son, en ese caso,

$$\lambda_1 = 0\,, \quad \lambda_2 = 1 \quad \text{y} \quad \lambda_3 = 2\,.$$

Es a partir de esta propiedad como trataremos de resolver los dos problemas planteados.

5.1.7 DEFINICIÓN. Sea E un e.v. sobre $\mathbb{K}$ y $f \in \mathcal{L}(E)$; sea $x \in E$. Decimos que x es un *vector propio* de f cuando

$$\textbf{(vcp)} \quad (\exists \lambda \in \mathbb{K}) \quad f(x) = \lambda\, x\,.$$

5.1.8 Ejemplo. En el ejemplo (5.1.5), los vectores a_1, a_2 y a_3 son vectores propios de f. Hay una infinidad de vectores propios de f; compruébese, por ejemplo, que lo son todos los vectores de la forma $\mu\, a_i$ cualesquiera que sean $\mu \in \mathbb{K}$ y $i = 1, 2, 3$.

5.1.9 El vector 0 es siempre un vector propio, puesto que

$$f(0) = 0 = \lambda 0$$

cualquiera que sea $\lambda \in \mathbb{K}$; todos los escalares hacen pues cierta la igualdad de **(vcp)** para el vector 0.

No ocurre lo mismo para vectores no nulos; si $x \neq 0$ y $f(x) = \lambda x$ y $f(x) = \mu x$, entonces $\lambda x - \mu x = 0$, luego $(\lambda - \mu)x = 0$ y, como $x \neq 0$, resulta que $\lambda = \mu$. Para cada vector propio x no nulo, existe exactamente un escalar λ tal que $f(x) = \lambda x$.

El método para encontrar los vectores propios de f es el estudio de los escalares que sirven para realizar las igualdades del tipo $f(x) = \lambda x$.

5.1.10 DEFINICIÓN. Sea E un e.v. sobre $\mathbb{K}$ y $f \in \mathcal{L}(E)$; sea $\lambda \in \mathbb{K}$. Decimos que λ es un *valor propio* de f cuando

$$\textbf{(vlp)} \quad (\exists x \in E,\, x \neq 0) \quad f(x) = \lambda\, x\,.$$

La condición $x \neq 0$ que exigimos en **(vlp)** es importante puesto que, si no, cualquier elemento de $\mathbb{K}$ sería un valor propio.

5.1.11 Ejemplo. En el ejemplo (5.1.5) los escalares $0, 1$ y 2 son valores propios de f.

5.1.12 Es también útil otra forma de considerar los valores propios. Si $f \in \mathcal{L}(E)$ y $\lambda \in \mathbb{K}$, representamos por $V(\lambda)$ el subconjunto de E dado por

$$V(\lambda) = \{x \in E \mid f(x) = \lambda x\}\,.$$

Ya hemos visto que $0 \in V(\lambda)$. Es sencillo probar que $V(\lambda)$ es un subespacio vectorial de E; es justamente el subespacio

$$\mathrm{Ker}(f - \lambda\, id_E)\,,$$

núcleo del endomorfismo $f - \lambda\, id_E$ (v. 1.1.11). Decimos que $V(\lambda)$ es el *subespacio propio* de f asociado a λ.

Podemos expresar **(vlp)** es las siguientes formas equivalentes: λ es un valor propio de f si y sólo si

(vlp1) $V(\lambda) \neq \{0\}$,

y también si y sólo si

(vlp2) el endomorfismo $f - \lambda\, id$ no es inyectivo .

Si $\lambda \neq \mu$, entonces

$$V(\lambda) \cap V(\mu) = \{0\} ,$$

como se ve utilizando los argumentos de (5.1.9).

5.1.13 Ejemplo. En el ejemplo (5.1.5) se tiene

$$\begin{aligned} V(0) &= \operatorname{Ker} f = \langle a_1 \rangle \\ V(1) &= \operatorname{Ker}(f - id) = \langle a_2 \rangle \\ V(2) &= \operatorname{Ker}(f - 2\, id) = \langle a_3 \rangle , \end{aligned}$$

como se comprueba con facilidad. Se puede comprobar también que, si λ es distinto de $0, 1$ y 2, entonces

$$V(\lambda) = \operatorname{Ker}(f - \lambda\, id) = \{0\} ,$$

lo que indica que no hay más valores propios que estos tres.

5.1.14 PROPOSICIÓN

Sea $f \in \mathcal{L}(E)$ y sean $\lambda_1, \ldots, \lambda_p$ valores propios de f, todos distintos entre sí. Si $x_1 \in V(\lambda_1), \ldots, x_p \in V(\lambda_p)$, y $x_1 \neq 0, \ldots, x_p \neq 0$, entonces $(x_1, \ldots, x_p)$ es un sistema libre.

Lo probaremos por inducción sobre el número p. El resultado es evidente cuando $p = 1$. Supongamos (hipótesis de recurrencia) que el resultado es cierto para $p-1$; sean entonces $x_1, \ldots, x_p$ p vectores que cumplen las hipótesis del enunciado. Si

$$\alpha_1 x_1 + \alpha_2 x_2 + \cdots + \alpha_p x_p = 0 ,$$

tenemos, por una parte, que

$$f(\alpha_1 x_1 + \alpha_2 x_2 + \cdots + \alpha_p x_p) = 0 ,$$

o sea,

$$\alpha_1 f(x_1) + \alpha_2 f(x_2) + \cdots + \alpha_p f(x_p) = 0 ,$$

luego

$$\alpha_1 \lambda_1 x_1 + \alpha_2 \lambda_2 x_2 + \cdots + \alpha_p \lambda_p x_p = 0 .$$

Por otra parte, $\lambda_1 0 = 0$, o sea,

$$\lambda_1(\alpha_1 x_1 + \alpha_2 x_2 + \cdots + \alpha_p x_p) = 0 ,$$

luego
$$\lambda_1\alpha_1x_1 + \lambda_1\alpha_2x_2 + \cdots + \lambda_1\alpha_px_p = 0\,.$$
Restando ambas expresiones obtenemos
$$(\lambda_2 - \lambda_1)\alpha_2x_2 + \cdots + (\lambda_p - \lambda_1)\alpha_px_p = 0\,,$$
y como, por la hipótesis de recurrencia, el sistema $(x_2, \ldots, x_p)$ de $p-1$ vectores es libre, resulta que
$$(\lambda_2 - \lambda_1)\alpha_2 = \cdots = (\lambda_p - \lambda_1)\alpha_p = 0\,,$$
luego que
$$\alpha_2 = \cdots = \alpha_p = 0$$
ya que todos los λ_i son distintos. Por fin, resulta que
$$\alpha_1x_1 = 0$$
y que $\alpha_1 = 0$. El sistema $(x_1, x_2, \ldots, x_p)$ es libre; el resultado es también cierto para p, y esto concluye la demostración.

5.1.15 Si E es un e.v. de dimensión n y $f \in \mathcal{L}(E)$, resulta inmediatamente de la proposición anterior que f posee a lo sumo n valores propios.

5.1.16 PROPOSICIÓN

Sea $f \in \mathcal{L}(E)$ y sean $\lambda_1, \ldots, \lambda_p$ valores propios de f, todos distintos entre sí. Entonces
$$V(\lambda_1) + \cdots + V(\lambda_p) = V(\lambda_1) \oplus \cdots \oplus V(\lambda_p)\,.$$

Si
$$0 = x_1 + \cdots + x_p$$
con $x_1 \in V(\lambda_1), \ldots, x_p \in V(\lambda_p)$, entonces $x_1 = \cdots = x_p = 0$, pues, en caso contrario, llamando $x_{i_1}, \ldots, x_{i_r}$ a aquellos vectores que no fuesen nulos, resultaría que
$$0 = x_{i_1} + \cdots + x_{i_r}\,;$$
esto es imposible, ya que el sistema $(x_{i_1}, \ldots, x_{i_r})$ es libre, como consecuencia de la proposición anterior.

Sea entonces x un vector de $V(\lambda_1) + \cdots + V(\lambda_p)$ y supongamos que
$$x = x_1 + \cdots + x_p \quad \text{y} \quad x = x_1' + \cdots + x_p'\,,$$
con $x_1, x_1' \in V(\lambda_1), \ldots, x_p, x_p' \in V(\lambda_p)$; se tiene
$$0 = (x_1 - x_1') + \cdots + (x_p - x_p')$$
y, como acabamos de ver, esto significa que
$$x_1 = x_1', \ldots, x_p = x_p'\,.$$
El subespacio suma es, pues, suma directa (v. 1.3.22).

5.1.17 TEOREMA

Sea E un e.v. de dimensión $n \neq 0$, $(a_1, \ldots, a_n)$ una base de E, $f \in \mathcal{L}(E)$ y $A = [f, (a_i)]$. Sea $\lambda \in \mathbb{K}$; las proposiciones siguientes son equivalentes:

(i) λ es un valor propio de f ;

(ii) el endomorfismo $f - \lambda\, id$ no es inversible;

(iii) $\det(A - \lambda I_n) = 0$.

Decir que $f - \lambda\, id$ no es inversible equivale a decir que no es inyectivo, puesto que E es de dimensión finita (v. 2.1.24); se tiene así la equivalencia de (i) y (ii) (v. 5.1.12).

La equivalencia de (ii) y (iii) proviene de que $A - \lambda I_n$ es la matriz de $f - \lambda\, id$ en la base $(a_1, \ldots, a_n)$.

5.1.18 DEFINICIÓN. Si A es una matriz cuadrada con elementos en $\mathbb{K}$, los escalares $\lambda \in \mathbb{K}$ tales que

$$\det(A - \lambda I_n) = 0$$

reciben el nombre de *valores propios* de la matriz A.

Lo que demuestra el teorema precedente es que los valores propios de un endomorfismo f de un espacio vectorial de dimensión finita son los valores propios de cualquier matriz asociada a f en una base del espacio.

Dos endomorfismos del e.v. E que se representen (en bases diferentes) por la misma matriz poseen entonces los mismos valores propios.

Otra consecuencia aún más importante es que dos matrices semejantes poseen los mismos valores propios (v. 2.5.29).

5.1.19 Ejemplo. La matriz

$$A = \begin{bmatrix} -1. & 3. & 6. \\ -6. & 8. & 16. \\ 2. & -2. & -4. \end{bmatrix}$$

asociada al endomorfismo del ejemplo (5.1.5) posee como valores propios $0, 1$ y 2. Compruébese que éstos son los únicos valores $\lambda \in \mathbb{R}$ para los que el determinante de

$$A - \lambda I_3 = \begin{bmatrix} -1. - \lambda & 3. & 6. \\ -6. & 8. - \lambda & 16. \\ 2. & -2. & -4. - \lambda \end{bmatrix}$$

es nulo.

5.1.20 Sea $A \in M_{\mathbb{K}}(n)$. Los valores propios de A son los del endomorfismo $f \in \mathcal{L}(\mathbb{K}^n)$ dado por A (v. 2.5.1), es decir, los escalares $\lambda \in \mathbb{K}$ para los que existe un vector $x \in \mathbb{K}^n$, $x \neq 0$, tal que

$$A\,x = \lambda\,x\,.$$

Se suele hablar de *vectores propios* de A para referirse a los vectores $x \in \mathbb{K}^n$ tales que

$$A\,x = \lambda\,x$$

para algún $\lambda \in \mathbb{K}$, o sea, a los vectores propios del endomorfismo dado por A.

Si $\lambda \in \mathbb{K}$, el subespacio

$$V(\lambda) = \{x \in \mathbb{K}^n \mid Ax = \lambda x\}$$

de $\mathbb{K}^n$ se suele denominar *subespacio propio* de A correspondiente a λ.

5.2 Polinomio característico.

5.2.1 Sea $A \in M_{\mathbb{K}}(n)$ una matriz cuadrada. Es sencillo averiguar cuando un elemento $\lambda \in \mathbb{K}$ es un valor propio de A; basta comprobar que

$$\det(A - \lambda I_n) = 0\,.$$

Parece, sin embargo, más dificil encontrar todos los valores propios de A, puesto que no podemos efectuar dicha comprobación para todos los elementos de $\mathbb{K}$, si éste posee una cantidad infinita de elementos (lo que es normal). Lo que se hace entonces es calcular $\det(A - \lambda I_n)$ sin precisar el valor de λ, y tratar de ver a continuación para qué valores de λ se anula.

5.2.2 Ejemplo. Para la matriz de elementos reales

$$A = \begin{bmatrix} 2. & 2. \\ 1. & 1. \end{bmatrix}$$

tenemos

$$A - \lambda I_n = \begin{bmatrix} 2. - \lambda & 2. \\ 1. & 1. - \lambda \end{bmatrix}$$

y

$$\det(A - \lambda I_n) = (2 - \lambda)(1 - \lambda) - 2 = \lambda^2 - 3\lambda\,,$$

expresión que se anula para $\lambda = 0$ y $\lambda = 3$; los valores propios de A son entonces 0 y 3.

5.2.3 Sea $A = [\alpha_i^j] \in M_{\mathbb{K}}(n)$ una matriz cuadrada. Siendo X un símbolo, consideramos la matriz

$$A - X\,I_n = \begin{bmatrix} \alpha_1^1 - X & \alpha_2^1 & \alpha_3^1 & \cdots & \alpha_n^1 \\ \alpha_1^2 & \alpha_2^2 - X & \alpha_3^2 & \cdots & \alpha_n^2 \\ \alpha_1^3 & \alpha_2^3 & \alpha_3^3 - X & \cdots & \alpha_n^3 \\ \vdots & \vdots & \vdots & & \vdots \\ \alpha_1^n & \alpha_2^n & \alpha_3^n & \cdots & \alpha_n^n - X \end{bmatrix}$$

cuyos elementos son polinomios en la indeterminada X con coeficientes en $\mathbb{K}$. Aun cuando los elementos diagonales de $A - X\,I_n$ no pertenecen al cuerpo $\mathbb{K}$, calculamos el determinante de esta matriz operando con los polinomios $\alpha_i^i - X$ como si se tratase de elementos de $\mathbb{K}$. (Una explicación más formalizada sobrepasa el propósito de este libro.)

$\det A - X\,I_n$ es un polinomio en X, con coeficientes en $\mathbb{K}$, y de grado n.

5.2.4 DEFINICIÓN. Sea $A \in M_{\mathbb{K}}(n)$ una matriz cuadrada. Llamamos *polinomio característico* de A y representamos por $p_A(X)$ al polinomio

$$p_A(X) = \det A - X\,I_n\,.$$

Si $\lambda \in \mathbb{K}$, tenemos que

$$p_A(\lambda) = \det A - \lambda\,I_n\,.$$

5.2.5 Ejemplo. Si

$$A = \begin{bmatrix} 2. & 2. \\ 1. & 1. \end{bmatrix}$$

entonces

$$p_A(X) = X^2 - 3X\,,$$

como vimos en (5.2.2).

Si

$$A = \begin{bmatrix} -2. & 4. & 5. \\ -3. & 5. & 5. \\ 0 & 0 & 1. \end{bmatrix}$$

entonces

$$\begin{aligned} p_A(X) &= \begin{vmatrix} -2. - X & 4. & 5. \\ -3. & 5. - X & 5. \\ 0 & 0 & 1. - X \end{vmatrix} \\ &= (-2 - X)(5 - X)(1 - X) + 12(1 - X) \\ &= -X^3 + 4X^2 - 5X + 2\,, \end{aligned}$$

5.2.6 Algunas consideraciones sencillas nos van a permitir calcular ciertos términos de $p_A(X)$. Los términos de grado n y $n-1$ provienen exclusivamente del producto de los elementos diagonales de $A - X\,I_n$, o sea, de

$$(\alpha_1^1 - X)(\alpha_2^2 - X)\cdots(\alpha_n^n - X)\,.$$

Esto se debe a que en el resto de los $n!$ productos de la fórmula **(3)** de (3.2.4) interviene algún α_i^j con $i \neq j$, luego no intervienen $\alpha_i^i - X$ y $\alpha_j^j - X$; el resto de los productos son entonces polinomios de grado menor o igual que $n-2$.

Como

$$\begin{aligned}&(\alpha_1^1 - X)(\alpha_2^2 - X)\cdots(\alpha_n^n - X) = \\ &= (-1)^n X^n + (-1)^{n-1}(\alpha_1^1 + \alpha_2^2 + \cdots + \alpha_n^n)X^{n-1} + \cdots\,,\end{aligned}$$

resulta que el término de grado n de $p_A(X)$ es

$$(-1)^n X^n\,,$$

y el de grado $n-1$ es

$$(-1)^{n-1}(\alpha_1^1 + \alpha_2^2 + \cdots + \alpha_n^n)X^{n-1}\,,$$

o bien

$$(-1)^{n-1}\,\mathrm{tr}\,A\,X^{n-1}\,,$$

donde el escalar

$$\mathrm{tr}\,A = \alpha_1^1 + \alpha_2^2 + \cdots + \alpha_n^n\,,$$

suma de los elementos diagonales de A, es lo que llamamos *traza* de la matriz A.

Por otra parte, el término independiente de $p_A(X)$ vale

$$p_A(0) = \det(A - 0\,I_n) = \det A\,;$$

$p_A(X)$ es entonces de la forma

$$p_A(X) = (-1)^n X^n + (-1)^{n-1}\,\mathrm{tr}\,A\,X^{n-1} + \cdots + \det A\,.$$

5.2.7 PROPOSICIÓN

Si $A, A' \in M_{\mathbb{K}}(n)$ son dos matrices semejantes, entonces

$$p_A(X) = p_{A'}(X)$$

y, en particular,

$$\mathrm{tr}\,A = \mathrm{tr}\,A'\,.$$

Existe una matriz P, $n \times n$ e inversible, tal que

$$A' = P^{-1} A\,P\,,$$

luego

$$A' - X\,I_n = P^{-1}A\,P - X\,I_n = P^{-1}A\,P - X\,I_n(P^{-1}P)$$

y, como $X\,I_n$ conmuta con cualquier matriz de $M_{\mathbb{K}}(n)$ (v. 2.4.28),

$$A' - X\,I_n = P^{-1}A\,P - P^{-1}X\,I_n\,P = P^{-1}(A - X\,I_n)P\,.$$

Entonces

$$p_{A'}(X) = \det(A' - X\,I_n) = \det(P^{-1})p_A(X)\det P = p_A(X)\,.$$

5.2.8 Sea E un e.v. de dimensión $n \neq 0$ y $f \in \mathcal{L}(E)$. Llamamos *polinomio característico* de f y *traza* de f al polinomio característico y la traza de una matriz cualquiera asociada a f; los representamos por $p_f(X)$ y $\operatorname{tr} f$. La proposición precedente nos garantiza que esta definición no depende de la matriz asociada a f que se elija.

Recordando la definición de $\det f$ (v. 3.2.21) se tiene que

$$p_f(X) = (-1)^n X^n + (-1)^{n-1}\operatorname{tr} f\,X^{n-1} + \cdots + \det f\,.$$

5.2.9 TEOREMA

---

a) Sea $A \in M_{\mathbb{K}}(n)$ y $\lambda \in \mathbb{K}$ ($\mathbb{R}$ o $\mathbb{C}$). Entonces λ es un valor propio de A si y sólo si es una raíz de $p_A(X)$.

b) Sea E un e.v. de dimensión $n \neq 0$ sobre $\mathbb{K}$ ($\mathbb{R}$ o $\mathbb{C}$), $f \in \mathcal{L}(E)$ y $\lambda \in \mathbb{K}$. Entonces λ es un valor propio de f si y sólo si es una raíz de $p_f(X)$.

---

a) λ es un valor propio de A si y sólo si

$$p_A(\lambda) = \det(A - \lambda\,I_n) = 0\,,$$

o sea, si y sólo si λ es raíz de $p_A(X)$.

b) es una consecuencia de a), pues los valores propios de f son los de una matriz cualquiera asociada a f.

5.2.10 Se deduce del teorema precedente que una matriz $n \times n$ posee a lo sumo n valores propios, y que un endomorfismo de un e.v. de dimensión n posee a lo sumo n valores propios.

5.2.11 Cuando $\mathbb{K} = \mathbb{R}$, es importante suponer en el teorema precedente que $\lambda \in \mathbb{R}$. En el caso real pueden existir raíces de $p_f(X)$ que no sean elementos de $\mathbb{R}$; dichas raíces no serán, sin embargo, valores propios de f.

5.2.12 Ejemplo. El endomorfismo f de $\mathbb{R}^2$ dado por la matriz

$$A = \begin{bmatrix} 0 & -1. \\ 1. & 0 \end{bmatrix}$$

tiene como polinomio característico

$$p_f(X) = X^2 + 1\,,$$

polinomio cuyas raíces son los números complejos $+i$ y $-i$, que no son valores propios de f.

Sin embargo, el endomorfismo g de $\mathbb{C}^2$ dado por la misma matriz posee $+i$ y $-i$ como valores propios.

Obsérvese que

$$\begin{bmatrix} 0 & -1. \\ 1. & 0 \end{bmatrix} \begin{bmatrix} 1 \\ -i \end{bmatrix} = i \begin{bmatrix} 1 \\ -i \end{bmatrix}, \quad \text{o sea,} \quad g(1,-i) = i\,(1,-i)$$

y que

$$\begin{bmatrix} 0 & -1. \\ 1. & 0 \end{bmatrix} \begin{bmatrix} 1 \\ i \end{bmatrix} = -i \begin{bmatrix} 1 \\ i \end{bmatrix}, \quad \text{o sea,} \quad g(1,i) = -i\,(1,i)\,,$$

luego que $(1,-i)$ y $(1,i)$ son vectores (de $\mathbb{C}^2$) propios de g. Compruébese que no existe ningún vector (de $\mathbb{R}^2$) propio de f, excepción hecha del vector $(0,0)$.

5.2.13 PROPOSICIÓN

Sea E un e.v. sobre $\mathbb{K}$ de dimensión $n \neq 0$, $f \in \mathcal{L}(E)$; sea $\lambda \in \mathbb{K}$ un valor propio de f. Hemos visto que λ es raíz de $p_f(X)$.

Si k es el orden de multiplicidad de λ como raíz, entonces

$$1 \leq \dim V(\lambda) \leq k\,.$$

Como λ es un valor propio de f, entonces $V(\lambda) \neq \{0\}$, luego $\dim V(\lambda) \geq 1$.

Representemos por h la dimensión de $V(\lambda)$; vamos a probar que $h \leq k$. Para ello completamos una base $(a_1, \ldots, a_h)$ de $V(\lambda)$ hasta formar una base $(a_1, \ldots, a_h, a_{h+1}, \ldots, a_n)$ de E. La matriz $A = [f, (a_i)]$ es de la forma

$$A = \begin{array}{c} \begin{array}{cc} \quad h \quad & \quad n-h \quad \end{array} \\ \left[\begin{array}{ccc|c} \lambda & & & \\ & \ddots & & A' \\ & & \lambda & \\ \hline & 0 & & A'' \end{array}\right] \end{array} \begin{array}{l} \\ h \\ \\ n-h \end{array}$$

luego

$$A - X\, I_n = \left[\begin{array}{ccc|c} \lambda - X & & & \\ & \ddots & & A' \\ & & \lambda - X & \\ \hline & 0 & & A'' - X\, I_{n-h} \end{array}\right],$$

y entonces (v. 3.2.14)

$$\begin{aligned} p_f(X) &= p_A(X) \\ &= \det(A - X\, I_n) \\ &= \det \begin{bmatrix} \lambda - X & & \\ & \ddots & \\ & & \lambda - X \end{bmatrix} \det(A'' - X\, I_{n-h}) \\ &= (\lambda - X)^h q(X)\,, \end{aligned}$$

donde $q(X)$ es un polinomio de grado $n - h$ (el polinomio característico de A''). Resulta así que λ es raíz de $p_f(X)$ con multiplicidad $\geq h$, o sea, $k \geq h$.

5.2.14 Ejemplo. Es posible, sin embargo, que

$$\dim V(\lambda) < k\,.$$

Así, si f es el endomorfismo de $\mathbb{R}^2$ dado por la matriz

$$A = \begin{bmatrix} 1. & 1. \\ 0 & 1. \end{bmatrix},$$

1 es un valor propio de f y es una raíz de multiplicidad 2 del polinomio

$$p_f(X) = (1 - X)^2\,.$$

Pero

$$V(1) = \{(x, y) \in \mathbb{R}^2 \mid y = 0\} = \langle (1, 0) \rangle\,,$$

luego $\dim V(1) = 1$.

5.3 Diagonalización: condiciones.

5.3.1 Estamos ahora en condiciones de responder a las dos cuestiones planteadas en (5.1.6). Comencemos por proporcionarnos una forma breve de expresar el problema.

5.3.2 DEFINICIÓN. Sea E un e.v. de dimensión $n \neq 0$ sobre $\mathbb{K}$ y $f \in \mathcal{L}(E)$. Decimos que el endomorfismo f es *diagonalizable* cuando existe una base $(a_1, \ldots, a_n)$ de E tal que la matriz $[f, (a_i)]$ que representa a f en dicha base es una matriz diagonal.

Sea $A \in M_{\mathbb{K}}(n)$ una matriz cuadrada (recordemos que $\mathbb{K} = \mathbb{R}$ o $\mathbb{C}$). Decimos que A es *diagonalizable* en $\mathbb{K}$ cuando A es semejante a una matriz diagonal de $M_{\mathbb{K}}(n)$, esto es, cuando existe $P \in M_{\mathbb{K}}(n)$ inversible y tal que $P^{-1}A\,P$ es una matriz diagonal con elementos de $\mathbb{K}$.

5.3.3 Las dos nociones precedentes están profundamente relacionadas. Si f es un endomorfismo diagonalizable, $(a_1, \ldots, a_n)$ una base de E y $A = [f, (a_i)]$, entonces la matriz A es diagonalizable. En efecto, sabemos que existe una base $(b_1, \ldots, b_n)$ de E tal que la matriz

$$B = [f, (b_i)]$$

es diagonal; como

$$B = P^{-1}A\,P\,,$$

donde P es la matriz de paso de la base (a_i) a la base (b_i), resulta que A es diagonalizable.

5.3.4 Recíprocamente, si $A \in M_{\mathbb{K}}(n)$ es una matriz diagonalizable, cualquier endomorfismo que se represente por A es diagonalizable (en particular lo es el endomorfismo de $\mathbb{K}^n$ dado por A). En efecto; sea E un e.v. de dimensión n sobre $\mathbb{K}$, $f \in \mathcal{L}(E)$ y $(a_1, \ldots, a_n)$ una base de E, de tal manera que

$$[f, (a_i)] = A\,.$$

Puesto que A es diagonalizable, existe $P \in M_{\mathbb{K}}(n)$ inversible y tal que $P^{-1}A\,P$ es diagonal. Llamemos $(b_1, \ldots, b_n)$ a la base de E tal que $P = [(b_i), (a_i)]$; entonces

$$[f, (b_i)] = P^{-1}A\,P\,,$$

lo que demuestra que f es diagonalizable.

5.3.5 Un endomorfismo f de E es diagonalizable si y sólo si existe una base $(a_1, \ldots, a_n)$ de E formada por vectores propios de f.

En efecto; si $(a_1, \ldots, a_n)$ es una base de E formada por vectores propios, existen escalares $\lambda_1, \ldots, \lambda_n$ tales que

$$\begin{aligned} f(a_1) &= \lambda_1 a_1 \\ &\cdots\cdots\cdots\cdots\cdots \\ f(a_n) &= \lambda_n a_n\ ; \end{aligned}$$

pero entonces

$$[f, (a_i)] = \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix},$$

luego f es diagonalizable.

Recíprocamente, si f es diagonalizable, existe una base $(a_1, \ldots, a_n)$ de E tal que

$$[f, (a_i)] = \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix}$$

y, como esto significa que

$$\begin{aligned} f(a_1) &= \lambda_1 a_1 \\ \ldots&\ldots\ldots\ldots \\ f(a_n) &= \lambda_n a_n \ , \end{aligned}$$

resulta que los elementos de la base son vectores propios.

5.3.6 TEOREMA

Sea E un e.v. sobre $\mathbb{K}$ ($\mathbb{R}$ o $\mathbb{C}$), de dimensión $n \neq 0$; sea $f \in \mathcal{L}(E)$. El endomorfismo f es diagonalizable si y sólo si se verifican las dos propiedades siguientes

(d1) $p_f(X)$ posee n raíces en $\mathbb{K}$, iguales o distintas (más exactamente, posee raíces en $\mathbb{K}$ cuyos órdenes de multiplicidad suman n),

propiedad que siempre se verifica cuando $\mathbb{K} = \mathbb{C}$, y

(d2) para cada raíz $\lambda \in \mathbb{K}$ de $p_f(X)$,

$$\dim V(\lambda) = k \, ,$$

donde k es el orden de multiplicidad de λ.

El mismo resultado es cierto para una matriz $A \in M_{\mathbb{K}}(n)$, substituyendo $p_f(X)$ por $p_A(X)$.

Supongamos en primer lugar que f es diagonalizable; existe entonces una base $(a_1, \ldots, a_n)$ de E tal que $[f, (a_i)]$ es diagonal, o sea,

$$[f, (a_i)] = \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix} .$$

Resulta así que

$$p_f(X) = (\lambda_1 - X)(\lambda_2 - X) \cdots (\lambda_n - X) \, ,$$

luego $p_f(X)$ posee n raíces, $\lambda_1, \lambda_2, \ldots, \lambda_n$, en $\mathbb{K}$ (no necesariamente diferentes), lo que prueba **(d1)**. Si λ es una de estas raíces y k su orden de multiplicidad, entonces

$$\lambda_{i_1} = \lambda_{i_2} = \cdots = \lambda_{i_k} = \lambda$$

para k índices $i_1, \ldots, i_k \in \{1, \ldots, n\}$. Los correspondientes vectores de la base verifican, pues,

$$\begin{aligned} f(a_{i_1}) &= \lambda_{i_1} a_{i_1} = \lambda a_{i_1} \\ &\ldots\ldots\ldots\ldots\ldots\ldots \\ f(a_{i_k}) &= \lambda_{i_k} a_{i_k} = \lambda a_{i_k} \ , \end{aligned}$$

es decir,

$$a_{i_1}, \ldots, a_{i_k} \in V(\lambda) \, ,$$

luego

$$\dim V(\lambda) \geq k \, .$$

Como por otra parte $\dim V(\lambda) \leq k$ (v. 5.2.13), tenemos la igualdad

$$\dim V(\lambda) = k \, ,$$

lo que prueba **(d2)**.

Recíprocamente, supongamos que se verifican **(d1)** y **(d2)**. Por la primera sabemos que existen elementos $\lambda_1, \ldots, \lambda_p$ que son raíces de $p_f(X)$ y que sus órdenes de multiplicidad, $k_1, \ldots, k_p$, suman

$$k_1 + \cdots + k_p = n \, .$$

Los λ_i son valores propios de f y cada subespacio $V(\lambda_i)$ es de dimensión k_i, puesto que se cumple **(d2)**. La suma directa (v. 5.1.16)

$$V(\lambda_1) \oplus \cdots \oplus V(\lambda_p)$$

es un subespacio de dimensión $k_1 + \cdots + k_p = n$ (v. 1.5.7), o sea,

$$V(\lambda_1) \oplus \cdots \oplus V(\lambda_p) = E \, .$$

Reuniendo bases de $V(\lambda_1), \ldots, V(\lambda_p)$, obtenemos entonces una base de E, base que estará formada por vectores propios. El endomorfismo f es, pues, diagonalizable.

El resultado para matrices es evidente, teniendo en cuenta que A es diagonalizable si y sólo si lo es el endomorfismo de $\mathbb{K}^n$ dado por A (v. 5.3.3 y 5.3.4).

5.3.7 La demostración del teorema precedente nos indica además un procedimiento para obtener una base $(a_1, \ldots, a_n)$ en la que un endomorfismo diagonalizable f tenga asociada una matriz diagonal. Tal base se puede formar reuniendo bases de los subespacios $V(\lambda_i)$, para los valores propios λ_i de f.

5.3.8 COROLARIO

Sea E un e.v. sobre $\mathbb{K}$ ($\mathbb{R}$ o $\mathbb{C}$), de dimensión $n \neq 0$; sea $f \in \mathcal{L}(E)$. Si $p_f(X)$ tiene n raíces distintas en $\mathbb{K}$, entonces f es diagonalizable.

El mismo resultado es cierto para una matriz $A \in M_{\mathbb{K}}(n)$, substituyendo $p_f(X)$ por $p_A(X)$.

Si $p_f(X)$ posee n raíces distintas, $\lambda_1, \ldots, \lambda_n$, en $\mathbb{K}$, se verifica la propiedad **(d1)** del teorema precedente. Además cada λ_i es de multiplicidad 1, luego (v. 5.2.13)

$$1 \leq \dim V(\lambda_i) \leq 1\,,$$

o sea, $\dim V(\lambda_i) = 1$; se verifica también **(d2)**.

5.3.9 Ejemplo. Sea $f \in \mathcal{L}(\mathbb{R}^3)$ el endomorfismo dado por la matriz

$$A = \begin{bmatrix} -2. & 4. & 5. \\ -3. & 5. & 5. \\ 0 & 0 & 1. \end{bmatrix}.$$

Tenemos que

$$\begin{aligned} p_f(X) &= \begin{vmatrix} -2.-X & 4. & 5. \\ -3. & 5.-X & 5. \\ 0 & 0 & 1.-X \end{vmatrix} \\ &= (1-X)\,(\,(5-X)(-2-X)+12) \\ &= (1-X)(X^2-3X+2) \end{aligned}$$

y las raíces de este polinomio son

$$1 \text{ con multiplicidad } 2 \quad \text{y} \quad 2 \text{ con multiplicidad } 1\,;$$

se verifica entonces **(d1)**. Vamos a estudiar los subespacios propios.

$V(1) = \mathrm{Ker}(f - id)$, luego $(x, y, z) \in V(1)$ cuando

$$(A - I_3) \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix},$$

o sea,

$$\begin{bmatrix} -3. & 4. & 5. \\ -3. & 4. & 5. \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix};$$

la solución de este sistema es

$$(x, y, z) = (4\lambda + 5\mu, 3\lambda, 3\mu)$$

y por consiguiente

$$V(1) = \langle (4,3,0), (5,0,3) \rangle .$$

$V(1)$ es de dimensión 2. Se cumple también **(d2)**, pues ya sabemos que la dimensión de $V(2)$ no puede ser sino 1. El endomorfismo f es diagonalizable.

Para encontrar una base en la que corresponda a f una matriz diagonal, tenemos que calcular también $V(2)$.

$V(2) = \operatorname{Ker}(f - 2\, id)$, luego $(x, y, z) \in V(2)$ cuando

$$(A - 2\, I_3) \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} ,$$

o sea,

$$\begin{bmatrix} -4. & 4. & 5. \\ -3. & 3. & 5. \\ 0 & 0 & -1. \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} ;$$

la solución de este sistema es

$$(x, y, z) = (\lambda, \lambda, 0) ,$$

luego

$$V(2) = \langle (1,1,0) \rangle .$$

En la base de $\mathbb{R}^3$

$$(\,(4,3,0), (5,0,3), (1,1,0)\,)$$

f se representa por la matriz diagonal

$$\begin{bmatrix} 1. & & \\ & 1. & \\ & & 2. \end{bmatrix} .$$

La matriz A es diagonalizable en $\mathbb{R}$ y

$$P^{-1} A\, P = \begin{bmatrix} 1. & & \\ & 1. & \\ & & 2. \end{bmatrix} ,$$

siendo P la matriz

$$P = \begin{bmatrix} 4. & 5. & 1. \\ 3. & 0 & 1. \\ 0 & 3. & 0 \end{bmatrix} .$$

Compruébese este hecho, teniendo en cuenta que

$$P^{-1} = \frac{1}{3.} \begin{bmatrix} 3. & -3. & -5. \\ 0 & 0 & 1. \\ -9. & 12. & 15. \end{bmatrix} .$$

5.3.10 Ejemplo. Sea $f \in \mathcal{L}(\mathrm{I\!R}^3)$ el endomorfismo dado por la matriz

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

Su polinomio característico vale

$$\begin{aligned} p_f(X) &= \begin{vmatrix} 2.-X & -2. & 1. \\ 1. & 3.-X & 1. \\ 0 & 1. & 2.-X \end{vmatrix} \\ &= (2-X)^2(3-X) + 1 - (2-X) + 2(2-X) \\ &= (2-X)^2(3-X) + (3-X) \\ &= (3-X)(\,(2-X)^2+1) \\ &= (3-X)(X^2-4X+5)\,; \end{aligned}$$

las raíces son 3, $2+i$ y $2-i$. $p_f(X)$ posee únicamente una raíz en $\mathrm{I\!R}$, luego **(d1)** no se cumple y f no es diagonalizable.

5.3.11 Ejemplo. Sea $f \in \mathcal{L}(\mathrm{I\!R}^3)$ el endomorfismo dado por la matriz

$$A = \begin{bmatrix} 3. & 2. & 4. \\ 0 & 1. & 0 \\ -2. & 0 & -3. \end{bmatrix}.$$

Tenemos que

$$\begin{aligned} p_f(X) &= \begin{vmatrix} 3.-X & 2. & 4. \\ 0 & 1.-X & 0 \\ -2. & 0 & -3.-X \end{vmatrix} \\ &= (1-X)(\,(3-X)(-3-X)+8) \\ &= (1-X)(X^2-1) \end{aligned}$$

con raíces

$$1 \text{ con multiplicidad } 2 \quad \text{y} \quad -1 \text{ con multiplicidad } 1\,;$$

se verifica entonces **(d1)**.

$V(1) = \mathrm{Ker}(f - id)$, luego $(x,y,z) \in V(1)$ cuando

$$\begin{bmatrix} 2. & 2. & 4. \\ 0 & 0 & 0 \\ -2. & 0 & -4. \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix};$$

la solución de este sistema es

$$(x,y,z) = (-2\lambda, 0, \lambda)$$

y por consiguiente

$$V(1) = \langle (-2,0,1) \rangle\,;$$

$V(1)$ es de dimensión 1 y no se verifica **(d2)**. f no es diagonalizable.

5.3.12 Ejemplo. Sea $f \in \mathcal{L}(\mathbb{R}^3)$ el endomorfismo del ejemplo (5.1.5), dado por la matriz

$$A = \begin{bmatrix} -1. & 3. & 6. \\ -6. & 8. & 16. \\ 2. & -2. & -4. \end{bmatrix}.$$

Vimos en dicho ejemplo que f es diagonalizable, puesto que encontramos una base en la que la matriz de f es diagonal.

El polinomio característico de f es

$$\begin{aligned} p_f(X) &= \begin{vmatrix} -1.-X & 3. & 6. \\ -6. & 8.-X & 16. \\ 2. & -2. & -4.-X \end{vmatrix} \\ &= (-1-X)(8-X)(-4-X)+96+72 \\ &\qquad -12(8-X)+18(-4-X)+32(-1-X) \\ &= -X^3+3X^2-2X \\ &= -X(X^2-3X+2) \end{aligned}$$

de raíces 0, 1 y 2. Utilizando el corolario (5.3.8) se puede llegar también a la conclusión de que f es diagonalizable.

5.3.13 Ejemplo. Los endomorfismos f y g del ejemplo (5.2.12) se comportan de manera diferente a pesar de venir dados por la misma matriz

$$A = \begin{bmatrix} 0 & -1. \\ 1. & 0 \end{bmatrix}.$$

El endomorfismo $g \in \mathcal{L}(\mathbb{C}^2)$ es diagonalizable; en la base $(\,(1,-i),(1,i)\,)$ de $\mathbb{C}^2$, la matriz de g es

$$\begin{bmatrix} i & \\ & -i \end{bmatrix}.$$

Sin embargo, el endomorfismo $f \in \mathcal{L}(\mathbb{R}^2)$ no es diagonalizable.

La matriz A es diagonalizable en $\mathbb{C}$, puesto que

$$\begin{bmatrix} 1. & 1. \\ -i & i \end{bmatrix}^{-1} \begin{bmatrix} 0 & -1. \\ 1. & 0 \end{bmatrix} \begin{bmatrix} 1. & 1. \\ -i & i \end{bmatrix} = \begin{bmatrix} i & \\ & -i \end{bmatrix}$$

(compruébese este hecho). Sin embargo, A no es diagonalizable en $\mathbb{R}$, puesto que no existe ninguna matriz $P \in M_{\mathbb{R}}(2)$ (¡con elementos reales!) inversible y tal que $P^{-1}A\,P$ sea diagonal con elementos reales.

5.4 Forma triangular de endomorfismos y matrices.

5.4.1 Cuando no es posible diagonalizar un endomorfismo se recurre a representarlo por una matriz sencilla de otro tipo. La forma alternativa más importante es la llamada forma canónica o forma de Jordan; las matrices de este tipo son triangulares y casi-diagonales. Trataremos este tema en la sección 5.6. En algunas ocasiones basta con reducir el endomorfismo o la matriz a la forma triangular; esta será la posibilidad que estudiaremos ahora.

5.4.2 Sea E un e.v. sobre $\mathbb{K}$ ($\mathbb{R}$ o $\mathbb{C}$) de dimensión $n \neq 0$ y $f \in \mathcal{L}(E)$; sea $(a_1, \ldots, a_n)$ una base de E y $[f, (a_i)] = [\alpha_i^j]$ la matriz de f en dicha base. Si la matriz es triangular, los elementos diagonales, $\alpha_1^1, \ldots, \alpha_n^n$, son los valores propios de f. En efecto

$$p_f(X) = (\alpha_1^1 - X)(\alpha_2^2 - X) \cdots (\alpha_n^n - X)\,,$$

y basta aplicar (5.2.9).

Sea $A \in M_{\mathbb{K}}(n)$ y supongamos que existe una matriz inversible $P \in M_{\mathbb{K}}(n)$ tal que la matriz $A' = P^{-1} A\, P$ es triangular con elementos en $\mathbb{K}$. Entonces los elementos diagonales de A' son los valores propios de A, como se prueba con un razonamiento idéntico al anterior.

5.4.3 PROPOSICIÓN

---

a) Sea E un e.v. sobre $\mathbb{K}$ ($\mathbb{R}$ o $\mathbb{C}$) de dimensión $n \neq 0$ y $f \in \mathcal{L}(E)$. Para que exista una base $(a_1, \ldots, a_n)$ de E tal que la matriz $[f, (a_i)]$ es triangular superior, es condición necesaria y suficiente que se verifique

(d1) $p_f(X)$ posee n raíces en $\mathbb{K}$, iguales o distintas,

propiedad que siempre se verifica cuando $\mathbb{K} = \mathbb{C}$.

b) Sea $A \in M_{\mathbb{K}}(n)$, $\mathbb{K} = \mathbb{R}$ o $\mathbb{C}$. Para que A sea semejante en $\mathbb{K}$ a una matriz triangular superior, es condición necesaria y suficiente que se verifique

(d1) $p_A(X)$ posee n raíces en $\mathbb{K}$, iguales o distintas,

propiedad que siempre se verifica cuando $\mathbb{K} = \mathbb{C}$. Toda matriz cuadrada compleja es semejante a una matriz triangular superior de números complejos.

c) Los resultados precedentes son también ciertos cuando substituimos 'triangular superior' por 'triangular inferior'.

Veamos primero el resultado para matrices.

El razonamiento del apartado (5.4.2) prueba que la condición **(d1)** es necesaria.

Recíprocamente, supongamos que A verifica **(d1)**. Vamos a probar que entonces A es semejante a una matriz triangular superior. Lo probaremos por inducción sobre la dimensión n de la matriz. Para $n = 1$ el resultado es evidente, puesto que toda matriz 1×1 es triangular superior. Supongamos cierto el resultado para n (hipótesis de recurrencia) y sea $A \in M_{\mathbb{K}}(n+1)$ una matriz tal que $p_A(X)$ posee

$n+1$ raíces en $\mathbb{K}$. Llamemos λ a una de estas raíces; λ es un valor propio de A y del endomorfismo $f \in \mathcal{L}(\mathbb{K}^{n+1})$ dado por A, luego existe $a_1 \in \mathbb{K}^{n+1}$, $a_1 \neq 0$, tal que

$$f(a_1) = \lambda a_1 \,.$$

Existen vectores $a_2, \ldots, a_{n+1} \in \mathbb{K}^{n+1}$ tales que $(a_1, a_2, \ldots, a_{n+1})$ es una base de $\mathbb{K}^{n+1}$. La matriz $A' = [f, (a_i)]$ es

$$A' = Q^{-1} A\, Q \,,$$

donde $Q = [(a_i), (e_i)]$, y es de la forma

$$A' = \left[\begin{array}{c|ccc} \lambda & \alpha'_1 & \cdots & \alpha'_n \\ \hline 0 & & & \\ \vdots & & A_1 & \\ 0 & & & \end{array}\right]$$

con $A_1 \in M_{\mathbb{K}}(n)$. Calculando su polinomio característico, que coincide con el de A, resulta

$$p_A(X) = p_{A'}(X) = (\lambda - X)\det(A_1 - X\, I_n) = (\lambda - X) p_{A_1}(X) \,.$$

El polinomio $p_{A_1}(X)$ posee entonces n raíces en $\mathbb{K}$. Utilizando la hipótesis de recurrencia, existe una matriz inversible $Q_1 \in M_{\mathbb{K}}(n)$ tal que la matriz

$$A'_1 = Q_1^{-1} A_1\, Q_1$$

es triangular superior. Pongamos ahora

$$P = Q \left[\begin{array}{c|c} 1 & 0 \\ \hline 0 & Q_1 \end{array}\right] \,;$$

la matriz P es inversible y su inversa es

$$P^{-1} = \left[\begin{array}{c|c} 1 & 0 \\ \hline 0 & Q_1^{-1} \end{array}\right] Q^{-1} \,.$$

Además,

$$\begin{aligned}
P^{-1} A\, P &= \left[\begin{array}{c|c} 1 & 0 \\ \hline 0 & Q_1^{-1} \end{array}\right] Q^{-1} A\, Q \left[\begin{array}{c|c} 1 & 0 \\ \hline 0 & Q_1 \end{array}\right] \\
&= \left[\begin{array}{c|c} 1 & 0 \\ \hline 0 & Q_1^{-1} \end{array}\right] A' \left[\begin{array}{c|c} 1 & 0 \\ \hline 0 & Q_1 \end{array}\right] \\
&= \left[\begin{array}{c|c} 1 & 0 \\ \hline 0 & Q_1^{-1} \end{array}\right] \left[\begin{array}{c|ccc} \lambda & \alpha'_1 & \cdots & \alpha'_n \\ \hline 0 & & A_1 & \end{array}\right] \left[\begin{array}{c|c} 1 & 0 \\ \hline 0 & Q_1 \end{array}\right] \\
&= \left[\begin{array}{c|c} 1 & 0 \\ \hline 0 & Q_1^{-1} \end{array}\right] \left[\begin{array}{c|ccc} \lambda & \beta_1 & \cdots & \beta_n \\ \hline 0 & & A_1\, Q_1 & \end{array}\right] \\
&= \left[\begin{array}{c|ccc} \lambda & \beta_1 & \cdots & \beta_n \\ \hline 0 & & Q_1^{-1} A_1\, Q_1 & \end{array}\right] \\
&= \left[\begin{array}{c|ccc} \lambda & \beta_1 & \cdots & \beta_n \\ \hline 0 & & A'_1 & \end{array}\right]
\end{aligned}$$

y esta matriz es triangular superior, puesto que A'_1 lo es. Esto termina la recurrencia y la demostración de b).

Utilizando b) y el mismo argumento que en (5.3.4) se prueba sin dificultad el apartado a).

Finalmente, si $[f,(a_1,a_2,\ldots,a_n)]$ es triangular superior (inferior), entonces $[f,(a_n,\ldots,a_2,a_1)]$ es triangular inferior (superior). Se obtiene así el resultado c) para endomorfismos. El correspondiente resultado para matrices se sigue de éste con el mismo argumento que en (5.3.3).

5.4.4 Ejemplo. La matriz

$$A = \begin{bmatrix} 3. & 2. & 4. \\ 0 & 1. & 0 \\ -2. & 0 & -3. \end{bmatrix}$$

de $M_{\mathbb{R}}(3)$ del ejemplo (5.3.11) no es diagonalizable (ni en $\mathbb{R}$ ni en $\mathbb{C}$), pero verifica **(d1)**; es entonces semejante en $\mathbb{R}$ a una matriz triangular superior. Sus valores propios son, como vimos, 1 y -1 (el primero con multiplicidad 2); vimos también que

$$(-2,0,1) \in V(1)\,.$$

$(\,(-2,0,1),(0,1,0),(0,0,1)\,)$ es una base de $\mathbb{R}^3$; la matriz

$$Q = \begin{bmatrix} -2. & 0 & 0 \\ 0 & 1. & 0 \\ 1. & 0 & 1. \end{bmatrix}$$

es inversible, con

$$Q^{-1} = \frac{1}{-2.}\begin{bmatrix} 1. & 0 & 0 \\ 0 & -2. & 0 \\ -1. & 0 & -2. \end{bmatrix},$$

y se tiene

$$A' = Q^{-1}A\,Q = \left[\begin{array}{c|cc} 1. & -1. & -2. \\ \hline 0 & 1. & 0 \\ 0 & 1. & -1. \end{array}\right].$$

Ponemos

$$A_1 = \begin{bmatrix} 1. & 0 \\ 1. & -1. \end{bmatrix};$$

como

$$p_A(X) = p_{A'}(X) = (1-X)p_{A_1}(X)\,,$$

ya sabemos que los valores propios de A_1 son 1 y -1 (ahora ambos con multiplicidad 1). El subespacio $V(1)$ de $\mathbb{R}^2$, correspondiente a la matriz A_1, es

$$V(1) = \langle (2,1) \rangle\,.$$

$((2,1),(0,1))$ es una base de $\mathbb{R}^2$; la matriz

$$Q_1 = \begin{bmatrix} 2. & 0 \\ 1. & 1. \end{bmatrix}$$

es inversible, con

$$Q_1^{-1} = \frac{1}{2.}\begin{bmatrix} 1. & 0 \\ -1. & 2. \end{bmatrix},$$

y se tiene que

$$A_1' = Q_1^{-1} A_1 \, Q_1 = \begin{bmatrix} 1. & 0 \\ 0 & -1. \end{bmatrix}.$$

Calculamos entonces

$$\begin{aligned} P &= Q \left[\begin{array}{c|cc} 1. & 0 & 0 \\ \hline 0 & \multicolumn{2}{c}{\multirow{2}{*}{Q_1}} \\ 0 & & \end{array}\right] \\ &= \begin{bmatrix} -2. & 0 & 0 \\ 0 & 1. & 0 \\ 1. & 0 & 1. \end{bmatrix} \begin{bmatrix} 1. & 0 & 0 \\ 0 & 2. & 0 \\ 0 & 1. & 1. \end{bmatrix} \\ &= \begin{bmatrix} -2. & 0 & 0 \\ 0 & 2. & 0 \\ 1. & 1. & 1. \end{bmatrix}. \end{aligned}$$

La matriz P es inversible y

$$P^{-1} A \, P = \begin{bmatrix} 1. & -4. & -2. \\ 0 & 1. & 0 \\ 0 & 0 & -1. \end{bmatrix},$$

como se comprueba fácilmente.

En el tratamiento de este ejemplo hemos seguido paso a paso el desarrollo realizado en la demostración de la proposición anterior. De hecho se puede proceder de forma más rápida si en cada paso se reduce lo más posible el tamaño de la siguiente matriz a triangularizar. Esto es lo que haremos en el ejemplo que sigue.

5.4.5 Ejemplo. Consideremos la matriz

$$A = \begin{bmatrix} 2. & 1. & 1. & 2. \\ 2. & 1. & -1. & 0 \\ 1. & 0 & 0 & -1. \\ 1. & -2. & 0 & 1. \end{bmatrix}$$

de $M_{\mathbb{R}}(4)$. Su polinomio característico

$$p_A(X) = (X-2)^3(X+2)$$

posee raíces 2 (con multiplicidad 3) y -2.

Es fácil comprobar que $\dim V(2) = 1$ y que A no es diagonalizable (ni en $\mathbb{R}$ ni en $\mathbb{C}$); sin embargo, A es triangularizable en $\mathbb{R}$.

Tenemos que

$$(1,1,1,-1) \in V(2) \quad \text{y} \quad (1,-1,-1,-1) \in V(-2)\,.$$

$(\,(1,1,1,-1),(1,-1,-1,-1),(0,0,1,0),(0,0,0,1)\,)$ es una base de $\mathbb{R}^4$; la matriz

$$Q = \begin{bmatrix} 1. & 1. & 0 & 0 \\ 1. & -1. & 0 & 0 \\ 1. & -1. & 1. & 0 \\ -1. & -1. & 0 & 1. \end{bmatrix}$$

es inversible, con

$$Q^{-1} = \frac{1}{2.}\begin{bmatrix} 1. & 1. & 0 & 0 \\ 1. & -1. & 0 & 0 \\ 0 & -2. & 2. & 0 \\ 2. & 0 & 0 & 2. \end{bmatrix},$$

y se tiene

$$A' = Q^{-1} A\, Q = \begin{bmatrix} 2. & 0 & 0 & 1. \\ 0 & -2. & 1. & 1. \\ 0 & 0 & 1. & -1. \\ 0 & 0 & 1. & 3. \end{bmatrix}.$$

Ponemos

$$A_1 = \begin{bmatrix} 1. & -1. \\ 1. & 3. \end{bmatrix};$$

como

$$p_A(X) = p_{A'}(X) = (2-X)(-2-X)p_{A_1}(X)\,,$$

el único valor propio de A_1 es 2, ahora con multiplicidad 2. Para el subespacio $V(2)$ de $\mathbb{R}^2$, correspondiente a la matriz A_1, se tiene que

$$(1,-1) \in V(2)\,.$$

$(\,(1,-1),(0,1)\,)$ es una base de $\mathbb{R}^2$; la matriz

$$Q_1 = \begin{bmatrix} 1. & 0 \\ -1. & 1. \end{bmatrix}$$

es inversible, con

$$Q_1^{-1} = \begin{bmatrix} 1. & 0 \\ 1. & 1. \end{bmatrix}$$

y se tiene

$$A_1' = Q_1^{-1} A_1\, Q_1 = \begin{bmatrix} 2. & -1. \\ 0 & 2. \end{bmatrix}.$$

Ponemos entonces

$$
\begin{aligned}
P &= Q \left[\begin{array}{cc|c} 1. & & \\ & 1. & 0 \\ \hline \multicolumn{2}{c|}{0} & Q_1 \end{array}\right] \\
&= \begin{bmatrix} 1. & 1. & 0 & 0 \\ 1. & -1. & 0 & 0 \\ 1. & -1. & 1. & 0 \\ -1. & -1. & 0 & 1. \end{bmatrix} \begin{bmatrix} 1. & 0 & 0 & 0 \\ 0 & 1. & 0 & 0 \\ 0 & 0 & 1. & 0 \\ 0 & 0 & -1. & 1. \end{bmatrix} \\
&= \begin{bmatrix} 1. & 1. & 0 & 0 \\ 1. & -1. & 0 & 0 \\ 1. & -1. & 1. & 0 \\ -1. & -1. & -1. & 1. \end{bmatrix} .
\end{aligned}
$$

La matriz P es inversible y

$$
P^{-1} A\, P = \begin{bmatrix} 2. & 0 & -1. & 1. \\ 0 & -2. & 0 & 1. \\ 0 & 0 & 2. & -1. \\ 0 & 0 & 0 & 2. \end{bmatrix} ,
$$

como puede comprobarse.

5.5 Polinomios que anulan una matriz.

5.5.1 Vamos a dedicar esta sección a desarrollar brevemente una idea que tendremos que utilizar en la construcción de las formas canónicas. Conviene que el lector recuerde lo que significan $p(f)$ y $p(A)$ cuando p es un polinomio, f un endomorfismo y A una matriz cuadrada (v. 2.4.8 y 2.4.29).

5.5.2 Si A y A' son matrices semejantes, o sea, si $A' = P^{-1} A\, P$ para una matriz inversible P, y si $p(X)$ es un polinomio, entonces $p(A)$ y $p(A')$ son también matrices semejantes. Más exactamente, se tiene

$$
p(A') = P^{-1} p(A)\, P .
$$

La demostración de este hecho no presenta problemas. Se procede por inducción para probar que

$$
A'^n = P^{-1} A^n\, P
$$

para todo $n \in \mathbb{N}$, y luego se extiende el resultado a los polinomios.

5.5.3 Si $p(X)$ es un polinomio y

$$
A = \begin{bmatrix} \boxed{A_1} & & & \\ & \boxed{A_2} & & \\ & & \ddots & \\ & & & \boxed{A_k} \end{bmatrix}
$$

una matriz cuadrada diagonal por bloques (se entiende que los bloques A_i son cuadrados), entonces

$$p(A) = \begin{bmatrix} p(A_1) & & & \\ & p(A_2) & & \\ & & \ddots & \\ & & & p(A_k) \end{bmatrix}.$$

Esto se puede probar también por inducción para potencias cualesquiera de A, pasando luego a polinomios.

5.5.4 Sea A una matriz $n \times n$ con elementos en $\mathbb{K}$ ($\mathbb{R}$ o $\mathbb{C}$). No es evidente que A sea 'raiz' de algún polinomio, esto es, que

$$p(A) = 0$$

para algún polinomio $p \in \mathbb{K}[X]$ que no sea, claro está, el polinomio 0. Sin embargo existen infinitos polinomios que anulan A; si $p(A) = 0$ es claro que tambiés será $q(A) = 0$ para todo polinomio, $q(X) = r(X)\,p(X)$, que sea múltiplo de p. Basta pues encontrar un polinomio que anule A para tener una infinidad de ellos.

Veamos una primera forma de probar la existencia de un polinomio que anula A. Las $n^2 + 1$ matrices $I, A, A^2, \ldots, A^{n^2}$, constituyen un sistema ligado de $M(n)$, puesto que la dimensión de este espacio es n^2. En consecuencia, existen escalares $\alpha_0, \alpha_1, \alpha_2, \ldots, \alpha_{n^2}$, no todos nulos y tales que

$$\alpha_0 I + \alpha_1 A + \alpha_2 A^2 + \cdots + \alpha_{n^2} A^{n^2} = 0\,,$$

esto es, el polinomio

$$p(X) = \alpha_0 + \alpha_1 X + \alpha_2 X^2 + \cdots + \alpha_{n^2} X^{n^2}$$

verifica que $p(A) = 0$. Nótese que, como los α_i no son todos nulos, el polinomio $p(X)$ no es idénticamente nulo.

Este resultado puede ser un punto de partida, pero en sí mismo no es muy útil. En primer lugar, p puede ser de grado n^2 y veremos que se pueden encontrar polinomios que anulen A y tengan menor grado. Pero, además, no sabemos cuál es el polinomio p, ni aún conociendo la matriz A. Veremos inmediatamente un resultado que permite rebajar el grado del polinomio que anula A y calcular este polinomio.

Conviene recordar que, si A y A' son semejantes, lo son $p(A)$ y $p(A')$ para todo polinomio p. En consecuencia, dos matrices semejantes son anuladas por los mismos polinomios.

Se pueden hacer idénticas consideraciones para un endomorfismo $f \in \mathcal{L}(E)$ de un espacio E de dimensión finita. Para ello basta recordar que, si A representa a f en una base, $p(A)$ representa en la misma base a $p(f)$. Esto significa en particular que un endomorfismo y la matriz que lo representa son anulados por los mismos polinomios.

5.5.5 Ejemplo. Para la matriz

$$A = \begin{bmatrix} -2. & 4. & 5. \\ -3. & 5. & 5. \\ 0 & 0 & 1. \end{bmatrix}$$

del ejemplo (5.3.9), el lector podrá comprobar que

$$A^2 - 3A + 2\,I = 0\,,$$

o sea, $p(A) = 0$ para

$$p(X) = X^2 - 3X + 2 = (X-1)(X-2)\,.$$

Para llegar a esta conclusión no es siquiera necesario realizar el cálculo de $A^2 - 3A + 2\,I$. Basta recordar de (5.3.9) que A es semejante a

$$D = \begin{bmatrix} 1. & & \\ & 1. & \\ & & 2. \end{bmatrix}\,,$$

observar que $p(D) = (D - I)(D - 2\,I) = 0$, y utilizar el hecho de que $p(A)$ y $p(D)$ son semejantes.

Nótese que el polinomio p es un divisor del polinomio característico de A ya que

$$p_A(X) = -(X-1)\,p(X)\,.$$

En consecuencia

$$p_A(A) = -(A - I)\,p(A) = 0\,.$$

Probaremos ahora que esto es lo que ocurre para todas las matrices.

5.5.6 TEOREMA (de Cayley-Hamilton[1])

Toda matriz $A \in M_{\mathbb{K}}(n)$ ($\mathbb{K} = \mathbb{R}$ o $\mathbb{C}$) es raíz de su polinomio característico, es decir, $p_A(A) = 0$.

Todo endomorfismo $f \in \mathcal{L}(E)$, de un espacio finito-dimensional real o complejo, cumple que $p_f(f) = 0$, donde p_f es el polinomio característico de f.

La demostración para el caso real es indirecta y se basa en el caso complejo. Comenzaremos entonces por este último, suponiendo que f es un endomorfismo del espacio complejo E de dimensión $n \neq 0$. La prueba servirá al mismo tiempo para el enunciado con una matriz compleja A, denotando por f en ese caso el endomorfismo de $\mathbb{C}^n$ dado por A.

[1] Así llamado en honor del matemático inglés Arthur Cayley (1821-1895) y del astrónomo y matemático irlandés William Rowan Hamilton (1805-1865).

Sea $p_f(X)$ el polinomio característico de f y $\lambda_1, \lambda_2, \ldots, \lambda_n$ sus raíces (no necesariamente distintas). Entonces

$$p_f(X) = (-1)^n (X - \lambda_1)(X - \lambda_2) \cdots (X - \lambda_n)\,.$$

Por lo tanto (v. 2.4.8),

$$p_f(f) = (-1)^n (f - \lambda_1\, id_E) \circ (f - \lambda_2\, id_E) \circ \cdots \circ (f - \lambda_n\, id_E)\,,$$

lo que abreviaremos, poniendo $g_i = f - \lambda_i\, id_E$, como

$$p_f(f) = (-1)^n g_1 \circ g_2 \circ \cdots \circ g_n\,.$$

Los endomorfismos g_i conmutan entre sí, puesto que todos ellos son polinomios de f. (Además, esta conmutatividad se puede comprobar sin dificultad en este caso concreto.) Denotemos por $(a_1, a_2, \ldots, a_n)$ una base de E en la que la matriz de f sea triangular superior (v. 5.4.3), o sea,

$$[f, (a_i)] = \begin{bmatrix} \lambda_1 & \alpha_2^1 & \cdots & \alpha_n^1 \\ & \lambda_2 & \cdots & \alpha_n^2 \\ & & \ddots & \vdots \\ & & & \lambda_n \end{bmatrix}.$$

Vamos a probar por inducción que, para todo $k = 1, 2, \ldots, n$, el endomorfismo $g_1 \circ g_2 \circ \cdots \circ g_k$ se anula en los vectores $a_1, a_2, \ldots, a_k$. En primer lugar

$$g_1(a_1) = (f - \lambda_1\, id)(a_1) = f(a_1) - \lambda_1 a_1 = \lambda_1 a_1 - \lambda_1 a_1 = 0\,.$$

Si ahora suponemos que $g_1 \circ g_2 \circ \cdots \circ g_k$ (con $k \le n-1$) se anula en $a_1, a_2, \ldots, a_k$, entonces, para $i = 1, 2, \ldots, k$,

$$g_1 \circ \cdots \circ g_k \circ g_{k+1}(a_i) = g_{k+1} \circ g_1 \circ \cdots \circ g_k(a_i) = g_{k+1}(0) = 0\,,$$

mientras que para a_{k+1} , teniendo en cuenta que $f(a_{k+1}) = \alpha_{k+1}^1 a_1 + \cdots + \alpha_{k+1}^k a_k + \lambda_{k+1} a_{k+1}$, se tiene también que

$$\begin{aligned} g_1 \circ \cdots \circ g_k \circ g_{k+1}(a_{k+1}) &= g_1 \circ \cdots \circ g_k(\,(f - \lambda_{k+1}\, id)(a_{k+1})\,) \\ &= g_1 \circ \cdots \circ g_k(f(a_{k+1}) - \lambda_{k+1} a_{k+1}) \\ &= g_1 \circ \cdots \circ g_k(\alpha_{k+1}^1 a_1 + \cdots + \alpha_{k+1}^k a_k) \\ &= 0\,. \end{aligned}$$

Esto termina la prueba por inducción.

Para $k = n$, lo que acabamos de ver significa que $g_1 \circ g_2 \circ \cdots \circ g_n$ se anula en todos los vectores de la base $(a_1, a_2, \ldots, a_n)$; lo mismo ocurre con el endomorfismo $p_f(f) = (-1)^n g_1 \circ g_2 \circ \cdots \circ g_n$, que, por lo tanto, es nulo. O sea, $p_f(f) = 0$.

Si A es una matriz compleja $n \times n$ y f es el endomorfismo de $\mathbb{C}^n$ dado por A, se tiene que $p_A(f) = p_f(f) = 0$, luego $p_A(A) = 0$.

Ocupémonos ahora del caso real. Si A es una matriz real $n \times n$, es también una matriz compleja y, por lo tanto, $p_A(A) = 0$. Finalmente, si f es un endomorfismo de un espacio real E y A es la matriz de f en una base cualquiera, tenemos que $p_f(A) = p_A(A) = 0$, luego $p_f(f) = 0$.

5.5.7 Obsérvese que, en el caso real, p_A y p_f son polinomios con coeficientes reales. Lo que impide hacer en el caso real la misma demostración directa del caso complejo es que p_A y p_f pueden no admitir una descomposición

$$p_f(X) = (-1)^n (X - \lambda_1)(X - \lambda_2) \cdots (X - \lambda_n)$$

con números λ_i reales, porque son polinomios que pueden tener raíces complejas no reales.

5.5.8 Al margen de la utilización que muy pronto haremos del teorema precedente, existen algunas aplicaciones inmediatas que es interesante citar.

Si A es una matriz $n \times n$, $p_A(A) = 0$, o sea,

$$(-1)^n A^n + (-1)^{n-1} \operatorname{tr} A \cdot A^{n-1} + \cdots + (\det A)\, I = 0\,.$$

De la igualdad anterior se puede despejar A^n, con lo que se obtiene A^n como un polinomio de A de grado igual o menor a $n - 1$. Lo mismo se puede hacer a continuación con las potencias $A^{n+1}, A^{n+2}, \ldots$

Resulta así que cualquier potencia de A y, más generalmente, cualquier polinomio de A coincide con un polinomio de A de grado igual o menor que $n - 1$.

5.5.9 **Ejemplo.** La matriz

$$A = \begin{bmatrix} 5. & 4. \\ -2. & -1. \end{bmatrix}$$

tiene por polinomio característico

$$p_A(X) = X^2 - 4X + 3$$

luego

$$A^2 - 4A + 3\, I = 0$$

y

$$A^2 = 4A - 3\, I\,.$$

En consecuencia, cualquier potencia de A es de la forma

$$A^n = \alpha_n A + \beta_n\, I\,.$$

Por ejemplo,

$$A^3 = A^2 A = (4A - 3\, I)A = 4A^2 - 3A = 4(4A - 3\, I) - 3A = 13A - 12\, I\,.$$

Es fácil obtener una relación de recurrencia entre los coeficientes α_n y β_n y los α_{n+1} y β_{n+1}. Como

$$\begin{aligned} A^{n+1} &= A^n A = (\alpha_n A + \beta_n\, I)A = \alpha_n A^2 + \beta_n A \\ &= \alpha_n (4A - 3\, I) + \beta_n A = (4\alpha_n + \beta_n)A - 3\alpha_n\, I\,, \end{aligned}$$

resulta que

$$\alpha_{n+1} = 4\alpha_n + \beta_n \quad \text{y} \quad \beta_{n+1} = -3\alpha_n \,.$$

Sabemos además que $\alpha_0 = 0$, $\beta_0 = 1$, $\alpha_1 = 1$ y $\beta_1 = 0$. Se puede probar fácilmente por inducción que, para todo n,

$$\alpha_n = \frac{3^n - 1}{2} \quad \text{y} \quad \beta_n = \frac{-3^n + 3}{2} \,.$$

Por lo tanto, para todo $n \in \mathbb{N}$,

$$A^n = \frac{3^n - 1}{2} A + \frac{-3^n + 3}{2} I \,,$$

o sea,

$$A^n = \begin{bmatrix} 2 \cdot 3^n - 1 & 2 \cdot 3^n - 2 \\ -3^n + 1 & -3^n + 2 \end{bmatrix} .$$

5.5.10 Supongamos ahora que A es una matriz $n \times n$ inversible; denotemos por $p_A(X) = \alpha_n X^n + \alpha_{n-1} X^{n-1} + \cdots + \alpha_1 X + \alpha_0$ el polinomio característico de A. Sabemos que

$$\alpha_n A^n + \alpha_{n-1} A^{n-1} + \cdots + \alpha_1 A + \alpha_0 I = 0 \,.$$

Multiplicando por A^{-1} obtenemos

$$\alpha_n A^{n-1} + \alpha_{n-1} A^{n-2} + \cdots + \alpha_1 I + \alpha_0 A^{-1} = 0$$

y, como $\alpha_0 = \det A \neq 0$, podemos despejar A^{-1}

$$A^{-1} = -\frac{1}{\alpha_0} (\alpha_n A^{n-1} + \alpha_{n-1} A^{n-2} + \cdots + \alpha_1 I) \,.$$

Resulta así la inversa de A como un polinomio de A de grado $n - 1$.

5.5.11 Ejemplo. Para la matriz

$$A = \begin{bmatrix} 2. & 2. & 1. \\ 4. & 4. & 1. \\ 1. & 0 & 0 \end{bmatrix}$$

(es la misma que en el ejemplo 2.4.36) el polinomio característico vale

$$p_A(X) = -X^3 + 6X^2 + X - 2 \,,$$

luego

$$-A^3 + 6A^2 + A - 2I = 0 \,,$$
$$-A^2 + 6A + I - 2A^{-1} = 0 \,,$$

y, por tanto,

$$A^{-1} = \frac{1}{2}(-A^2 + 6A + I) \,.$$

Se puede comprobar que, efectivamente, se obtiene así la inversa que ya fue calculada en (2.4.36).

5.5.12 Sea A una matriz $n \times n$. El polinomio característico $p_A(X)$ anula A, pero puede no ser el polinomio de menor grado que lo hace.

Por ejemplo, si A es diagonalizable y tiene valores propios repetidos se produce esta situación. Es lo que ocurre con la matriz de los ejemplos (5.3.9) y (5.5.5). Su polinomio característico es

$$p_A(X) = -X^3 + 4X^2 - 5X + 2 = (1 - X)^2(2 - X)$$

y p_A anula A, pero también anula A el polinomio

$$p(X) = X^2 - 3X + 2 = (1 - X)(2 - X)\,,$$

como ya vimos.

La descripción de los polinomios que anulan una matriz cuadrada A se puede resumir en los dos puntos siguientes:

— Existe un único polinomio (salvo multiplicación por constantes) de entre los que anulan A que posee grado mínimo. Se denomina *polinomio minimal* de A.

— Los polinomios que anulan A son exactamente los múltiplos del polinomio minimal.

A pesar de que no es particularmente difícil, no nos detendremos en justificar estas afirmaciones, ya que no vamos a utilizarlas en el posterior desarrollo del tema.

Naturalmente, si $p(X)$ es un polinomio minimal de A, también lo es $\alpha\, p(X)$ para cualquier $\alpha \neq 0$. A veces se acuerda tomar como polinomio minimal el que tiene coeficiente 1 en el término de mayor grado, pero, generalmente, la expresión 'polinomio minimal' sirve para designar a cualquiera de ellos.

Dos matrices semejantes poseen el mismo polinomio minimal, puesto que son anuladas por los mismos polinomios (v. 5.5.2).

Como el polinomio característico anula la matriz, resulta que es un múltiplo del polinomio minimal (coincidente a veces con él).

El principal inconveniente del polinomio minimal es la inexistencia de un método sencillo y general para el cálculo de este polinomio, al contrario de lo que ocurre con el polinomio característico.

En los casos sencillos se puede encontrar por tanteo, sabiendo que divide al polinomio característico y que posee sus mismas raíces. Es decir, si el polinomio característico de A es

$$p_A(X) = (\lambda_1 - X)^{k_1}(\lambda_2 - X)^{k_2} \cdots (\lambda_p - X)^{k_p}$$

con $\lambda_1, \lambda_2, \ldots, \lambda_p$ distintos, entonces el polinomio minimal de A es de la forma

$$p(X) = (\lambda_1 - X)^{l_1}(\lambda_2 - X)^{l_2} \cdots (\lambda_p - X)^{l_p}$$

con $1 \leq l_i \leq k_i$.

5.5.13 Ejemplo. La matriz de los ejemplos (5.3.9) y (5.5.5) tiene por polinomio característico

$$p_A(X) = -X^3 + 4X^2 - 5X + 2 = (1-X)^2(2-X)\,,$$

luego el polinomio minimal es o bien $p_A(X)$ o bien $p(X) = (1-X)(2-X)$. Como ya hemos comprobado, este último anula A, luego es el polinomio minimal de A.

5.6 Forma canónica de endomorfismos y matrices.

5.6.1 Lo que pretendemos en esta sección es conseguir representar los endomorfismos mediante matrices diagonales por bloques

$$A = \begin{bmatrix} \boxed{A_1} & & & \\ & \boxed{A_2} & & \\ & & \ddots & \\ & & & \boxed{A_p} \end{bmatrix},$$

en las que cada bloque diagonal tenga la forma

$$\begin{bmatrix} \lambda & 1 & & & \\ & \lambda & \ddots & & \\ & & \ddots & \ddots & \\ & & & \lambda & 1 \\ & & & & \lambda \end{bmatrix}.$$

Una matriz diagonal por bloques con bloques de este tipo se llama *matriz de Jordan*[2]. Las matrices de la forma

$$\begin{bmatrix} \lambda & 1 & & & \\ & \lambda & \ddots & & \\ & & \ddots & \ddots & \\ & & & \lambda & 1 \\ & & & & \lambda \end{bmatrix}$$

se denominan *cajas elementales de Jordan.*

Así pues, lo que llamaremos matriz de Jordan es una matriz diagonal por bloques cuyos bloques diagonales sean cajas elementales.

En (5.1.3) vimos un ejemplo de matriz de Jordan con dos cajas elementales en la diagonal.

Nótese que una matriz de Jordan es triangular superior. Además los únicos elementos que pueden no ser nulos son los diagonales y los de la superdiagonal (los α_i^j con $i = j+1$); estos últimos serán 1 o 0 según correspondan al interior de

[2] del ingeniero francés Camille Jordan (1838-1921).

una caja elemental o a la separación entre dos de estas cajas. Como las matrices de Jordan son triangulares, sus elementos diagonales serán los valores propios de la matriz.

Ya vimos en (5.1.2) la idea esencial para conseguir matrices diagonales por bloques; estudiaremos ahora la forma de conseguir que los bloques diagonales sean cajas elementales de Jordan.

5.6.2 Concentremos nuestra atención en una caja elemental con diagonal nula, es decir en una matriz $n \times n$ de la forma

$$A = \begin{bmatrix} 0 & 1 & & & \\ & 0 & \ddots & & \\ & & \ddots & \ddots & \\ & & & 0 & 1 \\ & & & & 0 \end{bmatrix}.$$

Si f es un endomorfismo que se representa por esta matriz, entonces $f^n = 0$, o sea $A^n = 0$. En efecto, si $[f, (a_i)] = A$,

$$f(a_1) = 0\,, \quad f(a_2) = a_1\,, \ldots, f(a_n) = a_{n-1}\,,$$

luego

$$f(a_1) = 0\,, \quad f^2(a_2) = 0\,, \ldots, f^n(a_n) = 0\,,$$

y, en consecuencia,

$$f^n(a_i) = 0\,, \quad i = 1, 2, \ldots, n\,.$$

Por lo tanto $f^n = 0$.

Nótese que la base $(a_1, \ldots, a_n)$ en la que $[f, (a_i)] = A$ es una base formada por las imágenes sucesivas

$$f^{n-1}(a_n)\,, \ldots, f^2(a_n), f(a_n), a_n\,,$$

del vector a_n.

5.6.3 DEFINICIÓN. Sea E un e.v. sobre $\mathbb{K}$ y $f \in \mathcal{L}(E)$; decimos que f es un endomorfismo *nilpotente* cuando $f^p = 0$ para algún $p \in \mathbb{N}$. Obviamente, si $f^p = 0$, entonces $f^{p+1} = f^{p+2} = \cdots = 0$. Si f es nilpotente, existe un único $p \in \mathbb{N}$ tal que

$$f^{p-1} \neq 0 \quad \text{y} \quad f^p = 0\,;$$

se dice entonces que f es *nilpotente de orden* p. Si $f^k = 0$, el orden de f es igual o menor que k.

Si A es una matriz cuadrada y $A^p = 0$ para algún $p \in \mathbb{N}$, se dice que A es una matriz *nilpotente*. Si p es tal que

$$A^{p-1} \neq 0 \quad \text{y} \quad A^p = 0\,,$$

se dice que A es *nilpotente de orden* p.

Cuando A representa a f en alguna base, A^k representa en la misma base a f^k, luego f es nilpotente si y sólo si lo es A, y, en ese caso, lo son del mismo orden.

5.6.4 Hemos visto en (5.6.2) que toda caja elemental $n \times n$ de diagonal nula (y todo endomorfismo representado por una matriz de ese tipo) es nilpotente; y no es difícil ver que lo es de orden n (v. 5.6.24).

La afirmación más o menos recíproca de ésta será el primer resultado importante de esta sección (v. 5.6.8). Vamos a anteponer algunos resultados sencillos que tendremos que utilizar.

5.6.5 PROPOSICIÓN

Si $f \in \mathcal{L}(E)$, $x \in E$, $f^{k-1}(x) \neq 0$ y $f^k(x) = 0$, el sistema

$$(x, f(x), \ldots, f^{k-1}(x))$$

es libre.

En efecto, los vectores $x, f(x), \ldots, f^{k-1}(x)$ son no nulos y si

$$\alpha_0 x + \alpha_1 f(x) + \cdots + \alpha_{k-1} f^{k-1}(x) = 0\,,$$

aplicando sucesivamente $f^{k-1}, f^{k-2}, \ldots, f$ a ambos miembros de la igualdad, se obtiene que $\alpha_0 = 0$, $\alpha_1 = 0$, etc.

5.6.6 PROPOSICIÓN

Si $f \in \mathcal{L}(E)$, entonces

$$\{0\} = \operatorname{Ker} f^0 \subset \operatorname{Ker} f \subset \operatorname{Ker} f^2 \subset \cdots$$

y

$$E = \operatorname{Im} f^0 \supset \operatorname{Im} f \supset \operatorname{Im} f^2 \supset \cdots$$

La demostración es un ejercicio sencillo. Para $\operatorname{Ker} f^0$ e $\operatorname{Im} f^0$, recuérdese que $f^0 = id_E$.

5.6.7 Si f es nilpotente de orden p se tiene además

$$\operatorname{Ker} f^{p-1} \subsetneqq \operatorname{Ker} f^p = E = \operatorname{Ker} f^{p+1} = \operatorname{Ker} f^{p+2} = \cdots$$

y

$$\operatorname{Im} f^{p-1} \supsetneqq \operatorname{Im} f^p = \{0\} = \operatorname{Im} f^{p+1} = \operatorname{Im} f^{p+2} = \cdots$$

5.6.8 Obtención de la matriz de Jordan de un endomorfismo nilpotente.

Sea E un espacio de dimensión $n \neq 0$ y $f \in \mathcal{L}(E)$ un endomorfismo nilpotente de orden p.

Pongamos $K_i = \operatorname{Ker} f^i$, $i = 0, 1, \ldots, p$; sabemos que

$$\{0\} = K_0 \subset K_1 \subset K_2 \subset \cdots \subset K_{p-1} \subset K_p = E\,.$$

Los números $n_i = \dim K_i$, $i = 0, 1, \ldots, p$ verifican que

$$0 = n_0 \leq n_1 \leq n_2 \leq \cdots \leq n_{p-1} \leq n_p = n\,;$$

sus diferencias $d_i = n_i - n_{i-1}$, $i = 1, \ldots, p$, cumplen $d_1 + d_2 + \cdots + d_p = n$. Nótese que $d_1 = n_1$. Por otra parte, como $f^{p-1} \neq 0$, $K_{p-1} \neq E$ y $n_{p-1} < n$, luego $d_p \geq 1$.

Consideremos un suplementario, G_p , de K_{p-1} en $K_p = E$, o sea,

$$E = K_p = G_p \oplus K_{p-1}\,;$$

G_p será de dimensión d_p . Tomemos una base $(a_1, \ldots, a_{d_p})$ de G_p . Los vectores

$$f(a_1), \ldots, f(a_{d_p})$$

pertenecen a K_{p-1} , forman un sistema libre y

$$\langle f(a_1), \ldots, f(a_{d_p})\rangle \cap K_{p-2} = \{0\}\,.$$

La primera afirmación se comprueba sin dificultad. Para la segunda, supongamos que

$$\alpha_1 f(a_1) + \cdots + \alpha_{d_p} f(a_{d_p}) = 0\,;$$

entonces

$$f^{p-1}(\alpha_1 a_1 + \cdots + \alpha_{d_p} a_{d_p}) = f^{p-2}(\alpha_1 f(a_1) + \cdots + \alpha_{d_p} f(a_{d_p})\,) = f^{p-2}(0) = 0\,,$$

luego $\alpha_1 a_1 + \cdots + \alpha_{d_p} a_{d_p} \in K_{p-1}$ y entonces $\alpha_1 a_1 + \cdots + \alpha_{d_p} a_{d_p} = 0$, lo que significa que $\alpha_1 = \cdots = \alpha_{d_p} = 0$. Para probar la tercera de las afirmaciones, supongamos que

$$\alpha_1 f(a_1) + \cdots + \alpha_{d_p} f(a_{d_p}) = x \in K_{p-2}\,;$$

se obtiene como en el caso anterior que

$$f^{p-1}(\alpha_1 a_1 + \cdots + \alpha_{d_p} a_{d_p}) = 0$$

y que $\alpha_1 = \cdots = \alpha_{d_p} = 0$, lo que significa que $x = 0$.

Consideremos ahora el subespacio $\langle f(a_1), \ldots, f(a_{d_p})\rangle \oplus K_{p-2}$ de K_{p-1} ; nótese que forzosamente $d_p + n_{p-2} \leq n_{p-1}$, o sea, $d_p \leq d_{p-1}$. Denotemos por G_{p-1} un suplementario en K_{p-1} de dicho subespacio, es decir

$$K_{p-1} = G_{p-1} \oplus \langle f(a_1), \ldots, f(a_{d_p})\rangle \oplus K_{p-2}\,.$$

Ahora es posible que $G_{p-1} = \{0\}$, si es que $d_p = d_{p-1}$. Tomemos una base $(a_{d_p+1}, \ldots, a_{d_{p-1}})$ de G_{p-1}, que estará formada por $d_{p-1} - d_p$ vectores. Los vectores

$$f^2(a_1), \ldots, f^2(a_{d_p}), f(a_{d_p+1}), \ldots, f(a_{d_{p-1}})$$

pertenecen a K_{p-2}, forman un sistema libre y

$$\langle f^2(a_1), \ldots, f(a_{d_{p-1}})\rangle \cap K_{p-3} = \{0\}\,.$$

La comprobación de estos hechos es muy parecida a la que hemos detallado más arriba.

El proceso continúa de la misma manera hasta que se obtiene el último suplementario G_1,

$$K_1 = G_1 \oplus \langle f^{p-1}(a_1), \ldots, f(a_{d_2})\rangle$$

(recuérdese que $K_0 = \{0\}$), suplementario que tendrá dimensión $d_1 - d_2$. Se elige finalmente una base $(a_{d_2+1}, \ldots, a_{d_1})$ de G_1.

Consideremos ahora los $d_p + d_{p-1} + \cdots + d_2 + d_1 = n$ vectores que hemos ido obteniendo

$$\begin{array}{cccccccccc}
a_1 & \cdots & a_{d_p} & & & & & & & \\
f(a_1) & \cdots & f(a_{d_p}) & a_{d_p+1} & \cdots & a_{d_{p-1}} & & & & \\
\vdots & & \vdots & \vdots & & \vdots & & & & \\
f^{p-2}(a_1) & \cdots & f^{p-2}(a_{d_p}) & f^{p-3}(a_{d_p+1}) & \cdots & f^{p-3}(a_{d_{p-1}}) & \cdots & a_{d_2} & & \\
f^{p-1}(a_1) & \cdots & f^{p-1}(a_{d_p}) & f^{p-2}(a_{d_p+1}) & \cdots & f^{p-2}(a_{d_{p-1}}) & \cdots & f(a_{d_2}) & a_{d_2+1} \cdots & a_{d_1}
\end{array}$$

organizados en p filas y en $d_1 = n_1$ columnas. En cada columna figura un vector y sus imágenes sucesivas; nótese que, en cada caso, la imagen del vector por la siguiente potencia de f ya es nula. El número de columnas de las diferentes alturas es d_p, $d_{p-1}-d_p$, ..., d_1-d_2; sabemos que $d_p \geq 1$, pero las otras cantidades pueden ser nulas todas ellas.

En este conjunto de n vectores, la última fila constituye una base de K_1; al reunirla con la penúltima se obtiene una base de K_2, y así sucesivamente. La reunión de todas ellas forma una base de $K_p = E$.

Los vectores $a_i, f(a_i), f^2(a_i), \ldots, f^k(a_i)$, de una misma columna, forman una base de un subespacio que es invariante para f. Si se considera la restricción de f a este subespacio, y la base formada por los vectores de la columna en el orden

$$(f^k(a_i), \ldots, f^2(a_i), f(a_i), a_i)\,,$$

la matriz en esta base es una caja elemental de la forma

$$\begin{bmatrix} 0 & 1 & & & \\ & 0 & \ddots & & \\ & & \ddots & \ddots & \\ & & & 0 & 1 \\ & & & & 0 \end{bmatrix}.$$

En consecuencia, si se toma la base de E que hemos obtenido, en el orden

$$(f^{p-1}(a_1), \ldots, f(a_1), a_1, \ldots, f(a_{d_2}), a_{d_2}, a_{d_2+1}, \ldots, a_{d_1}),$$

la matriz del endomorfismo nilpotente f resulta ser

$$\begin{bmatrix} \boxed{A_1} & & \\ & \ddots & \\ & & \boxed{A_{d_1}} \end{bmatrix},$$

diagonal por bloques, con bloques diagonales de tamaño decreciente que son cajas elementales con ceros en la diagonal.

Nótese que la base puede ser costosa de obtener, pero el número de las cajas y los tamaños de las cajas resultan de un simple estudio de los rangos de las potencias de f, que proporcionan las dimensiones n_i y las diferencias d_i.

5.6.9 Ejemplo. El endomorfismo f de $\mathbb{R}^4$ dado por la matriz

$$A = \begin{bmatrix} -1. & 1. & 0 & 0 \\ -1. & 0 & 1. & 0 \\ -1. & 0 & 1. & 0 \\ -1. & 0 & 1. & 0 \end{bmatrix}$$

es nilpotente de orden 3, ya que se tiene que

$$A^2 = \begin{bmatrix} 0 & -1. & 1. & 0 \\ 0 & -1. & 1. & 0 \\ 0 & -1. & 1. & 0 \\ 0 & -1. & 1. & 0 \end{bmatrix}$$

y que $A^3 = 0$. Como $\operatorname{rg} A = 2$ y $\operatorname{rg} A^2 = 1$, tenemos (utilizando las notaciones del apartado precedente) que

$$\begin{array}{llll} n_0 = 0 & n_1 = 2 & n_2 = 3 & n_3 = 4 \\ & d_1 = 2 & d_2 = 1 & d_3 = 1 \ . \end{array}$$

La base de $\mathbb{R}^4$ que buscamos tendrá la estructura

$$\begin{array}{ll} a_1 & \\ f(a_1) & \\ f^2(a_1) & a_2\,, \end{array}$$

y al ordenarla como

$$(f^2(a_1), f(a_1), a_1, a_2)$$

la matriz de f es esta base será

$$\left[\begin{array}{ccc|c} 0 & 1 & & \\ & 0 & 1 & \\ & & 0 & \\ \hline & & & 0 \end{array}\right].$$

Busquemos ahora los vectores a_1 y a_2 apropiados. $K_2 = \operatorname{Ker} f^2$ es el subespacio de dimensión 3

$$K_2 = \{(x, y, z, t) \in \mathbb{R}^4 \mid -y + z = 0\} \,.$$

Basta tomar como a_1 cualquier vector no nulo que no pertenezca a K_2, ya que entonces

$$\mathbb{R}^4 = \langle a_1 \rangle \oplus K_2 \,.$$

Tomaremos, por ejemplo,

$$a_1 = (0, 1, 0, 0) \,.$$

Este vector tiene por imagen

$$f(a_1) = (1, 0, 0, 0) \,.$$

$K_1 = \operatorname{Ker} f$ es el subespacio de dimensión 2 con ecuación

$$\begin{array}{rcrcrcl} -x & + & y & & & = & 0 \\ -x & & & + & z & = & 0 \;. \end{array}$$

En este nivel no hay que escoger ningún vector ya que

$$K_2 = \langle f(a_1) \rangle \oplus K_1 \,,$$

como sabemos por el estudio de las dimensiones, aunque el lector puede comprobar directamente este hecho. La columna se completa con el vector

$$f^2(a_1) = f(f(a_1)\,) = (-1, -1, -1, -1) \,.$$

El vector a_2 es cualquier vector que forme con $f^2(a_1)$ una base de K_1. Tomaremos, por ejemplo,

$$a_2 = (0, 0, 0, 1) \,.$$

En la base

$$(\,(-1, -1, -1, -1), (1, 0, 0, 0), (0, 1, 0, 0), (0, 0, 0, 1)\,)$$

el endomorfismo f tendrá la matriz que hemos descrito antes. El lector podrá comprobar que, para

$$P = \begin{bmatrix} -1 & 1. & 0 & 0 \\ -1 & 0 & 1. & 0 \\ -1 & 0 & 0 & 0 \\ -1 & 0 & 0 & 1. \end{bmatrix} \,,$$

se tiene que

$$P^{-1} A\, P = \left[\begin{array}{ccc|c} 0 & 1 & & \\ & 0 & 1 & \\ & & 0 & \\ \hline & & & 0 \end{array} \right] \,.$$

5.6.10 Obtención de la matriz de Jordan de un endomorfismo con un sólo valor propio de multiplicidad igual a la dimensión del espacio.

Sea E un espacio sobre $\mathbb{K}$ ($\mathbb{R}$ o $\mathbb{C}$) de dimensión $n \neq 0$ y $f \in \mathcal{L}(E)$ un endomorfismo. Supondremos que f posee un solo valor propio $\lambda \in \mathbb{K}$ de multiplicidad n. El polinomio característico de f es entonces

$$p_f(X) = (-1)^n (X - \lambda)^n \,.$$

De acuerdo con el teorema de Cayley-Hamilton se tiene que $p_f(f) = 0$, o sea,

$$(f - \lambda\, id_E)^n = 0\,.$$

El endomorfismo $g = f - \lambda\, id_E$ es nilpotente de orden igual o menor que n.

El método que hemos explicado en (5.6.8) permite entonces obtener una base $(a_1, \ldots, a_n)$ en la que la matriz de g es diagonal por bloques

$$[g, (a_i)] = \begin{bmatrix} A_1 & & \\ & \ddots & \\ & & A_p \end{bmatrix},$$

con bloques diagonales A_i de la forma

$$\begin{bmatrix} 0 & 1 & & & \\ & 0 & \ddots & & \\ & & \ddots & \ddots & \\ & & & 0 & 1 \\ & & & & 0 \end{bmatrix}.$$

Como $f = g + \lambda\, id_E$, es inmediato comprobar que la matriz de f en la misma base es diagonal por bloques,

$$[f, (a_i)] = \begin{bmatrix} B_1 & & \\ & \ddots & \\ & & B_p \end{bmatrix},$$

con bloques diagonales B_i de la forma

$$\begin{bmatrix} \lambda & 1 & & & \\ & \lambda & \ddots & & \\ & & \ddots & \ddots & \\ & & & \lambda & 1 \\ & & & & \lambda \end{bmatrix}.$$

La matriz $[f, (a_i)]$ es pues una matriz de Jordan.

5.6.11 Ejemplo. El endomorfismo f de $\mathbb{R}^4$ dado por la matriz

$$A = \begin{bmatrix} 0 & 1. & 0 & 0 \\ -1. & 1. & 1. & 0 \\ -1. & 0 & 2. & 0 \\ -1. & 0 & 1. & 1. \end{bmatrix}$$

tiene como polinomio característico

$$p_f(X) = (X-1)^4\,,$$

luego un solo valor propio, 1, de multiplicidad 4.

El endomorfismo $g = f - id$ viene dado por la matriz

$$B = \begin{bmatrix} -1. & 1. & 0 & 0 \\ -1. & 0 & 1. & 0 \\ -1. & 0 & 1. & 0 \\ -1. & 0 & 1. & 0 \end{bmatrix},$$

o sea, es el mismo endomorfismo nilpotente que fue estudiado en el ejemplo (5.6.9).

Para la base

$$(\,(-1,-1,-1,-1),(1,0,0,0),(0,1,0,0),(0,0,0,1)\,)$$

de $\mathbb{R}^4$, la matriz de g es

$$\left[\begin{array}{ccc|c} 0 & 1 & & \\ & 0 & 1 & \\ & & 0 & \\ \hline & & & 0 \end{array}\right],$$

luego la matriz de $f = g + id$ en esa misma base es

$$\left[\begin{array}{ccc|c} 1 & 1 & & \\ & 1 & 1 & \\ & & 1 & \\ \hline & & & 1 \end{array}\right].$$

5.6.12 Nos queda por ver lo que ocurre en el caso general. Para ello necesitamos un resultado esencial, que es el que presentamos en (5.6.14).

Sea E un espacio sobre $\mathbb{K}$ de dimensión $n \neq 0$ y $f \in \mathcal{L}(E)$ un endomorfismo. Supongamos que $p_f(X)$ posee n raíces en $\mathbb{K}$, iguales o distintas. Denotemos por $\lambda_1, \ldots, \lambda_p \in \mathbb{K}$ las raíces de p_f y por $k_1, \ldots, k_p$ sus multiplicidades, que sumarán $k_1 + \cdots + k_p = n$.

Si $V(\lambda_1), \ldots, V(\lambda_p)$ son los correspondientes subespacios propios, la suma directa (v. 5.1.16)

$$V(\lambda_1) \oplus \cdots \oplus V(\lambda_p)$$

puede ser todo E cuando f es diagonalizable (v. 5.3.6) pero, cuando f no es diagonalizable, es un subespacio diferente de E.

Lo que se puede hacer es substituir cada subespacio propio $V(\lambda_i)$, que es $\operatorname{Ker}(f - \lambda_i\, id_E)$, por alguno de los núcleos $\operatorname{Ker}(f - \lambda_i\, id_E)^r$, que son mayores (v. 5.6.6); se suele dar el nombre de *subespacios propios generalizados* a estos núcleos. Veremos que los que sirven son justamente los

$$F_i = \operatorname{Ker}(f - \lambda_i\, id_E)^{k_i},$$

donde el exponente k_i es la multiplicidad de λ_i como raíz.

Comenzaremos por algún resultado sencillo, para terminar probando la proposición (5.6.14), que es el resultado de demostración más difícil de entre los que se proponen en esta sección.

Los subespacios generalizados, $\operatorname{Ker}(f - \lambda_i\, id_E)^r$, son invariantes para f. En efecto, si $x \in \operatorname{Ker}(f - \lambda_i\, id)^r$, entonces, teniendo en cuenta que f y $f - \lambda_i\, id$ conmutan, resulta para $f(x)$ que

$$(f - \lambda_i\, id)^r(f(x)) = (f - \lambda_i\, id)^r \circ f(x) = f \circ (f - \lambda_i\, id)^r(x) = f(0) = 0,$$

por lo que $f(x) \in \operatorname{Ker}(f - \lambda_i\, id)^r$.

El resultado que anunciamos ahora es una generalización de (5.2.13).

5.6.13 PROPOSICIÓN

Sea E un e.v. sobre $\mathbb{K}$ de dimensión $n \neq 0$, $f \in \mathcal{L}(E)$; sea $\lambda \in \mathbb{K}$ un valor propio de f. Denotemos por k el orden de multiplicidad de λ como raíz de $p_f(X)$. Si $r \in \mathbb{N}$, entonces el subespacio $\operatorname{Ker}(f - \lambda\, id_E)^r$ verifica que

$$1 \leq \dim \operatorname{Ker}(f - \lambda\, id_E)^r \leq k.$$

Pongamos $F = \operatorname{Ker}(f - \lambda\, id)^r$. Como

$$V(\lambda) = \operatorname{Ker}(f - \lambda\, id) \subset \operatorname{Ker}(f - \lambda\, id)^r = F$$

y λ es un valor propio, $F \neq \{0\}$ y $\dim F \geq 1$.

Denotemos por h la dimensión de F; vamos a probar que $h \leq k$. El endomorfismo inducido por f en F verifica que $(f - \lambda\, id)^r = 0$, a causa de la propia definición de F. Si g es el endomorfismo de F dado por $g = f - \lambda\, id$, g es nilpotente y existe una base $(a_1, \ldots, a_h)$ de F tal que

$$[g, (a_1, \ldots, a_h)] = \begin{bmatrix} 0 & \alpha_1 & & & \\ & 0 & \ddots & & \\ & & \ddots & \ddots & \\ & & & 0 & \alpha_{h-1} \\ & & & & 0 \end{bmatrix},$$

donde los α_i valen 0 o 1. Para esta misma base,

$$[f,(a_1,\ldots,a_h)] = \begin{bmatrix} \lambda & \alpha_1 & & & \\ & \lambda & \ddots & & \\ & & \ddots & \ddots & \\ & & & \lambda & \alpha_{h-1} \\ & & & & \lambda \end{bmatrix}.$$

Completemos $(a_1,\ldots,a_h)$ hasta formar una base $(a_1,\ldots,a_h,\ldots,a_n)$ de E. La matriz de f en esta base de E es de la forma

$$\left[\begin{array}{ccccc|c} \lambda & \alpha_1 & & & & \\ & \lambda & \ddots & & & \\ & & \ddots & \ddots & & A' \\ & & & \lambda & \alpha_{h-1} & \\ & & & & \lambda & \\ \hline & & 0 & & & A'' \end{array}\right],$$

por lo que, siguiendo el mismo razonamiento que en (5.2.13),

$$p_f(X) = (\lambda - X)^h q(X)\,,$$

lo que demuestra que k (multiplicidad de la raíz λ) es igual o mayor que h.

5.6.14 PROPOSICIÓN

Sea E un e.v. sobre $\mathbb{K}$ ($\mathbb{R}$ o $\mathbb{C}$) de dimensión $n \neq 0$ y $f \in \mathcal{L}(E)$. Supongamos que el polinomio característico $p_f(X)$ posee n raíces en $\mathbb{K}$, iguales o distintas. Denotemos por $\lambda_1,\ldots,\lambda_p$ las raíces diferentes de $p_f(X)$ y por $k_1,\ldots,k_p$ sus multiplicidades. Pongamos

$$F_i = \mathrm{Ker}(f - \lambda_i\, id_E)^{k_i}\,, \quad i = 1,\ldots,p\,.$$

Entonces se tiene que

$$\dim F_i = k_i\,, \quad i = 1,\ldots,p$$

y que los subespacios $F_1,\ldots,F_p$ son invariantes y verifican

$$E = F_1 \oplus \cdots \oplus F_p\,.$$

El polinomio característico de f será

$$p_f(X) = (-1)^n (X - \lambda_1)^{k_1} \cdots (X - \lambda_p)^{k_p}\,,$$

y el teorema de Cayley-Hamilton nos garantiza que $p_f(f) = 0$.

Recordemos la conmutatividad entre dos polinomios cualesquiera del endomorfismo f (v. 2.4.8); tendremos que utilizarla varias veces a lo largo de la demostración.

Denotemos por $p_i(X)$ los polinomios

$$p_i(X) = \frac{(-1)^n p_f(X)}{(X - \lambda_i)^{k_i}}, \quad i = 1, \ldots, p,$$

resultantes de suprimir el factor $(X - \lambda_i)^{k_i}$ en el polinomio característico. Se tiene para todo $i = 1, \ldots, p$ que

$$(X - \lambda_i)^{k_i} p_i(X) = (-1)^n p_f(X),$$

luego

$$\textbf{(1)} \qquad (f - \lambda_i \, id)^{k_i} \circ p_i(f) = 0.$$

Observemos también que, si $x \in F_i = \operatorname{Ker}(f - \lambda_i \, id)^{k_i}$, entonces

$$\textbf{(2)} \qquad p_j(f)(x) = 0, \quad \text{para } j \neq i,$$

ya que $p_j(f) = q(f) \circ (f - \lambda_i \, id)^{k_i}$, donde q es un polinomio.

Como $\lambda_1, \ldots, \lambda_p$ son diferentes, los polinomios $p_1(X), \ldots, p_p(X)$ son primos entre sí; en consecuencia, existen polinomios $q_1(X), \ldots, q_p(X)$ tales que

$$q_1(X)p_1(X) + \cdots + q_p(X)p_p(X) = 1.$$

Resulta así que

$$\textbf{(3)} \qquad q_1(f) \circ p_1(f) + \cdots + q_p(f) \circ p_p(f) = id_E.$$

Comencemos por probar que

$$E = F_1 + \cdots + F_p.$$

Sea $x \in E$. Por la igualdad **(3)**

$$x = q_1(f) \circ p_1(f)(x) + \cdots + q_p(f) \circ p_p(f)(x) = x_1 + \cdots + x_p,$$

donde hemos puesto

$$x_i = q_i(f) \circ p_i(f)(x), \quad i = 1, \ldots, p.$$

La imagen de x_i por $(f - \lambda_i \, id)^{k_i}$ vale

$$\begin{aligned}(f - \lambda_i \, id)^{k_i}(x_i) &= (f - \lambda_i \, id)^{k_i} \circ q_i(f) \circ p_i(f)(x) \\ &= q_i(f) \circ (f - \lambda_i \, id)^{k_i} \circ p_i(f)(x) \\ &= 0,\end{aligned}$$

como consecuencia de **(1)**. Así pues

$$x_i \in \operatorname{Ker}(f - \lambda_i\, id)^{k_i} = F_i\,.$$

Esto demuestra que la suma de los subespacios F_i es E.

Para probar que la suma es directa bastará con que mostremos que si

$$0 = x_1 + \cdots + x_p$$

con $x_1 \in F_1, \ldots, x_p \in F_p$, necesariamente $x_1 = \cdots = x_p = 0$. Veamos que eso es lo que ocurre efectivamente. Tomemos uno cualquiera de los x_i y apliquemos el endomorfismo $p_i(f)$ a la igualdad $0 = x_1 + \cdots + x_p$; el resultado es

$$0 = p_i(f)(x_1) + \cdots + p_i(f)(x_i) + \cdots + p_i(f)(x_p)$$

y, utilizando **(2)** y la igualdad precedente, se obtiene que

$$0 = p_i(f)(x_i)\,.$$

Utilizando ahora **(3)**,

$$x_i = q_1(f) \circ p_1(f)(x_i) + \cdots + q_i(f) \circ p_i(f)(x_i) + \cdots + q_p(f) \circ p_p(f)(x_i) = 0\,,$$

puesto que todos los términos son 0 a consecuencia de **(2)** y de la igualdad que acabamos de obtener. Por lo tanto, la suma es directa.

Ya hemos visto en (5.6.12) que los subespacios F_i son invariantes.

Las dimensiones de los F_i verifican

$$\dim F_1 + \cdots + \dim F_p = n\,;$$

además $\dim F_i \leq k_i$ para todo i, y, por hipótesis, $k_1 + \cdots + k_p = n$. En consecuencia debe ser

$$\dim F_i = k_i$$

para todo i.

5.6.15 El que la dimensión de $\operatorname{Ker}(f - \lambda_i\, id)^{k_i}$ coincida con el orden k_i de multiplicidad de λ_i significa que F_i es el mayor de todos los subespacios propios generalizados $\operatorname{Ker}(f - \lambda_i\, id)^r$. Por lo tanto no se puede aumentar el subespacio aumentando la potencia r por encima de la multiplicidad de λ_i.

Por otra parte, es posible que F_i se obtenga como $\operatorname{Ker}(f - \lambda_i\, id)^r$ para $r < k_i$; de hecho es algo que sucede con frecuencia.

5.6.16 Obtención de la matriz de Jordan de un endomorfismo (caso general).

Sea E un espacio sobre $\mathbb{K}$ ($\mathbb{R}$ o $\mathbb{C}$) de dimensión $n \neq 0$ y $f \in \mathcal{L}(E)$ un endomorfismo. Supongamos que el polinomio característico $p_f(X)$ posee n raíces en $\mathbb{K}$, iguales o distintas. Denotemos por $\lambda_1, \ldots, \lambda_p$ las raíces diferentes de $p_f(X)$ y por $k_1, \ldots, k_p$ sus multiplicidades. Hemos visto que los subespacios

$$F_i = \operatorname{Ker}(f - \lambda_i \, id_E)^{k_i}\,, \quad i = 1, \ldots, p\,,$$

son invariantes, de dimensión k_i, y $E = F_1 \oplus \cdots \oplus F_p$.

Eligiendo una base de E que sea reunión de bases de los subespacios F_i, obtenemos para f una matriz diagonal por bloques,

$$A = \begin{bmatrix} \boxed{A_1} & & \\ & \ddots & \\ & & \boxed{A_p} \end{bmatrix},$$

donde cada bloque A_i es la matriz del endomorfismo que f induce en F_i.

Ahora bien, en cada subespacio F_i podemos emplear el procedimiento de (5.6.8) para el endomorfismo $g = f - \lambda_i \, id$, que es nilpotente ya que $g^{k_i} = 0$ sobre F_i. Obtendremos así una base de F_i en la que la matriz de g será una matriz de Jordan con 0 en la diagonal, mientras que la matriz de f será igual pero con el valor λ_i en la diagonal.

Si son éstas las bases que se reúnen, cada uno de los bloques A_i es una matriz de Jordan con el valor λ_i en todos los lugares de la diagonal,

$$A = \left[\begin{array}{ccccc|cc} \lambda_i & 1 & & & & & \\ & \lambda_i & \ddots & & & & \\ & & \ddots & \ddots & & & \\ & & & \lambda_i & 1 & & \\ & & & & \lambda_i & & \\ \hline & & & & & \ddots & \\ & & & & & & \boxed{\lambda_i} \end{array}\right].$$

Por lo tanto, la matriz A será una matriz de Jordan con los valores $\lambda_1, \ldots, \lambda_p$ en la diagonal.

Hemos obtenido así el resultado que enunciamos seguidamente:

5.6.17 TEOREMA

---

a) Sea E un espacio sobre $\mathbb{K}$ ($\mathbb{R}$ o $\mathbb{C}$) de dimensión $n \neq 0$ y $f \in \mathcal{L}(E)$. Para que exista una base $(a_1, \ldots, a_n)$ de E en la que la matriz $[f, (a_i)]$ sea una matriz de Jordan es condición necesaria y suficiente que se verifique

(d1) $p_f(X)$ posee n raíces en $\mathbb{K}$, iguales o distintas,

propiedad que siempre se verifica cuando $\mathbb{K} = \mathbb{C}$.

b) Sea $A \in M_{\mathbb{K}}(n)$ ($\mathbb{K} = \mathbb{R}$ o $\mathbb{C}$). Para que A sea semejante a una matriz de Jordan es condición necesaria y suficiente que se verifique

(d1) $p_A(X)$ posee n raíces en $\mathbb{K}$, iguales o distintas,

propiedad que siempre se verifica cuando $\mathbb{K} = \mathbb{C}$. Toda matriz cuadrada compleja es semejante a una matriz de Jordan.

---

a) Ya hemos visto que la condición es suficiente, a la vez que explicábamos un método para buscar la base y la matriz de Jordan.

La condición es también necesaria, puesto que las matrices de Jordan son matrices triangulares.

b) El resultado para una matriz se obtiene considerando el endomorfismo de $\mathbb{K}^n$ dado por la matriz.

5.6.18 DEFINICIÓN. Si J es una matriz de Jordan que representa al endomorfismo f en una base, se dice que J es la *forma canónica de Jordan* de f.

Si J es una matriz de Jordan semejante a la matriz A, se dice que J es la *forma canónica de Jordan* de A.

Todas las matrices complejas (y, por lo tanto, todas las matrices reales) admiten una forma canónica. Ahora bien, para una matriz real es posible que su forma canónica no sea real, cuando las raíces de su polinomio característico no son todas reales; es ese caso, la matriz de paso tampoco será real.

Si un endomorfismo, o una matriz, es diagonalizable, entonces la correspondiente matriz diagonal es su forma canónica. El proceso de construcción que hemos visto conduce en ese caso a la matriz diagonal.

5.6.19 Ejemplo. El endomorfismo f de $\mathbb{R}^4$ dado por la matriz

$$A = \begin{bmatrix} 0 & -4. & 0 & -1. \\ 0 & 2. & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 4. & 8. & -12. & 4. \end{bmatrix}$$

tiene como polinomio característico

$$p_f(X) = X(X-2)^3\,,$$

con raíces 0 (de multiplicidad 1) y 2 (de multiplicidad 3). El rango de $A - 2\,I$ es 2, luego $\dim V(2) = 2$ y el endomorfismo no es diagonalizable. Pondremos

$$F_1 = \operatorname{Ker}(f - 2\,id)^3 \quad \text{y} \quad F_2 = \operatorname{Ker} f\,.$$

El endomorfismo $g = f - 2\,id$ viena dado por la matriz $B = A - 2\,I$,

$$B = \begin{bmatrix} -2. & -4. & 0 & -1. \\ 0 & 0 & 0 & 0 \\ 0 & 0 & -2. & 0 \\ 4. & 8. & -12. & 2. \end{bmatrix}.$$

Tenemos que

$$B^2 = \begin{bmatrix} 0 & 0 & 12. & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 4. & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \quad B^3 = \begin{bmatrix} 0 & 0 & -24. & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & -8. & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix};$$

por lo tanto, $F_1 = \operatorname{Ker} g^3 = \operatorname{Ker} g^2$ es el subespacio de dimensión 3 formado por los $(x, y, z, t) \in \mathbb{R}^4$ tales que $z = 0$. El endomorfismo g es nilpotente de orden 2 sobre F_1. Utilizando las notaciones de (5.6.8), el núcleo $K_1 = \operatorname{Ker} g$ tiene dimensión 2 y ecuación

$$\begin{aligned} 2x + 4y \quad + t &= 0 \\ z \quad &= 0\,. \end{aligned}$$

En consecuencia,

$$\begin{array}{ccc} n_0 = 0 & n_1 = 2 & n_2 = 3 \\ & d_1 = 2 & d_2 = 1 \end{array}$$

y la base de F_1 que buscamos tendrá la estructura

$$\begin{array}{cc} a_1 & \\ g(a_1) & a_2 \end{array}$$

y habrá que ordenarla como

$$(g(a_1), a_1, a_2)\,.$$

Como a_1 se puede tomar cualquier vector de F_1 que no esté en K_1 ; pondremos

$$a_1 = (0, 0, 0, 1)\,.$$

Entonces

$$g(a_1) = (-1, 0, 0, 2)\,.$$

Ahora basta elegir a_2 de manera que complete con $g(a_1)$ una base de K_1 ; tomaremos, por ejemplo,

$$a_2 = (2, -1, 0, 0)\,.$$

Esto termina el trabajo con F_1 ; la matriz del endomorfismo inducido por $g = f - 2\,id$ en F_1 , para la base $(g(a_1), a_1, a_2)$, es

$$\left[\begin{array}{cc|c} 0 & 1 & \\ & 0 & \\ \hline & & 0 \end{array}\right],$$

luego la matriz del endomorfismo indicido por f en F_1 , para esta misma base, es

$$\left[\begin{array}{cc|c} 2 & 1 & \\ & 2 & \\ \hline & & 2 \end{array}\right].$$

El subespacio $F_2 = \operatorname{Ker} f$ tiene dimensión 1 y ecuación

$$\begin{array}{rcl} t & = & 0 \\ y & = & 0 \\ x - 3z & = & 0\,. \end{array}$$

En este caso bastará tomar un vector cualquiera, no nulo, de F_2. Por ejemplo, el vector $(3, 0, 1, 0)$.

En la base de $\mathbb{R}^4$

$$(\,(-1, 0, 0, 2), (0, 0, 0, 1), (2, -1, 0, 0), (3, 0, 1, 0)\,)\,,$$

reunión de las bases que hemos elegido en F_1 y en F_2, la matriz de f es

$$\left[\begin{array}{cc|c|c} 2 & 1 & & \\ & 2 & & \\ \hline & & \boxed{2} & \\ \hline & & & 0 \end{array}\right],$$

como puede comprobar el lector.

5.6.20 Vamos a dar una alternativa al procedimiento que hemos desarrollado en el ejemplo precedente. No presenta ninguna diferencia de fondo con lo que ya hemos hecho; simplemente evita tener que trabajar constantemente con las ecuaciones de los subespacios $F_i = \operatorname{Ker}(f - \lambda_i\, id)^{k_i}$, lo que en ocasiones puede resultar molesto.

La variante consiste en reducir, en un primer paso, la matriz de f a la forma diagonal por bloques (con bloques cualesquiera). Para ello basta tomar una base del espacio que sea reunión de bases (cualesquiera) de los subespacios F_i. Se puede así obtener una matriz Q que transforma la matriz A de f en

$$Q^{-1} A\, Q = \begin{bmatrix} A_1 & & & \\ & \boxed{A_2} & & \\ & & \ddots & \\ & & & A_p \end{bmatrix},$$

sin que las matrices A_i sean de ninguna forma específica.

En una segunda fase, lo que se hace es reducir cada A_i a la forma canónica, buscando una matriz Q_i, inversible y del mismo tamaño que A_i, tal que $A_i' = Q_i^{-1} A_i\, Q_i$ sea una matriz de Jordan. Para ello basta seguir las indicaciones de (5.6.10), puesto que la matriz A_i posee un solo valor propio.

Poniendo entonces

$$P = Q \begin{bmatrix} Q_1 & & & \\ & \boxed{Q_2} & & \\ & & \ddots & \\ & & & Q_p \end{bmatrix},$$

se tiene que

$$P^{-1}A\,P = \begin{bmatrix} \boxed{A_1'} & & & \\ & \boxed{A_2'} & & \\ & & \ddots & \\ & & & \boxed{A_p'} \end{bmatrix},$$

como es fácil comprobar. La matriz obtenida ahora es la forma canónica de f y de A, y la matriz P proporciona la base de E en la que f se representa por dicha forma canónica.

5.6.21 Ejemplo. Veamos cómo se utilizan las recomendaciones del apartado precedente para el endomorfismo del ejemplo (5.6.19), que venía dado por la matriz

$$A = \begin{bmatrix} 0 & -4. & 0 & -1. \\ 0 & 2. & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 4. & 8. & -12. & 4. \end{bmatrix}.$$

El subespacio correspondiente al valor propio 2, $F_1 = \operatorname{Ker}(f - 2\,id)^3$, tiene por ecuación $z = 0$. El sistema

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

es una base de F_1.

El subespacio correspondiente al valor propio 0, $F_2 = \operatorname{Ker} f$, tiene por ecuación

$$\begin{aligned} t &= 0 \\ y \quad &= 0 \\ x - 3z \quad &= 0\,, \end{aligned}$$

y el vector $(3,0,1,0)$ forma una base de F_2.

Tomando entonces

$$Q = \begin{bmatrix} 1. & 0 & 0 & 3. \\ 0 & 1. & 0 & 0 \\ 0 & 0 & 0 & 1. \\ 0 & 0 & 1. & 0 \end{bmatrix}$$

se obtiene la matriz diagonal por bloques

$$Q^{-1}A\,Q = \left[\begin{array}{ccc|c} 0 & -4. & -1. & \\ 0 & 2. & 0 & \\ 4. & 8. & 4. & \\ \hline & & & 0 \end{array}\right].$$

Nos ocuparemos ahora de la matriz

$$A_1 = \begin{bmatrix} 0 & -4. & -1. \\ 0 & 2. & 0 \\ 4. & 8. & 4. \end{bmatrix}$$

que tiene el único valor propio 2 con multiplicidad 3. Consideremos el endomorfismo nilpotente g de $\mathbb{R}^3$ dado por la matriz

$$B_1 = A_1 - 2\,I = \begin{bmatrix} -2. & -4. & -1. \\ 0 & 0 & 0 \\ 4. & 8. & 2. \end{bmatrix}.$$

Como $(B_1)^2 = 0$, g es nilpotente de orden 2. Con las notaciones habituales, tenemos que $K_1 = \operatorname{Ker} g$ es el subespacio de $\mathbb{R}^3$ de dimensión 2 y ecuación

$$2x + 4y + z = 0\,.$$

Las dimensiones que interesan y sus diferencias son

$$\begin{array}{ccc} n_0 = 0 & n_1 = 2 & n_2 = 3 \\ & d_1 = 2 & d_2 = 1 \end{array}\ ,$$

y la base de R^3 que buscamos tendrá la estructura

$$\begin{array}{cc} a_1 & \\ g(a_1) & a_2 \end{array}$$

y habrá que ordenarla como

$$(g(a_1), a_1, a_2)\,.$$

Como a_1 se puede tomar cualquier vector de $\mathbb{R}^3$ que no esté en K_1 ; pondremos

$$a_1 = (1, 0, 0)\,.$$

Entonces

$$g(a_1) = (-2, 0, 4)\,.$$

Ahora, elegimos a_2 de manera que complete con $g(a_1)$ una base de K_1 ; por ejemplo,

$$a_2 = (2, -1, 0)\,.$$

Poniendo

$$Q_1 = \begin{bmatrix} -2. & 1. & 2. \\ 0 & 0 & -1. \\ 4. & 0 & 0 \end{bmatrix}$$

se obtiene que

$$Q_1^{-1} A_1 Q_1 = \left[\begin{array}{cc|c} 2 & 1 & 0 \\ 0 & 2 & 0 \\ \hline 0 & 0 & 2 \end{array}\right].$$

Pongamos finalmente

$$\begin{aligned} P &= Q \left[\begin{array}{c|c} Q_1 & \\ \hline & 1 \end{array}\right] \\ &= \begin{bmatrix} 1. & 0 & 0 & 3. \\ 0 & 1. & 0 & 0 \\ 0 & 0 & 0 & 1. \\ 0 & 0 & 1. & 0 \end{bmatrix} \left[\begin{array}{ccc|c} -2. & 1. & 2. & \\ 0 & 0 & -1. & \\ 4. & 0 & 0 & \\ \hline & & & 1. \end{array}\right] \\ &= \begin{bmatrix} -2. & 1. & 2. & 3. \\ 0 & 0 & -1. & 0 \\ 0 & 0 & 0 & 1. \\ 4. & 0 & 0 & 0 \end{bmatrix} . \end{aligned}$$

Con esta matriz se obtiene que

$$P^{-1} A\, P = \left[\begin{array}{cc|c|c} 2 & 1 & & \\ & 2 & & \\ \hline & & \boxed{2} & \\ \hline & & & 0 \end{array}\right] ,$$

es decir, se obtiene la forma canónica de f, como en el ejemplo (5.6.19). Nótese que existen diferentes bases de $\mathbb{R}^4$ en las que f tiene su forma canónica. Por ejemplo, ahora hemos obtenido la base

$$(\,(-2,0,0,4),(1,0,0,0),(2,-1,0,0),(3,0,1,0)\,)\,,$$

diferente de la que obtuvimos en (5.6.19). Esto no depende del método empleado, sino de la elección que se haga en cada caso de los vectores arbitrarios.

5.6.22 Ejemplo. La matriz compleja

$$A = \begin{bmatrix} 2\,i & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}$$

tiene como polinomio característico

$$p_A(X) = -(X-1)(X-i)^2\,,$$

con raíces 1 (simple) e i (doble). Denotaremos por f el endomorfismo de $\mathbb{C}^3$ dado por A y pondremos

$$F_1 = \operatorname{Ker}(f - id) \quad \text{y} \quad F_2 = \operatorname{Ker}(f - i\cdot id)^2\,.$$

El vector $(0,1,0)$ forma una base de F_1.

F_2 es el subespacio

$$F_2 = \{(x,y,z) \in \mathbb{C}^3 \mid y = 0\}\,,$$

mientras que $K_1 = \operatorname{Ker}(f - i \cdot id)$ tiene dimensión 1 (por eso A no es diagonalizable) y ecuación

$$\begin{aligned} ix \quad + z &= 0 \\ y \quad\quad &= 0\,. \end{aligned}$$

Denotemos por g el endomorfismo $g = f - i \cdot id$. La base de F_2 que interesa será de la forma $(g(a), a)$, donde a debe ser un vector de F_2 que no pertenezca a K_1. Tomaremos el vector $(1, 0, 0)$, cuya imagen por g es el vector $(i, 0, 1)$. En la base

$$(\,(0,1,0), (i,0,1), (1,0,0)\,)$$

de $\mathbb{C}^3$, f tendrá la forma canónica; o sea, para

$$P = \begin{bmatrix} 0 & i & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix},$$

la forma canónica de A se obtendrá como $P^{-1} A\, P$. El lector comprobará que, efectivamente,

$$P^{-1} A\, P = \left[\begin{array}{c|cc} 1 & & \\ \hline & i & 1 \\ & & i \end{array}\right].$$

5.6.23 Vamos a terminar esta sección estudiando la manera en que la forma canónica de Jordan sirve para el cálculo de polinomios de matrices (y, consecuentemente, de endomorfismos).

Sea A una matriz compleja $n \times n$, de la que denotaremos por J su forma de Jordan; sea $p(X)$ un polinomio con coeficientes complejos. Como $J = P^{-1} A\, P$ para una matriz inversible P, entonces

$$P(A) = P\, p(J)\, P^{-1}$$

(v. 5.5.2). Esto permite calcular $p(A)$ a través de $p(J)$.

La ventaja estriba en que, como J tiene una estructura sencilla, sus potencias son más fáciles de calcular, y lo mismo ocurre con los polinomios. Pero la sencillez de J no proviene sólo de que posee bastantes ceros, sino del tipo y disposición de sus elementos no nulos; esto permite realizar el cálculo de $p(J)$ aplicando una fórmula sencilla que vamos a obtener.

Si J es una matriz diagonal

$$J = \begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{bmatrix},$$

entonces

$$p(J) = \begin{bmatrix} p(\lambda_1) & & & \\ & p(\lambda_2) & & \\ & & \ddots & \\ & & & p(\lambda_n) \end{bmatrix},$$

como es fácil comprobar. (Para todas las comprobaciones de este tipo conviene empezar con potencias y pasar luego al caso de polinomios cualesquiera.)

Cuando J no es diagonal, la cosa es algo menos sencilla. Hay que tener en cuenta que

$$J = \begin{bmatrix} \boxed{J_1} & & & \\ & \boxed{J_2} & & \\ & & \ddots & \\ & & & \boxed{J_k} \end{bmatrix},$$

donde los bloques J_i son cajas elementales de Jordan. Por lo tanto (v. 5.5.3)

$$p(J) = \begin{bmatrix} \boxed{p(J_1)} & & & \\ & \boxed{p(J_2)} & & \\ & & \ddots & \\ & & & \boxed{p(J_k)} \end{bmatrix},$$

y la dificultad queda reducida a saber calcular $p(J)$ cuando J es una caja elemental.

5.6.24 Si

$$H = \begin{bmatrix} 0 & 1 & & & \\ & 0 & \ddots & & \\ & & \ddots & \ddots & \\ & & & 0 & 1 \\ & & & & 0 \end{bmatrix}$$

es una caja elemental $n \times n$ con la particularidad de tener ceros en la diagonal, es fácil ver que sus potencias vienen dadas por

$$H^k = \begin{bmatrix} 0 & \cdots & 1 & & & \\ & 0 & \cdots & 1 & & \\ & & \ddots & & \ddots & \\ & & & \ddots & & 1 \\ & & & & \ddots & \vdots \\ & & & & & 1 \end{bmatrix},$$

es decir, que su término general α_i^j es 1 cuando $i = j + k$ y 0 en los restantes casos. La idea para probar esto se puede ver en (5.6.2). En particular,

$$H^{n-1} = \begin{bmatrix} 0 & \cdots & \cdots & 1 \\ & \ddots & & \vdots \\ & & \ddots & \vdots \\ & & & 0 \end{bmatrix}$$

y $H^n = H^{n+1} = \cdots = 0$.

5.6.25 Sea ahora

$$J = \begin{bmatrix} \lambda & 1 & & & \\ & \lambda & \ddots & & \\ & & \ddots & \ddots & \\ & & & \lambda & 1 \\ & & & & \lambda \end{bmatrix}$$

una caja elemental cualquiera, de tamaño $n \times n$, y sea $p(X) = \alpha_0 + \alpha_1 X + \cdots + \alpha_r X^r$ un polinomio. La *fórmula de Taylor*[3] significa para este polinomio que

$$p(X) = p(\lambda) + \frac{p'(\lambda)}{1!}(X - \lambda) + \frac{p''(\lambda)}{2!}(X - \lambda)^2 + \cdots + \frac{p^{r)}(\lambda)}{r!}(X - \lambda)^r ,$$

donde $p', p'', \ldots, p^{r)}$ representan las derivadas sucesivas de p. Entonces

$$p(J) = p(\lambda)\, I + \frac{p'(\lambda)}{1!}(J - \lambda\, I) + \frac{p''(\lambda)}{2!}(J - \lambda\, I)^2 + \cdots + \frac{p^{r)}(\lambda)}{r!}(J - \lambda\, I)^r .$$

Basta ahora tener en cuenta cómo son las potencias de $H = J - \lambda\, I$, que han sido descritas en (5.6.24), para obtener la fórmula

$$p(J) = \begin{bmatrix} p(\lambda) & p'(\lambda) & \frac{p''(\lambda)}{2!} & \cdots & \cdots & \frac{p^{n-1)}(\lambda)}{(n-1)!} \\ & p(\lambda) & p'(\lambda) & \cdots & \cdots & \frac{p^{n-2)}(\lambda)}{(n-2)!} \\ & & \ddots & \ddots & & \vdots \\ & & & \ddots & \ddots & \vdots \\ & & & & \ddots & p'(\lambda) \\ & & & & & p(\lambda) \end{bmatrix} .$$

5.6.26 **Ejemplo.** Calculemos $I + A - A^8$ para la matriz

$$A = \begin{bmatrix} 0 & 1. & 0 & 0 \\ -1. & 1. & 1. & 0 \\ -1. & 0 & 2. & 0 \\ -1. & 0 & 1. & 1. \end{bmatrix}$$

del ejemplo (5.6.11). Vimos en aquel ejemplo que para

$$P = \begin{bmatrix} -1. & 1. & 0 & 0 \\ -1. & 0 & 1. & 0 \\ -1. & 0 & 0 & 0 \\ -1. & 0 & 0 & 1. \end{bmatrix}$$

[3] del matemático inglés Brook Taylor (1685-1731); la fórmula será sin duda conocida del lector en su versión con resto, pero le hacemos notar que, en el caso de un polinomio, se trata de una fórmula exacta siempre que se llegue en el desarrollo por lo menos hasta el grado del polinomio.

se tenía que

$$J = P^{-1}A\,P = \left[\begin{array}{ccc|c} 1 & 1 & & \\ & 1 & 1 & \\ & & 1 & \\ \hline & & & 1 \end{array}\right].$$

Pongamos $p(X) = 1 + X - X^8$,

$$J_1 = \begin{bmatrix} 1 & 1 & \\ & 1 & 1 \\ & & 1 \end{bmatrix} \quad \text{y} \quad J_2 = \begin{bmatrix} 1 \end{bmatrix}.$$

Los cálculos que hay que realizar son

$$p(J_1) = \begin{bmatrix} p(1) & p'(1) & p''(1)/2 \\ & p(1) & p'(1) \\ & & p(1) \end{bmatrix} = \begin{bmatrix} 1. & -7. & -28. \\ & 1. & -7. \\ & & 1. \end{bmatrix},$$

$$p(J_2) = \begin{bmatrix} p(1) \end{bmatrix} = \begin{bmatrix} 1. \end{bmatrix},$$

$$p(J) = \left[\begin{array}{c|c} p(J_1) & \\ \hline & p(J_2) \end{array}\right] = \left[\begin{array}{ccc|c} 1. & -7. & -28. & \\ & 1. & -7. & \\ & & 1. & \\ \hline & & & 1. \end{array}\right],$$

y, finalmente,

$$I + A - A^8 = p(A) = P\,p(J)\,P^{-1} = \begin{bmatrix} 8. & 21. & -28. & 0 \\ 7. & 29. & -35. & 0 \\ 7. & 28. & -34. & 0 \\ 7. & 28. & -35. & 1. \end{bmatrix}.$$

5.6.27 La fórmula de (5.6.25) que proporciona $p(J)$ cuando p es un polinomio y J una caja elemental $n \times n$, tiene una consecuencia inmediata: si J es

$$J = \begin{bmatrix} \lambda & 1 & & & \\ & \lambda & \ddots & & \\ & & \ddots & \ddots & \\ & & & \lambda & 1 \\ & & & & \lambda \end{bmatrix}$$

y $p(X)$ y $q(X)$ son dos polinomios tales que

$$p(\lambda) = q(\lambda), p'(\lambda) = q'(\lambda), \ldots, p^{\,n-1)}(\lambda) = q^{\,n-1)}(\lambda)\,,$$

entonces $p(J) = q(J)$.

Este resultado se puede extender a una matriz cualquiera. Sea A una matriz compleja $n \times n$, $\lambda_1, \ldots, \lambda_p$ los valores propios distintos de A,

$$p_A(X) = (-1)^n (X - \lambda_1)^{k_1} \cdots (X - \lambda_p)^{k_p}$$

su polinomio característico y $p(X)$ y $q(X)$ dos polinomios. Si coinciden los siguientes valores de los polinomios y sus derivadas

$$\begin{array}{l} p(\lambda_1) = q(\lambda_1), p'(\lambda_1) = q'(\lambda_1), \ldots, p^{\,k_1-1)}(\lambda_1) = q^{\,k_1-1)}(\lambda_1)\,, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ p(\lambda_p) = q(\lambda_p), p'(\lambda_p) = q'(\lambda_p), \ldots, p^{\,k_p-1)}(\lambda_p) = q^{\,k_p-1)}(\lambda_p)\,, \end{array}$$

entonces $p(A) = q(A)$.

Para probar este hecho se puede razonar sobre la forma canónica de A con ayuda del resultado para las cajas elementales.

Hay, sin embargo, una demostración directa muy breve, que es la que vamos a exponer. Denotemos por $r(X) = p(X) - q(X)$ el polinomio diferencia de ambos. Como para todo $i = 1, \ldots, p$

$$r(\lambda_i) = 0, r'(\lambda_i) = 0, \ldots, r^{\,k_i-1)}(\lambda_i) = 0\,,$$

resulta de la fórmula de Taylor (v. 5.6.25) en potencias de $X - \lambda_i$ que

$$r(X) = (X - \lambda_i)^{k_i} s_i(X)\,,$$

donde $s_i(X)$ es un polinomio; esto, para todo $i = 1, \ldots, p$. En consecuencia, $r(X)$ es divisible por $(X - \lambda_1)^{k_1}$, por $(X - \lambda_2)^{k_2}$, etc.; por lo tanto

$$r(X) = (X - \lambda_1)^{k_1} (X - \lambda_2)^{k_2} \cdots (X - \lambda_p)^{k_p} s(X)\,,$$

donde $s(X)$ es un polinomio. Como, por el teorema de Cayley-Hamilton, $p_A(A) = 0$, resulta que

$$r(A) = (-1)^n p_A(A) s(A) = 0\,.$$

Así pues

$$p(A) - q(A) = r(A) = 0$$

y se obtiene la igualdad.

Nótese que si A es una matriz $n \times n$ con valores propios $\lambda_1, \ldots, \lambda_n$ todos simples, la condición suficiente para la igualdad de $p(A)$ y $q(A)$ se reduce a que

$$p(\lambda_1) = q(\lambda_1), \ldots, p(\lambda_n) = q(\lambda_n)\,.$$

En el fondo, las ideas expuestas en este apartado son otra versión (más cómoda en su forma actual) de lo que ya vimos en (5.5.8).

5.6.28 Ejemplo. Para calcular cualquier potencia de la matriz

$$A = \begin{bmatrix} 5. & 4. \\ -2. & -1. \end{bmatrix},$$

cuyos valores propios son 3 y 1, podemos buscar para la potencia A^n un polinomio de la forma $\alpha_n A + \beta_n I$ que coincida con A^n. Es suficiente para ello que los polinomios

$$p(X) = X^n \quad \text{y} \quad q(X) = \alpha_n X + \beta_n$$

tomen los mismos valores en 3 y en 1. Resulta sencillo ver que esto equivale a que

$$\alpha_n = \frac{3^n - 1}{2} \quad \text{y} \quad \beta_n = \frac{-3^n + 3}{2}.$$

Por lo tanto,

$$\begin{aligned} A^n &= \frac{3^n - 1}{2} A + \frac{-3^n + 3}{2} I \\ &= \begin{bmatrix} 2 \cdot 3^n - 1 & 2 \cdot 3^n - 2 \\ -3^n + 1 & -3^n + 2 \end{bmatrix}. \end{aligned}$$

Compárese este procedimiento con el que desarrollamos en el ejemplo (5.5.9).

5.6.29 Sólo un breve comentario final sobre la definición y el cálculo de funciones $h(A)$ de una matriz $n \times n$, A, cuando h es una función diferente de un polinomio. El ejemplo más importante entre las funciones de una matriz es e^A (que es $h(A)$ para la función exponencial $h(z) = e^z$), a causa de su utilidad en la resolución de sistemas diferenciales lineales.

La generalización más obvia de los polinomios son las series de potencias. El lector sabe que una serie de potencias

$$\sum_{n=0}^{\infty} \alpha_n z^n$$

no converge generalmente para cualquier valor de $z \in \mathbb{C}$. Los mismos problemas surgen cuando se pretende definir la matriz suma de la serie

$$\sum_{n=0}^{\infty} \alpha_n A^n$$

como límite de los polinomios

$$p_r(A) = \sum_{n=0}^{r} \alpha_n A^n$$

cuando r tiende a infinito.

El lector podrá ver en cursos posteriores que, en el caso de la exponencial, no existe problema alguno para definir

$$e^A = \sum_{n=0}^{\infty} \frac{1}{n!} A^n ,$$

como no lo hay tampoco para la convergencia de la serie

$$e^z = \sum_{n=0}^{\infty} \frac{z^n}{n!}$$

para todo $z \in \mathbb{C}$. La función e^A, definida como acabamos de decir, verifica buena parte de las propiedades de la exponencial de números complejos. Por ejemplo, si $AB = BA$, se tiene que

$$e^{A+B} = e^A e^B ;$$

también

$$e^0 = I ;$$

asimismo, e^A es inversible y

$$(e^A)^{-1} = e^{-A} .$$

Una de las maneras de calcular la exponencial de una matriz (y también otras funciones) es utilizar la forma canónica de la matriz y las fórmulas de (5.6.23) y (5.6.25), que son también válidas cuando p es, en lugar de un polinomio, la función $p(z) = e^z$.

También es válido para la exponencial de una matriz (y para otras funciones) el comentario del apartado (5.6.27). Es decir, $e^A = q(A)$ para todo polinomio q cuyo valor, y el de las correspondientes derivadas, coincida con los de la función e^z, y sus derivadas (recuérdese que la exponencial es su propia derivada), en los valores propios de A.

El objetivo de estas consideracioneses únicamente introducir al lector en el 'ambiente' de ciertas ideas que tendrá que estudiar en cursos posteriores. Naturalmente, no hemos justificado las distintas afirmaciones, ya que ello nos llevaría a introducir y estudiar la convergencia de matrices, tema que cae fuera del propósito de este libro.

Capítulo 6

Formas bilineales y formas sesquilineales.

Nota. Recordemos que, en este capítulo, $\mathbb{K}$ representa siempre $\mathbb{R}$ o $\mathbb{C}$. Se evitan así ciertas anomalías que se presentan cuando el cuerpo es de característica 2 y se aprovechan algunas propiedades específicas de los números reales y complejos. El lector advertirá sin duda que casi todos los resultados de este capítulo son válidos cuando $\mathbb{K}$ es un cuerpo que no tiene característica 2, aunque no sea ninguno de los dos citados; excepciones a esta regla general son los resultados que se enuncian específicamente para $\mathbb{R}$ o para $\mathbb{C}$.

6.1 Formas bilineales sobre un espacio vectorial.

6.1.1 Sea E un e.v. sobre $\mathbb{K}$ ($\mathbb{R}$ o $\mathbb{C}$). Ya conocemos la noción de forma bilineal sobre E, introducida en el capítulo 3 en un contexto más general, el de las formas n-lineales. Recordemos que una aplicación

$$\begin{aligned} f : E \times E &\to \mathbb{K} \\ (x, y) &\to f(x, y)\,, \end{aligned}$$

que envía cada par ordenado (x, y) de vectores de E a un escalar $f(x, y)$, es una forma bilineal sobre E cuando verifica las cuatro propiedades siguientes (v. 3.1.6):

$$\begin{array}{ll} \textbf{(bln1)} & (\forall x, x', y \in E) \quad f(x + x', y) = f(x, y) + f(x', y)\,, \\ \textbf{(bln2)} & (\forall x, y, y' \in E) \quad f(x, y + y') = f(x, y) + f(x, y')\,, \\ \textbf{(bln3)} & (\forall \lambda \in \mathbb{K})(\forall x, y \in E) \quad f(\lambda x, y) = \lambda f(x, y)\,, \\ \textbf{(bln4)} & (\forall \lambda \in \mathbb{K})(\forall x, y \in E) \quad f(x, \lambda y) = \lambda f(x, y)\,. \end{array}$$

Vimos también que en el conjunto de las formas bilineales sobre E se introducen dos operaciones (v. 3.1.13) que le dan estructura de espacio vectorial sobre $\mathbb{K}$; este espacio vectorial se representa por $\mathcal{L}_2(E, \mathbb{K})$. Recordemos que dichas operaciones vienen dadas por

$$(f + g)(x, y) = f(x, y) + g(x, y)\,,$$

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

El cero de este espacio vectorial es la aplicación

$$0 : E \times E \to \mathbb{K}$$

que envía todo par $(x, y) \in E \times E$ al 0 de $\mathbb{K}$.

6.1.2 Ejemplo. Para el e.v. $\mathbb{K}^2$, la aplicación $f : \mathbb{K}^2 \times \mathbb{K}^2 \to \mathbb{K}$ dada por

$$f(\,(x^1, x^2), (y^1, y^2)\,) = x^1y^1 + x^2y^2$$

es una forma bilineal. Para el e.v. $\mathbb{K}^3$, la aplicación $f : \mathbb{K}^3 \times \mathbb{K}^3 \to \mathbb{K}$ dada por

$$f(\,(x^1, x^2, x^3), (y^1, y^2, y^3)\,) = x^1y^1 + x^2y^2 + x^3y^3$$

es también una forma bilineal. Más generalmente, para el e.v. $\mathbb{K}^n$, la aplicación $f : \mathbb{K}^n \times \mathbb{K}^n \to \mathbb{K}$ dada por

$$f(\,(x^1, \ldots, x^n), (y^1, \ldots, y^n)\,) = \sum_{i=1}^{n} x^iy^i = x^1y^1 + x^2y^2 + \cdots + x^ny^n$$

es una forma bilineal.

Es clásico, y nosotros lo hemos hecho varias veces en este libro, representar un vector genérico de $\mathbb{K}^2$ por $v = (x, y)$ o un vector genérico de $\mathbb{K}^3$ por $v = (x, y, z)$; con estas notaciones, las dos primeras fórmulas se convierten en

$$f(\,(x, y), (x', y')\,) = xx' + yy' \,,$$

$$f(\,(x, y, z), (x', y', z')\,) = xx' + yy' + zz' \,.$$

Utilizaremos indistintamente las dos notaciones, puesto que ambas son frecuentes en los diferentes textos.

Estas formas bilineales poseen una importancia distinta según que $\mathbb{K}$ sea $\mathbb{R}$ o sea $\mathbb{C}$. Si $\mathbb{K} = \mathbb{R}$, poseen gran interés, como veremos en el capítulo siguiente; en particular, en el caso de $\mathbb{R}^3$, el lector reconocerá la fórmula correspondiente al producto escalar ordinario de la física.

6.1.3 Ejemplo. Este ejemplo y el siguiente se sitúan en espacios que no son de dimensión finita y utilizan nociones de cálculo diferencial e integral que no han sido introducidas en este libro, pero que el lector conocerá sin duda, dado su carácter elemental. Bien es cierto que estos ejemplos no son imprescindibles para seguir el hilo teórico de la exposición; pero no es menos cierto que constituyen, junto con el (6.1.2), los ejemplos más interesantes de la teoría que vamos a explicar, lo que justifica su introducción.

Consideremos el espacio c_{00} formado por las sucesiones finitas de números reales (v. 1.2.7). Si $x = (x_n)_{n\in\mathbb{N}}$ e $y = (y_n)_{n\in\mathbb{N}}$ son dos sucesiones finitas de números reales, ponemos

$$f(x, y) = \sum_{n=1}^{\infty} x_ny_n \,;$$

debe tenerse en cuenta que la aparente serie (suma de infinitos términos) no posee más que una cantidad finita de términos no nulos, y es entonces la suma de un número finito de términos. Es fácil ver que f es una forma bilineal sobre c_{00}.

Planteado en c_{00}, este ejemplo no es muy interesante. Es más importante cuando se extiende a un espacio de sucesiones más grande. El ejemplo no se puede extender a todas las sucesiones de números reales. Vamos a generalizarlo para el mayor espacio de sucesiones al que se puede extender la fórmula que proporciona f.

Representemos por l^2 el subconjunto de $S_{\mathbb{R}}$ formado por las sucesiones $x = (x_n)_{n\in\mathbb{N}}$ de números reales tales que

$$\sum_{n=1}^{\infty}(x_n)^2 < \infty\,,$$

es decir, las sucesiones para las que la serie de términos positivos cuyo término general es $(x_n)^2$ sea convergente (sucesiones 'cuadrado-sumables'). Todas las sucesiones finitas verifican esta propiedad, luego $c_{00} \subset l^2$; pero existen también sucesiones no finitas que la cumplen.

Vamos a comenzar por demostrar que l^2 es un subespacio vectorial de $S_{\mathbb{R}}$. Sean $x = (x_n)_{n\in\mathbb{N}}$ e $y = (y_n)_{n\in\mathbb{N}}$ dos sucesiones de l^2 y pongamos

$$\alpha = \sum_{n=1}^{\infty}(x_n)^2 \quad \text{y} \quad \beta = \sum_{n=1}^{\infty}(y_n)^2\,.$$

Si λ y μ son dos números reales, es fácil comprobar que

$$(\lambda+\mu)^2 + (\lambda-\mu)^2 = 2(\lambda)^2 + 2(\mu)^2\,,$$

luego

$$(\lambda+\mu)^2 \le 2(\lambda)^2 + 2(\mu)^2\,.$$

Si $n \in \mathbb{N}$, se tiene entonces que

$$\begin{aligned}\sum_{k=1}^{n}(x_k+y_k)^2 &\le 2\sum_{k=1}^{n}(x_k)^2 + 2\sum_{k=1}^{n}(y_k)^2\\ &\le 2\sum_{k=1}^{\infty}(x_k)^2 + 2\sum_{k=1}^{\infty}(y_k)^2\\ &= 2(\alpha+\beta)\,.\end{aligned}$$

Como todas sus sumas parciales están acotadas por $2(\alpha+\beta)$, resulta que la serie

$$\sum_{k=1}^{\infty}(x_k+y_k)^2$$

es convergente, luego $x+y = (x_n+y_n)_{n\in\mathbb{N}} \in l^2$. Por otra parte, es sencillo ver que, si $x = (x_n)_{n\in\mathbb{N}} \in l^2$ y $\lambda \in \mathbb{R}$, entonces la sucesión $\lambda x = (\lambda x_n)_{n\in\mathbb{N}}$ pertenece a l^2. Así pues, l^2 es un subespacio de $S_{\mathbb{R}}$ y, por tanto, un espacio vectorial.

Vamos a ver a continuación que, si $x = (x_n)_{n\in\mathbb{N}}$ e $y = (y_n)_{n\in\mathbb{N}}$ son sucesiones de l^2, entonces la serie

$$\sum_{n=1}^{\infty} x_n y_n$$

es absolutamente convergente y, por tanto, convergente. Pongamos

$$\alpha = \sum_{n=1}^{\infty}(x_n)^2 \quad \text{y} \quad \beta = \sum_{n=1}^{\infty}(y_n)^2 .$$

Si λ y μ son dos números reales, entonces $(\lambda - \mu)^2 \geq 0$, luego

$$\lambda\,\mu \leq \frac{1}{2}(\lambda^2 + \mu^2) .$$

Para todo $k \in \mathbb{N}$, la desigualdad precedente para $\lambda = |x_k|$ y $\mu = |y_k|$ nos asegura que

$$|x_k y_k| = |x_k|\,|y_k| \leq \frac{1}{2}(|x_k|^2 + |y_k|^2) = \frac{1}{2}(\,(x_k)^2 + (y_k)^2) .$$

Si $n \in \mathbb{N}$, se tiene entonces que

$$\begin{aligned} \sum_{k=1}^{n} |x_k y_k| &\leq \frac{1}{2}\sum_{k=1}^{n}(x_k)^2 + \frac{1}{2}\sum_{k=1}^{n}(y_k)^2 \\ &\leq \frac{1}{2}\sum_{k=1}^{\infty}(x_k)^2 + \frac{1}{2}\sum_{k=1}^{\infty}(y_k)^2 \\ &= \frac{1}{2}(\alpha + \beta) . \end{aligned}$$

Esto prueba que la serie es absolutamente convergente.

Para todo $x, y \in l^2$, podemos poner entonces

$$f(x, y) = \sum_{n=1}^{\infty} x_n y_n ,$$

donde la serie es absolutamente convergente. Es fácil ver que f es una forma bilineal sobre l^2.

Un razonamiento sencillo prueba que la fórmula precedente no se puede extender a ningún subespacio F de $S_{\mathbb{R}}$ más grande que l^2. Si F es un subespacio de $S_{\mathbb{R}}$ para el que la fórmula que proporciona f tiene sentido y $x = (x_n)_{n\in\mathbb{N}} \in F$, entonces la serie

$$f(x, x) = \sum_{n=1}^{\infty}(x_n)^2$$

debe ser convergente, luego $x \in l^2$; o sea, $F \subset l^2$.

6.1.4 Ejemplo. Ya vimos en (1.1.3) que el conjunto de las aplicaciones de un conjunto X en $\mathbb{R}$ posee, con las operaciones allí definidas, una estructura de espacio vectorial sobre $\mathbb{R}$. Vamos a considerar ahora un intervalo $[a, b]$ de $\mathbb{R}$, con $a < b$, como conjunto X y el correspondiente espacio vectorial $\mathcal{F}([a, b], \mathbb{R})$. Los elementos de $\mathcal{F}([a, b], \mathbb{R})$ son, pues, las funciones

$$\begin{aligned} f : [a, b] &\to \mathbb{R} \\ t &\to f(t)\,, \end{aligned}$$

y las operaciones que se consideran están dadas por

$$\begin{aligned} (f + g)(t) &= f(t) + g(t)\,, \\ (\lambda f)(t) &= \lambda f(t)\,. \end{aligned}$$

El subconjunto de $\mathcal{F}([a, b], \mathbb{R})$ formado por las funciones que son continuas en $[a, b]$ es un subespacio vectorial; representamos este subespacio por $\mathcal{C}([a, b], \mathbb{R})$. Se trata de un espacio que no es de dimensión finita; de hecho contiene como subespacio al conjunto $\mathbb{R}[X]$ de las funciones polinómicas que no es de dimensión finita.

Si $f, g \in \mathcal{C}([a, b], \mathbb{R})$, entonces su producto

$$(fg)(t) = f(t)g(t)$$

es también una función continua, luego existe la integral de Riemann[1]

$$\int_a^b f(t)g(t)\, dt\,.$$

Consideremos la aplicación

$$\begin{aligned} P : \mathcal{C}([a, b], \mathbb{R}) \times \mathcal{C}([a, b], \mathbb{R}) &\to \mathbb{R} \\ (f, g) &\to P(f, g) \end{aligned}$$

dada por

$$P(f, g) = \int_a^b f(t)g(t)\, dt\,;$$

esta aplicación es una forma bilineal sobre $\mathcal{C}([a, b], \mathbb{R})$, como se demuestra utilizando las propiedades elementales de la integral de Riemann. Volveremos con frecuencia sobre este ejemplo.

La forma bilineal $P(f, g) = \int_a^b f(t)g(t)\, dt$ se puede extender a espacios más amplios que $\mathcal{C}([a, b], \mathbb{R})$, auunque no al espacio $\mathcal{F}([a, b], \mathbb{R})$ de todas las funciones, como es prácticamente evidente. Por ejemplo, se puede extender al subespacio de $\mathcal{F}([a, b], \mathbb{R})$ formado por las funciones que son integrables en $[a, b]$ en el sentido de Riemann. Sin embargo no es posible a este nivel describir el mayor subespacio de $\mathcal{F}([a, b], \mathbb{R})$ al que se puede extender la forma bilineal que hemos visto en este ejemplo.

[1] del matemático alemán Bernhard Riemann (1826-1866).

6.1.5 Si f es una forma bilineal sobre el e.v. E, entonces

$$(\forall y \in E) \quad f(0, y) = f(y, 0) = 0 \, ,$$

como se prueba sin dificultad.

6.1.6 Sea f una forma bilineal sobre el e.v. E. Recordemos que f es simétrica cuando (v. 3.1.17)

(sim) $(\forall x, y \in E) \quad f(y, x) = f(x, y) \, ,$

y que f es antisimétrica cuando

(ant) $(\forall x, y \in E) \quad f(y, x) = -f(x, y) \, .$

Como $\mathbb{K}$ (que es $\mathbb{R}$ o $\mathbb{C}$) no es de característica 2, son válidos los dos apartados de (3.1.21) y resulta que una forma bilineal f es antisimétrica si y sólo si es alternada, es decir, si verifica

(alt) $(\forall x \in E) \quad f(x, x) = 0 \, .$

Aunque la demostración de este hecho la hicimos en (3.1.21) para el caso más general de las formas n-lineales, vamos a repetirla en este caso, para el que resulta más sencilla.

Supongamos que f es alternada; si $x, y \in E$, resulta entonces que

$$0 = f(x + y, x + y) = f(x, x) + f(x, y) + f(y, x) + f(y, y) = f(x, y) + f(y, x) \, ,$$

luego

$$f(y, x) = -f(x, y)$$

y f es antisimétrica.

Recíprocamente, supongamos que f es antisimétrica, y sea $x \in E$; entonces

$$f(x, x) = -f(x, x)$$

luego

$$2f(x, x) = f(x, x) + f(x, x) = 0$$

y

$$f(x, x) = 0 \, ,$$

lo que demuestra que f es alternada.

6.1.7 **Ejemplo.** Ya vimos en (3.1.18), (3.1.19) y (3.1.20) ejemplos de formas simétricas y antisimétricas, y de formas que no eran ni simétricas ni antisimétricas. Las formas de los ejemplos (6.1.2), (6.1.3) y (6.1.4) son simétricas.

6.1.8 Los subconjuntos $\mathcal{S}_2(E, \mathbb{K})$ y $\mathcal{A}_2(E, \mathbb{K})$ de $\mathcal{L}_2(E, \mathbb{K})$, constituidos respectivamente por las formas simétricas y por las formas antisimétricas o alternadas, son subespacios vectoriales de $\mathcal{L}_2(E, \mathbb{K})$.

6.1.9 Sea f una forma bilineal sobre el e.v. E. Si F es un subespacio vectorial de E, podemos considerar la restricción

$$g : F \times F \to \mathbb{K}$$

de f a $F \times F$, es decir, la aplicación dada por

$$g(x, y) = f(x, y)$$

para $x, y \in F$. La aplicación g es una forma bilineal sobre F. Además, si f es simétrica, g es también simétrica, y si f es antisimétrica, g lo es también.

6.1.10 Como ocurría con las aplicaciones lineales (v. 2.5.5), una forma bilineal sobre un e.v. de dimensión finita se puede describir por una matriz.

Sea E un e.v. de dimensión $n \neq 0$ y $(a_1, \ldots, a_n)$ una base de E; sea f una forma bilineal sobre E. Si x e y son vectores de E, con coordenadas

$$x = (x^1, \ldots, x^n)_{(a_i)} \quad \text{e} \quad y = (y^1, \ldots, y^n)_{(a_i)} ,$$

entonces

$$f(x, y) = f(x^1 a_1 + \cdots + x^n a_n, y^1 a_1 + \cdots + y^n a_n) = \sum_{i,j=1}^{n} x^i y^j f(a_i, a_j) ,$$

donde recurrimos al símbolo $\sum$, cuyo significado explicamos en (1.1.7), como forma breve de escribir la suma de los n^2 términos que resultan. Observamos entonces que f queda completamente determinada por los valores que toma en los pares (a_i, a_j) de vectores de la base.

Esto nos lleva a una proposición completamente análoga a (2.1.6).

6.1.11 PROPOSICIÓN

Sea E un e.v. de dimensión $n \neq 0$ sobre $\mathbb{K}$ y $(a_1, \ldots, a_n)$ una base de E. Para $i, j \in \{1, \ldots, n\}$, consideramos escalares $\alpha_i^j \in \mathbb{K}$. Existe una única forma bilineal $f \in \mathcal{L}_2(E, \mathbb{K})$ que verifica

$$(\forall i, j) \quad f(a_i, a_j) = \alpha_i^j .$$

Dicha forma toma en el par (x, y) de vectores de E de coordenadas

$$x = (x^1, \ldots, x^n)_{(a_i)} \quad \text{e} \quad y = (y^1, \ldots, y^n)_{(a_i)}$$

el valor

$$f(x, y) = \sum_{i,j=1}^{n} \alpha_i^j x^i y^j .$$

En efecto; si definimos f por

$$f(x,y) = \sum_{i,j=1}^{n} \alpha_i^j x^i y^j \,,$$

se prueba sin dificultad que f es una forma bilineal sobre E y que

$$(\forall i,j) \quad f(a_i, a_j) = \alpha_i^j$$

(considérense las coordenadas de a_i y a_j en la base $(a_1, \ldots, a_n)$); o sea, f verifica las condiciones requeridas. Si $g \in \mathcal{L}_2(E, \mathbb{K})$ las verifica también, entonces, para $x, y \in E$,

$$g(x,y) = \sum_{i,j=1}^{n} x^i y^j g(a_i, a_j) = \sum_{i,j=1}^{n} x^i y^j \alpha_i^j = f(x,y) \,,$$

luego $g = f$, y esto demuestra la unicidad.

6.1.12 DEFINICIÓN. Sea E un e.v. de dimensión $n \neq 0$ sobre $\mathbb{K}$ y $(a_1, \ldots, a_n)$ una base de E; sea f una forma bilineal sobre E. Llamamos *matriz asociada a la forma f en la base* $(a_1, \ldots, a_n)$ a la matriz

$$[\alpha_i^j] \in M(n) \,,$$

cuyo elemento general es

$$\alpha_i^j = f(a_i, a_j) \,.$$

Representamos esta matriz por

$$[f, (a_i)] \,.$$

Obsérvese que, aunque empleamos el mismo símbolo en (2.2.17) que aquí, su significado es diferente según que f sea un endomorfismo de E o una forma bilineal sobre E.

6.1.13 Ejemplo. Para las formas bilineales sobre $\mathbb{K}^2$ y $\mathbb{K}^3$ del ejemplo (6.1.2) y las bases canónicas de $\mathbb{K}^2$ y $\mathbb{K}^3$ se tiene

$$[f, (e_1, e_2)] = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I_2 \,;$$

$$[f, (e_1, e_2, e_3)] = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = I_3 \,.$$

6.1.14 Ejemplo. Para la forma f sobre $\mathbb{K}^3$ dada por

$$f(\,(x^1,x^2,x^3),(y^1,y^2,y^3)\,) = 2x^1y^1 + x^3y^1 + 2x^2y^3 + x^3y^3$$

y la base canónica (e_1,e_2,e_3), se puede comprobar que

$$\begin{array}{lll} f(e_1,e_1)=2 & f(e_2,e_1)=0 & f(e_3,e_1)=1 \\ f(e_1,e_2)=0 & f(e_2,e_2)=0 & f(e_3,e_2)=0 \\ f(e_1,e_3)=0 & f(e_2,e_3)=2 & f(e_3,e_3)=1\ . \end{array}$$

Por lo tanto

$$[f,(e_1,e_2,e_3)] = \begin{bmatrix} 2. & 0 & 1. \\ 0 & 0 & 0 \\ 0 & 2. & 1. \end{bmatrix}.$$

6.1.15 PROPOSICIÓN

Sea E un e.v. de dimensión $n \neq 0$ sobre $\mathbb{K}$ y $(a_1,\ldots,a_n)$ una base de E. La aplicación

$$\begin{aligned} \mathcal{L}_2(E,\mathbb{K}) &\to M_{\mathbb{K}}(n) \\ f &\to [f,(a_i)] \end{aligned}$$

es un isomorfismo de espacios vectoriales.

La proposición (6.1.11) garantiza justamente que esta aplicación es biyectiva. La linealidad de la aplicación es, por otra parte, muy fácil de probar.

6.1.16 Resulta entonces que, si E es de dimensión $n \neq 0$, entonces $\mathcal{L}_2(E,\mathbb{K})$ es de dimensión n^2, como $M_{\mathbb{K}}(n)$ (v. 2.3.5).

6.1.17 Expresión matricial de una forma bilineal. Sea E un e.v. de dimensión $n \neq 0$ sobre $\mathbb{K}$, $(a_1,\ldots,a_n)$ una base de E, $f \in \mathcal{L}_2(E,\mathbb{K})$ una forma bilineal sobre E y

$$A = [\alpha_i^j] = [f,(a_i)]\,.$$

Hemos visto que

$$f(x,y) = \sum_{i,j=1}^{n} \alpha_i^j x^i y^j$$

para los vectores x e y de coordenadas

$$x = (x^1,\ldots,x^n)_{(a_i)} \quad \text{e} \quad y = (y^1,\ldots,y^n)_{(a_i)}\,.$$

Resulta así que

$$f(x,y) = \begin{bmatrix} y^1 & \cdots & y^n \end{bmatrix} \begin{bmatrix} \alpha_1^1 & \cdots & \alpha_n^1 \\ \vdots & & \vdots \\ \alpha_1^n & \cdots & \alpha_n^n \end{bmatrix} \begin{bmatrix} x^1 \\ \vdots \\ x^n \end{bmatrix},$$

puesto que el segundo miembro es la matriz 1×1 cuyo único elemento es

$$\sum_{i,j=1}^{n} \alpha_i^j x^i y^j .$$

Tenemos, pues, que

$$\textbf{(1)} \qquad \boxed{f(x,y) = [y,(a_i)]^t\,[f,(a_i)]\,[x,(a_i)]\,.}$$

Además, la matriz $[f,(a_i)]$ es la única que hace cierta la fórmula precedente; si $B = [\beta_i^j] \in M_{\mathbb{K}}(n)$ y si

$$(\forall x, y \in E) \quad f(x,y) = [y,(a_i)]^t\, B\, [x,(a_i)]\,,$$

entonces

$$(\forall i,j) \quad f(a_i,a_j) = [0 \ \cdots \overset{\overset{j}{\downarrow}}{1} \cdots \ 0]\ B \begin{bmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{bmatrix} \leftarrow i = \beta_i^j\,,$$

luego $B = [f,(a_i)]$.

Cuando $E = \mathbb{K}^n$ y la base fijada es la base canónica de $\mathbb{K}^n$, resulta que

$$\textbf{(2)} \qquad \boxed{f(x,y) = y^t\, A\, x\,,}$$

con la identificación habitual entre vectores de $\mathbb{K}^n$ y matrices columna (v. 2.2.13). La forma bilineal de $\mathbb{K}^n$ que se representa en la base canónica por una matriz A (y que, por lo tanto, viene dada por la fórmula **(2)**), recibe corrientemente el nombre de 'forma bilineal dada por A'.

6.1.18 Ejemplo. Para la forma sobre $\mathbb{K}^2$ de (6.1.2), la expresión **(2)** se reduce a

$$f(x,y) = y^t\, x = \begin{bmatrix} y^1 & y^2 \end{bmatrix} \begin{bmatrix} x^1 \\ x^2 \end{bmatrix},$$

y lo mismo en el caso de $\mathbb{K}^3$ o de $\mathbb{K}^n$.

6.1.19 Ejemplo. Para la forma f de (6.1.14), tenemos

$$f(x,y) = y^t\,[f,(e_1,e_2,e_3)]\,x = \begin{bmatrix} y^1 & y^2 & y^3 \end{bmatrix} \begin{bmatrix} 2. & 0 & 1. \\ 0 & 0 & 0 \\ 0 & 2. & 1. \end{bmatrix} \begin{bmatrix} x^1 \\ x^2 \\ x^3 \end{bmatrix}.$$

Compruébese, realizando el producto de matrices, que se obtiene exactamente la expresión que define f.

6.1.20 Cambio de base. Sea E un e.v. de dimensión $n \neq 0$ sobre $\mathbb{K}$, $(a_1,\ldots,a_n)$ y $(a'_1,\ldots,a'_n)$ dos bases de E y $f \in \mathcal{L}_2(E,\mathbb{K})$ una forma bilineal sobre E. Pongamos

$$A = [f,(a_i)] \quad \text{y} \quad A' = [f,(a'_i)]\,.$$

Si $x, y \in E$, tenemos

$$f(x,y) = [y,(a'_i)]^t\, A'\,[x,(a'_i)]$$

y también

$$f(x,y) = [y,(a_i)]^t\, A\,[x,(a_i)]\,.$$

Transformemos esta última igualdad, teniendo en cuenta que (v. 2.5.14)

$$[x,(a_i)] = P\,[x,(a'_i)] \quad \text{y} \quad [y,(a_i)] = P\,[y,(a'_i)]\,,$$

donde $P = [(a'_i),(a_i)]$ es la matriz de paso de $(a_1,\ldots,a_n)$ a $(a'_1,\ldots,a'_n)$; se obtiene

$$f(x,y) = (P\,[y,(a'_i)])^t\, A\,P\,[x,(a'_i)] = [y,(a'_i)]^t\, P^{\,t}\, A\,P\,[x,(a'_i)]\,.$$

El razonamiento de (6.1.17) acerca de la unicidad de la matriz permite entonces concluir que

$$\textbf{(3)} \qquad \boxed{A' = P^{\,t}\, A\, P\,.}$$

Obsérvese las semejanzas y diferencias de esta fórmula con la de (2.5.17).

Como P y $P^{\,t}$ son inversibles, resulta que dos matrices A y A' que representan a una forma bilineal (en bases diferentes) son equivalentes (v. 2.5.18) y, por lo tanto, poseen el mismo rango (v. 2.5.24). Sin embargo, A y A' no son en general semejantes (v. 2.5.25), puesto que $P^{\,t}$ no coincide generalmente con la inversa de P.

De dos matrices $A, A' \in M_{\mathbb{K}}(n)$ tales que

$$A' = P^{\,t}\, A\, P$$

para alguna matriz inversible P, se dice que son *congruentes*. Un razonamiento como el de (2.5.24) nos permite asegurar que dos matrices son congruentes si y sólo si representan (en bases distintas) una misma forma bilineal.

6.1.21 Sea E un e.v. de dimensión $n \neq 0$ sobre $\mathbb{K}$, $(a_1, \ldots, a_n)$ una base de E y $f \in \mathcal{L}_2(E, \mathbb{K})$. Sea

$$A = [\alpha_i^j] = [f, (a_i)] .$$

Si la forma f es simétrica, entonces

$$\alpha_j^i = f(a_j, a_i) = f(a_i, a_j) = \alpha_i^j ,$$

luego la matriz A es simétrica. Recíprocamente, si A es simétrica, teniendo en cuenta que toda matriz 1×1 coincide con su traspuesta, obtenemos que

$$\begin{aligned} f(x, y) &= [x, (a_i)]^t A [y, (a_i)] = ([x, (a_i)]^t A [y, (a_i)])^t \\ &= [y, (a_i)]^t A^t [x, (a_i)] = [y, (a_i)]^t A [x, (a_i)] = f(x, y) , \end{aligned}$$

luego f es simétrica. Es decir, f es simétrica si y sólo si $[f, (a_i)]$ es simétrica.

Lo mismo ocurre con las formas antisimétricas. Si f es antisimétrica, entonces

$$(\forall i, j) \quad \alpha_j^i = -\alpha_i^j ,$$

o sea,

$$A^t = -A .$$

Las matrices que verifican esta propiedad se llaman *antisimétricas*. Resulta, como en el caso precedente, que f es antisimétrica si y sólo si $[f, (a_i)]$ es antisimétrica.

6.2 Núcleo y rango de una forma bilineal.

6.2.1 En esta sección y en las que siguen vamos a considerar al mismo tiempo aplicaciones lineales y formas bilineales. Con objeto de facilitar la lectura de las distintas fórmulas y expresiones, procuraremos representar por f, g, h, etc. las formas bilineales, y por u, v, w, etc. las aplicaciones lineales.

6.2.2 El hecho de que (en el caso de dimensión finita) una forma bilineal f se represente por una matriz, hace pensar en asociar a f una aplicación lineal dada por la misma matriz. Como veremos a continuación, esto se puede hacer de manera independiente de la dimensión y de las bases.

Sea $f \in \mathcal{L}_2(E, \mathbb{K})$,

$$\begin{aligned} f : E \times E &\to \mathbb{K} \\ (x, y) &\to f(x, y) , \end{aligned}$$

una forma bilineal. Si fijamos un vector $x \in E$, obtenemos una aplicación,

$$\begin{aligned} E &\to \mathbb{K} \\ y &\to f(x, y) , \end{aligned}$$

que, gracias a las propiedades **(bln2)** y **(bln4)**, es una aplicación lineal, o sea, un elemento del dual E^*; se trata del procedimiento que ya mencionamos en (3.1.11).

Vamos a denotar este elemento de E^* por $u_f(x)$; diremos que $u_f(x)$ se obtiene fijando en f la variable de la izquierda. Con la notación del corchete de dualidad (v. 2.6.6), se tiene que

$$\textbf{(4)} \qquad \boxed{(\forall x, y \in E) \quad \langle y, u_f(x) \rangle = f(x, y)\,.}$$

Consideremos ahora la aplicación

$$\begin{aligned} u_f \,:\, E &\to E^* \\ x &\to u_f(x) \end{aligned}$$

que lleva cada vector $x \in E$ al elemento $u_f(x) \in E^*$ así obtenido. Resulta, gracias ahora a las propiedades **(bln1)** y **(bln3)**, que u_f es una aplicación lineal, o sea, que $u_f \in \mathcal{L}(E, E^*)$.

El mismo proceso se puede seguir fijando en f la variable de la derecha. Si $x \in E$, representamos por $v_f(x)$ el elemento de E^* dado por

$$\textbf{(5)} \qquad \boxed{(\forall x, y \in E) \quad \langle y, v_f(x) \rangle = f(y, x)}$$

y consideramos también la aplicación lineal

$$\begin{aligned} v_f \,:\, E &\to E^* \\ x &\to v_f(x)\,. \end{aligned}$$

Cuando f es simétrica, entonces $f(y, x) = f(x, y)$, luego $u_f = v_f$, como se deduce de la simple observación de las fórmulas **(4)** y **(5)**.

6.2.3 Ejemplo. Para la forma f sobre $\mathbb{K}^2$ de (6.1.2) y el vector $x = (2, 1)$, tenemos

$$\langle (y^1, y^2), u_f(x) \rangle = 2y^1 + y^2$$

y

$$\langle (y^1, y^2), v_f(x) \rangle = 2y^1 + y^2$$

(nótese que f es simétrica), es decir, $u_f(x)$ y $v_f(x)$ son el elemento de $(\mathbb{K}^2)^*$ que representamos por $(2, 1)^*$, con notación de (2.6.19).

6.2.4 Ejemplo. Para la forma f sobre $\mathbb{K}^3$ del ejemplo (6.1.14), dada por

$$f(\,(x^1, x^2, x^3), (y^1, y^2, y^3)\,) = 2x^1y^1 + x^3y^1 + 2x^2y^3 + x^3y^3\,,$$

y para el vector $x = (1, 1, 1)$, tenemos

$$\langle (y^1, y^2, y^3), u_f(x) \rangle = 3y^1 + 3y^3\,,$$

$$\langle (y^1, y^2, y^3), v_f(x) \rangle = 2y^1 + 2y^2 + 2y^3\,.$$

Con notación de (2.6.19), resulta así que

$$u_f(x) = (3, 0, 3)^* \quad \text{y} \quad v_f(x) = (2, 2, 2)^*\,,$$

6.2.5 Ejemplo. Para la forma bilineal P sobre $\mathcal{C}([a,b],\mathbb{R})$ de (6.1.4), tenemos que, si $f,g \in \mathcal{C}([a,b],\mathbb{R})$,

$$\langle g, u_P(f)\rangle = \int_a^b f(t)g(t)\,dt\,,$$

$$\langle g, v_P(f)\rangle = \int_a^b f(t)g(t)\,dt\,.$$

Así, si f representa la función constante 1, las correspondientes formas lineales $u_P(f)$ y $v_P(f)$ envían cada función g a $\int_a^b g(t)\,dt$.

6.2.6 Vamos a definir lo que llamaremos 'núcleo' de una forma bilineal. Recordando la definición de núcleo de una aplicación lineal, que dimos en (2.1.10), resulta tentadora la idea de considerar el conjunto

$$\{x \in E \mid f(x,x) = 0\}$$

como núcleo de la forma bilineal $f \in \mathcal{L}_2(E,\mathbb{K})$. El principal obstáculo para aceptar esta definición es que dicho conjunto no es, en general, un subespacio vectorial de E; así ocurre, en efecto, en el ejemplo que sigue. Sin embargo, este conjunto desempeña un papel importante en el estudio de la forma bilineal f, y coincide en ciertos casos (v. 6.5.9) con el núcleo de f.

6.2.7 Ejemplo. Para la forma bilineal sobre $\mathbb{R}^3$ dada por

$$f(\,(x^1,x^2,x^3),(y^1,y^2,y^3)\,) = x^1y^1 - x^2y^2$$

se tiene

$$\begin{aligned}
&\{(x^1,x^2,x^3) \in \mathbb{R}^3 \mid f(\,(x^1,x^2,x^3),(x^1,x^2,x^3)\,) = 0\} \\
&= \{(x^1,x^2,x^3) \mid (x^1)^2 - (x^2)^2 = 0\} \\
&= \{(x^1,x^2,x^3) \mid x^1 = x^2\} \cup \{(x^1,x^2,x^3) \mid x^1 = -x^2\}\,,
\end{aligned}$$

conjunto que es reunión de dos subespacios, pero que no es un subespacio.

6.2.8 Lo que haremos entonces es considerar como núcleo de f el de u_f o el de v_f, que son, ambos, subespacios de E. Estos núcleos son

$$\begin{aligned}
\operatorname{Ker} u_f &= \{x \in E \mid u_f(x) = 0\} = \{x \in E \mid (\forall y) \quad \langle y, u_f(x)\rangle = 0\} \\
&= \{x \in E \mid (\forall y) \quad f(x,y) = 0\}
\end{aligned}$$

y, análogamente,

$$\operatorname{Ker} v_f = \{x \in E \mid (\forall y) \quad f(y,x) = 0\}\,.$$

Tenemos así dos núcleos para f, en general diferentes. $\operatorname{Ker} u_f$ se suele denominar 'núcleo a la izquierda de f' y $\operatorname{Ker} v_f$, 'núcleo a la derecha de f'. La ambiguedad desaparece si f es simétrica, ya que entonces $u_f = v_f$; en ese caso obtenemos una definición única para el núcleo.

6.2.9 DEFINICIÓN. Sea $f \in \mathcal{S}_2(E, \mathbb{K})$ una forma bilineal y simétrica sobre E. El subespacio $\operatorname{Ker} u_f = \operatorname{Ker} v_f$ de E recibe el nombre de *núcleo* de la forma simétrica f y se representa por $\operatorname{Ker} f$. Entonces

$$\operatorname{Ker} f = \{x \in E \mid (\forall y) \quad f(x, y) = 0\} = \{x \in E \mid (\forall y) \quad f(y, x) = 0\} .$$

Decimos que la forma simétrica f es *degenerada* si $\operatorname{Ker} f \neq \{0\}$, y decimos que f es *no degenerada* si $\operatorname{Ker} f = \{0\}$.

6.2.10 PROPOSICIÓN

Sea $f \in \mathcal{S}_2(E, \mathbb{K})$ una forma bilineal y simétrica sobre E; las siguientes propiedades son equivalentes

(i) f es no degenerada;

(ii) para todo $x \in E$, $x \neq 0$, existe $y \in E$ tal que $f(x, y) \neq 0$;

(iii) para todo $x \in E$, $x \neq 0$, existe $y \in E$ tal que $f(y, x) \neq 0$;

(iv) la aplicación lineal $u_f = v_f$ de E en E^* es inyectiva.

La equivalencia se deduce sin dificultad de las definiciones precedentes.

6.2.11 En el caso en que E sea de dimensión finita, podemos definir también el 'rango' de una forma bilineal. Vamos a ver en primer lugar que la matriz de f coincide con la de u_f y es la traspuesta de la matriz de v_f.

6.2.12 PROPOSICIÓN

Sea E un e.v. de dimensión $n \neq 0$ sobre $\mathbb{K}$, $(a_1, \ldots, a_n)$ una base de E y $(a^{*1}, \ldots, a^{*n})$ la base de E^* dual de la anterior (v. 2.6.13). Si $f \in \mathcal{L}_2(E, \mathbb{K})$ es una forma bilineal y $u_f, v_f \in \mathcal{L}(E, E^*)$ son las aplicaciones lineales asociadas, se tiene

$$[u_f, (a_i), (a^{*i})] = [f, (a_i)] ,$$

$$[v_f, (a_i), (a^{*i})] = [f, (a_i)]^t .$$

Pongamos, por ejemplo,

$$[u_f, (a_i), (a^{*i})] = [\beta_i^j] \quad \text{y} \quad [f, (a_i)] = [\alpha_i^j] ;$$

entonces

$$\alpha_i^j = f(a_i, a_j) = \langle a_j, u_f(a_i) \rangle = \beta_i^j ,$$

puesto que el escalar que figura en el corchete es la coordenada número j de $u_f(a_i)$ en la base $(a^{*1}, \ldots, a^{*n})$ (v. 2.6.16); esto prueba la primera igualdad. La demostración para la segunda es análoga.

6.2.13 DEFINICIÓN. Sea E un e.v. de dimensión $n \neq 0$ sobre $\mathbb{K}$ y $f \in \mathcal{L}_2(E, \mathbb{K})$ una forma bilineal sobre E. Llamamos *rango* de f a cualquiera de los números siguientes:

— el rango de u_f ,

— el rango de v_f ,

— el rango de $[f, (a_i)]$, cualquiera que sea la base $(a_1, \ldots, a_n)$ de E;

nótese que estos números son todos iguales, como consecuencia de la proposición precedente. Representamos el rango de f por $\operatorname{rg} f$.

6.2.14 PROPOSICIÓN

Sea E un e.v. de dimensión $n \neq 0$ sobre $\mathbb{K}$ y $f \in \mathcal{S}_2(E, \mathbb{K})$ una forma bilineal y simétrica sobre E. Se tiene que

$$n = \operatorname{rg} f + \dim \operatorname{Ker} f \,.$$

Además, las siguientes propiedades son equivalentes

(i) f es no degenerada;
(ii) $\operatorname{rg} f = n$;
(iii) la aplicación $u_f = v_f$ es un isomorfismo entre los espacios vectoriales E y E^*.

La primera parte resulta de que

$$n = \operatorname{rg} u_f + \dim \operatorname{Ker} u_f \,.$$

En cuanto a la segunda, las tres propiedades equivalen a que $\operatorname{Ker} f = \{0\}$ (obsérvese que $\dim E = \dim E^*$).

6.2.15 En las condiciones de la proposición anterior, si fijamos una base $(a_1, \ldots, a_n)$ de E y consideramos la matriz $A = [f, (a_i)]$, basta calcular $\det A$ para saber si f es degenerada o no lo es; f es no degenerada cuando $\det A \neq 0$. El $\det A$ recibe el nombre de *discriminante* de f. Resulta que f posee distintos discriminantes puesto que, si $(a'_1, \ldots, a'_n)$ es otra base de E y $A' = [f, (a'_i)]$, entonces

$$A' = P^{\,t}\, A\, P \,,$$

donde P es la matriz de paso; como

$$\det A' = \det A\, (\det P)^2 \,,$$

en general $\det A' \neq \det A$. Ahora bien, como $\det P \neq 0$, resulta que $\det A \neq 0$ si y sólo si $\det A' \neq 0$. Para comprobar que f es no degenerada, sirve entonces cualquier discriminante de f.

6.2.16 Ejemplo. Las formas simétricas sobre $\mathbb{K}^2$, $\mathbb{K}^3$ y $\mathbb{K}^n$ de (6.1.2) poseen rangos 2, 3 y n, iguales a la dimensión del espacio sobre el que están definidas. Se trata, pues, de formas no degeneradas.

6.2.17 Ejemplo. La forma bilineal y simétrica sobre l^2 o sobre c_{00}

$$f(x, y) = \sum_{n=1}^{\infty} x_n y_n \,,$$

definida en el ejemplo (6.1.3), es no degenerada. En efecto, si $x \in l^2$ (o c_{00}) y $x \neq 0$, la sucesión x tendrá algún término no nulo, $x_k \neq 0$; entonces, para la sucesión $y = (y_n)_{n \in \mathbb{N}}$ con

$$y_k = 1 \quad \text{e} \quad y_n = 0 \quad \text{si} \quad n \neq k$$

se tiene que

$$f(x, y) = \sum_{n=1}^{\infty} x_n y_n = x_k \neq 0\,;$$

luego $x \notin \operatorname{Ker} f$. Esto prueba que $\operatorname{Ker} f = \{0\}$.

6.2.18 Ejemplo. La forma simétrica sobre $\mathcal{C}([a, b], \mathbb{R})$ del ejemplo (6.1.4) es también no degenerada. Probaremos este hecho en (6.5.13) de una menera muy sencilla.

6.2.19 Ejemplo. La forma bilineal sobre $\mathbb{R}^3$

$$f(\,(x^1, x^2, x^3), (y^1, y^2, y^3)\,) = x^1 y^1 - x^2 y^2$$

se representa en la base canónica por la matriz

$$A = \begin{bmatrix} 1 & & \\ & -1 & \\ & & 0 \end{bmatrix},$$

y es por lo tanto simétrica. Su rango es

$$\operatorname{rg} f = \operatorname{rg} A = 2\,,$$

luego es degenerada. Utilizando (6.2.12), resulta fácil calcular su núcleo:

$$\begin{aligned} \operatorname{Ker} f &= \operatorname{Ker} u_f \\ &= \{(x^1, x^2, x^3) \mid A \begin{bmatrix} x^1 \\ x^2 \\ x^3 \end{bmatrix} = 0\} \\ &= \{(x^1, x^2, x^3) \mid x^1 = x^2 = 0\}\,. \end{aligned}$$

Compárese este núcleo con el conjunto que obtuvimos en (6.2.7).

6.3 Formas cuadráticas.

6.3.1 DEFINICIÓN. Sea $f \in \mathcal{S}_2(E, \mathbb{K})$ una forma bilineal simétrica sobre el espacio vectorial E. La aplicación

$$\begin{aligned} q : E &\to \mathbb{K} \\ x &\to q(x) \end{aligned}$$

que envía cada vector $x \in E$ al escalar

$$q(x) = f(x, x)\,,$$

recibe el nombre de *forma cuadrática* asociada a f. Es posible considerar la misma definición para formas bilineales que no sean simétricas; nosotros no lo haremos ya que esto nos crearía dificultades suplementarias sin proporcionarnos ningún beneficio adicional (véanse, sin embargo, los ejercicios de esta sección).

A pesar de que q es una aplicación de E en $\mathbb{K}$, q no es generalmente una forma lineal sobre E. Se comprueba sin dificultad que q verifica

$$g(x + y) = q(x) + q(y) + 2f(x, y)$$

y

$$q(\lambda x) = \lambda^2 q(x)$$

(compárense estas propiedades con las **(ln1)** y **(ln2)** de 1.1.10). Más exactamente, si q es lineal, entonces

$$(\forall x, y) \quad f(x, y) = 0\,,$$

o sea, $f = 0$ y en consecuencia $q = 0$; este es el único caso en el que q es lineal.

6.3.2 Ejemplo. Las formas simétricas de (6.1.2) poseen como forma cuadrática asociada

$$\begin{aligned} q(x^1, x^2) &= (x^1)^2 + (x^2)^2\,, \\ q(x^1, x^2, x^3) &= (x^1)^2 + (x^2)^2 + (x^3)^2\,, \\ q(x^1, \ldots, x^n) &= (x^1)^2 + (x^2)^2 + \cdots + (x^n)^2\,. \end{aligned}$$

6.3.3 Ejemplo. La forma simétrica sobre $\mathbb{R}^3$

$$f(\,(x^1, x^2, x^3), (y^1, y^2, y^3)\,) = x^1 y^1 - x^2 y^2$$

posee como forma cuadrática asociada

$$q(x^1, x^2, x^3) = (x^1)^2 - (x^2)^2\,.$$

6.3.4 Ejemplo. Para la forma simétrica sobre l^2 o sobre c_{00} del ejemplo (6.1.3), la correspondiente forma cuadrática toma en cada sucesión x el valor

$$q(x) = \sum_{n=1}^{\infty} (x_n)^2\,.$$

6.3.5 Ejemplo. La forma simétrica sobre $\mathcal{C}([a,b],\mathbb{R})$ del ejemplo (6.1.4) posee como forma cuadrática asociada

$$q(f) = \int_a^b f(t)^2\, dt\,.$$

6.3.6 Sea $f \in \mathcal{S}_2(E,\mathbb{K})$ y q la forma cuadrática asociada. Como

$$q(x+y) = q(x) + q(y) + 2f(x,y)\,,$$

resulta que

$$\textbf{(6)} \qquad \boxed{f(x,y) = \frac{1}{2}(\,q(x+y) - q(x) - q(y)\,)\,.}$$

También se puede comprobar que

$$\textbf{(7)} \qquad \boxed{f(x,y) = \frac{1}{4}q(x+y) - \frac{1}{4}q(x-y)\,.}$$

Las expresiones **(6)** y **(7)**, que permiten reconstruir la forma simétrica f a partir de su forma cuadrática asociada q reciben el nombre de *fórmulas de polarización*. Nótese que, gracias a estas fórmulas, podemos garantizar que cada forma cuadrática proviene de una única forma bilineal simétrica (lo de 'simétrica' es aquí esencial para la unicidad), que suele llamarse 'forma polar' de q.

En particular, resulta que $q = 0$ si y sólo si $f = 0$.

6.3.7 Si $f \in \mathcal{S}_2(E,\mathbb{K})$ es una forma bilineal simétrica y q la forma cuadrática asociada, se suele llamar 'núcleo de q', y denotar por $\operatorname{Ker} q$, al núcleo de f; se dice que q es no degenerada si f es no degenerada.

Si E es de dimensión finita, del rango de f se dice que es el 'rango de q', y se lo denota por $\operatorname{rg} q$. Si además $(a_1,\ldots,a_n)$ es una base de E, representamos por $[q,(a_i)]$ la matriz $[f,(a_i)]$, y decimos que es la 'matriz de q' en la base $(a_1,\ldots,a_n)$. Nótese que, como f es simétrica, la matriz $[q,(a_i)]$ de la forma cuadrática q es siempre simétrica. Su elemento general α_i^j viene dado por

$$\alpha_i^j = f(a_i,a_j) = \frac{1}{4}q(a_i+a_j) - \frac{1}{4}q(a_i-a_j)\,,$$

como se deduce de la fórmula de polarización **(7)**.

6.3.8 Sea E un e.v. de dimensión $n \neq 0$ sobre $\mathbb{K}$ y $(a_1,\ldots,a_n)$ una base de E; sea $f \in \mathcal{S}_2(E,\mathbb{K})$ una forma bilineal simétrica sobre E y q la forma cuadrática asociada. La fórmula **(1)** de (6.1.17) nos dice que

$$f(x,y) = [y,(a_i)]^t\,[f,(a_i)]\,[x,(a_i)]\,;$$

por lo tanto

$$\text{(8)} \qquad \boxed{q(x) = [x,(a_i)]^t\,[q,(a_i)]\,[x,(a_i)]\,,}$$

o sea, si $[q,(a_i)] = [\alpha_i^j]$, entonces q toma en $x = (x^1,\ldots,x^n)_{(a_i)}$ el valor

$$q(x) = \begin{bmatrix} x^1 & \cdots & x^n \end{bmatrix} \begin{bmatrix} \alpha_1^1 & \cdots & \alpha_n^1 \\ \vdots & & \vdots \\ \alpha_1^n & \cdots & \alpha_n^n \end{bmatrix} \begin{bmatrix} x^1 \\ \vdots \\ x^n \end{bmatrix} = \sum_{i,j=1}^{n} \alpha_i^j x^i x^j\,.$$

Conviene agrupar los términos de esta suma en la forma

$$q(x) = \sum_{i=1}^{n} \alpha_i^i (x^i)^2 + \sum_{1\le i<j\le n} (\alpha_i^j + \alpha_j^i) x^i x^j\,,$$

o mejor aún, puesto que la matriz $[q,(a_i)]$ es simétrica,

$$\text{(9)} \qquad \boxed{q(x) = \sum_{i=1}^{n} \alpha_i^i (x^i)^2 + \sum_{1\le i<j\le n} 2\alpha_i^j x^i x^j\,.}$$

Si $A \in M_{\mathbb{K}}(n)$ es una matriz simétrica, la forma bilineal sobre $\mathbb{K}^n$ dada por A (v. 6.1.17)

$$f(x,y) = y^t\,A\,x\,,$$

es simétrica; la correspondiente forma cuadrática

$$\text{(10)} \qquad \boxed{q(x) = x^t\,A\,x}$$

es, en el lenguaje habitual, la 'forma cuadrática dada por A'.

6.3.9 Ejemplo. La forma cuadrática sobre $\mathbb{K}^n$ dada por la matriz unidad I_n es

$$q(x^1,\ldots,x^n) = \begin{bmatrix} x^1 & \cdots & x^n \end{bmatrix} \begin{bmatrix} x^1 \\ \vdots \\ x^n \end{bmatrix} = (x^1)^2 + \cdots + (x^n)^2\,,$$

es decir, la que obtuvimos en (6.3.2).

6.3.10 Ejemplo. La forma cuadrática sobre $\mathbb{R}^3$

$$q(x^1,x^2,x^3) = (x^1)^2 - (x^2)^2$$

de (6.3.3) es la forma cuadrática dada por la matriz

$$\begin{bmatrix} 1. & & \\ & -1. & \\ & & 0 \end{bmatrix}.$$

6.3.11 Comparando las expresiones de estas formas cuadráticas con la fórmula **(9)**, vemos que son nulos todos los términos de la segunda suma que aparece en **(9)**. Uno de nuestros objetivos es justamente hacer esto cierto para toda forma cuadrática sobre un espacio de dimensión finita; este objetivo es alcanzable siempre, como veremos en la sección siguiente, cuando se utiliza una base adecuada del espacio. Intentaremos también que los coeficientes de los términos de la primera suma de **(9)** sean 1, pero esto no será posible en todos los casos.

6.3.12 Uno de los problemas que se presentan es el saber qué aplicaciones

$$q : E \to \mathbb{K}$$

son formas cuadráticas. En (6.3.1) hemos visto que las formas lineales no lo son, excepción hecha de la forma 0. la fórmula **(9)** nos proporciona la solución de este problema cuando el espacio es de dimensión finita.

Si E es de dimensión finita y $(a_1, \ldots, a_n)$ una base de E, toda forma cuadrática q posee una expresión de la forma

$$q(x) = \sum_{i=1}^{n} \beta_i^i (x^i)^2 + \sum_{1 \leq i < j \leq n} \beta_i^j x^i x^j$$

para $x = (x^1, \ldots, x^n)_{(a_i)}$. De una expresión de este tipo se suele decir que es un 'polinomio homogéneo de segundo grado' en las indeterminadas $x^1, \ldots, x^n$.

Recíprocamente, si para $x = (x^1, \ldots, x^n)_{(a_i)}$ ponemos

$$q(x) = \sum_{i=1}^{n} \beta_i^i (x^i)^2 + \sum_{1 \leq i < j \leq n} \beta_i^j x^i x^j \ ,$$

entonces la función $q : E \to \mathbb{K}$ definida de esta manera es una forma cuadrática sobre E: la forma cuadrática que procede de la forma bilineal simétrica

$$f(x, y) = \sum_{i,j=1}^{n} \alpha_i^j x^i y^j \ ,$$

donde los coeficientes α_i^j (que son los elementos de $[f, (a_i)] = [q, (a_i)]$) vienen dados por

$$\alpha_i^i = \beta_i^i \quad (i = 1, \ldots, n)\, , \alpha_i^j = \beta_i^j/2 \quad (i < j)\, , \alpha_i^j = \beta_j^i/2 \quad (i > j)\, .$$

En efecto; no reviste ninguna dificultad comprobar que f es una forma bilineal; además, es simétrica, puesto que la matriz $[\alpha_i^j]$ es simétrica; finalmente, se comprueba sin problemas que $f(x, x) = q(x)$.

6.3.13 **Ejemplo.** La aplicación $q : \mathbb{R}^3 \to \mathbb{R}$ dada por

$$q(x^1, x^2, x^3) = (x^1)^2 + x^1 x^2 + x^1 x^3 - 2x^2 x^3$$

es una forma cuadrática, asociada a la forma bilineal simétrica f sobre $\mathbb{R}^3$ dada por

$$f(\,(x^1,x^2,x^3),(y^1,y^2,y^3)\,) = \\ = x^1y^1 + 0.5x^1y^2 + 0.5x^2y^1 + 0.5x^1y^3 + 0.5x^3y^1 - x^2y^3 - x^3y^2 \,.$$

Obsérvese que en la práctica se hace lo siguiente: los términos en $(x^i)^2$ se transforman en términos en x^iy^i con el mismo coeficiente, y los términos en x^ix^j se desdoblan en dos términos, uno en x^iy^j y otro en x^jy^i, cada uno de ellos con la mitad del coeficiente.

La matriz de f y de q en la base canónica de $\mathbb{R}^3$ es la matriz simétrica

$$[f,(e_i)] = [q,(e_i)] = \begin{bmatrix} 1. & 0.5 & 0.5 \\ 0.5 & 0 & -1. \\ 0.5 & -1. & 0 \end{bmatrix} .$$

6.4 Bases ortogonales.

6.4.1 DEFINICIÓN. Sea E un espacio vectorial de dimensión $n \neq 0$ sobre $\mathbb{K}$, $f \in \mathcal{S}_2(E,\mathbb{K})$ una forma bilineal simétrica sobre E y q la forma cuadrática asociada. Decimos que una base $(a_1,\ldots,a_n)$ de E es *ortogonal* para f (y para q) cuando

$$\textbf{(org)} \quad i \neq j \quad \Rightarrow \quad f(a_i,a_j) = 0 \,.$$

Esta propiedad equivale claramente a la siguiente

(org1) $[f,(a_i)]$ es una matriz diagonal.

Si la base $(a_1,\ldots,a_n)$ es ortogonal para f, entonces para $x = (x^1,\ldots,x^n)_{(a_i)}$ e $y = (y^1,\ldots,y^n)_{(a_i)}$ se tiene

$$f(x,y) = [y,(a_i)]^t\,[f,(a_i)]\,[x,(a_i)] = \sum_{i=1}^n \alpha_i^i x^i y^i \,,$$

puesto que $[f,(a_i)]$ es diagonal; los α_i^i son los elementos diagonales de $[f,(a_i)]$. Se tiene también

$$q(x) = \sum_{i=1}^n \alpha_i^i (x^i)^2 \,.$$

Recíprocamente, si f posee dicha forma (o q posee dicha forma), entonces $[f,(a_i)]$ es diagonal y la base $(a_1,\ldots,a_n)$ es ortogonal. Es decir, la base $(a_1,\ldots,a_n)$ es ortogonal si y sólo si

$$\textbf{(org2)} \;\; f(x,y) = \sum_{i=1}^n \alpha_i^i x^i y^i \text{ para } x{=}(x^1,\ldots,x^n)_{(a_i)} \text{ e } y{=}(y^1,\ldots,y^n)_{(a_i)} \,,$$

y también si y sólo si

$$\textbf{(org3)} \quad q(x) = \sum_{i=1}^{n} \alpha_i^i (x^i)^2 \text{ para } x = (x^1, \ldots, x^n)_{(a_i)} \,.$$

La bases ortogonales son, pues, las que resuelven el problema que apuntábamos en (6.3.11).

6.4.2 Ejemplo. La base canónica es una base ortogonal para las formas cuadráticas de los ejemplos (6.3.2) y (6.3.3).

6.4.3 Ejemplo. Para la forma cuadrática sobre $\mathbb{R}^3$

$$q(x^1, x^2, x^3) = (x^1)^2 + x^1x^2 + x^1x^3 - 2x^2x^3$$

del ejemplo (6.3.13), la base canónica no es ortogonal, ya que $[q, (e_i)]$ no es diagonal. Sin embargo, la base (a_1, a_2, a_3) formada por

$$a_1 = (1, 0, 0)\,, \quad a_2 = (-1, 2, 0) \quad \text{y} \quad a_3 = (-2, -5, 1)$$

es una base ortogonal para q, como se puede comprobar.

6.4.4 TEOREMA

Sea E un e.v. de dimensión $n \neq 0$ sobre $\mathbb{K}$ y $f \in \mathcal{S}_2(E, \mathbb{K})$ una forma bilineal simétrica sobre E. Existe una base de E que es ortogonal para f.

Utilizaremos el principio de inducción sobre la dimensión n de E. Si $n = 1$, cualquier base de E es ortogonal (la matriz de f es 1×1, luego es diagonal).

Supondremos (hipótesis de recurrencia) que el resultado es cierto cuando el espacio es de dimensión $n - 1$; veamos que entonces el resultado es cierto para un espacio E de dimensión n. Si $f = 0$, cualquier base de E verifica **(org)** y es ortogonal. Si $f \neq 0$, entonces la forma cuadrática asociada q verifica también que $q \neq 0$ (v. 6.3.6); existe, pues, un vector $a_1 \in E$ tal que

$$q(a_1) = f(a_1, a_1) \neq 0 \,,$$

lo que significa en particular que $a_1 \neq 0$. Consideremos entonces el subconjunto de E

$$F = \{x \in E \mid f(x, a_1) = 0\} \,;$$

resulta sencillo probar que F es un subespacio vectorial de E. Veamos que

$$E = \langle a_1 \rangle \oplus F \,.$$

Si $x \in \langle a_1 \rangle \cap F$, entonces $x = \lambda a_1$ y además

$$\lambda f(a_1, a_1) = f(\lambda a_1, a_1) = f(x, a_1) = 0 \,;$$

como $f(a_1, a_1) \neq 0$, resulta que $\lambda = 0$ y que $x = \lambda a_1 = 0$; esto prueba que $\langle a_1 \rangle \cap F = \{0\}$. Sea $x \in E$; pongamos

$$x_1 = \frac{f(x, a_1)}{f(a_1, a_1)}\, a_1 \quad \text{y} \quad x_2 = x - \frac{f(x, a_1)}{f(a_1, a_1)}\, a_1 \, ;$$

es exidente que $x_1 \in \langle a_1 \rangle$ y fácilmente se prueba que $x_2 \in F$; además, $x = x_1 + x_2$. Esto significa que $E = \langle a_1 \rangle + F$ y que $E = \langle a_1 \rangle \oplus F$. Resulta entonces que $\dim F = n - 1$. La restricción de f

$$g : F \times F \to \mathbb{K}$$

a $F \times F$ (v. 6.1.9) es una forma bilineal y simétrica sobre F. Gracias a la hipótesis de recurrencia, existe una base $(a_2, \ldots, a_n)$ de F que es ortogonal para g. Consideremos finalmente el sistema $(a_1, a_2, \ldots, a_n)$; es una base de E y es ortogonal para f (razónese por qué lo es), lo que finaliza la demostración.

6.4.5 Si $A \in M_{\mathbb{K}}(n)$ ($\mathbb{K} = \mathbb{R}$ o $\mathbb{C}$) es una matriz simétrica y consideramos la forma bilineal sobre $\mathbb{K}^n$ dada por A (que es una forma simétrica), el teorema precedente y la fórmula **(3)** de (6.1.20) relativa al cambio de base nos permiten afirmar que existe una matriz $P \in M_{\mathbb{K}}(n)$ inversible y tal que $P^t A\, P$ es diagonal. Es la llamada 'diagonalización por congruencia', que nada tiene que ver en principio con la 'diagonalización por semejanza' que estudiamos en el capítulo 5 (para las definiciones de congruencia y de semejanza, consúltense los apartados 6.1.20 y 2.5.25).

Es cierto que, si $A \in M_{\mathbb{R}}(n)$ es una matriz simétrica real, entonces A es diagonalizable por semejanza. Existen, no obstante, matrices simétricas complejas, $A \in M_{\mathbb{C}}(n)$, que no son diagonalizables; por ejemplo, se puede comprobar que la matriz

$$A = \begin{bmatrix} 2. & i \\ i & 0 \end{bmatrix}$$

no es diagonalizable.

6.4.6 Ejemplo (método de Gauss para el cálculo de una base ortogonal).

Aunque la justificación teórica del método que vamos e exponer no resulta conceptualmente difícil, es extremadamente compleja desde el punto de vista de las notaciones. Parece aconsejable, e este nivel, tratar de ver mediante un ejemplo cómo se puede utilizar.

Supongamos que se desea calcular una base ortogonal para la forma cuadrática sobre el espacio $\mathbb{R}^4$ dada por

$$q(x^1, x^2, x^3, x^4) = 2(x^1)^2 + 2(x^2)^2 + 2(x^4)^2 + 4x^1x^2 - 4x^1x^4 - 2x^2x^3 - 4x^2x^4 - 4x^3x^4 \,.$$

La matriz de q en la base canínica de $\mathbb{R}^4$ es

$$A = [q, (e_1, e_2, e_3, e_4)] = \begin{bmatrix} 2. & 2. & 0 & -2. \\ 2. & 2. & -1. & -2. \\ 0 & -1. & 0 & -2. \\ -2. & -2. & -2. & 2. \end{bmatrix} .$$

Como lo que buscamos es una base (a_1, a_2, a_3, a_4), ortogonal para q, debemos conseguir que en dicha base q posea la forma

$$q({x'}^1 a_1 + {x'}^2 a_2 + {x'}^3 a_3 + {x'}^4 a_4) = {\alpha'}^1_1({x'}^1)^2 + {\alpha'}^2_2({x'}^2)^2 + {\alpha'}^3_3({x'}^3)^2 + {\alpha'}^4_4({x'}^4)^2\,,$$

con ${\alpha'}^1_1, {\alpha'}^2_2, {\alpha'}^3_3, {\alpha'}^4_4 \in \mathbb{R}$. Vamos a reducir entonces la primitiva expresión de q a una forma como ésta.

Nos ocupamos en primer lugar de conseguir un término 'cuadrado' que agrupe a todos aquellos en los que interviene x^1 :

$$q(x^1, x^2, x^3, x^4) = 2(\,(x^1)^2 + 2x^1(x^2 - x^4)\,) + t(x^2, x^3, x^4)\,,$$

donde $t(x^2, x^3, x^4)$ es una expresión en la que no interviene x^1, y de la que nos ocuparemos más adelante. Ahora llevamos los términos que hemos seleccionado a la forma $(\,x^1 + (x^2 - x^4)\,)^2$; el proceso es el siguiente:

$$\begin{aligned}
q(x^1, x^2, x^3, x^4) &= 2(\,(x^1)^2 + 2x^1(x^2 - x^4) + (x^2 - x^4)^2\,) - 2(x^2 - x^4)^2 + t(x^2, x^3, x^4) \\
&= 2(x^1 + x^2 - x^4)^2 - 2(x^2 - x^4)^2 + t(x^2, x^3, x^4) \\
&= 2(x^1 + x^2 - x^4)^2 - 2(x^2 - x^4)^2 + 2(x^2)^2 + 2(x^4)^2 - 2x^2x^3 - 4x^2x^4 - 4x^3x^4 \\
&= 2(x^1 + x^2 - x^4)^2 - 2x^2x^3 - 4x^3x^4\,.
\end{aligned}$$

Repetimos ahora el proceso de formación de un 'cuadrado' con los términos restantes; si entre estos términos apareciese $(x^2)^2$, $(x^3)^2$ o $(x^4)^2$, estaríamos en un caso similar al precedente; no es así y el método debe ser diferente. Seleccionamos un término de la forma x^ix^j, por ejemplo $-2x^2x^3$, y procedemos al agrupamiento de todos los términos en los que intervienen x^2 y x^3:

$$\begin{aligned}
q(x^1, x^2, x^3, x^4) &= 2(x^1 + x^2 - x^4)^2 - 2x^2x^3 - 4x^3x^4 \\
&= 2(x^1 + x^2 - x^4)^2 - 2(x^2x^3 + x^2(0) + x^3(2x^4)\,) + t(x^4)
\end{aligned}$$

donde $t(x^4)$ es una expresión en la que no intervienen x^1, x^2 y x^3, y que en este caso es cero. Llevamos los términos seleccionados a la forma $(x^2 + 2x^4)(x^3 + 0)$, o sea $(x^2 + 2x^4)x^3$:

$$q(x^1, x^2, x^3, x^4) = 2(x^1 + x^2 - x^4)^2 - 2(x^2 + 2x^4)x^3\,.$$

Utilizando la identidad

$$ab = \frac{1}{4}(a + b)^2 - \frac{1}{4}(a - b)^2$$

obtenemos

$$q(x^1, x^2, x^3, x^4) = 2(x^1 + x^2 - x^4)^2 - 0.5(x^2 + x^3 + 2x^4)^2 + 0.5(x^2 - x^3 + 2x^4)^2\,.$$

Esto finaliza la primera parte del método; diremos para entendernos que hemos puesto q en la forma de una suma de cuadrados.

Nuestro objetivo aparece ahora mucho más nítido. Si conseguimos que las coordenadas de un vector en la base (a_1, a_2, a_3, a_4) , ${x'}^1$, ${x'}^2$, ${x'}^3$ y ${x'}^4$, se relacionen con las coordenadas x^1, x^2, x^3 y x^4 del mismo vector en la base canónica por las igualdades

$$\begin{aligned} x'^1 &= x^1 + x^2 - x^4 \\ x'^2 &= x^2 + x^3 + 2x^4 \\ x'^3 &= x^2 - x^3 + 2x^4 \\ x'^4 &= \lambda_1 x^1 + \lambda_2 x^2 + \lambda_3 x^3 + \lambda_4 x^4 \ , \end{aligned}$$

con $\lambda_1, \lambda_2, \lambda_3, \lambda_4 \in \mathbb{R}$ cualesquiera, habremos logrado que la expresión de q en la base (a_1, a_2, a_3, a_4) sea

$$q(x'^1 a_1 + x'^2 a_2 + x'^3 a_3 + x'^4 a_4) = 2(x'^1)^2 - 0.5(x'^2)^2 + 0.5(x'^3)^2 \, ,$$

y que dicha base sea ortogonal. Llamando P a la matriz de paso de (e_1, e_2, e_3, e_4) a (a_1, a_2, a_3, a_4), tenemos que

$$\begin{bmatrix} x'^1 \\ x'^2 \\ x'^3 \\ x'^4 \end{bmatrix} = P^{-1} \begin{bmatrix} x^1 \\ x^2 \\ x^3 \\ x^4 \end{bmatrix} ,$$

luego está claro que P^{-1} debe ser

$$P^{-1} = \begin{bmatrix} 1. & 1. & 0 & -1. \\ 0 & 1. & 1. & 2. \\ 0 & 1. & -1. & 2. \\ \lambda_1 & \lambda_2 & \lambda_3 & \lambda_4 \end{bmatrix} ;$$

los λ_i son arbitrarios, pero deben ser elegidos de tal manera que P^{-1} sea inversible. Se logra este objetivo haciendo $\lambda_1 = \lambda_2 = \lambda_3 = 0$ y $\lambda_4 = 1$, ya que entonces

$$P^{-1} = \begin{bmatrix} 1. & 1. & 0 & -1. \\ 0 & 1. & 1. & 2. \\ 0 & 1. & -1. & 2. \\ 0 & 0 & 0 & 1. \end{bmatrix}$$

y $\det P^{-1} = -2$. La matriz de paso es en este caso

$$P = (P^{-1})^{-1} = \begin{bmatrix} 1. & -0.5 & -0.5 & 3. \\ 0 & 0.5 & 0.5 & -2. \\ 0 & 0.5 & -0.5 & 0 \\ 0 & 0 & 0 & 1. \end{bmatrix} ,$$

y la base (a_1, a_2, a_3, a_4) está formada por los vectores

$$\begin{aligned} a_1 &= (\ 1.\ ,\ 0\ ,\ 0\ ,\ 0\ ) \\ a_2 &= (\ -0.5,\ 0.5,\ 0.5,\ 0\ ) \\ a_3 &= (\ -0.5,\ 0.5,\ -0.5,\ 0\ ) \\ a_4 &= (\ 3.\ ,\ -2.\ ,\ 0\ ,\ 1.\ ) \, . \end{aligned}$$

Esta base es ortogonal para q. Se puede comprobar que

$$[q,(a_1,a_2,a_3,a_4)] = P^t\,A\,P = \begin{bmatrix} 2. & & & \\ & -0.5 & & \\ & & 0.5 & \\ & & & 0 \end{bmatrix}$$

y que, por lo tanto,

$$q({x'}^1 a_1 + {x'}^2 a_2 + {x'}^3 a_3 + {x'}^4 a_4) = 2.({x'}^1)^2 - 0.5({x'}^2)^2 + 0.5({x'}^3)^2\,,$$

como habíamos previsto.

Advertencia: Retrocedamos al instante en que, puesta q en la forma

$$q(x^1,x^2,x^3,x^4) = 2(x^1+x^2-x^4)^2 - 0.5(x^2+x^3+2x^4)^2 + 0.5(x^2-x^3+2x^4)^2\,.$$

estamos obligados a completar la matriz

$$\begin{bmatrix} 1. & 1. & 0 & -1. \\ 0 & 1. & 1. & 2. \\ 0 & 1. & -1. & 2. \\ \lambda_1 & \lambda_2 & \lambda_3 & \lambda_4 \end{bmatrix}$$

con números λ_i que hagan dicha matriz inversible. En nuestro caso ha sido posible hacerlo, pero nada parece garantizar esta posibilidad.

Para el ejemplo

$$\begin{aligned} r(x^1,x^2,x^3,x^4) &= 2(x^1)^2 + (x^2)^2 - (x^4)^2 + 4x^1x^2 + 2x^2x^4 \\ &= (x^1+x^2)^2 + (x^1+x^2)^2 - (x^2-x^4)^2\,, \end{aligned}$$

la matriz

$$\begin{bmatrix} 1. & 1. & 0 & 0 \\ 1. & 1. & 0 & 0 \\ 0 & 1. & 0 & -1. \\ \lambda_1 & \lambda_2 & \lambda_3 & \lambda_4 \end{bmatrix}$$

no se puede prolongar hasta una matriz inversible.

Para el ejemplo

$$\begin{aligned} s(x^1,x^2,x^3,x^4) &= (x^1)^2 + (x^3)^2 + (x^4)^2 - 2x^1x^3 + 4x^2x^4 \\ &= (x^1-x^3)^2 + (x^2+x^4)^2 - (x^2-x^4)^2 + (x^4)^2\,, \end{aligned}$$

la matriz

$$\begin{bmatrix} 1. & 0 & -1. & 0 \\ 0 & 1. & 0 & 1. \\ 0 & 1. & 0 & -1. \\ 0 & 0 & 0 & 1. \end{bmatrix},$$

que resulta directamente, no es inversible y no sirve para continuar el proceso.

La justificación teórica del método de Gauss garantiza que esto no puede suceder a condición de respetar las siguientes reglas en la descomposición en suma de cuadrados

— siempre que exista un término en $(x^i)^2$, se debe agrupar este término y 'todos' aquellos en los que intervenga x^i;

— sólo en el caso en que no quede ningún término en $(x^i)^2$, se debe proceder al agrupamiento a partir de un $x^i x^j$; en este caso, se deben agrupar con $x^i x^j$ 'todos' los términos en los que intervengan x^i y x^j.

De acuerdo con estas reglas, las formas cuadráticas r y s de los ejemplos precedentes quedan reducidas a

$$r(x^1, x^2, x^3, x^4) = 2(x^1 + x^2)^2 - (x^2 - x^4)^2 ,$$

$$s(x^1, x^2, x^3, x^4) = (x^1 - x^3)^2 + (2x^2 + x^4)^2 - 4(x^2)^2 ,$$

que son descomposiciones que permiten proseguir el proceso en la forma indicada.

6.4.7 Ejemplo (método de las operaciones elementales para la diagonalización por congruencia).

Junto al método de Gauss, que acabamos de ver, hay otro procedimiento, más sencillo en ciertos casos, que aprovecha la estructura simétrica de la matriz de las formas bilineales simétricas. El método hace uso de las operaciones elementales, presentadas en (2.4.31), y que en su momento empleamos para invertir matrices (v. 2.4.35) o calcular determinantes (v. 3.3.5). Conservamos la notación empleada en (2.4.31).

Sea f una operación elemental de fila (v. 2.4.31), o sea, una de las f_j^λ, $f_{j,k}$ o $f_{j,k}^\lambda$. Denotemos por $F = f(I)$ la correspondiente matriz elemental. Representemos por c la misma operación elemental, pero sobre las columnas de la matriz, y pongamos $C = c(I)$. Cuando A es una matriz $n \times n$ arbitraria, sabemos (v. 2.4.33) que $f(A) = F\,A$. Pues bien, resulta también que

$$(\forall A \in M(n)) \quad c(A) = A\,C$$

(ahora la multiplicación debe ser a la derecha; es la forma de afectar a las columnas de A). Pero, además, resulta que $C = F^t$ y $F = C^t$, ya que $c(I)$ y $f(I)$ son transpuestas una de la otra, como se comprueba con facilidad.

Supongamos ahora que realizamos sobre A una operación elemental de fila y, sobre la matriz resultante, la misma operación elemental de columna. Si F y C son las correspondientes matrices elementales, habremos transformado A en la matriz $F\,A\,F^t = C^t\,A\,C$, que es congruente con A.

Pues bien, el método de las operaciones elementales para la diagonalización por congruencia consiste en transformar una matriz simétrica A en una matriz diagonal D mediante pares de operaciones elementales f_1, c_1 $\dots$, f_p, c_p, o sea,

$$c_p(f_p \dots (c_1(f_1(A))\dots) = D .$$

Describiremos esto mediante las correspondientes matrices elementales. Poniendo $F_i = f_i(I)$ y $C_i = c_i(I)$, lo que tenemos es

$$F_p \dots F_1\,A\,C_1 \dots C_p = D ,$$

o sea,

$$(C_1 \, \dots \, C_p)^t \, A \, C_1 \, \dots \, C_p = D \, .$$

De esa manera habremos transformado A en la matriz D, diagonal y congruente con A. La matriz inversible P que hace que $P^t A \, P = D$ es

$$P = C_1 \, \dots \, C_p = (F_p \, \dots \, F_1)^t \, .$$

Nótese que

$$P = C_1 \, \dots \, C_p = c_p(\dots c_1(I) \, \dots) \, ,$$

o sea, que se obtiene P realizando sobre la matriz unidad las operaciones elementales de columna (sólo las de columna) que hemos llevado a cabo en A. Alternativamente, como

$$P^t = F_p \, \dots \, F_1 = f_p(\dots f_1(I) \, \dots) \, ,$$

obtenemos la traspuesta de P realizando sobre la matriz unidad las operaciones elementales de fila (sólo las de fila) que hemos llevado a cabo en A.

En la práctica, se actúa sobre A y sobre I simultáneamente; en el lugar de A se obtiene la matriz diagonal congruente con A; en el lugar de I se obtiene la matriz de paso (realizando únicamente las transformaciones de columna) o la traspuesta de la matriz de paso (realizando únicamente las transformaciones de fila).

A estas alturas, el lector tendrá una cierta experiencia en 'triangularizar' matrices mediante operaciones elementales de fila. Por lo tanto no detallaremos de nuevo el proceso, ya que la técnica básica es aquí la misma, pues sólo hay que preocuparse de introducir ceros bajo la diagonal; las operaciones gemelas de columna se encargan (gracias a la simetría de A) de introducir ceros sobre ella.

De forma previa al ejemplo concreto, algunas consideraciones sobre la posibilidad de llevar a cabo el procedimiento, siempre en el supuesto de que A es simétrica. Explicaremos cómo reducir el problema a un orden inferior; la repetición del proceso permite llegar al final en un número finito de pasos. Se trata de emplear el elemento α_1^1 de A como pivote para introducir ceros en los restantes lugares de la primera columna (y al mismo tiempo en los de la primera fila). Si $\alpha_1^1 \neq 0$, esto se hace como ya vimos en (2.4.36): se suma a la segunda fila la primera multiplicada por $-\alpha_1^2/\alpha_1^1$, y otro tanto se hace con las correspondientes columnas; así se introducen ceros en los lugares α_1^2 y α_2^1. Luego se procede igual para substituir α_1^3 y α_3^1, y así sucesivamente. De esta manera A queda reemplazada por una matriz

$$\left[\begin{array}{c|ccc} \alpha_1^1 & 0 & \cdots & 0 \\ \hline 0 & & & \\ \vdots & & A_1 & \\ 0 & & & \end{array}\right] ,$$

con A_1 también simétrica y de un orden inferior. El mismo proceso se repite con A_1 (obsérvese que el nuevo proceso no altera los elementos de la matriz que quedan fuera de A_1) y así sucesivamente.

Cuando $\alpha_1^1 = 0$, el proceso no se puede hacer directamente y hay que substituir este elemento por otro. Supondremos que entre los $\alpha_1^2, \dots, \alpha_1^n$ hay alguno diferente de 0, pues en caso contrario ya podríamos pasar al orden inmediato inferior directamente. Hay dos posibilidades de reemplazamiento de α_1^1. Si alguno de los elementos diagonales $\alpha_2^2, \ldots, \alpha_n^n$ es no nulo, se substituye α_1^1 por dicho elemento mediante un cambio de filas y el correspondiente de columnas. Si $\alpha_2^2, \ldots, \alpha_n^n$ fuesen nulos, se considera el elemento α_1^j que sea distinto de 0 y se suma a la primera fila la fila número j (junto con la correspondiente operación sobre las columnas). Esto permite reemplazar α_1^1 por $\alpha_1^j + \alpha_j^1 = 2\,\alpha_1^j \neq 0$.

Pasemos ahora a ejemplos concretos, y nada mejor que comenzar con el mismo ejemplo que hemos resuelto con el método de Gauss en (6.4.6). Supongamos de nuevo que se desea calcular una base ortogonal para la forma cuadrática sobre el espacio $\mathbb{R}^4$ dada por

$$q(x^1,x^2,x^3,x^4) = 2(x^1)^2+2(x^2)^2+2(x^4)^2+4x^1x^2-4x^1x^4-2x^2x^3-4x^2x^4-4x^3x^4\,.$$

Ya dijimos que la matriz de q en la base canónica de $\mathbb{R}^4$ era

$$A = [q,(e_1,e_2,e_3,e_4)] = \begin{bmatrix} 2. & 2. & 0 & -2. \\ 2. & 2. & -1. & -2. \\ 0 & -1. & 0 & -2. \\ -2. & -2. & -2. & 2. \end{bmatrix}.$$

Por el momento, será esta matriz A con la que trabajemos, al mismo tiempo que lo haremos con la matriz unidad I_4. La letra 'f' representará operaciones de fila, que realizaremos simultáneamente con ambas matrices. La letra 'c' representará operaciones de columna, que sólo haremos a la primera de nuestras matrices. Recordemos que, a medida que A se va transformando en una matriz diagonal, I_4 se transforma en la traspuesta P^t de la matriz de paso.

$$\begin{bmatrix} 2. & 2. & 0 & -2. \\ 2. & 2. & -1. & -2. \\ 0 & -1. & 0 & -2. \\ -2. & -2. & -2. & 2. \end{bmatrix} \qquad \begin{bmatrix} 1. & 0 & 0 & 0 \\ 0 & 1. & 0 & 0 \\ 0 & 0 & 1. & 0 \\ 0 & 0 & 0 & 1. \end{bmatrix}$$

$$\text{con } f_{2,1}^{-1.} \qquad \begin{bmatrix} 2. & 2. & 0 & -2. \\ 0 & 0 & -1. & 0 \\ 0 & -1. & 0 & -2. \\ -2. & -2. & -2. & 2. \end{bmatrix} \qquad \begin{bmatrix} 1. & 0 & 0 & 0 \\ 0 & 1. & 0 & 0 \\ 0 & 0 & 1. & 0 \\ 0 & 0 & 0 & 1. \end{bmatrix}$$

$$\text{con } c_{2,1}^{-1.} \qquad \begin{bmatrix} 2. & 0 & 0 & -2. \\ 0 & 0 & -1. & 0 \\ 0 & -1. & 0 & -2. \\ -2. & 0 & -2. & 2. \end{bmatrix} \qquad \begin{bmatrix} 1. & -1. & 0 & 0 \\ 0 & 1. & 0 & 0 \\ 0 & 0 & 1. & 0 \\ 0 & 0 & 0 & 1. \end{bmatrix}$$

$$\text{con } f_{4,1}^{1.} \qquad \begin{bmatrix} 2. & 0 & 0 & -2. \\ 0 & 0 & -1. & 0 \\ 0 & -1. & 0 & -2. \\ 0 & 0 & -2. & 0 \end{bmatrix} \qquad \begin{bmatrix} 1. & -1. & 0 & 0 \\ 0 & 1. & 0 & 0 \\ 0 & 0 & 1. & 0 \\ 0 & 0 & 0 & 1. \end{bmatrix}$$

$$\text{con } c_{4,1}^{1.} \quad \begin{bmatrix} 2. & 0 & 0 & 0 \\ 0 & 0 & -1. & 0 \\ 0 & -1. & 0 & -2. \\ 0 & 0 & -2. & 0 \end{bmatrix} \qquad \begin{bmatrix} 1. & -1. & 0 & 1. \\ 0 & 1. & 0 & 0 \\ 0 & 0 & 1. & 0 \\ 0 & 0 & 0 & 1. \end{bmatrix}$$

Nótese que ahora se anula el 'pivote' α_2^2 natural; y que, además, $\alpha_3^3 = \alpha_4^4 = 0$. Consideramos el elemento no nulo $\alpha_2^3 = \alpha_3^2 = -1$ y procedemos a sumar a la segunda fila la tercera y a la segunda columna la tercera:

$$\text{con } f_{2,3}^{1.} \quad \begin{bmatrix} 2. & 0 & 0 & 0 \\ 0 & -1. & -1. & -2. \\ 0 & -1. & 0 & -2. \\ 0 & 0 & -2. & 0 \end{bmatrix} \qquad \begin{bmatrix} 1. & -1. & 0 & 1. \\ 0 & 1. & 0 & 0 \\ 0 & 0 & 1. & 0 \\ 0 & 0 & 0 & 1. \end{bmatrix}$$

$$\text{con } c_{2,3}^{1.} \quad \begin{bmatrix} 2. & 0 & 0 & 0 \\ 0 & -2. & -1. & -2. \\ 0 & -1. & 0 & -2. \\ 0 & -2. & -2. & 0 \end{bmatrix} \qquad \begin{bmatrix} 1. & -1. & 0 & 1. \\ 0 & 1. & 0 & 0 \\ 0 & 1. & 1. & 0 \\ 0 & 0 & 0 & 1. \end{bmatrix}$$

$$\text{con } f_{3,2}^{-0.5} \quad \begin{bmatrix} 2. & 0 & 0 & 0 \\ 0 & -2. & -1. & -2. \\ 0 & 0 & 0.5 & -1. \\ 0 & -2. & -2. & 0 \end{bmatrix} \qquad \begin{bmatrix} 1. & -1. & 0 & 1. \\ 0 & 1. & 0 & 0 \\ 0 & 1. & 1. & 0 \\ 0 & 0 & 0 & 1. \end{bmatrix}$$

$$\text{con } c_{3,2}^{-0.5} \quad \begin{bmatrix} 2. & 0 & 0 & 0 \\ 0 & -2. & 0 & -2. \\ 0 & 0 & 0.5 & -1. \\ 0 & -2. & -1. & 0 \end{bmatrix} \qquad \begin{bmatrix} 1. & -1. & 0.5 & 1. \\ 0 & 1. & -0.5 & 0 \\ 0 & 1. & 0.5 & 0 \\ 0 & 0 & 0 & 1. \end{bmatrix}$$

$$\text{con } f_{4,2}^{-1.} \quad \begin{bmatrix} 2. & 0 & 0 & 0 \\ 0 & -2. & 0 & -2. \\ 0 & 0 & 0.5 & -1. \\ 0 & 0 & -1. & 2. \end{bmatrix} \qquad \begin{bmatrix} 1. & -1. & 0.5 & 1. \\ 0 & 1. & -0.5 & 0 \\ 0 & 1. & 0.5 & 0 \\ 0 & 0 & 0 & 1. \end{bmatrix}$$

$$\text{con } c_{4,2}^{-1.} \quad \begin{bmatrix} 2. & 0 & 0 & 0 \\ 0 & -2. & 0 & 0 \\ 0 & 0 & 0.5 & -1. \\ 0 & 0 & -1. & 2. \end{bmatrix} \qquad \begin{bmatrix} 1. & -1. & 0.5 & 2. \\ 0 & 1. & -0.5 & -1. \\ 0 & 1. & 0.5 & -1. \\ 0 & 0 & 0 & 1. \end{bmatrix}$$

$$\text{con } f_{4,3}^{2.} \quad \begin{bmatrix} 2. & 0 & 0 & 0 \\ 0 & -2. & 0 & 0 \\ 0 & 0 & 0.5 & -1. \\ 0 & 0 & 0 & 0 \end{bmatrix} \qquad \begin{bmatrix} 1. & -1. & 0.5 & 2. \\ 0 & 1. & -0.5 & -1. \\ 0 & 1. & 0.5 & -1. \\ 0 & 0 & 0 & 1. \end{bmatrix}$$

$$\text{con } c_{4,3}^{2.} \quad \begin{bmatrix} 2. & 0 & 0 & 0 \\ 0 & -2. & 0 & 0 \\ 0 & 0 & 0.5 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \qquad \begin{bmatrix} 1. & -1. & 0.5 & 3. \\ 0 & 1. & -0.5 & -2. \\ 0 & 1. & 0.5 & 0 \\ 0 & 0 & 0 & 1. \end{bmatrix}$$

Para que P sea una de las matrices que hace que $P^t A\, P$ sea diagonal, podemos tomar

$$P = \begin{bmatrix} 1. & -1. & 0.5 & 3. \\ 0 & 1. & -0.5 & -2. \\ 0 & 1. & 0.5 & 0 \\ 0 & 0 & 0 & 1. \end{bmatrix}.$$

De proceder realizando con A los cambios de fila (únicamente esos), hubiésemos obtenido

$$P^t = \begin{bmatrix} 1. & 0 & 0 & 0 \\ -1. & 1. & 1. & 0 \\ 0.5 & -0.5 & 0.5 & 0 \\ 3. & -2. & 0 & 1. \end{bmatrix}.$$

En cualquier caso, de esta manera

$$P^t A\, P = \begin{bmatrix} 2. & & & \\ & -2. & & \\ & & 0.5 & \\ & & & 0 \end{bmatrix}.$$

El trabajo con la matriz A de q está concluido.

Falta interpretar el resultado en relación con la forma cuadrática q. La nueva base en la que la matriz de q es diagonal, o sea, la base ortogonal que obtenemos, viene dada por las columnas de P. Es decir, es la base (c_1, c_2, c_3, c_4) formada por los vectores

$$c_1 = (1., 0, 0, 0)\,,\ c_2 = (-1., 1., 1., 0)\,,\ c_3 = (0.5, -0.5, 0.5, 0)\,,\ c_4 = (3., -2., 0, 1.)\,;$$

en esta base la matriz de q es

$$[q, (c_1, c_2, c_3, c_4)] = \begin{bmatrix} 2. & & & \\ & -2. & & \\ & & 0.5 & \\ & & & 0 \end{bmatrix}.$$

Hay que notar que el resultado no coincide con el obtenido en (6.4.6). De hecho hubiera podido ser más dispar. Esto no quiere decir más que la base ortogonal no es única y que métodos diferentes pueden llevar lógicamente a resultados distintos (aunque ambos sean correctos).

6.4.8 Si f es una forma bilineal simétrica sobre un e.v. E de dimensión finita, el teorema (6.4.4) nos garantiza la existencia de una base ortogonal para f, o sea, una base $(a_1, \ldots, a_n)$ tal que $[f, (a_i)]$ es una matriz diagonal. Vamos a ver que es posible hacer ciertas precisiones sobre los elementos diagonales de $[f, (a_i)]$; el resultado es diferente según que $\mathbb{K} = \mathbb{R}$ o que $\mathbb{K} = \mathbb{C}$ y, por lo tanto, trataremos los dos casos de forma separada.

6.4.9 COROLARIO

Sea E un e.v. sobre $\mathbb{R}$ de dimensión $n \neq 0$, $f \in \mathcal{S}_2(E, \mathbb{K})$ una forma bilineal simétrica sobre E y q la forma cuadrática asociada. Sea $(a_1, \ldots, a_n)$ una base ortogonal para f; la matriz $[f, (a_i)]$ es entonces diagonal. Si representamos por s el número de elementos diagonales de $[f, (a_i)]$ que son mayores que cero (> 0) y por t el número de los que son menores que cero (< 0), entonces

$$s + t = \operatorname{rg} f\,.$$

Además, si $(a'_1, \ldots, a'_n)$ es también una base ortogonal para f y s' y t' representan el número de elementos mayores que cero y menores que cero en la diagonal de $[f, (a'_i)]$, tenemos que

$$s = s' \quad \text{y} \quad t = t'\,.$$

(Este resultado es conocido con el nombre de *ley de inercia* de Sylvester [2].)

La primera parte es sencilla; nótese que $[f, (a_i)]$ es una matriz diagonal y que posee exactamente $s + t$ elementos diagonales no nulos, luego

$$s + t = \operatorname{rg}[f, (a_i)] = \operatorname{rg} f\,.$$

Pasemos a la segunda parte; pongamos

$$[f, (a_i)] = \begin{bmatrix} \alpha_1^1 & & \\ & \ddots & \\ & & \alpha_n^n \end{bmatrix} \quad \text{y} \quad [f, (a'_i)] = \begin{bmatrix} {\alpha'}_1^1 & & \\ & \ddots & \\ & & {\alpha'}_n^n \end{bmatrix}.$$

Con el objeto de hacer las notaciones más sencillas supondremos que los s elementos α_i^i que son mayores que cero son los s primeros (respectivamente, los s' elementos ${\alpha'}_i^i$ que son mayores que cero son los s' primeros), y que los t elementos α_i^i que son menores que cero son los t siguientes (respectivamente, los t' elementos ${\alpha'}_i^i$ que son menores que cero son los t' siguientes). En otras palabras, supondremos que

$$\begin{array}{cc} \alpha_1^1 > 0, \ldots, \alpha_s^s > 0 & {\alpha'}_1^1 > 0, \ldots, {\alpha'}_{s'}^{s'} > 0 \\ \alpha_{s+1}^{s+1} < 0, \ldots, \alpha_{s+t}^{s+t} < 0 & {\alpha'}_{s'+1}^{s'+1} < 0, \ldots, {\alpha'}_{s'+t'}^{s'+t'} < 0 \\ \alpha_{s+t+1}^{s+t+1} = \ldots = \alpha_n^n = 0 & {\alpha'}_{s'+t'+1}^{s'+t'+1} = \ldots = {\alpha'}_n^n = 0\,. \end{array}$$

Recuérdese que

$$(\forall i) \quad \alpha_i^i = f(a_i, a_i) \quad \text{y} \quad {\alpha'}_i^i = f(a'_i, a'_i)$$

[2] James Joseph Sylvester, matemático nacido en Londres (1814-1897).

y que
$$i \neq j \quad \Rightarrow \quad f(a_i, a_j) = 0 \quad \text{y} \quad f(a'_i, a'_j) = 0\,.$$
Consideramos los subespacios de E
$$F = \langle a_1, \ldots, a_s \rangle \quad \text{y} \quad G' = \langle a'_{s'+1}, \ldots, a'_{s'+t'}, \ldots, a'_n \rangle\,;$$
es claro que $\dim F = s$ y $\dim G' = n - s'$. Vamos a probar que
$$F \cap G' = \{0\}\,.$$
Sea $x \in F \cap G'$. Por una parte,
$$x = \lambda^1 a_1 + \cdots + \lambda^s a_s\,,$$
luego
$$\begin{aligned} q(x) &= f(x, x) \\ &= f(\lambda^1 a_1 + \cdots + \lambda^s a_s\,, \lambda^1 a_1 + \cdots + \lambda^s a_s) \\ &= (\lambda^1)^2 f(a_1, a_1) + \cdots + (\lambda^s)^2 f(a_s, a_s) \\ &= (\lambda^1)^2 \alpha_1^1 + \cdots + (\lambda^s)^2 \alpha_s^s\,. \end{aligned}$$
Por otra,
$$x = \mu^{s'+1} a'_{s'+1} + \cdots + \mu^{s'+t'+1} a'_{s'+t'+1} + \cdots + \mu^n a'_n\,,$$
luego, con un razonamiento análogo,
$$q(x) = (\mu^{s'+1})^2 \alpha'^{s'+1}_{s'+1} + \cdots + (\mu^{s'+t'+1})^2 \alpha'^{s'+t'+1}_{s'+t'+1} + \cdots + (\mu^n)^2 \alpha'^n_n\,.$$
La primera igualdad implica que $q(x) \geq 0$, y la segunda que $q(x) \leq 0$, luego $q(x) = 0$ y
$$(\lambda^1)^2 \alpha_1^1 + \cdots + (\lambda^s)^2 \alpha_s^s = 0\,;$$
como $\alpha_1^1 > 0, \ldots, \alpha_s^s > 0$, resulta que
$$\lambda^1 = \cdots = \lambda^s = 0$$
y que $x = 0$. Tenemos así que
$$F + G' = F \oplus G'$$
y
$$\dim F \oplus G' = \dim F + \dim G' = s + n - s'\,;$$
como $F \oplus G'$ es una subespacio de E, obtenemos
$$s + n - s' \leq n\,,$$
de donde resulta que
$$s \leq s'\,.$$

El mismo razonamiento con los subespacios

$$G = \langle a_{s+1}, \dots, a_{s+t}, \dots, a_n \rangle \quad \text{y} \quad F' = \langle a'_1, \dots, a'_{s'} \rangle$$

nos lleva a que

$$s' \leq s$$

y por lo tanto a la igualdad $s = s'$.

Como, por otra parte,

$$s + t = \operatorname{rg} f = s' + t' ,$$

se tiene también que $t = t'$ y esto termina la demostración.

6.4.10 DEFINICIÓN. Sea E un espacio vectorial sobre $\mathbb{R}$ de dimensión $n \neq 0$, $f \in \mathcal{S}_2(E, \mathbb{K})$ una forma bilineal simétrica sobre E y q la forma cuadrática asociada. El par

$$(s, t)$$

formado por el número s de elementos mayores que cero (> 0) y el número t de elementos menores que cero (< 0) en la diagonal de la matriz diagonal $[f, (a_i)]$, donde $(a_1, \dots, a_n)$ es una base ortogonal para f, recibe el nombre de *signatura* de f (o de q).

Si la signatura de f es (s, t), entonces

$$\operatorname{rg} f = s + t .$$

6.4.11 Sea $A \in M_{\mathbb{R}}(n)$ una matriz real y simétrica. Sabemos que existe una matriz inversible $P \in M_{\mathbb{R}}(n)$ tal que $P^t A P$ es diagonal (v. 6.4.5). Además, para todas las matrices diagonales de la forma $P^t A P$ (o sea, congruentes con A), los números s y t de términos diagonales mayores que cero (> 0) y menores que cero (< 0) son los mismos. Del par (s, t) decimos entonces que es la *signatura* de A.

Si la signatura de A es (s, t), entonces

$$\operatorname{rg} A = s + t .$$

Nótese que si f es una forma bilineal simétrica sobre un e.v. real E, $(a_1, \dots, a_n)$ una base de E y $A = [f, (a_i)]$, entonces la signatura de A es la signatura de f.

6.4.12 Ejemplo. Para la forma cuadrática q del ejemplo (6.4.6), para su forma polar f y para la matriz

$$A = [q, (e_i)] = \begin{bmatrix} 2. & 2. & 0 & -2. \\ 2. & 2. & -1. & -2. \\ 0 & -1. & 0 & -2. \\ -2. & -2. & -2. & 2. \end{bmatrix} ,$$

la signatura es $(2,1)$ y el rango es 3. Estos hechos son inmediatos a la vista de la matriz

$$[q,(a_i)] = P^t\,A\,P = \begin{bmatrix} 2. & & & \\ & -0.5 & & \\ & & 0.5 & \\ & & & 0 \end{bmatrix},$$

pero no resultan nada evidentes cuando se dispone únicamente de la expresión de q en una base que no es ortogonal (por ejemplo, en la base canónica).

6.4.13 Ejemplo. La signatura de las formas cuadráticas del ejemplo (6.3.2) es $(2,0)$, $(3,0)$ y $(n,0)$ respectivamente.

6.4.14 Ejemplo. La forma cuadrática del ejemplo (6.3.10) posee signatura $(1,1)$.

6.4.15 Sea E un e.v. sobre $\mathbb{R}$, de dimensión $n \neq 0$, $f \in \mathcal{S}_2(E,\mathbb{K})$ una forma bilineal simétrica sobre E y q la forma cuadrática asociada. Representemos por (s,t) la signatura de f y por r $(= s+t)$ su rango. Para una base ortogonal $(b_1,\ldots,b_n)$, convenientemente ordenada, la matriz diagonal

$$[f,(b_i)] = \begin{bmatrix} \beta_1^1 & & \\ & \ddots & \\ & & \beta_n^n \end{bmatrix}$$

verifica que

$$\begin{gathered} \beta_1^1 > 0, \ldots, \beta_s^s > 0 \\ \beta_{s+1}^{s+1} < 0, \ldots, \beta_{s+t}^{s+t} < 0 \\ \beta_{s+t+1}^{s+t+1} = \ldots = \beta_n^n = 0\,. \end{gathered}$$

Definimos entonces una nueva base $(a_1,\ldots,a_n)$ de E poniendo

$$a_i = \begin{cases} b_i/\sqrt{\beta_i^i} & i = 1,\ldots,s\,, \\ b_i/\sqrt{-\beta_i^i} & i = s+1,\ldots,s+t\,, \\ b_i & i = s+t+1,\ldots,n\,. \end{cases}$$

Las raíces cuadradas que figuran en estas expresiones existen (como números reales), puesto que $\beta_i^i > 0$ para los primeros índices y $-\beta_i^i > 0$ para los segundos. Se comprobará sin dificultad que $(a_1,\ldots,a_n)$ es también ortogonal para f. Además, si $i = 1,\ldots,s$,

$$f(a_i,a_i) = f(b_i/\sqrt{\beta_i^i}\,,\, b_i/\sqrt{\beta_i^i}) = \frac{1}{\beta_i^i} f(b_i,b_i) = \frac{1}{\beta_i^i}\,\beta_i^i = 1\,;$$

si $i = s+1,\ldots,s+t$,

$$f(a_i,a_i) = f(b_i/\sqrt{-\beta_i^i}\,,\, b_i/\sqrt{-\beta_i^i}) = \frac{1}{-\beta_i^i} f(b_i,b_i) = \frac{1}{-\beta_i^i}\,\beta_i^i = -1\,;$$

y si $i = s + t + 1, \ldots, n$,

$$f(a_i, a_i) = f(b_i, b_i) = \beta_i^i = 0\,.$$

Entonces la matriz de f en la base $(a_1, \ldots, a_n)$ posee la forma siguiente

$$[f, (a_i)] = \begin{bmatrix} 1 & & & & & & & & \\ & \ddots \; {\scriptstyle s \text{ veces}} & & & & & & & \\ & & 1 & & & & & & \\ & & & -1 & & & & & \\ & & & & \ddots \; {\scriptstyle t \text{ veces}} & & & & \\ & & & & & -1 & & & \\ & & & & & & 0 & & \\ & & & & & & {\scriptstyle n-r \text{ veces}} & \ddots & \\ & & & & & & & & 0 \end{bmatrix},$$

y la expresión de q en dicha base es

$$q(x^1 a_1 + \cdots + x^n a_n) = (x^1)^2 + \cdots + (x^s)^2 - (x^{s+1})^2 - \cdots - (x^{s+t})^2\,.$$

En el caso real, ésta es la mejor respuesta que podemos dar a los planteamientos que hicimos en (6.3.11).

6.4.16 Ejemplo. Consideremos de nuevo la situación de los ejemplos (6.4.6) y (6.4.12). Para la base (b_1, b_2, b_3, b_4), formada por los vectores

$$\begin{aligned} b_1 &= a_1 = (\ \ 1.\ ,\ \ 0\ \ ,\ \ 0\ ,\ 0\ ) \\ b_2 &= a_2 = (\ -0.5,\ \ 0.5,\ \ 0.5,\ 0\ ) \\ b_3 &= a_3 = (\ -0.5,\ \ 0.5,\ -0.5,\ 0\ ) \\ b_4 &= a_4 = (\ \ 3.\ ,\ -2.\ ,\ \ 0\ ,\ 1.\ )\ , \end{aligned}$$

la matriz de q es

$$[q, (b_i)] = \begin{bmatrix} 2. & & & \\ & -0.5 & & \\ & & 0.5 & \\ & & & 0 \end{bmatrix}.$$

Ponemos entonces

$$c_1 = \frac{b_1}{\sqrt{2}} = \left(\frac{1}{\sqrt{2}}, 0, 0, 0\right),$$

$$c_2 = \frac{b_2}{\sqrt{0.5}} = \sqrt{2}\, b_2 = \sqrt{2}\left(-\frac{1}{2}, \frac{1}{2}, -\frac{1}{2}, 0\right) = \left(-\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}, 0\right),$$

$$c_3 = \frac{b_3}{\sqrt{0.5}} = \sqrt{2}\, b_3 = \sqrt{2}\left(-\frac{1}{2}, \frac{1}{2}, \frac{1}{2}, 0\right) = \left(-\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0\right),$$

$$c_4 = b_4 = (3, -2, 0, 1);$$

La base (c_1, c_2, c_3, c_4) es ortogonal y

$$[q, (c_i)] = \begin{bmatrix} 1. & & & \\ & 1. & & \\ & & -1. & \\ & & & 0 \end{bmatrix}.$$

6.4.17 Los resultados son diferentes en el caso complejo; esto es debido a que todo número complejo posee (dos) raíces cuadradas, mientras que sólo los números reales positivos poseen raíces cuadradas reales.

Sea E un e.v. sobre $\mathbb{C}$, de dimensión $n \neq 0$, $f \in \mathcal{S}_2(E, \mathbb{K})$ una forma bilineal simétrica sobre E y q la forma cuadrática asociada. Representemos por r el rango de f. El teorema (6.4.4) nos garantiza la existencia de bases ortogonales para f; en dichas bases la matriz de f es diagonal, y el número de elementos diagonales no nulos es necesariamente r. Para una base ortogonal, $(b_1, \ldots, b_n)$, convenientemente ordenada, la matriz diagonal

$$[f, (b_i)] = \begin{bmatrix} \beta_1^1 & & \\ & \ddots & \\ & & \beta_n^n \end{bmatrix}$$

verifica que

$$\begin{gathered} \beta_1^1 \neq 0, \ldots, \beta_r^r \neq 0 \\ \beta_{r+1}^{r+1} = \ldots = \beta_n^n = 0\,. \end{gathered}$$

Definimos entonces una nueva base $(a_1, \ldots, a_n)$ de E, poniendo

$$a_i = \begin{cases} b_i / \sqrt{\beta_i^i} & i = 1, \ldots, r\,, \\ b_i & i = r+1, \ldots, n\,, \end{cases}$$

donde $\sqrt{\beta_i^i}$ representa cualquiera de las dos raíces cuadradas del número complejo $\beta_i^i \neq 0$. Se comprueba como en (6.4.15) que

$$\begin{aligned} f(a_i, a_i) = 1 \quad &\text{si} \quad i = 1, \ldots, r\,, \\ \text{y} \quad f(a_i, a_i) = 0 \quad &\text{si} \quad i = r+1, \ldots, n\,. \end{aligned}$$

La base $(a_1, \ldots, a_n)$ es ortogonal para f; la matriz de f en esta base es de la forma

$$[f,(a_i)] = \begin{bmatrix} 1 & & & & & \\ & \ddots & {\scriptstyle r \text{ veces}} & & & \\ & & 1 & & & \\ & & & 0 & & \\ & & & {\scriptstyle n-r \text{ veces}} & \ddots & \\ & & & & & 0 \end{bmatrix},$$

y la expresión de q en dicha base es

$$q(x^1 a_1 + \cdots + x^n a_n) = (x^1)^2 + \cdots + (x^r)^2 \,.$$

6.5 Formas bilineales positivas y producto escalar (real).

6.5.1 Uno de los objetivos que se persiguen con el estudio de las formas bilineales simétricas es el de dar una noción de 'longitud' de un vector (la magnitud equivamente al módulo o valor absoluto de un número que, naturalmente, debe ser un número positivo). Si (x^1, x^2) es un vector de $\mathbb{R}^2$, se suele entender por longitud de dicho vector el número real positivo

$$\sqrt{(x^1)^2 + (x^2)^2}\,,$$

es decir,

$$\sqrt{q(x^1, x^2)}\,,$$

donde q representa la forma cuadrática del ejemplo (6.3.2); obsérvese que

$$q(x^1, x^2) = (x^1)^2 + (x^2)^2$$

es siempre un número real igual o mayor que cero, lo que resulta necesario para hablar de su raíz positiva.

Si q es una forma cuadrática sobre un e.v. E y $x \in E$, se puede generalizar esta idea considerando $\sqrt{q(x)}$ como la longitud del vector x. Para ello es necesario que $q(x)$ sea un número real positivo para todos los vectores x de E.

Conviene, pues, estudiar las formas bilineales simétricas $f \in \mathcal{S}_2(E, \mathbb{K})$ tales que

$$(\forall x \in E) \quad f(x,x) = q(x) \geq 0 \,.$$

Para $\mathbb{K} = \mathbb{R}$ existen numerosas formas simétricas que verifican esta propiedad, como veremos más adelante. Para $\mathbb{K} = \mathbb{C}$, la situación es muy diferente. Supongamos que E es un e.v. sobre $\mathbb{C}$ y $f \in \mathcal{S}_2(E, \mathbb{C})$ una forma bilineal y simétrica sobre E tal que

$$(\forall x \in E) \quad f(x,x) = q(x) \geq 0 \,;$$

entonces, para todo $x \in E$,

$$-q(x) = -f(x,x) = i^2 f(x,x) = f(i\,x, i\,x) = q(i\,x) \geq 0\,,$$

luego

$$(\forall x \in E) \quad q(x) \leq 0\,,$$

y resulta que

$$(\forall x \in E) \quad q(x) = 0\,,$$

o sea, $f = q = 0$. En el caso complejo, sólo la forma nula verifica tal propiedad.

Esta sección se limitará entonces al estudio de las formas bilineales simétricas reales que cumplen dicha propiedad. En el caso complejo se puede dar también una noción de longitud de un vector, pero se emplean para ello otras formas diferentes de las bilineales; las estudiaremos en la sección siguiente.

6.5.2 DEFINICIÓN. Sea $f \in \mathcal{S}_2(E, \mathbb{R})$ una forma bilineal simétrica sobre un e.v. real E y q la forma cuadrática asociada. Decimos que f y q son *positivas* cuando

$$\textbf{(pos)} \quad (\forall x \in E) \quad f(x,x) = q(x) \geq 0\,.$$

6.5.3 Ejemplo. Las formas sobre $\mathbb{R}^2$ y $\mathbb{R}^3$ dadas por

$$f(\,(x^1, x^2), (y^1, y^2)\,) = x^1 y^1 + x^2 y^2$$

y

$$f(\,(x^1, x^2, x^3), (y^1, y^2, y^3)\,) = x^1 y^1 + x^2 y^2 + x^3 y^3$$

y más generalmente la forma sobre $\mathbb{R}^n$ dada por

$$f(\,(x^1, \ldots, x^n), (y^1, \ldots, y^n)\,) = x^1 y^1 + \cdots + x^n y^n$$

(v. 6.1.2 y 6.3.2), son positivas ya que la correspondiente forma cuadrática verifica

$$q(x^1, \ldots, x^n) = (x^1)^2 + \cdots + (x^n)^2 \geq 0\,.$$

A pesar de tener el mismo aspecto, la forma cuadrática sobre $\mathbb{C}^2$ dada por

$$q(x^1, x^2) = (x^1)^2 + (x^2)^2$$

no verifica la misma propiedad, puesto que, por ejemplo

$$q(i\,, 0) = i^2 = -1 < 0\,.$$

6.5.4 Ejemplo. La forma f sobre $\mathbb{R}^3$ del ejemplo (6.3.3) no es positiva, puesto que

$$q(0,1,0) = -(1^2) = -1 < 0\,.$$

6.5.5 Ejemplo. La forma sobre l^2 o sobre c_{00}

$$f(x,y) = \sum_{n=1}^{\infty} x_n y_n$$

del ejemplo (6.1.3) es positiva puesto que

$$q(x) = \sum_{n=1}^{\infty} (x_n)^2 \geq 0$$

para toda sucesión $x \in l^2$ (o c_{00}).

6.5.6 Ejemplo. La forma sobre $\mathcal{C}([a,b], \mathbb{R})$

$$P(f,g) = \int_a^b f(t)g(t)\,dt$$

del ejemplo (6.1.4) es positiva. En efecto,

$$P(f,f) = \int_a^b f(t)^2\,dt \geq 0\,,$$

ya que $f(t)^2 \geq 0$ para todo $t \in [a,b]$.

6.5.7 Sea E un e.v. de dimensión $n \neq 0$ sobre $\mathbb{R}$ y $f \in \mathcal{S}_2(E, \mathbb{R})$ una forma bilineal simétrica sobre E. La signature de f (v. 6.4.10) nos proporciona un criterio para saber cuándo f es positiva; la propiedad **(pos)** es equivalente a la siguiente

(pos1) La signatura de f es de la forma $(r,0)$, con $r = \operatorname{rg} f$.

Probaremos esta equivalencia. Si f es positiva, consideramos una base $(a_1, \ldots, a_n)$ ortogonal para f y la correspondiente matriz diagonal

$$[f,(a_i)] = \begin{bmatrix} \alpha_1^1 & & \\ & \ddots & \\ & & \alpha_n^n \end{bmatrix}\;;$$

entonces

$$(\forall i) \quad \alpha_i^i = f(a_i, a_i) \geq 0\,,$$

es decir, no existen términos diagonales < 0, y la signatura de f es $(r,0)$ con $r = \operatorname{rg} f$. Recíprocamente, si la signatura de f es $(r,0)$, existe una base $(a_1, \ldots, a_n)$ para la que

$$q(x^1 a_1 + \cdots + x^n a_n) = (x^1)^2 + \cdots + (x^r)^2$$

(v. 6.4.15); entonces, cualquiera que sea $x \in E$, tenemos $q(x) \geq 0$, luego la forma es positiva.

6.5.8 TEOREMA (desigualdad de Schwarz[3])

Sea f una forma positiva sobre un e.v. real E. Entonces se verifica para todo $x, y \in E$ que

$$(f(x,y))^2 \leq f(x,x)\, f(y,y)\,.$$

Sean $x, y \in E$. Supongamos en primer lugar que $f(x,x) = 0$; cualquiera que sea $\lambda \in \mathbb{R}$, tenemos que

$$0 \leq f(\lambda x + y\,,\, \lambda x + y) = 2\lambda f(x,y) + f(y,y)\,;$$

esto no es posible más que si $f(x,y) = 0$. La desigualdad se cumple en este caso, puesto que ambos miembros valen 0.

Supongamos por el contrario que $f(x,x) \neq 0$; consideremos entonces el vector

$$z = y - \frac{f(x,y)}{f(x,x)}\, x\,.$$

Como f es positiva,

$$0 \leq f(z,z) = f(y,y) - \frac{(f(x,y))^2}{f(x,x)}$$

(compruébese esta última igualdad), y como $f(x,x) > 0$, se obtiene la desigualdad del enunciado.

6.5.9 COROLARIO

Sea f una forma positiva sobre un e.v. real E y q la forma cuadrática asociada. Entonces

$$\operatorname{Ker} f = \{x \in E \mid q(x) = 0\}\,.$$

Si $x \in \operatorname{Ker} f$, entonces (v. 6.2.9)

$$(\forall y \in E) \quad f(x,y) = 0$$

y en particular $f(x,x) = 0$. Recíprocamente, si $f(x,x) = q(x) = 0$, entonces basta aplicar el teorema precedente para ver que

$$(\forall y \in E) \quad f(x,y) = 0\,,$$

luego $x \in \operatorname{Ker} f$.

[3] del matemático alemán Hermann Amandus Schwarz (1843-1921). También es oportuno llamarla desigualdad de Cauchy-Schwarz, en atención al ingeniero y matemático francés Augustin-Louis Cauchy (1789-1857).

6.5.10 DEFINICIÓN. Sea E un espacio vectorial real; un *producto escalar* (real) sobre E es una forma positiva y no degenerada, es decir, una aplicación

$$f : E \times E \to \mathbb{R}$$

que es una forma bilineal simétrica y que verifica

$$(\forall x \in E) \quad f(x,x) \geq 0$$

y

$$\operatorname{Ker} f = \{0\}\,.$$

Si f es una forma bilineal y simétrica sobre E, entonces f es un producto escalar si y sólo si

$$\textbf{(ps)} \quad x \neq 0 \quad \Rightarrow \quad f(x,x) > 0\,.$$

En efecto, si f verifica la propiedad **(ps)**, entonces

$$(\forall x \in E) \quad f(x,x) \geq 0\,,$$

puesto que $f(0,0) = 0$, luego f es positiva; además, si $x \in \operatorname{Ker} f$, resulta del corolario precedente que $f(x,x) = q(x) = 0$, luego $x = 0$ por la propiedad **(ps)**; esto demuestra que $\operatorname{Ker} f = \{0\}$ y que f es no degenerada. Recíprocamente, si f es un producto escalar, entonces f es positiva y en particular

$$(\forall x \in E) \quad f(x,x) \geq 0\,;$$

pero f es también no degenerada, luego, si $x \neq 0$, $x \notin \operatorname{Ker} f$, lo que significa según el corolario precedente que $f(x,x) = q(x) \neq 0$; resulta así que

$$x \neq 0 \quad \Rightarrow \quad f(x,x) > 0\,.$$

6.5.11 Sea E un e.v. real de dimensión $n \neq 0$ y $f \in \mathcal{S}_2(E, \mathbb{R})$ una forma bilineal simétrica sobre E. Entonces f es un producto escalar (real) si y sólo si su signatura es $(n, 0)$, como se deduce inmediatamente de (6.5.7).

6.5.12 Ejemplo. La forma sobre l^2 o sobre c_{00}

$$f(x,y) = \sum_{n=1}^{\infty} x_n y_n$$

es un producto escalar, ya que es simétrica, no degenerada y positiva (v. 6.2.17 y 6.5.5).

6.5.13 Ejemplo. De la forma sobre $\mathcal{C}([a,b],\mathbb{R})$

$$P(f,g)=\int_a^b f(t)g(t)\,dt$$

sabemos que es positiva (v. 6.5.6). Veamos que se trata de un producto escalar. Sea $f\in\mathcal{C}([a,b],\mathbb{R})$, $f\neq 0$; la función

$$t\to f(t)^2$$

es igual o mayor que cero (≥ 0). Pero además, como $f\neq 0$, existe un punto $t_0\in[a,b]$ donde $f(t_0)\neq 0$ y en consecuencia $f(t_0)^2>0$. La continuidad de la función $t\to f(t)^2$ nos permite afirmar la existencia de un intervalo $[a',b']$ con $a'<b'$ y

$$t_0\in[a',b']\subset[a,b]$$

y de un número $\alpha>0$ tal que

$$(\forall\, t\in[a',b'])\quad f(t)^2\geq\alpha\,.$$

Resulta entonces que

$$P(f,f)=\int_a^b f(t)^2\,dt\geq\int_{a'}^{b'} f(t)^2\,dt\geq\int_{a'}^{b'}\alpha\,dt=(b'-a')\alpha>0\,;$$

esto demuestra que P es un producto escalar.

6.5.14 Ejemplo. La forma positiva sobre $\mathbb{R}^n$

$$f(\,(x^1,\ldots,x^n),(y^1,\ldots,y^n)\,)=x^1y^1+\cdots+x^ny^n$$

(v. 6.5.3), es un producto escalar (real). Diremos que es el *producto escalar ordinario* de $\mathbb{R}^n$, porque es el que se maneja habitualmente. El producto escalar ordinario de $\mathbb{R}^n$ se representa en la base canónica de $\mathbb{R}^n$ por la matriz unidad I_n .

El teorema que sigue demuestra que todo producto escalar (real) sobre un espacio de dimensión finita se representa por la matriz unidad en una base adecuada. Lo que proporciona un interés especial al producto escalar ordinario de $\mathbb{R}^n$ es que la base en que tal cosa ocurre es la base canónica de $\mathbb{R}^n$.

6.5.15 TEOREMA

Sea E un e.v. sobre $\mathbb{R}$ de dimensión $n\neq 0$ y f un producto escalar (real) sobre E. Existe una base $(a_1,\ldots,a_n)$ de E tal que

$$[f,(a_i)]=I_n\,.$$

La expresión de f en dicha base es entonces

$$f(x,y)=x^1y^1+\cdots+x^ny^n$$

para $x=(x^1,\ldots,x^n)_{(a_i)}$ e $y=(y^1,\ldots,y^n)_{(a_i)}$.

Como hemos visto en (6.5.11), la signatura de f es $(n,0)$; basta entonces aplicar el resultado de (6.4.15).

6.6 Matrices positivas y estrictamente positivas.

6.6.1 El objetivo fundamental de esta sección es obtener un criterio práctico para reconocer un producto escalar sobre un espacio de dimensión finita mediante el examen de la matriz que lo representa en una base cualquiera del espacio.

6.6.2 DEFINICIÓN. Sea $A \in M_{\mathbb{C}}(n)$ una matriz cuadrada de números complejos. Decimos que A es *positiva* cuando

$$(\forall x \in \mathbb{C}^n) \quad x^* A x \geq 0\,,$$

y que A es *estrictamente positiva* cuando

$$(\forall x \in \mathbb{C}^n, x \neq 0) \quad x^* A x > 0\,.$$

Es evidente que toda matriz estrictamente positiva es en particular positiva.

6.6.3 Si A es una matriz cuadrada de números reales, entonces son equivalentes las dos propiedades siguientes:

$$\begin{aligned} &\text{(i) } A \text{ es positiva}\,, \\ &\text{(ii) } A = A^t \text{ y } (\forall x \in \mathbb{R}^n) \quad x^t A x \geq 0\,; \end{aligned}$$

también son equivalentes las dos siguientes:

$$\begin{aligned} &\text{(i) } A \text{ es estrictamente positiva}\,, \\ &\text{(ii) } A = A^t \text{ y } (\forall x \in \mathbb{R}^n, x \neq 0) \quad x^t A x > 0\,. \end{aligned}$$

La implicación (i)⇒(ii) es en ambos casos una consecuencia directa de la definición y de (6.3.3) (recuérdese que una matriz real es hermítica si y sólo si es simétrica).

Veamos que (ii)⇒(i). Supongamos que $A = A^t$; si $z \in \mathbb{C}^n$, extonces $z = x + i\,y$ con $x, y \in \mathbb{R}^n$; por lo tanto

$$\begin{aligned} z^* A z &= (x + i\,y)^* A(x + i\,y) \\ &= (x - i\,y)^t A(x + i\,y) \\ &= (x^t - i\,y^t) A(x + i\,y) \\ &= x^t A x + y^t A y + i\,(x^t A y - y^t A x)\,. \end{aligned}$$

Ahora bien, la matriz $y^t A x$ es 1×1 y coincide con su traspuesta, luego

$$y^t A x = (y^t A x)^t = x^t A^t y = x^t A y$$

y resulta que

$$z^* A z = x^t A x + y^t A y\,.$$

Entonces, si $(\forall x \in \mathbb{R}^n)$ $x^t A x \geq 0$, resulta que $(\forall z \in \mathbb{C}^n)$ $z^* A z \geq 0$, lo que prueba que (ii)⇒(i) en el primer caso. Si $(\forall x \in \mathbb{R}^n, x \neq 0)$ $x^t A x > 0$, entonces para $z \in \mathbb{C}^n$ tal que $z \neq 0$, o bien $x \neq 0$, o bien $y \neq 0$, y en cualquier caso $z^* A z > 0$, lo que prueba que (ii)⇒(i) en el segundo caso.

6.6.4 PROPOSICIÓN

Si A es una matriz positiva (respectivamente, estrictamente positiva), sus elementos diagonales son todos números reales y, además, iguales o mayores que cero (resp. mayores que cero).

Pongamos $A = [\alpha_i^j]$. Representando por $(e_1, \ldots, e_n)$ la base canónica de $\mathbb{C}^n$ tenemos que

$$e_i^* A e_i = \alpha_i^i\,;$$

si A es positiva, entonces $\alpha_i^i \geq 0$, y si A es estrictamente positiva, entonces $\alpha_i^i > 0$.

6.6.5 COROLARIO

Una matriz diagonal es positiva (respectivamente, estrictamente positiva), si y sólo si sus elementos diagonales son todos números reales iguales o mayores que cero (resp. mayores que cero).

Una parte del enunciado es consecuencia inmediata de la proposición anterior. Veamos la otra; si

$$A = \begin{bmatrix} \alpha_1^1 & & \\ & \ddots & \\ & & \alpha_n^n \end{bmatrix}$$

es diagonal y $x = (x^1, \ldots, x^n) \in \mathbb{C}^n$, tenemos que

$$\begin{aligned} x^* A x &= \begin{bmatrix} \overline{x^1} & \cdots & \overline{x^n} \end{bmatrix} \begin{bmatrix} \alpha_1^1 & & \\ & \ddots & \\ & & \alpha_n^n \end{bmatrix} \begin{bmatrix} x^1 \\ \vdots \\ x^n \end{bmatrix} \\ &= \alpha_1^1\, x^1\, \overline{x^1} + \cdots + \alpha_n^n\, x^n\, \overline{x^n} \\ &= \alpha_1^1 |x^1|^2 + \cdots + \alpha_n^n |x^n|^2\,. \end{aligned}$$

Si los α_i^i son iguales o mayores que cero, resulta que $x^* A x \geq 0$. Si los α_i^i son mayores que cero y $x \neq 0$, resulta que $x^* A x > 0$.

6.6.6 Salvo los dos resultados precedentes, poco se puede decir acerca de los elementos de las matrices positivas. Los ejemplos que siguen muestran que no se debe confundir matriz positiva con matriz de elementos positivos.

6.6.7 Ejemplo. La matriz

$$\begin{bmatrix} 1 & i \\ -i & 1 \end{bmatrix}$$

es positiva. En efecto, si $(x^1, x^2) \in \mathbb{C}^2$, entonces

$$\begin{aligned}
\begin{bmatrix} \overline{x^1} & \overline{x^2} \end{bmatrix} \begin{bmatrix} 1 & i \\ -i & 1 \end{bmatrix} \begin{bmatrix} x^1 \\ x^2 \end{bmatrix} &= \begin{bmatrix} \overline{x^1} & \overline{x^2} \end{bmatrix} \begin{bmatrix} x^1 + i\,x^2 \\ -i\,x^1 + x^2 \end{bmatrix} \\
&= x^1\,\overline{x^1} + i\,x^2\,\overline{x^1} - i\,x^1\,\overline{x^2} + x^2\,\overline{x^2} \\
&= \overline{x^1}(x^1 + i\,x^2) - i\,\overline{x^2}(x^1 + i\,x^2) \\
&= (x^1 + i\,x^2)\,(\overline{x^1} - i\,\overline{x^2}) \\
&= (x^1 + i\,x^2)\,\overline{(x^1 + i\,x^2)} \\
&= |x^1 + i\,x^2|^2 \\
&\geq 0\,.
\end{aligned}$$

La matriz

$$\begin{bmatrix} 1 & i \\ -i & 2 \end{bmatrix}$$

es estrictamente positiva, ya que, si $(x^1, x^2) \in \mathbb{C}^2$, entonces

$$\begin{aligned}
\begin{bmatrix} \overline{x^1} & \overline{x^2} \end{bmatrix} \begin{bmatrix} 1 & i \\ -i & 2 \end{bmatrix} \begin{bmatrix} x^1 \\ x^2 \end{bmatrix} &= x^1\,\overline{x^1} + i\,x^2\,\overline{x^1} - i\,x^1\,\overline{x^2} + 2\,x^2\,\overline{x^2} \\
&= |x^1 + i\,x^2|^2 + |x^2|^2\,.
\end{aligned}$$

Si $|x^1 + i\,x^2|^2 + |x^2|^2 = 0$, entonces $x^2 = 0$ y $x^1 = -i\,x^2 = 0$, luego $(x^1, x^2) = 0$. Es decir, si $(x^1, x^2) \neq 0$, entonces

$$\begin{bmatrix} \overline{x^1} & \overline{x^2} \end{bmatrix} \begin{bmatrix} 1 & i \\ -i & 2 \end{bmatrix} \begin{bmatrix} x^1 \\ x^2 \end{bmatrix} > 0\,.$$

6.6.8 Ejemplo. La matriz real

$$\begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix}$$

es estrictamente positiva, ya que es simétrica y, si $(x^1, x^2) \in \mathbb{R}^2$, $(x^1, x^2) \neq 0$, entonces

$$\begin{aligned}
\begin{bmatrix} x^1 & x^2 \end{bmatrix} \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix} \begin{bmatrix} x^1 \\ x^2 \end{bmatrix} &= \begin{bmatrix} x^1 & x^2 \end{bmatrix} \begin{bmatrix} 2x^1 - x^2 \\ -x^1 + 2x^2 \end{bmatrix} \\
&= 2(x^1)^2 - x^1x^2 - x^1x^2 + 2(x^2)^2 \\
&= (x^1 - x^2)^2 + (x^1)^2 + (x^2)^2 \\
&> 0\,.
\end{aligned}$$

6.6.9 Ejemplo. La matriz

$$\begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}$$

no es positiva a pesar de su 'aspecto positivo', ya que

$$\begin{bmatrix} 1 & -1 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ -1 \end{bmatrix} = \begin{bmatrix} 1 & -1 \end{bmatrix} \begin{bmatrix} -1 \\ 1 \end{bmatrix} = -2 < 0 \,.$$

6.6.10 TEOREMA

Sea E un e.v. sobre $\mathbb{R}$ (sobre $\mathbb{C}$) de dimensión $n \neq 0$, $(a_1, \ldots, a_n)$ una base de E, f una forma bilineal (sesquilineal) sobre E y $[f, (a_i)]$ la matriz de f en dicha base.

a) La forma f es positiva si y sólo si la matriz $[f, (a_i)]$ es positiva.

b) La forma f es un producto escalar si y sólo si la matriz $[f, (a_i)]$ es estrictamente positiva.

Pondremos por comodidad $A = [f, (a_i)]$.

a) Caso real: si f es positiva, es simétrica (por definición), luego $A = A^t$; además

$$(\forall x \in E) \quad [x, (a_i)]^{\,t} A \,[x, (a_i)] = f(x, x) \geq 0 \,,$$

luego

$$(\forall x \in \mathbb{R}^n) \quad x^t A \, x \geq 0 \,;$$

resulta entonces de (6.6.3) que A es positiva.

Recíprocamente, si A es positiva, entonces $A = A^t$ y

$$(\forall x \in \mathbb{R}^n) \quad x^t A \, x \geq 0 \,.$$

Esto significa que f es simétrica y que

$$(\forall x \in E) \quad f(x, x) = [x, (a_i)]^{\,t} A \,[x, (a_i)] \geq 0 \,,$$

luego f es positiva.

Caso complejo: si f es positiva, entonces

$$(\forall x \in E) \quad [x, (a_i)]^{\,*} A \,[x, (a_i)] = f(x, x) \geq 0 \,,$$

luego

$$(\forall x \in \mathbb{C}^n) \quad x^* A \, x \geq 0$$

y A es positiva.

Recíprocamente, si A es positiva, entonces

$$(\forall x \in \mathbb{C}^n) \quad x^* A \, x \geq 0 \,,$$

luego

$$(\forall x \in E) \quad f(x, x) = [x, (a_i)]^{\,*} A \,[x, (a_i)] \geq 0$$

y f es positiva.

b) La demostración sigue paso a paso la del apartado precedente.

6.6.11 Para saber si una forma f es un producto escalar basta entonces comprobar si su matriz en una base cualquiera es estrictamente positiva. Los resultados que siguen nos llevarán a un criterio práctico para saber si una matriz es estrictamente positiva.

6.6.12 PROPOSICIÓN

Sea $A \in M_{\mathbb{C}}(n)$ una matriz cuadrada de elementos complejos. La matriz A es estrictamente positiva si y sólo si existe una matriz $P \in M_{\mathbb{C}}(n)$ inversible y tal que

$$A = P^* P .$$

Supongamos que A es estrictamente positiva. La forma sesquilineal f sobre $\mathbb{C}^n$ dada por A, esto es, tal que $[f, (e_i)] = A$, es un producto escalar. Por lo tanto, existe una base $(a_1, \ldots, a_n)$ de $\mathbb{C}^n$ tal que

$$[f, (a_i)] = I_n .$$

A e I_n estarán relacionadas por

$$A = P^* I_n P = P^* P ,$$

donde P es la matriz de paso de $(a_1, \ldots, a_n)$ a la base canónica $(e_1, \ldots, e_n)$; la matriz P es además inversible.

Recíprocamente, supongamos ahora que

$$A = P^* P$$

con P inversible. Si $x \in \mathbb{C}^n$, entonces

$$x^* A x = x^* (P^* P) x = (P x)^* P x .$$

Si $x \neq 0$, entonces $P x \neq 0$, pues en caso contrario $x = P^{-1} 0 = 0$; esto es,

$$P x = \begin{bmatrix} y^1 \\ \vdots \\ y^n \end{bmatrix}$$

con algún $y^i \neq 0$. Resulta entonces que

$$x^* A x = \begin{bmatrix} \overline{y^1} & \cdots & \overline{y^n} \end{bmatrix} \begin{bmatrix} y^1 \\ \vdots \\ y^n \end{bmatrix} = |y^1|^2 + \cdots + |y^n|^2 > 0 ,$$

luego A es estrictamente positiva.

6.6.13 Si A es una matriz estrictamente positiva y de elementos reales, podemos asegurar que existe $P \in M_{\mathbb{R}}(n)$ inversible y tal que

$$A = P^t \, P \, .$$

Basta repetir la primera parte de la demostración precedente, con $\mathbb{R}^n$ en lugar de $\mathbb{C}^n$.

6.6.14 COROLARIO

Si A es una matriz estrictamente positiva, entonces

$$\det A > 0 \, .$$

Como $A = P^* \, P$, con P inversible, resulta que

$$\det A = \det P^* \, \det P = \overline{\det P} \, \det P = |\det P|^2 > 0 \, .$$

6.6.15 Para enunciar el teorema de caracterización de las matrices estrictamente positivas nos será útil la definición siguiente. Siendo $A = [\alpha_i^j]$ una matriz cuadrada $n \times n$, consideramos las matrices extraídas formadas por las k primeras columnas y las k primeras filas,

$$A_k = \begin{bmatrix} \alpha_1^1 & \cdots & \alpha_k^1 \\ \vdots & & \vdots \\ \alpha_1^k & \cdots & \alpha_k^k \end{bmatrix} , \qquad k = 1, \ldots, n \, ,$$

que son matrices $k \times k$. Los determinantes

$$\det A_k \, , \qquad k = 1, \ldots, n \, ,$$

reciben el nombre de *menores principales* de A.

6.6.16 **Ejemplo.** Los menores principales de la matriz

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

son

$$\det A_1 = 2. \, ,$$

$$\det A_2 = \begin{vmatrix} 2. & -1. \\ -1. & 1. \end{vmatrix} = 1. \, ,$$

$$\det A_3 = \begin{vmatrix} 2. & -1. & 1. \\ -1. & 1. & 0 \\ 1. & 0 & 3. \end{vmatrix} = 2.$$

6.6.17 TEOREMA

Sea $A \in M_{\mathbb{C}}(n)$ una matriz cuadrada de elementos complejos. La matriz A es estrictamente positiva si y sólo si $A = A^*$ y los menores principales de A son todos mayores que cero (> 0).

Supondremos en primer lugar que A es estrictamente positiva. Ya sabemos que entonces $A = A^*$. La forma sesquilineal f sobre $\mathbb{C}^n$ dada por A es un producto escalar; recordemos que $[f, (e_i)] = A$. Si $k = 1, 2, \ldots, n$, denotemos por A_k la matriz extraída de A formada por las k primeras columnas y las k primeras filas. Consideramos el subespacio

$$F_k = \langle e_1, \ldots, e_k \rangle$$

de $\mathbb{C}^n$ y la restricción

$$\begin{aligned} f_k : F_k \times F_k &\to \mathbb{C} \\ (x, y) &\to f_k(x, y) = f(x, y) \end{aligned}$$

de f a $F_k \times F_k$. Entonces f_k es un producto escalar sobre F_k y

$$[f_k, (e_1, \ldots, e_k)] = A_k ,$$

luego A_k es estrictamente positiva y por lo tanto $\det A_k > 0$ (v. 6.6.14).

Demostraremos ahora que toda matriz A tal que $A = A^*$ y que sus menores principales son mayores que cero (> 0) es estrictamente positiva. Utilizaremos el principio de inducción sobre la dimensión de A. Si A es 1×1 y cumple dichas condiciones, entonces $A = [\alpha]$ con $\alpha > 0$, y es sencillo ver que A es estrictamente positiva. Suponemos ahora que toda matriz $(n-1) \times (n-1)$ que cumple dichas condiciones es estrictamente positiva; sea entonces A una matriz $n \times n$ tal que $A = A^*$ y que sus menores principales son todos mayores que cero (> 0). Pongamos

$$A = \begin{bmatrix} \alpha_1^1 & \cdots & \alpha_n^1 \\ \vdots & & \vdots \\ \alpha_1^n & \cdots & \alpha_n^n \end{bmatrix} ;$$

la matriz extraída

$$A_{n-1} = \begin{bmatrix} \alpha_1^1 & \cdots & \alpha_{n-1}^1 \\ \vdots & & \vdots \\ \alpha_1^{n-1} & \cdots & \alpha_{n-1}^{n-1} \end{bmatrix}$$

cumple evidentemente las mismas condiciones y es, por la hipótesis de inducción, estrictamente positiva. La proposición (6.6.12) nos asegura entonces que existe $Q \in M_{\mathbb{C}}(n-1)$ inversible y tal que

$$A_{n-1} = Q^* Q .$$

Nuestro objetivo consiste en encontrar números complejos

$$\lambda_1, \ldots, \lambda_{n-1}, \lambda_n$$

tales que la matriz

$$P = \left[\begin{array}{c|c} & \lambda_1 \\ Q & \vdots \\ & \lambda_{n-1} \\ \hline 0 \ \cdots \ 0 & \lambda_n \end{array} \right]$$

sea inversible y que

$$P^* P = A \, ,$$

o sea,

$$\left[\begin{array}{c|c} & 0 \\ Q^* & \vdots \\ & 0 \\ \hline \overline{\lambda_1} \ \cdots \ \overline{\lambda_{n-1}} & \overline{\lambda_n} \end{array} \right] \left[\begin{array}{c|c} & \lambda_1 \\ Q & \vdots \\ & \lambda_{n-1} \\ \hline 0 \ \cdots \ 0 & \lambda_n \end{array} \right] = \left[\begin{array}{c|c} & \alpha_n^1 \\ A_{n-1} & \vdots \\ & \alpha_n^{n-1} \\ \hline \alpha_1^n \ \cdots \ \alpha_{n-1}^n & \alpha_n^n \end{array} \right] .$$

Existen números complejos $\lambda_1, \ldots, \lambda_{n-1}$ únicos tales que

$$Q^* \begin{bmatrix} \lambda_1 \\ \vdots \\ \lambda_{n-1} \end{bmatrix} = \begin{bmatrix} \alpha_n^1 \\ \vdots \\ \alpha_n^{n-1} \end{bmatrix} ,$$

ya que Q^* es inversible y el sistema admite una única solución. Además, dichos números complejos verifican también que

$$\begin{bmatrix} \overline{\lambda_1} & \cdots & \overline{\lambda_{n-1}} \end{bmatrix} Q = (Q^* \begin{bmatrix} \lambda_1 \\ \vdots \\ \lambda_{n-1} \end{bmatrix})^* = \begin{bmatrix} \alpha_n^1 \\ \vdots \\ \alpha_n^{n-1} \end{bmatrix}^* = \begin{bmatrix} \alpha_1^n & \cdots & \alpha_{n-1}^n \end{bmatrix} ,$$

ya que $A = A^*$. Consideremos también el número complejo μ dado por

$$\mu = \alpha_n^n - |\lambda_1|^2 - \cdots - |\lambda_{n-1}|^2 \, ;$$

entonces

$$|\lambda_1|^2 + \cdots + |\lambda_{n-1}|^2 + \mu = \alpha_n^n \, ,$$

luego

$$\left[\begin{array}{c|c} & 0 \\ Q^* & \vdots \\ & 0 \\ \hline \overline{\lambda_1} \ \cdots \ \overline{\lambda_{n-1}} & 1 \end{array} \right] \left[\begin{array}{c|c} & \lambda_1 \\ Q & \vdots \\ & \lambda_{n-1} \\ \hline 0 \ \cdots \ 0 & \mu \end{array} \right] = \left[\begin{array}{c|c} & \alpha_n^1 \\ A_{n-1} & \vdots \\ & \alpha_n^{n-1} \\ \hline \alpha_1^n \ \cdots \ \alpha_{n-1}^n & \alpha_n^n \end{array} \right] = A \, .$$

Se tiene entonces que

$$\mu |\det Q|^2 = \mu\, \overline{\det Q}\, \det Q = \mu \det Q^* \det Q = \det A > 0\,,$$

y esto implica que $\mu > 0$. Ponemos entonces

$$\lambda_n = \sqrt{\mu}\,;$$

λ_n es real, mayor que cero y $\overline{\lambda_n} = \lambda_n$, luego $\lambda_n \overline{\lambda_n} = \mu$. Para los números complejos

$$\lambda_1, \ldots, \lambda_{n-1}, \lambda_n\,,$$

así elegidos, la matriz

$$P = \left[\begin{array}{c|c} & \lambda_1 \\ Q & \vdots \\ & \lambda_{n-1} \\ \hline 0 \;\cdots\; 0 & \lambda_n \end{array}\right]$$

es inversible y $P^* P = A$. Esto demuestra (v. 6.6.12) que A es estrictamente positiva, lo que termina la prueba del teorema.

6.6.18 Ejemplo. Se puede comprobar utilizando el teorema precedente que las matrices

$$\begin{bmatrix} 1 & i \\ -i & 2 \end{bmatrix} \quad \text{y} \quad \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix}$$

de (6.6.7) y (6.6.8) son estrictamente positivas.

6.6.19 Ejemplo. La matriz del ejemplo (6.6.16) es estrictamente positiva, ya que es real y simétrica (luego hermítica) y sus menores principales son todos mayores que cero (> 0).

Libros cuya lectura se recomienda.

Se me ocurren al menos dos razones para convencer al estudiante de la necesidad de simultanear el estudio del presente libro con otras lecturas. En primer lugar, el tema objeto del libro no se agota con lo escrito en los ocho capítulos precedentes; por el contrario, es muchísimo más amplio. Además, y esta es la segunda razón, es frecuente que conceptos o demostraciones difíciles se asimilen más fácilmente cuando se leen diferentes versiones sobre los mismos.

Si el lector desea tener otra visión de los temas relativos a la teoría de conjuntos, puede acudir al libro

Paul R. HALMOS [HALMOS 1]
Teoría intuitiva de los conjuntos
CECSA (México) (1982),

en el que la exposición es excelente, si bien la presentación y traducción lo convierten en un libro de lectura poco agradable. Para quien no comprenda bien estas cuestiones, la lectura de la primera parte de

Michael SPIVAK [SPIVAK]
Calculus. Cálculo Infinitesimal
REVERTÉ (Barcelona) (1987)

podría ser de gran utilidad.

Estos mismos temas y además los relativos a grupos, anillos, cuerpos, polinomios, etc. se encuentran asimismo en los textos de QUEYSANNE y GODEMENT que citamos más adelante.

Entre la gran cantidad existente de textos que abordan el álgebra lineal desde diferentes puntos de vista y con objetivos distintos, he preferido seleccionar los que me parecen más convenientes de aquellos que están publicados en castellano y que serán, por consiguiente, más accesibles al lector. De un nivel similar al nuestro y cubriendo buena parte de los temas que hemos visto, conviene destacar

Juan de BURGOS [BURGOS]
Álgebra lineal
McGRAW-HILL (Madrid) (1996),

Roger GODEMENT [GODEMENT]
Álgebra
TECNOS (Madrid) (1974),

Kenneth HOFFMAN y **Ray KUNZE** [HOFFMAN-KUNZE]
Álgebra lineal
PRENTICE / HALL HISPANOAMERICANA (México) (1984) ,

Michel QUEYSANNE [QUEYSANNE]
Álgebra básica
VICENS-VIVES (Barcelona) (1979) ,

Gilbert STRANG [STRANG]
Álgebra lineal y sus aplicaciones
ADDISON-WESLEY IBEROAMERICANA (Madrid) (1989) .

Cualquiera de ellos puede servir para realizar una lectura algo diferente de los temas básicos del libro.

Veamos ahora dónde podrá el lector ampliar sus conocimientos sobre algunas secciones del libro. Nosotros no hemos tocado los aspectos numéricos de la resolución de sistemas lineales; si lo desea, el lector puede acudir a alguno de los numerosos libros de cálculo numérico. Por ejemplo, los libros

Richard L. BURDEN y **J. Douglas FAIRES** [BURDEN-FAIRES]
Análisis numérico
GRUPO EDITORIAL IBEROAMÉRICA (México) (1996) ,

S. D. CONTE y **Carl DE BOOR** [CONTE-DE BOOR]
Análisis numérico elemental: un enfoque algorítmico
McGRAW-HILL (México) (1985) ,

son de nivel introductorio y pueden leerse sin mayores problemas.

Es posible que el lector esté interesado en ampliar sus conocimientos sobre la forma canónica de Jordan y sus aplicaciones. Si desea conocer algo más sobre el polinomio minimal y otras cuestiones teóricas, puede consultar el libro [HOFFMAN-KUNZE] ya citado, y también el de

A. I. MALTSEV [MALTSEV]
Fundamentos de Álgebra lineal
MIR (Moscú) (1978) ,

que proporciona una visión bastante completa. Otra dirección posible es la de las aplicaciones de la forma canónica al cálculo de funciones de una matriz y a la resolución de sistemas diferenciales lineales. Muchos libros de ecuaciones diferenciales tratan estos temas. Naturalmente, no puedo olvidar

Jesús ROJO [ROJO]
Ecuaciones y sistemas diferenciales lineales. Una introducción
A C (Madrid) (1991) ,

y asimismo el más voluminoso y completo

Sylvia NOVO , **Rafael OBAYA** y **Jesús ROJO** [NOVO-OBAYA-ROJO]
Ecuaciones y sistemas diferenciales
McGRAW-HILL (Madrid) (1995) .

Pero también conviene citar referencias de otros autores, como

Francisco MARCELLÁN, **Luis CASASÚS** y **Alejandro ZARZO** [MARCELLÁN-CASASÚS-ZARZO]
Ecuaciones diferenciales. Problemas lineales y aplicaciones
McGRAW-HILL (Madrid) (1991),

Morris W. HIRSCH y **Stephen SMALE** [HIRSCH-SMALE]
Ecuaciones diferenciales, sistemas dinámicos y álgebra lineal
ALIANZA EDITORIAL (Madrid) (1983).

Los mismos temas matriciales y también muchos otros forman el contenido del extraordinario libro

Richard E. BELLMAN [BELLMAN]
Introducción al Análisis Matricial
REVERTÉ (Barcelona) (1965),

que constituye un lugar de consulta obligado para quien se interese por la utilización de la teoría de matrices.

Los temas que nosotros abordamos en la parte final se pueden ver en los capítulos 1 y 2 de

George BACHMAN y **Lawrence NARICI** [BACHMAN-NARICI]
Análisis Funcional
TECNOS (Madrid) (1981),

donde, en unas cincuenta páginas, se hace una exposición bastante clara (el resto del libro aborda temas de nivel más avanzado). También se pueden ver en

Paul R. HALMOS [HALMOS 2]
Espacios vectoriales de dimensión finita
CECSA (México) (1971),

con los defectos (imputables al editor) a que aludíamos al comentar el libro de teoría de conjuntos del mismo autor y editor. Sobrepasar el simple nivel introductorio en el que nosotros nos hemos quedado requiere consultar algún libro de análisis funcional. El [BACHMAN-NARICI], ya citado, puede servir, aunque mencionaremos también

Balmohan V. LIMAYE [LIMAYE]
Functional Analysis
WILEY / WILEY EASTERN LIMITED (New Delhi) (1981),

Arch W. NAYLOR and **George R. SELL** [NAYLOR-SELL]
Linear Operator Theory in Engineering and Science
SPRINGER (New York) (1982),

advirtiendo al lector que son libros, sobre todo el primero, de un nivel muy superior al de los primeros cursos de la licenciatura. Otra posibilidad interesante es aprender algo sobre la utilización práctica de las ideas que se han estudiado en el último capítulo del libro. Para ello,

Francis B. HILDEBRAND [HILDEBRAND]
Métodos de la Matemática Aplicada
EUDEBA (Buenos Aires) (1973)

se puede leer sin excesivas dificultades; en

Hervé REINHARD [REINHARD]
Equations differentielles. Fondements et applications
DUNOD (París) (1989)

se pueden ver técnicas espectrales para la resolución de ecuaciones diferenciales e integrales, aunque a un nivel superior.

Los libros [BURGOS], [GODEMENT], [QUEYSANNE], [HOFFMAN-KUNZE] y [HALMOS 2], que hemos citado antes, poseen una buena colección de ejercicios, aunque no se trate de ejercicios resueltos. Existen el castellano bastantes libros de problemas resueltos; sin embargo, algunos de los más populares entre los estudiantes de carreras técnicas no resultarán excesivamente provechosos al lector de este libro.

Naturalmente, mi libro favorito de problemas es

Jesús ROJO e **Isabel MARTÍN** [ROJO-MARTÍN]
Ejercicios y problemas de álgebra lineal
McGRAW-HILL (Madrid) (1994).

Contiene una buena colección de problemas siguiendo el mismo programa que el presente texto. Los enunciados están acompañados por un resumen teórico y seguidos de las soluciones detalladas. Otros libros de ejercicios resueltos de los que el lector puede obtener provecho son

I. PROSKURIAKOV [PROSKURIAKOV]
Problemas de álgebra lineal
MIR (Moscú) (1986),

H. D. IKRAMOV [IKRAMOV]
Problemas de álgebra lineal
MIR (Moscú) (1990).

Ambos poseen muchísimos ejercicios de carácter numérico y también problemas teóricos y semiteóricos. Los dos contienen soluciones numéricas de los ejercicios y también indicaciones para resolver la mayor parte de los problemas más difíciles.