Manual de

GEOMETRÍA I

Miguel Sánchez Caja

Depto. Geometría y Topología
Universidad de Granada

Año 2022

Nota introductoria

El presente *Manual de Geometría I* desarrolla los contenidos usuales de Álgebra Lineal en el primer cuatrimestre del primer curso de los grados de Matemáticas y Física, con el nivel de rigor, abstracción y generalidad propio de estos grados.

Para adaptar su contenido a los estudiantes que ingresan por primera vez en la Universidad, el Tema 1 incluye preliminares en los que cada estudiante tendrá que detenerse más o menos según su formación. Las cinco primeras secciones de este tema incluyen:

1. Cuestiones elementales intuitivas sobre lógica y conjuntos, que pueden suponerse a las de otras asignaturas de matemáticas.

2. La definición y primeras propiedades de objetos algebraicos como grupos y cuerpos.

3. Un repaso de sistemas de ecuaciones lineales (SEL) por el método de Gauss, en el que se deja constancia del uso de las propiedades algebraicas de los cuerpos.

4. El enunciado y demostración de las propiedades del grupo de permutaciones que se usarán posteriormente.

5. La definición directa del determinante de una matriz cuadrada, junto a la demostración de algunas de sus propiedades relevantes.

A estas secciones se les añade una última §1.6 donde se aplican los determinantes a la resolución de SEL. Su objetivo es que el estudiante pueda ejercitarse en ellas a lo largo de todo el curso. No obstante, para ello es necesario usar algunas propiedades de determinantes y rangos que en ese punto aún no se han demostrado. Por esta razón, en la primera subsección §1.6.1 se anticipan explícitamente los resultados que se usan sin demostrar, así como los temas en los que se acabarán demostrando. Se espera así que el estudiante pueda resolver ejercicios usando correctamente herramientas con las que, probablemente, tenga alguna familiaridad. No obstante, insistimos en que a lo largo del resto del libro ninguna de estas propiedades se usará hasta que se demuestre formalmente.

Los Temas 2 y 3 desarrollan la teoría de espacios vectoriales y aplicaciones lineales. Incluyen también los correspondientes conceptos de espacios afines. Estos conceptos se desarrollan en Física sólo en el primer curso, por lo que resultan necesarios para los alumnos de este grado. En Matemáticas, se estudiarán con detalle en cursos superiores. No obstante, su ubicación dentro del presente manual puede servir para fijar ideas posteriormente.

Los conceptos de espacio vectorial cociente y dual pueden presentar dificultades por su nivel de abstracción. El cociente se estudia completamente en el Tema 3, §3.5, mostrándose como aplicación suya el primer teorema de isomorfía. El Tema 4 se dedica específicamente al dual. Aunque no sea imprescendible aquí, adoptamos los convenios de índices superiores e inferiores usuales en Física y que acabarán siendo necesarios cuando se estudien tensores en otros cursos. En el Tema 5 se estudian los determinantes desde el punto de vista tensorial. Creemos que es conveniente incluirlo aquí porque, por una parte, los determinantes cierran todas las demostraciones de las herramientas que se usan y, por otra, proporcionan una interpretación geométrica intuitiva en términos de elementos de volumen. Además, proporcionarán un ejemplo interesante de tensor que puede servir de motivación para su inmediatamente posterior estudio.

El libro contiene demostraciones o resultados que pueden resultar algo repetitivos. El lector, más que estudiar pasivamente resultados, debe tratar de demostrarlos por sí mismo, considerando así muchas proposiciones y corolarios como ejercicios resueltos.

Agradecimientos

Estos apuntes han sido elaborados teniendo en cuenta trabajos previos de otros miembros del Departamento de Geometría y Topología de la Universidad de Granada, que el autor agradece afectuosamente. Entre ellos, desde luego, el excelente libro del Prof. Alfonso Romero Sarabia, *Álgebra lineal y Geometría I* (Ed. La Madraza, 1991), las notas del curso del Prof. César Rosales Lombardo o el trabajo de impartición conjunta del curso con los Profesores Antonio Alarcón y Ana María Hurtado.

Índice general

Tema 1

Preliminares

1.1. Fundamentos: lógica, conjuntos y aplicaciones

Los contenidos de esta parte, que están en los fundamentos de toda la matemática, se verán de modo extendido en otras asignaturas. Para nuestros propósitos resulta suficiente el Tema 0 del libro de A. Romero [3], en el cual se basa esta sección de los apuntes.

Como notación, usaremos $\mathbb{N}, \mathbb{Z}, \mathbb{Q}, \mathbb{R}$ para denotar los números naturales, enteros, racionales y reales, respectivamente.

1.1.1. Algunos elementos y notación de lógica

La Lógica, que subyace en los fundamentos de las Matemáticas, estudia las reglas que permiten hacer razonamientos válidos. Nos restringiremos aquí a algunas ideas intuitivas sobre ella y su lenguaje básico.

Una *proposición* es un enunciado que puede ser verdadero o falso. Partiendo de dos proposiciones p, q, podemos formar otras proposiciones más complejas usando ciertas operaciones lógicas; la verdad o falsedad de cada nueva proposición dependerá de la de p y q, así como de la operación lógica. Concretamente:

- Conjunción: $p \wedge q$, que se lee "p y q".

  Esta proposición es verdadera cuando lo son simultáneamente las dos, p y q, y falsa en caso contrario, esto es, cuando es falsa al menos una de las proposiciones p, q.

- Disyunción: $p \vee q$, que se lee "p ó q".

  Esta proposición es verdadera cuando lo es al menos una de los dos proposiciones p, q y falsa en caso contrario, esto es, cuando son falsas tanto p como q.

- Negación $\neg p$, que se lee "no p".

  Este enunciado es verdadero cuando p es falso, y es falso cuando p es verdadero.

  Nota. En Lógica se supone siempre que si una proposición es verdadera entonces su negación es falsa (y viceversa, de hecho, se entiende que $\neg(\neg p) = p$). Esto no sólo coincide con lo que sugiere nuestra intuición, sino que tiene una raíz más profunda: no es difícil demostrar, usando reglas lógicas simples, el llamado *principio de explosión*[1]: en el caso de que exista una proposición p tal que p y $\neg p$ sean ciertas, entonces cualquier otra proposición q sería cierta también.

Ejercicio. Discútase si las siguientes afirmaciones son proposiciones y, en caso afirmativo, cuál es su negación en el lenguaje cotidiano: (a) "Puede que llueva mañana", (b) "Esta frase tiene cinco palabras".

- Condicional: $p \Rightarrow q$, que se lee de varios modos alternativos: "si p entonces q", "p implica q", "p es suficiente (o condición suficiente) para q", "q es necesario (o condición necesaria) para p", "q se deduce de p". [2]

  La proposición $p \Rightarrow q$ es verdadera cuando se da uno de los siguientes dos casos[3]: (i) p y q son verdaderas, o (ii) p es falsa, sea q verdadera o no. Así, se tiene:

  (a) Si $p \Rightarrow q$ es verdadera, y p también lo es, entonces q es verdadera. A este modo de razonar se le llama *modus ponendo ponens* ("modo que, afirmando, afirma", entendiendo que al afirmar p se afirma q).

  (b) Si la proposición $p \Rightarrow q$ es verdadera, y q es falsa (esto es, $\neg q$ es verdadera) entonces p es falsa (esto es, $\neg p$ es verdadera). A este modo de razonar se le llama *modus tollendo tollens* ("modo que, negando, niega", entendiendo que al negar q se niega p).

  (c) De los dos puntos anteriores, de deduce que $p \Rightarrow q$ es verdadera si y sólo si lo es $\neg q \Rightarrow \neg p$. A esta última implicación se le llama el *contrarrecíproco* de la primera.

  Para ejercitarse, el lector puede tomar p como la proposición "llueve", q como "la calle está mojada" y aplicar todo lo dicho a la proposición $p \Rightarrow q$ ("si llueve entonces la calle está mojada").

- Condicional recíproco: $p \Leftarrow q$, que se lee "sólo si p entonces q".

  Puede tomarse como una notación alternativa para $q \Rightarrow p$, por lo que se reduce al caso anterior (y puede leerse de modos alternativos similares a los de ese caso).

- Equivalencia: $p \Leftrightarrow q$, que se lee "p equivale a q" o "p si y sólo si q"'.

  Puede tomarse como una notación abreviada de $(p \Rightarrow q) \wedge (p \Leftarrow q)$. Así, $p \Leftrightarrow q$ es verdadera cuando (i) p y q son verdaderas, o (ii) p y q son falsas.

  Como hemos visto, la proposición $p \Rightarrow q$ era equivalente a su contrarrecíproco; esto lo podemos escribir ahora:

  $$(p \Rightarrow q) \iff (\neg p \Leftarrow \neg q),$$

  (por supuesto, el recíproco $p \Leftarrow q$ no es equivalente).

Como notación lógica que también usaremos, se tienen los símbolos:

- Cuantificador universal $\forall$, que se lee "para todo".
- Cuantificador existencial $\exists$, que se lee "existe"; si se escribe $\exists!$ se lee "existe un único".

Estos símbolos se aplicarán a menudo en la próxima sección, lo que servirá como ejemplo de su uso.

[1] En terminología clásica, *ex contradictione quodlibet* (o *ex falso quodlibet*): de la contradicción (se sigue) lo que se quiera.

[2] Para la proposición p se usan nombres como antecedente, prótasis o hipótesis, y para la q consecuente, apódosis o tesis.

[3] Esto no se corresponde totalmente con la idea intuitiva de lo que es una implicación, pero no entraremos en una tal discusión cuyas sutilezas exceden nuestros objetivos.

1.1.2. Conjuntos

Conceptos generales

De manera intuitiva, suficiente para nuestros objetivos, entenderemos por *conjunto* una colección cualquiera de objetos; de cada uno de estos objetos se dirá que es un *elemento* del conjunto. Los elementos de un conjunto pueden determinarse por extensión (esto es, enumerando todos y cada uno de ellos) o por comprensión (enunciando una propiedad que los caracterice inequívocamente). Si X es un conjunto, de cualquiera de sus elementos diremos que *pertenece* a X, lo cual se denota $x \in X$ (equivalentemente, se dice que X *contiene* a x, y se denota $X \ni x$).[4] Diremos que dos conjuntos X, Y son iguales, lo que denotaremos $X = Y$, cuando tienen los mismos elementos, esto es, cuando todo elemento de X también lo es de Y y todo elemento de Y también lo es de X. Así, en notación lógica, $X = Y$ significa:

$$(\forall x \in X, x \in Y) \wedge (\forall y \in Y, y \in X).$$

Dados dos conjuntos A y X, diremos que A es un subconjunto de X si todo elemento de A es un elemento de X, esto es, cuando:

$$\forall x \in A, x \in X, \qquad \text{o, igualmente,} \qquad x \in A \Rightarrow x \in X.$$

En este caso, escribiremos $A \subset X$, que se lee A *incluido* en X (o X incluye a A). Resulta inmediato de las deficiniones que, para cualesquiera conjuntos[5] X, Y:

$$X = Y \qquad \Longleftrightarrow \qquad (X \subset Y) \wedge (Y \subset X)$$

Todo conjunto X admite dos subconconjuntos a los que llamaremos *impropios* el propio conjunto X y el conjunto vacío (esto es, el que no tiene ningún elemento), el cual denotaremos $\emptyset$. Este último se puede admitir sin incurrir en contradicción y resulta útil desde el punto de vista lógico, como se verá más adelante. A cualquier subconjunto de X que no sea impropio, se le llamará *propio*.

Operaciones entre conjuntos

Dado un conjunto X, al conjunto cuyos elementos son todos los subconjuntos de X le llamaremos *conjunto de las partes de* X, y lo denotaremos $\mathcal{P}(X)$. Se tiene así: $A \in \mathcal{P}(X)$ si y sólo si $A \subset X$.

Ejercicio. Dado un número natural n, se define el conjunto $X_n = \{1, 2, \ldots, n\}$. Construir $\mathcal{P}(X_n)$ para $n = 1, 2, 3$. ¿Cuántos elementos tiene? Generalízese para cualquier valor de n (véase nota 1.32).

Dados $A, B \in \mathcal{P}(X)$, se definen los siguientes subconjuntos de X:

- Intersección, $A \cap B := \{x \in X : (x \in A) \wedge (x \in B)\}$.
- Unión, $A \cup B := \{x \in X : (x \in A) \bigvee (x \in B)\}$.
- Diferencia, $A \setminus B := \{x \in X : (x \in A) \wedge (x \notin B)\}$.[6]

[4] *Nota.* Para definir los conjuntos, implícitamente se ha supuesto una clase de objetos preexistente. Algunos de estos objetos podrían ser conjuntos previamente definidos, pero no se permite que un conjunto sea elemento de sí mismo (pues no estaría entre los objetos preexistentes) u otras construcciones autorreferentes similares. Lo contrario daría lugar a la conocida *paradoja de Russell*: si a los conjuntos se les permitiera contenerse a sí mismos, entonces podríamos construir el conjunto X de todos los conjuntos que no son elementos de sí mismos; se llega entonces a un absurdo al tratar de determinar si X es un elemento de sí mismo o no (piénsese cuidadosamente qué contradicción ocurriría si se supone que lo es y si se supone que no lo es).

[5] Obsérvese que nuestro uso del símbolo $\subset$ es el mismo que el del símbolo $\subseteq$ (el cual remarca que se puede dar la igualdad). El símbolo $\subsetneq$ denota que se da la inclusión pero no la igualdad (esto es, que la inclusión es *estricta*).

[6] En ocasiones a la diferencia también se le denota $A - B$, y se lee A menos B.

A la diferencia $X \setminus A$ también se le llama *complementario de A en X*.

Ejercicio. Representar los conjuntos así definidos usando diagramas de Venn (búsquese qué son, si no se conocen).

Nota. A partir de las definiciones anteriores no es difícil definir la intersección o la unión de una colección finita, o incluso infinita, de subconjuntos de X.

Dados dos conjuntos X, Y se define su *producto cartesiano* $X \times Y$ como

$$X \times Y : \{(x,y) : (x \in X) \wedge (y \in Y)\},$$

donde (x,y) es el *par ordenado* formado por x e y (en este orden). Aquí, por par ordenado se entiende un conjunto de dos elementos en el que importa el orden en el que se enumeran; dado (x,y), a x se le llama primer elemento del par, y a y segundo elemento. El producto cartesiano $X \times Y$ es, pues, el conjunto de todos los pares ordenados cuyo primer elemento pertenece a X y cuyo segundo elemento pertenece a Y. [7]

Ejercicio. Explicar la diferencia existente entre los objetos que aparecen en cada caso (a), (b), (c) siguiente: (a) x, $\{x\}$, (x,x); (b) $\{x,y\}$, $\{y,x\}$; (c) (x,y), (y,x).

Ejercicio. Si X e Y tienen un número finito n y m de elementos, respectivamente, ¿cuántos elementos tiene $X \times Y$?

Relaciones binarias

Para definir el concepto de relación binaria en un conjunto X, observemos primero que podemos construir el producto cartesiano $X \times X$ (como un caso particular de la definición general $X \times Y$). Llamaremos *relación binaria* en X a cualquier subconjunto R de $X \times X$. Cuando un par $(x,y) \in X \times X$ pertenezca a X diremos: "x está relacionado con y", y escribiremos $x\mathcal{R}y$.

Algunas posibles propiedades que puede que verifique una relación binaria son:

- *Reflexiva*: todo elemento está relacionado consigo mismo, esto es: $x \in X \Rightarrow x\mathcal{R}x$.
- *Simétrica*: $x\mathcal{R}y \Rightarrow y\mathcal{R}x$.
- *Transitiva*: $(x\mathcal{R}y) \wedge (y\mathcal{R}z) \Rightarrow x\mathcal{R}z$.
- *Antisimétrica:* $(x\mathcal{R}y) \wedge (y\mathcal{R}x) \Rightarrow x = y$.

[7] *Nota.* Aunque la definición anterior de par ordenado es suficiente para nuestros propósitos, es de señalar que la noción de "orden" no ha sido introducida todavía dentro de la teoría de conjuntos. No obstante, se puede definir el par ordenado sin hacer mención explícita al orden, concretamente, $(x,y) := \{\{x\},y\}$. Obsérvese que, conjuntistamente $\{\{x\},y\} = \{y,\{x\}\}$, pero en cualquiera de estas dos expresiones x "se distingue" de y (por lo que se sitúa a x como primera componente del par (x,y)): de hecho, no se considera x en sí mismo como elemento del par, sino el conjunto cuyo único elemento es x.

A partir de esta definición, es fácil formalizar los conjuntos ordenados de un número finito de elementos, que se denominarán ternas, cuadruplas, quintuplas o, en general, n-úplas, según tengan 3, 4, 5 ó n elementos, respectivamente. Así, p. ej., $\mathbb{R}^n := \mathbb{R} \times \cdots \times^{(n)} \mathbb{R}$.

Ejercicio. Pónganse ejemplos de dos relaciones que sean a la vez simétricas y antisimétricas, una de ellas que verfique además la propiedad reflexiva, y la otra que no la verifique.

Una relación binaria R se dice que es *de orden* si verifica las propiedades reflexiva, antisimétrica y transitiva. Si en una relación de orden se verifica además:

$$\forall x, y \in X, \quad (x\mathcal{R}y) \vee (y\mathcal{R}x)$$

la relación se llama *de orden total*.

Ejercicio. Compruébese que las siguientes relaciones binarias son relaciones de orden:

(i) En el conjunto $\mathbb{R}$ de los números reales, la relación "ser menor o igual a" (esto es, $\mathcal{R}$ es la relación $\leq$).

(ii) En el conjunto $\mathcal{P}(X)$ de las partes de un conjunto X, la relación 'estar incluido en" (esto es, $\mathcal{R}$ es la relación $\subset$).

¿Es alguna de estas dos relaciones de orden total?

Relaciones de equivalencia

De entre las relaciones binarias, estaremos especialmente interesados en las del siguiente tipo.

Definición 1.1. *Dado un conjunto X, diremos que una relación binaria $R \subset X \times X$ es* de equivalencia *si satisface las propiedades reflexiva, simétrica y transitiva. En este caso, el símbolo $\mathcal{R}$ se sustituirá por el símbolo $\sim$.*

Ejercicio. (i) Se considera en $X = \{1,2,3,4,5,6\}$ la relación binaria $R = \{(1,1),(1,2),(3,5),(6,5)\}$. Se pide completar R hasta una relación de equivalencia, es decir, hallar la única relación binaria R' que verifica las siguientes propiedades: (a) R' incluye a R (esto es, $R \subset R'$), (b) R' es de equivalencia, y (c) R' es la relación binaria más pequeña que verifica estas dos propiedades (más pequeña en el sentido de que si R'' es otra relación binaria que verifica (a) y (b) entonces $R' \subset R''$).

(ii) Discutir si, dado un conjunto arbitrario X y una relación binaria R cualquiera en X, existe una relación binaria R' que complete R hasta una relación de equivalencia (esto es, que satisfaga las propiedades (a), (b) y (c) del apartado anterior).

Definición 1.2. *Sea $(X,\sim)$ un conjunto dotado con una relación de equivalencia. Para cada $x \in X$ se define la* clase de equivalencia de *x por*

$$[x] := \{y \in X : y \sim x\},$$

esto es, $[x]$ es el conjunto de todos los elementos de X que están relacionados[8] *con x. Llamaremos* conjunto cociente $X/\sim$ *al conjunto de todas las clases de equivalencia.*

Obsérvese que cada clase $[x]$ es un subconjunto de X y un elemento de $X/\sim$, y que $X/\sim$ es un subconjunto de $\mathcal{P}(X)$.

Ejercicio. Se considera en $\mathbb{R}^3$ la relación de equivalencia "estar a la misma distancia del origen". Determinar el conjunto cociente.

[8]Debido a la propiedad simétrica, no importa si "estar relacionado" se interpreta en el sentido $x \sim y$ ó $y \sim x$ (o ambos).

Proposición 1.3. *Las clases de equivalencia de* $(X, \sim)$ *verifican:*

(i) $x \in [x]$. *Por tanto, ninguna clase de equivalencia es vacía, y la unión de todas las clases de equivalencia es* X.

(ii) si $y \in [x]$ *entonces* $[x] = [y]$. *En consecuencia, las clases de equivalencia son disjuntas dos a dos, esto es:*

$$[x] \neq [y] \quad \Rightarrow \quad [x] \cap [y] = \emptyset.$$

Demostración. (i) Como por la propiedad reflexiva $x \sim x$, se sigue $x \in [x]$. Así, la clase de un elemento nunca puede ser vacía (pues el propio elemento está en la clase) y la unión de todas las clases es el conjunto X.

(ii) Para demostrar $[y] \subset [x]$, sea $z \in [y]$, esto es, $z \sim y$. Como, por hipótesis, $y \sim x$, aplicando la propiedad transitiva se sigue $z \sim x$, esto es, $z \in [x]$, como se quería.

Para demostrar $[x] \subset [y]$, sea $z \in [x]$, esto es, $z \sim x$. Por hipótesis, $y \sim x$, y por la propiedad simétrica $x \sim y$. Aplicando la propiedad transitiva a $z \sim x$, $x \sim y$, se sigue $z \sim y$, esto es, $z \in [y]$, como se quería.

Para la última afirmación, razonando por el contrarrecíproco si $\exists z \in [x] \cap [y]$ entonces $[z] = [x]$ y $[z] = [y]$ (aplíquese dos veces la afirmación anterior) por lo que $[x] = [y]$, como se quería. □

Como consecuencia de la proposición 1.3, las clases de equivalencia que forman $X/\sim$ satisfacen:

(1) No son vacías.

(2) Son disjuntas dos a dos.

(3) Su unión es todo X.

En general, dado el conjunto X, una *partición* P de X es una colección de subconjuntos de X que satisfacen las propiedades (1), (2) y (3). La proposición anterior implica por tanto que el conjunto cociente $X/\sim$ es una partición de X. El siguiente ejercicio muestra que, de hecho, toda partición de X puede verse como el conjunto cociente para una $\sim$ canónicamente determinada.

Ejercicio. Sea P una partición de X, y se considera la relación binaria en X: "x está relacionado con y si y sólo si x, y pertenecen a un mismo elemento de P". Demostrar que esta relación binaria es de equivalencia y que su conjunto cociente coincide con P.

Nota. A cada elemento y de un clase $[x]$ se le llama *representante de la clase*. El hecho de que si $y \in [x]$ entonces $[y] = [x]$ sugiere que todos los representantes de la clase son "igualmente buenos" a la hora de denotar esa clase.

1.1.3. Aplicaciones entre conjuntos

Concepto

Sean X, Y dos conjuntos. Una *aplicación* f de X a Y es una regla que asigna a cada elemento de X un único elemento de Y. La aplicación se representa $f : X \to Y$, y escribiremos $x \mapsto y$ o bien $f(x) = y$ si y es el único elemento de Y que se asigna a x mediante f. Al conjunto X se le llama *dominio* (o conjunto inicial) de la aplicación, a Y el *codominio* (o conjunto final) y a $f(x)$ *imagen por* f *de* x. La *imagen* de la aplicación f, $\mathrm{Im} f$, es el subconjunto del codominio que contiene las imágenes de todos los elementos de X, esto es:

$$\mathrm{Im} f := \{f(x) : x \in X\} = \{y \in Y \mid \exists x \in X : y = f(x)\} \subset Y.$$

El *grafo* o *gráfico* de la aplicación f se define como el siguiente subconjunto del producto cartesiano $X \times Y$:

$$G(f) := \{(x, f(x)) | x \in X\} = \{(x,y) \in X \times Y | y = f(x)\}$$

Observación. Una aplicación queda caracterizada totalmente por su dominio, codominio y grafo (esto es, si estos tres objetos son iguales para dos aplicaciones f, $\bar{f}$ entonces $f = \bar{f}$).

No obstante, dos aplicaciones pueden ser distintas incluso si sus dominios y grafos coinciden. De hecho, son diferentes las aplicaciones $f_1 : \mathbb{N} \to \mathbb{N}, x \mapsto x^2$, y $f_2 : \mathbb{N} \to \mathbb{R}, x \mapsto x^2$. También es distinta la aplicación $f_3 : \mathbb{R} \to \mathbb{R}, x \mapsto x^2$, a pesar de que tanto su codominio como su regla para asignar imágenes coincidan con los de f_2.

Proposición 1.4. *(i) El grafo $G(f)$ satisface:*

(a) $\forall x \in X,\ \exists y \in Y : (x,y) \in G(f)$.

(b) $(x,y), (x,y') \in G(f) \Rightarrow y = y'$.

Recíprocamente, si C es un subconjunto de $X \times Y$ que verifica las propiedades (a) y (b) del grafo entonces existe una única aplicación $f : X \to Y$ tal que $C = G(f)$.

Demostración. (a) Dado x, su imagen $f(x)$ es el elemento $y \in Y$ buscado.

(b) Por la definición del grafo, $y = f(x)$ y también $y' = f(x)$. Como por la definición de aplicación la imagen de x es única, se sigue $y = y'$. Para la última afirmación, obsérvese primero que C satisface la siguiente propiedad, que equivale a (a) y (b):

$$\forall x \in X,\ \exists! y \in Y :\ (x,y) \in C. \tag{1.1}$$

Por tanto, basta con definir $f : X \to Y$ como la aplicación que asigna a cada $x \in X$ el único $y \in Y$ determinado por la propiedad (1.1). Esta propiedad asegura directamente tanto que f es una aplicación como que C es su grafo. La unicidad de f se debe a que cualquier otra aplicación $\bar{f} : X \to Y$ tal que $G(\bar{f}) = C$ tendría el mismo dominio, codominio y grafo que f, por lo que sería igual a f. □ [9]

Generación de aplicaciones a partir de otras

Dada una aplicación $f : X \to Y$ y un subconjunto A de su dominio ($A \subset X$), se define la *restricción de f a A*, denotada $f|_A$, por:

$$f|_A : A \to Y, \qquad x \mapsto f(x),$$

esto es, la regla que asigna a cada elemento x su imagen es la misma que la de f, pero $f|_A$ la aplica sólo a los elementos de A. [10]

La aplicación $f : X \to Y$ también permite definir aplicaciones entre los correspondientes conjuntos de las partes. Concretamente, la aplicación *imagen directa*,

$$f_* : \mathcal{P}(X) \to \mathcal{P}(Y). \qquad A \mapsto f_*(A) := \{f(x) : x \in A\}$$

[9] *Nota.* La proposición 1.4 permite eliminar la ligera imprecisión que se cometió al definir el concepto de aplicación usando la expresión "una regla que asigna". En efecto, se puede definir de manera totalmente formal una aplicación como una terna (X,Y,C) en la que X, Y son conjuntos y C es un subconjunto de $X \times Y$ que satisface (1.1).

Si se adoptara esta definición de aplicación, la proposición 1.4 dejaría de tener sentido (esencialmente, estaría incluida en la definición). No obstante, la demostración de su última parte serviría como motivación para recuperar la idea intuitiva de que, gracias a las propiedades de C, la aplicación f "asigna" a cada elemento de X uno de Y.

[10] Análogamente, si se considera cualquier $B \subset Y$ que incluya $\mathrm{Im} f$, se puede obtener una aplicación $f|^B$ restringiendo el codominio, esto es, $f|^B : X \to B, x \mapsto f(x)$.

y la imagen recíproca,

$$f^* : \mathcal{P}(Y) \to \mathcal{P}(X). \qquad B \mapsto f^*(B) := \{x \in X : f(x) \in B\}.$$

Nota. En ocasiones se abusa de la notación escribiendo $f(A)$ en lugar de $f_*(A)$, y $f^{-1}(B)$ en lugar de $f^*(B)$. Esto no es formalmente correcto, por lo que no debe de hacerse (al menos mientras no se tenga la suficiente práctica para entender inequívocamente cuándo se está abusando de la notación).

Observación. De la definición se sigue:
(a) $f_*(A) \subset Y$, por lo que $f_*(A) \in \mathcal{P}(Y)$,
(b) $f^*(B) \subset X$, por lo que $f^*(B) \in \mathcal{P}(X)$,
(c) $f_*(X) = \mathrm{Im} f$, $f^*(Y) = X$,
(d) cuando $B \subset Y$ verifica $B \cap \mathrm{Im}\ f = \emptyset$ entonces $f^*(B) = \emptyset$ (el cual es un subconjunto de X y, de hecho, de cualquier otro conjunto).

Definición 1.5. *Se consideran las aplicaciones $f : X \to Y, g : Y \to Z$, en la que el codominio de f coincide con el dominio de g. Se define su* composición $g \circ f$ *por:*

$$g \circ f : X \to Z, \qquad x \mapsto (g \circ f)(x) := g(f(x)),$$

esto es, cada $(g \circ f)(x)$ se obtiene aplicando a x primero f (que por este motivo aparece notacionalmente a la derecha en $g \circ f$) y luego aplicando g a la imagen $f(x)$.

Ejercicio. (i) Sean $f, g : \mathbb{R} \to \mathbb{R}$ definidas por:

$$f(x) = x^2 + 1, \qquad g(x) = e^x, \qquad \forall x \in \mathbb{R}.$$

Calcular $g \circ f$ y $f \circ g$.
(ii) Sean $f : X \to Y$ y $g : X' \to Y'$ dos aplicaciones entre conjuntos. Supongamos que se puede definir $g \circ f$, ¿bajo qué condiciones se puede definir también $f \circ g$?
(iii) Compruébese que la definición de composición se puede extender naturalmente al caso de funciones $f : X \to Y, g : Y' \to Z$ tales que $\mathrm{Im}\ f \subset Y'$.

Proposición 1.6. *Se consideran las aplicaciones $f : X \to Y, g : Y \to Z, h : Z \to W$. Entonces*

$$(h \circ g) \circ f = h \circ (g \circ f). \tag{1.2}$$

Demostración. Claramente, ambos miembros de la igualdad están bien definidos como aplicaciones y tienen igual dominio y codominio. La regla de asignación de imágenes también coincide pues:

$$[(h \circ g) \circ f] = (h \circ g)(f(x)) = h(g(f(x))), \quad [h \circ (g \circ f)](x) = h((g \circ f)(x)) = h(g(f(x))), \qquad \forall x \in X \;\square$$

Nota. La proposicion 1.6 se puede enunciar diciendo que la composición es *asociativa*. La proposicion también justifica que la notación $h \circ g \circ f$ es inequívoca, pues se puede usar para indicar cualquiera de las dos expresiones (1.2).

Tipos notables de aplicaciones y descomposición canónica

Definición 1.7. *Sea $f : X \to Y$ una aplicación. Se dice que f es:*

- Inyectiva*: si elementos distintos de X tienen imágenes distintas en Y, esto es, para todo $x, x' \in X$ tales que $x \neq x'$ se tiene $f(x) \neq f(x')$ (equivalentemente, cuando $f(x) = f(x') \Rightarrow x = x'$).*
- Suprayectiva, sobreyectiva, sobre *o* exhaustiva*: si todo elemento de Y es imagen de alguno de X, esto es, $\forall y \in Y$, $\exists x \in X : f(x) = y$ (equivalentemente, si* Im f*=*Y*).*
- Biyectiva*: si f es inyectiva y suprayectiva, esto es, $\forall y \in Y$, $\exists! x \in X$: $y = f(x)$.*

  En este caso, se define la aplicación inversa *de f, que denotaremos f^{-1}, estableciendo que $f^{-1}(y)$ es el único $x \in X$ tal que $f(x) = y$.*

Ejemplos generales de estos tipos de aplicaciones son:

- La inclusión i de un subconjunto: dado $A \subset X$, $i : A \to X$, $x \mapsto x$. La inclusión i es inyectiva.
- La proyección canónica π sobre un conjunto cociente: dado $(X, \sim)$,

  $$\pi : X \to X/\sim, \qquad x \mapsto \pi(x) := [x].$$

  La proyección π es suprayectiva.
- La aplicación identidad I_X: dado un conjunto X, $I_X : X \to X, x \mapsto x$ (caso particular e la inclusión cuando $A = X$). La identidad I_X es biyectiva.

Ejercicio. Tómese como $f : D \to \mathbb{R}$ cada una de las siguientes cuatro funciones elementales: $x \mapsto x^2, e^x, \sqrt{x}$, $\ln(x)$, siendo en cada caso $D \subset \mathbb{R}$ su dominio natural. Determínese cuáles de ellas son inyectivas, suprayectivas o biyectivas. ¿Es alguna de estas aplicaciones la inversa de otra?

Proposición 1.8. *La composición de dos aplicaciones inyectivas (respectivamente, suprayectivas, biyectivas es inyectiva (respectivamente, suprayectiva, biyectiva).*

Demostración. Hágase como ejercicio. □

Podemos caracterizar la aplicación inversa como sigue.

Lema 1.9. *Sean $f : X \to Y$, $g : Y \to Z$.*

- *Si $g \circ f$ es inyectiva entonces f es inyectiva.*
- *Si $g \circ f$ es suprayectiva entonces g es suprayectiva.*

Por tanto, si $g \circ f$ es biyectiva entonces f es inyectiva y g es suprayectiva.

Demostración. Para la primera afirmación, si f no fuera inyectiva existirían $x, x' \in X$, $x \neq x'$ tales que $f(x) = f(x')$. En consecuencia, $(g \circ f)(x) = g(f(x)) = g(f(x')) = (g \circ f)(x')$, y $g \circ f$ no es inyectiva.

Para la segunda, sea $z \in Z$. Como g es suprayectiva, existe $x \in X$ tal que $z = (g \circ f)(x) = g(f(x))$. Por tanto, z es la imagen por g de $f(x) \in Y$. □

Proposición 1.10. *Sean $f : X \to Y$, $g : Y \to X$ dos aplicaciones tales que*

$$g \circ f = I_X, \qquad f \circ g = I_Y.$$

Entonces f, g son biyectivas y $g = f^{-1}$.

Demostración. Aplicando el lema anterior a la primera igualdad se obtiene la inyectivdad de f y la suprayectividad de g y, aplicándolo a la segunda, la inyectivdad de g y la suprayectividad de f, de donde se sigue la biyectividad de ambas.

Claramente, g y f^{-1} tienen igual dominio y codominio. Sea $y \in Y$ y $x = f^{-1}(y)$, esto es, $f(x) = y$. Se tiene entonces $g(y) = g(f(x)) = (g \circ f)(x) = I_X(x) = x = f^{-1}(y)$, por lo que g y f^{-1} coinciden. □

Una aplicación arbitraria no tiene por qué ser inyectiva o suprayectiva. No obstante, toda aplicación se puede escribir como composición de una aplicación inyectiva, una biyectiva y una suprayectiva, del modo canónico que se describe a continuación.

Teorema 1.11 (Descomposición canónica de una aplicación)**.** *Sea $f : X \to Y$ una aplicación. Se define la relación de equivalencia en X: $x \sim x' \Leftrightarrow f(x) = f(x')$, y se considera su proyección canónica $\pi : X \to X/\sim$, que es sobre. Asimismo, se considera* $\operatorname{Im} f$ *y su inclusión $i : \operatorname{Im} f \to Y$, que es inyectiva.*

Se tiene entonces que la aplicación

$$b : (X/\sim) \longrightarrow \operatorname{Im} f, \qquad b([x]) := f(x)$$

está bien definida, es biyectiva y verifica:

$$f = i \circ b \circ \pi.$$

Demostración. En primer lugar, obsérvese que, efectivamente, la relación binaria $\sim$ es (trivialmente) una relación de equivalencia.

Mostrar que b está correctamente definida como aplicación requiere una cierta discusión. Al definirse la imagen por b de toda la clase $[x]$ como la imagen por f del representante x (éste es el significado de $b([x]) := f(x)$), uno podría preguntarse qué ocurriría si tomáramos un representante distinto x' para esa clase. Esto es, si $x' \sim x$ entonces $[x'] = [x]$ y también se tendría $b([x]) = b([x']) = f(x')$. No obstante, el hecho de que $x \sim x'$ significa precisamente $f(x') = f(x)$ (por la definición de la relación de equivalencia), la imagen que de $[x]$ se obtiene por b es independiente del representante escogido, esto es, b está bien definida como aplicación.

La inyectividad de b se sigue entonces porque si $b([x]) = b([y])$ entonces $f(x) = f(y)$ y, por tanto, $x \sim y$, esto es, $[x] = [y]$.

Para la suprayectividad de b, si $z \in \operatorname{Im} f$ entonces existe un $x \in X$ tal que $z = f(x)$. Por tanto $b([x]) = z$, esto es, $z \in \operatorname{Im} b$.

Finalmente, es inmediato que f y $i \circ b \circ \pi$ tienen igual dominio y codominio. Por tanto, su igualdad se deduce de

$$i \circ b \circ \pi(x) = i(b(\pi(x))) = i(b([x]) = i(f(x)) = f(x), \qquad \forall x \in X. \qquad \square$$

Como un último comentario, obsérvese que f es inyectiva si y sólo si π lo es, y que f es suprayectiva si y sólo si i lo es.

1.2. Grupos, cuerpos y matrices

Abstrayendo las operaciones y propiedades conocidas en el conjunto $\mathbb{R}$ de los números reales, introducimos en esta sección el concepto de cuerpo y las estructuras algebraicas previas (grupos, anillos) necesarias para su comprensión.

Nuestro objetivo aquí sólo es proporcionar el concepto de cuerpo para poder definir en general el de espacio vectorial, así como el de grupo, por ser una estructura geométrica básica. En otras asignaturas del grado se desarrollarán con detalle todas las estructuras que veremos aquí. Así, aunque el lector tenga ahora en mente sólo la intuición de un cuerpo como $\mathbb{R}$, podrá entender con facilidad que las propiedades de los espacios vectoriales construidos sobre $\mathbb{R}$ se extienden con generalidad a espacios vectoriales sobre cualquier cuerpo.

1.2.1. Grupos

Sea G un conjunto. Una *operación* o *ley de composición interna* es cualquier aplicación del tipo $\cdot : G\times G \to G$. Usaremos la notación[11]:

$$\begin{array}{rcl} \cdot : G\times G & \to & G \\ (x,y) & \mapsto & x\cdot y \end{array}$$

Obsérvese que las operaciones usuales $+, \cdot$ en $\mathbb{R}$ (esto es, la suma $+$ y el producto usual $\cdot$), verifican nuestra definición y convenios.

Algunas posibles propiedades que puede tener la operación $\cdot$ son:

1. *Asociativa*: $x\cdot(y\cdot z) = (x\cdot y)\cdot z$, $\forall x,y,z\in G$.

2. *Elemento neutro*: $\exists e\in G$ (al que llamaremos *elemento neutro*) tal que $e\cdot x = x\cdot e = x$, $\forall x\in G$.

3. *Elemento simétrico* (supuesta la existencia de un elemento neutro $e\in G$):

   $\forall x\in G$, $\exists \bar{x}\in G$ (al que llamaremos *elemento simétrico* de x) tal que $x\cdot\bar{x} = \bar{x}\cdot x = e$.

4. *Conmutativa*: $x\cdot y = y\cdot x$, $\forall x,y\in G$.

Definición 1.12. *Un grupo es el par $(G,\cdot)$ formado por un conjunto G y una operación $\cdot$ en G que verifica las tres primeras propiedades anteriores. Si además verifica la cuarta, el grupo se dirá* conmutativo *o* abeliano.

Proposición 1.13. *En todo grupo $(G,\cdot)$:*

(a) El elemento neutro e es único,

(b) Se puede "simplificar", esto es, $x\cdot y = x\cdot y' \Rightarrow y = y'$ (simplificación por la izquierda) y análogamente $x\cdot y = x'\cdot y \Rightarrow x = x'$ (simplificación por la derecha).

(c) El simétrico $\bar{x}$ de cada elemento x es único. Más aún, si dos elementos x,y del grupo verifican $x\cdot y = e$ entonces x es el simétrico de y (e y es el simétrico de x).

Demostración. (a) Si e' fuera otro elemento neutro: $e' = e'\cdot e = e$, la primera igualdad usando que e es neutro, y la segunda que e' lo es.

[11] Obsérvese el convenio de denotar $x\cdot y$ a la imagen $\cdot((x,y))$. Se pueden usar otros símbolos (p.ej., $\star$), para denotar operaciones con el mismo convenio.

(b) Para demostrar la primera implicación, operando ambos miembros por el simétrico $\bar{x}$ de x por la izquierda, se tiene:

$$\bar{x}\cdot(x\cdot y)=\bar{x}\cdot(x\cdot y')$$

por la propiedad asociativa:

$$(\bar{x}\cdot x)\cdot y=(\bar{x}\cdot x)\cdot y'$$

por la elemento simétrico:

$$e\cdot y=e\cdot y'$$

y por la elemento neutro $y=y'$, como se quería demostrar. La otra implicación se demuestra análogamente, operando ambos miembros por el simétrico $\bar{y}$ de y por la derecha.

(c) Consecuencia sencilla de (a) y (b). □

Suele denotarse al (único) elemento simétrico de x como x^{-1}, y llamársele también *elemento inverso*; en algunos grupos particulares se le da un nombre distinto. Algunos ejemplos de grupos son:

- $(\mathbb{R},+)$. Claramente, el neutro es $e=0$, y el elemento simétrico de x es $-x$, al cual se le llama elemento *opuesto* de x.
- $\mathbb{R}\setminus\{0\}$, con la restricción natural del producto usual de $\mathbb{R}$; obviamente, $e=1$, al simétrico de x se le denota x^{-1} (ó $1/x$) y se le llama *inverso*.
- Dado un conjunto X consideremos el conjunto Biy (X) formado por todas las aplicaciones biyectivas de X en X. Claramente, la composición $\circ$ es una operación en Biy (X), y se verifica que el par (Biy $(X),\circ)$ es un grupo cuyo neutro es la aplicación identidad.

Cuestión: ¿Cuáles de los grupos anteriores son conmutativos?

Ejercicio 1.14. *Demostrar que, en todo grupo $(G,\cdot)$ se verifica:*

$$\overline{xy}=\bar{y}\bar{x},\qquad \forall x,y\in G.$$

¿qué ocurre si el grupo es conmutativo?

1.2.2. Anillos y cuerpos

Consideremos ahora un conjunto A dotado de dos operaciones, que denotaremos $+$, . ; esto es, una terna $(A,+,.)$.

Definición 1.15. *Diremos que $(A,+,.)$ es un* anillo *si verifica las siguientes propiedades:*

1. *$(A,+)$ es un grupo conmutativo.*

   *(*Convenio*: si no se especifica lo contrario, al elemento neutro se le denotará* 0 *y a la operación $+$ se le llamará* suma *en A.)*

2. *$(A,.)$ verifica la propiedad asociativa.*

   *(*Convenio: *si no se especifica lo contrario, a la operación . se le llamará* producto *en A.)*

3. $(A,+,.)$ *verifica la* propiedad distributiva de la suma respecto al producto*, esto es*[12]*:*

$$\begin{array}{rcll} x.(y+z) & = & x.y+x.z & \textit{(distributiva por la izquierda) y} \\ (y+z).x & = & y.x+z.x & \textit{(distributiva por la derecha),} \end{array}$$

para todo $x,y,z \in A$.

Observación 1.16. Es fácil comprobar que en todo anillo se verifica $0\ .\ x = x\ .\ 0 = 0$ (obsérvese que $0.x = (0+0).x = 0.x + 0.x$ y simplifíquese por el simétrico de $0.x$).

Casos particulares de anillos son:

- Anillo unitario: anillo cuyo producto tiene elemento neutro, esto es, $\exists 1 \in A : x\ .\ 1 = 1\ .\ x = x$, para todo $x \in A$.

  Al elemento neutro para el producto se le llamará *elemento unidad* (y se le denotará 1).

- *Anillo unitario no trivial*: anillo unitario en el que $1 \neq 0$.

  Nota: si un anillo unitario no verifica esta propiedad, es decir, $1 = 0$ (es un anillo trivial), se tiene que $A = \{0\}$ (pues $x = 1.x = 0.x = 0$ para todo $x \in A$).

- *Anillo conmutativo*: anillo en el que el producto es conmutativo[13], esto es, $x.y = y.x, \forall x,y \in A$.

Observación 1.17. Es sencillo comprobar que en un anillo unitario $(-1)\cdot x = -x$ (pues sumados a $1\cdot x$ dan el neutro para la operación $+$) así como $(-1)(-x) = x$ (pues se puede aplicar la propiedad anterior a $-x$ y se sabe que para la operación de grupo $+$ se sabe $-(-x) = x$).

Proposición 1.18. *Sea $(A,+,.)$ un anillo unitario no trivial y sea*

$$G(A) = \{z \in A : \exists z^{-1} \in A \ / \ z.z^{-1} = z^{-1}.z = 1\}$$

(esto es, $G(A)$ contiene los elementos z que admiten un elemento simétrico para el producto, al cual denotaremos z^{-1} y llamaremos inverso*). Entonces la operación . es restringible a $G(A)$ y $(G(A),.)$ es un grupo (en particular, el inverso es único).*

Demostración. La restringibilidad de . es inmediata porque, si $x,y \in G(A)$ la asociatividad del producto implica directamente $(x.y)^{-1} = y^{-1}.x^{-1}$. Trivialmente, la restricción hereda la propiedad asociativa y $1 \in G(A)$, por lo que $(G(A),.)$ es un grupo con elemento neutro 1. □

Ejercicio 1.19. En el ambiente de la proposición anterior, demuéstrese:

(a) La operación $+$ nunca es restringible a $G(A)$ (por tanto, $G(A)$ no puede tener una estructura de grupo con $+$).

(b) Si $x,y \in A$ verifican $x \neq 0 \neq y$ pero $x.y = 0$ (0 admite divisores) entonces $x,y \notin G(A)$.

Remarquemos que $0 \notin G(A)$ por la observación 1.16. Con estos conceptos previos podemos ya definir el de cuerpo, que usaremos continuadamente.

Definición 1.20. *Un* cuerpo $(K,+,.)$ *es un anillo unitario no trivial en el que el producto verifica la propiedad elemento simétrico para todo $x \in K\setminus\{0\}$, esto es $G(K) = K\setminus\{0\}$.*

El cuerpo $(K,+,.)$ se dirá conmutativo *si el producto . verifica la propiedad conmutativa.*

[12] *Convenio*: extenderemos la notación sobre prelación de operaciones usual en $\mathbb{R}$, de modo que $(x.y)+(x.z)$ se denotará simplemente $x.y+x.z$.

[13] Recuérdese que la operación suma $+$ era siempre conmutativa en un anillo.

Es fácil comprobar que un conjunto dotado de dos operaciones $(K,+,.)$ es un *cuerpo si y sólo si*:

1. $(K,+)$ es un grupo conmutativo
2. $K\setminus\{0\}$ es un grupo con la restricción natural de la operación producto definida en todo[14] K.
3. Se verifica la propiedad distributiva de la suma respecto al producto.

Algunos ejemplos de anillos y cuerpos son los siguientes:

- $(\mathbb{Z},+,.)$ es un anillo unitario (no trivial), pero no un cuerpo.

  Como se señalará más adelante, estas mismas propiedades las tiene el conjunto de matrices reales cuadradas de un orden fijo $n\times n$, con su suma y producto usuales.

- Sea $p\in\mathbb{N}$ un número natural[15]. El conjunto $p\mathbb{Z}(:=\{0,p,-p,2p,-2p,...\})$ de los múltiplos de p, con su suma y producto naturales (restricción de los de $\mathbb{Z}$) forma un anillo. Si $p\neq 1$ este anillo no es unitario.

- El conjunto $\mathbb{Z}/p\mathbb{Z}$ de los enteros módulo $p\mathbb{Z}$ (esto es, el conjunto cociente, que denotaremos $\mathbb{Z}/p\mathbb{Z}$, para la relación de equivalencia en $\mathbb{Z}$ “dos enteros están relacionados si y sólo si su resto al dividir por p es el mismo”, el cual está formado por p clases de equivalencia), tiene una estructura natural de anillo unitario conmutativo (con operaciones en el cociente inducidas por las de $\mathbb{Z}$). Para p primo este anillo es un cuerpo. En el caso $p=2$, este cuerpo $K=\mathbb{Z}/2\mathbb{Z}$ tendría dos elementos $K=\{[0],[1]\}$ y las operaciones se definirían por:

  $$\begin{array}{cccc} [0]+[0]=[0], & [0]+[1]=[1], & [1]+[0]=[1], & [1]+[1]=[0] \\ {[0]}.[0]=[0], & [0].[1]=[0], & [1].[0]=[0], & [1].[1]=[1] \end{array}$$

  Puede comprobarse directamente como ejercicio que $K=\{[0],[1]\}$ con las operaciones anteriores es un cuerpo conmutativo[16].

- $(\mathbb{Q},+,.),(\mathbb{R},+,.)$ son cuerpos conmutativos.

Ejercicio 1.21. *Calcúlese $G(A)$ para cada uno de los anillos unitarios anteriores.*

En adelante, trabajaremos siempre con **cuerpos conmutativos**, denotados simplemente por K, salvo mención explícita de lo contrario.

1.2.3. El cuerpo $\mathbb{C}$ de los números complejos

Por su gran importancia en Matemáticas (incluso a un nivel muy elemental) definiremos aquí el cuerpo $\mathbb{C}$ de los números complejos, el cual se estudirará con detalle en otras asignaturas del grado.

Al igual que los números reales se definen como una extensión de los racionales la cual, permite, por ejemplo, considerar como un número (real) a la raíz de 2 (esto es, que la ecuación $x^2=a$ siempre admita alguna solución cuando $a\geq 0$), los complejos se definen a partir de los reales para dotar de

[14] En particular, esta propiedad impedirá que $(K,+,.)$ pueda ser un anillo unitario trivial.

[15] Como convenio, supondremos que los naturales *no* incluyen el 0.

[16] En este cuerpo, al sumar el elemento unidad consigo misma se obtiene 0. En general, se dice que la *característica* de un cuerpo K es $p\in\mathbb{N}$ si p es el menor natural tal que al operar p veces la unidad consigo misma se obtiene 0, esto es, $1+\cdots+^{(p)}1=0$; en caso de que esto no ocurra para ningún natural, se dice que la característica de K es 0. Para p primo, los cuerpos $\mathbb{Z}/p\mathbb{Z}$ tienen característica p, mientras que los cuerpos con los que trabajaremos usualmente $(\mathbb{Q},\mathbb{R},\mathbb{C})$ tienen característica 0.

sentido a la raíz de números negativos (esto es, que la ecuación $x^2 = a$ siempre admita alguna solución incluso cuando $a < 0$). Para ello, se introduce la llamada *unidad imaginaria* i, que desempeña el papel de ser una solución de $\sqrt{-1}$, y se extienden las operaciones $+, .$ de $\mathbb{R}$ de manera natural a cualquier combinación de números reales y la unidad imaginaria, teniendo siempre en cuenta que $i^2 = -1$.

Formalmente, se define el cuerpo $\mathbb{C}$ de los números complejos como $(\mathbb{R}^2, +, .)$ donde para cada número complejo $z = (x, y) \in \mathbb{R}^2$ a x (resp. y) se le llama la *parte real* (resp. *parte imaginaria*) de z, y las operaciones se definen por:

$$\begin{array}{rcl} (x,y)+(x',y') & := & (x+x', y+y') \\ (x,y).(x',y') & := & (x.x'-y.y', x.y'+y.x') \end{array}$$

para todo $(x, y), (x', y') \in \mathbb{R}^2$ (en estas definiciones, los símbolos $+$, . a la izquierda de las igualdades denotan, respectivamente, la suma y producto de complejos que se está definiendo sobre $\mathbb{R}^2$, mientras que a la derecha representan la suma y producto usuales en $\mathbb{R}$). Se define la *unidad imaginaria* como

$$i := (0, 1).$$

Usando i, simplificamos la notación escribiendo x en lugar de $(x, 0)$ de modo que $(0, y) = y.i$ (que también se denotará indistintamente como $y\, i$, así como por $i\, y$). Así, podemos escribir cada número complejo (x, y) como

$$(x, y) \equiv x + iy,$$

de modo que las definiciones anteriores se reescriben:

$$\begin{array}{rcl} (x+iy)+(x'+iy') & = & (x+x')+i(y+y') \\ (x+iy).(x'+iy') & = & (x.x'-y.y')+i(x.y'+y.x'). \end{array}$$

La definición formal de los complejos permite visualizarlos como puntos de un plano (*plano de Argand*). Intuitivamente, el papel de multiplicar por i un complejo z equivale a rotar en el plano $\mathbb{R}^2$ a z $\pi/2$ radianes (90 grados sexagesimales), usando el sentido positivo de giro (opuesto al del movimiento de las agujas del reloj). Dado un complejo $z = x + iy$ se define su *conjugado*

$$\bar{z} := x - iy$$

(simétrico a z con respecto al eje de abscisas) y su *módulo* (que generaliza el valor absoluto en $\mathbb{R}$)

$$|z| := \sqrt{x^2 + y^2}\ (= \sqrt{z.\bar{z}}).$$

No es difícil demostrar que todo complejo $z = x + iy$ admite una *representación polar*, del tipo

$$z = r(\cos\theta + i \operatorname{sen}\theta), \tag{1.3}$$

donde r no es más que el módulo de z y θ su *argumento*[17] (este último, cuando $x > 0$, puede calcularse como $\theta = \arctan(y/x)$, y resulta sencilla dar fórmulas para los casos $x \leq 0$). Dado un segundo número complejo expresado en forma polar $z' = r'(\cos\theta' + i \operatorname{sen}\theta')$, de las propiedades del producto y suma:

$$z.z' = rr'\left((\cos\theta\cos\theta' - \operatorname{sen}\theta\operatorname{sen}\theta') + i(\operatorname{sen}\theta\cos\theta' + \cos\theta\operatorname{sen}\theta')\right) = rr'\left(\cos(\theta+\theta') + i\operatorname{sen}(\theta+\theta')\right)$$

(la última igualdad usando las fórmulas conocidas del seno y coseno de una suma). Esto es, el producto de dos números complejos es otro complejo cuyo módulo es el producto de los módulos, y su argumento puede escogerse como la suma de los argumentos.

Usando las propiedades anteriores no es difícil comprobar:

[17] Obsérvese que cuando $z \neq 0$, el argumento está definido unívocamente salvo la suma de un número entero de veces 2π.

$\mathbb{C}$ *es un cuerpo conmutativo.*

De hecho, la suma y producto definidos en $\mathbb{C}$ heredan de manera natural las propiedades de cuerpo que verifican la suma y el producto en $\mathbb{R}$. Es de señalar que, para cada $z = a + ib \in \mathbb{C} \setminus \{0\}$ su inverso para el producto es (compruébese directamente):

$$z^{-1} = \frac{a}{a^2+b^2} - i\frac{b}{a^2+b^2} \qquad \left(\text{mnemotécnicamente} \quad z^{-1} = \frac{1}{z} = \frac{\bar{z}}{z\bar{z}} = \frac{\bar{z}}{|z|^2}\right)$$

Para la forma polar (1.3), se tiene $z^{-1} = r^{-1}(\cos(-\theta) + i\,\mathrm{sen}(-\theta))$.

Nota 1 *(Expresión exponencial).* Definiendo,

$$e^{i\theta} := \cos\theta + i\,\mathrm{sen}\,\theta$$

(lo que se justifica cuando se estudia el desarrollo en serie de potencias de la exponencial, $e^w = \sum_{n=0}^{\infty} w^n/n!$, y se compara con el de las funciones seno y coseno), la representación polar se reescribe

$$z = r.e^{i\theta}.$$

Dado un segundo complejo $z' = r'.e^{i\theta'}$ se tiene entonces $z.z' = (r.r').e^{i(\theta+\theta')}$, consistentemente con fórmulas anteriores. Así, para cada $n \in \mathbb{N}$, el conjunto $\{e^{i2\pi k/n} : k = 0, 1, \ldots n-1\}$ está formado por las *raíces n-ésimas de la unidad* (las n soluciones complejas de la ecuación $z^n = 1$).

Nota 2 *(Cuaterniones).* De un modo similar a como se construyen los números complejos a partir de los reales, se define el cuerpo de los números *cuaterniones* (o *cuaternios*). En este caso, el conjunto es $\mathbb{R}^4$, donde se definen

$$i := (0,1,0,0),\; j := (0,0,1,0),\; k := (0,0,0,1)$$

de modo que podemos reescribir cada $(x,y,z,t) \in \mathbb{R}^4$ como:

$$(x,y,z,t) = x + iy + jz + kt.$$

Se define la suma usual, componente a componente, mientras que para el producto se opera de manera natural imponiendo las relaciones:

$$i^2 = j^2 = k^2 = i.j.k = -1$$

(usando la notación $i^2 = i.i$, etc.) De estas relaciones no es difícil demostrar $i.j = k$; $j.i = -k$; $j.k = i$; $k.j = -i$; $k.i = j$; $i.k = -j$. Las operaciones resultantes $+$, . dotan a $\mathbb{R}^4$ de una estructura de cuerpo, el cual *no es conmutativo*. A $\mathbb{H} := (\mathbb{R}^4, +, .)$ se le llama el *cuerpo de los cuaterniones*[18].

[18] Este cuerpo también se puede construir a partir de $\mathbb{C}$ de un modo análogo a como construimos $\mathbb{C}$ a partir de $\mathbb{R}$ (esto es, como “complejos de los números complejos”). Para ello denotamos $z + jw$ al par de complejos (z,w), consideramos la suma componente a componente, e introducimos el siguiente producto (que extiende al de $\mathbb{C}$): $(z_1 + jw_1).(z_2 + jw_2) := (z_1.z_2 - \bar{w}_1.w_2) + j(\bar{z}_1 w_2 + w_1.z_2)$. Denotando por k a $i.j$, esto es, $k := (i + j0).(0 + j1) = 0 + j(-i)$, se puede reobtener $\mathbb{H}$.

1.2.4. Matrices y sus estructuras algebraicas

Sea K un cuerpo *conmutativo* y $m,n \in \mathbb{N}$. Una MATRIZ de orden $m \times n$ con coeficientes en K es una colección de $m \cdot n$ elementos de K dispuestos rectangularmente y entre paréntesis en m filas "horizontales" y n columnas "verticales". Usualmente, denotaremos a las matrices por letras mayúsculas como A, B, C..., mientras que os elementos o entradas de la matriz aparecerán en letras minúsculas y con subíndices. De este modo, toda matriz de orden $m \times n$ es de la forma:

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

donde cada a_{ij} es un elemento de K. De forma abreviada escribiremos a veces $A = (a_{ij})$ y también $a_{ij} = (A)_{ij}$. Nótese que cada elemento de A tiene dos subíndices: el primero nos indica la fila y el segundo la columna. Por ejemplo, el elemento a_{23} es el que se encuentra en la segunda fila y en la tercera columna. Diremos que dos matrices $A = (a_{ij})$ y $B = (b_{ij})$ de órdenes $m \times n$ y $m' \times n'$ son IGUALES si $m = m'$, $n = n'$ y $a_{ij} = b_{ij}$, para cualesquiera $i = 1, \ldots, m$ y $j = 1, \ldots, n$. Dada $A \in M_{m\times n}$ su matriz TRASPUESTA $A^t \in M_{n\times m}$ es la que se obtiene cambiando las filas de A por columnas, esto es $(A^t)_{ji} = a_{ij}$.

La MATRIZ NULA $0_{m\times n}$ (o, simplemente, 0) es la matriz cuyos elementos coinciden todos con el neutro 0 de K. Una matriz con una sola fila ($m = 1$) se llama MATRIZ FILA O VECTOR FILA. Una matriz con una sola columna ($n = 1$) se llama MATRIZ COLUMNA O VECTOR COLUMNA.

Una matriz con el mismo número de filas que de columnas ($m = n$) se llama MATRIZ CUADRADA de orden n. La DIAGONAL PRINCIPAL de una matriz cuadrada $A = (a_{ij})$ está formada por los elementos a_{ii} con $i = 1, \ldots, n$, a la suma de estos elementos se le llama TRAZA de A, y la denotaremos $\mathrm{tr}(A)$ $(= \sum_{i=1}^{n} a_{ii})$. Una MATRIZ DIAGONAL es una matriz cuadrada tal que todos los elementos que no están en la diagonal principal son cero. Esto significa que $a_{ij} = 0$ para cada $i, j = 1, \ldots, n$ con $i \neq j$. Ejemplos de este tipo de matrices son la MATRIZ (CUADRADA) NULA de orden n, que denotaremos 0_n (0_n=$0_{n\times n}$) y la MATRIZ IDENTIDAD de orden n, que denotaremos por I_n (todos los elementos de la diagonal principal son iguales a 1 y el resto se anulan). Para facilitar algunas expresiones, se introduce el símbolo dependiente de dos índices δ_{ij}, llamado *delta de Kronecker*, el cual es igual a 1 si $i = j$ y a 0 en caso contrario. Por ejemplo, la matriz identidad I_n verifica:

$$(I_n)_{ij} = \delta_{ij} = \begin{cases} 1 & \text{si } i = j \\ 0 & \text{si } i \neq j \end{cases} \qquad \forall i, j \in \{1, \ldots, n\}.$$

Desde un punto de vista formal, una matriz puede definirse simplemente como una aplicación $A : \{1, \ldots, m\} \times \{1, \ldots, n\} \to K$, $(i, j) \mapsto a_{ij}$. Denotaremos por $M_{m\times n}(K)$ al conjunto de todas las matrices de orden $m \times n$ con coeficientes en K. Cuando $m = n$ (matrices cuadradas de orden n) simplificaremos la notación $M_{n\times n}(K)$ por $M_n(K)$.

Dadas dos matrices $A = (a_{ij})$ y $B = (b_{ij})$ en $M_{m\times n}(K)$ definimos su *suma* $A + B$ como la matriz en $M_{m\times n}(K)$ cuyo elemento ij es $a_{ij} + b_{ij}$ (con la suma de K). Esto significa que $A + B = (a_{ij} + b_{ij})$. Por otro lado, dado $a \in K$ y $A = (a_{ij})$ en $M_{m\times n}(K)$, definimos su *producto* $a \cdot A$ como la matriz en $M_{m\times n}(K)$ cuyo elemento ij es $a \cdot a_{ij}$ (con el producto de K). Esto significa que $a \cdot A = (a \cdot a_{ij})$.

Nota 1.22. Es un ejercicio sencillo comprobar que $(M_{m\times n}(K), +)$ tiene una estructura de grupo conmutativo. De hecho, más adelante mostraremos con detalle que, si consideramos también el producto por escalares de K, se obtiene una estructura de lo que llamaremos *espacio vectorial* (ejemplo 2.13).

Dadas dos matrices $A=(a_{ij})\in M_{m\times n}(K), B=(b_{kl})\in M_{n\times p}(K)$ (sobre el mismo cuerpo K y tales que el número de columnas de la primera coincide con el de filas de la segunda) se define su PRODUCTO $A\cdot B=(c_{il})\in M_{n\times p}$ por:

$$(A\cdot B)_{il}=c_{il}=\sum_{j=1}^{n}a_{ij}\cdot b_{jl},\qquad \forall i\in\{1,\cdots,m\},\quad \forall l\in\{1,\cdots p\}.$$

Sus propiedades fundamentales se resumen en el siguiente resultado.

Proposición 1.23. *Sean $A,A'\in M_{m\times n}$, $B,B'\in M_{n\times p}$, $C\in M_{p\times q}$:*

1. *Asociativa: $(A\cdot B)\cdot C=A\cdot(B\cdot C)$.*
2. *Neutro por la derecha: $A\cdot I_n=A$.*
3. *Neutro por la izquierda: $I_m\cdot A=A$.*
4. *Distributiva por la derecha: $A\cdot(B+B')=A\cdot B+A\cdot B'$.*
5. *Distributiva por la izquierda: $(A+A')\cdot B=A\cdot B+A'\cdot B$.*
6. *Producto por 0 (a derecha e izquierda): $A\cdot 0_{n\times p}=0_{m\times p}$, $0_{q\times m}\cdot A=0_{q\times n}$.*
7. *Relación con el producto de K: $a\cdot(A\cdot B)=(a\cdot A)\cdot B=A\cdot(a\cdot B)$.*
8. *Linealidad[19] de la aplicación trasposición ($^t: M_{m\times n}(K)\to M_{n\times m}(K), A\mapsto A^t$):*
   8a. $(A+A')^t=A^t+(A')^t$ (traspuesta de la suma igual a suma de traspuestas).
   8b. $(a\cdot A)^t=a\cdot A^t$ (traspuesta del producto por un escalar igual a escalar por la traspuesta).
9. *Involutividad: $(A^t)^t=A$.*
10. *Relación del producto con la trasposición: $(A\cdot B)^t=B^t\cdot A^t$.*

Demostración. 1. $((A\cdot B)\cdot C)_{il}=\sum_{k=1}^{p}(A\cdot B)_{ik}\cdot c_{kl}=\sum_{k=1}^{p}(\sum_{j=1}^{n}a_{ij}\cdot b_{jk})\cdot c_{kl}=\sum_{k=1}^{p}\left(\sum_{j=1}^{n}(a_{ij}\cdot b_{jk})\cdot c_{kl}\right)$
(donde se usa dos veces la definición del producto de matrices y la distributividad de $\cdot$ en K)
$=\sum_{k=1}^{p}\left(\sum_{j=1}^{n}a_{ij}\cdot(b_{jk}\cdot c_{kl})\right)=\sum_{j=1}^{n}\left(\sum_{k=1}^{p}a_{ij}\cdot(b_{jk}\cdot c_{kl})\right)$
(donde se usa la asociatividad de $\cdot$ en K y la conmutatividad de $+$ en K)
$=\sum_{j=1}^{n}a_{ij}\cdot\left(\sum_{k=1}^{p}(b_{jk}\cdot c_{kl})\right)=\sum_{j=1}^{n}a_{ij}\cdot(B\cdot C)_{jl}=(A\cdot(B\cdot C))_{il}$
(donde se usa la distributividad de $\cdot$ en K y dos veces la definición del producto de matrices).

2. $(A\cdot I)_{il}=\sum_{j=1}^{n}a_{ij}\cdot\delta_{jl}=a_{il}$, donde todos los sumandos en la última expresión se anulan salvo cuando $j=l$, en cuyo caso se tiene $1=\delta_{ll}$.

3-9. Ejercicio inmediato.

10. $((A\cdot B)^t)_{ji}=(A\cdot B)_{ij}=\sum_{k=1}^{n}a_{ik}\cdot b_{kj}=\sum_{k=1}^{n}(A^t)_{ki}\cdot(B^t)_{jk}=\sum_{k=1}^{n}(B^t)_{jk}\cdot(A^t)_{ki}=(B^t\cdot A^t)_{ji}$, donde en las igualdades se usan, resp., las definiciones de traspuesta, de producto de matrices, de traspuesta, la conmutatividad de $\cdot$ en K, y la definición de producto de matrices. □

Observación 1.24. En la proposición anterior, si $m\neq p$ entonces, aunque tenga sentido multiplicar $A\cdot B$ no lo tiene $B\cdot A$. Si $m=p$ entonces $A\cdot B\in M_m(K)$ y $B\cdot A\in M_n(K)$, por lo que los dos productos son necesariamente distintos si $m\neq n$. Pero incluso cuando $m=n$ el producto puede ser diferente:

$$A=\begin{pmatrix}0&1\\0&0\end{pmatrix},\qquad B=\begin{pmatrix}1&0\\0&0\end{pmatrix}\qquad\Longrightarrow\qquad A\cdot B=0_2,\qquad B\cdot A=A(\neq 0_2).$$

[19]Este nombre se justificará con precisión en el Tema 3.

Proposición 1.25. *$M_n(K)$ dotados de la suma y producto de matrices definidos, es un anillo unitario. Para $n \geq 2$, este anillo no es ni conmutativo ni un cuerpo.*

Demostración. Como se ha dicho, $(M_{m\times n}(K), +)$ es siempre un grupo conmutativo, por lo que, en particular, esto ocurre para $m = n$. Las propiedades 1, 4 y 5 de la proposición 1.23 completan la estructura de anillo y las 2,3 el que sea unitario.

La observación 1.24 no sólo muestra con un contraejemplo que el producto no es conmutativo en $M_2(K)$ (para cualquier cuerpo conmutativo K) sino también que 0_2 admite divisores, lo que impide que $(M_2(K), +, \cdot)$ sea un cuerpo (véase el ejercicio 1.19). Este ejemplo puede extenderse trivialmente para matrices cuadradas de orden $n \geq 2$ (añadiendo 0 más allá de la primera submatriz cuadrada de orden 2), lo que completa el resultado. □

Definición 1.26. *Diremos que $A \in M_n(K)$ es* regular *(o* invertible*) si admite inversa para el producto matricial. Al conjunto de todas estas matrices lo llamaremos* grupo lineal general de orden n sobre K, *y lo denotaremos $GL(n,K)$*

Consistentemente con la proposición 1.18, $GL(n,K) = G(M_n(K))$ y tiene estructura de grupo para el producto matricial. Como justificaremos más adelante (sección 1.6.1), si $X \in M_n(K)$ basta con demostrar que existe una matriz cuadrada $\bar{X}$ que verifique bien $X \cdot \bar{X} = I_n$, bien $\bar{X} \cdot X = I_n$, para que $\bar{X}$ sea la inversa de X y, por tanto, $X \in GL(n,K)$.

Ejercicio 1.27. Demuéstrese que si $A \in GL(n,K)$ entonces $A^t \in GL(n,K)$ siendo $(A^t)^{-1} = (A^{-1})^t$.

1.3. Sistemas de Ecuaciones Lineales (Método de Gauss)

1.3.1. Definiciones

Sea K un cuerpo (conmutativo). Una ECUACIÓN LINEAL CON COEFICIENTES EN K es una expresión del tipo

$$a_1 x_1 + \ldots + a_n x_n = b,$$

en la que n es un número natural y $a_1, \ldots, a_n, b$ son elementos de K. Los elementos a_i se llaman *coeficientes*, el número b se llama *término independiente* y las variables x_i son las *incógnitas*. Cuando haya pocas incógnitas las representaremos simplemente por las letras $x, y, z, t, w, \ldots, etc$. Una SOLUCIÓN de la ecuación consiste en una n-úpla $(\alpha_1, \ldots, \alpha_n)$ de elementos de K de forma que al sustituir cada incógnita x_i por el elemento α_i se cumple la ecuación, esto es:

$$a_1 \alpha_1 + \ldots + a_n \alpha_n = b.$$

Se llama SISTEMA DE ECUACIONES LINEALES CON COEFICIENTES EN K (lo escribiremos simplemente SEL para abreviar) a una familia de m ecuaciones lineales sobre K todas ellas con las mismas n incógnitas. Siempre escribiremos un SEL de forma que las incógnitas aparezcan en el mismo orden en todas las ecuaciones. Por tanto, un SEL tiene el siguiente aspecto:

$$\left\{ \begin{array}{ccccccccc} a_{11} x_1 & + & a_{12} x_2 & + & \ldots & + & a_{1n} x_n & = & b_1 \\ a_{21} x_1 & + & a_{22} x_2 & + & \ldots & + & a_{2n} x_n & = & b_2 \\ \vdots & & \vdots & & & & \vdots & & \vdots \\ a_{m1} x_1 & + & a_{m2} x_2 & + & \ldots & + & a_{mn} x_n & = & b_m \end{array} \right. .$$

Una SOLUCIÓN de un SEL es una n-úpla $(\alpha_1,\ldots,\alpha_n)$ de elementos de K que sean solución a la vez de todas las ecuaciones del SEL. También diremos que $x_1=\alpha_1,\ldots,x_n=\alpha_n$ es solución. El problema principal que abordaremos es el de determinar si un SEL tiene soluciones y, en caso afirmativo, calcularlas todas.

Los sistemas de ecuaciones lineales se clasifican según la cantidad de soluciones que tienen. Un SEL se dice INCOMPATIBLE (SI) si no tiene ninguna solución. De lo contrario se dice COMPATIBLE (SC). Un sistema compatible puede ser:

1. SISTEMA COMPATIBLE DETERMINADO (SCD) si tiene una única solución.
2. SISTEMA COMPATIBLE INDETERMINADO (SCI) si tiene más de una solución.

Discutir un SEL consiste en decidir de qué tipo es según la clasificación anterior. Resolver un SEL consiste en calcular todas las soluciones del SEL si las hay. Ya con una ecuación tan sencilla como $ax=b$ se manifiestan todas las opciones posibles de sistemas.

Diremos que un SEL es HOMOGÉNEO si todos sus términos independientes son iguales a cero. Es evidente que un SEL de este tipo siempre será compatible, pues se tiene por lo menos la solución $x_1=0,\ldots,x_n=0$.

1.3.2. Sistemas escalonados

Para resolver cualquier SEL aprenderemos primero a resolver sistemas que sean especialmente sencillos. Estos son los llamados sistemas escalonados.

Diremos que un SEL es ESCALONADO (para una ordenación de las incógnitas) si la primera incógnita de cada ecuación no aparece (esto es, se halla multiplicada por 0) en ninguna de las siguientes ecuaciones.

Nota. Para que la definición anterior tenga sentido es preciso recordar nuestro convenio de que *siempre escribiremos un SEL de forma que las incógnitas aparezcan en el mismo orden en todas las ecuaciones*. De lo contrario, un mismo SEL podría ser a la vez escalonado y no serlo. Por ejemplo, el SEL dado por:

$$\left\{\begin{array}{cccccc} x & +y & +z & = & 0 \\ & y & -z & = & 0 \\ & y & & = & 8 \end{array}\right.$$

no está escalonado, mientras que el mismo SEL escrito como:

$$\left\{\begin{array}{cccccc} x & +y & +z & = & 0 \\ & -z & +y & = & 0 \\ & y & & = & 8 \end{array}\right.$$

sí cumpliría la definición de SEL escalonado. El problema es que este último SEL no está escrito según nuestro convenio y, por tanto, no tiene sentido plantearse si está escalonado o no. Para considerar el anterior sistema como un SEL deberemos escribir en las tres ecuaciones primero la incógnita x, luego la z y por último la y. Es de observar que, en un SEL escalonado, si el número m de ecuaciones fuera mayor que el número n de incógnitas entonces las últimas $m-n$ ecuaciones deben ser del tipo:

$$0x_1+\cdots+0x_n=b_j \qquad \forall j\in\{n+1,\ldots,m\}.$$

Para discutir y/o resolver un SEL escalonado procedemos así:

1. Si el SEL contiene una ecuación del tipo $0x_1 + \cdots + 0x_n = b$ con $b \neq 0$ entonces el sistema es incompatible. De lo contrario el SEL será compatible y pasamos al paso 2.

2. Identificamos las *incógnitas principales* y las *incógnitas secundarias* del SEL. Las principales son las que aparecen como primera incógnita en alguna de las ecuaciones del sistema. Las secundarias son las restantes. Si todas las incógnitas son principales entonces el SEL será compatible determinado. De lo contrario se tratará de un SEL compatible indeterminado. Esto completaría la discusión del SEL.

3. Si no hay incógnitas secundarias entonces la única solución del SEL se calculará despejando directamente las incógnitas principales de abajo hacia arriba. En caso de que haya incógnitas secundarias, entonces se asigna a cada una de ellas un parámetro distinto, y se despejan las incógnitas principales en función de estos parámetros de abajo hacia arriba.

En particular, se deduce del procedimiento anterior que todo SEL escalonado compatible con coeficientes en un cuerpo infinito tiene una única solución o infinitas.

Como ejemplo, vamos a resolver el SEL con coeficientes reales siguiente:

$$\left\{ \begin{array}{ccccccc} x & +3y & +2z & +w & = & -5 \\ & y & +z & +w & = & -2 \\ & & & -w & = & 1 \end{array} \right. ,$$

que es claramente escalonado. La variable z es secundaria mientras que x, y, w son principales. Por tanto, ponemos $z = \lambda$ donde λ es un parámetro real, y la dejamos pasándola al miembro derecho:

$$\left\{ \begin{array}{cccccc} x & + & 3y & +w & = & -5-2\lambda \\ & & y & +w & = & -2-\lambda \\ & & & -w & = & 1 \end{array} \right. \tag{1.4}$$

Como una primera opción, éste es un sistema que se puede resolver directamente, sin más que despejar de abajo hacia arriba las incógnitas x, y, w en función de λ. De hecho, se obtiene directamente $w = -1$ en la tercera ecuación, y sustituyendo en la segunda:

$$y = -1 - \lambda$$

Sustituyendo ahora los valores obtenidos de w, y en la primera:

$$x = -3y - w - 5 - 2\lambda = 3 + 3\lambda + 1 - 5 - 2\lambda = -1 + \lambda.$$

Anticipando la forma de operar del método de Gauss (que veremos a continuación), podemos también resolver (1.4) reduciéndolo progresivamente a un sistema con una ecuación e incóginta menos. De hecho, sumando la última ecuación a las dos primeras, el sistema (1.4) resulta equivalente a resolver:

$$\left\{ \begin{array}{cccccc} x & + & 3y & & = & -4-2\lambda \\ & & y & & = & -1-\lambda \\ & & & -w & = & 1 \end{array} \right. \tag{1.5}$$

y multiplicando ahora la ecuación (1.5) por -3 y sumándosela a la primera, así como la última por -1:

$$\left\{ \begin{array}{ccccc} x & & & = & -1+\lambda \\ & y & & = & -1-\lambda \\ & & w & = & -1 \end{array} \right. \tag{1.6}$$

el cual proporciona directamente la solución.

En cualquier caso, deducimos que estamos ante un SCI con soluciones $x=-1+\lambda, y=-1-\lambda, z=\lambda, \quad w=-1$, para todo $\lambda\in\mathbb{R}$, esto es, $\{(-1+\lambda,-1-\lambda,\lambda,-1) : \lambda\in\mathbb{R}\}$.

Observación 1.28. Es inmediato comprobar que un SEL escalonado es incompatible si y sólo si una de sus ecuaciones es del tipo $0=b$ con $b\neq 0$. En el caso compatible, es determinado si y sólo si todas sus incógnitas son principales.

1.3.3. Sistemas equivalentes. Método de Gauss

Diremos que dos SEL son EQUIVALENTES si tienen el mismo número de incógnitas y exactamente las mismas soluciones. Si el número de incógnitas es n, lo que queremos decir es que una n-úpla $(\alpha_1,\ldots,\alpha_n)$ de elementos de K es una solución del primer SEL si y sólo si lo es también del segundo.

El método de Gauss es un proceso mediante el que cualquier SEL se transforma en un SEL escalonado equivalente. De este modo, las soluciones del nuevo sistema (calculadas por el procedimento descrito en la sección anterior) coinciden con las soluciones del SEL original.

La conversión de un SEL en otro escalonado se lleva a cabo mediante las siguientes TRANSFORMACIONES ELEMENTALES sobre el SEL:

(I) Intercambiar el orden de dos ecuaciones cualesquiera del SEL.

(II) Multiplicar cualquier ecuación del SEL por un elemento de K distinto de cero.

(III) Sustituir una ecuación del SEL por el resultado de sumarle a dicha ecuación otra ecuación del SEL (eventualmente multiplicada antes por un elemento de K, usando el caso (II)).

Evidentemente, debemos asegurarnos de que este tipo de transformaciones sobre el SEL no alteran las soluciones. Esto se prueba en el siguiente resultado:

Proposición: *Las transformaciones anteriormente descritas transforman un SEL en otro SEL equivalente con el original.*

Demostración: En sencilla aunque pesada de escribir. Llamemos S a un SEL dado por:

$$\left\{\begin{array}{ccccccccc}
a_{11}x_1 & + & a_{12}x_2 & + & \ldots & + & a_{1n}x_n & = & b_1 \\
\vdots & & \vdots & & & & \vdots & & \vdots \\
a_{i1}x_1 & + & a_{i2}x_2 & + & \ldots & + & a_{in}x_n & = & b_i \\
\vdots & & \vdots & & & & \vdots & & \vdots \\
a_{j1}x_1 & + & a_{j2}x_2 & + & \ldots & + & a_{jn}x_n & = & b_j \\
\vdots & & \vdots & & & & \vdots & & \vdots \\
a_{m1}x_1 & + & a_{m2}x_2 & + & \ldots & + & a_{mn}x_n & = & b_m
\end{array}\right. .$$

Denotemos por $E_1,\ldots,E_i,\ldots,E_j,\ldots,E_m$ a las ecuaciones del SEL.

Si intercambiamos dos ecuaciones de orden, por ejemplo E_i y E_j, entonces el SEL resultante S' es el siguiente:

$$\left\{\begin{array}{ccccccccc}
a_{11}x_1 & + & a_{12}x_2 & + & \dots & + & a_{1n}x_n & = & b_1 \\
\vdots & & \vdots & & & & \vdots & & \vdots \\
a_{j1}x_1 & + & a_{j2}x_2 & + & \dots & + & a_{jn}x_n & = & b_j \\
\vdots & & \vdots & & & & \vdots & & \vdots \\
a_{i1}x_1 & + & a_{i2}x_2 & + & \dots & + & a_{in}x_n & = & b_i \\
\vdots & & \vdots & & & & \vdots & & \vdots \\
a_{m1}x_1 & + & a_{m2}x_2 & + & \dots & + & a_{mn}x_n & = & b_m
\end{array}\right. ,$$

que es claramente equivalente con S. Así una transformación de tipo (I) nos lleva a un SEL equivalente.

Si multiplicamos una ecuación E_i de S por un elemento $a \in K$ con $a \neq 0$, entonces el SEL resultante S' es el siguiente:

$$\left\{\begin{array}{ccccccccc}
a_{11}x_1 & + & a_{12}x_2 & + & \dots & + & a_{1n}x_n & = & b_1 \\
\vdots & & \vdots & & & & \vdots & & \vdots \\
(a\,a_{i1})x_1 & + & (a\,a_{i2})x_2 & + & \dots & + & (a\,a_{in})x_n & = & a\,b_i \\
\vdots & & \vdots & & & & \vdots & & \vdots \\
a_{m1}x_1 & + & a_{m2}x_2 & + & \dots & + & a_{mn}x_n & = & b_m
\end{array}\right. .$$

Veamos que S es equivalente a S'. Sea $(\alpha_1,\dots,\alpha_n)$ una solución de S. Como S y S' tienen todas las ecuaciones iguales salvo quizás la i-ésima, entonces $(\alpha_1,\dots,\alpha_n)$ es solución de todas las ecuaciones de S' salvo quizás E_i'. Pero también es solución de E_i', pues como $a_{i1}x_1 + a_{i2}x_2 + \dots + a_{in}x_n = b_i$, al multiplicar esta igualdad por a, y usar la propiedad asociativa del producto en K, se llega a que $(a\,a_{i1})x_1 + (a\,a_{i2})x_2 + \dots + (a\,a_{in})x_n = a\,b_i$. Esto prueba que toda solución de S lo es también de S'. Supongamos ahora que $(\alpha_1,\dots,\alpha_n)$ es una solución de S'. Al igual que antes, se sigue que $(\alpha_1,\dots,\alpha_n)$ es solución de todas las ecuaciones de S salvo quizás la i-ésima. Pero también es solución de E_i, pues como $(a\,a_{i1})x_1 + (a\,a_{i2})x_2 + \dots + (a\,a_{in})x_n = a\,b_i$, al multiplicar esta igualdad por a^{-1} (que existe por ser K un cuerpo y $a \neq 0$), y usar la propiedad asociativa del producto en K, se llega a que $a_{i1}x_1 + a_{i2}x_2 + \dots + a_{in}x_n = b_i$. Esto prueba que S y S' son equivalentes. De aquí se concluye que una transformación de tipo (II) nos lleva a un SEL equivalente.

Finalmente, consideremos el tipo (III). Sea S' el SEL que resulta de S cuando la ecuación E_j se sustituye por $a\,E_i + E_j$ con $a \in K$ (realmente bastaría con $a = 1$, pues en caso contrario aplicaríamos primero el tipo (II)). Por las propiedades distributivas en K el SEL obtenido es:

$$\left\{\begin{array}{ccccccccc}
a_{11}x_1 & + & a_{12}x_2 & + & \dots & + & a_{1n}x_n & = & b_1 \\
\vdots & & \vdots & & & & \vdots & & \vdots \\
a_{i1}x_1 & + & a_{i2}x_2 & + & \dots & + & a_{in}x_n & = & b_i \\
\vdots & & \vdots & & & & \vdots & & \vdots \\
(a\,a_{i1}+a_{j1})x_1 & + & (a\,a_{i2}+a_{j2})x_2 & + & \dots & + & (a\,a_{in}+a_{jn})x_n & = & a\,b_i+b_j \\
\vdots & & \vdots & & & & \vdots & & \vdots \\
a_{m1}x_1 & + & a_{m2}x_2 & + & \dots & + & a_{mn}x_n & = & b_m
\end{array}\right. .$$

Sea $(\alpha_1,\dots,\alpha_n)$ una solución de S. Como S y S' tienen todas las ecuaciones iguales salvo quizás la j-ésima, entonces $(\alpha_1,\dots,\alpha_n)$ es solución de todas las ecuaciones de S' salvo quizás de E_j'. Como $(\alpha_1,\dots,\alpha_n)$ cumple E_i y E_j, entonces $a_{i1}x_1 + a_{i2}x_2 + \dots + a_{in}x_n = b_i$ y $a_{j1}x_1 + a_{j2}x_2 + \dots + a_{jn}x_n =$

b_j. Si multiplicamos la primera igualdad por a y la sumamos a la segunda igualdad, obtenemos justamente que $(\alpha_1,\ldots,\alpha_n)$ cumple E'_j. Así, toda solución de S lo es también de S'. Recíprocamente, supongamos que $(\alpha_1,\ldots,\alpha_n)$ es una solución de S'. Al igual que antes, se sigue que $(\alpha_1,\ldots,\alpha_n)$ es solución de todas las ecuaciones de S salvo quizás de E_j. Como $(\alpha_1,\ldots,\alpha_n)$ cumple E'_i y E'_j entonces $a_{i1}x_1+a_{i2}x_2+\ldots+a_{in}x_n=b_i$ y $(a\,a_{i1}+a_{j1})x_1+(a\,a_{i2}+a_{j2})x_2+\ldots+(a\,a_{in}+a_{jn})x_n=a\,b_i+b_j$. Si multiplicamos la primera igualdad por $-a$ y la sumamos a la segunda igualdad, obtenemos justamente que $(\alpha_1,\ldots,\alpha_n)$ cumple E_j. Esto prueba que S y S' son equivalentes y concluye la prueba.

El MÉTODO DE GAUSS para SEL se basa en realizar sobre el SEL tantas transformaciones de tipo (I), (II) o (III) como sea necesario hasta llegar a un SEL escalonado. Por el teorema anterior, el SEL así obtenido es equivalente con el original, por lo que al resolverlo estamos calculando exactamente las soluciones del original.

Ejemplo: Apliquemos el método de Gauss para resolver el SEL con coeficientes en $\mathbb{R}$ dado por:

$$\left\{\begin{array}{cccccc} x & +3y & +2z & = & -5 \\ 3x & +y & -2z & = & 1 \\ 2x & +y & -z & = & 0 \end{array}\right. ,$$

que claramente no está escalonado. Para comenzar a escalonar el SEL necesitamos eliminar la incógnita x de la segunda ecuación y de la tercera ecuación. Para ello, utilizamos transformaciones de tipo (III). Sustituimos la segunda ecuación por el resultado de sumarle la primera multiplicada por -3. Del mismo modo, sustituimos la tercera ecuación por el resultado de sumarle la primera multiplicada por -2. De este modo, llegamos al SEL dado por:

$$\left\{\begin{array}{ccccc} x & +3y & +2z & = & -5 \\ & -8y & -8z & = & 16 \\ & -5y & -5z & = & 10 \end{array}\right. ,$$

que es equivalente con el original pero todavía no está escalonado, ya que la incógnita y es la primera de la segunda ecuación y de la tercera ecuación. En estos momentos observamos que los coeficientes y el término independiente de la segunda ecuación son múltiplos de -8. De igual modo, los coeficientes y el término independiente de la tercera ecuación son múltiplos de 5. Aplicamos tansformaciones de tipo (II) para simplificar las ecuaciones segunda y tercera: dividimos la segunda ecuación por -8 y la tercera por 5. Al hacerlo, llegamos al siguiente SEL equivalente con el original:

$$\left\{\begin{array}{ccccc} x & +3y & +2z & = & -5 \\ & y & +z & = & -2 \\ & -y & -z & = & 2 \end{array}\right. ,$$

que sigue sin estar escalonado ya que la incógnita y es la primera de la segunda ecuación y de la tercera ecuación. Finalmente, para eliminar la incógnita y en la tercera ecuación realizamos una transformación de tipo (III): sustituimos la tercera ecuación por el resultado de sumarla con la segunda ecuación. Así llegamos a este SEL equivalente con el original:

$$\left\{\begin{array}{ccccc} x & +3y & +2z & = & -5 \\ & y & +z & = & -2 \end{array}\right. ,$$

que ya está escalonado. Razonando como en el SEL escalonado que resolvimos en la sección anterior, se tratan x,y como incógnitas principales y z como secundaria. Escribimos entonces $z=\lambda$, y

$$\left\{\begin{array}{ccccc} x & +3y & = & -5 & -2\lambda \\ & y & = & -2 & -\lambda \end{array}\right.$$

éste se puede resolver sin más que susituir en la primera ecuación el valor de y obtenido en la segunda, o bien simplificándolo aún más, multiplicando la segunda fila por -3 y sumándosela a la primera:

$$\left\{\begin{array}{ccccc} x & & = & 1 & +\lambda \\ & y & = & -2 & -\lambda \end{array}\right.$$

En cualquier caso, deducimos que el SEL original es un SCI con soluciones:

$$x = 1+\lambda, \quad y = -2-\lambda, \quad z = \lambda, \quad \text{con } \lambda \in \mathbb{R}.$$

Alternativamente, el conjunto de soluciones del SEL se escribe como:

$$\{(1+\lambda, -2-\lambda, \lambda) \, / \, \lambda \in \mathbb{R}\}.$$

El ejemplo anterior muestra cómo se pueden aplicar transformaciones elementales en un ejemplo concreto para transformar un SEL en otro equivalente escalonado. Para aplicaciones teóricas necesitaremos poder garantizar que esto siempre se puede conseguir para cualquier SEL. Esto se hace por inducción, obteniéndose:

Teorema: *Todo SEL es equivalente a un SEL escalonado mediante una cantidad finita de transformaciones elementales.*

Corolario: *Si un SEL con coeficientes en un cuerpo infinito es compatible indeterminado, entonces tiene infinitas soluciones.*

Demostración: Por el teorema previo el SEL será equivalente a otro escalonado. Y sabemos que un SCI escalonado sobre un cuerpo infinito tiene infinitas soluciones.

1.3.4. Expresión matricial del método de Gauss

En la práctica, uno se da cuenta de que al aplicar el método de Gauss se están escribiendo superfluamente las incógnitas: lo que de verdad importa en todo el proceso son los coeficientes del SEL y sus términos independientes. Para abreviar la escritura se suele desarrollar el método de Gauss con la ayuda de las *matrices* (sección 1.2.4). Consideremos un SEL general dado por:

$$\left\{\begin{array}{ccccccccc} a_{11}x_1 & + & a_{12}x_2 & + & \ldots & + & a_{1n}x_n & = & b_1 \\ a_{21}x_1 & + & a_{22}x_2 & + & \ldots & + & a_{2n}x_n & = & b_2 \\ \vdots & & \vdots & & & & \vdots & & \vdots \\ a_{m1}x_1 & + & a_{m2}x_2 & + & \ldots & + & a_{mn}x_n & = & b_m \end{array}\right. .$$

Se llama MATRIZ DE COEFICIENTES del SEL a la matriz de orden $m \times n$ dada por:

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}.$$

Se llama MATRIZ AMPLIADA del SEL a la matriz de orden $m \times (n+1)$ dada por:

$$(A|b) = \left(\begin{array}{cccc|c} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} & b_m \end{array}\right).$$

Nótese que cada fila de $(A|b)$ codifica una de las ecuaciones del SEL. De hecho, los elementos de la columna j-ésima con $j = 1,\ldots,n$ son los coeficientes del SEL que acompañan a la incógnita x_j. Además, la última columna contiene los términos independientes del SEL. La barra vertical en la matriz ampliada indica dónde aparece el signo igual en cada ecuación del SEL. Es claro que el SEL queda determinado de forma única por su matriz ampliada.

Las transformaciones elementales que se aplican a un SEL para convertirlo en escalonado tienen su contrapartida en el ambiente matricial. Así, establecemos tres tipos de TRANSFORMACIONES ELEMENTALES POR FILAS de una matriz, a saber:

(A) Intercambiar la posición de dos filas (esto se corresponde con una transformación de tipo (I) para el SEL).

(B) Multiplicar una fila por un elemento de K distinto de cero (esto se corresponde con una transformación de tipo (II) para el SEL).

(C) Sustituir una fila por el resultado de sumarle a dicha fila otra fila que ha sido previamente multiplicada por un elemento de K (esto se corresponde con una transformación de tipo (III) para el SEL).

La implementación matricial del método de Gauss tiene tres pasos:

1. Escribir la matriz ampliada del SEL.

2. Efectuar transformaciones elementales por filas hasta obtener la matriz ampliada asociada a un SEL escalonado.

3. Escribir el SEL escalonado equivalente al que se ha llegado y resolverlo.

Ejemplo: Queremos resolver el SEL dado por:

$$\left\{\begin{array}{rrrcr} x & +3y & +2z & = & -5 \\ 3x & +y & -2z & = & 1 \\ 2x & +y & -z & = & 0 \end{array}\right. .$$

Realizamos los siguientes pasos:

1. Escribimos la matriz ampliada del SEL. En nuestro caso particular, obtenemos:

$$\left(\begin{array}{rrr|r} 1 & 3 & 2 & -5 \\ 3 & 1 & -2 & 1 \\ 2 & 1 & -1 & 0 \end{array}\right).$$

Nótese que cada una de las tres primeras columnas proporciona los coeficientes que acompañan a cada una de las incógnitas, mientras que la cuarta columna es la de términos independientes.

2. Realizamos sobre la la matriz ampliada tantas transformaciones elementales por filas como sea necesario hasta obtener la matriz ampliada de un SEL escalonado. En nuestro caso se obtiene:

$$\left(\begin{array}{rrr|r} 1 & 3 & 2 & -5 \\ 3 & 1 & -2 & 1 \\ 2 & 1 & -1 & 0 \end{array}\right) \underset{\substack{(-3)F_1+F_2 \\ (-2)F_1+F_3}}{\sim} \left(\begin{array}{rrr|r} 1 & 3 & 2 & -5 \\ 0 & -8 & -8 & 16 \\ 0 & -5 & -5 & 10 \end{array}\right) \underset{\substack{-\frac{1}{8}F_2 \\ \frac{1}{5}F_3}}{\sim}$$

$$\left(\begin{array}{rrr|r} 1 & 3 & 2 & -5 \\ 0 & 1 & 1 & -2 \\ 0 & -1 & -1 & 2 \end{array}\right) \underset{F_2+F_3}{\sim} \left(\begin{array}{rrr|r} 1 & 3 & 2 & -5 \\ 0 & 1 & 1 & -2 \\ 0 & 0 & 0 & 0 \end{array}\right).$$

3. En este punto se puede escribir el SEL obtenido y resolver. En nuestro ejemplo el SEL al que llegaríamos es:
$$\left\{\begin{array}{ccccc} x & +3y & +2z & = & -5 \\ & y & +z & = & -2 \end{array}\right. ,$$
ya solucionado en la sección anterior. Reescribiendo matricialmente el proceso:
$$\left(\begin{array}{ccc|c} 1 & 3 & 2 & -5 \\ 0 & 1 & 1 & -2 \\ 0 & 0 & 0 & 0 \end{array}\right) \overset{F_1-3F_3}{\sim} \left(\begin{array}{ccc|c} 1 & 0 & -1 & 1 \\ 0 & 1 & 1 & -2 \\ 0 & 0 & 0 & 0 \end{array}\right) \sim \left(\begin{array}{ccc|c} 1 & 0 & -1 & 1 \\ 0 & 1 & 1 & -2 \end{array}\right),$$
donde la última ecuación se ha suprimido obteniéndose un sistema equivalente, que se puede resolver directamente. Merece comentarse que, continuando matricialmente, podemos tomar la coordenada secundaria como un parámetro $z = \lambda \in \mathbb{R}$ y reescribir la matriz del sistema escalonado, para cada λ, como
$$\left(\begin{array}{cc|c} 1 & 0 & 1+\lambda \\ 0 & 1 & -2-\lambda \end{array}\right),$$
generándose directamente la solución $x = 1+\lambda, y = -2-\lambda$, esto es,
$$\{(1+\lambda, -2-\lambda, \lambda) \,/\, \lambda \in \mathbb{R}\}.$$

Nota: ¿Es posible realizar transformaciones elementales *por columnas* para resolver un SEL? La respuesta es no y la razón es muy sencilla: este tipo de transformaciones no convierten un SEL dado en otro equivalente. Por ejemplo, consideremos el SEL con coeficientes reales dado por:
$$\left\{\begin{array}{ccccc} x & + & y & = & 2 \\ x & - & y & = & 0 \end{array}\right. ,$$
que es un SCD con solución $x = y = 1$. Su matriz ampliada es:
$$(A|b) = \left(\begin{array}{cc|c} 1 & 1 & 2 \\ 1 & -1 & 0 \end{array}\right).$$
Al multiplicar por 2 la segunda columna obtendríamos la matriz:
$$\left(\begin{array}{cc|c} 1 & 2 & 2 \\ 1 & -2 & 0 \end{array}\right),$$
que es la matriz ampliada del SEL con coeficientes reales dado por:
$$\left\{\begin{array}{ccccc} x & + & 2y & = & 2 \\ x & - & 2y & = & 0 \end{array}\right. ,$$
el cual es un SCD con solución $x = 1$, $y = 1/2$, distinta a la anterior.

Ejercicio: Discutir y resolver el siguiente SEL con coeficientes reales:
$$\left\{\begin{array}{ccccccccccc} x & + & 2y & & & + & s & & & = & 7 \\ & & y & + & z & - & 2s & + & t & = & 2 \\ 2x & - & y & + & z & + & 4s & + & t & = & 0 \end{array}\right. .$$

Ejercicio: Discutir y resolver el siguiente SEL con coeficientes complejos:

$$\left\{\begin{array}{rcrcrcl} ix & + & 2iy & + & z & = & 0 \\ x & - & 3y & + & 2z & = & 0 \\ 2ix & & & + & (i+1)z & = & i \end{array}\right. .$$

Ejercicio: Discutir y resolver el siguiente SEL con coeficientes reales:

$$\left\{\begin{array}{rcrcrcr} & - & 4y & - & z & = & -7 \\ x & + & y & + & z & = & 2 \\ x & - & 2y & + & z & = & -2 \\ -x & + & 2y & & & = & 3 \end{array}\right. .$$

1.3.5. Cálculo de la matriz inversa por el método de Gauss-Jordan

Aunque más adelante se verá otro método de cálculo de la matriz inversa (y se discutirá con detalle cuándo una matriz admite inversa), conviene estudiar ahora este problema por el procedimiento de Gauss-Jordan, basado en las propiedades estudiadas de los SEL.

Sea $A \in M_n(K)$ una matriz regular. Calcular su inversa equivale a solucionar la ecuación de n^2 incógnitas[20] (x_{ij}):

$$A \cdot X = I_n, \qquad \text{donde} \qquad X = \begin{pmatrix} x_{11} & \cdots & x_{1n} \\ \cdots & \cdots & \cdots \\ x_{n1} & \cdots & x_{nn} \end{pmatrix}.$$

Obsérvese que la ecuación puede interpretarse diciendo que al multiplicar A por cada una de las columnas de X se obtiene cada una de las columnas de I_n, por lo que se tiene el SEL de n^2 ecuaciones y n^2 incógnitas:

$$\left(\begin{array}{c|c|c} A & \cdots & 0 \\ \hline \cdots & \cdots & \cdots \\ \hline 0 & \cdots & A \end{array}\right) \left(\begin{array}{c} x_{11} \\ \cdots \\ x_{n1} \\ \hline \cdots \\ \hline x_{1n} \\ \cdots \\ x_{nn} \end{array}\right) = \left(\begin{array}{c} 1 \\ \cdots \\ 0 \\ \hline \cdots \\ \hline 0 \\ \cdots \\ 1 \end{array}\right)$$

Por ser A regular, este sistema debe ser compatible y determinado[21]. Por tanto, puede resolverse por el método de Gauss-Jordan realizando transformaciones elementales por filas. Obsérvese que las mismas transformaciones elementales que se usen para las primeras n filas se pueden usar para cada uno de los siguientes grupos de n filas. Por esta razón, reescribiremos la matriz ampliada del sistema original como:

$$(A|I_n) = \left(\begin{array}{ccc|ccc} a_{11} & \cdots & a_{1n} & 1 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{n1} & \cdots & a_{nn} & 0 & \cdots & 1 \end{array}\right)$$

Bastará entonces con realizar transformaciones elementales por filas hasta obtener la matriz identidad I_n en la parte izquierda de la matriz (la cual sería la matriz de Hermite para cada uno de los n sistemas

[20] Véase la discusión en la sección 1.6.1, especialmente el punto 2.

[21] De hecho, consistentemente con la sección 1.6.1: (a) por ser A regular $|A| \neq 0$, (b) teniendo en cuenta el ejercicio 1.58, la primera matriz de la fórmula anterior tiene determinante $|A|^n \neq 0$, por lo que también es regular y (c) multiplicando por la inversa de esta matriz a la izquierda de ambos miembros de la igualdad se obtendrá la (única) solución del SEL.

de n ecuaciones con n^2 incógnitas en que se descompone el original). En ese momento, cada una de las columnas a la derecha será la solución del SEL de n^2 ecuaciones para la correspondiente columna de incógnitas de X, esto es, se obtendrá a la derecha la matriz buscada A^{-1}. (Por supuesto, si esto no pudiera llevarse a cabo la matriz A no sería regular).

En el siguiente ejemplo, el lector puede ir comprobando en un caso particular el procedimiento general antes descrito. Estudiemos si $A \in \mathrm{GL}(n,K)$, donde

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

Para ello, obsérvese que

$$(A|I_2) = \left(\begin{array}{cc|cc} 1 & 2 & 1 & 0 \\ -1 & 2 & 0 & 1 \end{array}\right) \sim \left(\begin{array}{cc|cc} 1 & 2 & 1 & 0 \\ 0 & 4 & 1 & 1 \end{array}\right)$$

Así, A es regular. Usando transformaciones elementales de abajo hacia arriba

$$(A|I_2) \sim \left(\begin{array}{cc|cc} 1 & 2 & 1 & 0 \\ 0 & 1 & 1/4 & 1/4 \end{array}\right) \sim \left(\begin{array}{cc|cc} 1 & 0 & 1/2 & -1/2 \\ 0 & 1 & 1/4 & 1/4 \end{array}\right)$$

En consecuencia, la inversa de A es:

$$A^{-1} = \begin{pmatrix} 1/2 & -1/2 \\ 1/4 & 1/4 \end{pmatrix}.$$

1.4. El grupo de permutaciones

Seguimos aquí como referencias el libro de M. Castellet e I. Llerena "Álgebra Lineal y Geometría" Ed. Reverté, Barcelona, 2000) y el de Ángel Primo Martínez "Matemáticas. Curso de Orientación Universitaria", Reverté SM, 1987.[22]

1.4.1. Conceptos básicos

Sea $n \in \mathbb{N}$ y $A = \{1,\ldots,n\}$. Una *permutación* de A es cualquier aplicación biyectiva $\sigma : A \to A$; denotaremos S_n al conjunto de todas las permutaciones de[23] A. Como se explicó en el tema de preliminares, la composición de dos aplicaciones biyectivas es biyectiva y el conjunto de todas las aplicaciones biyectivas de un conjunto en sí mismo tiene estructura de grupo con la composición. En resumen:

Proposición 1.29. $(S_n,\circ)$ *tiene estructura de grupo, al que se llamará* grupo de las permutaciones (de n elementos)*, y se simplificará la notación denotándolo solamente* S_n*.*

En el contexto de permutaciones, a la composición se le suele llamar *producto*, y se denota:

$$\sigma = \begin{pmatrix} 1 & 2 & \cdots & n \\ \sigma(1) & \sigma(2) & \cdots & \sigma(n) \end{pmatrix}$$

[22]Desarrollos más extensos en álgebra pueden consultarse en los libros de S. Mc Lane y G. Birkhoff, N. Jacobson ó J.K. Goldhaber y G. Ehrlich.

[23] La definición y todo el desarrollo se puede generalizar trivialmente a cualquier conjunto $A = \{a_1,\ldots,a_n\}$ con n elementos (se escribe entonces S_A ó $\mathrm{Biy}(A)$); este caso general se obtiene sin más que considerar $1,\ldots,n$ como los subíndices de $a_1,\ldots,a_n$.

Ejemplo 1.30. Son permutaciones en S_3:

$$\sigma = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}, \; \tau = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix}, \; \sigma \circ \tau = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix}, \; \tau \circ \sigma = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}.$$

Nota. Obviamente, sólo la segunda fila en nuestra representación de la permutación resulta relevante. Así, por ejemplo, σ sería representable sólo por $(3\ 1\ 2)$ o la n-úpla $(3,1,2)$. En ese caso, tendría que tenerse presente que la permutación asociada es la que aplica cada i a la componente i-ésima de la n-úpla (que no debe confundirse con la permutación inversa, la cual aplicaría cada i al ordinal de su posición dentro n-úpla). No obstante, no seguiremos esa notación, que se podría confundir con la de los ciclos más adelante.

Ejercicio 1.31. *(1) Constrúyanse explícitamente todos los elementos de S_1, S_2 y S_3.*

(2) Demuéstrese por inducción que el cardinal (número de elementos) de S_n es $n! := 1 \cdot 2 \cdot \cdots \cdot (n-1) \cdot n$.

(3) Compruébese que S_n es conmutativo si y sólo si $n \leq 2$.

Nota 1.32. Aunque fuera de nuestros objetivos, el lector puede extender el ejercicio anterior comprobando las siguientes propiedades elementales de combinatoria para un conjunto A de n elementos[24].

(1) *Variaciones*. A una m-úpla (subconjunto ordenado) formada por $m \leq n$ elementos distintos de A se le llama *variación (sin repetición) de n elementos tomados de m en m* y al número de todas las variaciones construibles con los elementos de A lo denotaremos $V_{n,m}$. Se verifica:

$$V_{n,m} = n \cdot (n-1) \cdots (n-m+1) = \frac{n!}{(n-m)!}. \tag{1.7}$$

$V_{n,m}$ coincide con el número de aplicaciones inyectivas de un conjunto de m elementos en uno de n elementos. Cuando $m = n$ cada variación es una permutación y se tiene $V_{n,n} = n!$ (por convenio, se define $0! = 1$ de modo que la fórmula (1.7) siga siendo válida).

(2) *Combinaciones*. A un subconjunto (no ordenado) formado por $m \leq n$ elementos (distintos) de A se le llama *combinación (sin repetición) de n elementos tomados de m en m* y al número de todas las combinaciones construibles con los elementos de A lo denotaremos $C_{n,m}$. Se verifica:

$$C_{n,m} = \frac{V_{n,m}}{m!} = \frac{n!}{m!(n-m)!} =: \begin{pmatrix} n \\ m \end{pmatrix} \qquad \text{(léase “número combinatorio } n \text{ sobre } m\text{”).}$$

De la fórmula anterior es inmediato comprobar $C_{n,m} = C_{n,n-m}$ (que puede obtenerse también percatándose de que cada vez que formamos un subconjunto con m elementos, los restantes elementos de A forman un subconjunto de $n-m$ elementos) y $C_{n,m} = C_{n-1,m-1} + C_{n-1,m}$ (si nos fijamos en un elemento, el primer sumando es el número de combinaciones en las que participa y el segundo el número en las que no participa), lo cual permite calcular recursivamente $C_{n,m}$ (*triángulo de Tartaglia*).

(3) *Binomio de Newton*. Si $x, y \in K$ (cuerpo conmutativo):

$$(x+y)^n = \sum_{m=0}^{n} \begin{pmatrix} n \\ m \end{pmatrix} x^{n-m} \cdot y^m$$

(procédase por inducción usando las propiedades de los números combinatorios). En particular,

$$2^n = \sum_{m=0}^{n} \begin{pmatrix} n \\ m \end{pmatrix} = C_{n,0} + \ldots + C_{n,n},$$

que es el cardinal de $\mathcal{P}(A)$, esto es, el número de subconjuntos incluidos en A.

[24] Ampliable en Wikipedia `https://es.wikipedia.org/wiki/Combinatoria`

Un elemento $j \in \{1,\ldots,n\}$ se llamará *fijo* por la permutación $j \in S_n$ si $\sigma(j) = j$, y *no fijo* en caso contrario. Si j no es fijo, podemos construir la sucesión:

$$j, \sigma(j), \sigma^2(j), \ldots \sigma^r(j) \ldots$$

(donde $\sigma^2 := \sigma \circ \sigma$ y $\sigma^r := \sigma \circ \sigma^{r-1}$). Dado que $A = \{1,\ldots,n\}$ es finito, en algún momento $\sigma^k(j)$ coincidirá con los anteriores.

Lema 1.33. *Sea $\sigma \in S_n$ y $j \in \{1,\ldots,n\}$ no fijo. El primer $r > 1$ tal que $\sigma^r(j) \in \{j, \sigma(j)\ldots, \sigma^{r-1}(j)\}$ debe cumplir $\sigma^r(j) = j$.*

Demostración. En efecto, si $\sigma^r(j) = \sigma^h(j)$ y $r > h \geq 1$ entonces, componiendo $r-h$ veces con la inversa σ^{-1} se tiene $\sigma^{r-h}(j) = j$ (esto es, j ya habría salido antes). ∎

Sean pues $j, \sigma(j), \sigma^2(j), \ldots \sigma^{r-1}(j)$ distintos con $\sigma^r(j) = j$. Diremos que $\sigma \in S_n$ es un *ciclo de orden r* si deja fijos todos sus elementos salvo r de ellos que se pueden escribir como en la sucesión anterior. Esto es, un ciclo es una permutación cuyos elementos no fijos se reordenan circularmente:

$$j \to \sigma(j) \to \sigma^2(j) \cdots \to \sigma^{r-1}(j) \to j.$$

Como notación simplificada para ciclos, se escribe la r-úpla

$$\sigma = (j, \sigma(j), \sigma^2(j), \ldots \sigma^{r-1}(j))$$

(obsérvese que en esta notación no aparece explícitamente el valor n de S_n; en cualquier caso n no puede ser menor que ninguno de los elementos de la r-úpla). A un ciclo de orden 2 se le llama *trasposición*. Una trasposición (j,k), $1 \leq j < k \leq n$ se dice *de índices contiguos* cuando $k = j+1$. A un ciclo de orden n (en S_n) se le llama *permutación cíclica*.

Ejemplo 1.34. En S_5 haciendo corresponder $1 \to 2 \to 4 \to 3 \to 5 \to 1$ se tiene la permutación:

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 4 & 5 & 3 & 1 \end{pmatrix}, \qquad \text{esto es} \qquad (1,2,4,3,5) \in S_5$$

es un ciclo de orden 5 (o permutación cíclica). Análogamente $1 \to 2 \to 3 \to 4 \to 1$ proporciona:

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 3 & 4 & 1 & 5 \end{pmatrix}, \qquad \text{esto es} \qquad (1,2,3,4) \in S_5,$$

que es un ciclo de orden 4.

Ejemplo 1.35. Las permutaciones de σ y τ del ejemplo 1.30 son ciclos:

$$\sigma = (1,3,2) \in S_3, \qquad\qquad \tau = (2,3) \in S_3,$$

siendo además τ una trasposición. Es fácil comprobar que las seis permutaciones de S_3 son ciclos (admitiendo la identidad como el ciclo trivial de orden 1). Claramente, esto no ocurre en S_n si $n \geq 4$.

Se dice que dos ciclos $\sigma, \sigma' \in S_n$ son *disjuntos* cuando ningún $j \in \{1,\ldots,n\}$ es, a la vez, un elemento no fijo para σ y σ' (esto es, los elementos no fijos de σ son fijos para σ' y viceversa). Claramente, dos ciclos disjuntos conmutan (es decir, $\sigma \circ \sigma' = \sigma' \circ \sigma$). En adelante, al usar la notación para ciclos, omitiremos el símbolo de composición siempre que no haya lugar a confusión.

Ejemplo 1.36. Los ciclos $\sigma = (1,2), \sigma = (3,4) \in S_4$ son disjuntos. Se tiene:

$$\sigma \circ \sigma' = (1,2)(3,4) = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 1 & 4 & 3 \end{pmatrix} = (3,4)(1,2) = \sigma' \circ \sigma$$

Proposición 1.37. *(1) Toda permutación $\sigma \in S_n$ se puede expresar como producto (composición) de ciclos disjuntos dos a dos.*

(2) Todo ciclo de orden r se puede expresar como producto de $r-1$ trasposiciones.

(3) Toda permutación se puede expresar como producto de trasposiciones.

Demostración. (1) Si σ no es la identidad, se toma un elemento no fijo, j_1, y se genera el ciclo

$$\sigma_1 := (j_1, \sigma(j_1), \sigma^2(j_1), \ldots, \sigma^{r_1 - 1}(j_1))$$

(donde $\sigma^{r_1}(j_1) = j_1$ y todos los elementos anteriores son distintos). Si $\sigma = \sigma_1$ se obtiene el resultado; en caso contrario existirá un elemento j_2 que es fijo para σ_1 y no fijo para σ, y que generará análogamente un ciclo

$$\sigma_2 := (j_2, \sigma(j_2), \sigma^2(j_2), \ldots . \sigma^{r_2 - 1}(j_2)).$$

Este segundo ciclo, necesariamente, es disjunto de σ_1, pues si $\sigma^{k_1}(j_1) = \sigma^{k_k}(j_2)$ entonces $\sigma^{k_1 - k_2}(j_1) = j_2$ (y j_2 no sería fijo para σ_1).

Procediendo inductivamente, una vez obtenidos m ciclos disjuntos $\sigma_1, \ldots, \sigma_m$, si todo elemento no fijo j de σ es también no fijo para alguno de los ciclos, entonces $\sigma = \sigma_1 \circ \cdots \circ \sigma_m$ (pues j sólo puede ser no fijo para uno de los ciclos σ_{m_j} y, por construcción, $\sigma_{m_j}(j) = \sigma(j)$). En caso contrario, se puede construir un nuevo ciclo disjunto σ_{m+1}, hasta agotar todos los elementos no fijos de σ en un número finito de pasos.

(2) Basta con comprobar (inductivamente):

$$(j_1, j_2, \ldots, j_{r-1}, j_r) = (j_1, j_2)(j_2, j_3) \ldots (j_{r-2}, j_{r-1})(j_{r-1}, j_r).$$

(3) Inmediato de los dos puntos anteriores. ∎

Observación 1.38. Merece también comentarse que toda trasposición $(j, j+h)$, $1 \le j < j+h \le n$ se puede expresar como producto de $2h-1$ trasposiciones de índices contiguos, concretamente:

$$(j, j+h) = \underbrace{(j, j+1)(j+1, j+2) \ldots (j+h-2, j+h-1)}\underbrace{(j+h, j+h-1) \ldots (j+2, j+1)(j+1, j)}.$$

Ello se comprueba inductivamente. Informalmente, el producto de las h trasposiciones de índices contiguos a la derecha hace “avanzar” h posiciones a la imagen del elemento j (hasta “dejar atrás” a $j+h$, ocupando esa posición) y, a continuación, el de las $h-1$ trasposiciones a la izquierda hace “retroceder” a la imagen del elemento $j+h$ (hasta ocupar la posición j-ésima).

Ejemplo 1.39. Como casos particulares:

- $\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 3 & 6 & 4 & 8 & 5 & 7 & 2 & 1 \end{pmatrix} = (1,3,4,8)(2,6,7) = (2,6,7)(1,3,4,8)$
- $(1,3,4,8) = (1,3)(3,4)(4,8)$
- $(2,6) = (2,3)(3,4)(4,5)\ (6,5)(5,4)(4,3)(3,2)$.

1.4.2. Paridad y signo

La permutación identidad I y, por tanto, cualquier otra, se puede expresar de muchas maneras como producto de trasposiciones. Nuestro objetivo es comprobar que la *paridad* del número p de trasposiciones (esto es, el carácter par o impar de p) es la misma en todas esas expresiones.

Lema 1.40. *La permutación identidad I no se puede expresar como producto de un número impar de trasposiciones.*

Demostración. Consideremos el siguiente producto

$$P := \Pi_{1 \leq i < j \leq n}(j - i)$$

donde $i, j \in \{1, \ldots, n\}$ (obsérvese $P \neq 0$). Dada una permutación, $\sigma \in S_n$ definimos:

$$\sigma P := \Pi_{1 \leq i < j \leq n}(\sigma(j) - \sigma(i)).$$

La demostración del lema se basa en el siguiente aserto, aparentemente anecdótico, que se comprobará aparte por discusión de casos:

Aserto. Si σ es una trasposición entonces $\sigma P = -P$.

Supongamos entonces que I se escribe como una composicón

$$I = \tau_s \circ \cdots \circ \tau_2 \circ \tau_1$$

donde las τ_i son s trasposiciones. Por la definición de σP es claro que $IP = P$; sin embargo, aplicando el aserto sucesivamente a cada una de las s trasposiciones se tiene:

$$P = (-1)^s P.$$

Y, como $P \neq 0$, necesariamente, s es par. $\square$

Discusión del aserto. Sea $\sigma = (h,k), h < k$ y comprobemos $\sigma P = -P$. Por conveniencia del lector, el razonamiento general se irá ilustrando con el caso $n = 5$, $\sigma = (2,4)$. Así, $P = P_5$ es:

$$P_5 = (2-1)(3-1)(4-1)(5-1)(3-2)(4-2)(5-2)(4-3)(5-3)(5-4).$$

Veamos el efecto de $\sigma = (h,k)$ sobre cada factor $(j-i)$ con $1 \leq i < j \leq n$:

- *Ningún efecto.* Si $i \neq h,k,\ j \neq h,k$ entonces $(j-i) = (\sigma(j) - \sigma(i))$.

  Para P_5, $\sigma = (2,4)$: los factores $(3-1), (5-1), (5-3)$ quedan inalterados.

- *Cambio de orden en la expresión de P dos factores* (no afecta a P). Ocurre en dos subcasos:

  - Para cada $i < h$, cuando $j = h$ y cuando $j = k$: cambian de orden $(h-i)$, $(k-i)$.
    Para P_5, $\sigma = (2,4)$: cambian de orden $(2-1)$ y $(4-1)$.
  - Para cada $j > k$, cuando $i = h$ y cuando $i = k$: cambian de orden $(j-h)$, $(j-k)$.
    Para P_5, $\sigma = (2,4)$: cambian de orden $(5-2)$ y $(5-4)$.

- *Cambio de signo y de orden de dos factores* (no afecta a P). Para cada l tal que $h < l < k$ el factor con $i = h, j = l$ y el factor con $i = l, j = k$ (esto es, $(l-h)$ y $(k-l)$) se cambian el uno por otro, cambian además de signo.

  Para P_5, $\sigma = (2,4)$: se cambia $(3-2)$ por $(3-4)$, y $(4-3)$ por $(2-3)$.

- *Cambio de signo de un único factor* (y, por tanto, del signo de todo P). Para $i = h, j = k$ se cambia el factor $(k-h)$ por $(h-k)$.

  Para P_5, $\sigma = (2,4)$: se cambia $(4-2)$ por $(2-4)$.

En resumen, sólo en el último de los casos se produce un cambio de signo que afecte a P. ∎

Teorema 1.41. *Si σ se expresa como composición de p trasposiciones y de q trasposiciones,*

$$\sigma = \tau_p \circ \tau_{p-1} \circ \cdots \circ \tau_2 \circ \tau_1, \qquad \sigma = \rho_q \circ \rho_{q-1} \circ \cdots \circ \rho_2 \circ \rho_1,$$

entonces p y q tienen la misma paridad, esto es, o ambos son pares o ambos impares.

Demostración. Para toda trasposición τ se tiene $\tau \circ \tau = I$, esto es, $\tau^{-1} = \tau$. Por tanto, usando la primera expresión de σ se tiene $\sigma^{-1} = \tau_1 \circ \tau_2 \circ \cdots \circ \tau_{p-1} \circ \tau_p$. En consecuencia:

$$I = \sigma^{-1} \circ \sigma = (\tau_1 \circ \tau_2 \circ \cdots \circ \tau_{p-1} \circ \tau_p) \circ (\rho_q \circ \rho_{q-1} \circ \cdots \circ \rho_2 \circ \rho_1).$$

Esto es, I es producto de $p+q$ trasposiciones y basta usar que, por el lema, $p+q$ es par. ∎

Este teorema asegura la consistencia de la siguiente definición.

Definición 1.42. *Una permutación σ se llama* par *si se puede expresar como producto de un número par de trasposiciones, e* impar *si como producto de un número impar de trasposiciones.*

Proposición 1.43. *La* aplicación signo *(o signatura) sig definida por:*

$$sig : S_n \to \{+1, -1\}, \qquad sig(\sigma) = \begin{cases} +1 & si\ \sigma\ es\ par \\ -1 & si\ \sigma\ es\ impar, \end{cases}$$

verifica las siguientes propiedades[25]*:*

(i) $sig(\sigma \circ \tau) = sig(\sigma) \cdot sig(\tau)$

(ii) $sig(I) = 1$, $sig\ (\sigma^{-1}) = sig(\sigma)$.

(iii) Si σ en un ciclo de orden r, entonces $sig(\sigma) = (-1)^{r-1}$.

Demostración. (i) Inmediato de que si σ y τ se expresan como producto de p y q trasposiciones, entonces $\sigma \circ \tau$ se expresa como producto de $p+q$ trasposiciones.

(ii) Inmediato.

(iii) Úsese la prop. 1.37 (2). ∎

Ejercicio 1.44. *Demuéstrese que en cada grupo S_n con $n \geq 2$ el número de permutaciones pares es igual al de las impares.*

[25] La propiedad (i) es equivalente a decir que *sig* es un *homomorfismo de grupos*, y la (ii) se puede ver como un caso particular de una propiedad general para cualquier homomorfismo de grupos.

1.4.3. Número de inversiones

Describimos a continuación una forma sistemática de calcular la signatura de una permutación.

Definición 1.45. *Dada $\sigma \in S_n$, diremos que dos elementos $i < j$ de A presentan un* permanencia *si $\sigma(i) < \sigma(j)$ y una* inversión *si $\sigma(i) > \sigma(j)$. El* número de inversiones de σ *es el número total de inversiones obtenidos con todos los pares (i, j) tales que $1 \leq i < j \leq n$, y se denotará $[\sigma]$.*

Así, para calcular $[\sigma]$ basta con contar el número de total de alteraciones del orden natural en la segunda fila de $\begin{pmatrix} 1 & 2 & \cdots & n \\ \sigma(1) & \sigma(2) & \cdots & \sigma(n) \end{pmatrix}$.

Ejemplo 1.46. Para calcular el número de inversiones de

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 3 & 4 & 2 & 1 & 5 \end{pmatrix}$$

basta con contar para cada elemento k de la segunda fila ($k = \sigma(i)$) cuántos elementos l ($l = \sigma(j)$) menores que k (esto es, $\sigma(i) > \sigma(j)$) aparecen a su derecha (de modo que $i < j$) y sumar. En concreto:

Para $k = 5$: cero.
Para $k = 4$: dos (2 y 1)
Para $k = 3$: dos (2 y 1)
Para $k = 2$: uno (el 1)
Para $k = 1$: cero (necesariamente)

Por tanto, $[\sigma] = 0 + 2 + 2 + 1 + 0 = 5$.

Proposición 1.47. *Toda permutación $\sigma \in S_n$ se puede expresar como producto de $[\sigma]$ trasposiciones. Por tanto, $sig(\sigma) = (-1)^{[\sigma]}$.*

Demostración. Razonando por inducción sobre n, el resultado es trivial para $n = 1$, y supongámoslo válido para $n-1$. Sea i tal que $\sigma(i) = n$, por lo que que i presenta inversiones con $j = i+1, \ldots, n$. Compongamos σ con las correspondientes $n-i$ trasposiciones, obteniendo

$$\tilde{\sigma} := ((\sigma(n), \sigma(i))(\sigma(n-1), \sigma(i)) \ldots (\sigma(i+2), \sigma(i))(\sigma(i+1), \sigma(i))) \circ \sigma \tag{1.8}$$

Como $\sigma(i) = n$, se tiene $[\tilde{\sigma}] = [\sigma] - (n-i)$ y también $\tilde{\sigma}(n) = n$. Esta última igualdad permite usar la hipótesis de inducción y expresar $\tilde{\sigma}$ como producto de $[\tilde{\sigma}]$ trasposiciones. Sustituyendo esta expresión en (1.8) y despejando se obtiene una expresión de σ como producto de $[\tilde{\sigma}] + (n-i) = [\sigma]$ trasposiciones. La última afirmación es inmediata de la definición de la aplicación sig. ∎

Ejemplo 1.48. Usando el procedimiento de la demostración anterior, expresemos $\sigma = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix}$ como composición de $[\sigma] = 3$ trasposiciones. Observemos primero $\sigma(1) = 3$ y

$$\tilde{\sigma} = (3,1)(3,2) \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}.$$

Se tiene entonces $\tilde{\sigma}(3) = 3$ y $[\tilde{\sigma}] = 3 - 2 = 1$. Repitiendo el proceso para $\tilde{\sigma}$ se tiene:

$$I = (2,1)(3,1)(3,2) \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix} \qquad \text{esto es,} \qquad \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix} = (3,2)(3,1)(2,1).$$

Nota 1.49. Es de remarcar que $[\sigma]$ no es el mínimo número de trasposiciones necesario para expresar σ (de hecho, en el ejemplo anterior σ era directamente una trasposición, $\sigma = (1,3)$). El mínimo es siempre $\leq n-1$, como puede comprobarse escribiendo la permutación como composición de ciclos disjuntos, y cada ciclo de orden r como composición de $r-1$ trasposiciones (proposición 1.37).

1.5. Determinante de una matriz cuadrada

El concepto de determinante de una matriz de cuadrada $A \in M_n(K)$ sobre un cuerpo K se desarrollará más adelante en relación con los tensores determinantes y endomorfismos. No obstante, por su utilidad práctica lo introducimos a a continuación, y mostramos algunas de sus propiedades elementales. En toda esta sección, K será un cuerpo conmutativo de característica distinta de 2. Anticipando la nomenclatura que usaremos en espacios vectoriales, llamaremos *escalares* a los elementos del cuerpo K.

1.5.1. Noción de determinante

Definición 1.50. *Sea $A = (a_{ij}) \in M_n(K)$. Se define el* determinante de A *como*

$$\det A := \sum_{\sigma \in S_n} \mathrm{sig}(\sigma)\, a_{\sigma(1)\,1} \dots a_{\sigma(n)\,n} \tag{1.9}$$

Usaremos también las notaciones:

$$\det A = |A| = \begin{vmatrix} a_{11} & \cdots & a_{1n} \\ \cdots & \cdots & \cdots \\ a_{n1} & \cdots & a_{nn} \end{vmatrix}$$

(el símbolo $|\cdot|$ no debe confundirse aquí con "valor absoluto" o "módulo"). Obsérvese que en cada sumando de la definición aparece un único elemento de cada fila y de cada columna. El siguiente resultado muestra que el papel de filas y columnas es intercambiable en la definición de determinante.

Proposición 1.51. $\det A = \det A^t$*, para toda $A \in M_n(K)$.*

Demostración. Sea $A = (a_{ij})$ y $A^t = (b_{ij})$ (esto es, $b_{ij} = a_{ji}$). Se tiene entonces:

$$\begin{aligned} \det A &= \textstyle\sum_{\sigma \in S_n} \mathrm{sig}(\sigma)\, a_{\sigma(1)1} \dots a_{\sigma(n)n} \\ &= \textstyle\sum_{\sigma \in S_n} \mathrm{sig}(\sigma)\, a_{1\,\sigma^{-1}(1)} \dots a_{n\,\sigma^{-1}(n)} = \sum_{\sigma \in S_n} \mathrm{sig}(\sigma^{-1})\, a_{1\,\sigma^{-1}(1)} \dots a_{n\,\sigma^{-1}(n)} \\ &= \textstyle\sum_{\tau \in S_n} \mathrm{sig}(\tau)\, a_{1\,\tau(1)} \dots a_{n\,\tau(n)} = \sum_{\tau \in S_n} \mathrm{sig}(\tau)\, b_{\tau(1)\,1} \dots b_{\tau(n)\,n} \\ &= \det A^t \end{aligned}$$

donde en la segunda línea primero se reordenan (usando la conmutatividad de K) los factores de cada sumando por orden creciente de superíndice, y a continuación se usa $\mathrm{sig}(\sigma) = \mathrm{sig}(\sigma^{-1})$, en la tercera se toma $\tau = \sigma^{-1}$ (al ser biyectiva la aplicación $\sigma \mapsto \sigma^{-1}$ en S_n, resulta equivalente sumar en $\sigma \in S_n$ que en $\tau = \sigma^{-1} \in S$), y a continuación se cambia cada a_{ij} por b_{ji}. ∎

Observación 1.52. Como consecuencia de esta proposición, todas las propiedades de los determinantes que se obtengan "por columnas" serán también válidas "por filas" y viceversa. Para incluir ambos casos, llamaremos *línea* a cualquier fila o columna de A.

1.5.2. Propiedades elementales de los determinantes

Resumimos a continuación algunas propiedades de los determinantes, que se deducen de su definición y de las propiedades estudiadas de la signatura de las permutaciones.

1. De la definición de $|A|$ a partir de permutaciones, se tiene directamente (hágase como ejercicio):
$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21},$$
$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31} - a_{12}a_{21}a_{33} - a_{11}a_{23}a_{32},$$
(regla de Sarrus).

2. $|A^t| = |A|$ (proposición 1.51, que permite usar la observación 1.52).

3. a) Si una línea de A se multiplica por el escalar $a \in K$ entonces el determinante de la matriz resultante también queda multiplicado por a. En particular: $|aA| = a^n|A|$.

   Dem. Como en cada sumando de la fórmula (1.9) aparece un único elemento de cada fila y de cada columna, necesariamente en cada sumando aparecerá un único elemento de la línea multiplicada por a. Usando la conmutatividad de K podemos suponer que a es el primer factor de cada sumando, por lo que basta con "sacar factor común" a aplicando la propiedad distributiva en el cuerpo K.

   b) Si cada uno de los elementos a_k de una línea se expresa como la suma de dos escalares $a_k = b_k + c_k$ entonces $|A|$ es igual a la suma de dos determinantes, uno de ellos con cada escalar a_k reemplazando por el primer escalar de la suma, b_k, y el otro con el segundo, c_k.

   Dem. Al reemplazar cada a_k por $b_k + c_k$ en (1.9), se obtiene que en cada sumando se reemplaza por un único factor por $(b_k + c_k)$. El resultado se obtiene entonces sin más que aplicar la propiedad distributiva en K.

4. Si se intercambian dos líneas del mismo tipo (filas o columnas) de A, el determinante de la matriz resultante cambia de signo. Más aún, dada una permutación $\sigma_0 \in S_n$, si A^{σ_0} es la matriz que se obtiene cambiando cada fila i-ésima de A por la $\sigma_0(i)$-ésima, entonces:

   $|A^{\sigma_0}| = \mathrm{sig}(\sigma_0)\ |A|$.

   Dem. Razonemos el caso en que se intercambian las filas i-ésima y j-ésima $1 \leq i < j \leq n$. Llamemos τ a la permutación que se obtiene como trasposición de i y j, y sea A^τ a la correspondiente matriz obtenida permutando las filas i, j:

   $\det A^\tau := \sum_{\sigma\in S_n} \mathrm{sig}(\sigma)\ a_{\sigma(\tau(1))\,1} \dots a_{\sigma(\tau(n))\,n} = -\sum_{\sigma\in S_n} \mathrm{sig}(\sigma\circ\tau)\ a_{[(\sigma\circ\tau)(1)]\,1} \dots a_{[(\sigma\circ\tau)(n)]\,n}$
   $= -\sum_{\sigma'\in S_n} \mathrm{sig}(\sigma')\ a_{\sigma'(1)\,1} \dots a_{\sigma'(n)\,n} = -\det A.$

   donde en la primera igualdad se ha usado que, por ser τ una trasposición, $\mathrm{sig}(\sigma\circ\tau) = -\mathrm{sig}(\sigma)$ y en la segunda se ha llamado σ' a $\sigma\circ\tau$ y se ha usado que, fijada τ, la aplicación $\sigma \mapsto \sigma'$ es una biyección en S_n. La afirmación sobre A^{σ_0} es inmediata si σ_0 es una trasposición y el caso general se reduce a éste sin más que escribir σ_0 como composición de un número par o impar de trasposiciones.

5. Si dos líneas paralelas de $|A|$ son iguales entonces $|A| = 0$.

   Dem. Como consecuencia directa de 4 se tiene $\det A = -\det A$.

6. Si dos líneas paralelas de $|A|$ son proporcionales (en particular, si una es nula) entonces $|A| = 0$.

   Dem. Consecuencia directa de 3a y 5.

7. Si una línea de A se puede expresar como *combinación lineal*[26] del resto de líneas del mismo tipo (filas o columnas), entonces $|A| = 0$.

   Dem. Consecuencia directa de 3b y 6.

8. Si a una línea de A se le suma una combinación lineal del resto de líneas del mismo tipo (filas o columnas), entonces el valor del determinante no varía.

   Dem. Consecuencia directa de 3b y 7.

9. det (I_n)=1.

   Dem. Recordemos que, usando la delta de Kronecker $(I_n)_{ij} = \delta_{ij}$ por lo que:

   $\det(I_n) = \sum_{\sigma\in S_n} \text{sig}(\sigma)\, \delta_{\sigma(1)\,1} \dots \delta_{\sigma(n)\,n} = \delta_{11} \dots \delta_{nn} = 1.$

10. Si C es otra matriz cuadrada de orden n, entonces $|A \cdot C| = |A| \cdot |C|$. En particular $|A \cdot C| = |C \cdot A|$.

    Nota. La demostración de esta propiedad no es inmediata a partir de la definición de los determinantes (1.9). De hecho, posponemos su demostración hasta que estudiemos tensores determinantes[27] (teorema 5.17) aunque podemos usar la propiedad con fines prácticos en adelante, como discutiremos en la sección 1.6.1.

Observación 1.53. De las dos últimas propiedades se deduce que, si $A \in M_n(K)$ es regular, se tiene:

$$1 = \det(A \cdot A^{-1}) = \det A \cdot \det(A^{-1})$$

En particular, $\det A \neq 0$. Más aún, de la propiedad 7 se sigue entonces que si una línea de A es combinación del resto entonces A no puede ser regular.

Ejercicio 1.54. Escríbanse explícitamente las propiedades anteriores para una matriz cuadrada y el caso en que las líneas que se consideran son columnas.

Ejercicio 1.55. ¿En cuáles de las propiedades anteriores se ha usado que la característica de K es distinta de 2? ¿En cuáles que K es conmutativo?

1.5.3. Desarrollo del determinante por los adjuntos de una línea. Regla de Chio

Dada la matriz $A = (a_{ij})$ llamaremos:

- *Menor complementario* del elemento (i, j): determinante α_{ij} de la matriz que se obtiene suprimiendo la fila i-ésima y la columna j-ésima de la matriz A.

- *Adjunto* del elemento (i, j): escalar $\Delta_{ij} := (-1)^{i+j}\alpha_{ij}$.

[26] Anticipamos de nuevo la nomenclatura de los espacios vectoriales. Aquí queremos decir simplemente que la línea mencionada al principio se puede escribir como la suma de todas las otras líneas del mismo tipo tras haber multiplicado todos los elementos de cada una de estas líneas por un escalar (que puede ser distinto para líneas distintas).

[27] Una demostración distinta puede consultarse en el libro de L. Merino y E. Santos.

Proposición 1.56. *Para cada $i_0, j_0 \in \{i, \dots n\}$ se verifican*

$$|A| = \sum_{j=1}^{n} a_{i_0 j}\Delta_{i_0 j} = \sum_{i=1}^{n} a_{i j_0}\Delta_{i j_0}$$

(desarrollo de $|A|$ por los adjuntos de la fila i_0-ésima y columna j_0-ésima, respectivamente)[28].

Demostración. Consideremos primero el caso $i_0 = 1$. Si $S[j] := \{\sigma \in S_n : \sigma(1) = j\}$ se tiene:

$$\begin{aligned}
|A| = & \ \textstyle\sum_{\sigma \in S_n} (-1)^{[\sigma]} \ a_{1\sigma(1)} \dots a_{n\sigma(n)} \\
= & \ \textstyle\sum_{j=1}^{n} a_{1j} \left(\sum_{\sigma \in S[j]} (-1)^{[\sigma]} \ a_{2\sigma(2)} \dots a_{n\sigma(n)} \right) \\
= & \ \textstyle\sum_{j=1}^{n} (-1)^{j-1} a_{1j} \left(\sum_{\sigma \in S[j]} (-1)^{[\sigma]-(j-1)} \ a_{2\sigma(2)} \dots a_{n\sigma(n)} \right) \\
= & \ \textstyle\sum_{j=1}^{n} (-1)^{j-1} a_{1j} \alpha_{1j} = \sum_{j=1}^{n} a_{1j} (-1)^{j+1} \alpha_{1j} = \sum_{j=1}^{n} a_{1j} \Delta_{1j},
\end{aligned}$$

donde en la primera línea de la fórmula se usa la definición de $|A|$ ($= |A^t|$) y en las otras:

- Segunda línea: se agrupan los términos de la primera según el valor de $j = \sigma(1)$
- Tercera línea: se multiplica y divide por $(-1)^{j-1}$.
- Cuarta línea. Para la primera igualdad se usa la expresión del menor complementario:

  $\alpha_{1j} = \sum_{\sigma \in S[j]} (-1)^{[\sigma]-(j-1)} \ a_{2\sigma(2)} \dots a_{n\sigma(n)}$

  De hecho, en cada sumando los valores $\sigma(2), \dots, \sigma(n)$ pueden verse como una permutación del conjunto $A = \{1, \dots, j-1, j+1, \dots, n\}$ (véase la nota a pie de página 23) cuyo número de inversiones se calcula descontando a $[\sigma]$ las $j-1$ inversiones correspondientes a $\sigma(1) = j$.

  En la segunda igualdad se usa $(-1)^{j-1} = (-1)^{j+1}$ y en la última la definición de adjunto.

Para el caso $i_0 > 1$, haciendo $i_0 - 1$ trasposiciones de índices contiguos de filas para llevar la i_0-ésima hasta la primera posición, se llega a una nueva matriz A_0 cuyo determinante es $|A_0| = (-1)^{i_0-1}|A|$. Los menores complementarios de la primera fila de A_0 son iguales a los menores complementarios de la fila i_0-ésima de A. Sin embargo, los adjuntos de A_0 difieren en el factor $(-1)^{i_0-1}$. Por eso, desarrollando $|A_0|$ por los adjuntos adjuntos de su primera línea (usando el caso resuelto anteriormente) y multiplicando la expresión por $(-1)^{i_0-1}$, se obtiene precisamente el desarrollo del determinante de A por los adjuntos de su fila i_0-ésima.

Finalmente el caso del desarrollo por la columna j_0-ésima se reduce al del desarrollo por filas sin más que trasponer la matriz y usar que los menores complementarios del elemento (i, j) de la matriz A y (j, i) de la matriz A^t coinciden (por la propiedad 2 del principio de esta sección). ∎

Usando el desarrollo por adjuntos, el cálculo de un determinante de orden n se reduce al de n determinantes de orden $n-1$. En principio, esto permite reducir el cálculo de un determinante de orden $n > 3$ al caso $n = 3$, el cual resulta computable por la regla de Sarrus. Sin embargo, este procedimiento no sería nada eficiente: se necesitaría calcular $n!/3! = n!/6$ determinantes de orden 3. Empero, las otras propiedades de los determinantes permiten simplificar esos cálculos.

Explicamos a continuación la llamada *regla de Chio* para el cálculo de determinantes de orden superior. Esencialmente, esta regla consiste en conseguir que todos los elementos de una línea del

[28] Al desarrollo de un determinante por los adjuntos de una línea también se le llama *desarrollo de Laplace* y puede generalizarse para calcular el determimante de una matriz a partir de menores complementarios de submatrices cuadradas de cualquier orden; véase, por ejemplo, el libro de M. Castellet e I. Llerena.

determinante de A sean nulos excepto uno de ellos, al cual llamaremos *elemento base* o *pivote*, que se podrá tomar como 1. Así, el desarrollo por los adjuntos de esa línea asegura que $|A|$ coincide con el adjunto del elemento base. Explícitamente:

1. Se escoge una línea que contenga un 1 o un -1 y el mayor número posible ceros. En el caso de que ningún elemento sea 1 ó -1, se puede escoger uno de los elementos no nulos[29] a de la línea (escogida de entre las de más ceros), y dividir todos los elementos de esa línea por a; esto hace que que el determinante quede dividido por a. lo que se tendrá en cuenta multiplicando por ese escalar el resultado.[30]

2. Una vez escogido el elemento base, se consigue que el resto de los elementos de su línea, fila o columna, sean nulos multiplicando sucesivamente la columna o fila del elemento base por un escalar apropiado y restándolo (o sumándolo) a las otras líneas paralelas.

   Por ejemplo, si el elemento base es el $(1,1)$ y se escoge como línea su fila, se multiplicará sucesivamente su columna por $a_{12},\dots,a_{1n}$ y se restará a las siguientes columnas:

$$|A| = \begin{vmatrix} 1 & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} = \begin{vmatrix} 1 & 0 & \cdots & 0 \\ a_{21} & a_{22}-a_{12}a_{21} & \cdots & a_{2n}-a_{1n}a_{21} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2}-a_{12}a_{n1} & \cdots & a_{nn}-a_{1n}a_{n1} \end{vmatrix}$$

3. $|A|$ es entonces el adjunto del elemento base o su opuesto, según el elemento base fuera 1 ó -1:

$$|A| = 1\cdot\Delta_{11} + 0\cdot\Delta_{12} + \cdots + 0\cdot\Delta_{1n} = \Delta_{11}.$$

Ejemplo 1.57. Calculemos el siguiente determinante usando la regla de Chio:

$$\begin{vmatrix} 5 & 3 & 2 & 3 \\ 4 & 3 & 1 & 1 \\ 1 & 1 & 2 & 1 \\ 0 & 2 & 2 & 3 \end{vmatrix}$$

Los pasos son entonces: (1) tomamos como elemento base el $(3,1)$ (pues es un 1 y su columna tiene el mayor número de ceros), y desarrollaremos por los adjuntos de su columna, (2) para conseguir que en la columna del elemento base los otros elementos sean 0, multiplicamos la tercera fila por 5 y 4 y se la restamos a, respectivamente, la primera y la segunda, y (3) el determinante se reduce entonces al adjunto del elemento $(3,1)$ de la nueva matriz. Esto es:

$$\begin{vmatrix} 5 & 3 & 2 & 3 \\ 4 & 3 & 1 & 1 \\ 1 & 1 & 2 & 1 \\ 0 & 2 & 2 & 3 \end{vmatrix} = \begin{vmatrix} 0 & -2 & -8 & -2 \\ 0 & -1 & -7 & -3 \\ 1 & 1 & 2 & 1 \\ 0 & 2 & 2 & 3 \end{vmatrix} = (-1)^{1+3} \begin{vmatrix} -2 & -8 & -2 \\ -1 & -7 & -3 \\ 2 & 2 & 3 \end{vmatrix} = \begin{vmatrix} -2 & -8 & -2 \\ -1 & -7 & -3 \\ 2 & 2 & 3 \end{vmatrix}$$

Este último determinante ya es resoluble por la regla de Sarrus. No obstante, puede ser más sencillo seguir aplicando la regla de Chio, tomando ahora como base el elemento $(2,1)$ y desarrollando por

[29] Por supuesto, si todos los elementos de la línea son 0, el determinante también es 0, y no se debe hacer nada.

[30] En ocasiones, también se puede obtener un 1 multiplicando una línea por un escalar y sumándosela a otra. En otras ocasiones se puede simplemente reajustar el siguiente paso con facilidad.

los adjuntos de su columna:

$$\begin{vmatrix} -2 & -8 & -2 \\ -1 & -7 & -3 \\ 2 & 2 & 3 \end{vmatrix} = \begin{vmatrix} 0 & 6 & 4 \\ -1 & -7 & -3 \\ 0 & -12 & -3 \end{vmatrix} = (-1)(-1)^{2+1} \begin{vmatrix} 6 & 4 \\ -12 & -3 \end{vmatrix} = 30.$$

Ejercicio 1.58. Se considera la matriz por cajas $\left(\begin{array}{c|c} A & 0_{n\times m} \\ \hline C & D \end{array}\right)$ donde $A \in M_n(K), D \in M_m(K), C \in M_{m\times n}(K)$. Demuéstrese:

$$\det \left(\begin{array}{c|c} A & 0 \\ \hline C & D \end{array}\right) = |A|.|D|$$

1.5.4. Inversa de una matriz regular

A continuación vamos a demostrar que toda matriz con determinante no nulo es regular, proporcionando explícitamente su matriz inversa [31].

Proposición 1.59. *Dada $A = (a_{ij}) \in M_n(K)$ una matriz cuadrada, y sea $\Delta \in M_n(K)$ la matriz cuyo elemento (i, j) es el adjunto del elemento (i, j) de A, para todo $i, j \in \{1, \ldots, n\}$. Se verifica:*

1. $\sum_{k=1}^{n} a_{ik}\Delta_{jk} = \sum_{k=1}^{n} a_{ki}\Delta_{kj} = \delta_{ij}|A|$*, para todo* $i, j \in \{1, \ldots, n\}$.
2. $A \cdot \Delta^t = \Delta^t \cdot A = |A| I_n$.
3. *Si $|A| \neq 0$ entonces A es regular y su inversa viene dada por:*

$$A^{-1} = \frac{\Delta^t}{|A|}.$$

Demostración. 1. En el caso $i = j$, el resultado se obtiene porque el primer miembro no es más que el desarrollo de $|A|$ por los adjuntos de su fila i-ésima, mientras que el segundo miembro lo es por los de su columna j-ésima. En el caso $i \neq j$ el primer miembro se anula (y análogamente el segundo) porque es igual al desarrollo por adjuntos de una matriz con dos filas repetidas; concretamente, la matriz que se obtiene reemplazando en A su fila j-ésima por su fila i-ésima (nótese que el adjunto del elemento (j,k) es el mismo para esas dos matrices).

2. Obsérvese que al calcular el elemento (i, j) del primer miembro (y análogamente del segundo):

$$(A \cdot \Delta^t)_{ij} = \sum_{k=1}^{n} a_{ik}(\Delta^t)_{kj} = \sum_{k=1}^{n} a_{ik}\Delta_{jk} = \delta_{ij}|A|,$$

la última igualdad por el apartado anterior.

3. Como $|A| \neq 0$, basta con multiplicar A por $\Delta^t/|A|$ por la derecha y la izquierda y aplicar la igualdad anterior. ∎

[31] De la observación 1.53 sabíamos (aunque aún no lo hayamos demostrado) que, si una matriz A es regular, entonces $|A| \neq 0$. Por tanto, sabemos ahora que A es regular si y sólo si $|A| \neq 0$ (véase la observación 1.6.1 sobre resultados no demostrados en este primer capítulo).

Ejemplo 1.60. Calculemos la inversa de la matriz

$$A=\begin{pmatrix} 1 & 2 & -1 \\ 3 & 4 & -3 \\ 2 & 2 & 2 \end{pmatrix}.$$

Calculamos en primer lugar su determinante (el cual debe resultar distinto de 0)

$$\begin{vmatrix} 1 & 2 & -1 \\ 3 & 4 & -3 \\ 2 & 2 & 2 \end{vmatrix} = \begin{vmatrix} 1 & 2 & -1 \\ 0 & -2 & 0 \\ 0 & -2 & 4 \end{vmatrix} = -8.$$

Computamos a continuación la matriz de sus adjuntos:

$$\Delta=\begin{pmatrix} 14 & -12 & -2 \\ -6 & 4 & 2 \\ -2 & 0 & -2 \end{pmatrix}.$$

Para calcular la inversa, basta con trasponer y dividir por el determinante:

$$A^{-1}=\frac{1}{|A|}\cdot\Delta^t=\frac{1}{(-8)}\cdot\begin{pmatrix} 14 & -6 & -2 \\ -12 & 4 & 0 \\ -2 & 2 & -2 \end{pmatrix}=\begin{pmatrix} -7/4 & 3/4 & 1/4 \\ 3/2 & -1/2 & 0 \\ 1/4 & -1/4 & 1/4 \end{pmatrix}.$$

1.6. Aplicaciones de los determinantes a los SEL

Debido a la importancia práctica de la resolución de los SEL a lo largo del curso, vamos a continuación a describir los conceptos y técnicas principales asociadas a ellos. Como detallamos en el próximo apartado, para ello anticiparemos algunos conceptos y resultados que se estudiarán con más profundidad y se demostrarán más adelante.

1.6.1. Conceptos y resultados anticipados

Ya habíamos anticipado el significado de que una línea (fila o columna) de una matriz sea combinación lineal del resto (sección 1.5.2, propiedad 7). Anticipamos ahora también que un conjunto de líneas se dice *linealmente independiente* cuando ninguna de ellas se puede expresar como combinacion lineal del resto[32]. Llamaremos entonces *rango por columnas* (resp. *por filas*) de una matriz $A \in M_{m\times n}(K)$ al número máximo de columnas (resp. filas) linealmente independientes de A.

Los resultados que demostraremos a lo largo del curso pero que anticipamos ahora para poder obtener resultados prácticos son:

- Tema 2. (a) Una matriz cuadrada $A \in M_n(K)$ es regular si y sólo si su rango (por columnas) es igual a n. Esto se demostrará al estudiar las bases de un espacio vectorial en el Tema 2 (obs. 2.158). Es de remarcar que este resultado es una primera caracterización importante de $\mathrm{GL}(n,K)$.

  (b) Si se tiene un conjunto linealmente independiente de r líneas de A y, al añadirle una más, las $r+1$ líneas son linealmente dependientes, entonces la última línea se puede expresar como combinación lineal de las r primeras (Prop. 2.94). Esta propiedad subyace al cálculo del rango de una matriz que describiremos más adelante.

[32]En particular, si alguna línea está repetida el conjunto se considera linealmente independiente, porque una de las líneas repetidas se puede expresar como combinación lineal de la otra con coeficiente 1 y el resto de columnas con coeficiente 0.

- Tema 3. Si $A \in M_{m\times n}(K)$, $C \in M_{n\times p}(K)$ entonces el rango (por columnas) de $A\cdot C$ es menor o igual al mínimo entre el rango de A y el de C. Esto será una consecuencia inmediata del concepto y propiedades del rango de una aplicación lineal (corolario 3.20). Como una consecuencia, si $m = p = n$ y $A\cdot C = I_n$ entonces los rangos de A y C son iguales a n, por lo que, necesariamente, A y C son regulares y $C = A^{-1}$ (corolario 3.58).

  Esto es, como anunciamos debajo de la definición 1.26, dada $A \in M_n(K)$ basta con demostrar que existe una matriz cuadrada X tal que o bien $A\cdot X = I_n$ o bien $X\cdot A = I_n$ para asegurar que A sea regular y X su inversa.

- Tema 4. Para cualquier $A \in M_{m\times n}(K)$ el rango de A por filas es igual a su rango por columnas (teorema 4.30). Esto se demostrará a partir del concepto de aplicación traspuesta y las propiedades de su rango.

  Como consecuencia de esta propiedad, habitualmente se habla solo del rango de la matriz, sin especificar si es por filas o columnas.

- Tema 5. Si $A,C \in M_n(K)$ entonces $|A\cdot C| = |A|.|C|$ (teorema 5.17). Esto se demostrará tras estudiar el concepto de determinante de un endomorfismo (véase la observación **??**) y ya se enunció como propiedad 10 en la subsección 1.5.2.

  Como consecuencia, toda matriz regular debe tener determinante $\neq 0$. Esto, unido al cálculo explícito de la inversa de una matriz A con $|A| \neq 0$, proporciona la importante caracterización siguiente: una matriz cuadrada A es regular si y sólo si $|A| \neq 0$.

A lo largo del resto de esta sección, se dará una primera justificación de los resultados sobre rangos y SEL partiendo de la base de que estas propiedades anticipadas son ciertas. El objetivo es sólo que el lector pueda ir ejercitándose con la práctica de determinantes, ya que en los capítulos siguientes *no se presupone ninguna de estas propiedades* hasta que se demuestre formalmente. Se invita al lector a que, una vez terminado el curso, vuelva a leer la presente sección y compruebe que, efectivamente, sabe demostrar todo lo que aqui se dice.

1.6.2. Cálculo del rango de una matriz $m \times n$

Dada una matriz cualquiera $A \in M_{m\times n}(K)$, se llama *submatriz* de A a cualquier matriz que se obtenga suprimiendo k filas y l columnas de A, donde $k \in \{0,1,\dots m\}$, $l \in \{0,1,\dots,n\}$. Llamaremos *menor de orden h* de A, donde $h \in \{1,\dots,\mathrm{Min}\{m,n\}\}$), al determinante de cualquier submatriz cuadrada $h\times h$ de A (obtenida suprimiendo $m-h$ filas y $n-h$ columnas).

Proposición 1.61. *El rango r de la matriz $A \in M_{m\times n}(K)$ es igual al mayor orden h de los menores de A distintos de 0.*

Demostración. Para $r \geq h$, basta con darse cuenta de que las h columnas (o filas) de un menor distinto de 0 tienen que ser linealmente independientes y, por tanto, las correspondientes h lineas de la matriz A también son independientes.

Para la desigualdad $r \leq h$, obsérvese que A contiene r columnas independientes, las cuales forman una submatriz $S \in M_{m\times r}$. Ahora bien el rango de S (por columnas) coincide con su rango por filas (esto es, con el rango de S^t) y, por tanto, r filas de S son independientes. Estas r filas forman una submatriz cuadrada C de S y, por tanto, también de A y verifican $|C| \neq 0$ al ser independientes sus líneas. ∎

Definición 1.62. *A los menores de A distintos de 0 de mayor orden (esto es, de orden igual al rango r de A) se les llamará* menores principales.

El rango de A se reduce entonces a encontrar un menor principal. Aunque a menudo esto se pueda hacer con simple inspección, debe desarrollarse un método sistemático. Obsérvese que si, por ejemplo, se ha obtenido un menor de orden $h < \text{Min}\{m,n\}$ distinto de 0, no resultaría nada eficiente tener que calcular todos los menores de orden $h+1$ para poder demostrar que el rango de A es h. Explicamos a continuación un método apropiado, ilustrándolo con el cálculo del rango de la matriz:

$$A_0 = \begin{pmatrix} -5 & -3 & 1 & -1 & 0 \\ -1 & 0 & 2 & 1 & 0 \\ 1 & 1 & 1 & 1 & 0 \\ 2 & 0 & -4 & -2 & 0 \\ -3 & -3 & -3 & 2 & 0 \end{pmatrix}.$$

1. Se suprimen todas las líneas que sean nulas, o proporcionales a otra (o que se aprecie fácilmente que se pueden escribir como combinación lineal del resto).

   En nuestro ejemplo se puede suprimir la última columna (por ser nula) y la cuarta fila (por ser proporcional a la segunda). El problema se reduce entonces a calcular el rango de[33]:

   $$A_1 = \begin{pmatrix} -5 & -3 & 1 & -1 \\ -1 & 0 & 2 & 1 \\ 1 & 1 & 1 & 1 \\ -3 & -3 & -3 & 2 \end{pmatrix}.$$

2. Escogemos un elemento de la matriz dada A que sea distinto de 0 (con lo que podemos asegurar que el rango es al menos 1). A continuación, *orlamos* ese menor, esto es, formamos un menor de orden superior en una unidad, añadiéndole una nueva fila y columna de la matriz dada A. Si uno de esos menores es no nulo, podemos asegurar que el rango es al menos 2, y repetir sucesivamente el proceso.

   En el ejemplo, primero escogemos el elemento el -5 (elemento $(1,1)$) que es distinto de 0, y a continuación lo orlamos con los elementos de las segundas fila y columna para obtener:
   $$\begin{vmatrix} -5 & -3 \\ -1 & 0 \end{vmatrix} = -3,$$
   también distinto de 0, por lo que el rango es al menos 2. Obsérvese que se anula el siguiente menor que se obtiene orlando de la manera más sencilla:
   $$\begin{vmatrix} -5 & -3 & 1 \\ -1 & 0 & 2 \\ 1 & 1 & 1 \end{vmatrix} = 0.$$

3. En el caso de que tengamos un menor C no nulo de orden h y, al orlarlo, se obtenga uno nulo, podemos escoger como sigue los menores de orden $h+1$, con el objetivo de detectar si una línea de la matriz original depende de las líneas seleccionadas por C. Fijemos la fila f con la que orlamos C (fijando la columna sería análogo) y consideremos la submatriz S que se obtiene con todas las filas seleccionadas por C y la fila f. En S, orlemos C con f y cada una de las columnas de S. Si uno de esos menores es 0, podemos asegurar que la columna de S con que

[33] No obstante, el lector puede comprobar que el procedimiento sigue siendo válido sin suprimir la cuarta fila, aunque resulte más lento.

se orla C es combinación lineal de las columnas de S seleccionadas por C. En consecuencia, si todos esos menores son 0, el rango (por columnas) de S será h. Como este rango es igual al rango por filas de S, la fila f será combinación de las filas de A que contienen a las de C, por lo que la podemos suprimir.

En nuestro ejemplo nos fijamos ahora en la submatriz

$$S = \begin{pmatrix} -5 & -3 & 1 & -1 \\ -1 & 0 & 2 & 1 \\ 1 & 1 & 1 & 1 \end{pmatrix},$$

y observamos:

$$\begin{vmatrix} -5 & -3 & 1 \\ -1 & 0 & 2 \\ 1 & 1 & 1 \end{vmatrix} = \begin{vmatrix} -5 & -3 & -1 \\ -1 & 0 & 1 \\ 1 & 1 & 1 \end{vmatrix} = 0.$$

En consecuencia, el rango de S es dos (cada una de sus dos últimas columnas se escribe como combinación lineal de las dos primeras). Por tanto, su rango por filas también es dos, y la tercera fila de S se tiene que poder expresar como combinación lineal de las dos primeras. Suprimimos esta fila de la matriz y el problema se reduce al cálculo del rango de:

$$A_2 = \begin{pmatrix} -5 & -3 & 1 & -1 \\ -1 & 0 & 2 & 1 \\ -3 & -3 & -3 & 2 \end{pmatrix}.$$

Orlando de nuevo, se obtiene:

$$\begin{vmatrix} -5 & -3 & 1 \\ -1 & 0 & 2 \\ -3 & -3 & -3 \end{vmatrix} = 0$$

(esto es, la tercera columna de A_2 es combinación lineal de las dos primeras) y

$$\begin{vmatrix} -5 & -3 & -1 \\ -1 & 0 & 1 \\ -3 & -3 & 2 \end{vmatrix} \neq 0.$$

Esta última desigualdad nos proporciona un menor principal de A_2 y, por tanto, de A_0. En consecuencia, el rango de A_0 es 3.

Observación 1.63. Consideremos las transformaciones elementales (por filas) introducidas para el estudio de los SEL en la sección 1.3.4. Es fácil darse cuenta que estas transformaciones no alteran el rango (por filas) de la matriz a la que se aplican[34] (y, por tanto, tampoco alterarán el rango si se realizan por columnas). Además, es un ejercicio sencillo comprobar que el rango de una matriz escalonada coincide con el número de filas no nulas que contiene. Así, el procedimiento de Gauss permite también calcular el rango de una matriz (y puede, eventualmente, combinarse con el descrito en términos de menores principales).

1.6.3. SEL: regla de Cramer y Teorema de Rouché-Frobenius

Apliquemos a continuación los resultados previos a los SEL, empezando con los de Cramer.

[34] Esto se puede razonar tanto directamente (a partir la definición de rango) como por aplicación de la proposición 1.61, ya que las transformaciones elementales no cambian el signo (positivo, negativo o nulo) de los correspondientes menores.

Proposición 1.64. (Regla de Cramer.) *Un SEL de Cramer, esto es, del tipo*

$$A\cdot x=b \qquad \textit{donde} \quad x=\begin{pmatrix} x_1\\ \vdots \\ x_n\end{pmatrix},\ b=\begin{pmatrix} b_1\\ \vdots \\ b_n\end{pmatrix},\ A=\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n}\\ a_{21} & a_{22} & \cdots & a_{2n}\\ \cdots & \cdots & \cdots & \cdots\\ a_{n1} & a_{n2} & \cdots & a_{nn}\end{pmatrix}\in M_n(K),$$

$|A|\neq 0$ *es compatible y determinado, y su única solución* $x=A^{-1}b$ *se expresa:*

$$x_1=\frac{\begin{vmatrix} b_1 & a_{12} & \cdots & a_{1n}\\ b_2 & a_{22} & \cdots & a_{2n}\\ \cdots & \cdots & \cdots & \cdots\\ b_n & a_{n2} & \cdots & a_{nn}\end{vmatrix}}{|A|},\quad x_2=\frac{\begin{vmatrix} a_{11} & b_1 & \cdots & a_{1n}\\ a_{21} & b_2 & \cdots & a_{2n}\\ \cdots & \cdots & \cdots & \cdots\\ a_{n1} & b_n & \cdots & a_{nn}\end{vmatrix}}{|A|},\quad \cdots \quad x_n=\frac{\begin{vmatrix} a_{11} & a_{12} & \cdots & b_1\\ a_{21} & a_{22} & \cdots & b_2\\ \cdots & \cdots & \cdots & \cdots\\ a_{n1} & a_{n2} & \cdots & b_n\end{vmatrix}}{|A|}$$

(obsérvese que en el numerador de la incógnita x_j*, se reemplaza la columna* j*-ésima de la matriz* A *por la columna de términos independientes).*

Demostración. Por ser A regular, resulta inmediato que es compatible y determinado con única solución $x=A^{-1}b$. Usando la forma explícita de la matriz inversa:

$$x=A^{-1}\cdot b=\frac{1}{|A|}\Delta^t\cdot b=\frac{1}{|A|}\begin{pmatrix}\sum_{i=1}^n \Delta_{i1}b_i\\ \vdots\\ \sum_{i=1}^n \Delta_{in}b_i\end{pmatrix},$$

y la expresión de la solución se obtiene sin más que tener en cuenta que cada componente $\sum_{i=1}^n \Delta_{ij}b_i=\sum_{i=1}^n b_i\Delta_{ij}$ no es más que el desarrollo por los adjuntos de la columna j-ésima en el numerador de la expresión de cada x_j. ∎

Ejemplo 1.65. Discutiremos y resolveremos el SEL

$$\left.\begin{array}{rl} 2x_1-x_2+x_3 & =3\\ 2x_2-x_3 & =1\\ -x_1+x_2 \quad & =1\end{array}\right\}$$

El determinante de la matriz de los coeficientes es:

$$\begin{vmatrix} 2 & -1 & 1\\ 0 & 2 & -1\\ -1 & 1 & 0\end{vmatrix}=\begin{vmatrix} 2 & 1 & 0\\ 0 & 2 & -1\\ -1 & 1 & 0\end{vmatrix}=\begin{vmatrix} 2 & 1\\ -1 & 1\end{vmatrix}=3.$$

En consecuencia, éste es un sistema compatible y determinado, cuya solución se puede obtener por la regla de Cramer:

$$x_1=\frac{1}{3}\begin{vmatrix} 3 & -1 & 1\\ 1 & 2 & -1\\ 1 & 1 & 0\end{vmatrix}=\frac{1}{3}\begin{vmatrix} 4 & 1 & 0\\ 1 & 2 & -1\\ 1 & 1 & 0\end{vmatrix}=1$$

$$x_2=\frac{1}{3}\begin{vmatrix} 2 & 3 & 1\\ 0 & 1 & -1\\ -1 & 1 & 0\end{vmatrix}=\frac{1}{3}\begin{vmatrix} 2 & 4 & 0\\ 0 & 1 & -1\\ -1 & 1 & 0\end{vmatrix}=2$$

$$x_3 = \frac{1}{3}\begin{vmatrix} 2 & -1 & 3 \\ 0 & 2 & 1 \\ -1 & 1 & 1 \end{vmatrix} = \frac{1}{3}\begin{vmatrix} 0 & 1 & 5 \\ 0 & 2 & 1 \\ -1 & 1 & 1 \end{vmatrix} = 3.$$

Consideremos ahora un SEL arbitrario,

$$A \cdot x = b, \qquad A \in M_{m\times n}. \tag{1.10}$$

Reescribiéndolo

$$\begin{pmatrix} a_{11} \\ a_{21} \\ \cdots \\ a_{m1} \end{pmatrix} x_1 + \begin{pmatrix} a_{12} \\ a_{22} \\ \cdots \\ a_{m2} \end{pmatrix} x_2 + \begin{pmatrix} a_{1n} \\ a_{2n} \\ \cdots \\ a_{mn} \end{pmatrix} x_n = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}$$

resulta evidente que el sistema es compatible si y sólo si la columna de los términos independientes se puede escribir como combinación lineal de las columnas de A, proporcionando entonces los coeficientes una solución del sistema. Se tiene así (*Teorema de Rouché-Frobenius*[35]):

- *El SEL* (1.10) *es compatible si y sólo si el rango r de la matriz A del sistema coincide con el de la matriz ampliada* $(A|b)$. En este caso:

  1. Podemos tomar un menor principal de A (que será también principal para $(A|b)$) y suprimir las filas de la matriz ampliada que no coincidan con las filas de ese menor (pues se corresponderán con ecuaciones que se expresan como combinación lineal de las no suprimidas).
  2. Las columnas del menor principal determinan r incógnitas que tomaremos como principales del sistema. Las otras $n-r$ se consideran secundarias, se las puede tomar como parámetros $\lambda_1, \dots \lambda_{n-r}$ y aplicarles la regla de Cramer.

     En particular, el SEL será determinado si y sólo si $n = r$.

Esto es, suponiendo que esas incógnitas principales son las r primeras (o renombrando las incógnitas para que así ocurra) el SEL original resulta equivalente a:

$$\begin{gathered} a_{11}x_1 + \cdots + a_{1r}x_r = b_1 - a_{1(r+1)}\lambda_1 \cdots - a_{1(r+(n-r)=n))}\lambda_{n-r} \\ \cdots \\ a_{r1}x_1 + \cdots + a_{rr}x_r = b_r - a_{r(r+1)}\lambda_1 \cdots - a_{r(r+(n-r)=n)}\lambda_{n-r} \end{gathered}$$

Para cada valor de los parámetros, este sistema es de Cramer, por lo que puede usarse directamente la regla de Cramer para solucionarlo:

$$x_1 = \frac{\begin{vmatrix} b_1 - a_{1(r+1)}\lambda_1 \cdots - a_{1n}\lambda_{n-r} & a_{12} & \cdots & a_{1r} \\ \cdots & \cdots & \cdots & \cdots \\ b_r - a_{r(r+1)}\lambda_1 \cdots - a_{rn}\lambda_{n-r} & a_{r2} & \cdots & a_{rr} \end{vmatrix}}{\begin{vmatrix} a_{11} & \cdots & a_{1r} \\ \cdots & \cdots & \cdots \\ a_{r1} & \cdots & a_{rr} \end{vmatrix}}, \dots$$

[35] Habiendo sido demostrado por múltiples matemáticos cercanos en el tiempo, recibe otros nombres como Rouché-Capelli en países angloparlantes e Italia, o Rouché-Fontené en Francia.

$$\ldots, x_r = \frac{\begin{vmatrix} a_{11} & \cdots & a_{1(r-1)} & b_1 - a_{1(r+1)}\lambda_1 \cdots - a_{1n}\lambda_{n-r} \\ \cdots & \cdots & \cdots & \cdots \\ a_{r1} & \cdots & a_{r(r-1)} & b_r - a_{r(r+1)}\lambda_1 \cdots - a_{rn}\lambda_{n-r} \end{vmatrix}}{\begin{vmatrix} a_{11} & \cdots & a_{1r} \\ \cdots & \cdots & \cdots \\ a_{r1} & \cdots & a_{rr} \end{vmatrix}}$$

Ejemplo 1.66. Discutiremos y resolveremos el SEL

$$\left.\begin{array}{rl} x_1 + x_2 + x_3 + x_4 & = 1 \\ -x_1 + x_2 - x_3 + x_4 & = 0 \\ -x_1 + 5x_2 - x_3 + 5x_4 & = 2 \end{array}\right\},$$

cuya matriz ampliada es:

$$\left(\begin{array}{rrrr|r} 1 & 1 & 1 & 1 & 1 \\ -1 & 1 & -1 & 1 & 0 \\ -1 & 5 & -1 & 5 & 2 \end{array}\right).$$

Para calcular el rango de la matriz del sistema, se considera el menor $\begin{vmatrix} 1 & 1 \\ -1 & 1 \end{vmatrix} = 2(\neq 0)$. El rango es, por tanto, al menos 2. Orlando este menor, se tiene:

$$\begin{vmatrix} 1 & 1 & 1 \\ -1 & 1 & -1 \\ -1 & 5 & -1 \end{vmatrix} = 0, \qquad \begin{vmatrix} 1 & 1 & 1 \\ -1 & 1 & 1 \\ -1 & 5 & 5 \end{vmatrix} = 0.$$

Por tanto, el menor escogido es un menor principal de la matriz del sistema, cuyo rango es 2. Para la matriz ampliada, orlamos el menor principal usando la columna de términos independientes. Como

$$\begin{vmatrix} 1 & 1 & 1 \\ -1 & 1 & 0 \\ -1 & 5 & 0 \end{vmatrix} = 0,$$

el rango de la matriz ampliada también es dos, y el sistema resulta ser compatible e indeterminado. Para resolverlo, teniendo en cuenta el menor principal escogido, se suprime la tercera fila y se escogen x_1, x_2 como incógnitas principales. Así, las otras dos se toman como parámetros,

$$x_3 = \lambda_1 \qquad x_4 = \lambda_2,$$

y el sistema se reescribe:

$$\left.\begin{array}{rl} x_1 + x_2 & = 1 - \lambda_1 - \lambda_2 \\ -x_1 + x_2 & = \lambda_1 - \lambda_2 \end{array}\right\}.$$

Su solución usando la regla de Cramer es entonces:

$$x_1 = \frac{1}{2} \cdot \begin{vmatrix} 1 - \lambda_1 - \lambda_2 & 1 \\ \lambda_1 - \lambda_2 & 1 \end{vmatrix} = \frac{1}{2} - \lambda_1,$$

$$x_2 = \frac{1}{2} \cdot \begin{vmatrix} 1 & 1 - \lambda_1 - \lambda_2 \\ -1 & \lambda_1 - \lambda_2 \end{vmatrix} = \frac{1}{2} - \lambda_2.$$

Tema 2

Espacios vectoriales

Los espacios vectoriales son estructuras básicas en matemáticas que aparecen no sólo en Geometría, sino también en Álgebra o en Cálculo, que abstraen propiedades conocidas a nivel elemental de $\mathbb{R}^n$. A pesar de su sencillez, desempeñan un papel clave en el estudio de figuras complicadas como curvas o superficies; de hecho, el método que se sigue para estudiar estas figuras consiste en asociarles en cada punto el espacio vectorial que “mejor las aproxime” en la cercanía del punto (la recta o el plano tangente en ese punto) y en analizar cómo este espacio vectorial cambia al movernos sobre la figura.

Los espacios vectoriales también son muy importantes en otras ciencias, como la Física, donde es habitual el empleo de espacios vectoriales para modelar magnitudes como las velocidades o las fuerzas; de hecho, la definición formal de espacio vectorial que estudiaremos resulta esencial para distinguir entre las propiedades intrínsecas de esas magnitudes (asociadas a su naturaleza física) y las extrínsecas (dependientes de las coordenadas con las que un observador las mida).

2.1. Definición, primeras propiedades y ejemplos

2.1.1. Concepto de espacio vectorial

Cuando hablamos de vectores pensamos en segmentos de recta orientados (flechas), con propiedades tales como la “dirección”, el “sentido” o el “módulo”. Sabemos desde la enseñanza secundaria que en el plano o en el espacio dos vectores se pueden sumar obteniéndose otro vector. También se puede multiplicar un número real por un vector y el resultado es un nuevo vector. Además, estas operaciones presentan algunas propiedades con similitudes a las que se vieron en la definición de anillo. Pues bien, los espacios vectoriales suponen la abstracción matemática de los conjuntos que admiten una suma interna y un producto por elementos de un cuerpo de modo que se cumplen esas mismas propiedades. La definición de espacio vectorial *se centra sólo en esa suma y producto*, sin tener en cuenta otros conceptos con el que el alumno pueda estar familiarizado (como el de módulo, producto escalar o vectorial), y que se irán desarrollando conforme se profundice más en Geometría.

Definición 2.1. *Sea* $(K,+,\cdot)$ *un cuerpo conmutativo. Un* espacio vectorial *(e.v.) sobre* $(K,+,\cdot)$ *o* K-espacio vectorial *es una terna* $(V,+,\cdot K)$ *formada por un conjunto* V *(a posteriori necesariamente no vacío) dotado de una operación (ley de composición interna) en* V*,*

$$+ : V \times V \to V$$

y una aplicación

$$\cdot : K \times V \to V$$

o ley de composición externa *de K sobre V tales que:*

I) *$(V,+)$ es un grupo conmutativo. Esto es, la operación $+$ en V, que asocia a cada par $(u,v) \in V \times V$ un único $u+v \in V$ verifica las propiedades ya estudiadas de grupo abeliano:*

- *(i) Asociativa: $(u+v)+w = u+(v+w)$, para cualesquiera $u,v,w \in V$.*
- *(ii) Elemento neutro: existe un elemento $e \in V$ tal que $v+e = e+v = v$, para todo $v \in V$. Al neutro lo denotaremos 0 e incluso $\vec{0}$ cuando se le quiera distinguir explícitamente del neutro 0 de $(K,+)$.*
- *(iii) Elemento simétrico: para cada $v \in V$, existe $w \in V$ tal que $v+w = w+v = 0$. Al elemento simétrico de v se le denotará $-v$ y se le llamará* opuesto de *v.*
- *(iv) Conmutativa: $u+v = v+u$, para todo $u,v \in V$.*

II) *La ley de composición externa $\cdot : K \times V \to V$, que asocia a cada $a \in K$ y cada $v \in V$ un único vector que denotaremos $a \cdot v \in V$, verifica las propiedades:*

- *(i) Pseudoasociativa: $(a \cdot b) \cdot v = a \cdot (b \cdot v)$, para todo $a,b \in K$ y todo $v \in V$,*
- *(ii) Unimodular: $1 \cdot v = v$, para todo $v \in V$, donde 1 es el elemento neutro de $(K \setminus \{0\}, \cdot)$*
- *(iii) Distributiva respecto de la suma en K: $(a+b) \cdot v = a \cdot v + b \cdot v$, para cualesquiera $a,b \in K$ y $v \in V$,*
- *(iv) Distributiva respecto de la suma en V: $a \cdot (u+v) = a \cdot u + a \cdot v$, para todo $a \in K$ y cualesquiera $u,v \in V$.*

Observación 2.2. (1) En la definición anterior se usan con frecuencia los símbolos $+$ y $\cdot$ para representar dos conceptos distintos: para $+$, las operaciones en $(V,+)$ y $(K,+)$; para $\cdot$ la segunda operación del cuerpo $(K,+,\cdot)$ y la ley de composición externa en el espacio vectorial. Como ejercicio, explíquese qué significa cada símbolo cada vez que aparece en la definición anterior. ¿Hay alguna posibilidad de confusión en la notación?

(2) El lector puede comprobar directamente que $\mathbb{R}^2$ es un espacio vectorial con la suma en $\mathbb{R}^2$ usual

$$(x,y)+(x',y') := (x+x', y+y') \qquad \forall (x,y),(x',y') \in \mathbb{R}^2$$

y la ley de composición externa de $\mathbb{R}$ sobre $\mathbb{R}^2$:

$$a \cdot (x,y) := (a \cdot x, a \cdot y) \qquad \forall x \in \mathbb{R}, \quad \forall (x,y) \in \mathbb{R}^2.$$

Obsérvese que, en cada caso, en el miembro izquierdo los símbolos $+$, $\cdot$ denotan las operaciones que se están definiendo en el espacio vectorial, mientras que en el derecho esos mismos símbolos denotan la suma y productos usuales en $\mathbb{R}$.

(3) En adelante, al igual que hablaremos simplemente del "cuerpo K" cuando se sobreentiendan las operaciones (sin especificarlas explícitamente), hablaremos del espacio vectorial $V(K)$ sin especificar sus leyes de composición externa e interna, e incluso V si se sobreentiende cuál es el cuerpo (ya sea uno concreto o uno genérico). Así, el ejemplo del punto 2) anterior se denotará $\mathbb{R}^2(\mathbb{R})$ o sólo $\mathbb{R}^2$.

(4) Nótese que la operación $\cdot$ en un anillo parece tener propiedades muy similares a la ley de composición externa en un e.v. Sin embargo, en un anillo el producto $\cdot$ es una operación en él (es una ley de composición interna porque multiplican dos elementos del anillo, obteniéndose entonces otro elemento del anillo) mientras que en un e.v. el producto $\cdot$ no es interno (se multiplica un elemento de K por otro de V resultando un elemento de V). Esto justifica que la primera propiedad del producto $\cdot$ de un e.v. no sea una propiedad asociativa propiamente dicha, y se le llame "pseudoasociativa".

(5) No se ha usado en la definición que el cuerpo K sea conmutativo y, de hecho, podría mantenerse esta definición cuando no lo es. No obstante, en ese caso, tendríamos dos opciones distintas para la propiedad pseudoasociativa: $(a.b).v = a.(b.v)$ y $(a.b).v = b.(a.v)$. Para darse cuenta de que esta segunda opción sería tan razonable como la primera, bastaría con denotar la ley de composición externa escribiendo primero el elemento de V y luego el de K. Concretamente, poniendo: $v \cdot a$, la segunda opción se escribiría $v \cdot (a \cdot b) = (v \cdot a) \cdot b$. Según la opción que se escogiera para la propiedad pseudoasociativa se hablaría de un espacio vectorial *por la izquierda* (nuestra opción por defecto) o *por la derecha* (la anterior opción alternativa). No obstante, nosotros no tendremos en cuenta esta sutileza (ni otras que aparecerían más adelante), al suponer siempre que los cuerpos sean conmutativos.

Definición 2.3. *Si V es un e.v. sobre K, llamaremos* escalares *a los elementos de K y* vectores *a los elementos de V; usualmente, usaremos las letras $a, b, c, \ldots$ para referirnos a los escalares y las letras $u, v, w, \ldots$ para referirnos a los vectores.*

La operación $+$ de V se llamará suma de vectores *y la ley de composición externa $\cdot$ de K sobre V se llamará* producto de escalares por vectores. *Cuando $K = \mathbb{R}$ (resp. $K = \mathbb{C}$) diremos que V es un* e.v. real *(resp.* e.v. complejo*).*

2.1.2. Propiedades elementales

Es de remarcar que, como en todo espacio vectorial $V(K)$, se cumple que $(V, +)$ es un grupo abeliano, todas las propiedades que conocemos de grupos se mantienen para la suma de vectores. En particular (compruébese como un ejercicio de repaso)

Proposición 2.4. *En todo espacio vectorial $V(K)$:*

(a) El elemento neutro 0 de la suma de vectores es único.

(b) La suma de vectores permite simplificar:

$$v + w = v + w' \Rightarrow w = w' \qquad w + v = w' + v \Rightarrow w = w', \qquad \forall v, w, w' \in V.$$

(c) Para cada vector v, su elemento opuesto $-v$ es único, y de la igualdad $v + w = 0$ se deduce $w = -v, v = -w$ (en particular $-(-v) = v$).

Como notación, dados $u, v \in V$ escribiremos $u - v$ para referirnos al vector $u + (-v)$. Esto se llamará *diferencia de vectores* y consiste en sumar al primero el opuesto del segundo. Obsérvese que, cuando consideramos además el producto por escalares, expresiones como

$$-(a \cdot v), \qquad (-a) \cdot v, \qquad a \cdot (-v),$$

son conceptualmente distintas (¿qué significa el signo $-$ en cada caso?) Sin embargo, la siguiente proposición nos tranquilizará al demostrar, en particular, que las tres coinciden, por lo que se puede escribir de manera no ambigua $-a \cdot v$. De hecho, será fácil darse cuenta de que otros abusos de notación usuales como *no* escribir explícitamente el símbolo del producto escalar (esto es, poner av en vez de $a \cdot v$) o no distinguir notacionalmente entre el neutro 0 de $(K, +)$ y $\vec{0}$ de $(V, +)$, no deben dar lugar a ambigüedad o confusión.

Proposición 2.5. *Sea $V(K)$ un e.v. Se verifican las siguientes propiedades:*

1) $a \cdot 0 = 0$ *(esto es,* $a \cdot \vec{0} = \vec{0}$*), para todo* $a \in K$.

2) $0 \cdot v = 0$ *(esto es,* $0 \cdot v = \vec{0}$*), para todo* $v \in V$.

3) Dados $a \in K$ *y* $v \in V$*, si* $a \cdot v = 0$*, entonces* $a = 0$ *ó* $v = 0$.

4) $-(a \cdot v) = (-a) \cdot v = a \cdot (-v)$*, para todo* $a \in K$ *y todo* $v \in V$.

5) $-v = (-1) \cdot v$*, para todo* $v \in V$.

6) $-(u+v) = -u - v$ *(esto es,* $-(u+v) = (-u) + (-v)$*), para todo* $u, v \in V$.

7) $(-a) \cdot (-v) = a \cdot v$*, para todo* $a \in K$ *y todo* $v \in V$.

8) Si $a \cdot u = a \cdot v$ *con* $a \in K$*,* $a \neq 0$ *y* $u, v \in V$*, entonces* $u = v$.

9) Si $a \cdot v = b \cdot v$ *con* $a, b \in K$ *y* $v \in V$ *con* $v \neq 0$*, entonces* $a = b$.

Demostración. 1). Usando la propiedad de neutro para $\vec{0}$ en $(V, +)$ y la distributiva

$$a \cdot \vec{0} = a \cdot (\vec{0} + \vec{0}) = a \cdot \vec{0} + a \cdot \vec{0},$$

por lo que el resultado se obtiene simplificando en $(V, +)$ (para más claridad en la simplificación, obsérvese que en el miembro izquierdo se puede sustituir $a \cdot \vec{0}$ por $a \cdot \vec{0} + \vec{0}$).

2) Usando ahora que 0 es el neutro en $(K, +)$:

$$0 \cdot v = (0 + 0) \cdot v = 0 \cdot v + 0 \cdot v$$

por lo que basta de nuevo con simplificar.

3) Sean $a \in K$ y $v \in V$ con $a \cdot v = 0$. Del enunciado se sigue que basta con demostrar: si $a \neq 0$ entonces $v = 0$. Como K es un cuerpo, podemos tomar el inverso a^{-1}. Multiplicando por a^{-1} a ambos lados de $a \cdot v = 0$, y usando la propiedad 1), se tiene:

$$0 = a^{-1} \cdot (a \cdot v) = (a^{-1} \cdot a) \cdot v = 1 \cdot v = v,$$

donde se ha usado la propiedad pseudoasociativa en la segunda igualdad, y la unimodular en la última.

4) Tenemos que demostrar dos igualdades. Usando la proposición 2.4(c), para la primera igualdad basta con comprobar que al sumar $a \cdot v$ con $(-a) \cdot v$ se obtiene $\vec{0}$, y para la segunda que también ocurre al sumarlo con $a \cdot (-v)$. Claramente:

$$a \cdot v + (-a) \cdot v = (a + (-a)) \cdot v = 0 \cdot v = \vec{0},$$
$$a \cdot v + a \cdot (-v) = a \cdot (v - v) = a \cdot \vec{0} = \vec{0},$$

donde en la primera igualdad de cada caso se usa una de las propiedades distributivas, y en la última igualdad en se usan 2) y 1), resp.

5) Tómese $a = 1$ en la primera igualdad 4) y úsese la propiedad unimodular.

6) Por 4) y una de las distributivas, se tiene:

$$-(u+v) = (-1) \cdot (u+v) = (-1) \cdot u + (-1) \cdot v = -u + (-v) = -u - v$$

(para la última igualdad, obsérvese que en cualquier anillo unitario se tiene $(-1)\cdot a = -a$ por lo que $(-1)\cdot(-a) = -(-a) = a$).

7) Por la propiedad 5) y la pseudoasociativa, se tiene:

$$(-a)\cdot(-v) = (-a)\cdot((-1)\cdot v) = ((-a)\cdot(-1))\cdot v = a\cdot v.$$

8) Multiplíquese por a^{-1} a ambos lados de $a\cdot u = a\cdot v$ y úsese la pseudoasociativa y la unimodular.

9) La igualdad $a\cdot v = b\cdot v$ equivale a $a\cdot v-(b\cdot v) = 0$. Partiendo de esta igualdad, usando 4) y una de las distributivas:

$$0 = a\cdot v-(b\cdot v) = a\cdot v+(-b)\cdot v = (a+(-b))\cdot v = (a-b)\cdot v$$

Como por hipótesis $v\neq 0$, de la propiedad 3), se sigue $a-b=0$, es decir, $a=b$. ∎

Observación 2.6. Una consecuencia inmediata de la propiedad 9) es la siguiente: *si V es un e.v. sobre un cuerpo infinito K, entonces o bien V tiene un único vector (el nulo) o bien contiene infinitos*. En efecto, supongamos que existe $v\in V$ con $v\neq 0$. Afirmamos que *la familia $L(v) := \{a\cdot v : a\in K\}$ es infinita.* De hecho, si ocurriese $a\cdot v = b\cdot v$, entonces por la propiedad 9) se tendría $a=b$. Por tanto, para diferentes valores de $a\in K$ los vectores $a\cdot v$ son distintos. Como K es infinito, entonces $L(v)$ también lo es. Y como $L(v)\subset V$, entonces V es infinito.

Ejercicio 2.7. *Demostrar que si V es un e.v. sobre K entonces se cumplen las igualdades:*

$$a\cdot(u-v) = a\cdot u-a\cdot v,$$
$$(a-b)\cdot v = a\cdot v-b\cdot v,$$

para todo $a,b\in K$ y todo $u,v\in V$.

Ejercicio 2.8. *(1) Obsérvese que el enunciado de la propiedad 6) de la proposición 2.5 sólo involucra propiedades de $(V,+)$. ¿Se puede deducir directamente de las propiedades de este grupo?*

(2) Demostrar que en un e.v. la conmutatividad de la suma de vectores es una consecuencia de los otros axiomas que aparecen en la definición.

Ejercicio 2.9. *Sea V un e.v. sobre K. Usar inducción sobre $m\in\mathbb{N}$, $m\geq 3$, para comprobar:*

$$(v_1+v_2+\ldots+v_{m-1})+v_m = v_1+(v_2+\ldots+v_{m-1}+v_m),$$

para cada familia finita $\{v_1,v_2,\ldots,v_{m-1},v_m\}$ de vectores de V. Gracias a esta propiedad y a la conmutatividad de la suma en V se sigue que podemos sumar los vectores de $\{v_1,v_2,\ldots,v_{m-1},v_m\}$ en el orden que queramos. Así, tiene sentido escribir:

$$v_1+v_2+\ldots+v_{m-1}+v_m$$

para referirnos a dicha suma.

Supongamos ahora que $a,a_1,\ldots,a_m\in K$ y $v\in V$. Demostrar por inducción:

$$a\cdot(v_1+\ldots+v_m) = a\cdot v_1+\ldots+a\cdot v_m,$$
$$(a_1+\ldots+a_m)\cdot v = a_1\cdot v+\ldots+a_m\cdot v.$$

2.1.3. Algunos ejemplos de espacios vectoriales

A continuación mostraremos diversos ejemplos que reflejan la riqueza y abundancia de espacios vectoriales en distintas ramas de las matemáticas. En cada uno indicaremos escalares, vectores, igualdad de vectores, suma de vectores y producto de escalares por vectores. Además, mostraremos cuáles son el vector cero y los vectores opuestos.

Ejemplo 2.10 (Los espacios K^n)**.** Generalizando el ejemplo de $\mathbb{R}^2(\mathbb{R})$ considerado en la observación 2.2, podemos sustituir $\mathbb{R}$ por un cuerpo (conmutativo) K cualquiera y el exponente 2 por cualquier número natural. Veámoslo explícitamente.

Sea K un cuerpo y $n \in \mathbb{N}$. Generalizando el concepto de par ordenado, una *n-úpla* en K es una colección de n elementos de K separados por comas y delimitados por paréntesis, esto es, un elemento del producto cartesiano $K \times \cdots \times^{(n)} K$ (formalmente, el producto cartesiano de n conjuntos se puede definir inductivamente a partir del ya visto para dos). Así, una n-úpla es de la forma $(x_1, \ldots, x_n)$ donde $x_i \in K$ para cada $i = 1, \ldots, n$. Diremos que dos n-úplas $(x_1, \ldots, x_n)$ e $(y_1, \ldots, y_n)$ son iguales si $x_i = y_i$ para cada $i = 1, \ldots, n$. Representaremos por K^n al conjunto de todas las n-úplas en K. Sobre este conjunto definimos las siguientes operaciones realizadas “coordenada a coordenada” mediante la suma y el producto de K:

$$(x_1, \ldots, x_n) + (y_1, \ldots, y_n) := (x_1 + y_1, \ldots, x_n + y_n), \quad a \cdot (x_1, \ldots, x_n) := (a \cdot x_1, \ldots, a \cdot x_n).$$

Usando las propiedades de la suma y del producto en K es fácil verificar que las operaciones anteriores convierten a K^n en un e.v. sobre K, que se denotará $K^n(K)$ o simplemente K^n. Claramente, el vector nulo está dado por $(0, \ldots, 0)$, mientras que el opuesto de $(x_1, \ldots, x_n)$ es $(-x_1, \ldots, -x_n)$. (Estos espacios serán muy importantes pues veremos que, en cierto sentido, son los modelos *de dimensión finita*.)

Ejercicio 2.11. *Sean V_1 y V_2 dos espacios vectoriales sobre un mismo cuerpo K. Demostrar que el producto cartesiano $V_1 \times V_2 = \{(v_1, v_2) : v_1 \in V_1,\, v_2 \in V_2\}$ es un espacio vectorial sobre K cuando se define la suma y el producto por elementos de K como:*

$$\begin{aligned} (u_1, u_2) + (v_1, v_2) &= (u_1 + v_1, u_2 + v_2), \\ a \cdot (v_1, v_2) &= (a \cdot v_1, a \cdot v_2), \end{aligned}$$

para todo $(u_1, u_2), (v_1, v_2) \in V_1 \times V_2$ y todo $a \in K$.

Ejercicio 2.12. *En $\mathbb{R}^3$ se considera la suma de elementos coordenada a coordenada y el producto por escalares reales dado por:*

$$a \star (x, y, z) = (a \cdot x, a \cdot y, 2016 \cdot a \cdot z),$$

para todo $a \in \mathbb{R}$ y $(x, y, z) \in \mathbb{R}^3$. Determínese si $\mathbb{R}^3$ con estas operaciones satisface las propiedades de un espacio vectorial real.

Ejemplo 2.13 (Espacios de matrices)**.** Sean $m, n \in \mathbb{N}$, y consideremos el conjunto $M_{m\times n}(K)$ de todas las matrices de orden $m \times n$ con coeficientes en K, del que ya hicimos un estudio preliminar. Veamos con detalle que $M_{m\times n}(K)$ tiene estructura de e.v. con la suma $A + B = (a_{ij} + b_{ij})$, para todo $A = (a_{ij}), B = (b_{ij}) \in M_{m\times n}(K)$ y el producto $a \cdot A = (a \cdot a_{ij})$, para todo $a \in K$ y $A = (a_{ij}) \in M_{m\times n}(K)$.

Para comprobar que la suma de matrices es asociativa, sean $A = (a_{ij}), B = (b_{ij}), C = (c_{ij}) \in M_{m\times n}(K)$. Entonces:

$$\begin{aligned} (A + B) + C &= (a_{ij} + b_{ij}) + (c_{ij}) = ((a_{ij} + b_{ij}) + c_{ij}) \\ &= (a_{ij} + (b_{ij} + c_{ij})) = (a_{ij}) + (b_{ij} + c_{ij}) = A + (B + C), \end{aligned}$$

donde hemos usado cómo se suman matrices y la propiedad asociativa de la suma en K.

Veamos que la suma de matrices tiene neutro. Definimos la *matriz nula* de orden $m \times n$ como la matriz de $M_{m\times n}(K)$ cuyas entradas coinciden todas con el elemento $0 \in K$. La denotaremos $0_{m\times n}$. Dada cualquier matriz $A = (a_{ij})$ en $M_{m\times n}(K)$, se tiene:

$$A + 0_{m\times n} = (a_{ij} + 0) = (a_{ij}) = A,$$
$$0_{m\times n} + A = (0 + a_{ij}) = (a_{ij}) = A,$$

donde hemos usado cómo se suman matrices y que 0 es el neutro para la suma en K.

Veamos que la suma de matrices verifica la propiedad elemento simétrico. Sea $A = (a_{ij})$ una matriz en $M_{m\times n}(K)$. La *matriz opuesta* de A es la matriz en $M_{m\times n}(K)$ definida por $-A = (-a_{ij})$. Se tiene:

$$A + (-A) = \big(a_{ij} + (-a_{ij})\big) = (0) = 0_{m\times n}, \tag{2.1}$$
$$-A + A = (-a_{ij} + a_{ij}) = (0) = 0_{m\times n}, \tag{2.2}$$

donde hemos usado cómo se suman matrices y que $-a_{ij}$ es el opuesto de a_{ij} en K.

Veamos que la suma de matrices es conmutativa. Sean $A = (a_{ij})$ y $B = (b_{ij})$ dos matrices en $M_{m\times n}(K)$. Entonces:

$$A + B = (a_{ij} + b_{ij}) = (b_{ij} + a_{ij}) = B + A,$$

donde hemos usado cómo se suman matrices y la propiedad conmutativa de la suma en[1] K.

Comprobemos ahora la propiedad pseudoasociativa. Dados $a, b \in K$ y $A = (a_{ij}) \in M_{m\times n}(K)$, se tiene:

$$(a \cdot b) \cdot A = ((a \cdot b) \cdot a_{ij}) = (a \cdot (b \cdot a_{ij})) = a \cdot (b \cdot a_{ij}) = a \cdot (b \cdot A),$$

donde hemos usado cómo se multiplican elementos de K por matrices y la propiedad asociativa del producto de K.

Comprobemos la propiedad unimodular. Dada $A = (a_{ij}) \in M_{m\times n}(K)$, se tiene:

$$1 \cdot A = (1 \cdot a_{ij}) = (a_{ij}) = A,$$

donde hemos usado que 1 es el neutro para el producto en K.

Comprobemos por último las dos propiedades distributivas. Sean $a, b \in K$ y tomemos también $A = (a_{ij})$ y $B = (b_{ij})$ en $M_{m\times n}(K)$. Se tiene:

$$\begin{aligned}
(a+b) \cdot A &= ((a+b) \cdot a_{ij}) = (a \cdot a_{ij} + b \cdot a_{ij}) = (a \cdot a_{ij}) + (b \cdot a_{ij}) = a \cdot A + b \cdot A,\\
a \cdot (A+B) &= a \cdot (a_{ij} + b_{ij}) = (a \cdot (a_{ij} + b_{ij})) = (a \cdot a_{ij} + a \cdot b_{ij})\\
&= (a \cdot a_{ij}) + (a \cdot b_{ij}) = a \cdot A + a \cdot B,
\end{aligned}$$

donde hemos usado cómo se suman matrices, cómo se multiplican elementos de K por matrices, y la propiedad distributiva en K (obsérvese que, aunque en K sólo hay una propiedad distributiva, aquí se usa "distribuyendo la suma a la izquierda" en el primer caso y "a la derecha" en el segundo).

Todo lo anterior demuestra que $M_{m\times n}(K)$ (dotado de las operaciones naturales antes definidas y que no se especifican ahora en la notación) es un e.v. sobre K. Obsérvese que el vector nulo es la matriz $0_{m\times n}$. Además, el vector opuesto de A es la matriz $-A$ antes definida.

[1] Obsérvese que hemos demostrado la propiedad conmutativa al final, pues esta es una propiedad adicional a la estructura de grupo. No obstante, si ésta se demuestra antes entonces se puede simplificar la demostración de la propiedad elemento neutro y elemento simétrico (por ejemplo, bajo conmutatividad bastaría con demostrar (2.1) para que también se verificara (2.2)).

Ejemplo 2.14. *[Espacios de funciones]* El siguiente ejemplo es relevante en Análisis Matemático. Sea X un conjunto y K un cuerpo. Una *función* de X en K es cualquier aplicación $f : X \to K$ (la cual asociará a cada elemento $x \in X$ un único elemento $f(x) \in K$, siendo X el dominio y K el codominio de f). Como para cualesquiera aplicaciones con el mismo dominio y codominio, $f, g : X \to K$ son iguales si y sólo si $f(x) = g(x)$ para todo $x \in X$ y; en este caso, escribimos $f = g$. Representaremos por $F(X,K)$ al conjunto de todas las funciones de X en K. Veamos que $F(X,K)$ admite una estructura natural de e.v. sobre K.

Dadas dos funciones $f, g : X \to K$ definimos su suma como la función

$$f+g : X \to K, \qquad (f+g)(x) := f(x)+g(x), \qquad x \in X.$$

Dado $a \in K$ y una función $f : X \to K$ definimos su producto como la función

$$a \cdot f : X \to K, \qquad (a \cdot f)(x) := a \cdot f(x), \qquad \forall x \in X.$$

Usando las propiedades de anillo unitario de K se comprueba fácilmente que $F(X,K)$ es un e.v. sobre K con las operaciones anteriores. De hecho, el vector nulo es la *función cero* $0 : X \to K$ que asocia a cada $x \in X$ el elemento $0 \in K$. Además el vector opuesto de una funcián $f : X \to K$ es la función de X en K que asocia a cada $x \in X$ el opuesto $-f(x)$ de $f(x)$ en K.

Como caso particular, si $K = \mathbb{R}$ y $X \subset \mathbb{R}$, entonces $F(X,\mathbb{R})$ es el e.v. real de todas las *funciones reales de una variable real* definidas en X. Cuando $X = \mathbb{N}$ obtenemos el e.v. real de las *sucesiones de números reales*.

Ejercicio 2.15. *Sea X un conjunto no vacío y V un espacio vectorial sobre un cuerpo K. Denotamos por $F(X,V)$ al conjunto de las aplicaciones $f : X \to V$. En $F(X,V)$ se definen la suma y el producto por elementos de K siguientes:*

$$\begin{array}{lll}(f+g)(x) = f(x)+g(x), & \forall x \in X, & \forall f, g \in F(X,V), \\ (a \cdot f)(x) = a \cdot f(x), & \forall x \in X, & \forall a \in K, \quad \forall f \in F(X,V).\end{array}$$

Demostrar que, con estas operaciones, $F(X,V)$ es un espacio vectorial sobre K.

Ejemplo 2.16 (Espacios de polinomios)**.** Existen dos maneras naturales de definir los polinomios sobre un cuerpo (conmutativo) K, del cual supondremos además que tiene característica 0 (esto es, sumando su unidad iteradamente nunca se obtiene 0). La primera de ellas es como el subconjunto de todas las *aplicaciones polinómicas* de $F(K,K)$. éstas son las aplicaciones $p \in F(K,K)$ que pueden expresarse como un *polinomio*, esto es, mediante una expresión del tipo

$$p(x) = a_0 + a_1 \cdot x + \ldots + a_n \cdot x^n, \qquad \forall x \in K, \tag{2.3}$$

donde $n \in \mathbb{N} \cup \{0\}$ (el valor de n puede variar con p, pero siempre es finito) y $a_i \in K$ para cada $i = 0, \ldots, n$ (se entiende que x^i significa operar x consigo mismo i veces en $(K, \cdot)$). Es fácil comprobar que las operaciones suma y producto por escalares ya definidas en $F(X,K)$ para el caso particular $X = K$, se inducen de manera natural en el conjunto de todas las aplicaciones polinómicas (esto es, si p, q son aplicaciones polinómicas y $a \in K$ entonces $p+q$ y $a \cdot p$ son aplicaciones polinómicas). Además, la aplicación cero se puede escribir como un polinomio (el polinomio nulo, que se puede escribir como $0 = 0 + 0x + \ldots + 0x^n$ para cualquier n) y el opuesto de una aplicación polinómica resulta ser también una aplicación polinómica. En consecuencia, se comprueba con facilidad que el conjunto de todas los aplicaciones polinómicas tiene una estructura de espacio vectorial. De hecho, cuando consideremos más adelante el espacio de los polinomios $K[x]$ lo consideraremos siempre en este sentido.

Además, por simplicidad, supondremos siempre que K es el cuerpo $\mathbb{R}$ ó $\mathbb{C}$, lo que haremos explícito denotándolo $\mathbb{K}$.

La segunda forma de definir los polinomios es interesante especialmente en álgebra; la explicaremos brevemente sin ahondar en consideraciones más penetrantes. Dado un cuerpo K, una *expresión polinómica* con coeficientes en K es una expresión del tipo:

$$p(x) = a_0 + a_1 x + \ldots + a_n x^n, \tag{2.4}$$

donde $n \in \mathbb{N} \cup \{0\}$, $a_i \in K$ para cada $i = 0, \ldots, n$ y donde $x, x^2, \ldots, x^n$ se consideran como símbolos. Se define la expresión polinómica nula como $p_0 = 0$, y diremos que una expresión polinómica no nula tiene grado n si se escribe como en (2.4) con $a_n \neq 0$. En adelante, se adopta el convenio para cualquier expresión polinómica distinta de la nula que $a_n \neq 0$ (esto es, si $a_n = 0$ no se escribe el término correspondiente $a_n x^n$ en (2.4)). Bajo este convenio, representaremos por $K[x]$ al conjunto de todas las expresiones polinómicas con coeficientes en K. Obsérvese que, cuando se consideraban funciones polinómicas, p denotaba una tal función y $p(x)$ su valor en el punto $x \in \mathbb{K}$. Al considerar expresiones polinómicas, $p(s)$ denota directamente la expresión (2.4). Para denotar polinomios es habitual considerar la notación $p(x)$, aunque luego se traten como funciones polinómicas.

Comprobemos que $K[x]$, considerado como conjunto de expresiones polinómicas, tiene una estructura natural de e.v. sobre K. Dados $p(x) = a_0 + a_1 x + \ldots + a_n x^n$ y $q(x) = b_0 + b_1 x + \ldots + b_m x^m$ en $K[x]$ con $n \leq m$, definimos:

$$p(x) + q(x) = (a_0 + b_0) + (a_1 + b_1) x + \ldots + (a_n + b_n) x^n + b_{n+1} x^{n+1} + \ldots + b_m x^m,$$

es decir, sumamos los monomios del mismo grado; una expresión análoga se toma si $n \geq m$. En el caso $n = m$, debe tenerse presente que tal vez $a_n + b_n = 0$, por lo que en este caso se suprime el correspondiente monomio de la expresión (esta precaución hay que tenerla en cuenta de nuevo para la nueva expresión polinómica si $a_{n-1} + b_{n-1} = 0$, y así sucesivamente). Dados $a \in K$ y un polinomio $p(x) = a_0 + a_1 x + \ldots + a_n x^n$ en $K[x]$, definimos:

$$a \cdot p(x) = (a \cdot a_0) + (a \cdot a_1) x + \ldots + (a \cdot a_n) x^n,$$

en el caso $a = 0$ se entiende consistentemente $a.p(x) = p_0$. No es difícil comprobar que, con las operaciones anteriores, $K[x]$ es un e.v. sobre K. El vector cero es la expresión polinómica nula p_0, mientras que el opuesto de una expresión polinómica $p(x) = a_0 + a_1 x + \ldots + a_n x^n$ es la expresión $-a_0 - a_1 x - \ldots - a_n x^n$.

Observación 2.17. Es interesante observar que en los los ejemplos anteriores la suma y el producto de escalares por vectores se construyen a partir de la suma y del producto en K. De hecho, las propiedades de las operaciones de e.v. se deducen a partir de las propiedades de anillo unitario que cumple K.

Observación 2.18. En la definición de e.v., la importancia del cuerpo K queda de manifiesto en lo siguiente: un mismo conjunto puede ser un e.v. sobre diferentes cuerpos para la misma operación suma. Por ejemplo, sabemos que, de manera natural, $\mathbb{C}$ admite una estructura de espacio vectorial sobre $\mathbb{C}$, la cual denotamos $\mathbb{C}(\mathbb{C})$ (véase el ejemplo 2.10 con $K = \mathbb{C}$ y $n = 1$). No obstante, de manera natural, $\mathbb{C}$ también admite una estructura de e.v. real, que denotaremos $\mathbb{C}(\mathbb{R})$, definida como sigue. En primer lugar, consideramos siempre su suma natural, esto es, dados $z = x + y\,i$ y $w = x' + y'\,i$ en $\mathbb{C}$, la suma está dada por:

$$z + w = (x + x') + (y + y')\,i.$$

Claramente, $(\mathbb{C}, +)$ es un grupo abeliano (de hecho, esto lo sabíamos de las propiedades de $\mathbb{C}$ como cuerpo). Definimos ahora el producto por escalares reales como la ley de composición externa $\mathbb{R} \times$

$\mathbb{C} \to \mathbb{C}$, que a cada $a \in \mathbb{R}$ y $z = x + y\,i \in \mathbb{C}$ proporciona $a \cdot z := (a \cdot x) + (a \cdot y)\,i$ (esta ley de composición externa no es más que la restricción a $\mathbb{R} \times \mathbb{C}$ del producto en $\mathbb{C}$, el cual servía a su vez como ley de composición externa para $\mathbb{C}(\mathbb{C})$). Con estas operaciones es fácil comprobar que $\mathbb{C}$ es un e.v. real.

Esta observación es general para los cuerpos $\mathbb{Q} \subset \mathbb{R} \subset \mathbb{C}$: manteniendo la *misma operación suma* $+$, todo espacio vectorial real $(V, +, \cdot\mathbb{R})$ genera un espacio vectorial sobre $\mathbb{Q}$, sin más que *restringir el dominio de la ley de composición externa* $\cdot : \mathbb{R} \times V \to V$ al subconjunto $\mathbb{Q} \times V \subset \mathbb{R} \times V$; análogamente, todo espacio vectorial sobre $\mathbb{C}$ genera uno sobre $\mathbb{R}$ (y, por tanto, sobre $\mathbb{Q}$). Más adelante veremos que hay mucha diferencia entre estas estructuras de e.v. sobre diferentes cuerpos.

2.1.4. Espacios afines

En Geometría Elemental, se suelen considerar como un ejemplo de espacio vectorial el de los vectores (segmentos orientados) de un plano que tienen un punto prefijado, el cual sirve de “origen” común para todos los vectores. La estructura geométrica de ese plano (sin prefijar el punto) y su relación con la de espacio vectorial resulta manifiesta en el siguiente concepto.

Definición 2.19. *Un* espacio afín *sobre un cuerpo conmutativo K es una terna* $(\mathcal{A}, \vec{\mathcal{A}}(K), \varphi)$ *donde* $\mathcal{A}$ *es un conjunto, a cuyos elementos llamaremos* puntos*,* $\vec{\mathcal{A}}(K)$ *es un espacio vectorial sobre K, y* φ *es una aplicación*

$$\varphi : \mathcal{A} \times \mathcal{A} \to \vec{\mathcal{A}}, \qquad (P,Q) \mapsto \varphi(P,Q), \text{ que denotaremos simplificadamente } \overline{PQ},$$

para los que se verifica:

1. *Para cada $O \in \mathcal{A}$, la aplicación $\mathcal{A} \to \vec{\mathcal{A}}$, $Q \mapsto \overline{OQ}$, es biyectiva.*
2. *Para cada $P, Q, R \in \mathcal{A}$, $\overline{PQ} + \overline{QR} = \overline{PR}$.*

A $\vec{\mathcal{A}}(K)$ le llamaremos el *espacio vectorial asociado* (o *espacio de direcciones*). En el contexto de espacios afines, al par $(P,Q) \in \mathcal{A} \times \mathcal{A}$ se le llama *vector ligado de origen P y extremo Q* y a $\overline{PQ}$ su correspondiente *vector libre*. Cuando no haya posibilidad de confusión, abusaremos de la nomenclatura llamando espacio afín a $\mathcal{A}$ (en lugar de a la tripleta formada por $\mathcal{A}, \vec{\mathcal{A}}$ y la aplicación $\mathcal{A} \times \mathcal{A} \to \vec{\mathcal{A}}$).

Proposición 2.20. *Sea $\mathcal{A}$ un espacio afín.*

1) Fijado un punto $O \in \mathcal{A}$, para cada $v \in \vec{A}$, existe un único $P \in \mathcal{A}$ tal que $\overline{OP} = v$ (esta última igualdad también la denotaremos: $P = O + v$).

2) $\overline{PQ} = 0$ si, y sólo si, $P = Q$.

3) $\overline{PQ} = -\overline{QP}$.

4) Identidad afín del paralelogramo: Si $\overline{PQ} = \overline{RS}$, entonces $\overline{PR} = \overline{QS}$.

Demostración. 1) No es más que una reformulación de la propiedad 1 en la definición.

2) Para la condición suficiente, aplicando la propiedad 2 al caso $P = Q = R$ se sigue: $\overline{PP} = \overline{PP} + \overline{PP}$ y, simplificando en $\vec{A}$, $\overline{PP} = 0$.

Para la condición necesaria, podemos decir ahora que $\overline{PQ} = 0$ implica $\overline{PQ} = \overline{PP}$ y, por la unicidad del extremo establecida en el apartado 1), se tiene $P = Q$.

3) Por la propiedad 2 de la definición, $\overline{PQ} + \overline{QP} = 0$ (la última igualdad por el apartado anterior), de donde $\overline{QP}$ es el opuesto en $\mathcal{A}$ de $\overline{PQ}$.

4) Aplicando dos veces la propiedad 2 se tiene: $\overline{PR} = \overline{PQ} + \overline{QR} = \overline{PQ} + (\overline{QS} + \overline{SR})$, (¡hágase un dibujo!) por lo que la igualdad pedida se obtiene de que $\overline{SR} = -\overline{PQ}$ (usando la hipótesis $\overline{PQ} = \overline{RS}$ y el apartado anterior) y usando las propiedades de la suma de vectores en $\vec{\mathcal{A}}$. ∎

El siguiente ejercicio justifica que todo espacio vectorial $V(K)$ se puede considerar como un espacio afín cuyo espacio de direcciones es el propio $V(K)$.

Ejercicio 2.21. *Dado un e.v. $V(K)$ se considera la terna formada por $(V, V(K), -)$, donde el primer elemento de la terna es el conjunto donde se define $V(K)$, el segundo elemento es el propio e.v. $V(K)$ y el símbolo $-$ denota la operación en V:*

$$V \times V \to V, \qquad (v, w) \mapsto v - w \ \ (:= v + (-w)).$$

Compruébese que esta terna es un espacio afín con espacio vectorial asociado $V(K)$.

Intuitivamente, un espacio vectorial puede verse como un espacio afín si lo consideramos como espacio de puntos "olvidándonos de quién era el 0". Análogamente, un espacio afín es intuitivamente un "espacio vectorial del que nos olvidamos de quién es su vector 0".

En la próxima sección veremos que, de manera natural, las rectas y planos de $\mathbb{R}^3$ pueden considerarse siempre como espacios afines (y sólo cuando además pasan por el origen, también como espacios vectoriales).

Nota 2.22 (Geometría Elemental y vectores libres)**.** En la presente nota se discuten las nociones relacionadas de vectores libres, ligados y equipolentes con las que el estudiante puede haber tenido contacto Geometría Elemental. En ella pretendemos justificar desde un punto de vista meramente intuitivo que cuando se modela un plano o el espacio ordinario (digamos, que percibimos sensorialmente), en el que presuponemos sólo que hay una noción natural de paralelismo (que no trataremos de formalizar) la estructura natural de este modelo es la de espacio afín. Por su puesto, el espacio $\mathcal{A}$ de puntos sería el propio plano o espacio, mientras que el espacio vectorial asociado es el de vectores libres, los cuales se obtienen a partir de vectores ligados por la llamada *relación de equipolencia*.

Así, nos centraremos en un plano Π en el que suponemos que las propiedades intuitivas de paralelismo conocidas de Geometría Elemental pueden llevarse a cabo. Escogido un punto $P_0 \in \Pi$, llamaremos *vector ligado* de origen P_0 y extremo $P \in \Pi$ al segmento orientado $\overline{P_0P}$ que empieza en P_0 y termina en P (formalmente este segmento queda caracterizado como el par $(P_0, P) \in \Pi \times \Pi$). En el conjunto V_{P_0} de los vectores ligados en P_0, se definen

1. La operación suma mediante:
$$\overline{P_0P} + \overline{P_0Q} := \overline{P_0R},$$
donde R es el único punto de Π obtenido como cuarto vértice para el paralelogramo construido a partir de los lados $\overline{P_0P}$ y $\overline{P_0Q}$ (entendiéndose que los otros tres vértices son P_0, P, Q, y que el paralelogramo puede degenerar en un segmento o un punto).

2. La ley de composición externa (producto por escalares) del cuerpo K, que tomaremos igual a $\mathbb{Q}$ ó $\mathbb{R}$, se construye considerando progresivamente los casos en que $a \in K$ es natural, entero, racional o irracional. Así, si $a = 2$ se construye $a \cdot \overline{P_0P}$ trasladando paralelamente el segmento $\overline{P_0P}$ hasta hacer coincidir su origen con P, y para $a = n \in \mathbb{N}$ esta operación se repite del modo obvio n veces. Para $a = -1$ definimos $-\overline{P_0P}$ trasladando paralelamente el vector ligado $\overline{PP_0} \in V_P$ hasta hacer coincidir su origen con P_0; de manera natural se define entonces $a \cdot \overline{P_0P}$ para $a \in \mathbb{Z}$.

> En Geometría Elemental, el teorema de Tales permite dividir un segmento en $n \in \mathbb{N}$ partes iguales (usando sólo el paralelismo y sin "medir su longitud"), lo cual conduce a una definición natural de $(1/n) \cdot \overline{P_0P}$ para $n \in \mathbb{N}$ y, entonces, a definir $a \cdot \overline{P_0P}$ para $a \in \mathbb{Q}$. Si se quiere extender intuitivamente este producto por escalares para todo $a \in \mathbb{R}$, podemos aproximar a por números racionales $r_n \in \mathbb{Q}$, y tomar $r \cdot \overline{P_0P}$ como límite de los $r_n \cdot \overline{P_0P}$.

Las propiedades intuitivas del paralelismo permiten justificar que, con las anteriores operaciones, $(V_{P_0}, +, \cdot K)$ admite una estructura natural de espacio vectorial para $K = \mathbb{R}$ ó $\mathbb{Q}$.

Para pasar de la estructura de espacio vectorial de los vectores ligados a cada punto a la de espacio afín de todo el plano, definiremos el concepto de vector libre como sigue. Consideramos el conjunto de todos los vectores ligados en cualquier punto $P_0 \in \Pi$:

$$V_{\text{lig}} := \{\overline{P_0P} : P_0, P \in \Pi\}.$$

A continuación, se define la siguiente relación binaria $\sim_e$, llamada *de equipolencia*, en V_{lig}:

$$\overline{P_0P} \sim_e \overline{Q_0Q} \quad \Longleftrightarrow \quad \text{el cuadrilátero que generan con } \overline{P_0Q_0} \text{ y } \overline{PQ} \text{ es un paralelogramo.}$$

Las propiedades del paralelismo permiten asegurar que $\sim_e$ es una relación de equivalencia. A cada clase de equivalencia $[\overline{P_0P}]$ se le llama *vector libre* y el conjunto cociente

$$V_{\text{lib}} := V_{\text{lig}} / \sim_e$$

es el *conjunto de los vectores libres* del plano Π. En este espacio se define la operación suma

$$[\overline{P_0P}] + [\overline{Q_0Q}] := [\overline{R_0R_P} + \overline{R_0R_Q}]$$

donde R_0 es cualquier punto escogido en Π, y R_P, R_Q se selecionan como los (únicos) puntos de Π tales que:

$$\overline{R_0R_P} \in [\overline{P_0P}], \qquad \overline{R_0R_Q} \in [\overline{Q_0Q}].$$

De nuevo, las propiedades intuitivas del paralelismo nos permiten asegurar que esta definición es consistente. De hecho, el vector libre obtenido como suma resulta independiente del punto R_0 que se escoja en Π. Análogamente, se define el producto por escalares:

$$a \cdot [\overline{P_0P}] := [a \cdot \overline{P_0P}]$$

para todo $a \in \mathbb{R}$ y $[\overline{P_0P}] \in V_{\text{lib}}$ (el cual, de nuevo, resulta ser independiente de que se tome cualquier otro representante $\overline{R_0R_P}$ de la clase $[\overline{P_0P}]$). Con las operaciones anteriores, el conjunto V_{lib} de los vectores libres resulta ser un espacio vectorial sobre K. Más aún, la terna $(\Pi, V_{\text{lib}}(K), \varphi)$ donde

$$\varphi : \Pi \times \Pi \to V_{\text{lib}}, \qquad (P, Q) \mapsto \varphi((P, Q)) := [\overline{PQ}]$$

resulta ser un espacio afín. Como notación habitual, se escribe simplemente $\overline{PQ}$ en lugar de la clase $[\overline{PQ}]$, en consistencia con la notación para φ en la definición de espacio afín, o bien se usan letras con flechas $\vec{u}, \vec{v}$, las cuales recuerdan que no se trata de puntos ni pares de puntos (sino, formalmente, clases de equivalencia).

Es de remarcar que, en Física, se suelen definir los vectores (libres o ligados) usando, además de la noción de paralelismo, la de *longitud* o *módulo* par el segmento orientado. Esto presupone que se dispone de alguna herramienta para medir esos segmentos. Si se parte del concepto de longitud, se podría simplificar nuestra construcción anterior (por ejemplo, para definir el producto por un esdalar positivo $a > 0$ bastaría con multiplicar la longitrud por a). No obstante, es de remarcar que el concepto de longitud no es necesario para todo el desarrollo que haremos aquí de espacios vectoriales, sino que es un concepto adicional (y muy importante) que se introducirá en espacios (vectoriales o afines) *métricos*.

No insistiremos en ejemplos como los de la nota anterior, que se irán aclarando a lo largo del Grado en Matemáticas.

2.2. Subespacios vectoriales

Una vez definido el concepto de espacio vectorial vamos a introducir otra de las nociones fundamentales de esta asignatura: la de subespacio vectorial. Intuitivamente, si $V(K)$ es un e.v y U es un subconjunto de V, parece lógico decir que U es un subespacio vectorial de V si U hereda de forma natural la estructura de e.v. de V, es decir, si se puede definir en U una estructura de e.v. sobre K a partir de la estructura ya existente en V. En esta sección nos ocuparemos de estudiar esta noción, así como de mostrar varios ejemplos y métodos de construcción de subespacios vectoriales.

2.2.1. Definición, caracterizaciones y ejemplos

Consideraremos en adelante un e.v. V sobre un cuerpo K y, como siempre, denotaremos por $+$ y por $\cdot$ a la suma de vectores en V y al producto de escalares por vectores, resp.

Definición 2.23. *Sea $U \subset V$ un subconjunto de V. Diremos que U es un* subespacio vectorial *(s.v.) o una* subvariedad lineal *de $V(K)$ si se verifican:*

(i) La operación suma $+ : V \times V \to V$ en V se puede restringir a U, esto es: para todo $u, v \in U$ se tiene $u + v \in U$ (se dice entonces que U es cerrado para $+$*).*

En consecuencia, se induce una operación $+_U : U \times U \to U, (u, u') \mapsto u + u'$, que, en adelante, denotaremos con el mismo símbolo $+$ que la suma en U.

(ii) El producto por escalares $\cdot : K \times V \to V$ en $V(K)$ se puede restringir a U, esto es, para todo $a \in K$ y $u \in U$, entonces $a \cdot u \in U$ (U es cerrado para $\cdot$).

En consecuencia, se induce una ley de composición externa $\cdot_U : K \times U \to U, (a, u) \mapsto a \cdot u$ que, en adelante, denotaremos con el mismo símbolo $\cdot$ que la suma en $V(K)$.

(iii) Las operaciones en U definidas en (i) y (ii) cumplen todas las propiedades de un e.v.

En resumen, U es un s.v. de V si las operaciones de $V(K)$ se pueden restringir a U y, entonces, U es un e.v. con estas operaciones restringidas [2].

La definición anterior es fácil de entender pero no resulta muy práctica, pues hay que comprobar numerosas propiedades. No obstante, vamos a demostrar que es suficiente con verificar (i) y (ii) (esto es, que se puedan restringir las operaciones a U) y que U no es vacío para tener un s.v., lo cual resultará mucho más sencillo de aplicar sobre ejemplos concretos.

[2] Se comprobará en la demostración de la Proposición 2.24 que en ese caso el neutro de $(U, +)$ tiene que coincidir con el de $(V, +)$ y el simétrico de un elemento de $(U, +)$ coincidirá con el simétrico de ese elemento en $(V, +)$. Ahí se demostrará usando las propiedades de subespacios vectoriales, pero es fácil darse cuenta de que estas propiedades son generales cuando se estudia un *subgrupo* $(U, +)$ de un grupo arbitrario $(V, +)$.

Proposición 2.24. *Sea $V(K)$ un e.v. y $U \subset V$, $U \neq \emptyset$. Se verifica que U es un s.v. de $V(K)$ si y sólo si se cumplen estas dos propiedades:*

(i) Si $u, v \in U$, entonces $u + v \in U$,

(ii) Si $a \in K$ y $u \in U$, entonces $a \cdot u \in U$.

Demostración. Supongamos primero que U es un s.v. de V. Necesariamente, U cumple (i), (ii) y (iii) de la definición de s.v. por lo que, en particular, cumple las propiedades (i) y (ii) del enunciado.

Recíprocamente, supongamos que U cumple (i) y (ii). Para demostrar que U es un s.v. de V basta entonces con demostrar que se cumple (iii), es decir, que las restricciones de $+$ y $\cdot$ al subconjunto U cumplen todas las propiedades de la definición de e.v.

Empecemos con demostrar la estructura de grupo conmutativo de $(U, +)$. La asociatividad de $(U, +)$ significa que $(u + v) + w = u + (v + w)$, para todo $u, v, w \in U$. Ahora bien, esto es inmediato, pues sabemos que esta igualdad se cumplía para cualesquiera $u, v, w \in V$ estén o no en U), por ser V un e.v. Análogamente se comprueba que la suma en U es conmutativa.

Ahora nos preguntamos si la suma en U tiene un neutro, es decir, buscamos $e \in U$ tal que $u + e = e + u = u$, para cada $u \in U$. Si demostramos que el vector 0 de V tiene que estar en U, entonces bastará con tomar $e = 0$ (al ser 0 el neutro de la suma en todo V, también lo será en U). Para probar que $0 \in U$, basta con tomar cualquier $u \in U$ (obsérvese que U no era vacío por hipótesis) y observar que $0 = 0 \cdot u$ en $V(K)$. Usando (ii) se sigue que $0 \in U$.

Veamos que la suma de U tiene opuestos. Sea $u \in U$. Queremos encontrar $v \in U$ tal que $u + v = v + u = 0$. Si probamos que $-u$ (el opuesto de u en V) está en U, entonces bastará con tomar $v = -u$ para acabar. Pero esto es consecuencia de (ii), pues $-u = (-1) \cdot u$.

Las propiedades que quedan por comprobar (pseudoasociativa, unimodular y distributivas) se demuestran igual que la asociativa y la conmutativa de la suma. Concretamente, como estas propiedades son válidas en V, y las operaciones que consideramos en U se obtienen al restringir las de V, entonces también se cumplen en U. ∎

Podemos incluso unificar las dos propiedades anteriores como sigue.

Proposición 2.25. *Sea $V(K)$ un e.v. y $U \subset V$, $U \neq \emptyset$. Se verifica que U es un s.v. de $V(K)$ si y sólo si se cumplen la siguiente propiedad:*

$$a \cdot u + b \cdot v \in U, \quad \text{para todo } a, b \in K \text{ y todo } u, v \in U. \tag{2.5}$$

Demostración. Basta con comprobar que esta propiedad se verifica si y sólo si se cumplen las propiedades (i) y (ii) en la proposición 2.24 anterior.

($\Rightarrow$) Vamos a deducir a partir de (i), (ii) la propiedad escrita en (2.5). Sean $a, b \in K$ y $u, v \in U$. Por la propiedad (ii) sabemos que $a \cdot u$ y $b \cdot v$ pertenecen a U. Por la propiedad (i) sabemos que la suma de $a \cdot u$ y $b \cdot v$ también pertenece a U, esto es, $a \cdot u + b \cdot v \in U$, como se quería.

($\Leftarrow$) Recíprocamente, supongamos que se cumple (2.5) y comprobemos que se satisfacen (i) y (ii). Para (i), sean $u, v \in U$, y nos preguntamos si $u + v \in U$. Esto se cumple porque $u + v = 1 \cdot u + 1 \cdot v$, y esta última expresión pertenece a U gracias a (2.5). Para (ii) sea ahora $a \in K$ y $u \in U$ y comprobemos que $a \cdot u \in U$. Esto se cumple porque $a \cdot u = a \cdot u + 0 = a \cdot u + 0 \cdot u$, y esta última expresión pertenece a U gracias a (2.5). ∎

Ejercicio 2.26. *Demuéstrese que un subconjunto U no vacío de $V(K)$ es un s.v. si y sólo si verifica:*

$$a \cdot u + v \in U, \quad \textit{para todo } a \in K \textit{ y todo } u, v \in U. \tag{2.6}$$

Observación 2.27 (Criterio del vector nulo)**.** Obsérvese que si $V(K)$ es un espacio vectorial entonces V no puede ser el conjunto vacío (pues $V(,+)$ debe tener un elemento neutro). Aunque se deduce de la definición de subespacio ningún s.v. puede ser vacío esta propiedad se debe imponer en las proposiciones 2.24, 2.25). Más aún, en la demostración de la proposición 2.24 se puso de manifiesto el siguiente hecho: si U es un s.v. de V, entonces $0 \in U$. Esto significa que todos los subespacios vectoriales de un e.v. deben contener forzosamente al vector cero. En particular, podemos afirmar que si $0 \notin U$, entonces U no es un s.v. de V.

Ejemplo 2.28. El criterio anterior nos dice que la propiedad de ser s.v. es restrictiva. De hecho, dentro de un e.v. hay "pocos" subespacios vectoriales. Por ejemplo, en $\mathbb{R}^2$ algunos subconjuntos sencillos como la circunferencia de ecuación $x^2 + y^2 = 1$ o la recta de ecuación $y = x + 1$ no son subespacios vectoriales, pues no contienen al vector nulo de $\mathbb{R}^2$.

Por supuesto, sería un error aplicar el criterio del vector nulo afirmando que si $0 \in U$ entonces U es un s.v. Esta afirmación NO es cierta en general, como muestra el siguiente contraejemplo. En $\mathbb{R}^2(\mathbb{R})$ tomamos el subconjunto $U = \{(0,0),(1,0)\}$. Es claro que U contiene al vector nulo de $\mathbb{R}^2$. Sin embargo, se puede comprobar de varias maneras que U no es un s.v. de $\mathbb{R}^2$. Por ejemplo, si tomamos $u = (1,0) \in U$, entonces $2 \cdot u = (2,0) \notin U$, con lo que falla la propiedad (ii) en la proposición 2.24 (la cual coincidía con la (ii) en la definición de s.v.). Con más generalidad, U no puede ser un subespacio vectorial porque tiene dos vectores y sabemos que todo e.v. real con más de un vector debe tener infinitos vectores (observación 2.6).

Algunos ejemplos de subespacios vectoriales

Ejemplo 2.29 (Subespacios vectoriales impropios)**.** Todo e.v $V(K)$ tiene siempre dos subespacios vectoriales especiales, que son $U = \{0\}$ y $U = V$. Estos subespacios se llaman[3] *impropios*. Todo s.v. de V que no sea uno de los anteriores se llamará *propio*. Por supuesto, los subespacios impropios de $V(K)$ son distintos en todo espacio vectorial excepto en el e.v. trivial $V = \{\,0\}$.

Ejemplo 2.30 (Rectas y planos)**.** Usando la Proposición 2.24 es fácil comprobar que en $\mathbb{R}^2(\mathbb{R})$ las rectas que pasan por el origen son subespacios vectoriales. Análogamente, las rectas y planos de $\mathbb{R}^3$ que pasan por el origen son s.v. de $\mathbb{R}^3$ (véase también el ejemplo 2.36). Las rectas y planos que no contienen al origen no son s.v. de $\mathbb{R}^3$ (serán *subespacios afines*).

El siguiente ejemplo servirá para poner de manifiesto la relevancia del cuerpo K cuando hablamos de subespacios vectoriales.

Ejemplo 2.31 (Dependencia con el cuerpo)**.** Sabemos que $\mathbb{C}$ admite las estructuras de e.v. real $\mathbb{C}(\mathbb{R})$ y de e.v. complejo $V(\mathbb{C})$. Veamos que las diferentes estructuras de e.v. se reflejan en los posibles subespacios vectoriales.

Sea $U = \{b\,i : b \in \mathbb{R}\}$ la familia de números complejos imaginarios (aquellos cuya parte real se anula). Veamos que *U es un s.v. de* $\mathbb{C}(\mathbb{R})$, usando la Proposición 2.24. Claramente U no es vacío ($0 \in U$), y sean $u = b\,i$ y $v = b'\,i$ vectores en U. Entonces $u + v = (b + b')\,i$, que también pertenece a U. Del mismo modo, si $u = b\,i \in U$ y $a \in \mathbb{R}$, entonces $a \cdot u = (a \cdot b)\,i$, que pertenece a U.

[3] No deben confundirse los subespacios impropios de $V(K)$ con los subconjuntos impropios de V. Estos son el vacío (que no es un subespacio impropio) y V.

Sin embargo, U *no es un s.v. de* $\mathbb{C}(\mathbb{C})$. En efecto; si lo fuera, se deberá cumplir que $a \cdot u \in U$ para cada $a \in \mathbb{C}$ y cada $u \in U$. Para ver que esta propiedad falla, basta tomar $a = i \in \mathbb{C}$ y $u = i \in U$; de hecho, se cumple $a \cdot u = i \cdot i = i^2 = -1 \notin U$.

Ejemplo 2.32 (Polinomios de grado menor o igual que n)**.** Recordemos (ejemplo 2.16) que $K[x]$ es el e.v. sobre K cuyos vectores son los polinomios en la variable x con coeficientes en un cuerpo K. Para cada $n \in \mathbb{N}$, denotamos:

$$\begin{aligned} K_n[x] &= \{p(x) \in K[x] \,:\, \text{grado}(p(x)) \leq n\} \\ &= \{a_0 + a_1 x + \ldots + a_n x_n \,:\, a_i \in K, \forall i = 0, \ldots, n\}. \end{aligned}$$

Veamos que $K_n[x]$ es un s.v. de $K[x]$. Usaremos la Proposición 2.25. Sean $a, b \in K$ y $p(x), q(x) \in K_n[x]$. Si escribimos $p(x) = a_0 + a_1 x + \ldots + a_n x_n$ y $q(x) = b_0 + b_1 x + \ldots + b_n x_n$, entonces resulta claro que:

$$a \cdot p(x) + b \cdot q(x) = (a \cdot a_0 + b \cdot b_0) + (a \cdot a_1 + b \cdot b_1)x + \ldots + (a \cdot a_n + b \cdot b_n)x^n,$$

que, obviamente, pertenece a $K_n[x]$.

Los siguientes tres ejemplos muestran espacios de sucesiones y funciones que aparecen en Análisis Matemático. Cada uno de ellos pueden verse como un subespacio de un e.v. $F(X, K)$ (ejemplo 2.14) para una elección apropiada de X con $K = \mathbb{R}$

Ejemplo 2.33 (Subespacio de sucesiones convergentes)**.** Sea $F(\mathbb{N}, \mathbb{R})$ el e.v. real de las sucesiones de números reales. Sea $U \subset F(\mathbb{N}, \mathbb{R})$ el subconjunto formado por las sucesiones convergentes. Propiedades elementales de Analisis Matemático demuestran que las proposiciones 2.24, 2.25 son aplicables a U, por lo que U resulta ser un s.v. de $F(\mathbb{N}, \mathbb{R})$.

Ejemplo 2.34 (Subespacios de funciones continuas y derivables)**.** Sea $F(I, \mathbb{R})$ el e.v. real de las funciones $f : I \to \mathbb{R}$, donde I es un intervalo de $\mathbb{R}$ que contenga más de un punto. Dentro de $F(I, \mathbb{R})$ consideramos los subconjuntos U y W formados, respectivamente, por las funciones continuas y por las derivables en todo I. Propiedades elementales de Analisis Matemático demuestran que las proposiciones 2.24, 2.25 son aplicables a U y W, por lo que ambos resultan ser son subespacios vectoriales de $F(I, \mathbb{R})$.

Ejemplo 2.35 (Subespacios de funciones integrables)**.** Sea $F([a, b], \mathbb{R})$ el e.v. real de las funciones $f : [a, b] \to \mathbb{R}$. Dentro de $F([a, b], \mathbb{R})$ consideramos el subconjunto U formado por las funciones acotadas e integrables. Propiedades elementales de Analisis Matemático demuestran que las proposiciones 2.24, 2.25 son aplicables a U, por lo que U resulta ser un s.v. de $F([a, b], \mathbb{R})$.

Ejemplos en SEL y matrices simétricas/antisimétricas

Por su importancia en nuestros objetivos de Geometría, detallemos algunos ejemplos de subespacios vectoriales como los obtenidos en K^n como solución a un SEL homogéneo o los subespacios de las matrices cuadradas formados por las matrices simétricas y por las antisimétricas.

Ejemplo 2.36 (Subespacio de soluciones de un SEL homogéneo)**.** Sea K un cuerpo. Consideremos un SEL de m ecuaciones y n incógnitas con coeficientes en K:

$$\left\{ \begin{array}{ccccccccc} a_{11}x_1 & + & a_{12}x_2 & + & \ldots & + & a_{1n}x_n & = & b_1 \\ a_{21}x_1 & + & a_{22}x_2 & + & \ldots & + & a_{2n}x_n & = & b_2 \\ \vdots & & \vdots & & & & \vdots & & \vdots \\ a_{m1}x_1 & + & a_{m2}x_2 & + & \ldots & + & a_{mn}x_n & = & b_m \end{array} \right. .$$

Sea U el subconjunto de K^n formado por todas las soluciones del SEL. Una condición necesaria para que U sea un s.v. de K^n es que $(0,\ldots,0)\in U$. Ahora bien, esto ocurrirá si y sólo si $b_1=\ldots=b_m=0$, es decir, el SEL es homogéneo. En tal caso, no es difícil verificar mediante la Proposición 2.25 que U es un s.v. de K^n, al cual llamaremos *subespacio de soluciones del* SEL.

Con más detalle, supongamos que $a,b\in K$ y $u=(\alpha_1,\ldots,\alpha_n), v=(\beta_1,\ldots,\beta_n)\in U$. Queremos dempstrar que el vector $a\cdot u+b\cdot v=(a\cdot\alpha_1+b\cdot\beta_1,\ldots,a\cdot\alpha_n+b\cdot\beta_n)$ está en U. Fijamos un índice $i=1,\ldots,m$ y comprobamos que se cumple la ecuación i-ésima:

$$\begin{aligned}&a_{i1}\,(a\cdot\alpha_1+b\cdot\beta_1)+\ldots+a_{in}\,(a\cdot\alpha_n+b\cdot\beta_n)\\&=a\,(a_{i1}\,\alpha_1+\ldots+a_{in}\,\alpha_n)+b\,(a_{i1}\,\alpha_1+\ldots+a_{in}\,\alpha_n)=0,\end{aligned}$$

donde se han usado propiedades del cuerpo K, así como que $u,v\in U$. Esto demuestra que $a\cdot u+b\cdot v$ es una solución de todas las ecuaciones del SEL, es decir, $a\cdot u+b\cdot v\in U$.

Como conclusión, las soluciones de un SEL homogéneo no se distribuyen de cualquiera manera dentro de K^n, sino que lo hacen de forma que resulte un s.v. de[4] K^n. Volveremos a esta cuestión más adelante cuando hablemos de las ecuaciones cartesianas de un s.v. de K^n.

Observación 2.37. En general, si tenemos un sistema de m ecuaciones *no lineales* con coeficientes en K y n incógnitas que sea homogéneo en el sentido de que incluya al vector nulo como una solución, entonces el conjunto de soluciones U puede ser o no un s.v. de K^n. Por ejemplo $U=\{(x,y)\in\mathbb{R}^2 : x^2-y=0\}$ no es un s.v. de $\mathbb{R}^2$. Para comprobarlo, basta con observar que los vectores $u=(1,1)$ y $v=(2,4)$ están en U, mientras que $u+v=(3,5)$ no lo está. Por otro lado, $U=\{(x,y)\in\mathbb{R}^2 : x^2+y^2=0\}=\{(0,0)\}$, que sí es un s.v. (impropio) de $\mathbb{R}^2$.

Ejemplo 2.38 (Matrices simétricas y antisimétricas como subespacios de $M_n(K)$)**.** Recordemos que, dada una matriz $A=(a_{ij})$ en $M_{m\times n}(K)$, se definía la *matriz traspuesta* de A como la matriz A^t en $M_{n\times m}(K)$ cuyas filas se obtienen escribiendo ordenadamente las columnas de A, es decir, si $A=(a_{ij})$ entonces $A^t=(a_{ji})$. Así, la trasposición de matrices definía una aplicación

$$^t : M_{m\times n}(K)\to M_{n\times m}(K), \qquad A\mapsto A^t$$

que cumple las siguientes propiedades (proposición 1.23):

1) $(A+B)^t=A^t+B^t$, para cada $A,B\in M_{m\times n}(K)$,

2) $(a\cdot A)^t=a\cdot A^t$, para cada $a\in K$ y cada $A\in M_{m\times n}(K)$,

3) $(A^t)^t=A$, para cada $A\in M_{m\times n}(K)$.

A continuación trabajaremos en el espacio $M_n(K)$ de las matrices cuadradas de orden n con coeficientes en K. Obsérvese que para matrices cuadradas se tiene[5]:

$$^t : M_n(K)\to M_n(K), \qquad t\circ t=I_{M_n(K)} \text{ (aplicación identidad en } M_n(K)), \quad \text{esto es, } t^{-1}=t.$$

[4]Un ejemplo de las dificultades que se encontrarían si prescindiéramos de la hipótesis de que el cuerpo K es conmutativo (digamos, considerando que $K^n(K)$ es un espacio vectorial por la izquierda, véase la observación 2.2(5)) sería que no se obtiene el mismo SEL si se multiplican los coeficientes a la derecha de las incógnitas que si se hace a la izquierda. Este tipo de discusiones excede nuestros objetivos.

[5]A una aplicación que coindice con su inversa se le llama *involutiva*. Así, para matrices cuadradas, la trasposición es pues una aplicación involutiva. Para matrices no cuadradas esto no es estrictamente cierto ya que, aunque se verifique $(A^t)^t=A$, los dos símbolos de trasposición denotan ahí aplicaciones con distinto dominio y codominio. Pese a todo, a la propiedad $(A^t)^t=A$ se le suele llamar *involución*, como hicimos en la proposición 1.23.

Diremos que una matriz $A = (a_{ij})$ en $M_n(K)$ es *simétrica* si $A^t = A$ (nótese que esto sólo tiene sentido para matrices cuadradas), esto es, $a_{ji} = a_{ij}$, para cada $i, j = 1, \dots, n$. Para estas matrices no hay ninguna restricción sobre los elementos diagonales a_{ii} que constituyen la *diagonal principal* de la matriz cuadrada, pero la parte de la matriz "por debajo de la diagonal principal" tiene que coincidir con la parte "por encima de la diagonal principal". Ejemplos triviales de matrices simétricas son la matriz nula 0_n y la matriz identidad I_n o la matriz $\begin{pmatrix} 1 & 2 \\ 2 & 0 \end{pmatrix} \in M_2(\mathbb{R})$. De hecho, toda matriz diagonal $A = (a_{ij})$ (esto es, tal que $a_{ij} = 0$ si $i \neq j$) es simétrica.

Diremos que una matriz $A = (a_{ij})$ en $M_n(K)$ es *antisimétrica* si $A^t = -A$, esto es, $a_{ji} = -a_{ij}$, para cada $i, j = 1, \dots, n$. Para estas matrices la parte de la matriz "por debajo de la diagonal principal" es opuesta a la parte "por encima de la diagonal principal". Esto implica $a_{ii} = 0$ para todo $i = 1, \dots, n$ en el caso de que K sea un cuerpo con *característica distinta de 2*, esto es, en donde se tenga $1 + 1 \neq 0$ (como ocurre en $\mathbb{R}, \mathbb{Q}$ ó $\mathbb{C}$) pues de $a_{ji} = -a_{ij}$ se deduce $2a_{ij} = 0$. Las matrices 0_n y $\begin{pmatrix} 0 & -2 \\ 2 & 0 \end{pmatrix} \in M_2(\mathbb{R})$ son antisimétricas, mientras que I_n no lo es.

En general, una matriz $A \in M_n(K)$ no será simétrica ni antisimétrica. Denotaramos los conjuntos de matrices simétricas y antisimétricas:

$$S_n(K) = \{A \in M_n(K) : A^t = A\} \qquad A_n(K) = \{A \in M_n(K) : A^t = -A\}.$$

Es fácil demostrar que $S_n(K)$ *y* $A_n(K)$ *son s.v. de* $M_n(K)$. Lo comprobaremos solamente para $S_n(K)$ (la prueba para $A_n(K)$ es análoga y se deja como ejercicio). Para ello emplearemos la Proposición 2.25. Sean $a, b \in K$ y $A, B \in S_n(K)$. Queremos comprobar que $a \cdot A + b \cdot B \in S_n(K)$, es decir, $(a \cdot A + b \cdot B)^t = a \cdot A + b \cdot B$. Esto es cierto por la siguiente cadena de igualdades:

$$(a \cdot A + b \cdot B)^t = (a \cdot A)^t + (b \cdot B)^t = a \cdot A^t + b \cdot B^t = a \cdot A + b \cdot B,$$

donde hemos usado las propiedades 1) y 2) de la trasposición de matrices, así como las igualdades $A^t = A$ y $B^t = B$ que se verifican porque $A, B \in S_n(K)$.

2.2.2. Subespacio generado por una familia de vectores

En esta sección mostraremos un método para construir de forma rápida subespacios vectoriales de cualquier espacio vectorial y empezaremos también a vislumbrar la idea fundamental en la teoría de espacios vectoriales consistente en que una cantidad "pequeña" (finita en muchos de los casos que nos interesarán) de vectores puede generar todo el espacio vectorial. Comencemos con un ejemplo.

Ejemplo 2.39. En $\mathbb{R}^3$ consideramos el subconjunto dado por $U = \{(x, y, z) \in \mathbb{R}^3 : x - 2y + 3z = 0\}$. Como U es el conjunto de soluciones de una ecuación lineal homogénea entonces U es un s.v. de $\mathbb{R}^3$ (intuitivamente, un plano en $\mathbb{R}^3$ que pasa por el origen). Nos preguntamos ahora cómo se pueden describir expresamente todos los vectores de U. Esto nos lleva a calcular todas las soluciones de la ecuación $x - 2y + 3z = 0$. Si ponemos $y = \lambda$ y $z = \mu$, entonces $x = 2\lambda - 3\mu$. Por tanto, todos los vectores de U son de la forma (x, y, z) con $x = 2\lambda - 3\mu$, $y = \lambda$ y $z = \mu$, con $\lambda, \mu \in \mathbb{R}$. Dicho de otro modo:

$$U = \{(2\lambda - 3\mu, \lambda, \mu) \in \mathbb{R}^3 \,|\, \lambda, \mu \in \mathbb{R}\},$$

que nos da una descripción de los vectores de U en función de los parámetros reales λ y μ. Ahora, si en la expresión de los vectores de U "separamos los parámetros", tenemos que:

$$(2\lambda - 3\mu, \lambda, \mu) = (2\lambda, \lambda, 0) + (-3\mu, 0, \mu) = \lambda \cdot (2, 1, 0) + \mu \cdot (-3, 0, 1).$$

La última igualdad nos permite expresar U como:

$$U = \{\lambda \cdot (2,1,0) + \mu \cdot (-3,0,1) \,|\, \lambda, \mu \in \mathbb{R}\}.$$

Esto significa que U está formado exactamente por los vectores de $\mathbb{R}^3$ que se obtienen a partir de sólo dos de ellos, el $(2,1,0)$ y el $(-3,0,1)$, cuando los multiplicamos por escalares reales y sumamos. Dicho de otro modo, usando solo los vectores $(2,1,0)$ y $(-3,0,1)$ podemos recuperar o generar todos los demás vectores de U mediante la suma y el producto por escalares (las operaciones de $\mathbb{R}^3$ como espacio vectorial real). Tiene sentido entonces decir que U es el *plano vectorial generado* por $(2,1,0)$ y $(-3,0,1)$. No es difícil dibujar esta situación en $\mathbb{R}^3$ (o pensarla en el espacio de vectores libres del espacio) para tener una visión geométrica de lo que estamos haciendo.

El anterior ejemplo motiva las siguientes definiciones.

Definición 2.40. *Sea $V(K)$ un e.v. y $S = \{v_1, \ldots, v_m\}$ un conjunto finito no vacío de vectores de V. Una* combinación lineal *(c.l.) de S es cualquier vector v de V obtenido al multiplicar cada v_i por un escalar $a_i \in K$ y después sumar los vectores resultantes, esto es, que se puede expresar como:*

$$v = a_1 \cdot v_1 + \ldots + a_m \cdot v_m, \quad \textit{donde } a_i \in K, \forall i \in \{1, \ldots, m\}.$$

Denotaremos $L(S)$ al subconjunto de V formado por los vectores obtenidos como c.l. de S, esto es[6]*:*

$$L(S) = L(v_1, \ldots, v_m) = \{a_1 \cdot v_1 + \ldots + a_m \cdot v_m \,:\, a_i \in K, \forall i = 1, \ldots, m\}.$$

Si $S \subset V$ es cualquier conjunto infinito de vectores, denotaremos $L(S)$ al subconjunto de V formado por todas las combinaciones lineales *de subconjuntos* finitos *(no vacíos) de vectores de S, esto es:*

$$L(S) = \{a_1 \cdot v_1 + \ldots + a_m \cdot v_m \,:\, m \in \mathbb{N}, v_i \in S, a_i \in K, \forall i = 1, \ldots, m\}.$$

Observación 2.41. Es fácil comprobar (¡hágase!): (a) cualquier subconjunto no vacío $S' \subset S$ verifica $L(S') \subset L(S)$ (en particular, en el caso finito $L(S)$ es también igual al conjunto de todas las combinaciones lineales de subconjuntos finitos de S) y (b) si en la expresiones de las combinaciones lineales hubiera dos vectores $v_i, v_j, i \neq j$ repetidos, se obtendría el mismo conjunto $L(S)$ al suprimir uno de ellos (por lo que abusaremos de notación permitiendo listar repetidamente los elementos[7] de S).

Ejemplo 2.42. Si tomamos los vectores de $\mathbb{R}^3$ dados por $u = (1,0,0)$ y $v = (0,0,1)$, entonces $L(\{u,v\}) = \{a \cdot u + b \cdot v \,|\, a,b \in \mathbb{R}\} = \{(a,0,b) \,:\, a,b \in \mathbb{R}\}$.

Nótese que el conjunto S no tiene por qué ser un s.v. de V (en el ejemplo anterior ni siquiera contenía al vector nulo). Sin embargo, demostraremos enseguida que $L(S)$ sí es un s.v. De hecho, es el s.v. de V "más pequeño" que contiene a S y, por tanto, el más próximo a S en cierto sentido.

Proposición 2.43. *Sea $V(K)$ un e.v. $S \subset V$ no vacío. Entonces $L(S)$ es un s.v. de V que contiene a S. Más aún, $L(S)$ es el s.v. más pequeño que contiene a S, en el sentido de que si U es un s.v. de V que verifica $S \subset U$, entonces $L(S) \subset U$.*

[6] Dicho de otro modo, un vector $v \in V$, cumple $v \in L(S)$ si y sólo si existen escalares $a_1, \ldots, a_m \in K$ tales que $v = a_1 \cdot v_1 + \ldots + a_m \cdot v_m$.

[7] Más formalmente, $L(S)$ se puede definir cuando S es una m-úpla de vectores; en este caso, resulta independiente de la reordenación de sus componentes y $L(S) = L(S')$ si S' es una $(m-1)$-úpla obtenida suprimiendo un vector repetido en S.

Demostración. Para comprobar que $L(S)$ es un s.v. de V usaremos la Proposición 2.25. Sean $a,b \in K$ y $u,v \in L(S)$. Por definición de $L(S)$ podemos expresar u y v como combinaciones lineales finitas de vectores de S. Esto significa que $u = a_1 \cdot u_1 + \ldots + a_m \cdot u_m$ para ciertos $m \in \mathbb{N}$, $u_1, \ldots, u_m \in S$ y $a_1, \ldots, a_m \in K$, mientras que $v = b_1 \cdot v_1 + \ldots + b_n \cdot v_n$ para ciertos $n \in \mathbb{N}$, $v_1, \ldots, v_n \in S$ y $b_1, \ldots, b_n \in K$. Por tanto:

$$\begin{aligned} a \cdot u + b \cdot v &= a \cdot (a_1 \cdot u_1 + \ldots + a_m \cdot u_m) + b \cdot (b_1 \cdot v_1 + \ldots + b_n \cdot v_n) \\ &= (a \cdot a_1) \cdot u_1 + \ldots + (a \cdot a_m) \cdot u_m + (b \cdot b_1) \cdot v_1 + \ldots + (b \cdot b_n) \cdot v_n, \end{aligned}$$

donde hemos usado la propiedad distributiva respecto de la suma de vectores y también la pseudoasociativa. Teniendo en cuenta la observación 2.41 (b), la expresión anterior es una c.l. de $m+n$ vectores de S. Concluimos por tanto que $a \cdot u + b \cdot v \in L(S)$, como se quería.

Veamos que $S \subset L(S)$. Dado $v \in S$, para comprobar que $v \in L(S)$ tenemos que expresar v como c.l. finita de vectores de S. Pero esto es obvio, ya que $v = 1 \cdot v$ por la propiedad unimodular.

Por último, supongamos que U es un s.v. de V con $S \subset U$. Queremos ver que $L(S) \subset U$. Sea $v \in L(S)$. Esto implica que $v = a_1 \cdot v_1 + \ldots + a_m \cdot v_m$ para ciertos $m \in \mathbb{N}$, $v_1, \ldots, v_m \in S$ y $a_1, \ldots, a_m \in K$. Como U es un s.v. de V y cada $v_i \in U$ se sigue que $v \in U$ (U es cerrado para la suma y el producto por escalares). Esto concluye la demostración. ∎

Ejemplo 2.44. Para $S = \{v\}$ a $L(\{v\})$ se le denotará también $L(v)$. En particular, cuando $v = 0$ se tien $L(0) = \{0\}$; cuando $v \neq 0$, al s.v. $L(v)$ se le llama *recta vectorial de V generada por v*.

La proposición y observación anteriores permiten redefinir $L(S)$ para cualquier $S \subset V$ como sigue.

Definición 2.45. *Sea V un e.v. sobre K y $S \subset V$ con $S \neq \emptyset$. Llamaremos* envolvente lineal *de S o* subespacio vectorial de V generado por *S al menor subespacio vectorial $L(S)$ que contiene a S, esto es, si $S = \emptyset$ entonces*[8] *$L(S) = \{0\}$ y, si $S \neq \emptyset$ entonces $L(S)$ es el conjunto de todas las combinaciones lineales (finitas) de vectores de S según la Definición 2.40.*

2.2.3. Operaciones con subespacios vectoriales

A continuación construiremos nuevos subespacios vectoriales a partir de subespacios vectoriales previamente dados. Recordemos en primer lugar los conceptos conjuntistas de unión e intersección.

Sobre los conceptos de intersección y unión conjuntista

Sea X un conjunto, y $P(X)$ el conjunto de sus partes, formado por todos los subconjuntos de X. Dados $A,B \in P(X)$ se define su intersección $A \cap B$ y su unión $A \cup B$ como los siguientes subconjuntos de X:

$$A \cap B := \{x \in X : x \in A \text{ y } x \in B\} \qquad A \cup B := \{x \in X : x \in A \text{ ó } x \in B\}.$$

Obsérvese que tanto $\cap$ como $\cup$ pueden verse como operaciones en $P(X)$.

Ejercicio 2.46. *Compruébese: (a) las operaciones unión e intersección verifican las propiedades asociativa y conmutativa, y (b) conjuntamente verifican las siguientes propiedades distributivas:*

$$A \cap (B \cup C) = (A \cap B) \cup (A \cap C), \qquad A \cup (B \cap C) = (A \cup B) \cap (A \cup C), \qquad \forall A,B,C \in P(X)$$

[8] Esto es, $L(\emptyset)$ no es el vacío, sino el subespacio impropio $\{0\}$, ya que este es el subespacio vectorial más pequeño de V (y, como todo conjunto, contiene a $\emptyset$).

¿Cuál es el elemento neutro para la operación intersección en $P(X)$? ¿Y para la unión? [9]

La asociatividad permite definir inductivamente la intersección y unión de un número finito de subconjuntos de X. También se puede definir directamente para una colección arbitraria (finita o no) de subconjuntos $\{A_i \in P(X) : i \in I\}$ (donde I es cualquier conjunto de índices para etiquetar los elementos de la colección) como sigue:

$$\cap_{i\in I} A_i := \{x \in X \,|\, x \in A_i, \forall i \in I\}, \qquad \cup_{i\in I} A_i := \{x \in X \,|\, \exists i \in I : x \in A_i\}.$$

Claramente, esta definición es consistente con la previa para dos o un conjunto finito de subconjuntos[10]. Tras estos preliminares, empezaremos hablando sobre la intersección de subespacios que resulta especialmente sencilla. Después estudiaremos la unión, que lleva aparejada el concepto de suma de subespacios y, finalmente, la suma directa. En nuestro estudio, consideraremos una familia finita de subespacios, aunque la extensión al caso infinito es sencilla. Según el nivel de dificultad que encuentre el lector, le recomendamos que haga un estudio detallado considerando sólo dos subespacios o bien con una familia infinita de ellos.

Intersección de subespacios

Definición 2.47. *Sea $V(K)$ un e.v. y $\{U_1, \ldots, U_m\}$ una familia finita de subespacios vectoriales de $V(K)$. La* intersección *de dicha familia es el subconjunto de V definido por su intersección conjuntista, esto es:*

$$\bigcap_{i=1}^{m} U_i := U_1 \cap \ldots \cap U_m = \{v \in V : v \in U_i \text{ para cada } i \in I\}.$$

Comprobemos que tal intersección resulta ser un s.v. de V.

Proposición 2.48. *Si $\{U_1, \ldots, U_m\}$ es una familia finita de subespacios vectoriales de $V(K)$, entonces la intersección $U = \cap_{i=1}^{m} U_i$ es un s.v. de V.*

Demostración. Usaremos la Proposición 2.25. Claramente, la intersección de subespacios es no vacía, pues todos ellos contienen al vector nulo. Sean $a, b \in K$ y $u, v \in U$. Queremos demostrar que $a \cdot u + b \cdot v \in U$, es decir, que $a \cdot u + b \cdot v \in U_i$, para cada $i = 1, \ldots, m$. Ahora bien, como $u, v \in U$, entonces $u, v \in U_i$, para cada $i = 1, \ldots, m$. Finalmente, como cada U_i es un s.v. de V se concluye que $a \cdot u + b \cdot v \in U_i$, para cada $i = 1, \ldots, m$. ∎

Observación 2.49. (1) La intersección $U = \cap_{i=1}^{m} U_i$ es el *mayor* s.v. de V que está contenido a la vez en cada U_i. En efecto, si W es otro s.v. de V de manera que $W \subset U_i$, para cada $i = 1, \ldots, m$, entonces es obvio que $W \subset U$, por definición de intersección.

(2) Con una prueba formalmente análoga se demuestra que la intersección de cualquier familia (no necesariamente finita) de subespacios vectoriales de V es un s.v. de V.

[9] La estructura precisa de las operaciones intersección, $A \cap B$, y unión, $A \cup B$, junto a la de tomar complementario en X, denotado $X \setminus A$ ó $X - A$, se estudian en lógica de conjuntos dentro de las *álgebras de Boole*.

[10] Esta definición formal, aunque muy intuitiva, puede no obstante resultar un poco chocante cuando se lleva al límite. Así, si se tomara $I = \emptyset$ se obtendría $\cap_{i\in\emptyset} A_i = X$, $\cup_{i\in\emptyset} A_i = \emptyset$.

Suma de subespacios vectoriales

Como la intersección de subespacios vectoriales es un subespacio vectorial, tiene sentido plantearse si lo mismo ocurre con la unión $\bigcup_{i=1}^{m} U_i$ de subespacios vectoriales $U_1,\ldots,U_m$ de $V(K)$:

$$\bigcup_{i=1}^{m} U_i := U_1 \cup \ldots \cup U_m = \{v \in V \,:\, v \in U_i \text{ para algún } i \in I\}.$$

En general, *la unión de subespacios vectoriales de un mismo e.v. $V(K)$ no es un subespacio vectorial* (aunque siempre habrá ejemplos particulares en que esta unión sí sea un s.v.).

Ejemplo 2.50. Cada eje de coordenadas es un s.v. de $\mathbb{R}^2$ (es una recta que pasa por el origen). Sin embargo, la unión de los dos ejes no lo es, pues si tomamos u y v vectores no nulos con cada uno de ellos en un eje distinto, se tiene entonces $u+v$ no permanece en la unión de los ejes. Por otra parte, en $\mathbb{R}^3$ la unión de un plano U que pasa por el origen y de una recta $W \subset U$ que pasa por el origen es igual a U, que es trivialmente un s.v. de $\mathbb{R}^3$.

Desde un punto de vista numérico el contraejemplo anterior es el siguiente. En $\mathbb{R}^2$ consideramos $U_1 = \{(x,y) \in \mathbb{R}^2 \,:\, y = 0\}$ y $U_2 = \{(x,y) \in \mathbb{R}^2 \,:\, x = 0\}$. Ya sabemos que U_1 y U_2 son s.v. de $\mathbb{R}^2$ (son conjuntos de soluciones de ecuaciones lineales y homogéneas). Para comprobar que $U_1 \cup U_2$ no es un s.v. de $\mathbb{R}^2$, tomando $u = (1,0) \in U_1 \cup U_2$ y $v = (0,1) \in U_1 \cup U_2$, obtenemos que $u+v = (1,1)$, que no pertenece a $U_1 \cup U_2$.

Dado que la unión de subespacios vectoriales no tiene por qué ser un nuevo s.v., tiene sentido plantearse la siguiente cuestión: ¿es siempre posible encontrar un s.v. U de V que contenga a todos los U_i y que tenga "algo que ver" con ellos? Obviamente, si tomamos $U = V$ tenemos un s.v. que contiene a todos los U_i. Ahora bien, esto no es satisfactorio, pues al elegir V no estamos teniendo en cuenta la forma concreta de los U_i. Lo ideal para U es que fuese el s.v. "más pequeño que contenga a todos los U_i" (en el sentido de la proposición 2.45) para que, de algún modo, "esté próximo" a los U_i. Esta idea motiva la siguiente definición.

Definición 2.51. *Sea $V(K)$ un e.v. y $\{U_1,\ldots,U_m\}$ una familia finita de subespacios vectoriales de V. Definimos la* suma *de la familia como el subconjunto de V dado por:*

$$\sum_{i=1}^{m} U_i = U_1 + \ldots + U_m = L\Big(\bigcup_{i=1}^{m} U_i\Big).$$

Esto es (véase la proposición 2.43), $\sum_{i=1}^{m} U_i$ es un s.v. de V que contiene a todos los U_i y que, además, es el subespacio vectorial más pequeño que verifica esta propiedad (si U es cualquier otro s.v. de V con $U_i \subset U$ para cada $i = 1,\ldots,m$, entonces $\sum_{i=1}^{m} U_i \subset U$).

Demostremos a continuación que cada $u \in \sum_{i=1}^{m} U_i$ se puede expresar como una suma finita de vectores $u_i \in U_i$, lo que justifica la notación empleada para el subespacio suma.

Proposición 2.52. *Sea $V(K)$ un e.v. y $\{U_1,\ldots,U_m\}$ una familia finita de subespacios vectoriales de $V(K)$. Entonces:*

$$\sum_{i=1}^{m} U_i = \{u_1 + \ldots + u_m \,|\, u_i \in U_i, \forall\, i = 1,\ldots,m\}.$$

Por tanto, dado $u \in V$, se cumple que $u \in \sum_{i=1}^{m} U_i$ si y sólo si $u = u_1 + \ldots + u_m$, donde $u_i \in U_i$ para cada $i = 1,\ldots,m$.

Demostración. Denotaremos por U al subconjunto de vectores de V siguiente:

$$U = \{u_1 + \ldots + u_m \,|\, u_i \in U_i, \forall\, i = 1, \ldots, m\}.$$

Para demostrar que $\sum_{i=1}^m U_i = U$, procederemos por doble inclusión.

($\subset$). Para demostrar $\sum_{i=1}^m U_i \subset U$, es suficiente con ver que U es un s.v. de V con $U_i \subset U$ para cada $i = 1, \ldots, m$ (recuérdese la proposición 2.43). Sean $a, b \in K$ y $u, v \in U$. Por definición de U tenemos que $u = u_1 + \ldots + u_m$ y $v = v_1 + \ldots + v_m$ con $u_i, v_i \in U_i$, para cada $i = 1, \ldots, m$. Veamos que $a \cdot u + b \cdot v \in U$. Para ello:

$$\begin{aligned} a \cdot u + b \cdot v &= a \cdot (u_1 + \ldots + u_m) + b \cdot (v_1 + \ldots + v_m) \\ &= (a \cdot u_1 + b \cdot v_1) + \ldots + (a \cdot u_m + b \cdot v_m). \end{aligned}$$

Si llamamos $w_i = a \cdot u_i + b \cdot v_i$, entonces $a \cdot u + b \cdot v = w_1 + \ldots + w_m$, donde cada $w_i \in U_i$ (por ser U_i un s.v. de V). Así, $a \cdot u + b \cdot v \in U$, lo que prueba que U es un s.v. de V. Además, si $u \in U_j$ entonces $u = 0 + \ldots + 0 + u + 0 + \ldots + 0$, donde cada u ocupa la posición j-ésima y cada sumando está en el correspondiente s.v. U_i. Por tanto, $u \in U$. Esto muestra que $U_j \subset U$ para cada $j = 1, \ldots, m$.

($\supset$). Finalmente, veamos que $U \subset \sum_{i=1}^m U_i$. Dado $u \in U$, se tiene que $u = u_1 + \ldots + u_m$ con $u_i \in U_i$ para cada $i = 1, \ldots, m$. Recordemos que $\sum_{i=1}^m U_i = L(\cup_{i=1}^m U_i)$, por lo que $\sum_{i=1}^m U_i$ está formado por las combinaciones lineales finitas de vectores de $\cup_{i=1}^m U_i$. En particular, como podemos escribir $u = 1 \cdot u_1 + \ldots + 1 \cdot u_m$, se sigue que $u \in \sum_{i=1}^m U_i$. ∎

Observación 2.53. La definición y propiedades de la suma de una familia de subespacios vectoriales se extiende de de cualquier familia (no necesariamente finita) de subespacios de manera formalmente análoga, sin más que tener en cuenta que las sumas que se hagan de elementos del espacio vectorial serán siempre finitas.

Ejemplo 2.54. En $\mathbb{R}^3$ si tomamos como U_1 y U_2 dos ejes coordenados, entonces es fácil comprobar que $U_1 + U_2$ es el plano vectorial que los contiene. Veamos este ejemplo numéricamente.

Sean U_1 y U_2 los subespacios de $\mathbb{R}^3$ dados por $U_1 = \{(x, y, z) \in \mathbb{R}^3 : x = z = 0\}$ y $U_2 = \{(x, y, z) \in \mathbb{R}^3 : y = z = 0\}$. Sabemos que $U_1 + U_2 = \{u_1 + u_2 \,|\, u_1 \in U_1, u_2 \in U_2\}$. Ahora, si $u \in U_1$ entonces $u_1 = (0, y, 0)$, mientras que si $u_2 \in U_2$ entonces $u_2 = (x, 0, 0)$. Así, $u_1 + u_2 = (x, y, 0)$. Esto prueba $U_1 + U_2 \subset \{(x, y, z) \in \mathbb{R}^3 : z = 0\}$. De hecho, se tiene $U_1 + U_2 = \{(x, y, z) \in \mathbb{R}^3 : z = 0\}$. Para probar la inclusión $\{(x, y, z) \in \mathbb{R}^3 : z = 0\} \subset U_1 + U_2$ nótese que si $v = (x, y, 0)$, entonces $v = (0, y, 0) + (x, 0, 0) = u_1 + u_2$ donde $u_1 \in U_1$ y $u_2 \in U_2$.

Ejercicio 2.55. *Demostrar que $U_1 + U_2 = \mathbb{R}^2$, donde U_1 y U_2 son los subespacios dados por $U_1 = \{(x, y) \in \mathbb{R}^2 : y = 0\}$ y $U_2 = \{(x, y) \in \mathbb{R}^2 : x = 0\}$.*

Suma directa de subespacios vectoriales

Sea $V(K)$ un e.v. y $\{U_1, \ldots, U_m\}$ una familia finita de subespacios vectoriales de $V(K)$. Dado $u \in \sum_{i=1}^m U_i$, sabemos que $u = u_1 + \ldots + u_m$ donde $u_i \in U_i$, para cada $i = 1, \ldots, m$. Ahora bien, esta forma de expresar u como suma de vectores en cada sumando U_i no tiene por qué ser única, es decir, podrán existir otros vectores $u'_i \in U_i$, distintos de los anteriores, y tales que $u = u'_1 + \ldots + u'_m$. Veamos un ejemplo de que esto puede efectivamente ocurrir.

Ejemplo 2.56. Sean $U_1 = \{(x,y,z) \in \mathbb{R}^3 : x = 0\}$ y $U_2 = \{(x,y,z) \in \mathbb{R}^3 : z = 0\}$. No es difícil demostrar que $U_1 + U_2 = \mathbb{R}^3$, pues todo vector $v = (x,y,z)$ de $\mathbb{R}^3$ se puede expresar como $u_1 + u_2$, donde $u_1 = (0,y,z) \in U_1$ y $u_2 = (x,0,0) \in U_2$. Veamos que el vector nulo de $\mathbb{R}^3$ se puede expresar de muchas formas distintas como suma de un vector de U_1 y otro de U_2. En efecto, tenemos por ejemplo $(0,0,0) = (0,1,0)+(0,-1,0)$ o $(0,0,0) = (0,2,0)+(0,-2,0)$. De hecho $(0,0,0) = (0,a,0)+(0,-a,0)$, para cada $a \in \mathbb{R}$.

Ahora nos planteamos cómo se puede evitar esta falta de unicidad, es decir, bajo qué condiciones la expresión de $u \in \sum_{i=1}^m U_i$ como suma de vectores $u_i \in U_i$ es única. Veremos más adelante que esta cuestión tiene que ver con descomponer un e.v. en piezas más pequeñas.

Proposición 2.57. *Sea $V(K)$ un e.v. y $\{U_1,\ldots,U_m\}$ una familia finita de subespacios vectoriales de $V(K)$ con $m \geq 2$. Son equivalentes las siguientes afirmaciones:*

(i) Para cada $u \in \sum_{i=1}^m U_i$ existen vectores unívocamente determinados $u_i \in U_i$ tales que $u = u_1 + \ldots + u_m$.

(ii) $(\sum_{i=1}^j U_i) \cap U_{j+1} = \{0\}$, para cada $j = 1,\ldots,m-1$.

Veremos primero cómo se demuestra la proposición en el caso $m = 2$, que será el más frecuente. Así motivaremos también la prueba del caso general.

Demostración en el caso $m = 2$. En este caso, la afirmación (i) significa que, para cada vector $u \in U_1 + U_2$, existen vectores unívocamente determinados $u_1 \in U_1$ y $u_2 \in U_2$ tales que $u = u_1 + u_2$. La afirmación (ii) se reduce a la igualdad $U_1 \cap U_2 = \{0\}$.

Veamos que (i) implica (ii). Sea $u \in U_1 \cap U_2$. Queremos ver que $u = 0$. Nótese que $u = u + 0$ con $u \in U_1$ y $0 \in U_2$. Además, $u = 0 + u$ con $0 \in U_1$ y $u \in U_2$. Como por hipótesis u se expresa de forma única como suma de un vector de U_1 y otro de U_2, entonces $u = 0$.

Veamos que (ii) implica (i). Sea $u \in U_1 + U_2$ y supongamos que hay dos expresiones de u como suma de un vector de U_1 y otro de U_2. Así, existen $u_1, u_1' \in U_1$ y $u_2, u_2' \in U_2$ tales que $u = u_1 + u_2$ y $u = u_1' + u_2'$. Queremos ver que $u_1' = u_1$ y $u_2' = u_2$. Tenemos $u_1 + u_2 = u_1' + u_2'$, que equivale a $u_1' - u_1 = u_2 - u_2'$. Llamemos $v := u_1' - u_1 = u_2 - u_2'$. De estas igualdades, y usando que U_1, U_2 son subespacios vectoriales de V, se sigue que $v \in U_1 \cap U_2$. Como suponemos que $U_1 \cap U_2 = \{0\}$ entonces $v = 0$. Como $v = u_1' - u_1$ deducimos que $u_1 = u_1'$. Y como $v = u_2 - u_2'$ concluimos que $u_2 = u_2'$. ∎

Demostración en el caso general. Veamos que (i) implica (ii). Fijamos $j \in \{1,\ldots,m-1\}$. Queremos ver que $(\sum_{i=1}^j U_i) \cap U_{j+1} = \{0\}$. Tomemos un vector u en dicha intersección. Como $u \in \sum_{i=1}^j U_i$ entonces $u = u_1 + \ldots + u_j$ con $u_i \in U_i$ para cada $i = 1,\ldots,j$. A partir de aquí podemos expresar u de dos formas distintas como suma de vectores de U_i, a saber, $u = u_1 + \ldots + u_j + 0 + u_{j+2} + \ldots + u_m$ y $u = 0 + \ldots + 0 + u + 0 + \ldots + 0$. Por la hipótesis (i) todos los vectores de las expresiones anteriores son nulos. En particular, $u = 0$.

Recíprocamente, veamos que (ii) implica (i) por inducción sobre el número de m de sumandos. El resultado ya ha sido demostrado para $m = 2$ (o, si, se prefiere, resulta trivial para $m = 1$), por lo que supondremos ahora como hipótesis de inducción que resulta cierto para $m-1$. Sea $u \in \sum_{i=1}^m U_i$ y supongamos que podemos expresar u de dos formas como suma de vectores en U_i. Así, para cada $i = 1,\ldots,m$, existen $u_i, u_i' \in U_i$ tales que $u = u_1 + \ldots + u_m$ y $u = u_1' + \ldots + u_m'$. Queremos ver que $u_i' = u_i$ para cada $i = 1,\ldots,m$. Tenemos $u_1 + \ldots + u_m = u_1' + \ldots + u_m'$, que equivale a

$$(u_1 - u_1') + \ldots + (u_{m-1} - u_{m-1}') = u_m' - u_m.$$

Poniendo $v := u'_m - u_m$, como U_m es un s.v. de V, se tiene $v \in U_m$ y, análogaamente $u_i - u'_i \in U_i$ para cada $i = 1,\ldots,m-1$. Por tanto, $v \in \sum_{i=1}^{m-1} U_i$ y aplicando la hipótesis (ii) con $j = m-1$ se sigue $v = 0$. Esto implica que $u'_m = u_m$ y $(u_1 - u'_1) + \ldots + (u_{m-1} - u'_{m-1}) = 0$. Esta última igualdad supone escribir el vector 0 como suma de vectores pertenecientes a U_i con $i = 1,\ldots,m-1$. Podemos ahora aplicar la hipótesis de inducción a $\{U_1,\ldots,U_{m-1}\}$ (obsérvese que como los m subespacios satisfacián la hipótesis (ii), los primeros $m-1$ subespacios también la verifican), y suponer que verifican (i). Escribiendo el vector 0 como $0 + \ldots + 0$, y considerando cada 0 como un vector de U_i para cada $i = 1,\ldots,m-1$, se deduce $u'_i - u_i = 0$ para cada $i = 1,\ldots,m-1$, lo que concluye el resultado. ∎

Observación 2.58. Una consecuencia inmediata de la equivalencia establecida en la proposición anterior es que la condición *(ii)* resulta ser independiente de la ordenación de los subespacios U_i (pues es equivalente a *(i)*, que obviamente resulta independiente de tal ordenación).

Definición 2.59. *Sea $V(K)$ un e.v. y $\{U_1,\ldots,U_m\}$ subespacios vectoriales de V. Diremos que la suma de subespacios $\sum_{i=1}^m U_i$ es directa si se verifican las condiciones alternativas de la proposición anterior (esto es, $(\sum_{i=1}^j U_i) \cap U_{j+1} = \{0\}$ para cada $j = 1,\ldots,m-1$ o, equivalentemente, para cada $u \in \sum_{i=1}^m U_i$, existen vectores únicos $u_i \in U_i$, para $i = 1,\ldots,m$, tales que $u = u_1 + \ldots + u_m$). En este caso, denotaremos al subespacio suma $U_1 \oplus \ldots \oplus U_m$ o, abreviadamente, $\oplus_{i=1}^m U_i$.*

En el caso de que se dé la igualdad $V = \oplus_{i=1}^m U_i$ se dice que V es suma directa de $\{U_1,\ldots,U_m\}$.

En el caso particular $m = 2$, la expresión $U = U_1 \oplus U_2$ significa que $U = U_1 + U_2$ y $U_1 \cap U_2 = \{0\}$. Equivalentemente, para cada $u \in U$ existen vectores únicos $u_1 \in U_1$ y $u_2 \in U_2$ tales que $u = u_1 + u_2$. En tal caso, diremos también que U_2 es un subespacio *complementario* o *suplementario* de U_1 en U.

Ejemplo 2.60. En el Ejemplo 2.56 teníamos una situación en la que $\mathbb{R}^3 = U_1 + U_2$. Sin embargo, no es cierto que $\mathbb{R}^3 = U_1 \oplus U_2$, pues vimos que el vector nulo se expresa de infinitas formas distintas como suma de un vector de U_1 y otro de U_2. Por otro lado, se tiene que $U_1 \cap U_2 \neq \{0\}$, pues $(0,1,0) \in U_1 \cap U_2$.

Ejercicio 2.61. *Sean U_1 y U_2 los subespacios de $\mathbb{R}^2$ del Ejercicio 2.55. Demostrar que $\mathbb{R}^2 = U_1 \oplus U_2$.*

Ejemplo 2.62. [Suma directa de matrices simétricas y antisimétricas] Sea $n \in \mathbb{N}$ y K un cuerpo con característica distinta de 2 ($2 := 1 + 1 \neq 0$, como en $K = \mathbb{R}$, $\mathbb{Q}$ o $\mathbb{C}$). Siguiendo con las propiedades de matrices vistas en el ejemplo 2.38, vamos a demostrar que:

$$M_n(K) = S_n(K) \oplus A_n(K).$$

Para ello hay que comprobar dos propiedades: $M_n(K) = S_n(K) + A_n(K)$ y $S_n(K) \cap A_n(K) = \{0_n\}$.

Veamos que $M_n(K) = S_n(K) + A_n(K)$. La inclusión $S_n(K) + A_n(K) \subset M_n(K)$ es obvia. Probar que $M_n(K) \subset S_n(K) + A_n(K)$ significa mostrar que toda matriz $A \in M_n(K)$ se puede expresar como $B + C$, verificándose $B \in S_n(K)$ y $C \in A_n(K)$. Nótese que, como $2 \neq 0$, entonces A se escribe como:

$$A = \frac{1}{2}(A + A^t) + \frac{1}{2}(A - A^t).$$

En la igualdad anterior el símbolo $1/2$ representa el inverso de 2 en K (que existe al suponer $2 \neq 0$). Si llamamos $B = (1/2)(A + A^t)$ y $C = (1/2)(A - A^t)$, entonces se tiene $A = B + C$. Veamos que $B \in S_n(K)$ y $C \in A_n(K)$ a partir de propiedades de la trasposición de matrices:

$$B^t = \frac{1}{2}(A + A^t)^t = \frac{1}{2}\left(A^t + (A^t)^t\right) = \frac{1}{2}(A^t + A) = B,$$
$$C^t = \frac{1}{2}(A - A^t)^t = \frac{1}{2}\left(A^t - (A^t)^t\right) = \frac{1}{2}(A^t - A) = -C.$$

Por último, comprobemos que $S_n(K) \cap A_n(K) = \{0_n\}$. Sea $A \in S_n(K) \cap A_n(K)$. Nos preguntamos si $A = 0_n$. Como $A \in S_n(K)$, entonces $A^t = A$. Y como $A \in A_n(K)$ tenemos $A^t = -A$. Encadenando ambas igualdades se sigue que $A = -A$ y, por tanto, $A + A = O_n$. Así, $2A = O_n$, y como $2 \neq 0$, llegamos a $A = 0_n$, como se quería.

2.2.4. Subespacios afines

Definición 2.63. *Sea $(\mathcal{A}, \vec{A}, \varphi)$ un espacio afín. Un* subespacio afín *$\mathcal{S}$ es un subconjunto (no vacío) de $\mathcal{A}$ tal que el conjunto*

$$\vec{\mathcal{S}} = \{\varphi(P,Q) = \overline{PQ} : P,Q \in \mathcal{S}\}$$

es un subespacio vectorial de $\mathcal{A}$ y $(\mathcal{S}, \vec{\mathcal{S}}, \varphi_{\mathcal{S}})$ es un espacio afín, donde

$$\varphi_{\mathcal{S}} : \mathcal{S} \times \mathcal{S} \to \vec{\mathcal{S}}, \qquad (P,Q) \mapsto \overline{PQ}$$

se obtiene como restricción de φ.

La siguiente caracterización de los subespacios afines es útil desde el punto de vista práctico.

Proposición 2.64. *Sea $(\mathcal{A}, \vec{A}, \varphi)$ un espacio afín y $\mathcal{S} \subset \mathcal{A}$ un subconjunto no vacío. Equivalen:*
(i) $\mathcal{S}$ es un subespacio afín.
(ii) El conjunto $\vec{\mathcal{S}} = \{\overline{PQ} : P,Q \in \mathcal{S}\}$ es un subespacio vectorial de $\vec{\mathcal{A}}$.
(iii) Existe un $P \in \mathcal{S}$ tal que $\vec{\mathcal{S}}_P := \{\overline{PQ} : Q \in \mathcal{S}\}$ es un subespacio vectorial de $\mathcal{A}$.
En este caso, $\vec{\mathcal{S}}_P = \vec{\mathcal{S}}$ (y, por tanto, $\vec{\mathcal{S}}_P$ es independiente de P).

Demostración. De la propia definición, se tiene (i)$\Rightarrow$(ii). El recíproco (i)$\Leftarrow$(ii) se verifica porque: (1) para cada $O \in \mathcal{S}$, la aplicación $\mathcal{S} \to \vec{\mathcal{S}}$, $Q \mapsto \overline{OQ}$ es biyectiva (es suprayectiva porque para cualquier $\overline{PR}$ con $P,R \in \mathcal{S}$ el punto $O + \overline{PR}$ debe pertenecer a $\mathcal{S}$ y es inyectiva porque si no tampoco lo sería la aplicación $\mathcal{A} \to \vec{\mathcal{A}}$, $Q \mapsto \overline{OQ}$) y (2) para cada $P,Q,R \in \mathcal{S}$, $\overline{PQ} + \overline{QR} = \overline{PR}$ (pues esta propiedad se verificaba para todo $P,Q,R \in \mathcal{A}$).

(ii)$\Rightarrow$(iii) $\vec{\mathcal{S}}$ no es vacío (pues $\overline{PP} = 0 \in \mathcal{S}$) y, para todo $u,v \in \vec{\mathcal{S}}$, $a,b \in K$ se tiene por hipótesis $au + bv \in \vec{\mathcal{S}}$ por lo que $P + (au + bv) \in \mathcal{S}$ y $au + bv \in \vec{\mathcal{S}}_P$.

(iii)$\Rightarrow$(ii) Basta con comprobar que se verifica $\vec{\mathcal{S}}_P = \vec{\mathcal{S}}$. La inclusión $\subset$ es trivial. Para $\supset$, sea $u = \overline{QR} \in \vec{\mathcal{S}}$. Entonces $\overline{PQ}, \overline{PR} \in \vec{\mathcal{S}}_P$. Se tiene entonces $u = \overline{QP} + \overline{PR} = -\overline{PQ} + \overline{PR}$, que pertenece a $\vec{\mathcal{S}}_P$ por ser un subespacio vectorial. ∎

Usando la caracterización (iii) es inmediato demostrar el siguiente resultado.

Proposición 2.65. *Sea $\mathcal{A}$ un espacio afín. Dados $P \in \mathcal{A}$ y un subespacio vectorial $\vec{W}$ de $\vec{\mathcal{A}}$, el conjunto: $\mathcal{W} = P + \vec{W} = \{P + w : w \in \vec{W}\}$ es el único subespacio afín $\mathcal{A}$ que contiene a P y tiene por espacio vectorial asociado $\vec{W}$.*

Ejemplo 2.66. Sabemos que en $K^n(K)$ las soluciones de cualquier SEL homogéneo forman un subespacio vectorial. Aunque lo detallaremos más adelante, es fácil comprobar directamente que las soluciones de cualquier SEL compatible forman un subespacio afín, cuyo espacio vectorial asociado es precisamente el subespacio del correspondiente SEL homogéneo. En particular, las rectas y planos de $\mathbb{R}^3$ (pasen o no por el origen) son subespacios afines de $\mathbb{R}^3$.

Intersección de subespacios afines

A diferencia del caso de los subespacios vectoriales, la intersección de dos subespacios afines $\mathcal{S}_1$ y $\mathcal{S}_2$ puede ser vacía (piénsese en dos rectas paralelas del plano). No obstante, si nos aseguramos de que esta intersección no es vacía, entonces $\mathcal{S}_1 \cap \mathcal{S}_2$ será también un subespacio afín. Esto lo formulamos a continuación para una familia arbitraria de subespacios afines $\mathcal{S}_\alpha$ (con α variando en algún conjunto de índices I).

Proposición 2.67. *Sea $\{\vec{\mathcal{S}}_\alpha : \alpha \in I\}$ un conjunto arbitrario de subespacios afines de $\mathcal{A}$ tal que $\cap_{\alpha\in I}\vec{S}_\alpha \neq \emptyset$. Entonces $\cap_{\alpha\in I}S_\alpha$ es un subespacio afín de espacio vectorial asociado $\cap_{\alpha\in I}\vec{S}_\alpha$.*

Demostración. Escogiendo cualquier $P = \cap_{\alpha\in I}S_\alpha$, es inmediato comprobar: $P + \cap_{\alpha\in I}\vec{S}_\alpha = \cap_{\alpha\in I}S_\alpha$. ∎

A continuación damos una caracterización sencilla de cuándo se intersecan dos subespacios afines.

Proposición 2.68. *Dos subespacios afines $P_1 + \vec{\mathcal{S}}_1$, $P_2 + \vec{\mathcal{S}}_2$ se intersecan si y sólo si $\overline{P_1P_2} \in \vec{\mathcal{S}}_1 + \vec{\mathcal{S}}_2$.*

En este caso, si P pertenece a la intersección $(P_1 + \vec{\mathcal{S}}_1) \cap (P_2 + \vec{\mathcal{S}}_2) = P + (\vec{\mathcal{S}}_1 \cap \vec{\mathcal{S}}_2)$.

Demostración. ($\Rightarrow$) Escogiendo P en la intersección: $\overline{P_1P_2} = \overline{P_1P} + \overline{PP_2} = -\overline{PP_1} + \overline{PP_2} \in \vec{\mathcal{S}}_1 + \vec{\mathcal{S}}_2$.

($\Leftarrow$) Sea $\overline{P_1P_2} = v_1 + v_2$ con $v_1 \in \vec{\mathcal{S}}_1$, $v_2 \in \vec{\mathcal{S}}_2$. Entonces: $P_2 = P_1 + (v_1 + v_2) = (P_1 + v_1) + v_2$ por lo que el punto $P_2 + (-v_2) = P_1 + v_1$ pertenece a la intersección. ∎

Ejercicio 2.69. *Compruébese que, si $\mathcal{S}, \mathcal{T}$ son subespacios afines de A, equivalen:*

1. $\mathcal{S} \cap \mathcal{T} \neq \emptyset$
2. *Para todo $P \in \mathcal{S}, Q \in \mathcal{T}$ se tiene $\overrightarrow{PQ} \in \vec{S} + \vec{T}$.*
3. *Existen $P \in \mathcal{S}, Q \in \mathcal{T}$ tales que $\overrightarrow{PQ} \in \vec{S} + \vec{T}$.*

Suma de subespacios afines

Definición 2.70. *Dado cualquier subconjunto no vacío $C \subset \mathcal{A}$ el* subespacio afín generado por C*, que se denotará $\langle C \rangle$, es el menor subespacio afín que contiene a C, esto es, la intersección de todos los subespacios afines de $\mathcal{A}$ que contienen C. En particular, si $\mathcal{S}$ y $\mathcal{S}'$ son dos subespacios afines, su* suma *$\mathcal{S} + \mathcal{S}'$ se define como el subespacio afín generado por $\mathcal{S} \cup \mathcal{S}'$.*

Nota: El propio espacio afín $\mathcal{A}$ es siempre un subespacio afín que contiene a C. Así, la proposición 2.67 asegura que, efectivamente, el *subespacio afín generado por C* tiene estructura de subespacio afín.

Proposición 2.71. *Dados dos subespacios afines $\mathcal{S}_1 = P_1 + \vec{\mathcal{S}}_1$, $\mathcal{S}_2 = P_2 + \vec{\mathcal{S}}_2$, su suma $\mathcal{S}$ tiene por subespacio director $L(\overline{P_1P_2}) + \vec{\mathcal{S}}_1 + \vec{\mathcal{S}}_2$.*

Demostración. Fácilmente se tiene $\mathcal{S}_1, \mathcal{S}_2 \subset P_1 + (L(\overline{P_1P_2}) + \vec{\mathcal{S}}_1 + \vec{\mathcal{S}}_2)$ por lo que, teniendo en cuenta la definición de subespacio afín suma, basta con comprobar:

$$P_1 + (L(\overline{P_1P_2}) + \vec{\mathcal{S}}_1 + \vec{\mathcal{S}}_2) \subset \mathcal{S} \qquad \text{y como } P_1 \in \mathcal{S} \text{ basta con:} \qquad L(\overline{P_1P_2}) + \vec{\mathcal{S}}_1 + \vec{\mathcal{S}}_2 \subset \vec{\mathcal{S}}.$$

De $\mathcal{S} = P_1 + \vec{\mathcal{S}}$ y $\mathcal{S}_1 \subset \mathcal{S}$ se sigue $\vec{\mathcal{S}}_1 \subset \vec{\mathcal{S}}$ y de $P_2 \in \mathcal{S}$ se sigue $\overline{P_1P_2} \subset \vec{\mathcal{S}}$. Razonando análogamente con P_2 se sigue $\vec{\mathcal{S}}_2 \subset \vec{\mathcal{S}}$. Por tanto, al ser $\vec{\mathcal{S}}$ un subespacio vectorial, $L(\overline{P_1P_2}) + \vec{\mathcal{S}}_1 + \vec{\mathcal{S}}_2 \subset \vec{\mathcal{S}}$. ∎

2.3. Bases, dimensión y coordenadas en un espacio vectorial

Una de las ideas principales en la teoría de espacios vectoriales consiste en expresar todos los vectores a partir de "unos pocos" mediante combinaciones lineales. El alumno ya debe de tener alguna familiaridad con esta idea, que aparece en la enseñanza preuniversitaria y cuando se trabaja con vectores libres. Por ejemplo, si en el e.v. real $\mathbb{R}^2$ (o en los vectores libres del plano) fijamos un "sistema de referencia" formado por dos vectores no colineales $\{e_1, e_2\}$ entonces cualquier otro vector v se expresa como $a \cdot e_1 + b \cdot e_2$ para ciertos $a, b \in \mathbb{R}$ (para ello, basta proyectar v sobre las rectas vectoriales $L(e_1)$ y $L(e_2)$). Esto significa que $\mathbb{R}^2 = L(e_1, e_2)$, es decir, con tan sólo dos vectores de $\mathbb{R}^2$ podemos generar todos los demás haciendo combinaciones lineales. Además, la forma de expresar v como c.l. de e_1 y e_2 es única, esto es, se tiene $\mathbb{R}^2 = L(e_1) \oplus L(e_2)$. Así, los escalares $a, b \in \mathbb{R}$ tales que $v = a \cdot e_1 + b \cdot e_2$ son únicos y se llaman coordenadas de v en el sistema de referencia. Con argumentos análogos, en el e.v. real $\mathbb{R}^3$ se tiene $\mathbb{R}^3 = L(e_1, e_2, e_3)$, donde $\{e_1, e_2, e_3\}$ es el sistema de referencia formado por cualesquiera tres vectores no coplanarios. Además, la forma de expresar $v \in \mathbb{R}^3$ como c.l. de $\{e_1, e_2, e_3\}$ es única, es, decir $\mathbb{R}^3 = L(e_1) \oplus L(e_2) \oplus L(e_3)$. Así, se puede hablar como antes de coordenadas asociadas a un vector de $\mathbb{R}^3$ respecto del sistema de referencia dado. En este apartado del tema desarrollaremos todas estas ideas en cualesquiera espacios vectoriales y afines.

2.3.1. Sistemas de generadores

Comenzaremos precisando la idea de generar todos los vectores de un e.v. a partir de unos pocos.

Definición 2.72. *Sea $V(K)$ y sea $S \subset V$ un subconjunto. Se dice que S es un* sistema de generadores *(s.d.g.) o* conjunto generador *de V si $V = L(S)$. En el caso de que S no sea vacío, esto equivale a que todo vector de V se expresa como c.l. (¡finita!) de vectores de S, es decir, para cada $v \in V$, existen $m \in \mathbb{N}$, vectores $v_1, \ldots, v_m \in S$ y escalares $a_1, \ldots, a_m \in K$, tales que $v = a_1 \cdot v_1 + \ldots + a_m \cdot v_m$.*

Es obvio que $V = L(V)$ y, por tanto, V es un s.d.g. de V. Esto no es muy interesante si nuestro objetivo es generar todos los vectores de V (probablemente infinitos) a partir de la menor cantidad posible de vectores. En particular, y como primera aproximación, nos interesarán los e.v. que tengan un s.d.g. finito.

Definición 2.73. *Sea V un e.v. sobre un cuerpo K. Diremos que V es* finitamente generado *(f.g.) si V admite un s.d.g. con un número finito de vectores, esto es, si existe $S = \{v_1, \ldots, v_m\} \subset V$ tal que $V = L(S)$ (por lo que para cada $v \in V$, existen $a_1, \ldots, a_m \in K$ tales que $v = a_1 \cdot v_1 + \ldots + a_m \cdot v_m$).*

Practiquemos un poco las definiciones anteriores con ejemplos concretos.

Ejemplo 2.74. Nos preguntamos si en $\mathbb{R}^2$ el conjunto $S = \{v_1, v_2\}$ con $v_1 = (1,0)$ y $v_2 = (1,1)$ es un s.d.g. Para ello tomamos un vector $v = (x,y) \in \mathbb{R}^2$ y estudiamos si es c.l. de S. Buscamos $a, b \in \mathbb{R}$ tales que $v = a \cdot v_1 + b \cdot v_2$. Al tomar componentes y operar esta igualdad se transforma en $(x,y) = (a+b, b)$. Por tanto, llegamos al SEL dado por las ecuaciones $a + b = x$ y $b = y$, que es compatible determinado con soluciones $b = y$, $a = x - y$. Esto prueba que S es un s.d.g. de $\mathbb{R}^2$. En particular, $\mathbb{R}^2$ es f.g. como e.v. real. En el caso particular $v = (-2,3)$ se tiene que $a = -5$ y $b = 3$, por lo que $v = -5 \cdot v_1 + 3 \cdot v_2$.

Ejemplo 2.75. Nos preguntamos si en $\mathbb{R}^2$ el conjunto $S = \{v_1, v_2, v_3\}$ con $v_1 = (1,2)$, $v_2 = (2,4)$ y $v_3 = (3,6)$ es un s.d.g. Para ello tomamos $v = (x,y) \in \mathbb{R}^2$ y estudiamos si es c.l. de S. Buscamos $a, b, c \in \mathbb{R}$ tales que $v = a \cdot v_1 + b \cdot v_2 + c \cdot v_3$. Tomando componentes y operando, esto equivale a que $(x,y) = (a + 2b + 3c, 2a + 4b + 6c)$. Llegamos así a un SEL con ecuaciones $a + 2b + 3c = x$ y

$2a+4b+6c=y$. Obviamente este SEL será compatible si y sólo si $2x=y$. Por tanto, si $v=(x,y)\in\mathbb{R}^2$ no cumple que $2x=y$, entonces v no es c.l. de S. Esto ocurre por ejemplo con el vector $v=(0,1)$. Concluimos que S no es un s.d.g. de $\mathbb{R}^2$.

Ahora mostraremos algunos s.d.g. para los espacios vectoriales que estudiamos en la primera sección de este tema.

Ejemplo 2.76. Sea $n\in\mathbb{N}$ y K un cuerpo. Consideremos K^n como e.v. sobre K. Para cada $i\in\{1,\dots,n\}$ definimos $e_i=(0,0,\dots,1,\dots,0,0)$, donde el 1 se encuentra en la i-ésima posición. Afirmamos que el conjunto $S=\{e_1,\dots,e_n\}$ es un s.d.g. de K^n. Esto se debe a que todo vector $v=(x_1,\dots,x_i,\dots,x_n)\in K^n$ se escribe como $x_1\cdot e_1+\dots+x_i\cdot e_i+\dots+x_n\cdot e_n$, que es una c.l. de S (nótese que los coeficientes de la combinación coinciden con las componentes del vector). En particular, K^n es f.g. como e.v. sobre K.

Ejemplo 2.77. Sean $m,n\in\mathbb{N}$ y K un cuerpo. Consideremos $M_{m\times n}(K)$ como e.v. sobre K. Para cada par de índices $i\in\{1,\dots,m\}$ y $j\in\{1,\dots,n\}$, definimos $E_{ij}=(e_{kl})$ como la matriz en $M_{m\times n}(K)$ tal que $e_{kl}=0$ si $(k,l)\neq(i,j)$ y[11] $e_{ij}=1$. Afirmamos que el conjunto $S=\{E_{ij}\,|\,i=1,\dots,m,\ j=1,\dots,n\}$ es un s.d.g. de $M_{m\times n}(K)$ como e.v. sobre K. Esto se debe a que toda matriz $A=(a_{ij})$ en $M_{m\times n}(K)$ se expresa como:

$$A=\sum_{i=1}^{m}\sum_{j=1}^{n}a_{ij}\cdot E_{ij}=a_{11}\cdot E_{11}+\dots+a_{1n}\cdot E_{1n}+\dots+a_{m1}E_{m1}+\dots+a_{mn}E_{mn},$$

que es una c.l. de S (nótese que los coeficientes de la combinación coinciden con las entradas de A). En particular, $M_{m\times n}(K)$ es f.g. como e.v. sobre K.

Ejemplo 2.78. Sea $\mathbb{K}=\mathbb{R}$ ó $\mathbb{C}$ un cuerpo y $\mathbb{K}[x]$ el e.v. sobre $\mathbb{K}$ de los polinomios con coeficientes en $\mathbb{K}$. Para cada $i\in\mathbb{N}\cup\{0\}$ definimos $p_i(x)=x^i$, entendiendo que $p_0(x)=1$. Afirmamos que el conjunto $S=\{p_i(x)\,|\,i\in\mathbb{N}\cup\{0\}\}$ es un s.d.g. de $\mathbb{K}[x]$. Esto se debe a que todo polinomio $p(x)=a_0+a_1x+\dots+a_nx^n$ se escribe como $a_0\cdot p_0(x)+a_1\cdot p_1(x)+\dots+a_n\cdot p_n(x)$, que es una c.l. finita de elementos de S. Nótese que S es un conjunto infinito numerable.

Lo anterior no demuestra que $\mathbb{K}[x]$ no es f.g., pues pudiera existir algún s.d.g. de $\mathbb{K}[x]$ que fuera finito. Veamos que esto es imposible. Supongamos que $S=\{p_1(x),\dots,p_m(x)\}$ fuese un s.d.g. de $\mathbb{K}[x]$. Para cada $i\in\{1,\dots,m\}$ denotamos $n_i=\text{grado}(p_i(x))$. Sea $n=\text{máx}\{n_1,\dots,n_m\}$. Es claro entonces que $S\subset\mathbb{K}_n[x]$ y, por tanto, $L(S)\subset\mathbb{K}_n[x]$. De este modo, el polinomio en $\mathbb{K}[x]$ dado por $p(x)=x^{n+1}$ no se expresa como c.l. de[12] S; de hecho, $p(x)\notin\mathbb{K}_n[x]$. Esto es una contradicción, por lo que se sigue que $\mathbb{K}[x]$ no es f.g. como e.v. sobre K.

Ejemplo 2.79. En el e.v. real $F(\mathbb{R},\mathbb{R})$ de las funciones $f:\mathbb{R}\to\mathbb{R}$ consideramos el subconjunto U de las funciones dos veces derivables y tales que $f''(x)+f(x)=0$ para cada $x\in\mathbb{R}$. No es difícil demostrar, usando la Proposición 2.25 y reglas de derivación conocidas de la enseñanza secundaria, que U es un s.v. de $F(\mathbb{R},\mathbb{R})$. El estudio de ecuaciones diferenciales[13] demuestra que, si $f\in U$, entonces existen $a,b\in\mathbb{R}$ tales que $f(x)=a\cdot\text{sen}(x)+b\cdot\cos(x)$, para cada $x\in\mathbb{R}$. Recíprocamente, es fácil comprobar que cada función f de este tipo pertenece a U. Así, deducimos que $U=L(\text{sen}(x),\cos(x))$. En particular, U es f.g. como e.v. real.

Ahora mostraremos la influencia de K cuando hablamos de s.d.g. y de espacios f.g.

[11] Usando la delta de Kronecker las matrices E_{ij} se reescriben entonces $(E_{ij})_{kl}=\delta_{ik}\cdot\delta_{jl}$.

[12] Una forma de justificar esto considerando los polinomios como aplicaciones polinómicas con coeficientes en $\mathbb{K}$ sería comprobar que al derivar $n+1$ veces $p(x)$ se obtiene $(n+1)!\neq 0$, mientras que al hacerlo con c.l. se S se obtiene 0.

[13] Asignatura Ecuaciones Diferenciales I.

Ejemplo 2.80. 1. Sea $\mathbb{C}$ el cuerpo de los números complejos. Sabemos que $\mathbb{C}$ es un e.v. complejo y también un e.v. real. Teniendo en cuenta el Ejemplo 2.76 con $K = \mathbb{C}$ y $n = 1$ se sigue que $S = \{1\}$ es un s.d.g. de $\mathbb{C}$ como e.v. complejo. Sin embargo, $S = \{1\}$ no es un s.d.g. de $\mathbb{C}$ como e.v. real, ya que ningún número imaginario puro se expresa como $a \cdot 1$ con $a \in \mathbb{R}$. De hecho, $L(S) = \mathbb{R}$ cuando vemos $\mathbb{C}$ como e.v. real.

2. Sabemos que $\mathbb{R}$ es f.g. como e.v. real. También podemos ver $\mathbb{R}$ como e.v. sobre $K = \mathbb{Q}$ (la suma sería la de números reales y el producto de un racional por un real sería el producto usual en $\mathbb{R}$). En un ejercicio de la relación de problemas se propone probar que $\mathbb{R}$ no es f.g. como e.v. sobre $\mathbb{Q}$.

Ya hemos comentado que nuestro objetivo es encontrar s.d.g. de un e.v. que sean "lo más pequeños posible". Esto nos conduce al siguiente problema.

Cuestión 2.81. Si V es un e.v. sobre K y V es f.g., ¿cuál es el menor número de vectores que puede tener un s.d.g. de V?

Nótese que el número mínimo de vectores de un s.d.g. es una medida del "tamaño" del e.v. pues es intuitivo que, cuantos más vectores sean estrictamente necesarios para generarlo, más grande será el espacio. Responderemos a la Cuestión 2.81 a lo largo de esta sección (observación 2.113). De momento nos podemos plantear cuándo es posible suprimir algún vector de un s.d.g. de forma que el resultado siga siendo un s.d.g. con menos vectores. Es obvio que esto no se puede hacer en general (piénsese por ejemplo en el s.d.g. de $\mathbb{R}^2$ dado por dos vectores no colineales $\{e_1, e_2\}$). El próximo resultado nos dice que un s.d.g. se puede "refinar" siempre que contenga vectores "que no aporten información". Probaremos este principio en el caso finito (el caso infinito queda como ejercicio).

Proposición 2.82 (Ampliación y reducción de s.d.g.)**.** *Sea V un e.v. sobre un cuerpo K y $S = \{v_1, \ldots, v_m\}$ un s.d.g. de V.*

(i) Si $S' \subset V$ y $S \subset S'$, entonces S' es un s.d.g. de V.

(ii) Si existe[14] $i \in \{1, \ldots, m\}$ tal que $v_i \in L(S - \{v_i\})$, entonces $S - \{v_i\} = \{v_1, \ldots, v_{i-1}, v_{i+1}, \ldots, v_m\}$ es un s.d.g. de V.

Demostración. (i) Como S es s.d.g. de V entonces $V = L(S)$. Nótese que $L(S')$ es un s.v. de V que contiene a S' y, por tanto, a S. Esto implica que $L(S) \subset L(S')$ por la Proposición 2.43. De aquí se sigue que $V \subset L(S')$. Y como $L(S') \subset V$ llegamos a $L(S') = V$.

(ii). Sea $v \in V$. Queremos ver que v se expresa como c.l. de $S - \{v_i\}$. Por un lado, como S es un s.d.g. de V, se tiene que:

$$v = a_1 \cdot v_1 + \ldots + a_{i-1} \cdot v_{i-1} + a_i \cdot v_i + a_{i+1} \cdot v_{i+1} + \ldots + a_m \cdot v_m.$$

Por otro lado, como $v_i \in L(S - \{v_i\})$, entonces:

$$v_i = b_1 \cdot v_1 + \ldots + b_{i-1} \cdot v_{i-1} + b_{i+1} \cdot v_{i+1} + \ldots + b_m \cdot v_m.$$

[14]Obsérvese que incluso en el caso $V = \{0\}, S = \{0\}$ el convenio $L(\emptyset) = \{0\}$ explicado en la Observación 2.45 es consistente aquí.

Sustituyendo la igualdad anterior en la expresión de v como c.l. de S, y usando propiedades de un e.v., llegamos a:

$$\begin{aligned} v = & a_1 \cdot v_1 + \ldots + a_{i-1} \cdot v_{i-1} \\ & + a_i \cdot (b_1 \cdot v_1 + \ldots + b_{i-1} \cdot v_{i-1} + b_{i+1} \cdot v_{i+1} + \ldots + b_m \cdot v_m) \\ & + a_{i+1} \cdot v_{i+1} + \ldots + a_m \cdot v_m \\ & = (a_1 + a_i \cdot b_1) \cdot v_1 + \ldots + (a_{i-1} + a_i \cdot b_{i-1}) \cdot v_{i-1} \\ & + (a_{i+1} + a_i \cdot b_{i+1}) \cdot v_{i+1} + \ldots + (a_m + a_i \cdot b_m) \cdot v_m, \end{aligned}$$

lo que nos dice que $v \in L(S - \{v_i\})$, como se quería demostrar. ∎

Ejercicio 2.83. *Compruébese que la proposición anterior se extiende al caso de que S sea infinito, obteniéndose así:*

Si $S \subset V$ es un s.d.g. se verifica (i) todo $S' \subset V$ que incluya S es un s.d.g., (ii) si $v \in S$ y $v \in L(S - \{v\})$ entonces $S - \{v\}$ es un s.d.g.

2.3.2. Conjuntos linealmente independientes

La Proposición 2.82 (ii) implica que si un s.d.g. contiene vectores que se escriben como c.l. de los demás generadores, entonces dichos vectores se pueden eliminar para obtener un s.d.g. más pequeño. Así, un s.d.g. será "irreducible" y, por tanto, tendrá el menor número posible de vectores cuando ningún generador sea c.l. del resto. Esto nos llevará a la noción de sistema linealmente independiente, pero antes pero antes la caracterizaremos como sigue.

Lema 2.84. *Sea $S = \{v_1, \ldots, v_m\}, m \in \mathbb{N}$ un subconjunto (finito no vacío) de V. Equivalen:*

(i) Ningún vector $v_i \in S$ verifica $v_i \in L(S \setminus \{v_i\})$ (esto es, en el caso $m \geq 2$ ningún vector de S se puede escribir como combinación lineal del resto de vectores de S y, en el caso $m = 1$, $v_1 \neq 0$).

(ii) La única combinación lineal de elementos de S igual a 0 es la que se obtiene con todos sus escalares nulos, esto es: dados $a_1, \ldots a_m \in K$ si

$$a_1 v_1 + \cdots + a_m v_m = 0$$

entonces

$$a_1 = \cdots = a_m = 0.$$

Demostración. (i) ⇒ (ii). Razonando por el contrarrecíproco, existen escalares $a_1, \ldots, a_m \in K$ no todos nulos tales que $a_1 \cdot v_1 + \ldots + a_m \cdot v_m = 0$. Sea $i \in \{1, \ldots, m\}$ tal que $a_i \neq 0$. Si en la igualdad anterior despejamos $a_i \cdot v_i$ tenemos:

$$a_i \cdot v_i = (-a_1) \cdot v_1 + \ldots + (-a_{i-1}) \cdot v_{i-1} + (-a_{i+1}) \cdot v_{i+1} + \ldots + (-a_m) \cdot v_m.$$

Como K es un cuerpo y $a_i \neq 0$ podemos multiplicar por a^{-1} y usar las propiedades de $\cdot$ para obtener:

$$v_i = (-a_i^{-1} \cdot a_1) \cdot v_1 + \ldots + (-a_i^{-1} \cdot a_{i-1}) \cdot v_{i-1} + (-a_i^{-1} \cdot a_{i+1}) \cdot v_{i+1} + \ldots + (-a_i^{-1} \cdot a_m) \cdot v_m,$$

por lo que $v_i \in L(S - \{v_i\})$ (obsérvese que en el caso $m = 1$ se tiene necesariamente $i = 1$ y se obtiene directamente $v_1 = 0$, por lo que $v_1 \in \{0\} = L(\emptyset) = L(S \setminus v_1)$).

(ii) ⇒ (i). Supongamos por el contrarrecíproco que existe $v_i \in S$ tal que $v_i \in L(S \setminus \{v_i\})$. Si $m = 1$ eso quiere decir $v_1 \in L(\emptyset) = \{0\}$, esto es, $v_1 = 0$, por lo que la combinación lineal $1 \cdot v_1 = 0$ contradice

(ii), como se quería. Si $m \geq 2$, existe $i \in \{1,\ldots,m\}$ tal que $v_i \in L(S - \{v_i\})$, por lo que podemos escribir:

$$v_i = a_1 \cdot v_1 + \ldots + a_{i-1} \cdot v_{i-1} + a_{i+1} \cdot v_{i+1} + \ldots + a_m \cdot v_m.$$

Sumando el opuesto de v_i en ambos lados de la igualdad llegamos a:

$$a_1 \cdot v_1 + \ldots + a_{i-1} \cdot v_{i-1} + (-1) \cdot v_i + a_{i+1} \cdot v_{i+1} + \ldots + a_m \cdot v_m = 0,$$

que es una expresión del vector nulo como c.l. de S con (al menos) el escalar (-1), que multiplica a v_i, distinto de 0, obteniéndose una contradicción con (ii). ▮

Observación 2.85. Obsérvese que si S verifica la propiedad (ii) entonces cualquier subconjunto $S' \subset S$ también la verifica. De hecho, reordenando los elementos de S para que los m' primeros sean los de S' (esto es, suponiendo $S' = \{v_1,\ldots,v_{m'}\}$ sin pérdida de generalidad) se tiene

$$a_1 \cdot v_1 + \ldots a_{m'} \cdot v_{m'} \quad = \quad a_1 \cdot v_1 + \cdots + a_{m'} \cdot v_{m'} + 0 \cdot v_{m'+1} + \cdots + 0 \cdot v_m \,.$$

Por tanto, si el primer miembro de esta expresión fuera igual a 0 entonces también lo sería el segundo y, por la independencia lineal de S, se obtendría $a_1 = \cdots = a_{m'} = 0$.

La observación y lema anteriores motivan la siguiente definición.

Definición 2.86. *Sea $V(K)$ un s.v. y $S \subset V$ cualquier subconjunto. Diremos que S es* linealmente independiente *(l.i.) si las únicas combinaciones lineales de elementos de S que son iguales a 0 son las que se obtienen con todos sus escalares nulos, esto es: para todo $\{v_1,\ldots,v_m\} \subset S$, y $a_1,\ldots a_m \in K$:*

$$a_1 \cdot v_1 + \cdots + a_m \cdot v_m = 0 \qquad \Rightarrow \qquad a_1 = \ldots a_m = 0. \tag{2.7}$$

En caso contrario, diremos que S es linealmente dependiente.

Observación 2.87. Podemos distinguir entonces los siguientes casos:

- Como caso límite, si $S = \emptyset$ entonces es l.i.
- Si S es finito no vacío, basta con comprobar la propiedad (2.7) para el caso en que $\{v_1,\ldots,v_m\} = S$ (por la observación 2.85 no hace falta tener en cuenta sus subconjuntos). Por tanto, la independencia lineal resulta equivalente a la propiedad (i) del lema 2.84.
- Si S es infinito, S será l.i. si y sólo si todo subconjunto finito S' de S es l.i. Como para S' la caracterización (i) del lema 2.84 es aplicable, esta caracterización sigue siendo válida incluso cuando S es infinito.

Como resumen, de las observaciones anteriores y de la definición de independencia lineal se tiene la siguiente extensión del lema 2.84 al caso S vacío y S infinito (¡compruébese!):

Teorema 2.88. *Sea $S \subset V$ cualquier subconjunto. Equivalen:*

(i) Ningún vector $v_i \in S$ verifica $v_i \in L(S \setminus \{v_i\})$.

(ii) S es linealmente independiente (según la definición 2.86).

En este caso, cualquier subconjunto $S' \subset S$ también es linealmente independiente.

Ejemplo 2.89. Consideremos algunos casos particulares de S:

1. En el caso de que S tenga un único vector, $S = \{v\}$, se sigue: S es l.i. si y sólo si $v \neq 0$.

   En efecto, sabemos por las propiedades de e.v. que la igualdad $a \cdot v = 0$ ocurre si y sólo si $a = 0$ ó $v = 0$. Así, si $v \neq 0$ entonces la igualdad sólo se da si $a = 0$, por lo que $\{v\}$ es l.i. Recíprocamente, si $v = 0$ la igualdad se da para todo $a \in K$, por lo que $\{0\}$ no es l.i.

2. Si S contiene al vector nulo, entonces será linealmente dependiente.

   En efecto, sabemos que $S' = \{0\}$ es linealmente dependiente y, por hipótesis $S' \subset S$, por lo que la última afirmación del teorema 2.88 implica la dependencia lineal de S (como ejercicio, demuéstrese directamente usando sólo la definición de independencia lineal).

3. Supongamos ahora $S = \{u, v\}$. Que S sea l.d. equivale a que u es proporcional a v o v es proporcional a u. Que u sea *proporcional* a v significa que $u = a \cdot v$ para cierto $a \in K$. Si u es proporcional a v y $u \neq 0$ entonces $v \neq 0$, $a \neq 0$ y $v = a^{-1} \cdot u$, es decir, v es proporcional a u. Sin embargo, la relación de proporcionalidad entre vectores no es simétrica en general; de hecho, el vector 0 es proporcional a todos los demás pero el único vector proporcional a 0 es 0. Por lo anterior, $S = \{u, v\}$ es l.i. si y sólo si $u \neq 0$, $v \neq 0$ y u no es proporcional a v. En tal caso, llamaremos a $L(u, v)$ el *plano (vectorial)* generado por u y v.

4. En el caso de los vectores libres del plano (nota 2.22), el punto anterior se suele resumir diciendo: $\{\vec{u}, \vec{v}\}$ es l.i. si y sólo $\vec{u}, \vec{v}$ no son colineales. Para una cantidad mayor de vectores, el concepto de dependencia lineal generaliza al de colinealidad. Así, para los vectores libres del espacio, el conjunto $S = \{\vec{u}, \vec{v}, \vec{w}\}$ en es l.d. si y sólo si los vectores de S son *coplanarios*, es decir, existe un plano vectorial que los contiene a todos.

Convenio 2.90. En adelante se considerará la independencia lineal permitiendo el caso de que en los conjuntos haya dos "elementos repetidos". Así, si, por ejemplo $S = \{v_1, \ldots v_i, \ldots, v_j, \ldots, v_m\}$ cumple $i < j$ y $v_i = v_j$ se entendería que ese conjunto es l.d. a causa de la combinación lineal[15] $1 \cdot v_i + (-1) \cdot v_j = 0$. En nuestra definición formal de conjunto esto no podría ser. No obstante, también podemos formalizar la idea subyacente a este caso con el sguiente convenio:

(a) Cuando se toma una combinación lineal de un conjunto finito S, se entiende implícitamente que este conjunto es un conjunto ordenado, esto es, que **S es una n-úpla de vectores**.

(b) Se definen las combinaciones lineales para para cualquier n-úpla de vectores tomando combinaciónes lineales de los elementos que la forman con el convenio de que, en el caso de que alguno de ellos esté repetido k veces, se permiten k repeticiones de ese vector en cada expresión de la correspondiente combinación lineal.

Así, en el ejemplo inicial, el conjunto S debe entenderse como la n-úpla $S = (v_1, \ldots v_i = v, \ldots, v_j = v, \ldots, v_m)$, que es linealmente dependiente por tener dos vectores repetidos.

Una propiedad relevante de los conjuntos l.i. es la siguiente.

Proposición 2.91. *Sea $S \subset V$ un conjunto l.i. Entonces todo $v \in L(S)$ se escribe de manera única como combinación lineal de los elementos de S, esto es:*

$$v = a_1 \cdot v_1 + \cdots + a_m \cdot v_m = a'_1 \cdot v_1 + \cdots + a'_m \cdot v_m \qquad \Rightarrow \qquad a_1 = a'_1, \ldots, a_m = a'_m. \tag{2.8}$$

[15] Esto es consistente con nuestro criterio sobre columnas linealmente independientes en una matriz en la nota a pie 32.

Demostración. Operando en el espacio vectorial se tiene de la expresión a la izquierda:

$$(a_1 - a'_1) \cdot v_1 + \cdots + (a_m - a'_m) \cdot v_m = 0,$$

que es una combinación lineal de elementos de S igual a 0. Al ser S l.i. $a_1 - a'_1 = \cdots = a_m - a'_m = 0$. ∎

Observación 2.92. (a) El recíproco de la proposición anterior es inmediato. De hecho, que el vector 0 se escriba como combinación lineal única de los elementos de S, es precisamente la definición de ser linealmente independiente.

(b) Merece la pena darse cuenta de que le expresión (2.8) recoge el significado de que todo vector de $L(S)$ se escribe como combinación lineal *única* de los elementos de S. Esto resulta claro en el caso de que S sea finito (teniendo en cuenta de que en el caso $S = \emptyset$ no hay nada que comprobar). En el caso de que S sea infinito, que se pueda escribir v como combinación lineal de dos maneras de elementos de S quiere decir que existen dos subconjuntos finitos S_1 y S_2 tales que $v \in L(S_1)$ y $v \in L(S_2)$. Ahora bien, $S' := S_1 \cup S_2$ es finito y tanto $L(S_1)$ como $L(S_2)$ están incluidos en $L(S')$. Por tanto, v se podría escribir como combinación lineal de elementos de S' de dos maneras. No obstante, la expresión (2.8) dice que esas dos maneras son la misma (y, de hecho, $v \in L(S_1 \cap S_2)$).

Resulta intuitivo pensar que el número máximo de vectores que puede tener una un conjunto l.i. es una medida del "tamaño" del e.v., pues indica en cuántas direcciones podemos movernos libremente dentro del espacio. Esto nos lleva a un problema similar al que planteamos para s.d.g. (cuestión 2.81).

Cuestión 2.93. Si V es un e.v. sobre K y V es f.g., ¿cuál es el mayor número de vectores que un conjunto l.i. de V puede tener?

De momento, damos una primera respuesta sobre cuándo podemos ampliar una familia l.i.

Proposición 2.94 (Ampliación de conjuntos l.i.)**.** *Sea $V(K)$ un e.v. y $S \subset V$ un conjunto l.i. Si $v \in V$ verifica $v \notin L(S)$, entonces el conjunto $S \cup \{v\}$ es también l.i.*

Demostración. Escribamos 0 como una combinación lineal finita de elementos de $S \cup \{v\}$. Si v no es uno de los vectores de la combinación, todos los coeficientes serán nulos por ser S independente. En consecuencia, basta con tomar cualquier c.l. finita del tipo $a_1 \cdot v_1 + \ldots + a_m \cdot v_m + a_{m+1} \cdot v = 0$, con $a_1, \ldots, a_{m+1} \in K$, $v_1, \ldots, v_m \in S$. Queremos ver que $a_i = 0$ para cada $i = 1, \ldots, m+1$. Como S es l.i. la demostración concluye si vemos que $a_{m+1} = 0$ (ya que la c.l. anterior nos daría una expresión del vector nulo como c.l. de vectores de S). Ahora bien, si $a_{m+1} \neq 0$ podríamos despejar v en la ecuación de arriba y tendríamos:

$$v = (-a_{m+1}^{-1} \cdot a_1) \cdot v_1 + \ldots + (-a_{m+1}^{-1} \cdot a_m) \cdot v_m.$$

Esto implicaría que $v \in L(S)$, lo que supone una contradicción. ∎

Observación 2.95. Combinando este resultado con la última afirmación del teorema 2.88 y las propiedades de los sistema de generadores se obtiene, para cualesquiera $S \subset S' \subset V$ y $v \in V$:

- Si S es un s.d.g. entonces S' es un s.d.g.

  Si S es l.i. y $S' = S \cup \{v\}$ entonces: S' es l.i. si y sólo si $v \notin L(S)$.

- Si S' es l.i. entonces S es l.i.

  Si S' es s.d.g. y $S' = S \cup \{v\}$ entonces: S es s.d.g. si y sólo si $v \in L(S)$.

Practiquemos a continuación las caracterizaciones del concepto de independencia lineal en algunos ejemplos concretos.

Ejemplo 2.96. En $\mathbb{R}^2$ el conjunto $S = \{v_1, v_2, v_3\}$ con $v_1 = (1,0)$, $v_2 = (1,1)$ y $v_3 = (0,1)$ es l.d. ya que $v_2 = v_1 + v_3$. En particular, el teorema 2.88 implica la existencia de una c.l. del tipo $a \cdot v_1 + b \cdot v_2 + c \cdot v_3 = 0$ con no todos los coeficientes nulos. Podemos calcularlas explícitamente: tomando componentes y operando, la igualdad anterior se transforma en $(a+b, b+c) = (0,0)$. Por tanto, llegamos al SEL homogéneo de ecuaciones $a+b=0$ y $b+c=0$. Este SEL es compatible indeterminado (está escalonado y tiene incógnitas secundarias), por lo que tiene infinitas solucioes. Así, para encontrar una expresión no trivial de 0 como c.l. de S basta encontrar una solución no trivial del SEL anterior, por ejemplo $a=1$, $b=-1$ y $c=1$. Se comprueba enseguida que, efectivamente, $v_1 - v_2 + v_3 = 0$. Posteriormente demostraremos que en $\mathbb{R}^2$ no puede haber conjuntos l.i. con más de dos vectores.

Ejemplo 2.97. Estudiemos si en $\mathbb{R}[x]$ el conjunto $S = \{p_1(x), p_2(x), p_3(x)\}$ donde $p_1(x) = x^2+x+1$, $p_2(x) = 2x+1$ y $p_3(x) = x^2+1$ es l.i. Dada una c.l. del tipo $a \cdot p_1(x) + b \cdot p_2(x) + c \cdot p_3(x) = 0$ nos preguntamos si $a=b=c=0$. Operando, la igualdad anterior se transforma en $(a+c)x^2 + (a+2b)x + (a+b+c) = 0$. Esto nos lleva al SEL homogéneo de ecuaciones[16] $a+c=0$, $a+2b=0$ y $a+b+c=0$. Este SEL es compatible determinado y, por tanto, su única solución es $a=b=c=0$. Concluimos que la familia S es l.i.

Comprobemos a continuación que varios s.d.g. importantes que estudiamos en la anterior subsección son l.i. en el espacio correspondiente.

Ejemplo 2.98. *Caso de K^n.* En K^n el conjunto $S = \{e_1, \ldots, e_n\}$ del Ejemplo 2.76 es l.i. Para verlo tomamos una c.l. del tipo $a_1 \cdot e_1 + \ldots + a_n \cdot e_n = 0$ y nos preguntamos si $a_i = 0$ para cada $i = 1, \ldots, n$. Al tomar componentes y operar, la igualdad anterior se transforma en $(a_1, \ldots, a_n) = (0, \ldots, 0)$, por lo que $a_1 = \ldots = a_n = 0$.

Caso de $M_{m\times n}(K)$. En $M_{m\times n}(K)$ el conjunto $S = \{E_{ij} \,|\, i = 1, \ldots, m,\ j = 1, \ldots, n\}$ definido en el Ejemplo 2.77 es l.i. Para verlo tomamos una c.l. del tipo

$$\sum_{i=1}^{m}\sum_{j=1}^{n} a_{ij} \cdot E_{ij} = a_{11} \cdot E_{11} + \ldots + a_{1n} \cdot E_{1n} + \ldots + a_{m1} E_{m1} + \ldots + a_{mn} E_{mn} = 0_{m\times n},$$

y nos preguntamos si $a_{ij} = 0$ para cada $i = 1, \ldots, m$ y cada $j = 1, \ldots, n$. Teniendo en cuenta cómo se definen las matrices E_{ij} la igualdad anterior equivale a la igualdad $A = 0_{m\times n}$, donde $A = (a_{ij})$. A partir de aquí se concluye lo que se quería.

Caso de $\mathbb{K}[x]$. En $\mathbb{K}[x]$ el conjunto $S = \{p_i(x) \,|\, i \in \mathbb{N} \cup \{0\}\}$ definido en el Ejemplo 2.78 es l.i. Hay que comprobar que cada subconjunto finito de S es l.i. Tomamos así cualquier $S' = \{p_{i_1}(x), \ldots, p_{i_m}(x)\} \subset S$ y una c.l. del tipo $a_1 \cdot p_{i_1}(x) + \ldots + a_m \cdot p_{i_m}(x) = 0$. Como $p_{i_k}(x) = x^{i_k}$, la igualdad previa implica inmediatamente que $a_i = 0$ para cada $i = 1, \ldots, m$.

Finalmente, veamos que el papel de K es importante cuando hablamos de independencia lineal.

[16] Si se consideran los polinomios como aplicaciones polinómicas, esto puede deducirse derivando esta igualdad dos veces, con lo que se obtiene $2(a+c) = 0$ así como $(a+2b)x + (a+b+c) = 0$, derivando a continuación esta igualdad una vez , de donde resulta $a+2b=0$ así como $a+b+c=0$.

Ejemplo 2.99. Sea $\mathbb{C}$ el cuerpo de los números complejos. Tomemos la familia $S = \{v_1, v_2\}$ con $v_1 = 1$ y $v_2 = i$. Si pensamos en $\mathbb{C}$ como e.v. complejo entonces S es l.d. pues $v_2 = i \cdot v_1$. Si vemos $\mathbb{C}$ como e.v. real entonces S es l.i. pues v_1 y v_2 son no nulos y no proporcionales (no existe $a \in \mathbb{R}$ tal que $v_2 = a \cdot v_1$).

Ejercicio 2.100. *Sea $S = \{v_1, \ldots, v_m\}$ es un subconjunto finito no vacío de $V(K)$. Demuéstrese:*
S es l.i. si y sólo si se tiene la suma directa $L\{v_1\} \oplus \cdots \oplus L\{v_m\}$.

2.3.3. Concepto de base y dimensión

Teorema de Steiniz

Ya hemos comentado que nos interesa encontrar s.d.g. de un espacio f.g. que tengan la menor cantidad posible de vectores, así como conjuntos l.i. que sean lo más grandes posibles, y planteamos dos cuestiones (la 2.81 y 2.93) al respecto. El próximo teorema establece una desigualdad entre los números de vectores de tales conjuntos. Más aún, el teorema servirá tanto para motivar la definición de *base* como para demostrar que dos bases de un espacio vectorial f.g. tienen el mismo número de vectores.

Teorema 2.101 (Teorema de Steinitz)**.** *Sea V un e.v. sobre un cuerpo K. Supongamos que tenemos dos subconjuntos finitos $S = \{u_1, u_2, \ldots, u_m\}$, que es un s.d.g. y $S' = \{v_1, v_2, \ldots, v_k\}$, que es l.i. Entonces, se cumple que $m \geq k$.*

Demostración. Se va a proceder por contradicción. Para ello basta con considerar el caso $k = m+1$ ya que, en el caso contrario, se obtendría la contradicción con el subconjunto de S' formado por sus primeros $m+1$ elementos, ya que este conjunto también es linealmente independiente. La idea de la contradicción que se persigue se desarrolla en los siguientes m pasos:

Paso 1: se toma v_1 y se se demuestra que, *salvo una reordenación de los elementos de S*, se puede sustituir el primero de ellos por v_1 de modo que $S_1 := \{v_1, u_2, \ldots, u_m\}$ sigue siendo un s.d.g.

Paso 2: se toma v_2 y se se demuestra que, *salvo una reordenación de los elementos de S_1*, se puede sustituir el segundo de ellos por v_2 de modo que $S_2 := \{v_1, v_2, u_3 \ldots, u_m\}$ sigue siendo un s.d.g.

(...) Paso m: se toma v_m y se se demuestra que, el m-ésimo (esto es, el último) de los elementos se $S_{m-1} := \{v_1, v_2, u_3 \ldots, v_{m-1}, u_m\}$, se puede sustituir por v_m de modo que $S_m := \{v_1, v_2, \ldots, v_m\}$ sigue siendo un s.d.g. Se tiene entonces que el vector $v_{m+1} \in S'$ se escribirá como c.l. de S_m, esto es, de los m primeros vectores de S', lo que contradice que S' es l.i.

Paso 1. Como S es un s.d.g. de V, la Proposición 2.82 (i) nos dice que $\overline{S}_1 = \{v_1, u_1, u_2, \ldots, u_m\}$ es un s.d.g. de V. Además, el vector v_1 de S' se expresa como una c.l. de S, es decir, existen escalares $a_1, a_2, \ldots, a_m \in K$ tales que:

$$v_1 = a_1 \cdot u_1 + a_2 \cdot u_2 + \ldots + a_m \cdot u_m.$$

Nótese que $v_1 \neq 0$ por ser S' l.i. Así, en la igualdad de arriba no todos los coeficientes del miembro derecho son nulos. Reordenando y renombrando si fuese necesario los vectores de S, podemos suponer $a_1 \neq 0$ (esto no resta generalidad al argumento). Despejamos entonces el vector u_1 en la ecuación de arriba y obtenemos:

$$u_1 = a_1^{-1} \cdot v_1 + (-a_1^{-1} \cdot a_2) \cdot u_2 + \ldots + (-a_1^{-1} \cdot a_m) \cdot u_m,$$

por lo que u_1 es una c.l. de $\{v_1, u_2, \ldots, u_m\}$, es decir, $u_1 \in L(\overline{S}_1 - \{u_1\})$. Aplicando la Proposición 2.82 (ii) deducimos que $S_1 = \overline{S}_1 - \{u_1\} = \{v_1, u_2, \ldots, u_m\}$ es un s.d.g. de V.

Ahora distinguimos dos casos. Si $m = 1$ entonces $S_1 = \{v_1\}$ y como existe el vector $v_2 \in S'$, éste tiene que ser una c.l. de S_1 (al ser S_1 un s.d.g. de V). Esto es una contradicción porque S' era l.i. Si $m \geq 2$ entonces continuamos con el siguiente paso.

Paso 2 (opcional para comprender mejor el procedimiento general). Como S_1 es un s.d.g. de V, la Proposición 2.82 (i) dice que $\overline{S}_2 = \{v_1, v_2, u_2, \ldots, u_m\}$ es un s.d.g. de V. Además, el vector v_2 de S' se podrá expresar como c.l. de S_1, es decir, existen $b_1, \ldots, b_m \in K$ tales que:

$$v_2 = b_1 \cdot v_1 + b_2 \cdot u_2 + b_3 \cdot u_3 \ldots + b_m \cdot u_m.$$

Necesariamente, alguno de los escalares b_i con $i > 1$ tiene que ser distinto de 0, porque, si no, se tendría una contradicción con que el conjunto $\{v_1, v_2\}$ (incluido en S') es l.i. Tras renombrar los vectores se puede suponer que $b_2 \neq 0$ y, despejando el vector u_2 en la ecuación de arriba, obtenemos:

$$u_2 = b_2^{-1} \cdot v_2 + (-b_2^{-1} \cdot b_1) \cdot v_1 + (-b_2^{-1} \cdot b_3) \cdot u_3 + \ldots + (-b_2^{-1} \cdot b_m) \cdot u_m,$$

por lo que u_2 es una c.l. de $\{v_1, v_2, \ldots, u_m\}$, es decir, $u_2 \in L(\overline{S}_2 - \{u_2\})$. Aplicando la Proposición 2.82 (ii) se llega a que $S_2 = \overline{S}_2 - \{u_2\} = \{v_1, v_2, u_3, \ldots, u_m\}$ es un s.d.g. de V.

Ahora distinguimos dos casos. Si $m = 2$ entonces $S_1 = \{v_1, v_2\}$ y el vector $v_3 \in S'$, será c.l. de S_2 (al ser S_2 un s.d.g. de V). Esto contradice que S' es l.i. Si $m \geq 2$ pasamos al siguiente paso.

Pasos sucesivos hasta m. Procediendo como en los casos anteriores, repetimos m veces el proceso para obtener que $S_m = \{v_1, v_2, \ldots, v_m\}$ es un s.d.g. de V. Como $m + 1 \leq k$, existe el vector $v_{m+1} \in S'$, el cual será una c.l. de S_m (al ser S_m un s.d.g. de V). Esto contradice que S' es l.i.

Para concretar más este procedimiento, supongamos que se han hecho $m' \in \{1, 2, \ldots, m-1, m\}$ pasos, y admitamos inductivamente que el conjunto $S_{m'} := \{v_1, \ldots, v_{m'}, v_{m'+1}, \ldots, v_m\}$ es un s.d.g. de V. Distinguimos dos casos.

Si $m' = m$ se tiene la contradicción buscada (el conjunto $S_m \cup \{v_{m+1}\}$ es l.i, por estar incluido en S', pero esto es absurdo ya que $v_{m+1} \in L(S_m)$, al ser este último conjunto un s.d.g. por la hipótesis de inducción).

En el caso $m' < m$, el vector $v_{m'+1}$ de S' se podrá expresar como c.l. de $S_{m'}$, es decir, existen $c_1, \ldots, c_m \in K$ tales que:

$$v_{m'+1} = c_1 \cdot v_1 + \cdots + c_{m'} \cdot u_{m'} + c_{m'+1} \cdot u_{m'+1} + \ldots + c_m \cdot u_m.$$

Necesariamente, alguno de los escalares b_i con $i > 1$ tiene que ser distinto de 0, porque, si no, se tendría una contradicción con que el conjunto $\{v_1, \ldots, v_{m'}\}$ (incluido en S') es l.i. Tras renombrar los últimos $m - m'$ vectores se puede suponer que $c_{m'+1} \neq 0$. Despejando el vector $u_{m'+1}$ en la ecuación anterior, se obtiene:

$$u_{m'+1} = c_{m'+1}^{-1} \cdot v_2 + (-c_{m'+1}^{-1} \cdot c_1) \cdot v_1 + \ldots + (-c_{m'+1}^{-1} \cdot c_m) \cdot u_m,$$

por lo que $u_{m'+1}$ es una c.l. de $\{v_1, \ldots, v_{m'+1}, \ldots, u_m\}$, es decir, $u_{m'+1} \in L(\overline{S}_{m'+1} - \{u_{m'+1}\})$. Aplicando la Proposición 2.82 se llega a que $S_{m'+1} := \overline{S}_{m'+1} - \{u_{m'+1}\} = \{v_1, ..., v_{m'+1}, u_{m'+2}, \ldots, u_m\}$ es un s.d.g. de V. Por tanto, supuesto que el proceso es válido para m' pasos, también lo es para $m' + 1$, y puede repetirse hasta llegar a $m' = m$. ∎

Ejercicio 2.102. *Encontrar cuatro subconjuntos finitos $S_i \subset \mathbb{R}^2$ con $i = 1, 2, 3, 4$, tales que S_1 sea s.d.g. pero no l.i., S_2 sea s.d.g. y l.i., S_3 no sea s.d.g. pero sí l.i. y S_4 no sea s.d.g. ni l.i. (esto ilustra que no existe relación entre el concepto de s.d.g. y el de conjunto l.i. más allá del teorema de Steinitz).*

Noción y existencia de bases

La idea de la demostración anterior sugiere la relevancia de conjuntos que sean a la vez s.d.g. y l.i.

Definición 2.103. *Sea V un e.v. sobre un cuerpo K. Una* base *de V es cualquier subconjunto $\mathcal{B} \subset V$ tal que $\mathcal{B}$ es un s.d.g. de V y $\mathcal{B}$ es l.i.*

Esta definición es válida para cualquier e.v. (aunque no sea f.g.). La idea es tener un s.d.g. irreducible, en el sentido de que ningún vector es c.l. del resto. La terminología "base" se justifica si pensamos que a partir de una cantidad mínima de vectores estamos generando todos los demás. A lo largo de esta asignatura nos interesarán especialmente los e.v. que tengan bases finitas.

Ejemplo 2.104. Los sistemas de referencia formados por dos vectores no colineales en el espacio de los vectores del plano o por tres no coplanarios en los del espacio son bases.

Ejemplo 2.105. Bases canónicas de algunos espacios vectoriales notables (véase el ejemplo 2.98):

En K^n, el conjunto $\mathcal{B}_u = \{e_1, \ldots, e_n\}$ definida en el Ejemplo 2.76 es una base, que llamaremos *base usual o canónica* de K^n.

En $M_{m\times n}(K)$, el conjunto $\mathcal{B}_u = \{E_{ij} \,|\, i = 1, \ldots, m,\ j = 1, \ldots, n\}$ definida en el Ejemplo 2.77 es una base, que llamaremos base usual o canónica *de $M_{m\times n}(K)$.*

En $\mathbb{K}[x]$, el conjunto $\mathcal{B}_u = \{p_i(x) \,|\, i \in \mathbb{N} \cup \{0\}\}$ con $p_i(x) = x^i$ es una base de $\mathbb{K}[x]$, que llamaremos *base usual o canónica* de $\mathbb{K}[x]$.

El disponer de una base en un e.v. es algo muy deseable, pues nos permite construir todos los vectores a partir de una cantidad mínima de ellos. En el caso límite del espacio vectorial trivial $V = \{0\}$, el conjunto vacío cumple formalmente los requisitos para ser base (pues es l.i. y $L(\emptyset) = \{0\} = V$). A continuación lo demostraremos para un caso mucho más interesante.

Teorema 2.106 (Existencia de bases)**.** *Sea V un e.v. sobre un cuerpo K.*

Si V es finitamente generado entonces admite al menos una base.

Más aún, si S es cualquier s.d.g. de V, entonces existe una base $\mathcal{B}$ de V incluida en S ($\mathcal{B} \subset S$).

Demostración. Escribiendo $S = \{v_1, \ldots, v_m\}$, La idea consiste en ir eliminando uno a uno generadores de S que se expresan como c.l. del resto hasta obtener, en una cantidad finita de pasos, un s.d.g. que es también l.i. Para formalizarlo con precisión, operaremos por inducción sobre el cardinal m de S.

El caso $m = 0$ resulta trivialmente cierto porque, si $S = \emptyset$ entonces $V = \{0\}$ (por ser S s.d.g.), que admite como base $\emptyset$ ($\subset S$). Aunque se puede realizar la inducción empezando $m = 0$, vale la pena discutir el caso $m = 1$, esto es, $S = \{v_1\}$. Si $V = \{0\}$ necesariamente $v_1 = 0$ y la base es $\emptyset$. Si $V \neq \{0\}$ $v_1 \neq 0$ (por ser S un s.d.g.) y, en consecuencia, S es l.i. Así, S resulta ser una base de V.

Supongamos ahora por hipótesis de inducción que todo s.d.g con $m-1$ elementos contiene una base, y consideremos un s.d.g. S de m elementos. Si S es un conjunto l.i., entonces S es una base de V y se concluye el resultado. De lo contrario, existe $i \in \{1, \ldots, m\}$ tal que $v_i \in L(S - \{v_i\})$. Gracias a la Proposición 2.82 (ii) sabemos que $S - \{v_i\}$ es un s.d.g. de V. Puesto que $S - \{v_i\}$ consta de $m-1$ vectores, la hipótesis de inducción garantiza que contiene una base. ∎

Nota 2.107. Resulta también cierto que los e.v. no finitamente generados admiten bases. No obstante, para demostrarlo se requiere el llamado *axioma de elección*, el cual excede los contenidos del presente curso. Por conveniencia del lector, lo esbozamos brevemente. De entre las distintas formulaciones del axioma de elección, consideramos la que proporciona el *Lema de Zorn*:

> Sea X cualquier conjunto no vacío con una relación de orden parcial $\leq$ que verifica:
>
> toda *cadena* $C \subset X$ (esto es, todo subconjunto C totalmente ordenado por $\leq$) posee una *cota superior* (esto es, un elemento s tal que $a \leq s$ para todo $a \in C$).
>
> Entonces, X admite un elemento *maximal* $M \in X$ (esto es, M satisface la propiedad: si $a \in X$ verifica $M \leq a$ entonces $M = a$).

Para demostrar la existencia de bases en $V(K)$, se toma como X el conjunto de todos los subconjuntos linealmente independientes de V (X no es vacío, pues $\emptyset \in X$) y como relación de orden parcial la relación de inclusión (no estricta) $\subset$. Un cadena C en X no es más que una colección totalmente ordenada por $\subset$ de subconjuntos l.i. de X, y la unión s de todos ellos sirve como cota superior (obsérvese que $s \in X$, esto es, s es l.i., porque cualquier subconjunto finito suyo estará incluido en algún elemento a de la cadena, el cual es l.i. por la definición de X). El elemento maximal M que proporciona el Lema de Zorn es necesariamente una base. En efecto, M es l.i. por pertenecer a X. M es un s.d.g. porque, dado cualquier $v \in V$, el conjunto $M \cup \{v\}$ es l.d. (pues si no $M \cup \{v\}$ sería l.i. por la proposición 2.94 y, por tanto, M no sería maximal), por lo que $v \in L(M)$ (contrarrecíproco de la proposición 2.94).

Dimensión

La demostración del teorema anterior nos hace pensar que un e.v. no trivial V que sea f.g. tendrá varias bases, pues a partir de cualquier s.d.g. se consigue una base eliminando sucesivamente generadores que se escriben como c.l. del resto. De hecho, es fácil darse cuenta de que si el cuerpo K es infinito entonces a partir de una base se pueden construir infinitas. Como consecuencia directa del teorema de Steinitz, comprobaremos que, aunque existan muchas bases, todas tienen algo en común: el número de vectores que poseen (obsérvese que esto no se ha visto aún, pues aunque dos bases son siempre irreducibles como s.d.g., en principio podría ocurrir que una tuviese más vectores que la otra).

Teorema 2.108 (de la dimensión)**.** *Todas las bases de un espacio f.g. son finitas y tienen el mismo número de vectores.*

Demostración. Veamos primero que todas las bases son finitas. Supongamos que hubiera dos bases $\mathcal{B}_1$ y $\mathcal{B}_2$ donde $\mathcal{B}_1$ tiene m vectores y $\mathcal{B}_2$ tiene infinitos. Como $\mathcal{B}_2$ es l.i. entonces cualquier subconjunto finito de $\mathcal{B}_2$ también lo es. Tomando un subconjunto $\mathcal{B}_2'$ en $\mathcal{B}_2$ con $m+1$ vectores y aplicando el teorema de Steinitz con $\mathcal{B}_1$ y $\mathcal{B}_2'$ tendríamos una contradicción. Esto implica que todas las bases de V son a la vez finitas o infinitas. Como V es f.g. hemos visto en la prueba del Teorema 2.106 que V tiene bases finitas. Por tanto, todas las bases de V son finitas. Finalmente, dadas dos bases finitas $\mathcal{B}_1$ y $\mathcal{B}_2$ con m y k vectores, respectivamente, aplicamos el teorema de Steinitz tomando $\mathcal{B}$ como s.d.g. y $\mathcal{B}'$ como conjunto l.i. Esto nos da $k \leq m$. Intercambiando los papeles de $\mathcal{B}$ y $\mathcal{B}'$ llegamos a $m \leq k$. ∎

Gracias al teorema anterior la siguiente definición tiene sentido.

Definición 2.109. *Se define la* dimensión *de un e.v.* $V(K)$*, que denotaremos* $\dim_K(V)$*, como sigue: (i) si* V *es f.g. y* $V \neq \{0\}$*, definimos* $\dim_K(V)$ *como el número (natural) de vectores de cualquier base de* V*, (ii) si* $V = \{0\}$ *definimos* $\dim_K(V) = 0$*, y (iii) si* V *no es f.g. diremos que el e.v. es* de dimensión infinita *y escribiremos* $\dim_K(V) = \infty$*.*

En particular, diremos que V *es una* recta vectorial *si* $\dim_K(V) = 1$*, y que* V *es un* plano vectorial *si* $\dim_K(V) = 2$*.*

Ejemplo 2.110. Tenemos $\dim_{\mathbb{R}}\mathbb{R}^n = n$, la dimensión del espacio vectorial de los vectores libres del plano es 2, y la de los vectores libres del espacio 3, $\dim_K(K^n) = n$, y $\dim_K(M_{m\times n}(K)) = m \cdot n$, $\dim_K(M_n(K)) = n^2$ y $\dim_{\mathbb{K}}(\mathbb{K}[x]) = \infty$. Podemos comprobar que $\mathbb{K}[x]$ no es f.g. o bien razonando como se hizo en el Ejemplo 2.78, o bien porque ya sabemos que $\mathbb{K}[x]$ tiene una base infinita (Ejemplo 2.105), por lo que si fuese f.g. todas sus bases serían finitas.

Veamos ahora que la dimensión de un espacio f.g. depende de forma esencial del cuerpo.

Ejemplo 2.111. Si vemos $\mathbb{C}$ como e.v. complejo entonces una base de $\mathbb{C}$ es $\mathcal{B}_{\mathbb{C}} = \{1\}$ y, por tanto, $\dim_{\mathbb{C}}(\mathbb{C}) = 1$. Si vemos $\mathbb{C}$ como e.v. real entonces una base está dada por $\mathcal{B}_{\mathbb{R}} = \{1, i\}$, de donde $\dim_{\mathbb{R}}(\mathbb{C}) = 2$. En la relación de problemas se generaliza este hecho: si V es un e.v. complejo con $\dim_{\mathbb{C}}(V) = n$, entonces V es un e.v. real con $\dim_{\mathbb{R}}(V) = 2n$.

No obstante, cuando se sobrentiende el cuerpo K en el que se trabaja, se puede suprimir el subíndice K de la expresión $\dim_K(V)$, esto es, escribir sólo $\dim(V)$.

Ahora podemos dar una interpretación de la dimensión de un espacio f.g. que resuelve de manera precisa dos cuestiones pendientes de las secciones anteriores.

Corolario 2.112. *Sea $V(K)$ un espacio vectorial f.g. Dados $S = \{v_1, \ldots, v_m\}$, un s.d.g. de V, y $S' = \{u_1, \ldots, u_k\}$ un conjunto l.i. en V, entonces:*

$$k \leq \dim_K(V) \leq m.$$

Demostración. V admite una base $\mathcal{B}$ con n vectores, siendo $n = \dim_K(V)$. Si aplicamos el teorema de Steinitz con S como s.d.g. de V y $\mathcal{B}$ como conjunto l.i. tenemos que $n \leq m$. Si aplicamos el teorema de Steinitz con $\mathcal{B}$ como s.d.g. de V y S' como conjunto l.i. deducimos que $k \leq n$, concluyéndose la demostración. ∎

Observación 2.113 (Respuesta a las cuestiones 2.81 y 2.93)**.** El corolario anterior implica que $\dim_K(V)$ es el mínimo de vectores que un s.d.g. de V puede tener. Esto resuelve la cuestión 2.81. Asimismo $\dim_K(V)$ es el mayor número de vectores que un conjunto l.i. en V puede tener. Esto resuelve la cuestión 2.93. Se sigue que ambas cuestiones tienen la misma respuesta y, por tanto, $\dim_K(V)$ es un valor representativo del "tamaño" de V.

En el teorema 2.106 hemos demostrado que, a partir de cualquier s.d.g., se puede obtener una base eliminando eventualmente algunos vectores. Ahora veremos que también se puede conseguir una base añadiendo vectores a un conjunto l.i.

Teorema 2.114 (ampliación de la base)**.** *Sea V un e.v. con $\dim_K(V) = n \in \mathbb{N}$. Si $S = \{v_1, \ldots, v_k\}$ es un conjunto l.i. y $k \leq n$, entonces existen $n-k$ vectores, $v_{k+1}, \ldots, v_n \in V$ tales que $\mathcal{B} = \{v_1, \ldots, v_k, v_{k+1}, \ldots, v_n\}$ es una base de V.*

Demostración. En el caso $k = n$, basta con comprobar que S es un s.d.g. En efecto, si no lo fuera, existiría un vector $v \in V \setminus \{v\}$ y $S \cup \{v\}$ sería l.i. (proposición 2.94) con $k = n+1$ vectores, en contradicción con el corolario anterior.

En el caso $k < n$, S no es un s.d.g. de V (de lo contrario S sería una base, en contradicción con el teorema 2.108), por lo que existe $v_{k+1} \in V$ con $v_{k+1} \notin L(S)$. En consecuencia, el conjunto $S \cup \{v_{k+1}\} = \{v_1, \ldots, v_k, v_{k+1}\}$ es l.i. (de nuevo por la proposición 2.94). Si $k+1 = n$ este conjunto será una base, aplicando el caso $k = n$ antes estudiado. En caso contrario, se repite el argumento un total de $n-k$

veces[17] hasta conseguir un conjunto $\mathcal{B} = \{v_1, \ldots, v_k, v_{k+1}, \ldots, v_n\}$ que es l.i. y tiene exactamente n vectores, la cual será una base aplicando de nuevo el caso $k = n$. ∎

Observación 2.115. Como caso particular, a partir de cualquier vector $v \in V$ con $v \neq 0$ se puede construir una base $\mathcal{B}$ de V con $v \in \mathcal{B}$. Como resulta patente en la demostración, la forma de ampliar un conjunto l.i. no es única, en general.

Ejemplo 2.116. En $\mathbb{R}^2$ el vector $u = (1, -2)$ es no nulo, y por tanto, existe una base $\mathcal{B}$ de $\mathbb{R}^2$ con $u \in \mathcal{B}$. Para construir $\mathcal{B}$ basta elegir un vector $v \in \mathbb{R}^2$ que no sea proporcional a u. Tomando por ejemplo $v = (0, 1)$ es fácil probar que $\mathcal{B} = \{u, v\}$ es una base de $\mathbb{R}^2$.

Nota 2.117. El teorema 2.114 se puede extender a dimensión infinita, esto es, cualquier conjunto linealmente independiente S se puede ampliar a una base. Su demostración puede llevarse a cabo razonando con argumentos similares a los de la nota 2.107.

Ejercicio 2.118. *Sea V un e.v. no trivial sobre un cuerpo K. Supongamos que V es f.g. y que S, S' son dos subconjuntos finitos de vectores de V de modo que $S' \subset S$, S es s.d.g. y S' es l.i. Demostrar que existe una base $\mathcal{B}$ de V tal que $S' \subset \mathcal{B} \subset S$.*

A continuación refinamos el corolario 2.112 mostrando que sus desigualdades son óptimas: cuando alguna de ellas se convierte en igualdad el conjunto correspondiente es una base de V. Esto resulta interesante desde el punto de vista práctico, pues *para demostrar que un subconjunto de un espacio vectorial f.g.* que tenga tantos vectores como la dimensión *es una base, basta con que sólo una de las dos condiciones de la definición (s.d.g ó l.i.) se verifique.*

Corolario 2.119. *Sea V un e.v. sobre un cuerpo K con $\dim_K(V) = n \geq 1$. Dada un subconjunto $\mathcal{B} = \{v_1, \ldots, v_n\}$ con n vectores de V, son equivalentes estas afirmaciones:*

(i) *$\mathcal{B}$ es una base de V.*

(ii) *$\mathcal{B}$ es un s.d.g. de V.*

(iii) *$\mathcal{B}$ es l.i. en V.*

Demostración. Es obvio que (i) implica (ii) y (iii) por definición de base. Veamos que (ii) implica (i). Supongamos que $\mathcal{B}$ es s.d.g. de V. Si $\mathcal{B}$ no fuese l.i. podríamos emplear el Teorema 2.106 para conseguir una base de V eliminando algunos vectores de $\mathcal{B}$. Así, obtendríamos una base de V con una cantidad de vectores menor que n, lo que contradice el teorema 2.108. Por último, (iii) implica (i) se sigue directamente del teorema 2.114 con $n = m$. ∎

Remarquemos que el resultado anterior no sería cierto en dimensión infinita.

Ejemplo 2.120. Consideremos el conjunto $\mathcal{B} = \{u, v\}$ del Ejemplo 2.116. Como $\mathcal{B}$ tiene 2 vectores y $\dim_{\mathbb{R}}(\mathbb{R}^2) = 2$ se sigue que $\mathcal{B}$ será una base de $\mathbb{R}^2$ si $\mathcal{B}$ es l.i. Pero esto es inmediato, ya que los vectores u y v no son proporcionales.

Finalmente, el siguiente resultado (sobre todo su apartado (3)) también tiene gran utilidad práctica.

[17]Formalmente, se puede proceder por inducción según el valor de $n - k$. Anteriormente se han demostrado los casos $n - k = 0, 1$.

Proposición 2.121. *Sea $V(K)$ un e.v. y $S = \{v_1, \ldots, v_m\} \subset V$ y $v \in V$ tal que $v = a_1v_1 + \ldots a_mv_m$ para ciertos $a_1, \ldots, a_m \in K$. Sea S_i el conjunto que se obtiene reemplazando en S al vector v_i por v, esto es, $S_i = \{v_1, \ldots, v_{i-1}, v, v_{i+1}, \ldots, v_m\}$. Se verifica:*

(1) En el caso de que S sea s.d.g., S_i es s.d.g. si $a_i \neq 0$.

(2) En el caso de que S sea l.i., S_i es l.i. si (y sólo si) $a_i \neq 0$.

(3) En el caso de que S sea una base, S_i es una base si y sólo si $a_i \neq 0$.

Demostración. (1) Reducción ya conocida de un s.d.g. (proposición 2.82 (ii)).

(2) La demostración puede hacerse directamente (como se propone en el ejercicio 2.122 a continuación). No obstante, la haremos ahora anticipando propiedades de subespacios vectoriales, las cuales se desarrollarán en la próxima subsección.

Observemos primero que, al ser S un conjunto l.i., $\dim_k(L(S)) = m$. Si $a_i \neq 0$ se tiene $v_i \in L(S_i)$ (porque v_i se puede despejar de la expresión de v). En consecuencia $S \subset L(S_i)$ y se tiene $L(S) = L(S_i)$ (pues $S_i \subset L(S)$ es inmediato). Por tanto, S_i es una base de $L(S)$, ya que es un s.d.g. de m elementos en un espacio vectorial de dimensión m. En particular, S_i es l.i. como subconjunto de $L(S)$ y, por tanto, también es l.i. como subconjunto de V (ya que la suma y el producto por escalares en U son los de V restringidos).[18]

(3) Inmediato de los casos anteriores. ∎

Ejercicio 2.122. *Pruébese el apartado (3) de la proposición anterior directamente de las definiciones de s.d.g, conjunto l.i. y base.*

2.3.4. Bases y dimensión de un subespacio. Fórmula de Grassmann

Sea $V(K)$ un e.v. Si U es un s.v. de V entonces U es un e.v. sobre K con las operaciones de V restringidas, por lo que, en particular, podemos hablar tanto de bases de U como de la dimensión de U sobre K, a la cual denotaremos $\dim_K(U)$.

Observación 2.123. Para cualquier subconjunto S de U, toda combinación lineal de S como subconjunto de U es también una combinación lineal como subconjunto de V, y viceversa. Esto es una consecuencia inmediata de que la suma y el producto por escalares en U son restricción de los de V. Así, trivialmente, si una combinación lineal de elementos de S, vista como c.l. en V, es distinta de 0, entonces también es distinta de 0, vista como una c.l. en U. En consecuencia, *todo subconjunto S de U que sea l.i. en U, también es l.i. en V.*

Ejemplo 2.124. Si $U = \{(x,y) \in \mathbb{R}^2 : x - y = 0\}$ entonces una base de U es $\mathcal{B}_U = \{(1,1)\}$ y, por tanto, $\dim_{\mathbb{R}}(U) = 1$.

Ejemplo 2.125. Si $S = \{v_1, \ldots, v_m\}$ es l.i. en V, entonces S es una base de $U = L(S)$ y $\dim_K(U) = m$. Si los vectores de S no son necesariamente l.i. entonces $\dim_K(U) \leq m$ en virtud del Corolario 2.112.

Ejemplo 2.126. En $\mathbb{K}_n[x]$ el conjunto $\mathcal{B}_u = \{p_i(x) \,|\, i \in \{0,1,\ldots,n\}\} = \{1, x, \ldots, x^n\}$ es una base de $\mathbb{K}_n[x]$, que llamaremos *base usual o canónica* de $\mathbb{K}_n[x]$ y representaremos por $\mathcal{B}_u$. En consecuencia, $\dim_{\mathbb{K}}(\mathbb{K}_n[x]) = n + 1$.

[18] Por otra parte, es trivial que si $a_i = 0$, entonces v se escribe como combinación lineal del resto de vectores de S_i ($v_i \in S_i \setminus \{v\}$), por lo que S_i es linealmente dependiente.

Más tarde veremos cómo calcular eficazmente la dimensión de un subespacio solución de un SEL homogéneo y de un subespacio $U = L(S)$ con S finito.

Ejercicio 2.127. *Calcúlese una base de cada subespacio de matrices $S_n(K)$ y $A_n(K)$ (con característica de $K \neq 2$), y compruébese $dim_K S_n(K) = n(n+1)/2$, $dim_K A_n(K) = n(n-1)/2$.*

Si U es un s.v. de V resulta natural plantearse estas cuestiones: ¿es U un espacio f.g. si lo es V? ¿Qué relación hay entre $\dim_K(U)$ y $\dim_K(V)$? ¿Y entre las bases de U y de V? Responderemos estas preguntas en el siguiente resultado.

Proposición 2.128. *Sea $V(K)$ un e.v. finitamente generado, $n = \dim_K(V)$, y U un s.v. de V. Entonces:*

(i) U es f.g., y su dimensión $m = \dim_K(U)$ satisface $m \leq n$, dándose la igualdad si y sólo si $U = V$.

(ii) Sea $0 \leq m \leq n$ y $\mathcal{B}_U = \{v_1, \ldots, v_m\}$ una base de U. Entonces existen $n-m$ vectores $v_{m+1}, \ldots, v_n \in V$ tales que $\mathcal{B} = \{v_1, \ldots, v_m, v_{m+1}, \ldots, v_n\}$ es una base de V, y el subespacio $W := L(v_{m+1}, \ldots, v_m)$ es un complementario de U en V, esto es, $V = U \oplus W$.

Demostración. (i) Si U no fuera f.g., razonando como en el teorema 2.114 podríamos construir un conjunto l.i. en U y, por tanto, l.i. en V, con un vector más que la dimensión n de V, lo que es absurdo. Igualmente, sería absurdo $m > n$, por lo que $m \leq n$. Si $m = n$ cualquier base $\mathcal{B}_U$ de U es un conjunto l.i. en V de n vectores. Por tanto, $\mathcal{B}_U$ es una base de V (corolario 2.119), y $U = L(\mathcal{B}_U) = V$.

(ii) La primera afirmación es una consecuencia inmediata del Teorema 2.114. Para la última se debe comprobar $V = U + W$ y $U \cap W = \{0\}$. Para la suma, basta con observar que de $\mathcal{B} \subset U \cup V$ se sigue $L(\mathcal{B}) \subset L(U \cup W)(= U + W)$ y, por ser $\mathcal{B}$ un s.d.g, $V \subset L(\mathcal{B}) \subset U + W$.

Sea ahora $v \in U \cap W$. El resultado se obtiene trivialmente si $U = \{0\}$ ó $W = \{0\}$, por lo que podemos suponer $0 < m < n$. Como $U = L(\mathcal{B}_U)$ existen escalares $a_1, \ldots, a_m \in K$ tales que $v = a_1 \cdot v_1 + \ldots + a_m \cdot v_m$. Y como $W = L(v_{m+1}, \ldots, v_n)$ existen $a_{m+1}, \ldots, a_n \in K$ tales que $v = a_{m+1} \cdot v_{m+1} + \ldots + a_n \cdot v_n$. Restando ambas igualdades para v obtenemos

$$0 = a_1 \cdot v_1 + \ldots + a_m \cdot v_m - a_{m+1} \cdot v_{m+1} + \ldots - a_n \cdot v_n,$$

que es una expresión del vector nulo como c.l. de $\mathcal{B}$. Puesto que $\mathcal{B}$ es l.i. se concluye $a_i = 0$, para todo $i = 1, \ldots, m$. Por tanto, $v = 0$. ∎

Observación 2.129. La demostración anterior nos da un método de construcción de subespacios complementarios. En efecto, hemos visto que si $\mathcal{B}_U = \{v_1, \ldots, v_m\}$ es una base de U y la completamos con vectores $\{v_{m+1}, \ldots, v_n\}$ hasta obtener una base de V, entonces el subespacio $W = L(v_{m+1}, \ldots, v_n)$ es un complementario de U en V. Nótese que cada completación de $\mathcal{B}_U$ hasta una base de V dará lugar a un complementario eventualmente distinto. Además, desde el punto de vista práctico, si conocemos ya una base $\mathcal{B}$ de $V(K)$ podemos entonces escoger los vectores para la ampliación de entre los de $\mathcal{B}$ (compruébese).

Ejemplo 2.130. Sea U el s.v. de $\mathbb{R}^2$ dado por $U = L(u)$ con $u = (1, -2)$. Es claro que $\mathcal{B}_U = \{u\}$ es una base de U y, por tanto, $\dim_K(U) = 1$. Dado $w = (0, 1)$, entonces $\mathcal{B} = \{u, v\}$ es una base de $\mathbb{R}^2$ y, por tanto, $W = L(w)$ es un complementario de U en $\mathbb{R}^2$. Asimismo, si $w' = (1, 0)$ y $W' = L(w)$ entonces W' es otro complementario. Así, se cumplen $\mathbb{R}^2 = U \oplus W = U \oplus W'$ (obsérvese que $W \neq W'$, esto es, ¡no se puede "simplificar" en la igualdad anterior!).

Definición 2.131. *Sea $V(K)$ un e.v. y $U \subset V$ un s.v. de V. Si $dim_K(U) = 1$ diremos que U es una* recta vectorial en V. *Si $dim_K(U) = 2$ diremos que U es un* plano vectorial en V.

Si V es f.g. llamamos codimensión *de U en V al número en $\mathbb{N} \cup \{0\}$ dado por:*

$$codim_K(U) = dim_K(V) - dim_K(U).$$

Si $codim_K(U) = 1$, es decir, $dim_K(U) = dim_K(V) - 1$ diremos que U es un hiperplano vectorial de V.

Observación 2.132. Es obvio que $U = V$ si y sólo si $\text{codim}_K(U) = 0$. Así, la codimensión de U en V es una medida de lo "próximo" que está U de coincidir con V.

Terminamos esta sección resolviendo la siguiente cuestión: ¿qué relación hay entre la dimensión del subespacio suma $U_1 + U_2$ y las dimensiones de los sumandos? ¿Cómo construir una base de $U_1 + U_2$ a partir de bases de U_1 y de U_2? ¿Qué papel desempeña $U_1 \cap U_2$?

Teorema 2.133 (Fórmula de Grassmann)**.** *Sea $V(K)$ un e.v. y sean U_1 y U_2 dos s.v. finitamente generados de V. Se verifica:*

$$\dim_K(U_1 + U_2) = \dim_K(U_1) + \dim_K(U_2) - \dim_K(U_1 \cap U_2).$$

Observación 2.134. La fórmula tiene sentido, pues todos los espacios que aparecen son f.g. Por otra parte, esta fórmula resulta natural si se piensa, p. ej., en el número de elementos de la unión de dos conjuntos finitos (o en propiedades como el volumen de dos objetos cuya intersección es no vacía).

Demostración. Sea $m = \dim_K(U_1 \cap U_2)$ y $n_i = \dim_K(U_i)$ para cada $i = 1,2$. Queremos demostrar que $\dim_K(U_1 + U_2) = n_1 + n_2 - m$. Nuestro objetivo es construir una base de $U_1 + U_2$ con exactamente $n_1 + n_2 - m$ vectores.

Sea $\mathcal{B}_{U_1 \cap U_2} = \{v_1, \ldots, v_m\}$ una base de $U_1 \cap U_2$ (si $U_1 \cap U_2 = \{0\}$, tomamos consistentemente $\mathcal{B}_{U_1 \cap U_2} = \emptyset$). Por la Proposición 2.128 (ii), como $U_1 \cap U_2$ es un s.v. de U_1 y de U_2, podemos encontrar vectores $v^i_{m+1}, \ldots, v^i_{n_i} \in U_i$ tales que $\mathcal{B}_i = \{v_1, \ldots, v_m, v^i_{m+1}, \ldots, v^i_{n_i}\}$ es una base de U_i. Sea $W_i = L(v^i_{m+1}, \ldots, v^i_{n_i})$, que es un s.v. de U_i (obsérvese que si $U_1 \cap U_2 = U_i$, tomamos consistentemnete $W_i = \{0\}$). Por la Proposición 2.128 (iv) sabemos que $U_i = (U_1 \cap U_2) \oplus W_i$, para cada $i = 1,2$. Además

$$U_1 \cap W_2 = \{0\} \qquad \text{(y análogamente } U_2 \cap W_1 = \{0\}) \tag{2.9}$$

ya que, por la construcción de W_2, $\{0\} = (U_1 \cap U_2) \cap W_2 = U_1 \cap (U_2 \cap W_2) = U_1 \cap W_2$ (la última igualdad porque $W_2 \subset U_2$). Consideremos ahora el conjunto de vectores:

$$\mathcal{B} = \mathcal{B}_1 \cup \mathcal{B}_2 = \{v_1, \ldots, v_m, v^1_{m+1}, \ldots, v^1_{n_1}, v^2_{m+1}, \ldots, v^2_{n_2}\}.$$

Una primera observación sobre este conjunto es que no se han escrito dos vectores repetidos en la expresión del miembro derecho (la única posibilidad no trivial sería $v^1_i = v^2_j$ para algún $i \in \{m+1, \ldots, n_1\}$ y $j \in \{m+1, \ldots, n_2\}$, pero en ese caso ese vector común $v := v^1_i = v^2_j$ pertenecería a $U_1 \cap U_2$, por lo que v se podría escribir como c.l. de $\{v_1, \ldots, v_m\}$, en contradicción con la independencia lineal de cada $\mathcal{B}_i$). Por tanto, en $\mathcal{B}$ hay $m + (n_1 - m) + (n_2 - m) = n_1 + n_2 - m$ vectores, y basta con demostrar:

- $\mathcal{B}$ es una base de $U_1 + U_2$.

Para comprobar que $\mathcal{B}$ un s.d.g. de $U_1 + U_2$, basta con observar: $L(\mathcal{B}) = L(\mathcal{B}_1 \cup \mathcal{B}_2) = U_1 + U_2$. Para comprobar que $\mathcal{B}$ es un conjunto l.i. en $U_1 + U_2$, dada una expresión del tipo:

$$\sum_{i=1}^{m} a_i \cdot v_i + \sum_{i=m+1}^{n_1} b_i \cdot v^1_i + \sum_{i=m+1}^{n_2} c_i \cdot v^2_i = 0,$$

nos preguntamos si todos los coeficientes son nulos. La igualdad anterior es equivalente a:

$$\sum_{i=1}^{m} a_i \cdot v_i + \sum_{i=m+1}^{n_1} b_i \cdot v_i^1 = -\sum_{i=m+1}^{n_2} c_i \cdot v_i^2, \tag{2.10}$$

donde el miembro izquierdo de la igualdad es un vector de U_1 y el derecho es un vector de W_2. Usando (2.9) los dos miembros de la igualdad (2.10) coinciden con el vector nulo, es decir:

$$\sum_{i=1}^{m} a_i \cdot v_i + \sum_{i=m+1}^{n_1} b_i \cdot v_i^1 = 0 = \sum_{i=m+1}^{n_2} c_i \cdot v_i^2.$$

A partir de aquí deducimos que todos los coeficientes se anulan ya que $\mathcal{B}_1$ es l.i. y $\{v_{m+1}^2, \ldots, v_{n_2}^2\}$ está incluido en $\mathcal{B}_2$, que es también l.i. ∎

Como consecuencia inmediata tenemos el siguiente resultado, cuya prueba queda como ejercicio.

Corolario 2.135. *Sea $V(K)$ un espacio vectorial finitamente generado. Dados dos subespacios vectoriales U_1 y U_2 de V, son equivalentes estas afirmaciones:*

(i) $\dim_K(U_1 + U_2) = \dim_K(U_1) + \dim_K(U_2)$.

(ii) $U_1 \cap U_2 = \{0\}$.

Ejercicio 2.136. *Sea $V(K)$ un espacio vectorial y $\{U_1, U_2, U\}$ subespacios de V, U finitamente generado, tales que $U_1, U_2 \subset U$. Demostrar que son equivalentes:*

(i) $U = U_1 \oplus U_2$.

(ii) $dim_K(U) = dim_K(U_1) + dim_K(U_2)$.

(iii) Para cada base de U_i, $i = 1, 2$, se tiene que $\mathcal{B}_1 \cup \mathcal{B}_2$ es una base de U.

¿Se puede obtener toda base de U como en el punto (iii)?

Ejercicio 2.137 (Generalización de la fórmula de Grassmann para una suma finita de subespacios)**.** . *Sea $V(K)$ un e.v. y sea $\{U_i : i = 1, \ldots, m\}$ una familia finita de s.v. finitamente generados de $V(K)$. Entonces:*

$$dim_K\left(\sum_{i=1}^{m} U_i\right) = \sum_{i=1}^{m} dim_K(U_i) - \sum_{i=1}^{m-1} dim_K\big((\sum_{j=1}^{i} U_j) \cap U_{i+1}\big).$$

En particular, equivalen: (a) $U_1 \oplus \ldots \oplus U_m$, y (b) $dim_K\left(\sum_{i=1}^{m} U_i\right) = \sum_{i=1}^{m} dim_K(U_i)$. (Sugerencia: téngase en cuenta el ejercicio 2.100(ii) y la definición 2.59).

2.3.5. Coordenadas respecto de una base

Una de las utilidades principales de tener una base en un e.v. es la de poder identificar cada vector con una familia de escalares que llamaremos coordenadas. Veamos en detalle cómo se realiza esta identificación.

Sea $V(K)$ un e.v. y $\mathcal{B} = \{v_1, \ldots, v_n\}$ una base de V. Como $\mathcal{B}$ es un s.d.g. de V sabemos que todo vector de V se expresa como c.l. de $\mathcal{B}$. Dicho de otro modo, para cada $v \in V$, existen escalares $a_1, \ldots, a_n \in K$ tales que $v = a_1 \cdot v_1 + \ldots + a_n \cdot v_n$. Como esta propiedad la cumple cualquier s.d.g. de V nos planteamos si, al ser $\mathcal{B}$ una base, tenemos alguna ventaja adicional. Esto es lo que veremos en el siguiente resultado.

Teorema 2.138 (existencia de coordenadas)**.** *Sea $V(K)$ un e.v. y $\mathcal{B} = \{v_1, \ldots, v_n\}$ una base de V. Entonces, para cada $v \in V$, existen escalares únicos $a_1, \ldots, a_n \in K$ tales que $v = a_1 \cdot v_1 + \ldots + a_n \cdot v_n$.*

Demostración. Sea $v \in V$. La existencia de escalares a_i que cumplen $v = a_1 \cdot v_1 + \ldots + a_n \cdot v_n$ se debe a que $\mathcal{B}$ es un s.d.g. de V. La unicidad de estos escalares es consecuencia de la independencia lineal de $\mathcal{B}$; aunque esta propiedad se vio en la proposición 2.91 para todo conjunto l.i., la detallamos nuevamente por su importancia.

Supongamos que tenemos dos expresiones de v como c.l. de $\mathcal{B}$, esto es,

$$v = a_1 \cdot v_1 + \ldots + a_n \cdot v_n,$$
$$v = b_1 \cdot v_1 + \ldots + b_n \cdot v_n.$$

para ciertos escalares $a_1, \ldots, a_n, b_1, \ldots, b_n \in K$. Queremos demostrar que $a_i = b_i$, para cada $i = 1, \ldots, n$. Si a la primera igualdad de arriba le restamos la segunda, llegamos a:

$$(a_1 - b_1) \cdot v_1 + \ldots + (a_n - b_n) \cdot v_n = 0,$$

que es una expresión del vector nulo como c.l. del conjunto $\mathcal{B}$. Como $\mathcal{B}$ es l.i. deducimos que $a_i - b_i = 0$ para cada $i = 1, \ldots, n$, como se quería demostrar. ∎

Observación 2.139. El recíproco del resultado anterior es cierto. De hecho, si $V(K)$ es un e.v. y $\mathcal{B} = \{v_1, \ldots, v_n\} \subset V$, son equivalentes las siguientes afirmaciones:

(i) $\mathcal{B}$ es una base de V.

(ii) $V = L(v_1) \oplus \ldots \oplus L(v_n)$.

(iii) Para cada $v \in V$, existen escalares únicos $a_i \in K$ tales que $v = a_1 \cdot v_1 + \ldots + a_n \cdot v_n$.

La demostración de estas equivalencias queda como ejercicio.

El teorema previo nos indica que, en presencia de una base, tenemos una "determinación numérica" de los vectores, pues a cada vector le asociamos de forma única tantos escalares como la dimensión. No obstante, debemos discutir un detalle sobre la ordenación de los escalares y la base. Obsérvese que en el teorema anterior los elementos de la base estaban ordenados por un subíndice. Esta ordenación coincidía con la ordenación física de los símbolos $v_1, \ldots v_n$, pero desde un punto de vista conjuntista se verifica, por ejemplo $\{v_1, v_2\} = \{v_2, v_1\}$. Como los escalares deben asignarse con precisión a los vectores, introducimos el siguiente concepto.

Definición 2.140. *Diremos que una n-úpla $B = (v_1, \ldots, v_n) \in V \times \ldots \times^{(n)} V$ es una base ordenada si el conjunto $\{v_1, \ldots, v_n\}$ es una base de $V(K)$.*

Con este concepto, pasamos a dar nombre a los escalares proporcionados por el teorema 2.138.

Definición 2.141. *Sea $V(K)$ un e.v. y $B = (v_1, \ldots, v_n)$ una base ordenada de $V(K)$. Dado $v \in V$, a la única n-úpla de escalares $(a_1, \ldots, a_n) \in K^n$ tal que $v = a_1 \cdot v_1 + \ldots + a_n \cdot v_n$ se le llamará* n-úpla de coordenadas *o, simplemente,* coordenadas, de v respecto de B. *En tal caso escribiremos la n-úpla como columna:*

$$v_B = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}.$$

Ejemplo 2.142. Las bases ordenadas $B = (v_1, v_2, \ldots, v_n)$ y $B' = (v_2, v_1, \ldots, v_n)$ dan lugar a la misma base conjuntista $\mathcal{B} = \{v_1, v_2, \ldots, v_n\} = \{v_2, v_1, \ldots, v_n\}$. Dado $v \in V$, si

$$v_B = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix}$$

son las coordenadas de v respecto de B, entonces:

$$v_{B'} = \begin{pmatrix} a_2 \\ a_1 \\ \vdots \\ a_n \end{pmatrix}$$

son las coordenadas de v respecto de B'; obviamente $v_B \neq v_{B'}$.

Convenio 2.143. Aunque las coordenadas de un vector constituyen una n-úpla que se asigna a una base ordenada (no a una base), usualmente se abusa del lenguaje y no se distingue entre bases y bases ordenadas. Esto puede hacerse cuando se proporcione la base $\mathcal{B}$ de modo que tenga una ordenación natural (digamos, el orden en el que se ha escrito), aunque no se esté diciendo explícitamente que está ordenada. Así, si mencionamos *coordenadas en una base* $\mathcal{B}$, presuponemos implícitamente alguna ordenación.

Observación 2.144. Dada una base ordenada $B = (v_1, \ldots, v_n)$ en V, las coordenadas v_B de un vector $v \in V$ proporcionan un elemento de K^n (identificando cada n-upla con un elemento de $M_{n\times 1}(K)$). La aplicación

$$f_B : V \to K^n, \qquad v \mapsto f_B(v) = v_B$$

es así una biyección entre V y K^n. Anticipando el lenguaje del próximo tema, f_B es un *isomorfismo de espacios vectoriales*: además de la biyectividad, cumple que $f_B(a \cdot u + b \cdot v) = a \cdot f_B(u) + b \cdot f_B(v)$, para todo $a, b \in K$ y todo $u, v \in V$. Esto equivale a la identidad $(a \cdot u + b \cdot v)_B = a \cdot u_B + b \cdot v_B$, que se puede demostrar como ejercicio.

Este isomorfismo f_B establece una “identificación” (dependiente de B) entre V y K^n como espacios vectoriales sobre K. Así, las propiedades de V que sólo dependan de la estructura de e.v. son equivalentes a las mismas propiedades en K^n. Por ejemplo, es fácil demostrar[19] directamente que $S = \{u_1, \ldots, u_m\}$ es un s.d.g. (resp. una familia l.i.) de V si y sólo si $S_B = \{(u_1)_B, \ldots, (u_m)_B\}$ es un s.d.g. (resp. una familia l.i.) de K^n.

Ejercicio 2.145. *Sea $V(K)$ un e.v. y $B = (v_1, \ldots, v_n)$ una base ordenada de V. ¿Qué coordenadas tiene el vector nulo respecto de B? ¿Y cada uno de los vectores v_i?*

Ejemplo 2.146. Una consecuencia directa de lo anterior es que $\mathbb{R}(\mathbb{Q})$ (recuérdese la observación 2.18) *no es finitamente generado*. En efecto, si existiera una base B con $n \in \mathbb{N}$ vectores, la aplicación $f_B : \mathbb{R} \to \mathbb{Q}^n$ sería biyectiva, contradiciendo el hecho sabido de que $\mathbb{R}$ *no* es numerable.

Ejercicio 2.147. *Demuéstrese que si $V(\mathbb{R})$ es un espacio vectorial con $\dim_{\mathbb{R}} V > 0$, entonces $V(\mathbb{Q})$ no es finitamente generado.*

[19] Demostraremos un resultado más general válido para todo isomorfismo de e.v. en el tema siguiente.

En los siguientes ejemplos se calculan las coordenadas en algunas de las bases ya estudiadas.

Ejemplo 2.148. Si en K^n tomamos la base (ordenada) usual $B_u = (e_1, \ldots, e_n)$ entonces cada vector $v = (x_1, \ldots, x_n) \in K^n$ se expresa como $v = x_1 \cdot e_1 + \ldots + x_n \cdot e_n$. Esto implica que:

$$v_{B_u} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix},$$

es decir, las coordenadas coinciden con las componentes del vector. Esta es una propiedad exclusiva de B_u que facilita mucho los cálculos.

Ejemplo 2.149. Si en $M_{m\times n}(K)$ tomamos $B_u = \{E_{ij} \,|\, i = 1, \ldots, m,\ j = 1, \ldots, n\}$, entonces cada $A = (a_{ij}) \in M_{m\times n}(K)$ se expresa como $A = \sum_{i=1}^m \sum_{j=1}^n a_{ij} \cdot E_{ij}$. Tomando B como una base ordenada (consistentemente con el convenio 2.143, ordenaremos B suponiendo que en primer lugar se tienen los elementos de la primera fila con su ordenación natural, después los de la segunda, etc.) esto implica que las componentes de la matriz coinciden con sus coordenadas en B_u (suponiendo análogamente que la $m\times n$-úpla de coordenadas se escribe como una matriz situando las n primeras coordenadas en la primera fila, las n siguientes en la segunda, etc.)

Ejemplo 2.150. Si en $\mathbb{K}_n[x]$ tomamos la base usual $B_u = (1, x, \ldots, x^n)$ se cumple que cada $p(x) = a_0 + a_1 x + \ldots + a_n x^n$ se expresa como $p(x) = a_0 \cdot p_0(x) + a_1 \cdot p_1(x) + \ldots + a_n \cdot p_n(x)$, donde $p_i(x) = x^i$, para cada $i = 0, 1, \ldots, n$. Así, las coordenadas de $p(x)$ respecto de B_u coinciden con los coeficientes del polinomio (ordenados de menor a mayor).

En general, el cálculo de las coordenadas respecto de una base es equivalente a la resolución de un SEL compatible determinado. Veámoslo con un ejemplo.

Ejemplo 2.151. En $\mathbb{R}^2$ consideramos el conjunto $B = \{v_1, v_2\}$ con $v_1 = (1,0)$ y $v_2 = (-1,1)$. Es obvio que B es una base de $\mathbb{R}^2$ pues $\dim_{\mathbb{R}}(\mathbb{R}^2) = 2$ y B es l.i. (v_1 y v_2 no son proporcionales). Calculemos las coordenadas de $v = (x,y) \in \mathbb{R}^2$ respecto de B (con su ordenación natural), esto es, los el único par de escalares $(a,b) \in \mathbb{R}^2$ tales que $v = a \cdot v_1 + b \cdot v_2$. Operando y tomando componentes llegamos a la igualdad $(x,y) = (a-b, b)$, de donde se tiene el SEL de ecuaciones $a - b = x$ y $b = y$. Este SEL es compatible determinado con soluciones $a = x + y$ y $b = y$. Por tanto:

$$v_B = \begin{pmatrix} x+y \\ y \end{pmatrix}.$$

Se comprueba rápidamente que, efectivamente, $v = (x+y) \cdot v_1 + y \cdot v_2$. Por ejemplo, si $v = (-2,7)$ entonces $v = 5 \cdot v_1 + 7 \cdot v_2$. Sin embargo, en B_u se tiene que $v = -2 \cdot e_1 + 7 \cdot e_2$.

En el ejemplo anterior apreciamos que un mismo vector puede tener coordenadas distintas en distintas bases. Nos preguntamos ahora lo siguiente: ¿qué relación existe entre las coordenadas de un vector en dos bases distintas de un mismo e.v.? Para responder esta cuestión necesitamos relacionar los vectores de ambas bases.

Para fijar ideas, supongamos que V es un e.v. sobre un cuerpo K con $\dim_K(V) = n$. Tomamos dos bases $B = (v_1, \ldots, v_n)$ y $B' = (v'_1, \ldots, v'_n)$ de V. Dado $v \in V$, la cuestión es: ¿qué relación hay entre v_B y $v_{B'}$? (p. ej., supongamos que conocemos v_B y queremos determinar $v_{B'}$ a partir de v_B). Para ello necesitamos conocer cómo los vectores de B se expresan como c.l. de los de B'.

Definición 2.152. *En las condiciones anteriores, la* matriz de cambio de base de B a B' *es la matriz en* $M_n(K)$*, que denotaremos* $M(I,B' \leftarrow B)$*, cuya columna j-ésima contiene las coordenadas del vector* v_j *de* B *respecto de* B'*. Lo simbolizamos así:*

$$M(I,B' \leftarrow B) = \big((v_1)_{B'}\,|\,(v_2)_{B'}\,|\,\dots\,|\,(v_n)_{B'}\big)(=: B_{B'}).$$

Esto es, si (a_{ij}) *es* $M(I,B' \leftarrow B)$ *entonces* $v_j = \sum_{i=1}^n a_{ij}v'_i$*, para todo* $j = 1,\dots,n$*.*

Observación 2.153. Por construcción, el rango (por columnas) de la matriz $M(I,B' \leftarrow B)$ es el máximo posible (esto es, n; recuérdese la observación 2.144). Recíprocamente, se se tiene una base B y una matriz A de rango n, existe una única base B' tal que $A = M(I,B' \leftarrow B)$.

Ahora ya podemos resolver el problema que nos habíamos planteado.

Proposición 2.154 (cambio de base)**.** *En las condiciones anteriores, se tiene:*

$$v_{B'} = M(I,B' \leftarrow B)\cdot v_B, \qquad \forall v \in V.$$

Esto es, si $v_B = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ *y* $v_{B'} = \begin{pmatrix} x'_1 \\ \vdots \\ x'_n \end{pmatrix}$ *entonces* $\begin{pmatrix} x'_1 \\ \vdots \\ x'_n \end{pmatrix} = M(I,B' \leftarrow B)\cdot \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$
o, equivalentemente,

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

Llamaremos a cualquier de las igualdades anteriores la expresión matricial del cambio de base de B a B'*. Además, esta expresión matricial caracteriza a* $M(I,B' \leftarrow B)$ *en el siguiente sentido: si* $A \in M_n(K)$ *verifica* $v_{B'} = A\cdot v_B$ *para todo* $v \in V$*, entonces* $A = M(I,B' \leftarrow B)$*.*

Demostración. Partimos de:

$$v_j = a_{1j}\cdot v'_1 + \dots + a_{nj}\cdot v'_n = \sum_{i=1}^n a_{ij}\cdot v'_i \qquad \text{para cada } j = 1,\dots,n,$$

así como de

$$v = x_1\cdot v_1 + \dots + x_n\cdot v_n, = \sum_{i=1}^n x_i\cdot v_i.$$

Buscamos expresar v como c.l. de B'. Sustituyendo cada v_j de la primera ecuación en la segunda obtenemos la igualdad:

$$\begin{aligned} v &= x_1\cdot(a_{11}\cdot v'_1 + \dots + a_{n1}\cdot v'_n) + \dots + x_n\cdot(a_{1n}\cdot v'_1 + \dots + a_{nn}\cdot v'_n) \\ &= \sum_{j=i}^n x_j(\sum_{i=1}^n a_{ij}v'_i) = \sum_{j=i}^n(\sum_{i=1}^n x_j a_{ij}v'_i) = \sum_{j=1}^n(\sum_{i=1}^n (a_{ij}x_j)v'_i) = \sum_{i=1}^n(\sum_{j=1}^n (a_{ij}x_j)v'_i) \\ &= (a_{11}\cdot x_1 + \dots + a_{1n}\cdot x_n)\cdot v'_1 + \dots + (a_{n1}\cdot x_1 + \dots + a_{nn}\cdot x_n)\cdot v'_n, \end{aligned}$$

que nos expresa v como c.l. de B'. Esto significa que los coeficientes de la c.l. anterior son las coordenadas de v respecto de B'. Por definición de producto de matrices, la entrada i-ésima de $v_{B'}$ es el producto de la fila i-ésima de $M(I,B' \leftarrow B)$ con el vector columna v_B.

Para la última afirmación, restando las igualdades que verifican A y $M(I,B' \leftarrow B)$ se tiene:

$$\big(A - M(I,B' \leftarrow B)\big)\cdot v_B = 0(\equiv 0_{n\times 1}).$$

Tomando $v_B = (v_j)_B$ donde v_j es el $j-$ésimo vector de la base B (esto es $(v_j)_B$ tiene todas las componentes 0 salvo la $j-$ésima, que es igual a 1) se obtiene que la columna $j-$ésima está compuesta de ceros y, al ocurrir esto para toda columna, $A - M(I, B' \leftarrow B) = 0 (\equiv 0_{n\times n})$. ∎

Ejemplo 2.155. En K^n, la matriz $M(I, B_u \leftarrow B) = \big((v_1)_{B_u} \,|\, (v_2)_{B_u} \,|\, \ldots \,|\, (v_n)_{B_u}\big)$ es aquella cuyas columnas contienen las componentes de los vectores de B.

Ejemplo 2.156. En $\mathbb{R}^2$ consideramos la base $B = (v_1, v_2)$ con $v_1 = (1, 0)$ y $v_2 = (-1, 1)$ del ejemplo 2.151. Ya vimos que, para cada $v = (x, y) \in \mathbb{R}^2$ se tiene:

$$v_B = \begin{pmatrix} x+y \\ y \end{pmatrix},$$

con un cálculo directo de las coordenadas. Vamos a llegar ahora al mismo resultado usando un cambio de base. Tomando B_u como base de referencia, la Proposición 2.154 nos dice que:

$$v_B = M(I, B \leftarrow B_u) \cdot v_{B_u}, \text{ donde } M(I, B \leftarrow B_u) = ((e_1)_B, (e_2)_B).$$

Necesitamos las coordenadas de e_1 y e_2 con respecto a B. Tras unos cálculos sencillos:

$$(e_1)_B = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \text{ y } (e_2)_B = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \text{ por lo que } M(I, B \leftarrow B_u) = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}.$$

Por último, teniendo en cuenta que en B_u las coordenadas de un vector están dadas por sus componentes, deducimos que:

$$v_B = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x+y \\ y \end{pmatrix},$$

con lo que llegamos al mismo resultado de antes.

Para terminar, veremos algunas propiedades generales de las matrices de cambio de base que se relacionan con las propiedades del producto de matrices. Recordemos primero que una matriz cuadrada, $A \in M_n(\mathbb{K})$ se dice *regular* si admite inversa para el producto, esto es, si existe otra matriz cuadrada, A^{-1} tal que $A \cdot A^{-1} = A^{-1} \cdot A = I_n$, donde $I_n \in M_n(K)$ es la *matriz identidad*, definida por $(I_n)_{jk} = \delta_{jk}$, para todo $j, k = 1, \ldots, n$.

Proposición 2.157. *Sea $V(K)$ un e.v. de dimensión $n \in \mathbb{N}$, y sean B, B', B'' tres bases ordenadas de $V(K)$. Se verifica:*
(i) $M(I, B \leftarrow B'') = M(I, B \leftarrow B') \cdot M(I, B' \leftarrow B'')$.
(ii) $M(I, B \leftarrow B)$ es la matriz identidad I_n.
(iii) $M(I, B' \leftarrow B)$ es una matriz regular *y su inversa verifica: $M(I, B' \leftarrow B)^{-1} = M(I, B \leftarrow B')$.*

Demostración. Para (i), sea v cualquier vector de V. La Proposición 2.154 nos dice:

$$v_{B''} = M(I, B'' \leftarrow B') \cdot v_{B'} \qquad v_{B'} = M(I, B' \leftarrow B) \cdot v_B.$$

Sustituyendo la segunda igualdad en la primera, y usando la asociatividad del producto de matrices:

$$v_{B''} = \big(M(I, B'' \leftarrow B') \cdot M(I, B' \leftarrow B)\big) \cdot v_B.$$

Por otra parte, también sabemos:

$$v_{B''} = M(I, B'' \leftarrow B) \cdot v_B$$

por lo que la igualdad requerida se deduce de esta igualdad caracteriza a $M(I,B'' \leftarrow B)$, según la última afirmación de la Proposición 2.154.

La afirmación (ii) es inmediata. Para (iii), aplicando la parte (i) y a continuación (ii):

$$M(I,B \leftarrow B') \cdot M(I,B' \leftarrow B) = M(I,B \leftarrow B) = I_n,$$

y para la igualdad $M(I,B' \leftarrow B) \cdot M(I,B \leftarrow B') = I_n$ basta con intercambiar el papel de las bases. ∎

Observación 2.158. Cada matriz $A \in M_n(K)$ puede verse como la matriz de coordenadas de n vectores $\bar{v}_1,\dots,\bar{v}_n$ de algún e.v. $V(K)$ de dimensión n en alguna base B de este espacio (p. ej. $V = K^n$, $B = B_u$), esto es $A = ((\bar{v}_1)_B,\dots,(\bar{v}_n)_B)$. Con esta interpretación, podemos obtener ahora como consecuencia sencilla de nuestro estudio el resultado: $A \in M_n(K)$ *es regular si y sólo si su rango (por columnas) es n.* $(\Leftarrow)$ Al ser el rango n, A es una matriz de cambio de base (observación 2.153), por lo que es regular por el apartado (iii) de la proposición anterior. $(\Rightarrow)$ Si $a_1(\bar{v}_1)_B + \cdots + a_n(\bar{v}_n)_B = 0$, se tiene:

$$\begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix} = A \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}.$$

Y, al ser A regular basta con multiplicar ambos miembros por A^{-1} para obtener $a_1 = 0, \cdots = a_n = 0$.

Ejemplo 2.159. Una de las aplicaciones usuales del resultado anterior es realizar con facilidad cambios de base en K^n a través de la base usual B_u, siempre que se sepa calcular la matriz inversa.

Por ejemplo, sea B una base de K^n y supongamos que queremos calcular $M(I,B \leftarrow B_u)$, el cambio de base de coordenadas de B_u a B. Por el corolario previo tenemos $M(I,B \leftarrow B_u) = M(I,B_u \leftarrow B)^{-1}$. La ventaja de $M(I,B_u \leftarrow B)$ frente a $M(I,B \leftarrow B_u)$ es que su cálculo suele ser inmediato, pues basta escribir por columnas las componentes de los vectores de B (eso sí, para terminar de obtener $M(I,B \leftarrow B_u)$ necesitamos saber calcular $M(I,B_u \leftarrow B)^{-1}$). Así, el resultado del ejemplo 2.156 se reobtendría:

$$M(I,B \leftarrow B_u) = M(I,B_u \leftarrow B)^{-1} = \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix}^{-1} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix},$$

Más aún, dadas cualesquiera dos bases B,B' de K^n se puede escribir:

$$M(I,B' \leftarrow B) = M(I,B \leftarrow B_u) \cdot M(I,B_u \leftarrow B),$$

donde $M(I,B \leftarrow B_u)$ se calcula por el procedimiento anterior, y $M(I,B_u \leftarrow B)$ es inmediata.

2.4. Aplicaciones: revisión del rango y ecuaciones de un subespacio

Terminaremos este tema aprovechando la teoría de espacios vectoriales que hemos estudiado para revisar algunas propiedades interesantes ya vistas en la teoría de matrices y que se pueden obtener sin necesidad del concepto de determinante. Deduciremos también algunos principios que serán de utilidad en el cálculo de bases y dimensiones de subespacios vectoriales.

2.4.1. Transformaciones elementales. Extracción de bases de un s.d.g.

Sea $V(K)$ un e.v. y $U = L(S)$ un s.v. no trivial generado por $S = \{v_1, \ldots, v_m\}$. Nuestro objetivo es revisar el significado de las transformaciones elementales y, mediante ellas, calcular una base (ordenada) de U así como $\dim_K(U)$.

Proposición 2.160. *Sea V un e.v. sobre K y $U = L(S)$ con $S = \{v_1, \ldots, v_i, \ldots, v_j, \ldots, v_m\} \subset V$. Entonces, se tiene:*

1. *Si se considera S como un conjunto ordenado (como en la nota 2.90), entonces cualquier reordenación S' de S es también un s.d.g.*

2. *Si $a \in K$ y $a \neq 0$, entonces $S'' = \{v_1, \ldots, a \cdot v_i, \ldots, v_j, \ldots, v_m\}$ es un s.d.g. de U.*

3. *Si $a \in K$ entonces $S''' = \{v_1, \ldots, v_i, \ldots, v_j + a \cdot v_i, \ldots, v_m\}$ es un s.d.g. de U.*

Demostración. El lema es consecuencia tanto de la Proposición 2.82 como de la 2.121. Usaremos la primera de ellas, aunque resulta más inmediato de la segunda.

Es obvio que, desde el punto de vista conjuntista, $S = S'$ y, por tanto[20], $L(S') = L(S) = U$. Por otro lado, si $a \in K$ y $a \neq 0$, se tiene que $S \cup \{a \cdot v_i\}$ es un s.d.g. de U. Como $v_i = a^{-1} \cdot (a \cdot v_i)$ entonces $(S \cup \{a \cdot v_i\}) - \{v_i\} = S''$ es un s.d.g. de U. Por último, dado $a \in K$, sabemos que $S \cup \{v_j + a \cdot v_i\}$ es un s.d.g. de U. Y como $v_j = (v_j + a \cdot v_i) + (-a) \cdot v_i$, entonces $(S \cup \{v_j + a \cdot v_i\}) - \{v_j\} = S'''$ es un s.d.g de U (pues la proposición 2.82(2) resulta aplicable). ∎

Veamos cómo aplicar el lema previo para calcular una base de $U = L(S)$, donde $S = \{v_1, \ldots, v_m\}$ es una familia de vectores en K^n. Sea M la matriz cuya fila i-ésima se identifica con v_i. El lema anterior nos dice que, aplicando a M transformaciones elementales por filas, obtenemos un nuevo s.d.g. de U. Procedemos como en el método de Gauss, realizando transformaciones hasta conseguir la matriz de un SEL escalonado. Entonces, las filas no nulas se identifican con una base de U, y el número de tales filas es $\dim_K(U)$.

Ejemplo 2.161. Calculemos una base y la dimensión del s.v. de $\mathbb{R}^4$ dado por $U = L(S)$, donde $S = \{(1,-2,1,0), (2,3,-2,1), (4,-1,0,1)\}$. Escribiendo los generadores por filas asociamos la matriz:

$$M = \begin{pmatrix} 1 & -2 & 1 & 0 \\ 2 & 3 & -2 & 1 \\ 4 & -1 & 0 & 1 \end{pmatrix}$$

[20] Se puede apreciar ahora que, en la definición de $L(S)$ para S finito, se fijaba implícitamente una ordenación de los elementos de S, si bien el resultado era obviamente independiente de esta ordenación por la conmutatividad de la suma.

Apliquemos transformaciones elementales por filas hasta conseguir la matriz de un SEL escalonado.

$$M \sim \begin{pmatrix} 1 & -2 & 1 & 0 \\ 0 & 7 & -4 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Como el vector nulo se puede eliminar de cualquier s.d.g., se sigue que $B = \{(1,-2,1,0),(0,7,-4,1)\}$ sigue siendo un s.d.g. Además, B es claramente l.i. (esto es una propiedad general para las filas no nulas de cualquier matriz escalonada, ¡compruébese!), por lo que B resulta ser una base y $\dim_{\mathbb{R}}(U) = 2$.

Es de señalar que las dos primeras filas de la matriz original M también proporcionan una base $B' = \{(1,-2,1,0),(2,3,-2,1)\}$ ya que B se generó a través de transformaciones elementales de estos dos vectores. El lector puede comprobar que todo esto resulta consistente con las propiedades que maneje del rango por determinantes de una matriz.

Observación 2.162. (1) Este procedimiento resulta extensible a cualquier s.d.g. finito S de un s.v. U de cualquier e.v. finitamente generado $V(K)$. Teniendo en cuenta la observación 2.144, basta con fijar una base B de $V(K)$, tomar las coordenadas de los elementos de S en esa base y considerar el subespacio en K^n generado por estas coordenadas. Así, la base que se obtiene para este subespacio proporciona las coordenadas en B de la base requerida de $L(S)$.

(2) Como comparación, el Teorema 2.106 afirma que existe una base B de U con $B \subset S$. Para construir B debemos ir eliminando de S sucesivamente un vector que se exprese como c.l. del resto de generadores. Este proceso puede ser muy tedioso, sobre todo si m es grande. En el procedimiento explicado, las transformaciones elementales permiten construir un s.d.g. de U en el que la independencia lineal sea muy sencilla de estudiar.

2.4.2. Rango de una matriz

Podemos revisar ahora la definición de rango de una matriz teniendo en cuenta lo aprendido de espacios vectoriales.

Definición 2.163. *Sea K un cuerpo y $m,n \in \mathbb{N}$. Dada una matriz $A = (a_{ij})$ en $M_{m\times n}(K)$, definimos $v_j = (a_{1j},\dots,a_{mj})$, que es el vector en K^m cuyas componentes coinciden con las entradas de la columna j-ésima de A. Definimos el* rango de A por columnas *o, simplemente,* rango de A *como:*

$$rg(A) = dim_K(U), \ con\ U = L(v_1,\dots,v_n).$$

Así, $\mathrm{rg}(A)$ resulta ser el número máximo de columnas l.i. en A cuando se interpretan como vectores de K^m. Es claro de esta definición que $\mathrm{rg}(A) \in \mathbb{N}\cup\{0\}$, $\mathrm{rg}(A) \leq n$ (al estar generado U por n vectores) y $\mathrm{rg}(A) \leq m$ (pues $U \subset K^m$). Además, $\mathrm{rg}(A) = 0$ si y sólo si $A = 0_{m\times n}$. Gracias a la proposición 2.160 y a la definición de rango deducimos lo siguiente.

Corolario 2.164. *El rango de A es invariante por transformaciones elementales por columnas de A.*

Observación 2.165. Las transformaciones elementales por columnas de A pueden verse como transformaciones elementales por filas de A^t. Veremos con detalle en la sección 4.5 que los rangos de A y A^t son iguales, por lo que se podrán realizar transformaciones elementales por filas y transformaciones elementales por columnas sin cambiar el rango de una matriz[21]. El caso particular de este resultado para matrices regulares se verá en la proposición 2.168.

[21] Esta libertad para el cálculo del rango contrasta con la necesidad de hacer las transformaciones elementales sólo por filas (o solo por columnas, si hubiéramos dispuesto los vectores por columnas) en el procedimiento de extraer una base de un s.d.g. visto entre la proposición 2.160 y el ejercicio 2.161. De hecho, si allí se llevaran a cabo transformaciones elementales por columnas cambiaríamos los vectores del s.d.g.

Ejercicio 2.166. *Calcular, en función del parámetro $a \in \mathbb{R}$, el rango de la matriz:*

$$A = \begin{pmatrix} a & 1 & a \\ 1 & 1 & 1 \end{pmatrix}.$$

Ahora podemos dar caracterizaciones de los conceptos fundamentales de la Sección 2.3 a través de condiciones sobre el rango.

Proposición 2.167. *Sea $S = \{v_1, \ldots, v_n\}$ una familia de vectores en K^m. Consideremos la matriz en $M_{m\times n}(K)$ dada por $A = \big((v_1)_{B_u} \,|\, (v_2)_{B_u} \,|\, \ldots \,|\, (v_n)_{B_u}\big)$. Entonces:*

(i) S es un s.d.g. de K^m (necesariamente entonces $n \geq m$) si y sólo si $\mathrm{rg}(A) = m$.

(ii) S es una familia l.i. en K^m (necesariamente entonces $n \leq m$) si y sólo si $\mathrm{rg}(A) = n$.

(iii) S es una base de K^m si y sólo si $n = m$ y $\mathrm{rg}(A) = n$.

Demostración. Sabemos que $\mathrm{rg}(A) = \dim_K(U)$ con $U = L(S)$. Si S es s.d.g. de K^m entonces $U = L(S) = K^m$ y $\mathrm{rg}(A) = m$. Recíprocamente, si $\mathrm{rg}(A) = m$, entonces $\dim_K(U) = m = \dim_K(K^m)$. Como U es un s.v. de K^m deducimos que $L(S) = U = K^m$. Esto demuestra (i).

Si S es l.i. entonces S es base de U y, por tanto, $\mathrm{rg}(A) = \dim_K(U) = n$. Recíprocamente, si $\mathrm{rg}(A) = n$, entonces $\dim_K(U) = n$. Como S es un s.d.g. de U con exactamente n vectores, el corolario 2.119 nos dice que S es un base de U. En particular, S es l.i. Esto demuestra (ii).

El apartado (iii) es consecuencia directa de (i) y (ii). ∎

2.4.3. Matrices regulares

Como una aplicación, caractericemos de modo sencillo a las matrices regulares.

Proposición 2.168. *Para toda matriz cuadrada $A \in M_n(K)$ equivale:*

(i) $\mathrm{rg}(A) = n$ *(esto es, sus n columnas son independientes como elementos de K^n).*

(ii) A es la matriz de cambio de base $M(I, B_u \leftarrow B)$ donde B es el subconjunto de K^n formado por las columnas de A con su ordenación natural.

(iii) A es una matriz regular, esto es, admite inversa.

(iv) A^t es una matriz regular.

(v) Las n columnas de A^t son independientes como elementos de K^n, esto es, el rango de A por filas es n.

Demostración. (i) $\Leftrightarrow$ (ii). Inmediata de la proposición 2.167.

(ii) $\Rightarrow$ (iii). La inversa de A es $M(I, B \leftarrow B_u)$ por la proposición 2.157(ii).

(iii) $\Rightarrow$ (i). Igualemos una c.l. de las columnas de A a 0, esto es, sean $a_1, \ldots, a_n \in K$ tales que:

$$\begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix} = a_1 \begin{pmatrix} a_{11} \\ \vdots \\ a_{n1} \end{pmatrix} + \ldots a_n \begin{pmatrix} a_{1n} \\ \vdots \\ a_{nn} \end{pmatrix} = A \cdot \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}$$

Como, por hipótesis, A admite inversa, basta con multiplicar el primer y el último miembro por A^{-1} para obtener $a_1 = 0, \ldots, a_n = 0$.

Esto concluye la equivalencia de las tres primeras hipótesis.

(iii) $\Leftrightarrow$ (iv). Para la implicación hacia la derecha, si A admite inversa entonces A^t debe tener por inversa $(A^t)^{-1} = (A^{-1})^t$ ya que, aplicando las propiedades elementales del producto de matrices: $A^t \cdot (A^{-1})^t = (A^{-1} \cdot A)^t = I_n^t = I_n$ y análogamente $(A^{-1})^t \cdot A^t = I_n$. Para la implicación hacia la izquierda, si A^t es regular entonces, usando la implicación anterior, también es regular $(A^t)^t = A$.

(iv) $\Leftrightarrow$ (v) Aplíquese (i) $\Leftrightarrow$ (iii) a A^t. ∎

Observación 2.169. (1) En la demostración anterior, la base B_u y el espacio vectorial $K^n(K)$ no representan ningún papel especial. Esto es, se puede escoger cualquier otro e.v. $V(K)$ de dimensión n así como cualquier base ordenada suya B, y los pasos de la demostración funcionarían igual reemplazándolos por $K^n(K)$ y B_u (recuérdese la observación 2.144).

(2) Merece la pena recordar que el grupo lineal general $GL(n,K)$ formado por todas las matrices regulares de orden n sobre K gozaba de estructura de grupo para el producto matricial. La proposición 2.157 permite ahora reinterpretar las propiedades de ese grupo a partir de las matrices cambio de base.

Ejercicio 2.170. *Estudiar si el conjunto* $S = \{(1,0,-1),(1,2,5),(-2,1,3)\}$ *es una base de* $\mathbb{R}^3$.

2.4.4. Ecuaciones paramétricas de un subespacio

Sisstemas de generadores y ecuaciones paramétricas

Sea U un s.v. de K^n con $\dim_K(U) = k \geq 1$. Tomemos una base $B_U = (u_1, u_2, \ldots, u_k)$ de U. Como $U = L(B_U)$, los vectores de U son exactamente de la forma $v = \lambda_1 \cdot u_1 + \lambda_2 \cdot u_2 + \ldots + \lambda_k \cdot u_k$ con $\lambda_i \in K$, para cada $i = 1, \ldots, k$. Escribiendo explícitamente $v = (x_1, x_2, \ldots, x_n)$ y $u_j = (c_{1j}, c_{2j}, \ldots, c_{nj})$ entonces la igualdad $v = \lambda_1 \cdot u_1 + \ldots + \lambda_k \cdot u_k$ se escribe

$$\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \lambda_1 \begin{pmatrix} c_{11} \\ c_{21} \\ \vdots \\ c_{n1} \end{pmatrix} + \cdots + \lambda_k \begin{pmatrix} c_{1k} \\ c_{2k} \\ \vdots \\ c_{nk} \end{pmatrix}$$

o, tomando las componentes de las filas:

$$\begin{aligned} x_1 &= c_{11} \cdot \lambda_1 + c_{12} \cdot \lambda_2 + \ldots + c_{1k} \cdot \lambda_k, \\ x_2 &= c_{21} \cdot \lambda_1 + c_{22} \cdot \lambda_2 + \ldots + c_{2k} \cdot \lambda_k, \\ &\vdots \qquad\qquad \vdots \qquad\qquad \vdots \\ x_n &= c_{n1} \cdot \lambda_1 + c_{n2} \cdot \lambda_2 + \ldots + c_{nk} \cdot \lambda_k, \end{aligned} \tag{2.11}$$

que son unas *ecuaciones paramétricas* para U. Nótese que hay n ecuaciones paramétricas con k parámetros $\lambda_1, \ldots, \lambda_k$. Los coeficientes que acompañan a cada λ_i son las coordenadas de u_i en la base usual B_u. Cuando damos cada parámetro un valor concreto, obtenemos un vector de U, y los valores de los parámetros son las coordenadas de ese vector en la base original B_U. En particular, el vector i-ésimo de la base B_U se recupera de las ecuaciones paramétricas sin más que tomar $\lambda_1 = \cdots = \lambda_{i-1} = 0$, $\lambda_i = 1$, $\lambda_{i+1} = \cdots = \lambda_k = 0$.

Obsérvese que si B_U sólo fuera un sistema de generadores de U podríamos escribir expresiones análogas. En este caso se dice que los parámetros no son independientes, y distintas elecciones de ellas pueden generar una mismo vector. Pasar de esas ecuaciones a unas con parámetros independientes (esto es, *ecuaciones paramétricas* según nuestra definición) se reduce a extraer una base del s.d.g. y expresar las ecuaciones en función exclusivamente de esa base.

Con más generalidad, podemos considerar un espacio vectorial $V(K)$ en el que se ha fijado previamente una base B, y un subespacio U del que se toma una base B_U (cuyos vectores no tienen por qué tener una relación particular con los de B). En este caso, las coordenadas de los vectores de U en esa base se expresan mediante la igualdad (2.11), donde ahora los (c_{ij}) denotan, para cada j las coordenadas del vector u_j en[22] B.

Las soluciones de un SEL homogéneo como ecuaciones paramétricas

Consideremos un SEL homogéneo:

$$A \cdot x = 0, \qquad A \in M_{m\times n}(K), \quad x \in M_{n\times 1}(K) \equiv K^n.$$

Nuestro estudio de SEL usando transformaciones elementales se puede resumir (compruébese) en:

(1) Las soluciones del SEL homogéneo forman un subespacio vectorial de K^n.

(2) Las transformaciones elementales de ecuaciones del SEL producen nuevos SEL equivalentes.

(3) Si se suprime del SEL una ecuación cuyos coeficientes sean combinación lineal del resto se obtiene un SEL equivalente.

(4) El SEL original es equivalente a uno escalonado de $r \in \{0,\dots,m\}$ ecuaciones, ninguna de ellas idénticamente nulas. En este SEL escalonado los coeficientes por filas son independientes y r coincide con el rango (por filas) de la matriz original A.

(5) Este SEL escalonado permite distinguir entre r incógnitas principales, las cuales consideraremos a continuación que son las r primeras $x_1,\dots,x_r$ (sin pérdida de generalidad pues en caso contrario, se podrían renombrar las incógnitas) y $n-r$ secundarias, $x_{r+1},\dots,x_n$.

(6) Todas las soluciones se construyen tomando: (a) valores arbitrarios para las incógnitas secundarias, las cuales se consideran entonces como parámetros $x_{r+1} = \lambda_1,\dots,x_n = \lambda_{n-r} \in K$, y (b) para las incógnitas principales, combinaciones lineales apropiadas de esos parámetros:

$$\begin{array}{c} x_1 = a_{11}\lambda_1 + \cdots + a_{1(n-r)}\lambda_{n-r} \\ \vdots \\ x_r = a_{r1}\lambda_1 + \cdots + a_{r(n-r)}\lambda_{n-r} \end{array} \qquad \text{esto es,} \qquad \begin{pmatrix} x_1 \\ \vdots \\ x_r \end{pmatrix} = \lambda_1 \begin{pmatrix} a_{11} \\ \vdots \\ a_{r1} \end{pmatrix} + \cdots + \lambda_{n-r} \begin{pmatrix} a_{1(n-r)} \\ \vdots \\ a_{n-r} \end{pmatrix}$$

(7) Las soluciones equivalen a las ecuaciones paramétricas de un subespacio de K^n de dimensión $n-r$:

$$\begin{array}{rcl} x_1 & = & a_{11}\lambda_1 + \cdots + a_{1(n-r)}\lambda_{n-r} \\ & \vdots & \\ x_r & = & a_{r1}\lambda_1 + \cdots + a_{r(n-r)}\lambda_{n-r} \\ x_{r+1} & = & \lambda_1 \\ & \vdots & \\ x_n & = & \lambda_{n-r} \end{array}.$$

2.4.5. Ecuaciones implícitas de un subespacio

Cuando un subespacio U de K^n viene dado por un SEL homogéneo $Ax = 0$ de m ecuaciones que sean linealmente independientes (esto es, el rango -por filas- de $(A) = m$) diremos que este SEL es un *sistema de ecuaciones cartesianas (o implícitas) para* U. Si se considera un espacio vectorial $V^n(K)$

[22] Recuérdese el isomorfismo $f_B : V \to K^n$, que establece una biyección entre vectores (así como s.d.g., conjuntos l.i. y subespacios vectoriales) de $V(K)$ y de $K^n(K)$ establecido en la observación 2.144.

en el que se ha fijado una base B y las coordenadas de los vectores de U en esa base vienen dadas por un SEL homogéneo en K^n de ecuaciones linealmente independientes, diremos que este SEL es un *sistema de ecuaciones implícitas en la base* B *para* U o, simplemente, unas *ecuaciones implícitas*[23] *de* U *en* B. Dadas unas ecuaciones implícitas de U en alguna base B, para calcular unas ecuaciones paramétricas basta con resolverlo, como se explicó en el apartado anterior. Describimos a continuación cómo solucionar el problema inverso.

Sea $B_U = (u_1, \ldots, u_k)$ una base de U, el cual es un s.v. de un e.v. $V(K)$, del cual se tiene una base B prefijada. Consideremos la matriz $C \in M_{n\times m}(K)$:

$$C = \big((u_1)_B \,|\, (u_2)_B \,|\, \ldots \,|\, (u_k)_B\big).$$

Sea $v \in V$ y escribamos $v_B = x = (x_1, \ldots, x_n)^t$. Como B_U es una base de U, deducimos que $v \in U$ si y sólo si el rango por columnas de $(C|x)$ es k. Las ecuaciones implícitas tienen la misión de asegurar que la línea de las incógnitas es combinación lineal de las de los vectores de B_U. Describamos con ejemplos algunos modos de asegurar esta propiedad.

Ejemplo 2.171. *(Cálculo de ecuaciones implícitas usando transformaciones elementales.)*[24]

(A) Calculemos las ecuaciones cartesianas para la recta $U = L((1,3,-2)) \subset \mathbb{R}^3$ (en la base usual B_u). Claramente, $B_U = \{(1,3,-2)\}$ es una base de U y la discusión anterior asegura que el sistema debe tener asociadas $n-k = 3-1 = 2$ ecuaciones cartesianas. Un vector $v = (x,y,z)$ de $\mathbb{R}^3$ pertenecerá a U si y sólo si es proporcional a $(1,3,-2)$. Esto equivale a que la matriz dada por[25]

$$\begin{pmatrix} C \\ \hline x \end{pmatrix} = \left(\begin{array}{ccc} 1 & 3 & -2 \\ \hline x_1 & x_2 & x_3 \end{array}\right)$$

cumpla rg $\left(\frac{C}{x}\right) = 1$. Para ello podemos escalonar,

$$\begin{pmatrix} C \\ \hline x \end{pmatrix} = \left(\begin{array}{ccc} 1 & 3 & -2 \\ \hline x_1 & x_2 & x_3 \end{array}\right) \sim \left(\begin{array}{ccc} 1 & 3 & -2 \\ \hline 0 & -3x_1 + x_2 & 2x_1 + x_3 \end{array}\right)$$

(multiplicando la primera fila por x y restándosela a la segunda). Esta matriz tendrá rango 1 si y sólo si la segunda fila es 0, esto es,

$$\boxed{\left.\begin{array}{r} -3x_1 + x_2 = 0 \\ 2x_1 + x_3 = 0 \end{array}\right\}} \tag{2.12}$$

(B) Ecuaciones cartesianas para el plano $W = L((1,3,-2),(1,1,1)) \subset \mathbb{R}^3$ (en la base usual B_u). Claramente, $B_W = \{(1,3,-2),(1,1,1)\}$ es l.i. y, por tanto, una base de W, por lo que el sistema debe

[23] Usualmente el término de ecuaciones *implícitas* se considera en este ambiente general, mientras que el de *cartesianas* se reserva para el caso particular de K^n con la base usual B_u, como hemos hecho en esta sección.

[24] Más adelante se describe el procedimiento de cálculo usando determinantes, que puede ser el más conveniente desde el punto de vista práctico general. Por completitud, en esta nota se describe otro procedimiento, cuyo significado geométrico se entenderá mejor tras estudiar el espacio dual.
Cada ecuación cartesiana es del tipo $ax_1 + bx_2 + cx_3 = 0$. Imponemos que el vector $(1,3,-2)$ la cumpla; esto nos lleva a la ecuación lineal $a + 3b - 2c = 0$, cuyo s.v. de soluciones U^* tiene dimensión 2. Una base de U^* viene dada por $B^* = \{(-3,1,0),(2,0,1)\}$. Así, las ecuaciones cartesianas que se obtienen para U son de nuevo $-3x_1 + x_2 = 0$ y $2x_1 + x_3 = 0$.

[25] Por supuesto, es irrelevante trabajar por columnas, como hacíamos al escribir coordenadas en la matriz $(C|x)$ antes, que por filas, como haremos ahora al escribir directamente los elementos de $K^n = \mathbb{R}^3$ en la matriz $\left(\frac{C}{x}\right)$.

tener asociadas $n-k=3-2=1$ ecuación cartesiana. Un vector $v=(x_1,x_2,x_3)$ de $\mathbb{R}^3$ pertenecerá W si es combinación lineal de B_W. Esto equivale a que la matriz dada por:

$$\left(\frac{C}{x}\right)=\left(\begin{array}{ccc}1 & 3 & -2\\ 1 & 1 & 1\\ \hline x_1 & x_2 & x_3\end{array}\right)$$

cumpla rg $\left(\frac{C}{x}\right)=2$. Para ello escalonamos

$$\left(\begin{array}{ccc}1 & 3 & -2\\ 1 & 1 & 1\\ \hline x_1 & x_2 & x_3\end{array}\right)\underset{\substack{-F_1+F_2\\(-x)F_1+F_3}}{\sim}\left(\begin{array}{ccc}1 & 3 & -2\\ 0 & -2 & 3\\ \hline 0 & -3x_1+x_2 & 2x_1+x_3\end{array}\right)\underset{(-3x+y)F_2+2F_3}{\sim}\left(\begin{array}{ccc}1 & 3 & -2\\ 0 & -2 & 3\\ \hline 0 & 0 & -5x_1+3x_2+2x_3\end{array}\right)$$

Esta matriz tendrá rango 2 si y sólo si la tercera fila es 0, esto es,

$$\boxed{-5x_1+3x_2+2x_3=0.} \tag{2.13}$$

Observación 2.172. Dados dos s.v. U,W de $V(K)$, podemos dar procedimientos generales para construir $U+W$ y $U\cap W$:

Suma $U+W$. Si se tienen ecuaciones paramétricas de U y W resulta inmediato construir bases B_U, B_W de cada subespacio. Entonces $B_U\cup B_W$ es un s.d.g. de $U+W$ y se puede extraer de esta unión una base (y construir unas ecuaciones paramétricas) de $U+W$.

Intersección $U\cap W$. Si se tienen ecuaciones implícitas de U y de W en una misma base B, el SEL formado por todas estas ecuaciones tiene por solución $U\cap W$, por lo que de este SEL se puede extraer unas ecuaciones implícitas de $U\cap W$.

No obstante, a la hora de la práctica otros procedimientos pueden ser más convenientes. Por ejemplo, supongamos que se dispone de una base $B_U=(u_1,\dots,u_r)$ de $V(K)$ y unas ecuaciones implícitas de W en una base B. Un vector arbitrario u de U se escribe entonces $u=x_1u_1+\cdots+x_ru_r$, donde $x_1,\dots x_r\in K$ y tendrá unas coordenadas u_B que dependen de los escalares x_i. Imponiendo que estas coordenadas u_B satisfagan las m ecuaciones implícitas de W se obtiene un SEL de m ecuaciones y r incógnitas, del que se pueden extraer unas ecuaciones implícitas para $U\cap W$.

Cálculo práctico de las ecuaciones implícitas usando determinantes.

El procedimiento de cálculo del rango de una matriz proporciona un método sistemático para el cálculo de las ecuaciones implícitas. En resumen, nos enfrentamos a los dos problemas siguientes, que resultan esencialmente equivalentes. Sea $V(K)$ un e.v. finitamente generado y U un subespacio vectorial suyo:

1. Dada una base de $B_U=(u_1,\dots,u_k)$, de cuyos elementos se conocen sus coordenadas en una base $B=(v_1,\dots,v_n)$ de $V(K)$ calcular unas ecuaciones implícitas para U en la base B.

   Procedimiento. Se escribe la matriz $A=((u_1)_B|\dots|(u_k)_B)$ y se halla un menor principal (que tendrá orden k, al ser B_U linealmente independiente). Se impone entonces que tenga rango k la matriz $(x|A)$, donde la columna añadida $x:=(x_1,\dots x_n)^t$ representa las n incógnitas (para las ecuaciones implícitas). Para conseguir esto, se orla el menor principal con los elementos de la columna añadida, y se impone que los $n-k$ menores así obtenidos sean iguales a 0 (generándose así las $n-k$ ecuaciones implícitas, necesariamente independientes, buscadas).

2. Dada unas ecuaciones paramétricas de U en la base B, calcular unas ecuaciones implícitas de U en esa base.

   Procedimiento. Sabemos que, si $\lambda_1,\dots,\lambda_k$ son los parámetros, podemos obtener una base de U mediante k elecciones independientes de ellos (canónicamente $(\lambda_1 = 1, \lambda_2 = 0, \dots, \lambda_{k-1} = 0, \lambda_k = 0), \dots, (\lambda_1 = 0, \lambda_2 = 0, \dots, \lambda_{k-1} = 0, \lambda_k = 1)$, lo que reduce el problema al caso anterior.

Ejemplo 2.173. Consideremos en $\mathbb{R}^3$ el subespacio U de ecuaciones paramétricas en la base usual:

$$\begin{array}{rcrl} x_1 & = & \lambda_1 & +2\lambda_2 \\ x_2 & = & \lambda_1 & -\lambda_2 \\ x_3 & = & -\lambda_1 & +2\lambda_2, \end{array}$$

y construyamos las implícitas en esa misma base. Tomando las elecciones canónicas de los parámetros, encontramos una base B_U de U, que proporciona la matriz A de coordenadas en la base usual:

$$A = \begin{pmatrix} 1 & 2 \\ 1 & -1 \\ -1 & 2 \end{pmatrix} \qquad \text{y construimos entonces} \qquad (x|A) = \begin{pmatrix} x_1 & 1 & 2 \\ x_2 & 1 & -1 \\ x_3 & -1 & 2 \end{pmatrix}$$

La submatriz formada por las dos primeras filas y columnas de A nos proporciona un menor principal de A. Orlando este menor en $(x|A)$ (de la única manera posible) se tiene la (única) ecuación implícita que define U:

$$0 = \begin{vmatrix} x_1 & 1 & 2 \\ x_2 & 1 & -1 \\ x_3 & -1 & 2 \end{vmatrix} = x_1 - 4x_2 - 3x_3$$

(obsérvese que el determinante se computa con facilidad desarrollando por los adjuntos de la columna de las incógnitas).

Ejemplo 2.174. Consideremos ahora el ejemplo 2.171, y calculemos las ecuaciones implícitas usando rangos y determinantes en los dos casos (A) y (B) allí considerados.

(A) Razonando como en ese ejemplo, vamos a asegurar ahora rg $\left(\frac{C}{x}\right) = 1$ usando determinantes. Para ello observamos que su elemento $(1,1)$ proporciona un menor principal de orden 1. Orlamos este menor e imponemos que los menores de orden 2 así obtenidos sean nulos, esto es:

$$\begin{vmatrix} 1 & 3 \\ x_1 & x_2 \end{vmatrix} = 0, \qquad \begin{vmatrix} 1 & -2 \\ x_1 & x_3 \end{vmatrix} = 0.$$

Desarrollando explícitamente los determinantes se obtiene un SEL equivalente al antes obtenido en (2.12) (de hecho, es el mismo SEL salvo reordenación del las incógnitas).

(B) Razonando análogamente, vamos a asegurar ahora rg $\left(\frac{C}{x}\right) = 2$ teniendo en cuenta que la submatriz cuadrada formada por las primeras dos filas y columnas nos da un menor principal de orden 2. Orlamos este menor e imponemos que el único menor de orden 3 así obtenido (esto es, el determinante) es nulo:

$$0 = \begin{vmatrix} 1 & 3 & -2 \\ 1 & 1 & 1 \\ x_1 & x_2 & x_3 \end{vmatrix} = 5x_1 - 3x_2 - 3x_3,$$

que es equivalente a la ecuación implícita (2.13) obtenida antes. Desde un punto de vista práctico, el cálculo de la ecuación implícita se simplifica desarrollando el determinante por los adjuntos de la línea de las incógnitas.

Ejemplo 2.175. Sea considera en U el siguiente subespacio vectorial generado por cuatro parámetros, de los que *no* se presupone que sean independientes:

$$\begin{array}{rcllll} x_1 & = & -\lambda_1 & +2\lambda_2 & +3\lambda_3 & +2\lambda_4 \\ x_2 & = & \lambda_1 & -2\lambda_2 & +\lambda_3 & +2\lambda_4 \\ x_3 & = & 2\lambda_1 & -4\lambda_2 & -\lambda_3 & +\lambda_4 \\ x_4 & = & 2\lambda_1 & -4\lambda_2 & +3\lambda_3 & +5\lambda_4 \end{array}$$

Nuestro objetivo será calcular unas ecuaciones implícitas para U (en la base usual). No obstante, como no se presupone que los parámetros sean independientes, comprobaremos en primer lugar el rango de la matriz obtenida mediante las elecciones canónicas de los parámetros (las cuales ahora sólo proporcionarán un sistema de generadores de U), esto es:

$$\begin{pmatrix} -1 & 2 & 3 & 2 \\ 1 & -2 & 1 & 2 \\ 2 & -4 & -1 & 1 \\ 2 & -4 & 3 & 5 \end{pmatrix} \tag{2.14}$$

Claramente, la segunda columna es proporcional a la primera[26]. Suprimiendo ésta, tomamos como menor no nulo $\begin{vmatrix} -1 & 3 \\ 1 & 1 \end{vmatrix} \neq 0$. Orlándolo, vemos que el rango es dos, puesto que:

$$\begin{vmatrix} -1 & 3 & 2 \\ 1 & 1 & 2 \\ 2 & -1 & 1 \end{vmatrix} = 0, \qquad y \qquad \begin{vmatrix} -1 & 3 & 2 \\ 1 & 1 & 2 \\ 2 & 3 & 5 \end{vmatrix} = 0.$$

En consecuencia, sólo los parámetros λ_1, λ_3 son independientes (en el sentido de que las correspondientes columnas primera y tercera de la matriz (2.14) son independientes, y la segunda y cuarta se pueden calcular como combinación linea de ellas). Para calcular las ecuaciones implícitas requeridas debemos asegurar que el rango de la matriz

$$\begin{pmatrix} x_1 & -1 & 3 \\ x_2 & 1 & 1 \\ x_3 & 2 & -1 \\ x_4 & 2 & 3 \end{pmatrix}.$$

es dos. Orlando el menor principal antes escogido, se obtienen las ecuaciones implícitas:

$$0 = \begin{vmatrix} x_1 & -1 & 3 \\ x_2 & 1 & 1 \\ x_3 & 2 & -1 \end{vmatrix} = -3x_1 + 5x_2 - 4x_3, \qquad 0 = \begin{vmatrix} x_1 & -1 & 3 \\ x_2 & 1 & 1 \\ x_4 & 2 & 3 \end{vmatrix} = x_1 + 9x_2 - 4x_4.$$

[26] El lector puede ejercitarse en el cálculo del rango sin tener en cuenta este hecho.

2.5. Espacios afines: independencia afín y sistemas de referencia

A continuación extenderemos espacios afines algunos de los conceptos estudiados.

2.5.1. Dimensión

Definición 2.176. *Dado un espacio afín $\mathcal{A}$, su* dimensión *dim $\mathcal{A}$ es la dimensión del espacio vectorial asociado $\vec{\mathcal{A}}$.*

En particular, para un subespacio afín $\mathcal{S}$ de $\mathcal{A}$ se tiene $\dim \mathcal{S} \leq \dim \mathcal{A}$. y, como consecuencia de las proposiciones 2.68 y 2.71:

Corolario 2.177. *Si $\mathcal{S}, \mathcal{T}$ son subespacios afines de $\mathcal{A}$, entonces:*

1. *Si $\mathcal{S} \cap \mathcal{T} \neq \emptyset$ entonces:* $\dim(\mathcal{S} \cap \mathcal{T}) = \dim(\vec{\mathcal{S}} \cap \vec{\mathcal{T}})$ *y* $\dim(\mathcal{S} + \mathcal{T}) = \dim(\vec{\mathcal{S}} + \vec{\mathcal{T}})$.

   Por tanto (al menos en dimensión finita):

   $\dim(\mathcal{S} + \mathcal{T}) = \dim \mathcal{S} + \dim \mathcal{T} - \dim(\mathcal{S} \cap \mathcal{T})$.

2. *Si $\mathcal{S} \cap \mathcal{T} = \emptyset$, entonces:* $\dim(\mathcal{S} + \mathcal{T}) = \dim(\vec{\mathcal{S}} + \vec{\mathcal{T}}) + 1$

   Por tanto (al menos en dimensión finita):

   $\dim(\mathcal{S} + \mathcal{T}) = \dim \mathcal{S} + \dim \mathcal{T} - \dim(\vec{\mathcal{S}} \cap \vec{\mathcal{T}}) + 1$.

2.5.2. Independencia afín

Sean $k+1$ puntos $\{P_0, P_1, \ldots, P_k\}$, $k \geq 0$, de un espacio afín $\mathcal{A}$, y denotemos $\mathcal{S} = \langle P_0, P_1, \ldots, P_k \rangle$ al subespacio afín que generan (definición 2.70) y $\vec{\mathcal{S}}$ a su espacio vectorial asociado. Sea $\vec{\mathcal{S}}_0 = L\{\overline{P_0P_1}, \ldots, \overline{P_0, P_k}\}$.

Proposición 2.178. *Con las definiciones anteriores, $\mathcal{S} = P_0 + \vec{\mathcal{S}}_0$ y, por tanto:*

1. $\vec{\mathcal{S}} = \vec{\mathcal{S}}_0$

2. $\vec{\mathcal{S}} = L\{\overline{P_0P_1}, \ldots, \overline{P_0P_k}\} = L\{\overline{P_iP_0}, \overline{P_iP_1} \ldots, \overline{P_iP_{i-1}}, \overline{P_iP_{i+1}} \ldots, \overline{P_i, P_k}\}$ *para todo $i = 1, \ldots, k$.*

3. *Equivalen:*

   a) $\{\overline{P_0P_i} : i = 1, 2, \ldots, k\}$ *es linealmente independiente.*

   b) *Para cada $j \in \{1, \ldots, k\}$, es linealmente independiente el conjunto* $\{\overline{P_jP_i} : i \in \{0, \ldots, j-1, j+1, \ldots k\}\}$

   c) *dim$\vec{\mathcal{S}} = k$.*

Demostración. Claramente, $P_0 + \vec{\mathcal{S}}_0$ es un subespacio afín que contiene a $\{P_0, P_1, \ldots, P_k\}$ y todo subespacio afín que contiene a este conjunto debe contener a $P_0 + \vec{\mathcal{S}}_0$, por lo que $P_0 + \vec{\mathcal{S}}_0 = \mathcal{S}$. De ahí resultan inmediatas las afirmaciones 1 y 2. Para 3, son inmediatas las equivalencias de a) y c) y de b) y c). ∎

Definición 2.179. *Sea $\mathcal{A}$ un espacio afín. Un conjunto de $k+1$ puntos de $\mathcal{A}$, $\{P_0, \ldots, P_k\}$, $k \in \mathbb{N} \cup \{0\}$, se dice que es* afínmente independiente *si el subespacio afín que generan $\langle P_0, P_1, \ldots, P_k \rangle$ tiene dimensión k. En caso contrario, el conjunto es* afínmente independiente.

Desde un punto de vista práctico, la proposición anterior nos dice que esto equivale a que $\{\overline{P_0P_1},\ldots,\overline{P_0,P_k}\}$ sea linealmente independiente.

Ejercicio 2.180. Sean $P,Q \in \mathcal{A}$. Demostrar:

- Si $P \neq Q$, entonces $\langle P,Q\rangle$ es un subespacio afín de dimensión 1, y es el único subespacio afín de dimensión 1 que los contiene.
- $\langle P,Q\rangle = \langle Q,P\rangle$.
- Si $R \in \langle P,Q\rangle$ y R es distinto de P y de Q entonces $\langle P,R\rangle = \langle P,Q\rangle$.
- Si $A,B,C \in \langle P,Q\rangle$ entonces $\overline{AC},\overline{AB}$ son linealmente dependientes.

En particular, si la dimensión n de $\mathcal{A}$ es finita y se tienen $n+1$ puntos independientes, fijado uno de ellos se genera una base del espacio vectorial director, lo que analizaremos en la próxima sección.

2.5.3. Sistemas de referencia afines

Definición 2.181. *Sea $\mathcal{A}$ un espacio afín. Un* sistema de referencia afín *o, simplemente,* referencia afín*, es el par $\mathcal{R} = (0,B)$ formado por un punto $0 \in \mathcal{A}$ al que llamaremos* origen *y una base ordenada B del espacio vectorial asociado $\vec{\mathcal{A}}$.*

Fijado el sistema de referencia afín, para cada $P \in \mathcal{A}$ el vector $\overline{OP}$ se puede expresar de manera única como combinación lineal de elementos de B.

Nos restringiremos a dimensión finita $n \in \mathbb{N}$, donde si $B = (v_1,\ldots,v_n)$ era equivalente dar un conjunto ordenado $(O,P_1,\ldots,P_n)$ de $n+1$ puntos afínmente independientes relacionados por la igualdad

$$P_i = O + v_i, \qquad \forall i = 1,\ldots,n.$$

Para cada punto $P \in \mathcal{A}$, la única n-úpla de escalares $P_{\mathcal{R}} = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} \in \mathbb{R}^n$ tales que

$$\overline{OP} = \sum_{i=1}^{n} a_i\, v_i = \sum_{i=1}^{n} a_i \overline{OP_i},$$

se le llama. *n-úpla de coordenadas afines* o, simplemente, *coordenadas afines* de Q respecto del sistema de referencia $\mathcal{R}$.

Teniendo en cuenta las propiedades conocidas de los SEL (que se resumían en el Teorema de Rouché Frobenius), y el estudio de las eucaciones implícitas de subespacios vectoriales, resulta fácil demostrar.

Proposición 2.182. Ecuaciones de un subespacio afín. *Se considera un sistema de referencia afín $\mathcal{R} = (O,P_1,\ldots,P_n)$ de un espacio afín $\mathcal{A}$.*

1. *Sea $M \in \mathcal{M}_{m\times n}(\mathbb{R})$ con $m \leq n$ y $\operatorname{rango}(M) = m$, y un vector (columna) $b \in \mathbb{R}^m$ tal que el SEL $M\cdot x = b$ es compatible. Entonces el conjunto $\mathcal{S} = \{P \in \mathcal{A} : MP_{\mathcal{R}} = b\}$ (formado por los puntos cuyas coordenadas en $\mathcal{R}$ son soluciones del SEL $M\cdot x = b$) es un subespacio afín de $\mathcal{A}$ y verifica: $\dim \mathcal{S} = n-m$.*

2. *Dado un subespacio afín $\mathcal{S}$ de $\mathcal{A}$, con* $\dim \mathcal{S} = k \in \{0,\ldots,n-1\}$, *existe una matriz* $M \in \mathcal{M}_{m\times n}(\mathbb{R})$ *con* $m = n-k < n$ *y* $\operatorname{rango}(M) = m$, *y un vector* $b \in \mathbb{R}^m$ *tales que* $\mathcal{S} = \{P \in \mathcal{A} : MP_{\mathcal{R}} = b\}$.

3. *En cualquiera de los dos casos anteriores, las ecuaciones implícitas del subespacio director, respecto de la base* $(\overline{OP}_1,\ldots,\overline{OP}_n)$ *son* $MP_{\mathcal{R}} = 0$.

2.5.4. Paralelismo en subespacios afines

Definición 2.183. *Sea $\mathcal{A}$ un espacio afín y sean $\mathcal{S}$ y $\mathcal{S}'$ dos subespacios afines de $\mathcal{A}$. Se dice que $\mathcal{S}$ es* paralelo *a $\mathcal{S}'$ si el subespacio asociado de $\mathcal{S}$ está incluido en el subespacio asociado de $\mathcal{S}'$. Si los subespacios asociados son iguales, se dice que $\mathcal{S}$ y $\mathcal{S}'$* son paralelos *y se escribe $\mathcal{S}\|\mathcal{S}'$.*

Si ni $\mathcal{S}$ es paralelo a $\mathcal{S}'$ ni $\mathcal{S}'$ es paralelo a $\mathcal{S}$, diremos que los subespacios son secantes *(o* se cortan*) si $\mathcal{S}\cap\mathcal{S}' \neq \emptyset$ y que se* cruzan *en caso contrario.*

Proposición 2.184. *Sea $\mathcal{A}$ un espacio afín y sean $\mathcal{S}$ y $\mathcal{S}'$ dos subespacios afines de $\mathcal{A}$.*

1. *Si $\mathcal{S}$ es paralelo a $\mathcal{S}'$, entonces o bien $\mathcal{S} \subset \mathcal{S}'$ o bien $\mathcal{S}\cap\mathcal{S}' = \emptyset$.*

2. *Si $\mathcal{S}\|\mathcal{S}'$, entonces o bien $\mathcal{S} = \mathcal{S}'$ o bien $\mathcal{S}\cap\mathcal{S}' = \emptyset$.*

3. (Postulado de las paralelas). *Sea $P \in \mathcal{A}\backslash\mathcal{S}$. Entonces, existe un único subespacio afín $\mathcal{S}'$ que pasa por P tal que $\mathcal{S}\|\mathcal{S}'$, el cual verifica $\mathcal{S}\cap\mathcal{S}' = \emptyset$.*

Demostración. 1. Si existe $P \in \mathcal{S}\cap\mathcal{S}'$ de $\vec{\mathcal{S}} \subset \mathcal{S}'$ se sigue: $\mathcal{S} = P+\vec{\mathcal{S}} \subset P+\vec{\mathcal{S}}' = \mathcal{S}'$.

2. En la demostración anterior, si $\vec{\mathcal{S}} = \vec{\mathcal{S}}'$ entonces $\mathcal{S} = \mathcal{S}'$.

3. Claramente, $P+\vec{\mathcal{S}}$ es el único subespacio afín que satisface las condiciones de $\mathcal{S}'$. ∎

Como un ejercicio opcional, el resto de esta sección permite demostrar un teorema clásico.

Definición 2.185. *Sean $A_1, A_2, A_3 \in \mathcal{A}$ tres puntos que caen sobre una misma recta afín tales que $A_1 \neq A_2$. La* razón simple *se define como el único escalar $(A_1A_2A_3) \in K$ tal que $\overrightarrow{A_1A_3} = (A_1A_2A_3)\overrightarrow{A_1A_2}$.*

Lema 2.186. *En un espacio afín de dimensión finita $n \geq 2$, todo hiperplano (subespacio afín de dimensión $n-1$) y toda recta no paralela al hiperplano se cortan exactamente en un punto.*

Teorema 2.187. (de Thales de las paralelas). *Sea $\mathcal{A}$ un espacio afín de dimensión finita $n \geq 2$, y H_1, H_2 y H_3 tres hiperplanos distintos y paralelos. Sean r_1 y r_2 dos rectas no paralelas a H_1 y consideremos los puntos de intersección $A_i = r_1 \cap H_i$, $B_i = r_2 \cap H_i$, $i = 1,2,3$. Entonces,*

$$(A_1A_2A_3) = (B_1B_2B_3)$$

Sugerencia: en el ambiente algebraico introducido, la demostración de este teorema clásico (uno de los dos tradicionalmente atribuidos a Thales) se reduce a comprobar $B_1 + (A_1A_2A_3)\overrightarrow{B_1B_2} \in H_3$.

Tema 3

Aplicaciones lineales

En este tema, nuestro objetivo será el de estudiar aplicaciones entre distintos espacios vectoriales (sobre el mismo cuerpo) que "preserven" o "respeten" las operaciones propias de cada espacio. Estas aplicaciones, a las que llamaremos *lineales*, conllevan un *rango*, que se relacionará con el de las matrices. También permiten introducir el concepto *isomorfismo de e.v.* el cual, en el caso finitamente generado, hace resaltar la importancia de los e.v. $K^n(K)$. Interpretaremos las matrices $m \times n$ a partir de aplicaciones lineales entre espacios $K^n(K)$, $K^m(K)$, lo que nos permitirá entender mejor tanto las propiedades de las matrices como las de las aplicaciones lineales. Finalmente, ganaremos en abstracción introduciendo los *espacios cocientes*. Aparte de ser un ejemplo de espacio vectorial con interés en sí mismo, estos espacios permiten dar una demostración alternativa del teorema del rango y reinterpretar la descomposición canónica (válida para cualquier aplicación entre conjuntos) en el caso de aplicaciones lineales. En adelante, siempre consideraremos dos e.v. $V(K)$, $V'(K)$ construidos sobre el mismo cuerpo conmutativo K.

3.1. El concepto de aplicación lineal

3.1.1. Definición y primeras propiedades

Definición 3.1. *Dados dos e.v.,* $V(K), V'(K)$*, una aplicación* $f : V \to V'$ *se dice que es* lineal *u* homomorfismo de espacios vectoriales *si verifica las siguientes dos propiedades:*

(i) *Es un homomorfismo de grupos, entre* $(V,+)$ *y* $(V',+)$*, esto es:*
$f(u+v) = f(u) + f(v)$, $\forall u, v \in V$.

(ii) $f(a \cdot v) = a \cdot f(v)$, $\forall u \in V$, $\forall a \in K$.

Nota 3.2. El nombre de *homomorfismo de grupos* para la primera propiedad es general para aplicaciones entre dos grupos cualesquiera. De hecho, las propiedades que obtengamos para las aplicaciones lineales y que no dependan del producto por escalares, resultan válidas también para cualquier homomorfismo de grupos.

Podemos resumir las dos condiciones anteriores en una única:

Proposición 3.3. *Una aplicación* $f : V \to V'$ *es lineal si y sólo si verifica:*

$$f(a \cdot u + b \cdot v) = a \cdot f(u) + b \cdot f(v) \qquad \forall u, v \in V, \quad \forall a, b \in K.$$

Demostración. ($\Rightarrow$) Aplicando primero (i) y después (ii):

$$f(a \cdot u + b \cdot v) = f(a \cdot u) + f(b \cdot v) = a \cdot f(u) + b \cdot f(v)$$

($\Leftarrow$) Aplicando la condición expresada:

$$\begin{aligned} f(u+v) = &\quad f(1 \cdot u + 1 \cdot v) = 1 \cdot f(u) + 1 \cdot f(v) = f(u) + f(v) \\ f(a \cdot v) = &\quad f(a \cdot v + 0 \cdot \vec{0}) = a \cdot v + 0 \cdot f(\vec{0}) = a \cdot v \end{aligned}$$

donde 0 y $\vec{0}$ son, resp., el escalar nulo de K y el vector nulo de V. ∎

Algunas propiedades simples de las aplicaciones lineales son las siguientes:

Proposición 3.4. *Sea $f : V \to V'$ una aplicación lineal. Se verifica:*

1. $f(0) = 0$ *(esto es, $f(\vec{0}) = \vec{0}'$ donde $\vec{0}$ y $\vec{0}'$ son, resp. los vectores nulos de V y V').*
2. $f(-v) = -f(v)$, $\forall v \in V$.
3. $f(\sum_{i=1}^{k} a_i \cdot u_i) = \sum_{i=1}^{k} a_i \cdot f(u_i)$, $\forall u_i \in V$, $\forall a_i \in K$, $\forall i \in \{1, \ldots, k\}$.
4. *Si $S \subset V$ entonces*[1] $f(L(S)) = L(f(S))$.
5. *Si U es un s.v. de V, entonces $f(U)$ es un s.v. de V'. En particular, $Im(f)$ es un s.v. de V'.*
6. *Si U' es un s.v. de V' entonces*[2] $f^{-1}(U')$ *es un s.v. de V.*

Demostración. 1. Claramente, $f(0) = f(0+0) = f(0) + f(0)$, donde en la última igualdad se usa la propiedad (i) de las aplicaciones lineales. El resultado se sigue sumando $-f(0)$ a ambos miembros.

2. Para demostrar que $f(-v)$ es el opuesto de $f(v)$, operamos ambos elementos y comprobamos que se obtiene el vector cero de V': $f(v) + f(-v) = f(v + (-v)) = f(0) = 0$ (la primera igualdad por la propiedad (i), la segunda por ser $-v$ el opuesto de v, y la tercera por el punto 1 anterior).

3. Inmediato por inducción sobre el número de sumandos.

4. Si S es un conjunto finito, $S = \{u_1, \ldots, u_m\}$, el punto 3 anterior afirma que la imagen de cualquier elemento en $L(S)$ es una combinación lineal de $f(S) = \{f(u_1), \ldots, f(u_m)\}$, y viceversa, lo que prueba la doble inclusión. Si S es infinito, el problema se reduce al caso finito, porque en las combinaciones lineales de S y $L(S)$ sólo interviene un conjunto finito de vectores.

5. Aplicando el apartado anterior[3]: $f(U) = f(L(U)) = L(f(U))$, que es un s.v. Para la última afirmación, basta con aplicar la primera a $f(V) = \mathrm{Im}(f)$.

[1] Abusando de la notación, para cada $C \subset V$ denotamos por $f(C)$ a lo que, rigurosamante, es $f_*(C) := \{f(v) : v \in C\}$ (recuérdese que $f_* : \mathcal{P}(V) \to \mathcal{P}(V')$ era la aplicación imagen directa entre los conjuntos de las partes de V y V').

[2] Abusando de la notación, para cada $C' \subset V'$ denotamos por $f^{-1}(C')$ a $f^*(C') := \{v \in V : f(v) \in C'\} \subset V$ (recuérdese que $f^* : \mathcal{P}(V') \to \mathcal{P}(V)$ era la aplicación imagen recíproca). Remarquemos que f no tiene por qué ser biyectiva, esto es, f^{-1} no denota la aplicación inversa aquí.

[3] Una demostración directa alternativa es la siguiente. Sean $u', w' \in f(U)$, por lo que existen $u, w \in U$ tales que $f(u) = u', f(w) = w'$, y sean $a, b \in K$. La pertenencia de $au' + bw'$ a $f(U)$ se deriva de la linealidad de f:

$$au' + bw' = af(u) + bf(w) = f(au + bw) \in f(U),$$

donde se usa $au + bw \in U$ (que es consecuencia de que U sea un s.v. y $u, w \in U$).

6. Sean $u, w \in f^{-1}(U')$, por lo que existen $u', w' \in U'$ tales que $f(u) = u', f(w) = w'$, y sean $a, b \in K$. Usando la linealidad de f,

$$f(au+bw) = af(u) + bf(w) = au' + bw' \in U'$$

(la pertenencia de $au' + bw'$ a U', consecuencia de que U' es un s.v. y $u', w' \in U'$). Por tanto, $au + bw \in f^{-1}(U')$, como se quería demostrar. ∎

Nota 3.5. Los puntos 1 y 2 de la proposición anterior se han demostrado usando solamente la propiedad (i) en la definición de aplicación lineal. Por esto, resultan válidas para cualquier homomorfismo de grupos (véase la nota 3.2 anterior). No obstante, ambos puntos se pueden demostrar también usando sólo la propiedad (ii) (hágase como ejercicio).

Como ejemplos sencillos y generales de aplicaciones lineales se tienen:

1. La aplicación identidad $I_V : V \to V$, $I_V(v) := v$, para todo $v \in V$ (que también denotaremos simplemente I), es claramente lineal en cualquier e.v. $V(K)$. Es de remarcar que las estructuras de e.v. que soporta V en el dominio y el codominio de I_V se entiende que son la misma.

2. La aplicación nula $f_0 : V \to V'$, $f_0(v) := 0$ para todo $v \in V$, (que también denotaremos simplemente 0) es claramente lineal para todo par de e.v., $V(K)$ y $V'(K)$.

   La aplicación nula puede verse como un caso particular de aplicación constante. Concretamente, sea $v'_0 \in V'$ un vector fijo y definamos la aplicación constante igual a v'_0 como $f_{v'_0} : V \to V'$, $f_{v'_0}(v) := v'_0$, para todo $v \in V$. Claramente, de entre todas la aplicaciones constantes de V a V' la aplicación nula es la única lineal (¿por qué?).

3. La homotecia de razón $\lambda \in K \setminus \{0, 1\}$, definida por $H_\lambda : V \to V$, $H_\lambda(v) = \lambda v, \forall v \in V$, es claramente lineal[4].

4. Si $U \subset V$ es un s.v. de $V(K)$, la aplicación inclusión $i : U \to V$, $i(u) = u$ para todo $u \in U$, es claramente lineal.

5. Generalizando el ejemplo anterior, si $f : V \to V'$ es lineal y $U \subset V$ es un s.v., la restricción de f a U, $f|_U : U \to V$ es una aplicación lineal.

   El ejemplo anterior es el caso particular del presente cuando f es I_V y, por tanto $i = (I_V)|_U$.

Ejemplo 3.6. [Aplicaciones lineales y K^n] (i) Se considera en $K^n(K)$ la $i-$ésima proyección $p_i : K^n \to K$ definida por $p_i((x_1, \dots, x_n)) = x_i$, para todo $(x_1, \dots, x_n) \in K^n$. Claramente, esta aplicación es lineal.

(ii) Sea $A \in M_{m\times n}(K)$. Considerando los elementos de K^n y K^m como columnas, la aplicación $f_A : K^n \to K^m$ dada por $f_A(x) = A \cdot x$, también es claramente lineal. Como veremos a lo largo del tema, toda aplicación lineal de K^n en K^m se puede escribir de este modo.

(iii) Fijemos n vectores $\{v_1, \dots, v_n\}$ de $V(K)$. La aplicación $f : K^n \to V$, dada por $f((x_1, \dots, x_n)) = \sum_{i=1}^n x_i v_i$ es lineal (generalizaremos este ejemplo en el teorema 3.8).

Ejercicio 3.7. *Sea $F(\mathbb{R}, \mathbb{R})$ el conjunto de todas las aplicaciones de $\mathbb{R}$ en $\mathbb{R}$ y U el conjunto de todas las aplicaciones de $\mathbb{R}$ en $\mathbb{R}$ que son derivables en todo punto (que es un s.v. de $F(\mathbb{R}, \mathbb{R})$. Demostrar que la aplicación* Der: $U \to F(\mathbb{R}, \mathbb{R})$*, que a cada función h le hace corresponder su función derivada h' (esto es, $h \mapsto h'$), es lineal.*

[4] Compruébese, (¿qué pasaría si K no fuera conmutativo?).

3.1.2. Construcción de aplicaciones lineales extendiendo por linealidad

Veremos a continuación que toda aplicación lineal queda determinada cuando se prescriben las imágenes de los vectores de una base, y construiremos explícitamente la aplicación.

Teorema 3.8. *Sea $V(K)$ un espacio vectorial de dimensión $n \in \mathbb{N}$ y $B = (v_1, \ldots, v_n)$ una base (ordenada) suya. Sea $(w'_1, \ldots, w'_n)$ una n-úpla de vectores de $V'(K)$. Entonces existe una única aplicación lineal $f : V \to V'$ que verifica $f(v_i) = w'_i$ para todo $i \in \{1, \ldots, n\}$.*

Demostración. Como una observación previa, si f existe entonces las condiciones que se le imponen la determinan de manera única. En efecto, como cada $v \in V$ se escribe como $v = \sum_{i=1}^{n} a_i v_i$ para ciertos escalares a_i (las coordenadas de v en B) unívocamente determinados, se tendrá

$$f(v) = f(\sum_{i=1}^{n} a_i v_i) = \sum_{i=1}^{n} a_i \bar{f}(v_i) = \sum_{i=1}^{n} a_i w'_i$$

(la segunda y tercera igualdades por las condiciones impuestas a $\bar{f}$, y la tercera por la definición de f). Esto demuestra que, si f existe, entonces es única y, de hecho, sólo puede estar definida por la expresión:

$$f(v) = \sum_{i=1}^{n} a_i w'_i.$$

Por tanto, basta con demostrar que esta definición de f satisface las condiciones requeridas.

Claramente, $f(v_i) = w'_i$, pues $(v_i)_B = (0, \ldots 0, 1, 0, \ldots, 0)^t$, donde el 1 aparece en la posición i-ésima. Para demostrar que f es lineal, si tomamos un segundo vector genérico $w = \sum_{i=1}^{n} b_i v_i$ y dos escalares cualesquiera $a, b \in K$ se tiene, sin más que usar las propiedades de los sumatorios:

$$av + bw = a\sum_{i=1}^{n} a_i v_i + b\sum_{i=1}^{n} b_i v_i = \sum_{i=1}^{n} (aa_i + bb_i) v_i,$$

Por tanto:

$$f(av + bw) = \sum_{i=1}^{n} (aa_i + bb_i) w'_i = a\sum_{i=1}^{n} a_i w'_i + b\sum_{i=1}^{n} b_i w'_i = af(v) + bf(w),$$

la primera y tercera igualdades por la definición de f, y la segunda por las propiedades de los sumatorios, con lo que se prueba la linealidad. ∎

Observación 3.9. (1) No se impone ninguna restricción sobre los vectores w'_i (por ejemplo, podrían ser todos iguales). No obstante, es inmediato de la construcción que el conjunto $\{w'_1, \ldots, w'_n\}$ es un sistema de generadores de $\mathrm{Im}(f)$. Además, veremos a lo largo del tema que esta familia es linealmente independiente (resp. un sistema de generadores, una base) en V' si y sólo si f es inyectiva (resp. suprayectiva, biyectiva)

(2) Al procedimiento constructivo de f en el teorema 3.8 se le llama *extensión por linealidad* (de los valores de f sobre la base B).

(3) En este teorema se ha supuesto que $V(K)$ es finitamente generado. No obstante, si se elimina esta restricción y se fijan una base (posiblemente infinita) B y una aplicación $B \to V'$, $v_i \mapsto w'_i$, entonces el mismo procedimiento de extensión por linealidad demuestra la existencia de una única aplicación lineal $f : V \to V'$ tal que $f(v_i) = w'_i$ para todo $v_i \in B$. Obsérvese que en ningún caso se ha supuesto que $V'(K)$ sea finitamente generado.

(4) Del teorema se concluye que dos aplicaciones lineales son iguales si coinciden en *alguna* base.

3.1.3. Tipos de aplicaciones lineales. Isomorfismos

El siguiente resultado proporciona un modo de construir aplicaciones lineales a partir de otras, y será usado en la discusión de los tipos de aplicaciones lineales.

Proposición 3.10. *Sean V, V', V'' tres e.v. sobre el mismo cuerpo K, y sean $f : V \to V'$, $g : V' \to V''$ aplicaciones lineales.*

(1) La composición $g \circ f : V \to V''$ es una aplicación lineal.

(2) Si f es biyectiva entonces la aplicación inversa $f^{-1} : V' \to V$ es lineal.

Demostración. (1) Para todo $v, w \in V$, $a, b \in K$:

$$g \circ f(av+bw) = g(f(av+bw)) = g(af(v)+bf(w)) = ag(f(v))+bg(f(w)) = a(g\circ f)(v)+b(g\circ f)(w)$$

(la linealidad de f y de g se usan, resp., en la segunda y tercera igualdades), como se quería.

(2) Para comprobar $f^{-1}(av'+bw') = af^{-1}(v')+bf^{-1}(w'), \forall v', w' \in V', a, b \in K$, basta con demostrar que las imágenes por f de ambos miembros de esta igualdad coinciden (la inyectividad de f implica entonces que son iguales), esto es:

$$f(af^{-1}(v')+bf^{-1}(w')) = af(f^{-1}(v'))+bf(f^{-1}(w')) = av'+bw' = f(f^{-1}(av'+bw'))$$

(la primera igualdad por la linealidad de f), como se quería. ∎

Definición 3.11. *Diremos que una aplicación lineal $f : V \to V'$ es un:*

- monomorfismo *si f es inyectiva*
- epimorfismo *si f es suprayectiva (esto es, sobre, exhaustiva).*
- isomorfismo *si f es biyectiva*

En el caso $V = V'$ diremos que f es un endomorfismo. *Si f es un endomorfismo biyectivo diremos que es un* automorfismo *(del e.v. $V(K)$).*

Corolario 3.12. *La composición de dos monomorfismos (resp. epimorfismos, isomorfismos) es un monomorfismo (resp. epimorfismo, isomorfismo).*

Demostración. Por la proposición 3.10(1) se sabe que la composición de aplicaciones lineales es lineal, y por las propiedades de las aplicaciones entre conjuntos que la composición de dos aplicaciones inyectivas (resp. suprayectivas, biyectivas) es inyectiva (resp. suprayectiva, biyectiva). ∎

Definición 3.13. *Se dice que $V(K)$ es isomorfo a $V'(K)$ si existe un isomorfismo de e.v. de V a V' (esto es, con dominio V y codominio V').*

Proposición 3.14. *La relación "ser isomorfo a" en la clase de todos los e.v., es una relación de equivalencia, esto es, verifica las propiedades:*

(i) Reflexiva: todo e.v. es isomorfo a sí mismo.

(ii) Simétrica: si $V(K)$ es isomorfo a $V'(K)$ entonces $V'(K)$ es isomorfo a $V(K)$.

(iii) Transitiva: si $V(K)$ es isomorfo a $V'(K)$ y $V'(K)$ es isomorfo a $V''(K)$ entonces $V'(K)$ es isomorfo a $V(K)$

Demostración. (i) Inmediato de que la aplicación identidad es un isomorfismo.

(ii) Inmediato de que (por lo proposición 3.10(2)), si $f : V \to V'$ es un isomorfismo entonces $f^{-1} : V' \to V$ es un isomorfismo.

(iii) Inmediato de que la composición de dos isomorfismos es un isomorfismo. ∎

El próximo resultado muestra que en la clase de equivalencia de cada e.v. finitamente generado existe un e.v. del tipo $K^n(K)$.

Proposición 3.15. *Sea $V(K)$ un e.v. de dimensión $n \in N$ y sea $B = (v_1, \ldots, v_n)$ una base. La aplicación:*

$$b_B : K^n \to V, \qquad b_B(x_1, \ldots, x_n) := \sum_{i=1}^{n} x_i v_i, \qquad \forall (x_1, \ldots, x_n) \in K^n$$

es un isomorfismo de e.v.

Demostración. Es inmediato que b_B es lineal, de hecho, es la extensión por linealidad de la aplicación que lleva la base usual B_u en B (respetando sus ordenaciones). La inversa de esta aplicación es la que asigna a cada vector sus coordenadas en la base B. Por tanto, la inyectividad es consecuencia de la unicidad de las coordenadas en la base (la indepencia lineal de B) y la suprayectividad de que tales coordenadas existan (B es un sistema de generadores). ∎

Observación 3.16. Del hecho de que b_B es lineal y biyectiva se deduce que b_B^{-1} es lineal también (proposición 3.10(2)). No obstante, es fácil de deducir directamente la linealidad de la aplicación que a cada vector de un e.v. le hace corresponder sus coordenadas en B (de hecho, estas propiedades se adelantaran en el tema anterior, observación 2.144).

3.2. Núcleo, imagen y rango

3.2.1. Propiedades del núcleo paralelas a las de la imagen

Definición 3.17. *Sea $f : V \to V'$ una aplicación lineal. El* núcleo *de f se define como el subconjunto:*

$$\mathrm{Nuc}(f) = f^{-1}(\{0\}) = \{v \in V : f(v) = 0\}.$$

Sabemos de la proposición 3.4(5) que $\mathrm{Im}(f)$ es un subespacio vectorial de V' y, (como en cualquier aplicación, sea lineal o no) f es suprayectiva si y sólo si $\mathrm{Im}(f) = V'$. Veamos a continuación resultados para el núcleo con cierto paralelismo a los de la imagen.

Proposición 3.18. *Sea $f : V \to V'$ una aplicación lineal:*

(i) $\mathrm{Nuc}(f)$ *es un subespacio vectorial de V.*

(ii) f es inyectiva si y sólo si $\mathrm{Nuc}(f) = \{0\}$.

Demostración. (i) Consecuencia inmediata de la proposición 3.4(6) (aplicada a $U' = \{0\}$).

(ii) Observemos en primer lugar que, por la proposición 3.4(1), siempre $0 \in \mathrm{Nuc}(f)$. Si suponemos que f es inyectiva, no puede existir otro vector $v \neq 0$ en $\mathrm{Nuc}(f)$, pues violaría la inyectividad ($f(0) = f(v)$, $0 \neq v$). Recíprocamente, supongamos que $\mathrm{Nuc}(f) = \{0\}$. Si $f(v) = f(w)$ para ciertos $u, w \in V$ entonces $f(v - w) = f(v) - f(w) = 0$. Esto es, $v - w \in \mathrm{Nuc}(f) = \{0\}$ y $v = w$, como se quería. ∎

Definición 3.19. *Sea $f : V \to V'$. Llamaremos* nulidad *de f a la dimensión de* $\mathrm{Nuc}(f)$*, y* rango *de f a la dimensión de* $\mathrm{Im}(f)$*, y los denotaremos, $n(f)$ y $r(f)$, resp.*

Trivialmente $n(f) \leq \dim V$, $r(f) \leq \dim V'$.

Corolario 3.20. *Si $f : V \to V'$ es lineal entonces $r(f) \leq \dim_K V$. Más aún:*
(a) si U es un s.v. de V entonces $\dim(f(U)) \leq \dim(U)$, y
(b) si $g : V' \to V''$ is lineal entonces $r(g \circ f) \leq$ Mínimo $\{r(f), r(g)\}$.

Demostración. Para la primera afirmación úsese la observación 3.9(1). Para (a) aplíquese lo anterior a $f|_U$. Para (b), aplíquese a $g|_{\mathrm{Im} f}$ teniendo en cuenta $\mathrm{Im}(g \circ f) = \mathrm{Im}(g|_{\mathrm{Im} f})$. ∎

Ejercicio 3.21. *Constrúyanse ejemplos donde las desigualdades del corolario sean estrictas.*

3.2.2. Caracterizaciones y existencia de monomorfismos y epimorfismos

Caracterizamos a continuación la inyectividad y suprayectividad mediante los conceptos de independencia lineal y sistema de generadores.

Teorema 3.22. *Para cualquier $f : V \to V'$ lineal:*

1. *Equivalen:*
   - a) *f es un monomorfismo.*
   - b) *Si $S \subset V$ es linealmente independiente entonces $f(S)$ es l.i.*
   - c) *Existe una base B de V tal que $f(B)$ es l.i.*
2. *Equivalen:*
   - a) *f es un epimorfismo.*
   - b) *Para todo sistema de generadores $S \subset V$, se tiene que $f(S)$ es un s.d.g. de V'.*
   - c) *Existe un sistema de generadores $S \subset V$, tal que $f(S)$ es un s.d.g. de V'.*
3. *Equivalen:*
   - a) *f es un isomorfismo.*
   - b) *Para toda base $B \subset V$, se tiene que $f(B)$ es una base de V'.*
   - c) *Existe una base $B \subset V$, tal que $f(B)$ es una base de V'.*

Demostración. En cada punto demostraremos cíclicamente a)$\Rightarrow$ b) $\Rightarrow$ c) $\Rightarrow$ a).

1. a) $\Rightarrow$ b). Sea $\{w'_1, \ldots, w'_m\} \subset f(S)$, donde $w'_i = f(w_i)$ para todo $i = 1, \ldots, m$ y $\{w_1, \ldots, w_n\} \subset S$. Tomemos una combinación lineal igualada a 0,

$$0 = a_1 w'_1 + \cdots + a_n w'_m = a_1 f(w_1) + \cdots + a_n f(w_m) = f(a_1 w_1 + \cdots + a_n w_m).$$

Como $f(0) = 0$ y f es inyectiva se sigue $a_1 w_1 + \cdots + a_n w_m = 0$, y por ser S l.i. todos los escalares son nulos, como se quería. b) $\Rightarrow$ c) Resulta inmediato, supuesto que se sabe de la existencia de alguna base B en V (lo que demostramos en el caso finitamente generado, y explicamos que ocurría en todo espacio vectorial). c) $\Rightarrow$ a) Supongamos que f no es inyectiva, y sea entonces $u \in \mathrm{Nuc}(f) \setminus \{0\}$. Escribiendo

u como combinación lineal de B, $u = a_1v_1 \cdots + a_nv_n$, con $\{v_1, \ldots, v_n\} \subset B$ y, necesariamete, no todos los escalares $a_1, \ldots, a_n \in K$ nulos, se tiene

$$0 = f(u) = f(a_1v_1 \cdots + a_nv_n) = a_1f(v_1) + \cdots + a_nf(v_n)$$

en contradicción con la independencia lineal de $f(B)$.

2. a)$\Rightarrow$ b) Sea $v' \in V'$. Por la suprayectividad, existe $v \in V$ tal que $f(v) = v'$. Por ser S un sistema de generadores, existen $w_1, \ldots, w_m \in S$ y escalares $a_1, \ldots, a_m \in K$ tales que $v = a_1w_1 + \cdots + a_mw_m$. Por tanto:

$$v' = f(v) = f(a_1w_1 + \cdots + a_mw_m) = a_1f(w_1) + \cdots + a_mf(w_m) \in L(f(S)),$$

como se quería. b) $\Rightarrow$ c) Tomando $S = V$ se sigue trivialmente. c) $\Rightarrow$ a) Como f es lineal, $\mathrm{Im}(f)$ es un subespacio vectorial de V' y $L(f(S)) = f(L(S)) \subset \mathrm{Im}(f)$. Por ser . Por ser S un s.d.g. de V se da la igualdad y, por ser $f(S)$ un s.d.g. de V', $\mathrm{Im}(f) = V'$.

3. Inmediato de las correspondientes implicaciones en los puntos 1 y 2. ∎

Como consecuencia, se obtiene el siguiente resultado que complementa al teorema 3.8.

Corolario 3.23. *Sea $f : V \to V'$ lineal, y sea B una base de V.*

1. *$f(B)$ es linealmente independiente si y sólo si f es inyectiva. En este caso, $\dim_K V \leq \dim_K V'$.*
2. *$f(B)$ es un sistema de generadores si y sólo si f es suprayectiva. En este caso, $\dim_K V \geq \dim_K V'$.*
3. *$f(B)$ es una base si y sólo si f es biyectiva. En este caso, $\dim_K V = \dim_K V'$.*

Demostración. Las únicas afirmaciones que no están contenidas en el teorema 3.22 son las referentes a las dimensiones. En el caso de que V tenga dimensión[5] $n \in \mathbb{N}$, la primera (resp. segunda, tercera) afirmación se deduce de que $f(B)$ es un subconjunto linealmente independiente (resp. un sistema de generadores, una base) de V con n elementos. En el caso de que V no sea finitamente generado entonces las expresiones $\dim_K V \leq \dim_K V'$ y $\dim_K V = \dim_K V'$ de los casos 1 y 3, resp., deben entenderse simplemente como que V' no es finitamente generado (lo que resulta obvio, pues $f(B)$ es infinito y linealmente independiente), mientras que la expresión $\dim_K V \geq \dim_K V'$ del punto 2 no implica ninguna restricción sobre si V' es finitamente generado o no[6]. ∎

El teorema 3.8 también se puede combinar con los resultados anteriores para concluir que las restricciones sobre las dimensiones en el corolario 3.23, las cuales eran necesarias para la existencia de un monorfismo o un epimorfismo, también son suficientes para su existencia. Por simplicidad, nos restringiremos a dimensión finita, aunque el resultado es cierto sin esta restricción.

Corolario 3.24. *Sean $V(K)$ y $V'(K)$ e.v. de dimensiones $n, m \in \mathbb{N}$, resp.*

1. *Existe un monomorfismo de V a V' si y sólo si $n \leq m$.*
2. *Existe un epimorfismo de V a V' si y sólo si $n \geq m$.*

[5] Si $V = \{0\}$ las afirmaciones siguen teniendo sentido tomando $B = \emptyset$ y $L(\emptyset) = 0$, como se explicó en el tema anterior.

[6] No obstante, las desigualdades entre las dimensiones en el caso de que V no sea finitamente generado también pueden interpretarse en el sentido más fuerte de que el cardinal de la base B es mayor, menor o igual que el de cualquier base de V'. Aunque bajo esta interpretación los resultados siguen siendo ciertos, su demostración excede nuestros objetivos.

3. *Existe un isomorfismo de V a V' si y sólo si $n = m$.*

Demostración. Las implicaciones hacia la derecha en los puntos 1, 2 y 3 están contenidas en el corolario 3.23 anterior. Para las implicaciones hacia la izquierda, sean B y B' bases de V y V', resp. Si $n \leq m$ (resp. $n \geq m$, $n = m$) existe una aplicación inyectiva (resp. suprayectiva, biyectiva) de B a B'. Extendiendo esta aplicación por linealidad (teorema 3.8) se obtiene una aplicación lineal de V a V' que es un monomorfismo (resp. epimorfismo, isomorfismo) por el corolario 3.23. ∎

Aunque inmediata, insistimos en la siguiente importante consecuencia.

Corolario 3.25. *Sean $V(K)$ y $V'(K)$ dos e.v., $V(K)$ es f.g.*
- *V y V' son isomorfos si y sólo si tienen igual dimensión.*
- *$V(K)$ es isomorfo o bien a $K^n(K)$, para un único $n \in \mathbb{N}$, o bien[7] a $\{0\}$.*
- *Si $f : V \to V'$ es un isomorfismo y U un s.v. de V, entonces $dim_K U = dim_K(f(U))$*

Demostración. Las dos primeras afirmaciones son inmediatas del punto 3 del corolario anterior. Para la tercera, como $f(U)$ es un s.v. de V' (proposición 3.4(5)), basta con observar que la aplicación $U \to f(U)$, $v \mapsto f(v)$ (obtenida restringiendo el dominio y codominio de f) es un isomorfismo. ∎

Observación 3.26. Esencialmente, este corolario reduce el estudio de los e.v. finitamente generados al caso de $K^n(K)$. De hecho, implícitamente se usa la última afirmación cuando, dada una base B, para calcular la dimensión de un s.v. de V, se toma un s.d.g. finito y se calcula el rango de la matriz obtenida con sus coordenadas.

3.2.3. Teorema del rango

Demostraremos a continuación una importante relación entre el núcleo y la imagen de una aplicación lineal o, con más precisión, entre su nulidad $n(f)$ y rango $r(f)$ (definición 3.19).

Teorema 3.27. *Sea $f : V \to V'$ lineal tal que $dim_K V = n \in \mathbb{N} \cup \{0\}$. Entonces,*

$$dim_K V = dim_K(\mathrm{Nuc}(f)) + dim_K(Im(f)), \qquad \text{esto es,} \qquad n = n(f) + r(f).$$

Demostración. Tomemos una base $B_{\mathrm{Nuc}f}\{v_1, \ldots, v_{n(f)}\}$ de Nuc(f) y ampliémosla a una base de $V(K)$, $B = \{v_1, \ldots, v_{n(f)}, v_{n(f)+1}, \ldots, v_n\}$. Sabemos que $f(B)$ es un sistema de generadores de Im(f) (vease la observación 3.9(1)) y, puesto que los vectores v_i con $i \leq n(f)$ se aplican en 0, el conjunto $\{f(v_{n(f)+1}), \ldots, f(v_n)\}$ es también un s.d.g de Imf. Como este conjunto tiene $n - n_f$ vectores, basta con comprobar que es linealmente independiente (pues se obtiene entonces el resultado $r(f) = n - n(f)$). Tomemos por tanto una combinación lineal suya igualada a 0:

$$0 = a_{n(f)+1} f(v_{n(f)+1}) + \cdots + a_n f(v_n) = f(a_{n(f)+1} v_{n(f)+1} + \cdots + a_n v_n),$$

la última igualdad aplicando la linealidad de f. La formula anterior implica que $a_{n(f)+1} v_{n(f)+1} + \cdots + a_n v_n \in$Nuc$(f)$ y, por tanto, este vector se puede escribir como combinación lineal de los elementos de $B_{\mathrm{Nuc}f}$. Esto es, para ciertos escalares $a_1, \ldots, a_{n(f)}$ se tiene:

$$a_{n(f)+1} v_{n(f)+1} + \cdots + a_n v_n = a_1 v_1 + \cdots + a_{n(f)} v_{n(f)}.$$

[7] La disyuntiva de casos se evita definiendo $K^0 := \{0\}$.

Pasando los términos del segundo miembro al primero, se tiene una combinación lineal de la base B igualada al vector 0. Por tanto, todos los coeficientes de la combinación deben ser nulos y, en particular, $a_{n(f)+1} = \cdots = a_n = 0$, como se quería demostrar. ∎

Como una aplicación de este teorema se tiene:

Corolario 3.28. *Sea $f : V \to V'$ una aplicación lineal entre espacios finitamente generados de igual dimensión. Equivalen:*

(i) f es biyectiva,

(ii) f es inyectiva, y

(iii) f es suprayectiva.

En consecuencia, si $g : V' \to V$ es lineal y cumple $g \circ f$ es un isomorfimo entonces g y f son isomorfismos. En este caso, si $g \circ f = I_V$ entonces $g = f^{-1}$.

Demostración. Sea n la dimensión común. De la igualdad $n = n(f) + r(f)$ se deduce: f es inyectiva (esto es, $n(f) = 0$), si y sólo si $r(f) = n$ (esto es, f es suprayectiva). Para la última afirmación, basta con observar que por la hipótesis f debe ser inyectiva y g sobre (o bien úsese el corolario 3.20). ∎

3.3. Expresión matricial de una aplicación lineal

Sabemos que, dada una matriz $A \in M_{m\times n}(K)$, se puede definir la aplicación lineal $f_A : K^n \to K^m$ sin más que multiplicar por A las n-úplas del dominio, consideradas como columnas (véase ejemplo 3.6(ii)). Teniendo en cuenta que, usualmente, consideramos los elementos de K^n y K^m como filas, mientras que escritos por columnas representan las coordenadas en la base usual B_u^n de K^n y B_u^m de K^m, podemos expresar f_A como sigue:

$$f_A : K^n \to K^m, \qquad (f_A(x))_{B_u^m} = A \cdot x_{B_u^n}. \tag{3.1}$$

Nuestro objetivo será mostrar cómo, aunque en espacios vectoriales arbitrarios no existe ninguna base preferida (a la que podamos llamar usual o canónica), cualquier aplicación lineal se puede representar de esta forma, una vez escogidas bases del dominio y codominio. A lo largo de toda esta sección V y V' serán siempre espacios vectoriales de dimensiones $n, m \in \mathbb{N}$, respectivamente.

3.3.1. Matriz asociada a una aplicación lineal

Definición 3.29. *Sea $f : V \to V'$ lineal y sean $B = (v_1, \ldots, v_n)$ y $B' = (v'_1, \ldots, v'_m)$ bases de V y V', resp. La* matriz de f en las bases B y B' *es la matriz en $M_{m\times n}(K)$, la cual, denotaremos $M(f, B, B')$ o, preferentemente, $M(f, B' \leftarrow B)$, cuya columna j-ésima contiene las coordenadas del vector $f(v_j)$ respecto de B', para todo $j \in \{1, \ldots, n\}$. Lo simbolizamos así:*

$$M(f, B' \leftarrow B) = \big(f(v_1)_{B'} \,|\, f(v_2)_{B'} \,|\, \ldots \,|\, f(v_n)_{B'}\big).$$

Con esta definición, como anunciamos, podemos generalizar la expresión vista para f_A.

Proposición 3.30 (Ecuaciones de f en B y B'). *En las condiciones anteriores, se tiene:*

$$f(v)_{B'} = M(f, B' \leftarrow B) \cdot v_B \qquad \forall v \in V.$$

Llamaremos a la igualdad anterior expresión matricial de f en las bases B y B'. *Esta igualdad caracteriza $M(f,B' \leftarrow B)$, esto es, si $A \in M_{m\times n}(K)$ verifica $f(v)_{B'} = A\cdot v_B$ para todo $v \in V$, entonces $A = M(f,B' \leftarrow B)$.*

Demostración. Tomemos escalares (a_{ij}) de modo que:

$$f(v_j) = \sum_{i=1}^{m} a_{ij}\cdot v'_i.$$

De esta forma, la j-ésima columna de $M(f,B' \leftarrow B)$ contiene exactamente a los escalares a_{ij} con $i = 1,\dots,n$. Escribamos ahora un vector genérico $v \in V$ como combinación lineal de B:

$$v = \sum_{j=1}^{n} x_j\cdot v_j$$

es decir, v_B contiene a los escalares x_i con $i = 1,\dots,n$. Buscamos calcular las coordenadas $(x'_1,\dots,x'_m)^t$ de $f(v)$ en B'. Teniendo encuenta las dos últimas fórmulas y usando la linealidad de f, se tiene:

$$\begin{aligned} f(v) = f\left(\sum_{j=1}^{n} x_j\cdot v_j\right) &= \sum_{j=1}^{n} x_j\cdot f(v_j) = \sum_{j=1}^{n} x_j\cdot\left(\sum_{i=1}^{m} a_{ij}\cdot v'_i\right)\\ &= \sum_{i=1}^{m}\left(\sum_{j=1}^{n} x_j\cdot a_{ij}\cdot v'_i\right) = \sum_{i=1}^{m}\left(\sum_{j=1}^{n} a_{ij}\cdot x_j\right)\cdot v'_i. \end{aligned}$$

Obsérvese que las coordenadas buscadas x'_i verifican

$$f(v) = \sum_{i=1}^{m} x'_i\cdot v'_i$$

y, por la unicidad de estas coordenadas, se pueden igualar los coeficientes de cada v'_i obteniéndose:

$$x'_i = \sum_{j=1}^{n} a_{ij}\cdot x_j$$

esto es, la expresión buscada $f(v)_{B'} = (a_{ij})\cdot v_B$.

Para la última afirmación, se razona como en el caso del cambio de base (prop. 2.154), esto es, basta con darse cuenta de que al ser $A\cdot v_B = M(I,B' \leftarrow B)\cdot v_B$ para todo v_B, la igualdad de cada columna se sigue tomando sucesivamente $v_B = (1,0,\dots,0)^t,\dots,(0,0,\dots,1)^t$. ∎

Observación 3.31. Dada una matriz $A \in M_{m\times n}$, la aplicación lineal asociada $f_A : K^n \to K^m$ verifica:

$$A = M(f_A, B_u^m \leftarrow B_u^n).$$

De hecho, la fórmula (3.1) es la expresión matricial de f_A en B_u^n y B_u^m. Más aún, toda aplicación lineal $f : K^n \to K^m$ es del tipo f_A, donde A se calcula simplemente como $A = M(f, B_u^m \leftarrow B_u^n)$.

El siguiente resultado puede verse como una generalización del caso K^n, K^m a cualesquiera e.v.

Corolario 3.32. *Sea $A \in M_{m\times n}(K)$, y $V(K), V(K')$ e.v. de dimensiones n,m resp. Fijadas bases B de V y B' de V' existe una única aplicación lineal $f : V \to V'$ tal que $A = M(f,B' \leftarrow B)$.*

Demostración. Es una aplicación inmediata del teorema 3.8 (compruébese como ejercicio). ∎

Observación 3.33. Un caso particular de aplicación lineal es la identidad $I_V : V \to V$. Claramente, $M(I_V, B \leftarrow B)$ es igual a la matriz identidad I_n para cualquier base B. Sin embargo, si tomamos dos bases distintas se tiene:

$$M(I_V, B' \leftarrow B) = \big((v_1)_{B'} \,|\, (v_2)_{B'} \,|\, \ldots \,|\, (v_n)_{B'}\big) = M(I, B' \leftarrow B),$$

esto es, las matrices de cambio de base pueden verse como casos particulares de las matrices de aplicaciones lineales.

El siguiente resultado muestra una primera propiedad de compatibilidad de las operaciones definidas entre matrices y entre aplicaciones lineales.

Proposición 3.34. *Sean $V(K), V'(K), V''(K)$ e.v. y B, B', B'', resp., bases ordenadas de cada uno.*
(i) Si $f : V \to V'$, $g : V' \to V''$ son lineales entonces:

$$M(g \circ f, B'' \leftarrow B) = M(g, B'' \leftarrow B') \cdot M(f, B' \leftarrow B).$$

(ii) f es un isomorfismo si y sólo si $M(f, B' \leftarrow B)$ es regular. En este caso

$$M(f^{-1}, B \leftarrow B') = M(f, B' \leftarrow B)^{-1}$$

Demostración. Para (i), la demostración es análoga a la del cambio de base estudiada en el tema anterior. Sea v cualquier vector de V. Por la proposición 3.30:

$$g(f(v))_{B''} = M(I, B'' \leftarrow B') \cdot f(v)_{B'} \qquad f(v)_{B'} = M(I, B' \leftarrow B) \cdot v_B.$$

Sustituyendo la segunda igualdad en la primera, y usando la asociatividad del producto de matrices:

$$g(f(v)) v_{B''} = \big(M(g, B'' \leftarrow B') \cdot M(f, B' \leftarrow B)\big) \cdot v_B.$$

Como la expresión matricial de $g \circ f$ (esto es, $((g \circ f)(v))_{B''} = M(g \circ f, B'' \leftarrow B) \cdot v_B$) caracteriza a $M(g \circ f, B'' \leftarrow B)$, se sigue la igualdad pedida.

Para (ii), si f es un isomorfismo, basta aplicar la parte (i) y luego la observación 3.33 para obtener:

$$M(f^{-1}, B \leftarrow B') \cdot M(f, B' \leftarrow B) = M(I_V, B \leftarrow B) = I_n,$$

Recíprocamente, si $M(f, B' \leftarrow B)$ es regular, la única aplicación lineal $h : V \to V'$ tal que $M(h, B \leftarrow B') = M(f, B' \leftarrow B)^{-1}$ (recuérdese el corolario 3.32), verifica:

$$M(h, B \leftarrow B') \cdot M(f, B' \leftarrow B) = I_n = M(I_V, B \leftarrow B),$$
$$M(f, B' \leftarrow B) \cdot M(h, B \leftarrow B') = I_n = M(I_{V'}, B' \leftarrow B')$$

de donde $h \circ f = I_V, f \circ h = I_{V'}$, esto es, f es biyectiva y h es la inversa de f. ∎

Corolario 3.35. *Sea $f : V \to V'$ lineal, $B, \bar{B}$ dos bases de V y $B', \bar{B}'$ dos bases de V'. Entonces:*

$$M(f, \bar{B}' \leftarrow \bar{B}) = M(I_{V'}, \bar{B}' \leftarrow B') \cdot M(f, B' \leftarrow B) \cdot M(I_V, B \leftarrow \bar{B})$$

Demostración. Trivialmente, $f = I_{V'} \circ f \circ I_V$ y por la proposición 3.34(i) el miembro derecho es $M(I_{V'} \circ f \circ I_V, \bar{B}' \leftarrow \bar{B})$. ∎

3.3.2. Rango y matrices equivalentes

El corolario 3.35 permite ver la relación entre la matriz de una misma apliación lineal cuando se cambian las bases. A onctinuación, estudiaremos con detalle lo que tienen en común todas esas bases. El siguiente ejercicio permite comprobar una primera relación con las transformaciones elementales.

Ejercicio 3.36. *Sea $f : V \to V'$ lineal, $B, \bar{B}$ dos bases de V y $B', \bar{B}'$ dos bases de V', $B = (v_1, \ldots, v_n)$, $B' = (v'_1, \ldots, v'_m)$. Calcular la relación entre $M(f, \bar{B}' \leftarrow \bar{B})$ y $M(f, B' \leftarrow B)$ en cada caso:*

1. *(Transformaciones elementales por columnas.) $\bar{B}' = B'$ y $\bar{B}$ se obtiene de B, para $1 \leq i < j \leq n$:*

   1.- Cambiando de orden los elementos v_i y v_j.

   2.- Cambiando v_1 por $a \cdot v_i$, siendo $a \neq 0$ ($a \in \mathbb{K}$).

   3.- Cambiando v_i por $a \cdot v_i + v_j$, siendo $a \in \mathbb{K}$.

2. *(Transformaciones elementales por filas.) $\bar{B} = B$ y $\bar{B}'$ se obtiene de B', para $1 \leq i < j \leq m$:*

   1.- Cambiando de orden los elementos v'_i y v'_j.

   2.- Cambiando v'_1 por $a^{-1} \cdot v'_i$, siendo $a \neq 0$ ($a \in \mathbb{K}$).

   3.- Cambiando v'_i por $v'_i - a \cdot v'_j$, siendo $a \in \mathbb{K}$.

Proposición 3.37. *El rango $r(f)$ de una aplicación lineal $f : V \to V'$ coincide con el rango (por columnas) de su matriz $M(f, B' \leftarrow B)$ en cualesquiera bases B de V y B' de V'.*

Demostración. Sean $B = (v_1, \ldots, v_n), B' = (v'_1, \ldots, v'_m)$, $r(f)$ es la dimensión de $L\{f(v_1), \ldots, f(v_n)\} \subset V'$. Como sabemos que la aplicación $b_{B'} : K^m \to V', (x_1, \ldots, x_m) \mapsto \sum_{i=1}^m x_i v'_i$ es un isomorfismo, esa dimensión coincide con la de $L\{b_{B'}^{-1}(f(v_1)), \ldots, b_{B'}^{-1}(f(v_n))\} \subset K^m$ (recuérdese el último punto del corolario 3.25). El resultado se concluye entonces porque cada $b_{B'}^{-1}(f(v_i))$ es precisamente la columna i-ésima de $M(f, B' \leftarrow B)$. ∎

Como consecuencia de la demostración del teorema del rango, se tiene entonces:

Proposición 3.38. *Sea $f : V \to V'$ lineal. Entonces existen bases B de V y B' de V' tales que:*

$$M(f, B' \leftarrow B) = \left(\begin{array}{c|c} I_r & O_{r \times (n-r)} \\ \hline O_{(m-r) \times r} & O_{(m-r) \times (n-r)} \end{array} \right),$$

donde $r \in \{0, 1, \ldots, \mathrm{Min}\{m, n\}\}$ es el rango de f (por lo que está unívocamente determinado).

Demostración. Procediendo como en la demostración del teorema del rango, tómese una base $B = (v_1, \ldots, v_r, v_{r+1}, \ldots v_n)$ donde ahora son los últimos $n - r$ vectores los que forman una base del núcleo (en aquel teorema se pusieron en las primeras posiciones). Sabemos entonces que $\{f(v_1), \ldots, f(v_r)\}$ es un conjunto l. i., por lo que se puede ampliar hasta una base $B' = (f(v_1), \ldots, f(v_r), v'_{r+1}, \ldots, v'_m)$ de todo V'. Es inmediato entonces comprobar que la matriz de f en estas bases B, B' es de la forma requerida. ∎

Observación 3.39. Obsérvese que, si tomamos otras dos bases, $\bar{B}, \bar{B}'$ de V y V' resp., por la fórmula del cambio de base para aplicaciones lineales (corolario 3.35), se tiene

$$M(f, \bar{B}' \leftarrow \bar{B}) = Q^{-1} \cdot \left(\begin{array}{c|c} I_r & O_{r\times(n-r)} \\ \hline O_{(m-r)\times r} & O_{(m-r)\times(n-r)} \end{array} \right) \cdot P$$

donde las matrices P y Q son de cambio de base y, por tanto, regulares. Concretamente, $P = M(I_V, B \leftarrow \bar{B}) \in Gl(n,K)$. y $Q = M(I_{V'}, B' \leftarrow \bar{B}') \in Gl(m,K)$

Esta relación entre matrices y aplicaciones lineales, motiva los siguientes conceptos.

Definición 3.40. *(i) El* núcleo de la matriz $A \in M_{m\times n}(K)$*, denotado* Nuc(A)*, es:*

$$\mathrm{Nuc}(A) = \{x \in M_{n\times 1} \,|\, A \cdot x = 0\},$$

que coincide con el núcleo de la aplicación $f_A : K^n \to K^m$ (tomando los elementos de K^n y K^m como columnas). Asimismo, la imagen de A *es* $Im(A) := Im(f_A)$.

(ii) Una segunda matriz C sobre K es equivalente *a A si existen e.v. V, V', bases $B, \bar{B}$ de V, $B', \bar{B}'$ de V' y una aplicación lineal $f : V \to V'$ tales que:*

$$A = M(f, B' \leftarrow B), \qquad C = M(f, \bar{B}' \leftarrow \bar{B})$$

(en particular, $A \in M_{m\times n}(K)$).

Con la definición anterior, es fácil comprobar que si dos matrices A y C son equivalentes entonces existen matrices regulares P, Q (de órdenes $n \times n$ y $m \times m$, resp.) tales que $C = Q^{-1}AP$. El siguiente resultado muestra, en particular, que el recíproco es cierto. Por tanto, la existencia de tales P y Q se puede considerar como una definición alternativa de que A y C sean equivalentes.

Proposición 3.41. *Sean $A, C \in M_{m\times n}$. Las siguientes afirmaciones equivalen:*

1. *A y C son equivalentes.*
2. *A y C tienen el mismo rango (por columnas).*
3. *A y C son equivalentes a una misma matriz $m \times n$ del tipo $\left(\begin{array}{c|c} I_r & O \\ \hline O & O \end{array} \right)$ (necesariamente, con r unívocamente determinado).*
4. *Existen matrices regulares $P \in Gl(n,K)$ y $Q \in Gl(m,K)$ tales que*[8] *$C = Q^{-1} \cdot A \cdot P$.*
5. *Escogidos dos e.v. V, V' de dimensiones n, m, resp. y una base B, B' en cada uno de ellos, resp., y construida la única aplicación lineal $f : V \to V'$ tal que*

$$A = M(f, B' \leftarrow B),$$

se pueden hallar dos nuevas bases $\bar{B}, \bar{B}'$ en V, V', resp. tales que

$$C = M(f, \bar{B}' \leftarrow \bar{B}).$$

En particular, puede escogerse $V = K^n$, $V' = K^m$, $B = B_u^n$, $B = B_u^m$.

[8]Obviamente, la propiedad es cierta escribiendo Q en lugar de Q^{-1}. No obstante, escogemos Q^{-1} por conveniencia, de modo que en la observación 3.39 P y Q representen cambios de bases "con barra" $\bar{B}, \bar{B}'$ a "sin barra" B, B' (su utilidad será manifiesta cuando se estudien las matrices semejantes).

Demostración. ($1 \Rightarrow 2$). Ambas matrices sirven para la expresión matricial de una misma aplicación lineal f y sabemos entonces (proposición 3.37): $\text{rango}(A) = \text{rango}(f) = \text{rango }(C)$.

($2 \Rightarrow 3$) Consideramos las aplicaciones lineales asociadas $f_A, f_C : K^n \to K^m$. Por la proposición 3.38 sabemos que A y C son equivalentes a matrices del tipo $\left(\begin{array}{c|c} I_{r_A} & O \\ \hline O & O \end{array}\right)$, $\left(\begin{array}{c|c} I_{r_C} & O \\ \hline O & O \end{array}\right)$, resp. El resultado se sigue de $r_A = \text{rango }(A) = \text{rango }(C) = r_C$.

($3 \Rightarrow 4$) Por la observación 3.39 existen matrices regulares $P_A, P_C \in Gl(n,K)$ y $Q_A, Q_C \in Gl(m,K)$ tales que:

$$A = Q_A^{-1} \cdot \left(\begin{array}{c|c} I_r & O \\ \hline O & O \end{array}\right) \cdot P_A, \qquad C = Q_C^{-1} \cdot \left(\begin{array}{c|c} I_r & O \\ \hline O & O \end{array}\right) \cdot P_C$$

de donde

$$Q_A \cdot A \cdot P_A^{-1} = Q_C \cdot C \cdot P_C^{-1}, \qquad \text{y, por tanto,} \qquad C = (Q_C^{-1} \cdot Q_A) \cdot A \cdot (P_A^{-1} \cdot P_C)$$

por lo que basta con tomar $P = P_A^{-1} \cdot P_C \in Gl(n,K)$ y $Q = Q_A^{-1} \cdot Q_C \in Gl(m,K)$.

($4 \Rightarrow 5$) Dado $A = M(f, B' \leftarrow B)$, construiremos las nuevas bases requeridas a partir de P y Q. Sean $\bar{B}$ y $\bar{B}'$ las únicas bases en V y V' que verifican respectivamente: $P = M(I_V, B \leftarrow \bar{B}), Q = M(I_{V'}, B' \leftarrow \bar{B}')$. Se tiene entonces:

$$C = Q^{-1} \cdot A \cdot P = M(I_{V'}, \bar{B}' \leftarrow B') \cdot M(f, B' \leftarrow B) \cdot M(I_V, B \leftarrow \bar{B}) = M(f, \bar{B}' \leftarrow \bar{B})$$

(la primera igualdad por hipótesis y la última por el cambio de bases para aplicaciones lineales).

($5 \Rightarrow 1$) Trivial de la definición de matrices equivalentes. ∎

Ejercicio 3.42. *Compruébese que cada una de las siguientes propiedades caracteriza que las matrices A y C sean equivalentes:*

(i) Las aplicaciones lineales asociadas $f_A, f_C : K^n \to K^m$ tienen el mismo rango.

(ii) C puede obtenerse a través de transformaciones elementales por filas y columnas de A (véase el ejercicio 3.36)[9].

El nombre *equivalente* en la definición 3.40(ii) es consistente con su uso en relaciones binarias:

Corolario 3.43. *La relación binaria "ser equivalente a" en el conjunto de todas las matrices sobre el mismo cuerpo conmutativo K, es una relación de equivalencia.*

Demostración. Del punto 2 de la proposición 3.41 se tiene que dos matrices cualesquiera son equivalentes si y sólo si tienen el mismo número de filas, el mismo número de columnas y el mismo rango. Estas propiedades verifican trivialmente las de una relación de equivalencia. ∎

Finalmente, demos una interpretación de los SEL.

Corolario 3.44. *Sea $A \in M_{m\times n}(K), b \in M_{m\times 1}(K)$ y el SEL $Ax = b$.*

(i) Las soluciones del SEL coinciden con $f_A^{-1}(b)$ (imagen recíproca de b).

(ii) El SEL es compatible si y sólo si $b \in Im(f_A)$. En particular, es compatible en el caso homogéneo $b = 0$ donde el conjunto de soluciones es $Nuc(f_A)$.

(iii) En el caso compatible, el conjunto de soluciones $f_A^{-1}(b)$ se expresa a partir de cualquier solución particular x_0 del sistema como el subespacio afín $x_0 +$ Nuc $(f_A) := \{x_0 + y \,|\, y \in \text{Nuc}(f_A)\}$.

[9]Para un estudio más detallado de las transformaciones elementales puede verse el libro de L. Merino y E. Santos.

(iv) La dimensión de Nuc (f_A) *(y, por tanto, la del subespacio afín* $f_A^{-1}(b)$ *en el caso compatible) coincide con* $n-$ *dim*(Imf_A).

(v) Cuando el SEL sea compatible, será también determinado si y sólo si f_A *es inyectiva.*

Demostración. La única afirmación no trivial es *(iii)*, esto es $f_A^{-1}(b) = \{x_0 + y \,|\, y \in \mathrm{Nuc}(f_A)\}$. Para la inclusión $\supset$, $f_A(x_0+y) = f_A(x_0) + f_A(y) = b + 0 = b$. Para $\subset$, si $x_0, x_0' \in f_A^{-1}(b)$ entonces $f(x_0' - x_0) = f(x_0') - f(x_0) = b - b = 0$, esto es, $y := x_0' - x_0 \in \mathrm{Nuc} f_A$. ∎

3.3.3. Estructura de espacio vectorial de Lin$(V(K), V'(K))$

Dados los e.v. $V(K), V'(K)$ (no necesariamente finitamente generados) consideremos el conjunto $\mathrm{Lin}(V(K), V'(K))$ de todas las aplicaciones lineales de V en V', que denotaremos simplemente,

$$\mathrm{Lin}(V, V') := \{f : V \to V' \,|\, f \text{ es lineal}\}.$$

Usando la estructura de e.v. en V', podemos definir unas operaciones suma y producto por escalares en $\mathrm{Lin}(V, V')$:

$$\begin{array}{rlllrl} + : \mathrm{Lin}(V,V') \times \mathrm{Lin}(V,V') \to \mathrm{Lin}(V,V'), & (f,g) \mapsto f+g & \text{donde} & (f+g)(v) := & f(v)+g(v), \\ \cdot : K \times \mathrm{Lin}(V,V') \to \mathrm{Lin}(V,V'), & (a,f) \mapsto a \cdot f & \text{donde} & (a \cdot f)(v) := & a \cdot f(v), \end{array}$$

para todo $v \in V$, $a \in K$. No obstante, antes se debe comprobar que las aplicaciones así definidas son lineales.

Lema 3.45. *Las definiciones anteriores son consistentes, esto es,* $f + g \in \mathrm{Lin}(V, V')$, $a \cdot f \in \mathrm{Lin}(V, V')$, *para todo* $f, g \in \mathrm{Lin}(V, V')$ *y todo* $a \in K$.

Demostración. Para la suma (el producto por escalares queda como ejercicio):

$$\begin{aligned} (f+g)(av+vw) = & \quad f(av+bw) + g(av+bw) = (af(v)+bf(w)) + (ag(v)+bg(w)) \\ = & \quad a(f(v)+g(v)) + b(f(v)+g(w)) = a \cdot (f+g)(v) + b \cdot (f+g)(w) \end{aligned}$$

para todo $v, w \in V$, $a, b \in K$, donde en la primera línea se usa primero la definición de la suma en $\mathrm{Lin}(V, V')$ y después la linealidad de f y g, y en la segunda se agrupan términos operando en V' y se vuelve a aplicar la definición de suma en $\mathrm{Lin}(V, V')$. ∎

Proposición 3.46. *Con estas operaciones,* $(\mathrm{Lin}(V, V'), + \cdot K)$ *es un espacio vectorial (al cual denotaremos simplemente* $\mathrm{Lin}(V, V')$*).*

Demostración. Queda como ejercicio, que puede comprobarse de al mneos dos maneras distintas: (a) demostrando directamente todas las propiedades de e.v., o (b) puesto que se sabe del tema anterior que el conjunto $F(X, V')$ de todas las aplicaciones de X en V' tiene una estructura natural de e.v., comprobando que $\mathrm{Lin}(V, V')$ es un subespacio vectorial de $F(X, V')$ para $X = V$. ∎

En el caso finitamente generado, hay una relación natural entre estas operaciones de $\mathrm{Lin}(V, V')$ y las de $M_{m \times n}(K)$.

Proposición 3.47. *Sean V, V' e.v. sobre K de dimensiones $n, m \in \mathbb{N}$ con bases B, B', resp. Entonces:*

$$M(f+g, B' \leftarrow B) = M(f, B' \leftarrow B) + M(g, B' \leftarrow B), \qquad M(a \cdot f, B' \leftarrow B) = a \cdot M(f, B' \leftarrow B)$$

para todo $f, g \in \mathrm{Lin}(V, V')$, $a \in K$.

Demostración. Sean $B = \{v_1, \ldots, v_n\}$ y $B' = \{v'_1, \ldots, v'_m\}$, y sea $b_{B'}^{-1} : V' \to K^m$ la aplicación que a cada $v' \in V'$ le hace corresponder sus coordenadas en B'. La expresión para la suma se deduce de que la columna j-ésima de $M(f+g, B' \leftarrow B)$ está formada por:

$$((f+g)(v_j))_{B'} = b_{B'}^{-1}((f+g)(v_j)) = b_{B'}^{-1}(f(v_j)) + b_{B'}^{-1}(g(v_j)),$$

que es la suma de las columnas j-ésimas de $M(f, B' \leftarrow B)$ y $M(g, B' \leftarrow B)$. Para el producto por escalares, razónese análogamente. ∎

La relación entre las estructuras de e.v. de matrices y aplicaciones lineales queda entonces resumida en el siguiente resultado.

Teorema 3.48. *Sean V, V' e.v. sobre K de dim. $n, m \in \mathbb{N}$. Fijadas bases B, B' de V, V' resp., la aplicación $F_{B' \leftarrow B}$ definida por:*

$$F_{B' \leftarrow B} : \mathrm{Lin}(V, V') \to M_{m \times n}, \qquad f \mapsto M(f, B' \leftarrow B),$$

es un isomorfismo de e.v. En consecuencia, $dim_k \mathrm{Lin}(V, V') = m.n$, y una base de $\mathrm{Lin}(V, V')$ es:

$$B_{B' \leftarrow B} := \left\{ f_{ij} \in \mathrm{Lin}(V, V') : M(f_{ij}, B' \leftarrow B) = E_{ij}, \, \forall i \in \{1, \ldots, m\}, \forall j \in \{1, \ldots, n\} \right\}$$

donde $E_{ij} \in M_{m \times n}$ es la matriz que tiene todos sus elementos nulos excepto un 1 en la posición (i, j) (esto es, $(E_{ij})_{kl} = \delta_{ik}\delta_{jl}$).

Demostración. La linealidad de $F_{B' \leftarrow B}$ es el contenido de la proposición 3.47 anterior, y su biyectividad ya se conocía (corolario 3.32). El resto es inmediato de que el conjunto de las matrices E_{ij} forma una base de $M_{m \times n}(K)$ y, por tanto, el conjunto de sus preimágenes por $F_{B' \leftarrow B}$ es una base de $\mathrm{Lin}(V, V')$. ∎

3.3.4. Anillo de endomorfismos y matrices semejantes

Recordemos que los endomorfismos son un caso especial de aplicaciones lineales $f : V \to V'$ en los que el dominio coincide con el codominio, $V = V'$. Denotaremos $\mathrm{End}(V) := \mathrm{Lin}(V, V)$, y todos los resultados vistos a lo largo de esta anterior resultan aplicables ahora. No obstante, se tiene ahora que la composición de dos endomorfismos siempre tiene sentido y es un nuevo endomorfismo, por lo que se puede definir la ley de composición interna en $\mathrm{End}(V)$:

$$\circ : \mathrm{End}(V) \times \mathrm{End}(V) \qquad (f, h) \mapsto f \circ h$$

Esto es similar al caso de las matrices cuadradas $M_n(K)$, para las cuales el producto proporcionaba una ley de composición interna. Recuérdese además que $M_n(K)$ gozaba de una estructura de anillo unitario con la suma y producto (de la cual carecían las matrices no cuadradas).

Proposición 3.49. $(\mathrm{End}(V), \circ, +)$ *es un anillo unitario (al cual llamaremos* anillo $\mathrm{End}(V)$*).*
$(\mathrm{End}(V), +, \cdot K)$ *es un espacio vectorial (al que denotaremos simplemente* $\mathrm{End}(V)$*).*

Demostración. Se puede comprobar fácilmente que se verifican todas las propiedades de anillo, siendo la aplicación identidad I_V su elemento unitario. La segunda afirmación es un caso particular de la estructura de e.v. de $\mathrm{Lin}(V,V')$. ∎

En el caso finitamente generado, cuando se estudian las matrices asociadas a endomorfismos se tiene una posibilidad nueva: escoger la misma base B en el dominio y codominio. Denotaremos:

$$M(f,B) := M(f, B \leftarrow B).$$

Se tiene entonces como una consecuencia inmediata del teorema 3.48:

Corolario 3.50. *Sea $V(K)$ un e.v. de dim. $n \in \mathbb{N}$. Fijada una base B de V la aplicación:*

$$F_B : \mathrm{End}(V) \to M_n, \qquad f \mapsto M(f,B),$$

es un isomorfismo de e.v. En consecuencia, $\dim_K \mathrm{End}(V) = n^2$, y una base de $\mathrm{End}(V)$ *es:*

$$B_B := \left\{ f_{ij} \in \mathrm{End}(V) : M(f_{ij}, B) = E_{ij},\ \forall i,j \in \{1,\ldots,n\} \right\}.$$

Observación 3.51. De la aplicación F_B se puede decir también que es un isomorfismo de anillos unitarios, al ser una aplicación biyectiva entre dos anillos unitarios que verifica:

$$F_B(f \circ h) = F_B(f) \cdot F_B(h), \qquad F_B(f+h) = F_B(f) + F_B(h), \qquad F_B(I_V) = I_n, \qquad \forall f,h \in \mathrm{End}(V),$$

esto es, F_B "preserva" (o "respeta") las estructuras de anillos unitarios de su dominio y codominio. [10]

La elección $B = B'$ sugiere el siguiente refinamiento de la relación "ser equivalente a" entre matrices.

Definición 3.52. *Dos matrices cuadradas $A, C \in M_n(K)$ son* semejantes *cuando existe un e.v. $V(K)$, dos bases $B, \bar{B}$ y un endomorfismo $f \in \mathrm{End}(V)$ que verifican:*

$$A = M(f,B) \qquad y \qquad C = M(f,\bar{B}).$$

Razonando como en la proposición 3.41 se tiene ahora:

Proposición 3.53. *Sean $A, C \in M_n$. Equivalen:*

1. *A y C son semejantes.*
2. *Existe una matriz regular $P \in Gl(n,K)$ tal que $C = P^{-1} \cdot A \cdot P$.*

[10] También podemos justificar ahora que si dos matrices cuadradas satisfacen $A \cdot C = I_n$ entonces A y C son regulares y $C \cdot A = I_n$, esto es, $C = A^{-1}$. De hecho, tomando, por ejemplo $V = K^n$, $B = B_u$ los endomorfismos f,h determinados por A y C resp. verifican $f \circ h = I_V$ por lo que h es biyectivo (en caso contrario, se llega a un absurdo aplicando la composición a un $v \in \mathrm{Nuc}(h) \setminus \{0\}$) y f es biyectivo (al absurdo se llegaría ahora tomando $v \in \mathrm{Nuc}(f) \setminus \{0\}$ y aplicando la composición a $h^{-1}(v)$). En consecuencia, existe la inversa h^{-1} que se puede componer a la derecha de ambos miembros en $f \circ h = I_V$, obteniéndose $f = h^{-1}$. Por tanto, A y C son regulares y $A = M(f,B) = M(h^{-1},B) = M(h,B)^{-1} = C^{-1}$. Demostraciones distintas se obtendrán más adelante, usando determinantes.

3. *Escogidos cualquier e.v. $V(K)$ de dimensión n, y una base B de V, y construido el único endomorfismo $f : V \to V$ tal que*
$$A = M(f,B),$$
se puede hallar una nueva base $\bar{B}$ en V tal que
$$C = M(f,\bar{B}).$$
En particular, puede escogerse $V = K^n$, $B = B_u$.

Demostración. ($1 \Rightarrow 2$). Particularizando la fórmula del cambio de base para aplicaciones lineales:

$$M(f,\bar{B}) = M(I_V,\bar{B} \leftarrow B) \cdot M(f,B) \cdot M(I_V,B \leftarrow \bar{B})$$

En nuestro caso, por hipótesis, $C = M(f,\bar{B})$ y $A = M(f,B)$, por lo que basta tomar $P = M(I_V,B \leftarrow \bar{B})$.

($2 \Rightarrow 3$) Dado $A = M(f,B)$, construiremos la nueva base requerida a partir de P. Sea $\bar{B}$ la única base en V que verifica: $P = M(I_V,B \leftarrow \bar{B})$. Se tiene entonces:

$$C = P^{-1} \cdot A \cdot P = M(I_V,\bar{B} \leftarrow B) \cdot M(f,B) \cdot M(I_V,B \leftarrow \bar{B}) = M(f,\bar{B})$$

(la primera igualdad por hipótesis y la última por el cambio de bases para aplicaciones lineales).

($3 \Rightarrow 1$) Trivial de la definición de matrices semejantes. ∎

Resulta fácil de comprobar ahora (p. ej, úsese el punto 2 de la proposición anterior):

Corolario 3.54. *La relación binaria "ser semejante a" en $M_n(K)$ es una relación de equivalencia.*

A diferencia del caso de matrices equivalentes, caracterizadas por el rango, no existe una caracterización sencilla de cuándo dos matrices son semejantes. Una condición necesaria la proporciona el siguiente resultado.

Proposición 3.55. *Sean $A,C \in M_n(K)$ dos matrices semejantes. Entonces sus trazas coinciden.*

Demostración. Recordemos primero que si $X,Y \in M_n(K)$ entonces $\mathrm{tr}(X \cdot Y) = \mathrm{tr}(Y \cdot X)$. Usando que $C = P^{-1} \cdot A \cdot P$ se tiene:

$$\mathrm{tr}(C) = \mathrm{tr}\left(P^{-1} \cdot (A \cdot P)\right) = \mathrm{tr}\left((A \cdot P) \cdot P^{-1}\right) = \mathrm{tr}\left(A \cdot (P \cdot P^{-1})\right) = \mathrm{tr}(A)$$

donde se usa la asociatividad del producto y la propiedad anterior de la traza con $X = P^{-1}$, $Y = AP$. ∎

Este resultado tiene la importante consecuencia de que permite dar una definición consistente para la traza de cualquier endomorfismo de un e.v. finitamente generado.

Definición 3.56. *Sea $V(K)$ finitamente generado y $f \in \mathrm{End}(V)$. La* traza *de f se define como la traza de la matriz $M(f,B)$, para cualquier base B de $V(K)$.*

El siguiente ejercicio ilustra la gran diferencia que existe en que dos matrices cuadradas sean semejantes o equivalentes; concretamente, lo muy restrictivo que supone para una matriz el ser semejante a una del tipo $\left(\begin{array}{c|c} I_r & O \\ \hline O & O \end{array}\right)$.

Ejercicio 3.57. *(1) Sea $h \in \mathrm{End}(V)$ que verifica $h \circ h = h$ (a estos endomorfismos se les llama* proyectores, *por las propiedades que se van a ver). Demostrar:*

(a) $V = Im(h) \oplus Nuc(h)$ (sea o no V finitamente generado). Además, si $v = u + w$ con $u \in Im(h)$ y $w \in Nuc(h)$ entonces $h(v) = u$.

(b) Si $dim_K V = n \in \mathbb{N}$, existe una base B tal que

$$M(h,B) = \left(\begin{array}{c|c} I_r & O_{r\times(n-r)} \\ \hline O_{(n-r)\times r} & O_{(n-r)\times(n-r)} \end{array}\right)$$

donde r es el rango de P. Por tanto, en cualquier otra base $\bar{B}$:

$$M(h,\bar{B}) = P^{-1} \cdot \left(\begin{array}{c|c} I_r & O_{r\times(n-r)} \\ \hline O_{(n-r)\times r} & O_{(n-r)\times(n-r)} \end{array}\right) \cdot P,$$

para alguna matriz regular $P \in Gl(n,\mathbb{R})$.

(2) Recíprocamente, si h es un endomorfismo cuya matriz en alguna base $\bar{B}$ satisface la igualdad anterior, entonces h verifica $h \circ h = h$.

(3) Como conclusión para dimensión $n \in \mathbb{N}$, un endomorfismo h verifica $h \circ h$ si y sólo existe alguna base tal que la ecuación en coordenadas de h para esa base es

$$x_1' = x_1, \ldots\ldots, x_r' = x_r, \quad x_{r+1}' = 0, \ldots\ldots, x_n' = 0.$$

(4) (Solución de la ecuación matricial $X^2 = X$.) *Una matriz cuadrada $X \in M_n(K)$ verifica $X \cdot X = X$ si y sólo si es semejante a una del tipo:*

$$\left(\begin{array}{c|c} I_r & O_{r\times(n-r)} \\ \hline O_{(n-r)\times r} & O_{(n-r)\times(n-r)}, \end{array}\right)$$

para algún $r \in \{0,1,\ldots,n\}$.

Finalmente, señalemos como consecuencia una propiedad útil de las matrices.

Corolario 3.58. *Sean $A, C \in M_n(K)$. Si $A \cdot C$ es regular entonces A y C son regulares. En particular, si $A \cdot C = I_n$ entonces $C = A^{-1}$.*

Demostración. Se consideran los endomorfismos f_A y f_C de K^n cuyas matrices en la base usual B_u de K^n son A y C, respectivamente. Como $M(f_A \circ f_C, B_u) = A \cdot C$, que es regular, entonces el rango de $f_A \circ f_C$ es n y, necesariamente, también son iguales a n el rango de f_A y el de f_C (corolario 3.20). Por tanto, f_A y f_C son isomorfismos y las matrices $A = M(f_A, B)$ y $C = M(f, B)$ son regulares. ∎

3.4. Aplicaciones afines y grupo afín

3.4.1. Aplicaciones afines y afinidades

Definición 3.59. *Sean $\mathcal{A},\mathcal{A}'$ dos espacios afines cobre el mismo cuerpo K. Una aplicación $f:\mathcal{A}\to\mathcal{A}'$ es* afín *si existe una aplicación lineal $\vec{f}:\vec{A}\to\vec{A}'$ de manera que para cualesquiera $O,P\in\mathcal{A}$, se cumple:*

$$f(P)=f(O)+\vec{f}(\overline{OP}). \tag{3.2}$$

En el caso de que f sea biyectiva, diremos que f es un isomorfismo afín *y, en este caso, $\mathcal{A}$ es* (afínmente) isomorfo *a $\mathcal{A}'$. Si además, $\mathcal{A}=\mathcal{A}'$, f es una* afinidad.

Observación 3.60. *Es fácil comprobar:*

(1) Equivalen para $f:\mathcal{A}\to\mathcal{A}'$: (i) $f(P)(=f(O+\overline{OP}))=f(O)+\vec{f}(\overline{OP})$, y (ii) $\vec{f}(\overline{OP})=\overline{f(O)f(P)}$

(2) Si $f:\mathcal{A}\to\mathcal{A}'$ verifica la relación (3.2) para un punto O y una aplicación lineal $\vec{f}$ entonces la verificará para todo punto $O'\in\mathcal{A}$ y para la misma aplicación lineal $\vec{f}$ (la cual queda determinada unívocamente).

Proposición 3.61. *Dos aplicaciones afines $f_1,f_2:\mathcal{A}\to\mathcal{A}'$ son iguales si y sólo si: (a) $\vec{f}_1=\vec{f}_2$ y (b) coinciden en un punto, esto es, $f_1(Q)=f_2(Q)$ para algún $Q\in\mathcal{A}$.*

Demostración. Si se verifican (i) y (ii), aplicando (3.2) para $O=Q$, se obtiene que coinciden en todo punto $P\in\mathcal{A}$. Recíprocamente, si f_1 y f_2 son iguales, coinciden en un punto (por coincidir en todos) y, por (3.2), $\vec{f}_1=\vec{f}_2$. ∎

Proposición 3.62. *(Existencia y determinación de aplicaciones afines).*

(1) Fijados $P\in\mathcal{A},P'\in\mathcal{A}'$ y una aplicación lineal $\phi:\vec{\mathcal{A}}\to\vec{\mathcal{A}}'$ existe una única aplicación afín $f:\mathcal{A}\to\mathcal{A}'$ tal que $f(P)=P'$ y $\vec{f}=\phi$.

(2) Si $\mathcal{A}$ tiene dimensión n, fijados $n+1$ puntos independientes $P_0,\dots,P_n\in\mathcal{A}$ y $n+1$ puntos cualesquiera $P'_0,\dots,P'_n\in\mathcal{A}'$ existe una única aplicación afín $f:\mathcal{A}\to\mathcal{A}'$ tal que $f(P_i)=P'_i$ para $i=0,\dots,n$.

Demostración. (1) Basta con comprobar que la aplicación

$$f(Q):=P'+\phi(\overline{PQ}), \qquad \forall Q\in\mathcal{A},$$

satisface las propiedades requeridas.

(2) Se reduce a aplicar el punto anterior anterior con $P=P_0$, $P'=P'_0$, y $\vec{f}:\vec{\mathcal{A}}\to\vec{\mathcal{A}}'$ la única aplicación lineal determinada por $\vec{f}(\overline{P_0P_1})=\overline{P'_0P'_1}$ para $i=1,\dots,n$. ∎

Algunas propiedades de las aplicaciones afines de demostración sencilla son las siguientes:

1. La composición de dos aplicaciones afines en una aplicación afín.

2. Una aplicación afín $f:\mathcal{A}\to\mathcal{A}'$ es inyectiva, exhaustiva o biyectiva si y sólo si lo es $\vec{f}$.

3. La inversa de una aplicación afín biyectiva $\vec{f}$ en una aplicación afín y $\overrightarrow{f^{-1}}=(\vec{f})^{-1}$.

   Como consecuencia de esta propiedad y la 1, la relación "ser afínmente isomorfo a" es una relación de equivalencia en la clase de todos los espacios afines sobre un mismo cuerpo.

4. Dos espacios afines de dimensión finita sobre el mismo cuerpo son isomorfos si y sólo si tienen la misma dimensión n. En particular, son isomorfos a K^n (dotado de su estructura afín natural).

5. Para cada $v \in \vec{\mathcal{A}}$, la *traslación a lo largo de* v, definida por $T_v : \mathcal{A} \to \mathcal{A}, P \mapsto P+v$, es una afinidad que verifica $\vec{f} = \mathrm{Id}$.

   Recíprocamente, si una aplicación afín $f : \mathcal{A} \to \mathcal{A}$ verifica $\vec{f} = \mathrm{Id}$, entonces f es una traslación a lo largo de un vector $v \in \vec{\mathcal{A}}$ a lo largo de un vector v que se puede escribir como $v = \overrightarrow{Pf(P)}$ para cualquier $P \in \mathcal{A}$.

6. La imagen por una aplicación afín f de un subespacio afín $\mathcal{S}$ es un subespacio afín de subespacio director $\vec{f}(\vec{\mathcal{S}})$. En particular, la imagen de f es un subespacio afín de dimensión igual al rango de $\vec{f}$.

3.4.2. Expresión matricial.

Sean $\mathcal{A}$ y $\mathcal{A}'$ dos espacios afines de dimensiones finitas, con referencias afines

$$R = (O, B = (v_1, \dots, v_n)), \qquad R' = (O', B' = (v'_1, \dots, v'_m)),$$

resp., y sea $f : \mathcal{A} \to \mathcal{A}'$ una afinidad. De la relación

$$f(X) = f(O) + \vec{f}(\overrightarrow{OX})$$

se sigue, para las coordenadas afines $x = X_B$, $x' = f(X)'_{B'}$:

$$\mathbf{x}' = b + M\mathbf{x}$$

donde b es la columna de las coordenadas de $f(O)$ en R' y $M = M(\vec{f}, B' \leftarrow B)$.

En forma más compacta, podemos escribir:

$$\begin{pmatrix} x' \\ 1 \end{pmatrix} = \begin{pmatrix} M & b \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ 1 \end{pmatrix}$$

Observación 3.63. (1) Con esta notación, la composición de aplicaciones afines se corresponde con el producto de matrices, de un modo completamente análoga al caso de las aplicaciones lineales.

(2) Tomando como f la aplicación identidad, también se obtienen las ecuaciones del *cambio de sistema de referencia afín*, análogas a las del caso vectorial.

3.4.3. El grupo afín

Nos restringimos a continuación al caso de afinidades. Resulta inmediato de las propiedades listadas de las aplicaciones afines;

Teorema 3.64. *Dado un espacio afín $\mathcal{A}$, sus afinidades forman un grupo con la composición.*

Si la dimensión n es finita, este grupo es isomorfo al de afinidades del espacio afín $\mathbb{K}^n$.

Con este resultado en mente y la expresión matricial vista para las aplicaciones afines, resulta natural definir entonces:

Definición 3.65. *El* grupo afín de dimensión $n \in \mathbb{N}$*, denotado* $GA(n)$*, es el grupo de las afinidades de* K^n*. Fijando en* K^n *la referencia afín usual* $\mathcal{R}_u = ((0,\ldots,0), B_u)$ *y expresando cada afinidad en esta referencia,* $GA(n)$ *se identifica con*

$$\left\{ \begin{pmatrix} M & b \\ 0 & 1 \end{pmatrix} : M \in GL(n), b \in \mathbb{K}^n \right\}$$

siendo $GL(n)$ *el grupo de matrices regulares (lineal general) de orden n, el cual a su vez era identificable naturalmente con el grupo de isomorfismos vectoriales de* $\mathbb{K}^n$*.*

De manera natural, $GL(n)$ y $(\mathbb{K}^n, +)$ pueden verse, respectivamente, como los subgrupos de $GA(n)$

$$\left\{ \begin{pmatrix} M & 0 \\ 0 & 1 \end{pmatrix} : M \in \mathrm{GL}(n) \right\} \qquad \left\{ \begin{pmatrix} I_n & b \\ 0 & 1 \end{pmatrix} : b \in \mathbb{K}^n \right\}.$$

Observación 3.66. Desde un punto de vista algebraico, $GA(n)$ es el *producto semidirecto* de los dos subgrupos anteriores.

Finalmente, decribimos otras afinidades notables, para lo cual introducimos el siguiente concepto.

Definición 3.67. *Dada una afinidad* f *de un espacio afín* $\mathcal{A}$*, diremos que* $P \in \mathcal{A}$ *es un* punto fijo *de* f *si* $f(P) = P$*, que* $C \subset \mathcal{A}$ *es invariante por* f *si* $f(C) = C$*.*

Claramente, si C está formado por puntos fijos entonces es invariante, pero el recíproco no es cierto.

Ejemplo 3.68. A una afinidad $f : \mathcal{A} \to \mathcal{A}$ tal que $\vec{f} = r \cdot I_{\vec{\mathcal{A}}}$ con $r \neq 0, 1$ se le llama homotecia de razón r (obsérvese que $r = 0$ se corresponde con la aplicación afín cuya imagen es un solo punto, la cual no es una afinidad por no serbiyectiva, y $r = 1$ con una traslación).

Una homotecia tiene un único punto fijo, $O \in \mathcal{A}$ (que se suele tomar como origen de coordenadas), el cual puede calcularse escogiendo un punto aunxiliar $P \in \mathcal{A}$ usando:

$$O = P + \frac{1}{1-r}\overline{Pf(P)}.$$

Esta igualdad se deduce de:

$$\begin{aligned} O = f(O) = f(P + \overline{PO}) = f(P) + \vec{f}(\overline{PO}) = f(P) + r\overline{PO} \quad &\Leftrightarrow \quad \overline{PO} = \overline{Pf(P)} + r\overline{PO} \Leftrightarrow \\ \Leftrightarrow \overline{PO} = \tfrac{1}{1-r}\overline{Pf(P)} \quad &\Leftrightarrow \quad O = P + \tfrac{1}{1-r}\overline{Pf(P)} \end{aligned}$$

Ejemplo 3.69. Sea $\mathcal{A}$ un espacio afín de dimensión $n \in \mathbb{N}$. A una afinidad $f : \mathcal{A} \to \mathcal{A}$ tal que $\vec{f}^2 = I_{\vec{\mathcal{A}}}$, con $f \neq I_{\vec{\mathcal{A}}}$ se le llama *simetría afín*.

En el caso $\vec{f} = -I_{\vec{\mathcal{A}}}$ entonces f es también una homotecia de razón $r = -1$, por lo que deja un único punto fijo O, y a la simetría se le llama *simetría central* o *inversión* respecto a O.

En general, se sabe de los ejercicios de aplicaciones lineales que de $\vec{f}^2 = I_{\vec{\mathcal{A}}}$ se deduce $\vec{\mathcal{A}} = V_1 \oplus V_{-1}$, donde $V_1 = \{v \in \vec{\mathcal{A}} : f(v) = v\}$ y $V_{-1} = \{v \in \vec{\mathcal{A}} : f(v) = -v\}$.

Además, no es difícil comprobar:

(a) el punto medio entre cualquier punto y su imagen $P + \overline{Pf(P)}/2$ es un punto fijo, y

(b) escogido un punto fijo O, el conjunto de los puntos fijos es $O + V_1$.

Como casos particulares se tiene:

(i) Si $\dim V_1 = 1$ entonces los puntos fijos de f forman una recta s o *eje de simetría*, y se dice que f es una *simetría (afín) axial* respecto a la recta s.

(ii) Si $\dim V_1 = n - 1$ entonces los puntos fijos de f forman un *hiperplano de simetría* H, y f es una *simetría (afín) especular* respecto al hiperplano H.

3.5. Espacios cocientes y teorema de isomorfía

Hemos visto que las soluciones de un SEL compatible $A\cdot x=b$ se pueden escribir como un conjunto del tipo x_0+ Nucf_A : $\{x_0+y|y\in \text{Nuc}(f_A)\}$ (corolario 3.43). Más prpiamente, es un subespacio afín que puede entenderse como un "subespacio (no vectorial) paralelo" a Nucf_A que pasa por x_0.

Veremos ahora que, en general, dado un e.v. $V(K)$ y un subespacio vectorial suyo U, tiene sentido considerar un nuevo espacio vectorial, que denotaremos V/U, cuyos elementos se escriben de modo natural mediante expresiones del tipo $v+U$ (esto es, son los "subespacios paralelos" a U). Como una aplicación de estos espacios vectoriales, veremos un teorema de isomorfía para aplicaciones lineales, que extiende al teorema del rango.

3.5.1. Espacios vectoriales cocientes

Sea $V(K)$ un e.v. y $U\subset V$ un s.v. Definimos la siguiente relación binaria $\sim$ en V:

$$v\sim w \iff v-w\in U$$

Proposición 3.70. *(1) La relación binaria $\sim$ es de equivalencia.*

(2) Para cada $v\in V$, su clase de equivalencia $[v]$ viene dada por $[v]=\{v+u : u\in U\}$ por lo que la denotaremos $v+U$, esto es:

$$v+U := \{v+u \mid u\in U\}.$$

Asimismo, al conjunto cociente $V/\sim$ formado por todas las clases anteriores, lo denotaremos V/U.

Demostración. (1) Propiedad reflexiva: como $v-v=0\in U$, se tiene $v\sim v$, para todo $v\in V$.

Propiedad simétrica: si $v\sim w$ entonces $v-w=u\in U$ y, al ser U un s.v., $w-v=-u\in U$, esto es, $w\sim v$.

Propiedad transitiva: si $v\sim v'$ y $v'\sim v''$ entonces $v-v'=u\in U$ y $v'-v''=u'\in U$ por lo que $v-v''=(v-v')+(v'-v'')=u+u'\in U$ (la pertenencia a U por ser un s.v.) esto es, $v\sim v''$.

(2) Para demostrar $[v]\subset\{v+u : u\in U\}$, obsérvese que si $v'\in[v]$ entonces $v'-v=u\in U$ por lo que $v'=v+u$ para un $u\in U$. Para la inclusión contraria, como $(v+u)-v=u\in U$ las clases de $v+u$ y v coinciden, esto es, $v+u\in[v]$. ∎

Nos podemos preguntar si V/U hereda una estructura de e.v. a partir de la de V, pues parece natural definir:

$$\begin{array}{rccc} +: & V/U\times V/U & \to & V/U \\ & (v+U,w+U) & \mapsto & (v+w)+U \end{array} \qquad \begin{array}{rccc} \cdot: & K\times V/U & \to & V/U \\ & (a,v+U) & \mapsto & (a\cdot v)+U \end{array} \tag{3.3}$$

No obstante, esta definición presenta la dificultad de que se lleva a cabo escogiendo los representantes v y w de cada clase, y podríamos preguntarnos qué ocurriría si tomáramos otros representantes $v'\in v+U$, $w'\in w+U$, esto es, si la clase de $v'+w'$ y $a\cdot v'$ serían las mismas clases que definen $v+w$ y $a\cdot v$, resp. El siguiente lema proporciona una respuesta afirmativa.

Lema 3.71. *Supongamos que v,v',w,w' satisfacen $v+U=v'+U$, $w+U=w'+U$. Entonces:*

$$(v+w)+U=(v'+w')+U \qquad y \qquad (a\cdot v)+U=(a\cdot v')+U,$$

esto es, las operaciones $+$ y $\cdot$ están consistentemente definidas por las expresiones en (3.3).

Demostración. Por hipótesis, $v - v' = u_v \in U, w - w' = u_w \in U$. Por tanto,

$$(v+w)-(v'+w') = (v-v')+(w-w') = u_v + u_w \in U, \qquad a\cdot v - a\cdot v' = a\cdot(v-v') = a\cdot u_v \in U,$$

esto es, $v'+w' \in (v+w)+U$ y también $a\cdot v' \in (a\cdot v)+U$, como se quería. ∎

Proposición 3.72. *Con las operaciones definidas en* (3.3)*,* $(V/U, + \cdot K)$ *es un espacio vectorial, al cual llamaremos* espacio cociente *de* V *sobre* U*, y denotaremos simplemente* V/U*.*

Demostración. Es una computación directa aplicando las definiciones de las operaciones y las propiedades de e.v. de $V(K)$. Lo demostramos para una de ellas y el resto queda como ejercicio.

Propiedad distributiva del producto por escalares respecto a la suma de vectores:

$$\begin{aligned} a\cdot((v+U)+(w+U)) = &\ a\cdot((v+w)+U) = (a\cdot(v+w))+U = ((a\cdot v)+(a\cdot w))+U \\ = &\ ((a\cdot v)+U)+((a\cdot w)+U) = a\cdot(v+U)+a\cdot(w+U) \end{aligned}$$

donde en las igualdades de la primera línea se usan, por orden, la definición de la suma en V/U, la del producto por escalares en V/U y la propiedad distributiva del producto por escalares en V, mientras que en la segunda se usan la definición de la suma en V/U y la del producto por escalares en V/U. ∎

3.5.2. Base y dimensión de V/U

Como en toda relación de equivalencia, la proyección natural

$$\pi : V \mapsto V/U, \qquad v \mapsto v+U$$

es una aplicación suprayectiva. No obstante, en nuestro caso podemos decir más.

Proposición 3.73. *La proyección* π *es un epimorfismo de e.v. En consecuencia, la imagen por* π *de cualquier sistema de generadores de* V *es un s.d.g. de* V/U*.*

Demostración. Basta con demostrar que π es lineal, puesto que sabemos es suprayectiva y que la última afirmación se cumple para todos los epimorfismos. Con este fin, para todo $v,w \in U$, y $a,b \in K$:

$$\pi(av+bw) = (av+bw)+U = (av+U)+(bw+U) = a\cdot(v+U)+b\cdot(w+U) = a\pi(v)+b\pi(w),$$

la primera y última igualdades por la definición de π, la segunda por la de la suma en V/U y la tercera por la del producto por escalares en V/U. ∎

El siguiente resultado nos permite computar de manera práctica la dimensión de V/U y una base suya, a partir de elementos de V.

Teorema 3.74. *Sea* V *un e.v. finitamente generado y* $U \subset V$ *un subespacio vectorial. Si* W *es un subespacio complementario de* U *(esto es,* $V = U \oplus W$*) y* $B_W = \{w_1, \ldots, w_k\}$ *es una base de* W *entonces*

$$\pi(B_W) = \{w_1+U, \ldots, w_k+U\}$$

es una base de V/U*. Por tanto,* $dim_K(V/U) = dim_K V - dim_K U$*.*

Demostración. Si ampliamos B_W hasta una base $B = \{w_1, \ldots, w_k, u_1, \ldots, u_m\}$ de V, donde $B_U = \{u_1, \ldots, u_m\}$ es una base de U. Sabemos por la proposición anterior que $\pi(B)$ es un s.d.g. de V/U y, como $u_1 + U, \ldots, u_m + U$ son todos iguales a la clase del 0, se pueden suprimir del s.d.g. sin que pierda esta característica, esto es, $\pi(B_W)$ es un s.d.g. de V/U. Para comprobar su independencia lineal, sean $a_1, \ldots, a_k \in K$ tales que:

$$0 + U = a_1(w_1 + U) + \cdots + a_k(w_k + U) = (a_1 w_1 + \cdots + a_k w_k) + U.$$

Esto quiere decir que $a_1 w_1 + \cdots + a_k w_k \in U$, por lo que se puede escribir como combinación lineal de la base B_U. Por tanto, existirán escalares $b_1, \ldots, b_m \in K$ tales que:

$$a_1 w_1 + \cdots + a_k w_k = b_1 u_1 + \cdots + b_k u_m, \quad \text{esto es,} \quad a_1 w_1 + \cdots + a_k w_k - b_1 u_1 + \cdots - b_k u_m = 0.$$

Como $B = B_W \cup B_U$ es una base, todos los coeficientes de esta combinación lineal son 0; en particular, $a_1 = \cdots = a_k = 0$, como se quería. ∎

Observación 3.75. Es fácil comprobar que el teorema anterior es válido incluso en dimensión infinita, en el sentido de que para cualquier suplementario W de U y cualquier base B_W de W su proyección $\pi(B_W)$ es una base de V/U. Como consecuencia, si U es finitamente generado pero V no lo es, entonces V/U tampoco es finitamente generado.

Observación 3.76. El teorema anterior permite construir explícitamente una base de V/U:
Paso 1: Se construye una base $B_U = \{u_1, \ldots, u_m\}$ de U.
Paso 2: Se amplía B_U hasta una base $B = \{u_1, \ldots, u_m, v_{m+1}, \ldots, v_n\}$ de V.
Paso 3: El conjunto $\{v_{m+1} + U, \ldots, v_n + U\}$ es una base de V/U.

Ejercicio 3.77. Se considera en $M_2(K)$ el subespacio U formado por las matrices de traza nula.

(a) Calcular una base $B_{V/U}$ de $M_2(K)/U$ y las coordenadas de $\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$ en esa base.

(b) Si B_0 es la base usual de $M_2(K)$ y $\pi : M_2(K) \to M_2(K)/U$ la proyección canónica, calcular las coordenadas de cada uno de los elementos de $\pi(B_0)$ en la base $B_{V/U}$ obtenida en el punto anterior.

3.5.3. Teorema de isomorfía

Recordemos que toda aplicación $f : X \to Y$ entre dos conjuntos admitía una descomposición canónica $f = i \circ b \circ p$. Para ello, se establece la relación de equivalencia en X: $x \sim x'$ si y sólo si $f(x) = f(x')$ y se define:

- $p : X \to X/\sim$ es la proyección canónica en el espacio cociente, que resulta ser suprayectiva.
- $b : (X/\sim) \to \mathrm{Im}(f)$, dada por $b([x]) = f(x)$, para todo $x \in X$ está consistentemente definida por cómo se estableció $\sim$, y resulta ser biyectiva.
- $i : \mathrm{Im}(f) \to Y$ es la inclusión natural, que resulta ser inyectiva.

Nuestro objetivo es comprobar que, si X, Y son espacios vectoriales $X = V(K)$, $Y = V'(K)$ y f es lineal, entonces la relación $\sim$ coincide con la que define el espacio cociente de V sobre $\mathrm{Nuc}(f)$, de modo que $V/\mathrm{Nuc}(f)$ resulta ser isomorfo a $\mathrm{Im}(f)$.

Teorema 3.78 (Primer teorema de isomorfía.)**.** *Sea $f : V \to V'$ una aplicación lineal, y sea $\sim$ la relación de equivalencia en V dada por $v \sim w$ si y sólo si $f(v) = f(w)$. Entonces*

$$v \sim w \Leftrightarrow v - w \in \mathrm{Nuc}(f), \qquad \textit{esto es}, \qquad [v] = v + \mathrm{Nuc}(f), \tag{3.4}$$

y se verifica $f = i \circ b \circ p$ donde:
(i) $p : V \to V/Nuc(f)$ es un epimorfismo de e.v.
(ii) $b : V/Nuc(f) \to Im(f)$ es un isomorfismo de e.v.
(iii) i :$Im(f)) \to V'$ es un monomorfismo de e.v.

Demostración. Para demostrar (3.4):

$$v \sim w \Leftrightarrow f(v) = f(w) \Leftrightarrow f(v - w) = 0 \Leftrightarrow v \in \mathrm{Nuc}(f),$$

donde en la segunda equivalencia se usa la linealidad de f.

La descomposición $f = i \circ b \circ p$ (incluyendo la consistencia de la definición de b) se sabe que ocurre para cada aplicación, sea lineal o no. No obstante, la linealidad de f asegura que también lo son tanto la inclusión i como la proyección p (esta última por la proposición 3.73). Basta por tanto con demostrar la linealidad de b, para lo cual:

$$\begin{aligned} b((a \cdot (v + \mathrm{Nuc}(f)) + c \cdot (w + \mathrm{Nuc}(f))) = &\quad b((av + cw) + \mathrm{Nuc}(f)) = f(av + cw) = af(v) + cf(w) \\ = &\quad a \cdot b(v + \mathrm{Nuc}(f)) + c \cdot b(w + \mathrm{Nuc}(f)) \end{aligned}$$

para todo $v + \mathrm{Nuc}(f), w + \mathrm{Nuc}(f) \in V/\mathrm{Nuc}(f)$, y todo $a, c \in K$, donde se ha usado sucesivamente la definición de las operaciones en V/U, la definición de b, la linealidad de f y de nuevo la definición de b. ∎

Observación 3.79. Por el punto (ii) se tiene que $V/\mathrm{Nuc}(f)$ es isomorfo a $\mathrm{Im}(f)$. Puesto que, en dimensión finita para V, sabemos (teorema 3.74) que la dimensión del primero es $\dim_K(V/\mathrm{Nuc}(f)) = \dim_K V - \dim{}_K \mathrm{Nuc}(f)$, se reobtiene ahora la fórmula del rango: $\dim_K V = \dim_K \mathrm{Nuc}(f) + \dim_K \mathrm{Im}(f)$.

Ejercicio 3.80. *Sea $V(K)$ un espacio vectorial y $U, W \subset V$ dos subespacios. Demostrar:*

1. *Si $V = U \oplus W$ entonces V/W es isomorfo a U*

2. *Con más generalidad (2º Teorema de isomorfía): $(U + W)/W$ es isomorfo a $U/(U \cap W)$.*

   (Suger. Compruébese que la aplicación $g : U \to (U + W)/W$ es un epimorfismo con $Nuc(g) = U \cap W$ y úsese el teorema 3.78.)

3. *(3er Teorema de isomorfía.) Si $U \subset W$, entonces $(V/U)/(W/U)$ es isomorfo a V/W.*

   (Suger. Compruébese que la aplicación $h : V/U \to V/W$ está bien definida y es un epimorfismo con $Nuc(h) = U/W$ y úsese el teorema 3.78.)

En el caso $\dim V < \infty$, compruébese la consistencia de los isomorfismos con el cálculo de dimensiones por la fórmula de Grassmann.

Tema 4

El espacio dual

El espacio dual es un caso particular del espacio de todas las aplicaciones lineales entre dos e.v. prefijados $\mathrm{Lin}(V,V')$ para el cual V' es el cuerpo $K(K)$ sobre el que está definido V. Por tanto, podremos aplicar a él todo lo que sabemos sobre $\mathrm{Lin}(V,V')$. Entre otras aplicaciones, este concepto permite dar una interpretación geométrica de la trasposición de una matriz y sus propiedades.

Restringiremos nuestro estudio al caso finitamente generado. Así, en adelante, $V(K)$ será un e.v. de dimensión $n \in \mathbb{N} \cup \{0\}$.

4.1. Concepto de espacio dual y forma lineal

Definición 4.1. *Dado un e.v. $V(K)$, su* espacio dual *es el espacio*

$$V^*(K) := \mathrm{Lin}(V,K) = \{\phi : V \to K : \ \phi \textit{ lineal}\}.$$

A cada elemento del dual $\phi \in V^*(K)$ se le llamará forma lineal de V.

De la propia definición se sigue inmediatamente:

Corolario 4.2. *Dado un e.v. V de dimensión n, el espacio dual $V^*(K)$, dotado de sus operaciones naturales, es un espacio vectorial de la misma dimensión n.*

Demostración. Basta con particularizar las propiedades generales de $\mathrm{Lin}(V,V')$ para concluir que $V^* = \mathrm{Lin}(V,\mathbb{R})$ es un e.v., cuya dimensión resulta ser: $\dim V \cdot \dim K = n \cdot 1 = n$. ∎

Si fijamos una base ordenada $B = (v_1,\dots,v_n)$ de $V(K)$ y tomamos $\{1\}$ como elección estándar de una base en $K(K)$ entonces para cada $\phi \in V^*(K)$ podemos calcular la matriz de la aplicación ϕ expresada en dichas bases $M(\phi,\{1\} \leftarrow B)$, la cual resulta ser igual a $(\phi(v_1)\dots\phi(v_n))$. De esta forma, si $v = \sum_{i=1}^n a_i v_i \in V$ entonces

$$\phi(v) = \sum_{i=1}^n \phi(v_i)a_i = (\phi(v_1)\dots\phi(v_n)) \cdot \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}$$

$$= M(\phi, \{1\} \leftarrow B) \cdot \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} \in K.$$

Por ejemplo, dado $\phi : \mathbb{R}^3 \to \mathbb{R}$, $\phi(x,y,z) = x - 2y + 3z$, se tiene $M(\phi, \{1\} \leftarrow B) = (1, -2, 3)$.

Ejercicio 4.3. *Demostrar que toda forma lineal $\phi \in (K^n)^*(K)$ se escribe como $\phi(x_1, \dots, x_n) = \sum_{i=1}^n b_i x_i$ para ciertos $b_1, \dots, b_n \in K$.*

Observación 4.4. Por el ejercicio anterior, el núcleo de una forma lineal $\phi \in (K^n)^*(K)$ es el conjunto de soluciones del sistema líneal homogéneo:

$$b_1 x_1 + \cdots + b_n x_n = 0,$$

cuyas soluciones, siempre que ϕ no sea la forma lineal nula, constituyen un subespacio vectorial de dimensión $n-1$. Estas propiedades se generalizan automáticamente a cualquier espacio vectorial V, sin más que fijar una base ordenada y calcular $M(\phi, \{1\} \leftarrow B)$.

La importancia del núcleo para las formas lineales se pone de manifiesto a continuación.

Proposición 4.5. *Sea $\phi \in V^*(K)$:*

(a) Si ϕ no es la forma lineal nula entonces $Nuc(\phi)$ es un hiperplano vectorial, esto es, $dim_K Nuc(\phi) = n-1$.

(b) Si $H \subset V$ es un hiperplano vectorial, existe un forma lineal ϕ tal que $Nuc(\phi) = H$.

(c) Dos formas lineales no nulas ϕ, ψ son proporcionales (esto es, $\psi = a\phi$ para algún $a \in K \setminus \{0\}$) si y sólo si sus núcleos coinciden.

Demostración. (a) Por el teorema del rango, $\dim_K \text{Nuc}(\phi) = n - \dim_K \text{Im}(\phi)$. Como $\text{Im}(\phi) \subset K$ y ϕ no es la forma lineal nula, $\text{Im}(\phi) = K$, de donde $\dim_K \text{Im}(\phi) = 1$ y se sigue el resultado[1].

(b) Sea $B_H = \{v_1, \dots, v_{n-1}\}$ una base de H y sea $v_n \in V$ cualquier vector que amplía B_H a una base B de V (esto es, $v_n \notin H$). La única forma lineal $\phi : V \to K$ tal que

$$\phi(v_1) = \cdots = \phi(v_{n-1}) = 0, \qquad \phi(v_n) = 1$$

satisface claramente $H \subset \text{Nuc}(\phi) \subsetneq V$. Como la dimensión de H coincide con la de $\text{Nuc}(\phi)$ (por el apartado (a)), $H = \text{Nuc}(\phi)$.

(c) Claramente, de $\psi = a\phi$ se sigue $\phi(v) = 0 \Rightarrow \psi(v) = 0$, esto es, $\text{Nuc}(\phi) \subset \text{Nuc}(\psi)$ y, como $a \neq 0$, también se verifica la inclusión contraria. Recíprocamente, llamemos H al núcleo común, $H = \text{Nuc}(\phi) = \text{Nuc}(\psi)$ y, como en el apartado anterior, construyamos una base B ampliando una B_H de H. Se tiene entonces:

$$\begin{aligned} M(\phi, \{1\} \leftarrow B) &= (0, \dots, 0, \phi(v_n)), \qquad & \phi(v_n) \neq 0 \\ M(\psi, \{1\} \leftarrow B) &= (0, \dots, 0, \psi(v_n)), \qquad & \psi(v_n) \neq 0 \end{aligned}$$

por lo que basta con tomar $a = \psi(v_n) \cdot \phi(v_n)^{-1}$. ∎

Ejercicio 4.6. *(1) Si $\phi, \psi \in V^*$, demostrar: $\psi = a\phi$ para algún $a \in K$ si y sólo si $Nuc(\phi) \subset Nuc(\psi)$.*

(2) Demostrar que las aplicaciones $\phi, \psi : \mathbb{R}^2 \to \mathbb{R}, \phi(x,y) = 2x + y, \psi(x,y) = x - 2y$ son formas lineales independientes. Representar gráficamente los núcleos de $\phi, \psi, 2\phi$ y $\phi + \psi$ así como las preimágenes de $1 \in \mathbb{R}$ para cada una de esas formas lineales.

[1] La observación 4.4 sugiere una demostración con un punto de vista alternativo.

4.2. Base dual

Consideremos los espacios vectoriales $V(K)$, $V^*(K)$ y fijemos una base $B = (v_1, \ldots, v_n)$ de $V(K)$.

Teorema 4.7. *Existe una única base $B^* = (\phi^1, \ldots, \phi^n)$ de $V^*(K)$ que verifica*

$$\phi^i(v_j) = \delta^i_j \quad \forall i, j \in \{1, \ldots, n\}, \qquad \text{(donde } \delta^i_j \text{es la delta de Kronecker)}. \tag{4.1}$$

A esta base B^* se le llamará base dual *de la base B.*

Demostración. Mediante el procedimiento de extensión lineal (teorema 3.8) se sigue inmediatamente que, para cada índice $i \in \{1, \ldots, n\}$ las relaciones (4.1) definen unívocamente una forma lineal. Por tanto, basta con demostrar que el conjunto B^* es linealmente independiente (recuérdese que $\dim_K V = n$). Sean $a_1, \ldots, a_n \in K$ tales que:

$$a_1\phi^1 + \cdots + a_n\phi^n = 0$$

donde el 0 a la derecha debe entenderse como la forma lineal nula ($0(v) = 0, \forall v \in V$). Aplicando ambos miembros de la igualdad a cada vector de la base v_j se obtiene

$$0 = \sum_{i=1}^{n} a_i\phi^i(v_j) = \sum_{i=1}^{n} a_i\delta^i_j = a_j$$

como se quería. ∎

Corolario 4.8. *Si $v \in V$ verifica $\phi(v) = 0$ para todo $\phi \in V^*$ entonces $v = 0$.*

Demostración. Si $v \neq 0$ se puede ampliar una base $B = (v = v_1, v_2, \ldots, v_n)$ de V, y el primer elemento ϕ^1 de B^* verifica $\phi^1(v) = 1 \neq 0$. ∎

Observación 4.9. Para los elementos de la base dual se tiene

$$\begin{array}{rcl} M(\phi^1, \{1\} \leftarrow B) & = & (1, 0, \ldots, 0) \\ \vdots & & \vdots \\ M(\phi^n, \{1\} \leftarrow B) & = & (0, 0, \ldots, 1). \end{array}$$

Esto es, la base dual coincide con la base $B_{B,B'}$ obtenida en general para $\mathrm{Lin}(V, V')$ en el teorema 3.48 para $V' = \mathbb{R}$, $B' = \{1\}$. De hecho, como en el caso de $\mathrm{Lin}(V, V')$, se tiene para cualquier forma lineal ϕ que $M(\phi, \{1\} \leftarrow B)$ coincide con las coordenadas (escritas por filas) de ϕ en la base B^*.

Esta última observación justifica el siguiente convenio.

Convenio. Usaremos la notación $M(\phi, B) := M(\phi, \{1\} \leftarrow B)$, y abusaremos del lenguaje llamando a esta matriz (tipo fila) *coordenadas de* ϕ *en* B (en lugar de coordenadas de ϕ en B^*, que propiamente es una matriz columna). Por este motivo, también denotaremos $\phi_B = M(\phi, B)$; esto es, $\phi_B = (\phi_{B^*})^t$.

Observación 4.10. Fijada la base $B = (v_1, \ldots, v_n)$ y su base dual $B^* = (\phi^1, \ldots, \phi^n)$ se verifica $\phi^j(v) = \phi^j(\sum_{i=1}^n a^i v_i) = \sum_{i=1}^n a^i \phi^j(v_i) = \sum_{i=1}^n a^i \delta^j_i = a^j$. Esto es,

$$v = \phi^1(v)v_1 + \cdots + \phi^n(v)v_n = \sum_{i=1}^{n} \phi^i(v)v_i, \quad \forall v \in V.$$

Análogamente, si $\phi \in V^*$ y $\phi = \sum_{i=1}^n b_i \phi^i$ entonces $\phi(v_j) = \sum_{i=1}^n b_i \delta^i_j = b_j$. Esto es,

$$\phi = \phi(v_1)\phi^1 + \cdots + \phi(v_n)\phi^n = \sum_{i=1}^n \phi(v_i)\phi^i, \quad \forall \phi \in V^*. \tag{4.2}$$

Convenio. Los siguientes convenios sobre las posiciones de índices (que se seguirán en adelante) resultan imprescindibles para el estudio sistemático y uso práctico de tensores. El lector sólo interesado en el espacio dual puede o no seguirlos.

Los índices de los vectores de la base B se escribirán como subíndices $B = (v_1, \ldots, v_n)$ mientras que las coordenadas de un vector v en la base B se escribirán como superíndices $(a^1, \ldots, a^n)$ de modo que se tiene: $v = \sum_i a^i v_i$.

Los índices de las formas de la base B^* se escribirán como superíndices $B = (\phi^1, \ldots, \phi^n)$ mientras que las coordenadas de una forma ϕ en la base B^* se escribirán como subíndices $(b_1, \ldots, b_n)$ de modo que se tiene: $\phi = \sum_i b_i \phi^i$.

Las matrices de cambio de base que se estudiarán a continuación serán consistentes con este convenio, levantando cuando sea preciso el índice de fila (para el cambio en V) o el de la columna (para el cambio en V^*); esta notación será también consistente con el convenio que se seguirá para aplicaciones lineales y trasposición, tomando esas aplicaciones como la identidad.

Cambio de base en el dual

Sean $B = (v_1, \ldots, v_n)$, $\overline{B} = (\bar{v}_1, \ldots, \bar{v}_n)$ dos bases de $V(K)$ y $B^* = (\phi^1, \ldots, \phi^n)$, $\overline{B}^* = (\overline{\phi}^1, \ldots, \overline{\phi}^n)$ sus respectivas bases duales. Supongamos que

$$\bar{v}_j = \sum_{i=1}^n a^i{}_j v_i, \qquad \forall j \in \{1, \ldots, n\} \tag{4.3}$$

es decir,

$$M(I_V, B \leftarrow \overline{B}) = (a^i{}_j)_{i,j} = \begin{pmatrix} a^1{}_1 & \ldots & a^1{}_n \\ \vdots & \ddots & \vdots \\ a^n{}_1 & \ldots & a^n{}_n \end{pmatrix}. \tag{4.4}$$

Siguiendo estos convenios de posición de índices, si $v \in V$ y escribimos $v = \sum_{i=1}^n a^i v_i = \sum_{j=1}^n \overline{a}^j \bar{v}_j$ se tiene $a^i = \sum_{i=1}^n a^i{}_j \overline{a}^j$. Esto es,

$$\begin{pmatrix} a^1 \\ \vdots \\ a^n \end{pmatrix} = M(I_V, B \leftarrow \overline{B}) \cdot \begin{pmatrix} \overline{a}^1 \\ \vdots \\ \overline{a}^n \end{pmatrix}.$$

Proposición 4.11. *Con la notación anterior:*

$$M(I_{V^*}, \overline{B}^* \leftarrow B^*) = M(I_V, B \leftarrow \overline{B})^t. \tag{4.5}$$

Demostración. Escribimos ahora $\phi^k = \sum_{j=1}^n b_j^{\ k} \bar{\phi}^j$ donde los elementos de la matriz $(b_j^{\ k})$ serán los de la matriz $M(Id, \overline{B}^* \leftarrow B^*)$ (el índice superior es ahora el segundo, por lo que indicará columna y no fila). Así, (4.5) se sigue de $b_j^{\ k} = \phi^k(\bar{v}_j) = a^k{}_j$ (la primera igualdad por (4.2)). ∎

Observación 4.12. Como consecuencia de (4.5), si $\phi = \sum_{i=1}^{n} b_i \phi^i = \sum_{j=1}^{n} \bar{b}_j \bar{\phi}^j$ se tiene

$$\bar{b}_j = \sum_{i=1}^{n} a^i{}_j b_i,$$

(o bien, directamente, $\bar{b}_j = \phi(\bar{v}_j) = \sum_{i=1}^{n} a^i{}_j \phi(v_i) = \sum_{i=1}^{n} a^i{}_j b_i$). Es de remarcar que, en las igualdades anteriores, si un índice aparece libre en un miembro de la igualdad (esto es, no se corresponde al de un sumatorio en ese miembro), *también aparece libre y en la misma posición (arriba o abajo)* en el otro miembro.

Ejercicio. Se consideran en $\mathbb{R}^3$ la base usual B_u y la base

$$\overline{B} = ((0,1,1),(1,0,1),(-1,-1,0)).$$

Hallar las coordenadas de la forma lineal $\phi(x,y,z) = x + 2z$ en B_u^* y $\overline{B}^*$, así como la matriz de cambio de base $M(I_{V^*}, B_u^* \leftarrow \overline{B}^*)$.

4.3. Teorema de Reflexividad

Sean $V(K)$ y $V^*(K)$ un espacio vectorial y su dual, respectivamente. Podemos considerar el dual de $V^*(K)$, o *bidual* de $V(K)$: $V^{**}(K) = (V^*(K))^*$. Estos tres espacios vectoriales tienen igual dimensión y, por tanto, son isomorfos. Sin embargo, mientras que no existe ningún isomorfismo canónico general entre $V(K)$ y $V^*(K)$, sí podemos definir uno entre $V(K)$ y $V^{**}(K)$. Ello, en la práctica, equivale a considerar ambos espacios como iguales y a que no nos resulte necesario recurrir a espacios como el dual del bidual $V^{***}(K)$ (que sería naturalmente isomorfo al dual $V^*(K)$), etc.

Lema 4.13. *Fijado un vector $v \in V$ la aplicación*

$$\begin{aligned} \Phi_v : V^* &\to K \\ \phi &\mapsto \phi(v) \end{aligned}$$

es lineal y, por tanto, pertenece al bidual $V^{**}(K)$.

Demostración. Aplicando la definición de Φ_v,

$$\Phi_v(a\phi + b\psi) = (a\phi + b\psi)(v) = a\phi(v) + b\psi(v) = a\Phi_v(\phi) + b\Phi_v(\psi)$$

para todo $\phi, \psi \in V^*$ y para todo $a, b \in K$. ∎

Teorema 4.14. *(de Reflexividad). La aplicación*

$$\begin{aligned} \Phi : V &\to V^{**} \\ v &\mapsto \Phi_v \end{aligned}$$

es un isomorfismo de espacios vectoriales.

Demostración. Por el lema anterior, la aplicación Φ está definida consistentemente. Para demostrar que es lineal, debemos comprobar que $\Phi_{av+bw}(=\Phi(av+bw))$ es igual a $a\Phi_v+b\Phi_w(=a\Phi(v)+b\Phi(w))$ para todo $v,w\in V$, $a,b\in K$. Para ello, aplicamos ambos a una forma lineal genérica $\phi\in V^*$:

$$\begin{aligned}\Phi_{av+bw}(\phi)= &\quad \phi(av+bw)=a\phi(v)+b\phi(w)\\ (a\Phi_v+b\Phi_w)(\phi)= &\quad a\Phi_v(\phi)+b\Phi_w(\phi)=a\phi(v)+b\phi(w),\end{aligned}$$

comprobándose que son iguales. Para demostrar la biyectividad de Φ, basta con comprobar su inyectidad (al tener V y V^{**} la misma dimensión finita)[2], esto es, $\mathrm{Nuc}(\Phi)=\{0\}$. Sea $v\in V$ tal que Φ_v es la forma nula del bidual. Esto quiere decir $\Phi_v(\phi)=0$, esto es, $\phi(v)=0$, para todo $\phi\in V^*$, lo que implica que v es 0 (véase el corolario 4.8). ∎

Corolario 4.15. *Toda base $B'=(\phi^1,\dots,\phi^n)$ de $V^*(K)$ es la base dual de una única base B de $V(K)$.*

Demostración. En efecto, dada la base del dual B', podemos tomar su base dual $B'^*\subset V^{**}$, y escribir $B'^*=(\Phi_{v_1},\dots,\Phi_{v_n})$ donde $B=(v_1,\dots,v_n)$ es la base $\Phi^{-1}(B'^*)$ de $V(K)$, proporcionada por el teorema de reflexividad. Se tiene entonces

$$\delta^i_j=\Phi_{v_j}(\phi^i)=\phi^i(v_j)\ , \qquad \forall i,j\in\{1,\dots,n\},$$

la primera igualdad por ser B'^* la base dual de B' y la segunda por la definición de Φ. En consecuencia, B' satisface la ecuación $\phi^i(v_j)=\delta^i_j$ que define la base dual de B, esto es, $B^*=B'$. ∎

Observación 4.16. El significado del Teorema de Reflexividad puede entenderse como sigue. Sean B, $\overline{B}$ dos bases de $V(K)$. Sabemos que existe un único isomorfismo $F:V\to V^*$ que, de manera ordenada, aplica B en B^* y, análogamente, un único isomorfismo $G:V^*\to V^{**}$ que aplica B^* en su base dual B^{**}. De igual modo, con $\overline{B}$ obtenemos isomorfismos $\overline{F}:V\to V^*$, $\overline{G}:V^*\to V^{**}$. En general, $F\neq\overline{F}$ y $G\neq\overline{G}$. Sin embargo, las composiciones $G\circ F,\overline{G}\circ\overline{F}:V\to V^{**}$ sí verifican $G\circ F=\overline{G}\circ\overline{F}$; de hecho, por el corolario 4.15 ambos coinciden con el isomorfismo que proporciona el Teorema de Reflexividad.

4.4. Anuladores

Como se comentó en la observación 4.4, el núcleo de una forma lineal siempre se puede ver como la solución de un sistema lineal homogéneo de una sola ecuación:

$$a_1x^1+\cdots+a_nx^n=0.$$

En el caso de que tengamos un sistema de m ecuaciones, cada fila de su matriz $A=(a^i{}_j)$ proporciona una forma lineal ϕ^i que “anula” a todas las soluciones del sistema. En general, todo subespacio vectorial U puede verse de este modo cuando escribimos unas ecuaciones implícitas para él. Desarrollamos a continuación estas ideas.

Definición 4.17. *Dado cualquier subconjunto $S\subset V$, se define su* anulador *en V^* como*

$$an(S)=\{\phi\in V^*:\phi(v)=0,\forall v\in S\}.$$

[2] Es de remarcar que esto no ocurre en dimensión infinita, por lo que el estudio del espacio dual es más complicado entonces.

Proposición 4.18. *El anulador verifica las siguientes propiedades:*

(1) $an(S)$ es un subespacio vectorial de V^*.

(2) Si $S \subset S'$ entonces $an(S) \supset an(S')$.

(3) $an(S) = an(L(S))$

(4) $an(\{0\}) = V^*$; $an(V) = \{0\}$ (donde $0 \in V^*$ es la forma lineal nula).

Demostración. (1) Sean $\phi, \psi \in \mathrm{an}(S)$, $a, b \in K$. Entonces, para cualquier vector v de S:

$$(a\phi + b\psi)(v) = a\phi(v) + b\psi(v) = a \cdot 0 + b \cdot 0 = 0,$$

esto es, $a\phi + b\psi \in \mathrm{an}(S)$

(2) Es inmediato (¡compruébese!).

(3) La inclusión $\supset$ es consecuencia del apartado anterior. Para la contraria, sea $\phi \in \mathrm{an}(S)$ y $w \in L(S)$, esto es, $w = \sum_{i=1}^m a_i v_i$ con $v_1, \ldots, v_m \in S$. Entonces,

$$\phi(w) = \phi(\sum_{i=1}^m a_i v_i) = \sum_{i=1}^m a_i \phi(v_i) = \sum_{i=1}^m a_i \cdot 0 = 0.$$

(4) Trivialmente, $\mathrm{an}(\{0\}) \subset V^*$, y la inclusión contraria se da porque $\phi(0) = 0$ para todo $\phi \in V^*$ (al ser ϕ una aplicación lineal).

Trivialmente, $\mathrm{an}(V) \supset \{0\}$, y la inclusión contraria equivale a decir que si $v \in V$ verifica $\phi(v) = 0$ para todo $\phi \in V$ entonces $v = 0$, lo que ya es conocido (corolario 4.8). ∎

Teorema 4.19. *Sea U un subespacio vectorial de $V(K)$ de dimensión m. Entonces:*

(1) $\dim_K(an(U)) = n - m$.

(2) Si Φ es el isomorfismo del teorema de reflexividad entonces $\Phi(U) = an(an(U))$, esto es, con la identificación natural de V y V^{**}, $U = an(an(U))$.

(3) Dado otro subespacio W, se tiene que $U = W$ si y sólo si $an(U) = an(W)$.

Demostración. (1) Sea $B = (v_1, \ldots, v_m, v_{m+1}, \ldots, v_n)$ una base de todo $V(K)$ obtenida ampliando una $B_U = (v_1, \ldots, v_m)$ de U. La base dual $B^* = (\phi^1, \ldots, \phi^m, \phi^{m+1}, \ldots, \phi^n)$ verifica $\{\phi^{m+1}, \ldots, \phi^n\} \subset \mathrm{an}(U)$ trivialmente. Puesto que este conjunto es linealmente independiente y consta de $n-m$ elementos, basta con comprobar que es un sistema de generadores de $\mathrm{an}(U)$. Para ello, sea $\phi \in \mathrm{an}(U)$ y escribámoslo como combinación lineal de B^*:

$$\phi = b_1\phi^1 + \ldots b_m\phi^m + b_{m+1}\phi^{m+1} + \cdots + b_n\phi^n.$$

Como cada b_i verifica $b_i = \phi(v_i)$ y los primeros m vectores de B pertenecen a U (que está contenido en el núcleo de ϕ), se sigue $b_1 = \cdots = b_m = 0$, como se quería demostrar.

(2) Como por el apartado anterior $\dim_K(\mathrm{an}(\mathrm{an}(U)) = n - \dim_K(\mathrm{an}(U)) = \dim_K(U)$, la cual coincide con $\dim_K(\Phi(U))$, basta con demostrar la inclusión $\subset$, que es inmediata (para cada $u \in U$ y todo $\phi \in \mathrm{an}(U)$ se tiene $\Phi_u(\phi) = \phi(u) = 0$, esto es, $\Phi_u \in \mathrm{an}(\mathrm{an}\ (U))$ como se quería)

(3) La implicación hacia la derecha es trivial. Hacia la izquierda, de $\mathrm{an}(U) = \mathrm{an}(W)$ se sigue $\mathrm{an}(\mathrm{an}(U)) = \mathrm{an}(\mathrm{an}(W))$ y por el apartado anterior $U = W$. ∎

Observación 4.20. Desde el punto de vista teórico, el punto (1) del teorema 4.19 permite construir explícitamente una base del anulador de S:

Paso 1: se determina una base $B_S = \{v_1, \ldots, v_m\}$ de $L(S)$.
Paso 2: se amplía hasta una base $B = (v_1, \ldots, v_m, v_{m+1}, \ldots, v_n)$ de $V(K)$.
Paso 3: Se construye la base dual $B^* = (\phi^1, \ldots, \phi^m, \phi^{m+1}, \ldots, \phi^n)$.
Paso 4: $\{\phi^{m+1}, \ldots, \phi^n\}$ es la base de an(S) que se buscaba.
Además, puede usarse que si se construyó la base B calculando sus coordenadas respecto a otra base prefijada B_0, esto es, si se conoce $M(I_V, B_0 \leftarrow B)$, entonces el Paso 4 se resuelve sin más que tener en cuenta $M(I_V, B_0 \leftarrow B)^t = M(I_{V^*}, B^* \leftarrow B_0^*)$, y calcular la inversa de esta última matriz.

No obstante, desde un punto de vista práctico, debe tenerse en cuenta lo siguiente. Siempre que se tenga un subespacio vectorial U determinado por unas ecuaciones implícitas (bien porque se esté trabajando en $K^n(K)$ o, con más generalidad, porque se haya prefijado una base B de V y se esté trabajando con coordenadas en esa base):

$$\begin{array}{cccc} a^1{}_1 x^1 + & \ldots & + a^1{}_n x^n & = 0 \\ \ldots & \ldots & \ldots & \ldots \\ a^m{}_1 x^1 + & \ldots & + a^m{}_n x^n & = 0 \end{array}$$

entonces la fila i-ésima puede verse como la ecuación del núcleo de la forma lineal ψ^i tal que

$$\psi^i_B = (a^i{}_1, \ldots, a^i{}_n)$$

de modo que una base de an(U) es, directamente,[3] $\{\psi^1, \ldots, \psi^m\}$.
Por otra parte, el punto (2) del teorema anterior asegura que, si $\{\psi^1, \ldots, \psi^m\}$ genera el anulador de U, entonces U es el anulador de $\{\psi^1, \ldots, \psi^m\}$.

Observación 4.21. En resumen, desde un punto de vista práctico, en el caso de que U esté determinado por ecuaciones implícitas, estas ecuaciones proporcionan directamente an(U). En el caso de que U está determinado por paramétricas, bastarña entonces con calcular sus ecuaciones implícitas, para lo cual podemos aplicar lo estudiado en la Sección 2.4.4 (usando eventualmente propiedades de los determinantes).

Ejemplo 4.22. En el Ejemplo 2.175, una base de an(U) está formada por las formas lineales $\{\phi, \psi\}$:

$$\phi(x_1, x_2, x_3, x_4) = -3x_1 + 5x_2 - 4x_3, \quad \psi(x_1, x_2, x_3, x_4) = x_1 + 9x_2 - 4x_4 \qquad \forall (x_1, x_2, x_3, x_4) \in \mathbb{R}^4.$$

Ejercicio 4.23. *Se consideran los siguientes subespacios vectoriales de* $\mathbb{R}^3$*:*

$$U = \{(x, y, z) \in \mathbb{R}^3 : x - y + z = 0\}, \qquad W = L(\{(1,1,1), (3,2,0)\}.$$

Calcular bases de an(U)*, an*(W)*, an*$(U + W)$ *y an*$(U \cap W)$*.*

Ejercicio 4.24. *Sean* U, W *dos subespacios vectoriales de* $V(K)$*. Demostrar:*
(1) an $(U + W)$ *=an*$(U) \cap$ *an*(W)*,*
(2) an $(U \cap W) =$ *an*$(U)+$ *an*(W)*.*
Nota: *es posible demostrar el apartado (2) a partir del apartado (1) (tómense los anuladores de cada miembro de (2) y aplíquense (1) y el teorema de reflexividad).*

[3] Por supuesto, si no estuviera garantizado que las ecuaciones son independientes, este conjunto de formas sería un sistema de generadores de an(U), del cual se podría extraer una base.

Ejercicio 4.25. *Dado cualquier subconjunto $\tilde{S} \subset V^*$ definimos su anulador en V como*

$$an(\tilde{S}) = \{v \in V : \phi(v) = 0, \forall \phi \in \tilde{S}\} \quad (= \cap_{\phi \in \tilde{S}} \mathrm{Nuc}(\phi)).$$

(1) Enunciar propiedades para $an(\tilde{S})$ análogas a las vistas para $an(S)$ en la proposición 4.18, comprobándolas de dos maneras: (a) directamente, y (b) aplicando el teorema de reflexividad.

(2) Comprobar que, en el caso finito, $\tilde{S} = \{\phi^1, \ldots, \phi^m\}$, su anulador puede escribirse como la solución de un SEL homogéneo de m ecuaciones y n incógnitas (véase la observación 4.4).

(3) Demostrar que si $\tilde{S} \subset V^*$, su anulador en V, $an_V(\tilde{S})$ y su anulador en V^{**}, $an_{V^{**}}(S)$ están relacionados por el teorema de reflexividad, esto es, $\Phi(an_V(\tilde{S})) = an_{V^{**}}(\tilde{S})$.

4.5. Trasposición de aplicaciones lineales

Sean $V(K)$, $V'(K)$ dos espacios vectoriales de dimensiones $n, m \in \mathbb{N} \cup \{0\}$ y $f : V \to V'$ una aplicación lineal. Para cada $\phi' \in V'^*(K)$ podemos considerar la composición de aplicaciones lineales $\phi' \circ f : V \to K$. Esta composición también lineal y, por tanto, $\phi' \circ f \in V^*(K)$.

Definición 4.26. *Dada una aplicación lineal $f : V \to V'$, su* aplicación traspuesta *f^t se define por:*

$$\begin{aligned} f^t : V'^* &\to V^* \\ \phi' &\mapsto f^t(\phi') := \phi' \circ f. \end{aligned}$$

Las propiedades de la trasposición se resumen en el siguiente resultado.

Proposición 4.27. *Sea $f : V \to V'$ lineales.*

(1) La aplicación traspuesta f^t es lineal.

(2) $I_V^t = I_{V^*}$, esto es, la traspuesta de la aplicación identidad en V es la identidad en V^*.

(3) Si V'' es otro e.v. y $h : V' \to V''$ es lineal entonces $(h \circ f)^t = f^t \circ h^t$.

(4) $(f+g)^t = f^t + g^t$ y $(a \cdot f)^t = a \cdot f^t$, para todo $g \in \mathrm{Lin}(V, V')$, $a \in K$. Más aún, la aplicación trasposición

$$\begin{aligned} t : \mathrm{Lin}(V, V') &\to \mathrm{Lin}(V'^*, V^*) \\ f &\mapsto f^t \end{aligned}$$

es un isomorfismo de espacios vectoriales.

(5) Con la identificación natural de V con V^{**} y V' con V'^{**},

$$f = (f^t)^t \qquad \text{esto es} \qquad f = (\Phi')^{-1} \circ (f^t)^t \circ \Phi$$

donde $\Phi : V \to V^{**}$, $\Phi' : V' \to V'^{**}$ son los isomorfismos naturales por reflexividad.

Demostración. (1) Veamos que, para todo $\phi', \psi' \in V'^*$ y $a, b \in K$ se tiene $f^t(a \cdot \phi' + b \cdot \psi') = a \cdot f^t(\phi') + b \cdot f^t(\psi')$. Como ambas expresiones pertenecen a V^*, comprobaremos que proporcionan el mismo resultado al aplicarlas a un $v \in V$ arbitrario:

$$\begin{aligned} [f^t(a \cdot \phi' + b \cdot \psi')](v) = &\quad (a \cdot \phi' + b \cdot \psi')(f(v)) = a \cdot \phi'(f(v)) + b \cdot \psi'(f(v)) \\ [a \cdot f^t(\phi') + b \cdot f^t(\psi')](v) = &\quad a \cdot [f^t(\phi')](v) + b \cdot [f^t(\psi')](v) = a \cdot \phi'(f(v)) + b \cdot \psi'(f(v)) \end{aligned}$$

donde en la primera igualdad de la primera línea y la última de la segunda se usa la definición de la traspuesta.

(2) Para todo $\phi \in V^*$, $I_V^t(\phi) = \phi \circ I_V = \phi$.

(3) Como $(h\circ f)^t, f^t\circ h^t : V''^* \to V^*$ debemos comprobar que, para todo $\phi'' \in V''^*$ se tiene $(h\circ f)^t(\phi'') = f^t\circ h^t(\phi'')$. Ambas expresiones pertenecen a V^*, y es fácil comprobar que coinciden:

$$\begin{aligned} (h\circ f)^t(\phi'') = &\quad \phi''\circ(h\circ f) \\ (f^t\circ h^t)(\phi'') = &\quad f^t(h^t(\phi'')) = f^t(\phi''\circ h) = (\phi''\circ h)\circ f, \end{aligned}$$

y la igualdad se sigue de la asociatividad de la composición.

(4) Para probar la linealidad basta con tomar también $b\in K$ y demostrar $(a\cdot f + b\cdot g)^t = a\cdot f^t + b\cdot g^t$. Como ambas se definen de $V'^* \to V^*$, debemos demostrar $(a\cdot f+b\cdot g)^t(\phi') = (a\cdot f^t + b\cdot g^t)(\phi')$ para $\phi'\in V'^*$:

$$\begin{aligned} (a\cdot f+b\cdot g)^t(\phi') = &\quad \phi'\circ(a\cdot f+b\cdot g) \\ (a\cdot f^t+b\cdot g^t)(\phi') = &\quad a\cdot f^t(\phi') + b\cdot g^t(\phi') = a\cdot(\phi'\circ f) + b\cdot(\phi'\circ g) \end{aligned}$$

y la igualdad de ambas expresiones se sigue de la linealidad de ϕ' (¡compruébese aplicándolas a un $v\in V$ arbitrario!).

Para demostrar que la aplicación trasposición es un isomorfismo, basta con comprobar que es inyectiva (al tener su dominio y codominio la misma dimensión), y para esto que su núcleo es 0, esto es, si $f^t = 0$ entonces $f = 0$. Ahora bien, si $f^t = 0$ se tiene $f^t(\phi') = \phi'\circ f = 0$ para todo $\phi'\in V'^*$. Así, para cada $v\in V$ se tiene $\phi'(f(v)) = 0$ para todo $\phi'\in V'^*$, por lo que $f(v) = 0$, como se quería.

(5) Debemos demostrar $\Phi'\circ f = (f^t)^t\circ\Phi$. Como ambas tienen a V como dominio y a V'^{**} como codominio, debemos demostrar que $\Phi'_{f(v)} (= \Phi'\circ f(v))$ coincide con $(f^t)^t(\Phi_v) (= (f^t)^t\circ\Phi(v))$ para todo $v\in V$. Como ambos pertenecen a V'^{**}, debemos comprobar que coinciden al aplicarlo a un $\phi'\in V'^*$ arbitrario:

$$\begin{aligned} \Phi'_{f(v)}(\phi') = &\quad \phi'(f(v)) \\ [(f^t)^t(\Phi_v)](\phi') = &\quad [\Phi_v\circ f^t](\phi') = \Phi_v[f^t(\phi')] = \Phi_v(\phi'\circ f) = (\phi'\circ f)(v), \end{aligned}$$

esto es, la igualdad se cumple aplicando las definiciones de trasposición e isomorfismos Φ, Φ'. ∎

Para trabajar con bases, fijemos la siguiente notación. Tomamos bases $B = (v_1,\ldots,v_n)$, $B' = (v'_1,\ldots,v'_m)$ en V y V', resp., y sus correspondientes bases duales $B^* = (\phi^1,\ldots,\phi^n), B'^* = (\phi'^1,\ldots,\phi'^m)$. Aunque no sea imprescindible, conviene subir uno de los índices de las matrices de las aplicaciones lineales, el de fila o el de columna según la aplicación sea entre V y V' o sus duales. Esto es:

• Al construir $M(f, B'\leftarrow B)$ subiremos (consistentemente con convenios anteriores) el primer índice (filas) para los elementos de esa matriz , esto es, escribiremos $f(v_j) = \sum_i a^i{}_j v'_i$. En consecuencia, para cada columna j, los $a^i{}_j$ son las coordenadas del vector $f(v_j)$ en la base B', con el correspondiente índice i arriba.

• Sin embargo, si $h^* : V^*\to V'^*$ es una aplicación lineal entre los correspondientes espacios duales, para los elementos de la matriz $M(h^*, B'^*\leftarrow B^*)$ subiremos el segundo índice (columnas), esto es, escribiremos $h^*(\phi^j) = \sum_i a_i{}^j \phi'^i$. En consecuencia, para cada columna j, los $a_i{}^j$ son las coordenadas de la forma $h^*(\phi^j)$ en la base B'^*, con el correspondiente índice i abajo.

Proposición 4.28. *La matriz de la aplicación traspuesta f^t verifica:*

$$M(f^t, B^*\leftarrow B'^*) = M(f, B'\leftarrow B)^t.$$

Demostración. (i) Si se denota por ${a^i}_j$ (resp. ${b_i}^j$) el elemento (i, j) de la matriz de f (resp. de f^t) se obtiene:

$$ {b_i}^j = [f^t(\phi'^j)](v_i) = (\phi'^j \circ f)(v_i) = \phi'^j(f(v_i)) = \phi'^j(\sum_{l=1}^n {a^l}_i v'_l) = \sum_{l=1}^n {a^l}_i \phi'^j(v'_l) = {a^j}_i $$

(obsérvese que el índice de columna y el de fila a la izquierda coincide, resp. con el de fila y columna a la derecha). ▮

Observación 4.29. Tomando $f = I_V$, de este resultado se reobtiene la fórmula del cambio de base: $M(I_{V^*}, B^* \leftarrow B'^*) = M(I_V^t, B^* \leftarrow B'^*) = M(I_V, B' \leftarrow B)^t$.

La proposición anterior permite afirmar que el rango de f coincide con el de su traspuesta, *supuesto conocido que el rango de una matriz coincide con el de su traspuesta*. No obstante, el siguiente resultado permite dar una demostración directa para aplicaciones lineales y, en consecuencia, deducir ese resultado para matrices.

Teorema 4.30. *Para cualquier aplicación lineal $f : V \to V'$:*
(i) an$(Im\ f) = Nuc(f^t)$.
(ii) rango$(f) =$ rango(f^t).
En consecuencia, para cualquier matriz $A \in M_{m\times n}(K)$ se tiene rango$(A) =$ rango(A^t).

Demostración. (i) Sea $\phi' \in V^*$, aplicando las definiciones:

$$ \phi' \in \text{an}(Imf) \quad \Leftrightarrow \quad \phi'(f(v)) = 0, \forall v \in V \quad \Leftrightarrow \quad \phi' \circ f = 0 \quad \Leftrightarrow \quad f^t(\phi') = 0 \quad \Leftrightarrow \quad \phi' \in \text{Nuc} f^t $$

(ii) Aplicando el teorema del rango a f^t, el apartado anterior y que $\dim(\text{an}(\text{Im}(f)) = m - \dim(\text{Im}(f))$:

$$ \text{rango}(f^t) := \dim(\text{Im} f^t) = m - \dim(\text{Nuc} f^t) = m - \dim(\text{an}(\text{Im}(f)) = \dim(\text{Im}(f)) = \text{rango}(f). $$

Finalmente, para demostrar la última afirmación, dada A podemos hallar una aplicación lineal f cuya matriz asociada sea A (corolario 3.32) y el rango r de f coincidirá con el de A (proposición 3.37). Según se acaba de demostrar, r coincide con el rango de f^t y, como A^t es la matriz asociada a f^t en bases apropiadas (proposición 4.28), el rango de A^t coincide con r. ▮

Observación 4.31. Este último resultado muestra que se puede desarrollar toda la teoría de aplicaciones lineales independientemente de la de matrices. Una vez hecho esto, las propiedades de las matrices se deducen inmediatamente de las de las aplicaciones lineales.

Corolario 4.32. *an$(Nuc\ f) = Im(f^t)$.*

Demostración. Las siguientes dos demostraciones tienen interés propio.

Dem. 1 Aplicando el apartado (i) del teorema 4.30 anterior a f^t, se sigue an(Im f^t)=Nuc($(f^t)^t$). Tomando anuladores en ambos miembros se tiene an(an((Im f^t))=an(Nuc($(f^t)^t$)), y el resultado se sigue aplicando el teorema de reflexividad en V^*. □

Dem. 2 Para la inclusión ($\supset$), dado $\phi \in \text{Im}(f^t)$, sea $\phi' \in V'^*$ tal que $\phi = f^t(\phi') = \phi' \circ f$. Claramente, esta composición se anula sobre Nuc(f) por lo que $\phi \in \text{an}(\text{Nuc}(f))$.

La inclusión ($\subset$) se obtiene directamente porque las dimensiones de ambos subespacios son iguales: dim(an(Nuc f))= dim(Im(f))= dim(Im(f^t)), la última igualdad por (ii) del teorema 4.30. □

Por completitud, incluimos también a continuación una demostración directa y constructiva de la inclusión ($\subset$).

Demostración alternativa de ($\subset$). Sea $\phi \in \mathrm{an}(\mathrm{Nuc} f)$, esto es, $\mathrm{Nuc}(f) \subset \mathrm{Nuc}(\phi)$. Podemos suponer $\phi \neq 0$ (pues el resultado sería trivial en este caso), con lo que $\mathrm{Nuc}(\phi)$ tiene dimensión $n-1$. Construyamos una base $B = (v_1, v_2, \ldots, v_r, v_{r+1}, \ldots, v_n)$ de V tomando primero una base de $\mathrm{Nuc}(f)$, que constituirá los últimos $n-r$ vectores de B, ampliándola a continuación a una base de $\mathrm{Nuc}(\phi)$, que constituirá los vectores de B a partir del segundo, y finalmente ampliando ésta a una base de V usando un vector v_1 (el cual, obviamente, no pertenecerá a $\mathrm{Nuc}(\phi)$) de modo que $\phi(v_1) = 1$. Necesariamente, $B'_{Imf} := \{f(v_1), \ldots, f(v_r)\} \subset V'$ es una base de $\mathrm{Im}(f)$, y podemos escoger una forma lineal ϕ' que cumpla: $\phi'(f(v_1)) = 1$ y ϕ' se anula sobre $f(v_2), \ldots, f(v_r)$ (por ejemplo, ϕ' se puede construir ampliando B'_{Imf} hasta una base ordenada B' de V y escogiendo el primer elemento de su dual B'^*). Se sigue entonces $\phi = f^t(\phi')$, pues al aplicar $f^t(\phi')$ sobre la base inicial B se obtiene:

$$\begin{cases} f^t(\phi')(v_1) = & \phi'(f(v_1)) = 1 \quad (= \phi(v_1)) & & \text{(por la elección de } \phi'\text{)} \\ f^t(\phi')(v_k) = & \phi'(f(v_k)) = 0 \quad (= \phi(v_k)) & \forall k \in \{2, \ldots, r\} & \text{(por la elección de } \phi'\text{)} \\ f^t(\phi')(v_k) = & \phi'(f(v_k)) = 0 \quad (= \phi(v_k)) & \forall k \in \{r+1, \ldots, n\} & \text{(porque } f(v_k) = 0\text{)} \end{cases}$$

En consecuencia, $\phi \in \mathrm{Im} f^t$, como se quería demostrar. ∎

Corolario 4.33. *(i) f es suprayectiva si y sólo si f^t es inyectiva.*
(ii) f es inyectiva si y sólo si f^t es suprayectiva.

Demostración. (i) f es sobre si y sólo si $\mathrm{Im} f = V'$, lo que equivale a $\mathrm{an}(\mathrm{Im} f) = \{0\}$ y, por el apartado (i) del teorema 4.30, a que $\mathrm{Nuc} f^t = \{0\}$, esto es, a que f^t sea inyectiva.

(ii) Razónese análogamente usando el corolario 4.32 (o bien aplicando el apartado (i) a f^t y usando el teorema de reflexividad). ∎

Observación 4.34. Si $n = \dim V$, $m = \dim V' \in \mathbb{N}$ es fácil comprobar:
(A) f es suprayectiva $\Leftrightarrow$ las columnas de $M(f, B' \leftarrow B)$ forman un sistema de generadores de K^m
$\Leftrightarrow$ el rango (por columnas) de $M(f, B' \leftarrow B)$ es m.
Análogamente:
(B) f es inyectiva $\Leftrightarrow$ las columnas de $M(f, B' \leftarrow B)$ son linealmente independientes en K^n
$\Leftrightarrow$ el rango (por columnas) de $M(f, B' \leftarrow B)$ es n.
Puesto que $M(f^t, B^* \leftarrow B'^*) = M(f, B' \leftarrow B)^t$ se tiene aplicando la anterior a f^t:
(A') f^t es es suprayectiva $\Leftrightarrow$ las filas de $M(f, B' \leftarrow B)$ forman un sistema de generadores de K^n
$\Leftrightarrow$ el rango (por filas) de $M(f, B' \leftarrow B)$ es n.
y análogamente:
(B') f^t es inyectiva $\Leftrightarrow$ las filas de $M(f, B' \leftarrow B)$ son linealmente independientes en K^n
$\Leftrightarrow$ el rango (por filas) de $M(f, B' \leftarrow B)$ es m.
Como por el teorema 4.30 el rango por filas de $M(f, B' \leftarrow B)$ coincide con su rango por columnas, se siguen: (A) $\Leftrightarrow$ (B'), (B) $\Leftrightarrow$ (A') (lo que proporciona una demostración alternativa del corolario 4.33).

Tema 5

Determinantes

El lector tendrá probablemente cierta familiaridad con diversas interpretaciones del determinante de una matriz cuadrada, al menos en los casos 2×2 y 3×3. Si se consideran las columnas de esas matrices como elementos de $\mathbb{R}^2$ y $\mathbb{R}^3$, el valor absoluto del determinante se corresponde, respectivamente, con el área del paralelogramo cuyos lados están generados por los dos vectores de $\mathbb{R}^2$ y el volumen del paralelepípedo de lados generados por los tres vectores de $\mathbb{R}^3$. En particular, esta área o volumen es cero si y sólo si los correspondientes vectores son linealmente dependientes (esto es, el paralelogramo degenera en un segmento o en un punto, y el paralelepípedo en un paralelogramo o en un segmento o en un punto). Todo ello es generalizable a dimensión superior. Nosotros desarrollaremos el concepto de determinante sin entrar en la discusión de los conceptos de *área* o *volumen* (los cuales es natural posponer hasta el estudio de *longitudes* en *espacios vectoriales métricos*); no obstante, daremos algunas interpretaciones intuitivas sobre determinantes (secciones 5.2.1, 5.4.1, 5.4.2). Asimismo, desarrollaremos aquí el concepto de determinante en relación con nuestro estudio de endomorfismos.

Recordemos que el concepto de determinante de una matriz cuadrada y bastantes de sus propiedades elementales se conocen desde la sección 1.5. Ahora revisaremos este concepto y terminaremos de demostrar sus principales propiedades. Nuestro estudio se desarrolla en cuatro pasos: (1) el concepto de tensor determinante, (2) el tensor determinante asociado a una base ordenada (lo que muestra, en particular, la existencia de tensores determinantes no idénticamente nulos), (3) el determinante de una matriz cuadrada, como caso particular del anterior, y (4) el determinante de un endomorfismo, con el que se gana en abstracción y, como recompensa, se obtiene simplicidad en algunas demostraciones.

Salvo mención explícita de lo contrario, $V(K)$ será en adelante un e.v. de dimensión $n\in\mathbb{N}$ sobre un cuerpo K de característica distinta de 2. Aunque no presupondremos el estudio de determinantes del tema 1, avisaremos con diversas notas la aparición de razonamientos análogos a los desarrollados en ese tema.

5.1. Tensores determinantes

Definición 5.1. *Sea $V(K)$ un e.v. de dimensión $n\in\mathbb{N}$. Diremos que una aplicación:*

$$D: V\times\cdots\times^{(n)} V\to K \qquad (5.1)$$

es un tensor determinante *si verifica:*

(i) Es multilineal*, esto es, lineal en cada una de sus n variables en el sentido:*

$$\begin{aligned} D(w_1,\ldots,a\cdot w_i,\ldots,w_n) = & \quad a\cdot D(w_1,\ldots,w_i,\ldots,w_n) \\ D(w_1,\ldots,w_i+w_i',\ldots,w_n) = & \quad D(w_1,\ldots,w_i,\ldots,w_n)+D(w_1,\ldots,w_i',\ldots,w_n) \end{aligned}$$

para todo[1] $w_1,...,w_n,w_i' \in V$, $a \in K$, $i \in \{1,\ldots,n\}$.

(ii) Es antisimétrico *(o* hemisimétrico*), esto es, para cada par de variables* $i,j \in \{1,\ldots,n\}$, $i<j$, *se verifica:*

$$D(w_1,\ldots,w_i,\ldots,w_j,\ldots,w_n) = -D(w_1,\ldots,w_j,\ldots,w_i,\ldots,w_n) \tag{5.2}$$

Nota 5.2. El determinante de una matriz det A visto en el tema 1 (definición 1.50) puede verse como una aplicación que, a la n-úpla de columnas de A (vistas como elementos de $K^n(K)$) le hace corresponder un escalar, lo que puede verse como un caso particular de (5.1). Las propiedades (i), (ii) que definen el tensor determinante pueden verse entonces como una abstracción de las correspondientes propiedades que se dedujeron para det A (las 3 y 4 de la sección 1.5.2).

Observación 5.3. Puesto que suponemos que la característica de K es distinta de 2, la propiedad de ser antisimétrico equivale (supuesta la multilinealidad de D) a la de ser *alternado*, esto es,

$$D(w_1,\ldots,w_i,\ldots,w_j,\ldots,w_n)=0, \qquad \text{si} \qquad w_i=w_j \tag{5.3}$$

para cualquier par de índices $i<j$ (se señalan sólo las posiciones i-ésima y j-ésima entre puntos suspensivos). Formalmente[2]:

($\Rightarrow$) Si D es antisimétrico se tiene aplicando (5.2) para $w=w_i=w_j$:

$$D(w_1,\ldots,w,\ldots,w,\ldots,w_n) = -D(w_1,\ldots,w,\ldots,w,\ldots,w_n)$$

por lo que $2D(w_1,\ldots,w,\ldots,w,\ldots,w_n)=0$, y como característica$(K)\neq 2$ se sigue (5.3).

($\Leftarrow$) Si D es alternado, se sigue de (5.3) poniendo $w=w_i+w_j$ en las posiciones i-ésima y j-ésima:

$$\begin{aligned} 0 = & \quad D(w_1,\ldots,w=w_i+w_j,\ldots,ww_i+w_j,\ldots,w_n) \\ = & \quad D(w_1,\ldots,w_i,\ldots,w_i,\ldots,w_n)+D(w_1,\ldots,w_j,\ldots,w_j,\ldots,w_n) \\ & \quad +D(w_1,\ldots,w_i,\ldots,w_j,\ldots,w_n)+D(w_1,\ldots,w_j,\ldots,w_i,\ldots,w_n), \end{aligned}$$

donde en la última igualdad se ha usado la linealidad respecto a las variables i,j. Usando de nuevo (5.3) los dos primeros sumandos del último miembro son nulos y se obtiene (5.2).

Es fácil comprobar que el conjunto $D(V)$ de todos los tensores determinantes sobre V tiene una estructura de espacio de espacio vectorial.

Ejercicio 5.4. *Compruébese, en el conjunto $D(V)$ de todos los tensores determinantes sobre V, la suma $D+D'$ y el producto por escalares $a\cdot D$ para $D,D' \in D(V), a \in K$ definidos por:*

$$\begin{aligned} (D+D')(w_1,\ldots,w_n) & \quad := \quad D(w_1,\ldots,w_n)+D'(w_1,\ldots,w_n) \\ (a\cdot D)(w_1,\ldots,w_n) & \quad := \quad a\cdot D(w_1,\ldots,w_n) \end{aligned}$$

para todo $w_1,\ldots,w_n \in V$, son también tensores determinantes y, con estas operaciones, $(D(V),+,\cdot K)$ resulta ser un espacio vectorial sobre[3] *K.*

[1] En general, un tensor m-covariante es una aplicación $V\times\cdots\times^{(m)} V \to K$ que es multilineal en este sentido.

[2] Este tipo de propiedad se discutió en 5 de la de la sección 1.5.2.

[3] De hecho, pueden verse como un subespacio vectorial del espacio vectorial de todas las aplicaciones $V\times\cdots\times^{(n)} V \to K$.

Proposición 5.5. *Todo tensor determinante D verifica:*

(1) Para cualquier permutación $\sigma \in S_n$:

$$D(w_{\sigma(1)},\dots,w_{\sigma(n)}) = \mathrm{sig}(\sigma)\cdot D(w_1,\dots,w_n)$$

(2) $D(w_1,\dots,w_i,\dots,w_n)=0$ si alguno de los vectores w_i es 0, o si se repiten dos de los argumentos de D (esto es, $w_i = w_j$ para algún par de índices $i \neq j$).

(3) Si uno de los vectores w_i es una combinación lineal del resto de vectores, el valor del determinante se anula, esto es:

$$D(w_1,\dots,\sum_{k\neq i} a^k w_k,\dots,w_n) = 0, \qquad \forall a^1,\dots,a^{k-1},a^{k+1},\dots a^n \in K.$$

(4) Si a uno de los vectores w_i se le añade una combinación lineal del resto de vectores, el valor del determinante no cambia, esto es:

$$D(w_1,\dots,w_i+\sum_{k\neq i} a^k w_k,\dots,w_n) = D(w_1,\dots,w_i,\dots,w_n).$$

(5) D queda determinado por su valor sobre una base $B=(v_1,\dots,v_n)$. Esto es, si D y D' son dos tensores determinantes sobre V tales que $D(v_1,\dots,v_n)=D'(v_1,\dots,v_n)$ entonces $D=D'$.

Demostración. [4] (1) Si σ es una permutación $\mathrm{sig}(\sigma)=-1$ y la propiedad se deduce de la antisimetría de D. En caso contrario se escribe σ como composición de k trasposiciones y se aplica k veces el resultado para trasposiciones.

(2) Para la primera afirmación, obsérvese $w_i = 0\cdot w_i$ y aplíquese la linealidad respecto a la variable i-ésima[5]. La segunda se ha demostrado en la observación 5.3).

(3) En efecto, por la linealidad en la i-ésima variable:

$$D(w_1,\dots,\sum_{k\neq i} a^k w_k,\dots,w_n) = \sum_{k\neq i} a^k D(w_1,\dots,w_k,\dots,w_n)$$

y cada sumando es nulo, porque el vector w_k aparece en las variables k-ésima e i-ésima de D.

(4) Aplicando la linealidad respecto a la variable i-ésima:

$$\begin{aligned} D(w_1,\dots,w_i+\textstyle\sum_{k\neq i} a^k w_k,\dots,w_n) &= D(w_1,\dots,w_i,\dots,w_n) + D(w_1,\dots,\textstyle\sum_{k\neq i} a^k w_k,\dots,w_n)\\ &= D(w_1,\dots,w_i,\dots,w_n). \end{aligned}$$

(la última igualdad por el apartado anterior.)

(5) Sean n vectores $w_j = \sum_{i=1}^n a^i{}_j v_i$ para $j=1,\dots,n$. Cambiando el subíndice mudo i por i_j en la expresión de cada w_j para evitar confusiones:

$$\begin{aligned} D(w_1,\dots,w_n) = &\quad D(\textstyle\sum_{i_1=1}^n a^{i_1}{}_1 v_{i_1},\dots,\sum_{i_n=1}^n a^{i_n}{}_n v_{i_n})\\ = &\quad \textstyle\sum_{i_1=1}^n \cdots \sum_{i_n=1}^n \left(a^{i_1}{}_1 \dots a^{i_n}{}_n\, D(v_{i_1},\dots,v_{i_n})\right). \end{aligned}$$

En esta expresión, será nulo cada sumando (escrito entre paréntesis grandes) para el cual alguno de los índices $i_1,\dots,i_n$ se repita. Por tanto, podemos restringirnos a los casos en que $i_1,\dots,i_n$ son una permutación, esto es, $i_1=\sigma(1),\dots,i_n=\sigma(n)$ para algún

$$\sigma = \begin{pmatrix} 1 & \dots & n \\ 1 & \dots & \sigma(n) \end{pmatrix} \in S_n.$$

[4] Las demostraciones de (1) a (4) son análogas a las vistas en la sección 1.5.2.

[5] Como una argumento alternativo, la propiedad se verifica porque, fijados los valores de w_j con $j \neq 1$, se tiene una aplicación lineal en la variable i-ésima, la cual aplicará el 0 en el 0.

Así, se pueden reemplazar los sumatorios[6] $\sum_{i_1=1}^n \cdots \sum_{i_n=1}^n \equiv \sum_{i_1,\dots i_n=1}^n$ por $\sum_{\sigma\in S_n}$ obteniéndose:

$$\begin{aligned} D(w_1,\dots,w_n) = & \sum_{\sigma\in S_n}\left(a^{\sigma(1)}{}_1 \dots a^{\sigma(n)}{}_n D(v_{\sigma(1)},\dots,v_{\sigma(n)})\right) \\ = & \sum_{\sigma\in S_n}\left(a^{\sigma(1)}{}_1 \dots a^{\sigma(n)}{}_n \operatorname{sig}(\sigma) D(v_1,\dots,v_n)\right), \\ = & \left(\sum_{\sigma\in S_n}(a^{\sigma(1)}{}_1 \dots a^{\sigma(n)}{}_n \operatorname{sig}(\sigma)\right) D(v_1,\dots,v_n) \end{aligned} \tag{5.4}$$

la segunda igualdad por el punto (1) anterior. Obsérvese que, para llegar al último miembro, sólo se han aplicado las propiedades de los tensores determinantes, de modo que para otro tal tensor D' se habría obtenido la misma expresión reemplazando el término $D(v_1,\dots,v_n)$ por $D'(v_1,\dots,v_n)$. Por tanto, si ambos términos coinciden, entonces D y D' coinciden sobre cualesquiera n-vectores. ∎

5.2. Tensor determinante en una base

La demostración del último punto de la proposición 5.5 sugiere, por una parte, la construcción de un tensor determinante (usando la fórmula (5.4)) y, por otra, que a partir de él se pueden construir todos los tensores determinantes.

Definición 5.6. *Dada una base ordenada $B = (v_1,\dots,v_n)$ de V, llamaremos* determinante en B *a la aplicación*

$$\begin{aligned} \det_B : \quad & V\times\cdots\times^{(n)} V && \to K \\ & (w_1,\dots,w_n) && \mapsto \sum_{\sigma\in S_n}\operatorname{sig}(\sigma)\, a^{\sigma(1)}{}_1 \dots a^{\sigma(n)}{}_n \end{aligned} \tag{5.5}$$

donde $w_j = \sum_{i=1}^n a^i{}_j v_i$ para $j = 1,\dots,n$.

Proposición 5.7. *El determinante en $B = (v_1,\dots,v_n)$ verifica:*

(1) $\det_B$ es un tensor determinante.

(2) $\det_B(v_1,\dots,v_n) = 1$, y es el único tensor determinante que verifica esta igualdad.

(3) Si D es cualquier tensor determinante, se tiene: $D = D(v_1,\dots,v_n)\det_B$.

(4) Si $\bar{B} = (\bar{v}_1,\dots,\bar{v}_n)$ es otra base ordenada, entonces:

$$\det_{\bar{B}} = \det_{\bar{B}}(v_1,\dots,v_n)\cdot\det_B \quad \textit{o, más mnemotécnicamente,} \quad \det_{\bar{B}} = \det_{\bar{B}}(B)\cdot\det_B \tag{5.6}$$

Demostración. (1) Multilinealidad: si se toman $a,b\in K$, $u_j = \sum_{i=1}^n b^i{}_j v_i \in V$ y en la expresión (5.5) que define $\det_B$ se reemplaza en la variable j-ésima el vector $w_j (= \sum_{i=1}^n a^i{}_j v_i)$ por $aw_j + bu_j = \sum_{i=1}^n (a a^i{}_j + b b^i{}_j) v_i$ se tiene:

$$\begin{aligned} \det_B(w_1,\dots,aw_j+bu_j,\dots,w_n) &= \sum_{\sigma\in S_n}\operatorname{sig}(\sigma)\, a^{\sigma(1)}{}_1 \dots (a a^{\sigma(j)}{}_j + b b^{\sigma(j)}{}_j)\dots a^{\sigma(n)}{}_n \\ &= a\sum_{\sigma\in S_n}\operatorname{sig}(\sigma)\, a^{\sigma(1)}{}_1 \dots a^{\sigma(j)}{}_j \dots a^{\sigma(n)}{}_n + b\sum_{\sigma\in S_n}\operatorname{sig}(\sigma)\, a^{\sigma(1)}{}_1 \dots b^{\sigma(j)}{}_j \dots a^{\sigma(n)}{}_n \\ &= a\det_B(w_1,\dots,w_j,\dots,w_n) + b\det_B(w_1,\dots,u_j,\dots,w_n) \end{aligned}$$

Antisimetría respecto a las variables $i<j$: si τ es la trasposición (i,j)

$$\begin{aligned} \det_B(w_{\tau(1)},\dots,w_{\tau(i)},\dots w_{\tau(j)},\dots,w_{\tau(n)}) &= \sum_{\sigma\in S_n}\operatorname{sig}(\sigma)\, a^{\sigma(1)}{}_{\tau(1)} \dots a^{\sigma(i)}{}_{\tau(i)} \dots a^{\sigma(j)}{}_{\tau(j)} \dots a^{\sigma(n)}{}_{\tau(n)} \\ &= \sum_{\sigma\in S_n}(-\operatorname{sig}(\sigma\circ\tau))\, a^{\sigma(\tau(1))}{}_1 \dots a^{\sigma(\tau(i))}{}_i \dots a^{\sigma(\tau(j))}{}_j \dots a^{\sigma(\tau(n))}{}_n \\ &= -\sum_{\rho\in S_n}\operatorname{sig}(\rho)\, a^{\rho(1)}{}_1 \dots a^{\rho(i)}{}_i \dots a^{\rho(j)}{}_j \dots a^{\rho(n)}{}_n \end{aligned}$$

[6] Se pasa de n^n sumandos, correspondientes a las *variaciones con repetición* de $\{1,\dots,n\}$ tomados de n en n, a tener $n!$ sumandos, correspondientes a las permutaciones (sin repetición).

donde en la segunda fila se ha usado que $\mathrm{sig}(\sigma) = -\mathrm{sig}(\sigma\circ\tau)$ (puesto que τ es una trasposición) y se han reordenado los factores en orden creciente de subíndices; en la tercera se ha tomado $\rho = \sigma\circ\tau$ y se ha usado que, al ser biyectiva en S_n la aplicación $\sigma \mapsto \sigma\circ\tau$, resulta equivalente sumar en $\sigma \in S_n$ que en $\rho = \sigma\circ\tau \in S$.

(2) En este caso $a^i{}_j = \delta^i_j$ (delta de Kronecker) por lo que todos los sumandos en (5.5) son nulos salvo el correspondiente a la permutación identidad, que es igual a 1. La unicidad es inmediata de la proposición 5.5(5).

(3) Puesto que el ambos miembros de la igualdad son tensores determinantes, por la proposición 5.5(5) basta con comprobar que ambos coinciden al aplicarse sobre $(v_1,\dots,v_n)$. Pero esto es inmediato usando el punto (2) para el cálculo del miembro derecho.

(4) Aplíquese el punto anterior a $D = \det_{\bar{B}}$ ∎

Como consecuencia, se tiene una caracterización sencilla (y usada en la práctica) de cuándo n vectores son linealmente independientes.

Corolario 5.8. *Sea B una base y $S = \{w_1,\dots,w_n\} \subset V$. S es linealmente independiente (y, por tanto, una base) si y sólo si* $\det_B(w_1,\dots,w_n) \neq 0$.

Demostración. ($\Leftarrow$) Si fuera linealmente dependiente, sabemos que $D(w_1,\dots,w_n) = 0$ para todo tensor determinante D (en particular, para $D = \det_B$).

($\Rightarrow$) Puesto que $\bar{B} = (w_1,\dots,w_n)$ es una base, se tiene $\det_B = \det_B(w_1,\dots,w_n)\det_{\bar{B}}$ por lo que si $\det_B(w_1,\dots,w_n)$ se anulara entonces $\det_B \equiv 0$. ∎

Finalmente, damos la siguiente consecuencia para el espacio vectorial $D(V)$ de todos los tensores determinantes (considerado en el ejercicio 5.4), la cual no era en absoluto evidente de su definición.

Corolario 5.9. *Fijada una base $B = (v_1,\dots,v_n)$, la aplicación*

$$D(V) \to K, \qquad D \mapsto D(v_1,\dots,v_n)$$

es un isomorfismo de e.v. En consecuencia, $\dim_K(D(V)) = 1$ y $D(V) = L\{\det_B\}$.

Demostración. La aplicación es lineal (¡compruébese!) trivialmente e inyectiva por la proposición 5.5(5). También es suprayectiva porque todo $a \in K$ es la imagen del tensor determinante $a\det_B$. ∎

Observación 5.10. Todo tensor determinante D distinto del nulo se puede expresar como $D = \det_{\bar{B}}$ para alguna base $\bar{B}$. Para comprobarlo si $B = (v_1,\dots,v_n)$ es cualquier base entonces $D(v_1,\dots,v_n) = a \neq 0$. Tomando la nueva base $\bar{B} = (\bar{v}_1 = a^{-1}\cdot v_1, \bar{v}_2 = v_2 \dots, \bar{v}_n = v_n)$ se tiene

$$D = D(\bar{v}_1,\bar{v}_2\dots,\bar{v}_n)\det_B = D(a^{-1}\cdot v_1, v_2\dots,v_n)\det_B = a^{-1}\cdot a\cdot\det_B = \det_B$$

(aplíquese la proposición 5.7(3))[7].

[7] Alternativamente, puede razonarse que si $\bar{B}$ se obtiene cambiando uno de los vectores v_j de B por $\bar{v}_j := \lambda v_j$ con $\lambda \neq 0$, entonces todas las coordenadas de un vector $v \in V$ coinciden en B y $\bar{B}$ excepto las j-ésimas a^j, $\bar{a}^j$ (resp.), que verifican $\bar{a}^j = \lambda^{-1}\cdot a^j$ (de modo que $a_j v_j = \bar{a}_j\bar{v}_j$). Por tanto, $\det_{\bar{B}} = \lambda^{-1}\det_B$.

5.2.1. Interpretación para $K = \mathbb{R}$: elemento de volumen

En geometría elemental, una base de $\mathbb{R}^2$ genera un paralelogramo y una de $\mathbb{R}^3$ un paralelepípedo. En general, en un e.v. real $V(\mathbb{R})$ podemos decir que una base $B = (u_1, \ldots, u_n)$ genera un *n-paralelelípedo* (o simplemente paralelepípedo) P_B definido por

$$P_B := \{a_1u_1 + \cdots + a_nu_n : a_1, \ldots, a_n \in [0,1]\}.$$

Fijado un tensor determinante no nulo D_0, llamaremos *volumen de P_B con respecto a D_0* a

$$\mathrm{vol}(P_B) := |D_0(u_1,...,u_n)| \; (= |D_0(B)|)$$

(nótese que se ha tomado el valor absoluto en $K = \mathbb{R}$). La expresión de P_B y de su volumen tiene sentido incluso si permitimos que los vectores de B sean linealmente dependientes. En este caso P_B estaría incluido en un subespacio de dimensión $< n$ y su volumen sería 0 (se podría considerar como un paralelepípedo degenerado). Para justificar intuitivamente el nombre de volumen, obsérvese:

1. Sabemos que $D_0 = \det_{B_0}$ para alguna base $B_0 = (v_1, \ldots, v_n)$ que satisfaga $D_0(B_0) = 1$ (la elección de B_0 no es única). En particular, P_{B_0} tiene volumen 1 con respecto a D_0.

   Así, D_0 se puede representar como un paralelepípedo P_{B_0} de $V(\mathbb{R})$ el cual fija el "elemento de volumen unitario".

2. Si construimos a partir de B una nueva base $\bar{B} = (m_1v_1, \ldots, m_nv_n)$ con $m_1, \ldots, m_n \in \mathbb{R} \setminus \{0\}$ entonces (por la multilinealidad de D_0)

   $$\mathrm{vol}(P_{\bar{B}}) = |D_0(\bar{B})| = |m_1 \cdots m_n \cdot D_0(B_0)| = |m_1 \cdots m_n| \cdot \mathrm{vol}(P_{B_0}).$$

   En el caso de que $m_1, \ldots, m_n$ sean enteros, esto es consistente con la intuición de que el paralelepípedo $P_{\bar{B}}$ ocuparía un volumen que sería cubierto por exactamente $|m_1 \cdot m_2 \cdots m_n|$ copias de P_{B_0}. De manera natural esta interpretación se extiende al caso de que los $m_1, \ldots, m_n$ sean racionales o cualesquiera reales.

3. Si construimos a partir de B una nueva base $\tilde{B}$ reemplazando el último elemento v_n de B_0 por $\tilde{v}_n := v_n + \sum_{i=1}^{n-1} a_iv_i$, donde $a_1, \ldots, a_{n-1} \in \mathbb{R}$, entonces (usando además la antisimetría de D_0)

   $$\mathrm{vol}(P_{\tilde{B}}) = |D_0(\tilde{B})| = |D_0(B)| = \mathrm{vol}(P_{B_0}).$$

   Esto es consistente con la intuición de que $P_{\tilde{B}}$ y P_{B_0} son paralelepípedos con una cara común (la formada por los $n-1$ primeros vectores de $\tilde{B}$ y B_0) y, al obtenerse $\tilde{v}_n$ desplazando v_n paralelamente a esa cara, P_B y P_{B_0} deben ocupar el mismo volumen[8].

Las consideraciones anteriores permiten visualizar el valor absoluto de D_0 del siguiente modo: se fija un paralelepído P_{B_0} (tal que $D_0 = P_{B_0}$) el cual servirá como *elemento de volumen* para medir el volumen de cualquier otro paralelepípedo (y, en general, de cualquier otro subconjunto de $V(\mathbb{R})$).

[8] Este tipo de intuición justifica las fórmulas "área= base $\times$ altura" para un paralelogramo en $\mathbb{R}^2$ y "volumen = (área de la base) $\times$ altura" en $\mathbb{R}^3$ de la geometría elemental (no obstante, nosotros ni siquiera hemos necesitado aquí ningún concepto de "perpendicularidad", el cual resulta necesario para el de "altura" en geometría elemental).

5.3. Determinante de una matriz cuadrada

Podemos ahora redefinir el determinante de una matriz cuadrada $A = (a^i{}_j)$ que anticipamos en el primer capítulo y deducir muchas de sus propiedades. Para ello, sea $B_u = (e_1, \ldots, e_n)$ la base usual de K^n y, fijada A, consideremos los vectores $w_j \in K^n$ tales que $(w_j)_{B_u}$ es la j-ésima columna de A (esto es, $w_j = (w_j)^t_{B_u}$). Comparando la fórmula del det A introducida en (1.9) y la del determinante en una base (5.5) aplicada a $B = B_u$ es inmediato:

$$\det A = \det_{B_u}(w_1, \ldots, w_n),$$

esto es, el determinante de A es igual a $\det_{B_u}$ aplicado a sus columnas, vistas como elementos de K^n. Con más generalidad:

Corolario 5.11. *Sea B una base de un espacio vectorial V, sean $w_1, \ldots, w_n \in V$ y sea A la matriz que se obtiene escribiendo por columnas $(w_1)_B, \ldots, (w_n)_B$. Entonces*

$$\det A = \det_B(w_1, \ldots, w_n).$$

En particular, si B y $\bar{B}$ son dos bases ordenadas de V:

$$\det M(I_V, \bar{B} \leftarrow B) = \det_{\bar{B}}(B) \qquad \text{y, por tanto,} \qquad \det_{\bar{B}} = \det M(I_V, \bar{B} \leftarrow B) \cdot \det_B$$

Demostración. Inmediata del razonamiento previo (úsese (5.6) para la última igualdad). ∎

5.4. Determinante de un endomorfismo

Veamos a continuación cómo todo endomorfismo f de V permite definir, para cada tensor determinante D, otro tensor determinante que denotaremos $f^*(D)$, y sus consecuencias.

Lema 5.12. *Sea $f \in \mathrm{End}(V)$.*

(1) Para cada $D \in D(V)$ la aplicación:

$$\begin{array}{rccl} f^*(D): & V \times \ldots^{(n)} \times V & \to & K \\ & (w_1, \ldots, w_n) & \mapsto & D((f(w_1), \ldots, f(w_n)) \end{array}$$

es también un tensor determinante.

(2) Más aún, la aplicación:

$$\begin{array}{rccl} f^*: & D(V) & \to & D(V) \\ & D & \mapsto & f^*(D) \end{array}$$

es una aplicación lineal, esto es, $f^* \in End(D(V))$.

(3) En consecuencia, f^* es un múltiplo de la aplicación identidad de $D(V)$.

Demostración. Los puntos (1) y (2) son una computación directa de la que damos los pasos, que puede justificar detalladamente el lector.

(1) Multilinealidad:

$$\begin{aligned} f^*(D) \quad & (w_1, \ldots, aw_i + bw'_i, \ldots, w_n) = \\ = \quad & D((f(w_1), \ldots, f(aw_i + bw'_i), \ldots, f(w_n)) = D((f(w_1), \ldots, af(w_i) + bf(w'_i), \ldots, f(w_n)) \\ = \quad & aD((f(w_1), \ldots, f(w_i), \ldots, f(w_n)) + bD((f(w_1), \ldots, f(w'_i), \ldots, f(w_n)) \\ = \quad & af^*D(w_1, \ldots, w_i, \ldots, w_n) + bf^*D(w_1, \ldots, w_i, \ldots, w_n) \end{aligned}$$

Antisimetría:

$$f^*D(w_1,\dots,w_i,\dots,w_j,\dots,w_n) = D(f(w_1),\dots,f(w_i),\dots,f(w_j),\dots,f(w_n))$$
$$= -D(f(w_1),\dots,f(w_j),\dots,f(w_i),\dots,f(w_n)) = -f^*D(w_1,\dots,w_j,\dots,w_i,\dots,w_n)$$

(2) Para demostrar $f^*(aD+bD') = af^*(D)+bf^*(D')$ se aplican a la misma n-úpla:

$$\begin{aligned}
f^*(aD+bD')(w_1,\dots,w_n) = &\ (aD+bD')((f(w_1),\dots,f(w_n)) \\
= &\ aD((f(w_1),\dots,f(w_n))+bD'((f(w_1),\dots,f(w_n)) \\
(af^*(D)+bf^*(D'))(w_1,\dots,w_n) = &\ af^*(D)(w_1,\dots,w_n)+bf^*(D')(w_1,\dots,w_n) \\
= &\ aD((f(w_1),\dots,f(w_n))+bD'((f(w_1),\dots,f(w_n)),
\end{aligned}$$

obteniéndose la misma expresión en ambos casos, como se quería.

(3) Inmediato por ser $f^*: D(V) \to D(V)$ lineal y (por el corolario 5.9) $\dim_K(D(V)) = 1$. ∎

El último punto permite dar una sorprendente definición del determinante de un endomofismo.

Definición 5.13. *Para cada $f \in \mathrm{End}(V)$, llamaremos* determinante de f *al único escalar* $\det f \in K$ *tal que (según el punto (3) del lema anterior) verifica*

$$f^* = \det f \cdot I_{D(V)}$$

Una primera propiedad de este determinante es, por supuesto, su consistencia con la del determinante de una matriz.

Proposición 5.14. *Sea $f \in \mathrm{End}(V)$ y B cualquier base de V. Entonces:*

$$\det f = \det M(f,B),$$

Demostración. De $f^* = \det f \cdot I_{D(V)}$ se sigue $f^*\det_B = \det f \cdot \det_B$. Cambiando el orden de los miembros de esta igualdad y aplicando ambos a $B = (v_1,\dots,v_n)$:

$$\det f = f^*\det_B(v_1,\dots,v_n) = \det_B(f(v_1),\dots,f(v_n)) = \det M(f,B),$$

donde en la última igualdad se aplica el corolario 5.11 (tomando $w_i = f(v_i)$). ∎

Observación 5.15. (1) En el resultado anterior la matriz de f debe calcularse *en un única base* (esto es, se está tomando $M(f,B) = M(f,B \leftarrow B)$). Esto es posible porque f es un endomorfismo; remarquemos que *no* hemos definido el determinante para una aplicación lineal entre dos e.v. distintos, aunque tengan la misma dimensión.

(2) Una consecuencia inmediata de la anterior proposición es que si dos matrices A y C son semejantes entonces sus determinantes son iguales. De hecho, por ser semejantes (según la definición[9] 3.52) existe un endomorfismo f de algún espacio vectorial V tal que $A = M(f,B)$ y $C = M(f,\bar{B})$ para dos bases $B,\bar{B}$ de V; por tanto, $\det A = \det f = \det C$.

(3) Podríamos preguntarnos si no habría sido más fácil definir directamente $\det f$ como el determinante de la matriz $M(f,B)$ y demostrar entonces que el valor de este determinante es independiente de la base B escogida. No obstante, la demostración de este hecho pasaría por saber algunas propiedades del determinante de matrices, entre ellas que el determinante del producto de dos matrices

[9] Recuérdense también las caracterizaciones en la proposición 3.53.

cuadradas es el producto de los determinantes de las dos matrices[10]. La demostración directa de este resultado, empero, no es sencilla[11], mientras que a partir de nuestra definición de $\det f$ va a resultar prácticamente inmediata.

Lema 5.16. *(1)* $I_V^* = I_{D(V)}$.
(2) Si $f,h \in \mathrm{End}(V)$ *entonces* $(f\circ h)^* = h^* \circ f^*$.

Demostración. (1) Dado $D \in D(V)$, para comprobar $I_V^*(D) = D$, sean $w_1,\dots,w_n \in V$:

$$I_V^*(D)(w_1,\dots,w_n) = D(I_V(w_1),\dots,I_V(w_n)) = D(w_1,\dots,w_n)$$

(2) De nuevo, para comprobar la igualdad, se aplican ambos términos al mismo $D \in D(V)$ y, entonces, a la misma n-úpla:

$$\begin{aligned}[(f\circ h)^*(D)](w_1,\dots,w_n) &= D((f\circ h)(w_1),\dots,(f\circ h)(w_n))\\ [(h^*\circ f^*)(D)](w_1,\dots,w_n) &= [h^*(f^*(D))](w_1,\dots,w_n) = [f^*(D)](h(w_1),\dots,h(w_n))\\ &= D(f(h(w_1)),\dots,f(h(w_n)))\end{aligned}$$

que claramente coinciden. Así, $(f\circ h)^*(D) = (h^*\circ f^*)(D)$ para todo D, como se quería. ∎

Teorema 5.17. *(1) El endomorfismo y la matriz identidad verifican:* $\det I_V = 1$, $\det I_n = 1$.
(2) La composición de endomorfismos y el producto de matrices verifican:

$$\det(f\circ h) = \det f\cdot\det h, \quad \forall f,h\in \mathrm{End}(V); \qquad \det(A\cdot C) = \det A\cdot\det C, \quad \forall A,C\in M_n(K).$$

(3) $f \in \mathrm{End}(V)$ *es biyectivo si y sólo si* $\det f \neq 0$. *En este caso,* $\det(f^{-1}) = (\det f)^{-1}$. *Consistentemente,* $A \in M_n(K)$ *es regular si y sólo si* $\det A \neq 0$. *En este caso,* $\det(A^{-1}) = (\det A)^{-1}$.
(4) $\det f = \det(f^t)$.

Demostración. (1) La primera igualdad es inmediata por el primer punto del lema previo y la definición de $\det I_V$. Para la segunda, usando la proposición 5.14: $\det I_n = \det M(I_V,B) = \det I_V = 1$ (o alternativamente, por computación directa en la definición 1.50).

(2) La igualdad para f,h se sigue de que, por la definición de determinante:

$$(f\circ h)^* = \det(f\circ h)\cdot I_{D(V)}, \qquad h^*\circ f^* = (\det h\cdot I_{D(V)})\circ(\det f\cdot I_{D(V)}) = (\det h\cdot\det f)\cdot I_{D(V)}$$

y de que las dos expresiones coinciden por el segundo punto del lema anterior.

Para la igualdad en el caso de matrices, fijemos cualquier e.v. $V(K)$ de dimensión n y cualquier base suya (p. ej. $V(K) = K^n(K), B = B_u$), y sean f,h tales que $A = M(f,B), C = M(h,B)$. Entonces:

$$\det(A\cdot C) = \det(M(f,B)\cdot M(h,B)) = \det M(f\circ h,B) = \det(f\circ h) = (\det f)(\det h) = (\det A)(\det C).$$

(3) Fijada cualquier base $B = (v_1,\dots,v_n)$ se tiene (véase la proposición 5.14):

$$\det f = \det_B(f(v_1),\dots,f(v_n))$$

[10]De hecho, sabemos que $M(f,\bar{B}) = P^{-1}M(f,B)P$ para la matriz regular $P = M(I_V, B \leftarrow \bar{B})$, por lo que la demostración se obtendría si se demuestra que $\det(P^{-1}M(f,B)P) = \det(P^{-1})\det(M(f,B))\det P$ y $\det(P^{-1}) = (\det P)^{-1}$.

[11]Véase el libro de Merino y Santos para una demostración a partir de propiedades de matrices elementales.

el cual es distinto de 0 si y sólo si $\{f(v_1),\ldots,f(v_n)\}$ es linealmente independiente (corolario 5.8) y, por tanto, si y sólo si f es biyectiva. Además, en este caso

$$1=\det I_V=\det(f\circ f^{-1})=(\det f)(\det f^{-1}),$$

esto es, $\det f^{-1}=(\det f)^{-1}$.

Para el caso de la matriz A, tomamos f como en el punto anterior de modo que $A=M(f,B)$ y:

$$A \text{ regular} \iff f \text{ biyectiva} \iff \det f\neq 0 \iff \det A(=\det M(f,B))\neq 0.$$

En este caso, $\det(A^{-1})=\det(M(f,B)^{-1})=\det M(f^{-1},B)=\det f^{-1}=(\det f)^{-1}=(\det A)^{-1}$.

(4) Como el determinante de una matriz coincide con el de su traspuesta (proposición 1.51):

$$\det f=\det M(f,B)=\det(M(f,B)^t)=\det M(f^t,B)=\det(f^t).$$

∎

5.4.1. Interpretación para $K=\mathbb{R}$: orientación

Podemos completar la interpretación del tensor determinante vista en la sección 5.2.1 dando una interpretación de su signo. En primer lugar, la propiedad $\det(A\cdot C)=\det A\cdot\det C$, aplicada a matrices de cambio de base, permite resolver fácilmente lo siguiente.

Ejercicio 5.18. *Se considera el conjunto $\mathbb{B}$ formado por todas las bases ordenadas de $V(\mathbb{R})$. Se Establece la relación binaria $\sim$ en $\mathbb{B}$:*

$$B\sim B'\Leftrightarrow \det_B(B')(=\det M(I,B\leftarrow B'))>0,\qquad \forall B,B'\in\mathbf{B}.$$

Demuéstrese que la relación $\sim$ es de equivalencia y existen exactamente dos clases de equivalencia.

A cada una de estas clases de equivalencia obtenidas en el ejercicio anterior se le llama una *orientación* de $V(\mathbb{R})$. En $\mathbb{R}^n(\mathbb{R})$, la base usual genera su *orientación usual*. Para $n\leq 3$ esta orientación (con el modo habitual de visualizar $\mathbb{R},\mathbb{R}^2$ y $\mathbb{R}^3$) se puede interpretar intuitivamente como sigue:

- $n=1$. Las bases son escalares distintos de cero, la orientación usual se corresponde con los escalares positivos o "clase de las bases que apuntan a la derecha", mientras que los escalares negativos forman la "clase de las bases que apuntan a la izquierda".

- $n=2$. La orientación usual se corresponde con las bases $B=(v_1,v_2)$ que definen el "giro en el sentido positivo" (opuesto a las agujas del reloj), mientras que las bases de la otra orientación definen el "giro en sentido negativo". Intuitivamente, para las bases del sentido de positivo se puede ir girando v_1 en vectores que sigan formando base con v_2 hasta llegar a un vector proporcional a v_2 y que apunte en el mismo sentido que v_2 (en esas bases, al girar en sentido contrario se llegaría antes a un vector linealmente dependiente con v_2 que apunta en sentido opuesto a v_2).

- $n=3$. La orientación usual se corresponde con las bases $B=(v_1,v_2,v_3)$ que definen el "sentido de giro de un sacacorchos" (girando de v_1 a v_2 en el plano generado por estos vectores, de los dos semiespacios que determina este plano, el vector v_3 cae en el que se desplazaría un sacacorchos al girar para abrir una botella).

5.4.2. Interpretación para $K = \mathbb{R}$: determinante de un endomorfismo

Consistentemente con las secciones 5.2.1y 5.4.1, podemos interpretar $\det f$ en un espacio vectorial real como sigue. Sea $B = (v_1, \ldots, v_n)$ una base (ordenada), P_B el paralelepípedo que determina, y $f(B) = (f(v_1), \ldots, f(v_n))$, que determina análogamente $P_{f(B)}$.

- $\det f \neq 0$ si y sólo si $f(B)$ es una base, esto es, si y sólo si $P_{f(B)}$ es un paralelepípedo (no degenerado).

  Como la anulación de $\det f$ es independiente de la base escogida, el que $P_{f(B)}$ sea no degenerado también resulta independiente de la base escogida.

- En este caso, $\det f > 0$ si y sólo si las bases B y $f(B)$ determinan la misma orientación.

  Como el signo de $\det f$ es independiente de la base escogida, esta propiedad equivale a que cualquier base $\bar{B}$ determine la misma orientación que $f(\bar{B})$.

  Análogamente, $\det f < 0$ equivale a que una base B determine la orientación distinta a la de $f(B)$ y, por tanto, a que cualquier base $\bar{B}$ determine la orientación distinta a la de $f(\bar{B})$.

- Finalmente, de las definiciones resulta inmediato:

  $$|\det f| = \mathrm{vol}(P_{f(B)})/\mathrm{vol}(P_B),$$

  que resulta independiente tanto de la base B escogida como del tensor determinante no nulo respecto al cual se esté calculando el volumen. Obsérvese que $f(P_B)$ es igual al paralelepípedo posiblemente degenerado $P_{f(B)}$. Por tanto, $|\det f|$ puede verse como el *factor de proporcionalidad, para todo paralelepípedo P_B entre el volumen de su imagen $f(P_B)$ y el volumen de P_B.*

El lector puede ejercitarse con las interpretaciones anteriores para comprobar ahora si la propiedad $\det(f \circ h) = \det f \cdot \det h$ (y, por tanto, $\det(A \cdot C) = \det A \cdot \det C$) le resulta intuitivamente evidente.

Bibliografía

[1] M. Castellet, M. e I. Llerena: *Álgebra lineal y Geometría*. Ed. Reverté, 1981.

[2] L. Merino, y E. Santos: *Álgebra lineal con métodos elementales*. Ed. Thomson, 2006.

[3] A. Romero: *Algebra Lineal y Geometría* Ed. La Madraza, 1991.

COMPLEMENTARIA

[4] J. Arvesú, J., R. Álvarez y F. Marcellán: Álgebra lineal y aplicaciones. Ed. Síntesis, 1999.

[5] J. Arvesú, J., R. Álvarez y F. Marcellán: *Problemas resueltos de Álgebra lineal.* Ed. Thomson, 2004.

[6] J. Burgos: *Álgebra lineal*. MacGraw-Hill, 1993.

[7] E. Hernández, M.L. Vázquez, y M. Á. Zurro: *Álgebra lineal y Geometría.* Pearson, 2012.

[8] S. Lang: *Linear Algebra*, 3rd edition. Springer-Verlag, 1987.

[9] A. Raya, A. Rider y R. Rubio: *Álgebra lineal y Geometría.* Ed. Reverté, 2007.

[10] J. Rojo y I. Martín: *Ejercicios y problemas de Álgebra lineal*. MacGraw-Hill, 1994.

[11] S. Wolfram: *Mathematica, a system for doing Mathematics by computer*. Addison- Wesley, 1991.

Nota sobre el uso lingüístico del género

En el presente libro se usa el género masculino y femenino del modo tradicional en la lengua española, que es consistente con la evolución fonética latina del masculino/ femenino/ neutro:

bonus/ bona/ bonum > bueno/ buena/ bueno.

Así, la expresión "el lector", usada a lo largo del texto, no presupone ningún sexo, pues para ello habría que haber especificado "el lector varón". En esta elección se ha tenido en cuenta que la expresión inclusiva más extendida en el ámbito educativo sería "el lector o la lectora". Sin embargo, éste es un recurso de creación reciente que deja patente la imposición de un género binario, lo cual *sí podría entenderse como una extralimitación sobre elección de sexo u orientación sexual* o como una falta de empatía con algunos colectivos totalmente impropia en la época actual. Así, se ha seguido tanto el criterio de la RAE[12] como las decisiones en otros países hablantes de lenguas latinas[13], sin apreciar el autor que exista una alternativa clara (que, en su caso, debería de obtenerse consensuadamente entre los países hispanohablantes)[14].

[12] "Lo que comúnmente se ha dado en llamar *lenguaje inclusivo* es un conjunto de estrategias que tienen por objeto evitar el uso genérico del masculino gramatical, mecanismo firmemente asentado en la lengua y que no supone discriminación sexista alguna." (Cuenta en twitter RAEinforma, 10/10/2021.)

[13] El gobierno francés, después de que la Academia francesa de la lengua (*L'Académie Française*) se posicionara en contra, prohibió este tipo de lenguaje en la docencia argumentando que las técnicas de desdoblamiento o evitación de género dificultan el aprendizaje y distraen la fluidez de razonamiento (Bulletin officiel de l'éducation nationale, de la jeunesse et des sports, 06/05/2021, `https://www.education.gouv.fr/bo/21/Hebdo18/MENB2114203C.htm`). En Italia la repercusión del debate no es comparable a la de España, y el criterio de la *Accademia della Crusca* se mantiene en la enseñanza.

[14] El autor confiesa no poder decantarse en una elección a la ligera. Por ejemplo, "la persona lectora" como regla llevaría a "las personas estudiantes", "las personas niñas", "las personas bebés", ad nauseam. La terminación "e", ciertamente, permitiría introducir un neutro ex novo que enriquezca el lenguaje. Sin embargo, mientras que unos (unas personas) lo usarían para evitar fórmulas repetitivas, otros (otras personas) preferirían mantener formas duplicativas binarias "la presidenta o el presidente" (posiblemente temiendo que la terminación "e" se pudiera identificar como masculina) y aun otros (otras personas) preferirían una afirmación ternaria: "todos, todas y todes". En cualquier caso, este tipo de modificación tendría gran calado y múltiples posibles ramificaciones, por lo que debería despertar algún tipo de consenso en una lengua de medio millardo de hablantes.