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, "Álgebra Lineal y Geometría" Ed. La Madraza (1991), 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? Tratar de generalizar para cualquier valor de n.

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

- Intersección, $A \cap B := \{x \in X : (x \in A) \wedge (x \in B)\}$

[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*).

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

  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. [6]

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:* $(x\mathcal{R}y) \wedge (y\mathcal{R}x) \Rightarrow x = y$.

[6] *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.

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[7] *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.

[7] Debido a la propiedad simétrica, no importa si "estar relacionado" se interpreta en el sentido $x \sim y$, $y \sim x$ 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. □ [8]

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. [9]

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

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

[8] *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.

[9] 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. $\square$

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 y cuerpos

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[10]:

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

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

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

(a) El elemento neutro e es único,

(b) Se puede "simplificar", esto es, $x\cdot y = x\cdot y' \Rightarrow y = y'$ y también $x\cdot y = x'\cdot y \Rightarrow x = x'$, y

(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 recíproco de y (e y es el recíproco 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.

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

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

[10] 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.

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 recíproco 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 1.13. *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 $(A,+,.)$. 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[11]:
$$\begin{array}{rcll} x.(y+z) & = & x.y+x.z & \text{(distributiva por la izquierda) y} \\ (y+z).x & = & y.x+z.x & \text{(distributiva por la derecha),} \end{array}$$
   para todo $x,y,z\in A$.

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

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[12], esto es, $x.y = y.x, \forall x, y \in A$.

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 $-(-x) = x$).

Con estos conceptos previos podemos ya definir el de cuerpo, que usaremos continuadamente.

Definición 1.14. *Un* cuerpo $(K,+,.)$ *es un anillo unitario no trivial en el que el producto verifica la propiedad elemento simétrico para todo* $x \in K \setminus \{0\}$.

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

Es fácil comprobar que $(K,+,.)$ es un *cuerpo conmutativo 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[13] 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. Estas mismas propiedades las tiene el conjunto de matrices reales cuadradas de un orden fijo $n \times n$, con la suma y producto usuales de matrices.

- Sea $p \in \mathbb{N}$ un número natural. El conjunto $p\mathbb{Z}$ 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[14].

- 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:

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

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

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

[14] Como convenio, supondremos que los naturales no incluyen el 0. (Obsérvese que en el ejemplo que estamos considerando arriba, si se tomara $p = 0$ se obtendría un anillo unitario trivial.)

Puede comprobarse directamente como ejercicio que $K = \{[0],[1]\}$ con las operaciones anteriores es un cuerpo conmutativo[15].

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

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 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

[15] En este cuerpo, al operar el elemento unidad consigo misma se obtiene 0. En genereal, 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 carcterística p, mientras que los cuerpos con los que trabajaremos usualmente $(\mathbb{Q},\mathbb{R},\mathbb{C})$ tienen característica 0.

$\pi/2$ 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

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

donde r no es más que el módulo de z y θ su *argumento*[16] (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 $z' = r'(\cos\theta' + i\,\mathrm{sen}\,\theta')$, de las propiedades del producto y suma:

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

(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:

$\mathbb{C}$ *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 $z^n = 1$).

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

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 $i.j = k$; $j.i = -k$; $i.k = -j$; $k.i = j$; $j.k = i$; $k.j = -i$. Las operaciones resultantes $+$, $.$ dotan a $\mathbb{R}^4$ de una estructura de cuerpo, el cual *no es conmutativo*, y a $\mathbb{H} := (\mathbb{R}^4, +, .)$ se le llama *cuerpo de los cuaterniones*.[17]

[17] 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 (extendiendo 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, $k := (i + j0).(0 + j1) = j(-i)$, se puede reobtener $\mathbb{H}$.

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, y representan las estructuras abstractas más simples en las que se estudia geometría. 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 magnitudes 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) Para nada 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, al suponer siempre que los cuerpos sean conmutativos. Empero, es de remarcar que todo nuestro estudio de espacios vectoriales se extiende sin dificultad al caso no conmutativo.

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:

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

(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 pripiedades 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-upla* 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-upla es de la forma $(x_1,\ldots,x_n)$ donde $x_i \in K$ para cada $i = 1,\ldots,n$. Diremos que dos n-uplas $(x_1,\ldots,x_n)$ 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-uplas 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 $m,n \in \mathbb{N}$. Denotaremos por $M_{m\times n}(K)$ al conjunto de todas las matrices de orden $m\times n$ con coeficientes en K. Cada matriz A es una "expresión rectangular entre paréntesis" de elementos de K distribuidos en m filas "horizontales" y n columnas "verticales"; escribiremos $A = (a_{ij})$, donde $i \in \{1,\ldots,m\}$ es el índice de la fila, y $j = \{1,\ldots,n\}$ el de la columna (formalmente, lo que importa de la matriz es que equivale a una aplicación de $\{1,\ldots,m\}\times\{1,\ldots,n\} \to K$, $(i,j) \mapsto a_{ij}$). Cuando $m = n$ (matrices cuadradas de orden n) denotaremos a $M_{m\times n}(K)$ como $M_n(K)$. Veamos que $M_{m\times n}(K)$ se puede dotar de estructura de e.v. sobre K de manera natural.

Dadas dos matrices $A=(a_{ij})$ y $B=(b_{ij})$ en $M_{m\times n}(K)$ definimos $A+B$ 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 $a\cdot A$ 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 $a\cdot A=(a\cdot a_{ij})$. Vamos a probar con detalle que $M_{m\times n}(K)$ es un e.v. sobre K con las operaciones anteriores.

Veamos que la suma de matrices es asociativa. Sean $A=(a_{ij})$, $B=(b_{ij})$ y $C=(c_{ij})$ tres matrices en $M_{m\times n}(K)$. 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:

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

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.

[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)).

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.

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.

Nota 2.19 (Vectores libres en el plano y el espacio)**.** Como un último ejemplo, consideremos los vectores libres que el alumno conocerá de manera intuitiva desde la enseñanza secundaria. Se define un vector *ligado* (o *fijo*) en el plano como un segmento de recta orientado, el cual está delimitado por dos puntos, llamados origen y extremo. Este segmento tiene una dirección (la de la recta que lo contiene) un sentido (proporcionado por el orden en el que se dan su origen y extremo) y una longitud. Cuando dos vectores ligados tienen igual dirección (están contenidos en rectas paralelas), sentido y longitud (aunque tal vez distinto origen y extremo), se dice que determinan un *vector libre*.

El conjunto $\overrightarrow{\mathbb{R}^2}$ de los vectores libres del plano, admite una suma y producto por escalares que le confieren una estructura de e.v. Concretamente, dados dos vectores libres $\overrightarrow{u}$ y $\overrightarrow{v}$, su suma $\overrightarrow{u}+\overrightarrow{v}$ es el vector libre asociado a un vector ligado cuyo origen es el origen de cualquier vector ligado asociado a $\overrightarrow{u}$ y cuyo extremo es el extremo de un vector ligado asociado a $\overrightarrow{v}$ para el cual se toma como origen de $\overrightarrow{v}$ el extremo de $\overrightarrow{u}$. Por otro lado, si $a\in\mathbb{R}$ y $\overrightarrow{v}$ es un vector libre, se define $a\cdot\overrightarrow{v}$ como el vector libre con la misma dirección de $\overrightarrow{v}$, el mismo sentido si $a\geq 0$ o sentido contrario si $a<0$, y cuya longitud es la de $\overrightarrow{v}$ multiplicada por $|a|$. Con estas operaciones, $\overrightarrow{\mathbb{R}^2}$ se convierte en un e.v. real. El vector nulo de $\overrightarrow{\mathbb{R}^2}$ es el vector libre de longitud 0 (el asociado a cualquier vector ligado donde el origen y el extremo coinciden). Además, el opuesto de un vector $\overrightarrow{v}$ es el vector libre asociado a cualquier vector con la misma dirección, la misma longitud y sentido contrario. Una construcción análoga lleva permite definir los vectores libres del espacio $\overrightarrow{\mathbb{R}^3}$.

Esta definición del plano y espacio de vectores libres es suficiente para nuestros propósitos, por lo que basta con que el lector la tenga en mente desde el punto intuitivo explicado. No obstante, merece la pena detenerse en dos cuestiones sobre este concepto. La primera es que el concepto de "longitud" (a veces llamado "módulo") del vector precisa de una manera de medir vectores, de modo que en nestr introducción de los vectores libres se presupone esta posibilidad de medición. No obstante, ello se podría esquivar a partir de la noción de "paralelismo" que se usa para la definición de la dirección de vector libre (p.ej., dos vectores ligados generan un mismo vector libre cuando al unir sus extremos con un segmento y sus extremos con otro se genera un paralelogramo). La segunda es que cuando se dice que dos vectores ligados que verifican ciertas propiedades "determinan un mismo vector libre" lo que se hace desde el punto de vista formal es definir los vectores libres como clases de equivalencia por una relación de equivalencia (a la que se le llama relación de equipolencia). La suma de vectores libres antes introducida es una suma de clases de equivalencia, llevada a cabo escogiendo representantes de esa clase (dos vectores ligados con el mismo origen), sumando esos representantes (un nuevo vector ligado con ese origen) y tomando a continuación su clase (un vector libre, que resulta ser independiente del origen escogido en la elección de los vectores ligados).

Por completitud, describimos intuitivamente con más detalle el proceso anterior, sin más pretensiones. Para ello consideraremos 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 define la suma como $\overline{P_0P} + \overline{P_0Q} := \overline{P_0R}$, 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). Para definir un producto por escalares (sobre $\mathbb{Q}$ ó $\mathbb{R}$), empezaremos definiendo $n \cdot \overline{P_0P}$ para $n \in \mathbb{N}$ trasladando paralelamente el segmento $\overline{P_0P}$ hasta hacer coincidir su origen con P y repitiendo análogamente la operación n veces. Definimos $-\overline{P_0P}$ trasladando paralelamente el vector ligado $\overline{PP_0} \in V_P$ hasta hacer coincidir su origen con P_0; esto permite además definir naturalmente $p \cdot \overline{P_0P}$ para $p \in \mathbb{Z}$. En Geometría elemental, el teorema de Tales permite dividir un segmento en $n \in \mathbb{N}$ partes, lo cual conduce a una definición natural de $(1/n) \cdot \overline{P_0P}$ para $n \in \mathbb{N}$ y, de manera inequívoca, a definir $r \cdot \overline{P_0P}$ para $r \in \mathbb{Q}$. Si se quiere extender intuitivamente este producto por escalares para todo $r \in \mathbb{R}$, podemos aproximar r por números racionales r_n, y tomar $r \cdot \overline{P_0P}$ como límite de los $r_n \cdot \overline{P_0P}$. Las propiedades elementales del paralelismo (eventualmente convenientemente axiomatizadas) permiten justificar que, con las anteriores operaciones, $(V_{P_0}, +, \cdot K)$ admite una estructura de espacio vectorial para $K = \mathbb{R}$ (y también para $K = \mathbb{Q}$, tanto por la construcción anterior como por la observación 2.18).

A partir de los anteriores espacios de vectores ligados, se llega al concepto de vector libre como sigue. Consideramos el conjunto de todos los vectores ligados

$$Vlig := \{\overline{P_0P} : P_0, P \in \Pi\}.$$

A continuación, se define la siguiente relación binaria, o *relación de equipolencia*, $\sim_e$ en $Vlig$:

$$\overline{P_0P} \sim_e \overline{Q_0Q} \quad \Longleftrightarrow \quad \text{el cuadrilátero que generan con } \overline{P_0Q_0} \text{ y } \overline{PQ} \text{ es un paralelogramo,}$$

para todo $\overline{P_0P}, \overline{Q_0Q} \in Vlig$. 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

$$Vlib := Vlig / \sim_e$$

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 de Π y R_P, R_Q se escogen de modo que

$$\overline{R_0R_P} \in [\overline{P_0P}], \qquad \overline{R_0R_Q} \in [\overline{Q_0Q}].$$

De nuevo, las propiedades del paralelismo aseguran que esta definición es consistente; de hecho, el vector libre suma resulta independiente del punto R_0 que se escoja en Π para tomar los vectores ligados en R_0 que representen las clases $[\overline{P_0P}]$ y $[\overline{Q_0Q}]$. Análogamente, se define el producto por escalares:

$$a \cdot [\overline{P_0P}] := [a \cdot \overline{P_0P}]$$

para todo $a \in \mathbb{R}$ y $[\overline{P_0P}] \in Vlib$ (el cual 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 $Vlib$ de los vectores libres resulta ser un espacio vectorial real. Como notación habitual, cada clase $[\overline{P_0P}]$ se escribe $\overline{P_0P}$ o simplemente se usan letras como hicimos al principio, $\overrightarrow{u}, \overrightarrow{v}$. La construcción es generalizable a los *vectores libres del espacio.*

Sin embargo, no insistiremos en ejemplos como éstos, al exigir su desarrollo conceptos más allá de los que hemos definido con precisión. Además, cuando se profundice en Geometría (Geometría III) se estudiará el concepto de *espacio afín* A, el cual permite una representación sencilla de $Vlib$ a partir del espacio vectorial director V asociado a A. No obstante, conviene tener presente ejemplos intuitivos como éste (o provenientes de la Física, como el espacio vectorial de las velocidades de un móvil) para una mejor comprensión del significado y aplicaciones de los espacios vectoriales.

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 2.20. *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.

La definición anterior aunque es clara 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.

Proposición 2.21. *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, unificándolas en una única como sigue.

Proposición 2.22. *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:*

$$a \cdot u + b \cdot v \in U, \quad \textit{para todo } a, b \in K \textit{ y todo } u, v \in U. \tag{2.5}$$

Demostración. Basta con comprobar que esta propiedad se verifica si y sólo si se cumplen las propiedades (i) y (ii) en la proposición 2.21 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 2.23. *Demuéstrese que un subconjunto U no vacío de $V(K)$ es un s.v. si y sólo si verifica:*

$$a \cdot u + v \in U, \quad \text{para todo } a \in K \text{ y todo } u, v \in U. \tag{2.6}$$

Observación 2.24 (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). Por tanto, ningún s.v. puede ser vacío (esta propiedad se deduce sin imponerla de la definición de subespacio, aunque se debe imponer en las proposiciones 2.21, 2.22). Más aún, en la demostración de la proposición 2.21 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 2.25. 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 $\mathbb{R}^2$.

Ejemplo 2.26. 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 probar 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 2.21 (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 2.27 (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 *impropios*. Todo s.v. de V que no sea uno de los anteriores se llamará *propio*. (Por supuesto, los subespacios impropios son distintos en todo espacio vectorial excepto en el e.v. trivial $V = \{\,0\}$.)

Ejemplo 2.28 (Rectas y planos)**.** Usando la Proposición 2.21 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 2.34). Las rectas y planos que no contienen al origen no son s.v. de $\mathbb{R}^3$: se llaman *subespacios afines* (y se estudiarán en otras partes de Geometría).

El siguiente ejemplo servirá para poner de manifiesto la relevancia del cuerpo K cuando hablamos de subespacios vectoriales.

Ejemplo 2.29 (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 2.21. 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.

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 2.30 (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 2.22. 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 2.31 (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 2.21, 2.22 son aplicables a U, por lo que U resulta ser un s.v. de $F(\mathbb{N}, \mathbb{R})$.

Ejemplo 2.32 (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 2.21, 2.22 son aplicables a U y W, por lo que ambos resultan ser son subespacios vectoriales de $F(I, \mathbb{R})$.

Ejemplo 2.33 (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 2.21, 2.22 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 2.34 (Subespacio de soluciones de un SEL homogéneo)**.** Sea K un cuerpo. Consideremos un SEL de m ecuaciones y n incógnitas con coeficientes en K:

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

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 $(0,\dots,0) \in U$. Ahora bien, esto ocurrirá si y sólo si $b_1 = \dots = b_m = 0$, es decir, el SEL es homogéneo. En tal caso, no es difícil verificar mediante la Proposición 2.22 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 $u = (\alpha_1,\dots,\alpha_n), v = (\beta_1,\dots,\beta_n) \in U$. Queremos dempstrar que el vector $a\cdot u + b\cdot v = (a\cdot\alpha_1 + b\cdot\beta_1,\dots,a\cdot\alpha_n + b\cdot\beta_n)$ está en U. Fijamos un índice $i = 1,\dots,m$ y comprobamos que se cumple la ecuación i-ésima:

$$\begin{aligned} & a_{i1}\,(a\cdot\alpha_1 + b\cdot\beta_1) + \dots + a_{in}\,(a\cdot\alpha_n + b\cdot\beta_n) \\ & = a\,(a_{i1}\,\alpha_1 + \dots + a_{in}\,\alpha_n) + b\,(a_{i1}\,\alpha_1 + \dots + a_{in}\,\alpha_n) = 0, \end{aligned}$$

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, $a\cdot u + b\cdot v \in U$.

La moraleja que podemos extraer de este ejemplo es la siguiente: 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 K^n. Volveremos a esta cuestión más adelante cuando hablemos de las ecuaciones cartesianas de un s.v. de K^n.

Observación 2.35. 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 $\mathbb{R}^2$.

Observación 2.36. En el caso de que el cuerpo no fuera conmutativo (por lo que $K^n(K)$ sería un espacio vectorial por la izquierda, véase la observación 2.2(5)) ¿seguiría siendo cierto que las soluciones de un SEL homogéneo forman un subespacio vectorial? (Suger.: obsérvese que no se tiene el mismo SEL si se multiplican los coeficientes a la derecha de las incógnitas que si se hace a la izquierda).

Ejemplo 2.37 (Matrices simétricas y antisimétricas como subespacios de $M_n(K)$)**.** Sean $m,n \in \mathbb{N}$ y K un cuerpo. Dada una matriz $A = (a_{ij})$ en $M_{m\times n}(K)$, se define 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. Esto significa que el escalar en la posición ij de A^t es a_{ji}, es decir, si $A = (a_{ij})$ entonces $A^t = (a_{ji})$. A partir de la definición es fácil comprobar que la trasposición de matrices define una aplicación

$$^t : M_{m\times n}(K) \to M_{n\times m}(K), \qquad A \mapsto A^t$$

que cumple las siguientes propiedades:

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 $A \in M_{m\times n}(K)$.

En efecto, para 1) si llamamos $C := A+B$, sabemos que $c_{ij} = a_{ij} + b_{ij}$. De este modo, el escalar ij de C^t es $c_{ji} = a_{ji} + b_{ji}$. Por tanto, $(A+B)^t = C^t = (a_{ji}) + (b_{ji}) = A^t + B^t$. Para 2) llamemos ahora $C = a \cdot A$. Sabemos que $c_{ij} = a \cdot a_{ij}$. Así, el elemento ij de C^t es $c_{ji} = a \cdot a_{ji}$. Por tanto, $(a \cdot A)^t = C^t = (a \cdot a_{ji}) = a \cdot (a_{ji}) = a \cdot A^t$. Finalmente para 3), llamemos $C = A^t$. Sabemos que $c_{ij} = a_{ji}$. Así, el elemento ij de C^t es $c_{ji} = a_{ij}$. De aquí se concluye que $C^t = C$.

Las propiedades 1) y 2) se resumen diciendo que "la trasposición de matrices es lineal" (esta terminología se aclarará en el tema siguiente).

En lo sucesivo 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

$$^t : M_n(K) \to M_n(K), \qquad t \circ t = I_{M_n(K)} \text{ (aplicación identidad en } M_n(K)), \quad \text{esto es, } t^{-1} = t.$$

(¿Es cierta esta última igualdad si las matrices no son cuadradas?)

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 significa que $a_{ji} = a_{ij}$, para cada $i, j = 1, \ldots, n$, es decir, 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. 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 significa que $a_{ji} = -a_{ij}$, para cada $i, j = 1, \ldots, n$, es decir, la parte de la matriz "por debajo de la diagonal principal" es opuesta a la parte "por encima de la diagonal principal". Además, debe ocurrir $a_{ii} = 0$ para todo $i = 1, \ldots, n$ 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 $\mathbb{C}$). La matriz 0_n es antisimétrica, mientras que I_n no lo es. En general, una matriz $A \in M_n(K)$ no será simétrica ni antisimétrica.

Para denotar a los conjuntos de matrices simétricas y antisimétricas escribiremos:

$$S_n(K) = \{A \in M_n(K) : A^t = A\} \qquad A_n(K) = \{A \in M_n(K) : A^t = -A\}.$$

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 2.22. 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. 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.

Comenzaremos con un ejemplo motivador. 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.

Lo anterior sirve como motivación para las siguientes definiciones.

Definición 2.38. *Sea $V(K)$ un e.v. y $S=\{v_1,\ldots,v_m\}$ una familia finita no vacía de vectores de V. Una* combinación lineal *(c.l.) de S es cualquier vector de V obtenido al multiplicar cada v_i por un escalar $a_i \in K$ y después sumar los vectores resultantes. Dicho de otro modo, es cualquier vector de V que se puede escribir como:*

$$a_1\cdot v_1+\ldots+a_m\cdot v_m, \quad \textit{donde } a_i \in K, \forall i=1,\ldots,m.$$

Llamaremos $L(S)$ al subconjunto de V formado por los vectores obtenidos como c.l. de S:

$$L(S) = L(v_1,\ldots,v_m) = \{a_1\cdot v_1+\ldots+a_m\cdot v_m : a_i \in K, \forall i=1,\ldots,m\}.$$

Por tanto, dado $v \in V$, se tiene que $v \in L(S)$ si y sólo si existen escalares $a_1,\ldots,a_m \in K$ tales que $v=a_1\cdot v_1+\ldots+a_m\cdot v_m$.

Cuando $S \subset V$ es cualquier familia no vacía de vectores definimos $L(S)$ como el subconjunto de V formado por todas combinaciones lineales *de subconjuntos* finitos *de vectores de S. Es decir:*

$$L(S) = \{a_1\cdot v_1+\ldots+a_m\cdot v_m : m \in \mathbb{N}, v_i \in S, a_i \in K, \forall i=1,\ldots,m\}.$$

Ejemplo 2.39. 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 2.40. *Sea V un e.v. sobre K y $S \subset V$ con $S \neq \emptyset$. Entonces $L(S)$ es un s.v. de V con $S \subset L(S)$. Además, $L(S)$ es el s.v. más pequeño que incluye a S, en el sentido de que si U es un s.v. de V con $S \subset U$, entonces $L(S) \subset U$.*

Demostración. Para comprobar que $L(S)$ es un s.v. de V usaremos la Proposición 2.22. 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. Obsérvese que 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. ∎

Ahora ya podemos introducir la siguiente definición.

Definición 2.41. *Sea V un e.v. sobre K y $S \subset V$ con $S \neq \emptyset$. Llamaremos* subespacio vectorial de V generado por *S (o* envolvente lineal *de S) al subespacio vectorial $L(S)$ formado por las combinaciones lineales (finitas) de vectores de S. Por convenio, definiremos además $L(\emptyset) = \{0\}$ (esto es, $L(\emptyset)$ no es el vacío, sino el subespacio trivial $\{0\}$).*

Observación 2.42. La proposición anterior permitía definir alternativamente a $L(S)$, para $S \neq \emptyset$, como el subespacio vectorial más pequeño que contiene a S (ya que $L(S)$ es un subespacio que contiene a S y cualquier otro s.v. que contenga a S también contendrá a $L(S)$). Esta definición alternativa tiene sentido incluso en el caso $S = \emptyset$, obteniéndose trivialmente $L(\emptyset) = \{0\}$, de acuerdo con nuestro convenio.

Ejemplo 2.43. Si $S = \{v\}$ entonces $L(S) = L(\{v\})$ satisface: cuando $v = 0$ entonces $L(\{v\}) = \{0\}$; cuando $v \neq 0$ entonces al s.v. $L(\{v\})$ se le llama *recta vectorial de V generada* por v.

Nota. En adelante, para subconjuntos finitos simplificaremos la notación suprimiendo las llaves. Así escribiremos sólo $L(v)$ (esto es, $L(v) = \{a \cdot v : a \in K\}$).

2.2.3. Operaciones con subespacios vectoriales

Nos ocuparemos ahora de construir un nuevo subespacio vectorial a partir de familias de subespacios vectoriales. Para ellos recordemos en primer lugar los conceptos conjuntistas de suma e intersección.

Sea X un conjunto, y $P(X)$ el conjunto de sus partes, esto es, el conjunto de 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 subconjuntos de X dados por:

$$A \cap B := \{x \in X : x \in A \text{ y } x \in B\} \qquad A \cup B := \{x \in X : x \in A \text{ ó } x \in B\}.$$

Obsérvese que tanto $\cap$ como $\cup$ pueden verse como operaciones en $P(X)$, cada una de las cuales verifica las propiedades asociativa y conmutativa.

Ejercicio 2.44. *Compruébese que se verifican estas propiedades (asociativa y conmutativa) así como las siguientes propiedades distributivas*

$$A \cap (B \cup C) = (A \cap B) \cup (A \cap C), \qquad A \cup (B \cap C) = (A \cup B) \cap (A \cup C), \qquad \forall A,B,C \in P(X)$$

¿Cuál es el elemento neutro para la operación intersección en $P(X)$? ¿Y para la unión? [2]

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[3]. Tras estos preliminares, empezaremos hablando sobre la intersección de subespacios que resulta especialemente sencilla. Después estudiaremos la unión, que lleva aparejada el concepto de suma de subespacios y, finalmente, la suma directa. En nuestro estudio, consideraremos un subconjunto finito de subespacios. Según el nivel de dificultad que encuentre el lector, le recomendamos que haga un estudio detallado considerando sólo dos subespacios o con una familia infinita de ellos.

Intersección de subespacios

Definición 2.45. *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:*

$$\bigcap_{i=1}^{m} U_i := U_1 \cap \ldots \cap U_m = \{v \in V : v \in U_i \textit{ para cada } i \in I\}.$$

Comprobemos que tal intersección resulta ser un s.v. de V.

Lema 2.46. *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 2.22. 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 $i = 1,\ldots,m$. ∎

[2]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*.

[3]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 $\cap_{i \in \emptyset} A_i = X$, $\cup_{i \in \emptyset} A_i = \emptyset$.

Observación 2.47. (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.

2.2.4. 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 2.48. 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 $\overrightarrow{\mathbb{R}^3}$.

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 2.42) para que, de algún modo, "esté próximo" a los U_i. Esta idea motiva la siguiente definición.

Definición 2.49. *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 2.40), $\sum_{i=1}^m U_i$ 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 2.50. *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 2.40). 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 $u \in \sum_{i=1}^m U_i$. ∎

Ejemplo 2.51. 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 2.52. *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 2.53. 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 2.54. *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 2.55. 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 2.56. *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 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 2.57. En el Ejemplo 2.53 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 2.58. *Sean U_1 y U_2 los subespacios de $\mathbb{R}^2$ del Ejercicio 2.52. Demostrar que $\mathbb{R}^2 = U_1 \oplus U_2$.*

Ejemplo 2.59. [Suma directa de matrices simétricas y antisimétricas] Sea $n \in \mathbb{N}$ y K un cuerpo con característica distinta de 2 (esto es, $2 := 1 + 1 \neq 0$, como ocurre, por ejemplo, si $K = \mathbb{R}$, $\mathbb{Q}$ o $\mathbb{C}$). Siguiendo con las propiedades de matrices vistas en el ejemplo 2.37, 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.

2.3. 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 famiiaridad con esta idea, que aparece en la enseñanza preuniversitaria. Por ejemplo, si en el e.v. real $\overrightarrow{\mathbb{R}^2}$ de los vectores libres en el plano fijamos un sistema de referencia cartesiano con vectores asociados $\{\vec{e}_1, \vec{e}_2\}$, entonces cualquier otro vector $\vec{v}$ se expresa como $a \cdot \vec{e}_1 + b \cdot \vec{e}_2$ para ciertos $a, b \in \mathbb{R}$ (para ello, basta proyectar $\vec{v}$ sobre las rectas vectoriales $L(\vec{e}_1)$ y $L(\vec{e}_2)$). Esto significa que $\overrightarrow{\mathbb{R}^2} = L(\vec{e}_1, \vec{e}_2)$, es decir, con tan sólo dos vectores de $\overrightarrow{\mathbb{R}^2}$ podemos generar todos los demás haciendo combinaciones lineales. Además, la forma de expresar $\vec{v}$ como c.l. de $\vec{e}_1$ y $\vec{e}_2$ es única, pues es sencillo comprobar que $\overrightarrow{\mathbb{R}^2} = L(\vec{e}_1) \oplus L(\vec{e}_2)$. Así, los números $a, b \in \mathbb{R}$ tales que $\vec{v} = a \cdot \vec{e}_1 + b \cdot \vec{e}_2$ son únicos y se llaman coordenadas de $\vec{v}$ en el sistema de referencia.

Con argumentos análogos, en el e.v. real $\overrightarrow{\mathbb{R}^3}$ de los vectores libres en el espacio se prueba que $\overrightarrow{\mathbb{R}^3} = L(\vec{e}_1, \vec{e}_2, \vec{e}_3)$, donde $\{\vec{e}_1, \vec{e}_2, \vec{e}_3\}$ son los vectores de un sistema de referencia cartesiano. Además, la forma de expresar $\vec{v} \in \overrightarrow{\mathbb{R}^3}$ como c.l. de $\{\vec{e}_1, \vec{e}_2, \vec{e}_3\}$ es única, pues no es difícil comprobar que $\overrightarrow{\mathbb{R}^3} = L(\vec{e}_1) \oplus L(\vec{e}_2) \oplus L(\vec{e}_3)$. Así, se puede hablar como antes de coordenadas asociadas a un vector de $\overrightarrow{\mathbb{R}^3}$ respecto del sistema de referencia dado.

En este apartado del tema desarrollaremos todas estas ideas en cualquier e.v.

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 2.60. *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 2.61. *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 2.62. Nos preguntamos si en $\mathbb{R}^2$ la familia $S = \{v_1, v_2\}$ 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 2.63. Nos preguntamos si en $\mathbb{R}^2$ la familia $S=\{v_1,v_2,v_3\}$ 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 2.64. Sea $n\in\mathbb{N}$ y K un cuerpo. Consideremos K^n como e.v. sobre K. Para cada $i\in\{1,\ldots,n\}$ definimos $e_i=(0,0,\ldots,1,\ldots,0,0)$, donde el 1 se encuentra en la i-ésima posición. Afirmamos que la familia $S=\{e_1,\ldots,e_n\}$ es un s.d.g. de K^n. Esto se debe a que todo vector $v=(x_1,\ldots,x_i,\ldots,x_n)\in K^n$ se escribe como $x_1\cdot e_1+\ldots+x_i\cdot e_i+\ldots+x_n\cdot e_n$, 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 2.65. 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 $i\in\{1,\ldots,m\}$ y $j\in\{1,\ldots,n\}$, 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[4] $e_{ij}=1$. Afirmamos que la familia $S=\{E_{ij}\,|\,i=1,\ldots,m,\ j=1,\ldots,n\}$ 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:

$$A=\sum_{i=1}^{m}\sum_{j=1}^{n}a_{ij}\cdot E_{ij}=a_{11}\cdot E_{11}+\ldots+a_{1n}\cdot E_{1n}+\ldots+a_{m1}E_{m1}+\ldots+a_{mn}E_{mn},$$

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 2.66. 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 la familia $S=\{p_i(x)\,|\,i\in\mathbb{N}\cup\{0\}\}$ es un s.d.g. de $\mathbb{K}[x]$. Esto se debe a que todo polinomio $p(x)=a_0+a_1x+\ldots+a_nx^n$ se escribe como $a_0\cdot p_0(x)+a_1\cdot p_1(x)+\ldots+a_n\cdot p_n(x)$, 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 $S=\{p_1(x),\ldots,p_m(x)\}$ fuese un s.d.g. de $\mathbb{K}[x]$. Para cada $i\in\{1,\ldots,m\}$ denotamos $n_i=\text{grado}(p_i(x))$. Sea $n=\text{máx}\{n_1,\ldots,n_m\}$. 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[5] 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.

[4] Para facilitar estas 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 $(I_n)_{ij}=\delta_{ij}$. Las matrices E_{ij} se reescriben entonces $(E_{ij})_{kl}=\delta_{ik}\cdot\delta_{jl}$.

[5] 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.

Ejemplo 2.67. 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 2.22 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[6] demuestra que, si $f\in U$, entonces existen $a,b\in\mathbb{R}$ tales que $f(x)=a\cdot\operatorname{sen}(x)+b\cdot\cos(x)$, para cada $x\in\mathbb{R}$. Recíprocamente, es fácil comprobar que cada función f de este tipo pertenece a U. Así, deducimos que $U=L(\operatorname{sen}(x),\cos(x))$. 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.

Ejemplo 2.68. 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 2.64 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 2.69. 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 1 a lo largo de esta sección. 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 $\overrightarrow{\mathbb{R}^2}$ dado por $\{\vec{e}_1,\vec{e}_2\}$). 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 2.70 (Ampliación y reducción de s.d.g.)**.** *Sea V un e.v. sobre un cuerpo K y $S=\{v_1,\dots,v_m\}$ 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[7] $i\in\{1,\dots,m\}$ tal que $v_i\in L(S-\{v_i\})$, entonces $S-\{v_i\}=\{v_1,\dots,v_{i-1},v_{i+1},\dots,v_m\}$ 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 2.40. De aquí se sigue que $V\subset L(S')$. Y como $L(S')\subset V$ llegamos a $L(S')=V$.

[6]Asignatura Ecuaciones Diferenciales I.

[7]Obsérvese que incluso en el caso $V=\{0\},S=\{0\}$ el convenio $L(\emptyset)=\{0\}$ explicado en la Observación 2.42 es consistente aquí.

(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:

$$v_i = b_1 \cdot v_1 + \ldots + b_{i-1} \cdot v_{i-1} + b_{i+1} \cdot v_{i+1} + \ldots + b_m \cdot v_m.$$

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 2.71. *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. Familias linealmente independientes

La Proposición 2.70 (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 2.72. *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) $\Rightarrow$ (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 2.73. 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 2.74. *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 2.75. 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 2.73 no hace falta tener en cuenta sus subconjuntos). Por tanto, la independencia lineal resulta equivalente a la propiedad (i) del lema 2.72.
- 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 2.72 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 2.72 al caso S vacío y S infinito (¡compruébese!):

Teorema 2.76. *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 2.74).

En este caso, cualquier subconjunto $S' \subset S$ también es linealmente independiente.

Ejemplo 2.77. 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 2.76 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 (nota 2.19), 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, se puede comprobar que una familia $S = \{\vec{u}, \vec{v}, \vec{w}\}$ en $\overrightarrow{\mathbb{R}^3}$ es l.d. si y sólo si los vectores de S son *coplanarios*, es decir, existe un plano vectorial en $\overrightarrow{\mathbb{R}^3}$ que los contiene a todos.

Nota 2.78. En ocasiones se habla de independencia lineal para conjuntos donde se permite que 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 $1 \cdot v_i + (-1) \cdot v_j = 0$. En nuestra definición formal de conjunto esto no puede. No obstante, también podemos incluir la idea subyacente a este caso como sigue. Primero, se entiende que S es un conjunto ordenado, esto es, una n-úpla de vectores. Se definen entonces las combinaciones lineales para toda n-úpla de vectores tomando combicaciónes lineales de sus elementos y, en el caso de que dos de ellos estén repetidos k veces, se permiten k repeticiones de esos vectores en cada expresión de la correspondiente combinación lineal. Así, en el ejemplo anterior, $S = (v_1, \ldots v_i = v, \ldots, v_j = v, \ldots, v_m)$ sería una n-úpla de vectores linealmente dependiente.

Una propiedad relevante de los conjuntos l.i. es la siguiente.

Proposición 2.79. *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:*

$$v = a_1 \cdot v_1 + \cdots + a_m \cdot v_m = a'_1 \cdot v_1 + \cdots + a'_m \cdot v_m \qquad \Rightarrow \qquad a_1 = a'_1, \ldots, a_m = a'_m. \tag{2.8}$$

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 2.80. (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 podrá 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, $v \in L(S_1 \cap S_2)$).

Ejercicio 2.81. *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 $L\{v_1\} \oplus \cdots \oplus L\{v_m\}$.

Resulta intuitivo pensar que el número máximo de vectores que puede tener una familia 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 2.69).

Cuestión 2.82. Si V es un e.v. sobre K y V es f.g., ¿cuál es el mayor número de vectores que una familia l.i. de V puede tener?

De momento, damos una primera respuesta sobre cuándo podemos reducir una familia l.i.

Proposición 2.83 (Ampliación de familias l.i.)**.** *Sea $V(K)$ un e.v. Si S es una familia l.i. y existe $v \in V$ con $v \notin L(S)$, entonces la familia $S \cup \{v\}$ 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 2.84. Combinando este resultado con la última afirmación del teorema 2.76 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 2.85. En $\mathbb{R}^2$ la familia $S = \{v_1, v_2, v_3\}$ 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 2.76 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. 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). 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 probaremos que en $\mathbb{R}^2$ no puede haber familias l.i. con más de dos vectores.

Ejemplo 2.86. Analicemos si en $\mathbb{R}^2$ la familia $S = \{v_1, v_2, v_3\}$ donde $v_1 = (1,0)$, $v_2 = (1,1)$ y $v_3 = (0,1)$ es l.i. Dada una c.l. del tipo $a \cdot v_1 + b \cdot v_2 + c \cdot v_3 = 0$ nos preguntamos si $a = b = c = 0$. 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 $b+c=0$. Así, la familia S será l.i. si y sólo este SEL es compatible determinado (para que la única solución sea la trivial). Sin embargo, el SEL es compatible indeterminado (está escalonado y tiene incógnitas secundarias). Concluimos que S es l.d. Para encontrar una expresión no trivial de 0 como c.l. de S basta considerar 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 probaremos que en $\mathbb{R}^2$ no puede haber familias l.i. con más de dos vectores.

Ejemplo 2.87. Estudiemos si en $\mathbb{R}[x]$ la familia $S = \{p_1(x), p_2(x), p_3(x)\}$ 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[8] $a+c=0$, $a+2b=0$ 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.

Ahora demostraremos que los s.d.g. que estudiamos en la anterior subsección son familias l.i. en el espacio correspondiente.

Ejemplo 2.88. En K^n la familia $S = \{e_1, \ldots, e_n\}$ del Ejemplo 2.64 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 $a_1 = \ldots = a_n = 0$.

Ejemplo 2.89. En $M_{m\times n}(K)$ la familia $S = \{E_{ij} \,|\, i = 1, \ldots, m,\ j = 1, \ldots, n\}$ definida en el Ejemplo 2.65 es l.i. Para verlo tomamos una c.l. del tipo

$$\sum_{i=1}^{m} \sum_{j=1}^{n} a_{ij} \cdot E_{ij} = a_{11} \cdot E_{11} + \ldots + a_{1n} \cdot E_{1n} + \ldots + a_{m1} E_{m1} + \ldots + a_{mn} E_{mn} = 0_{m\times n},$$

[8] 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 $a+b+c=0$.

y nos preguntamos si $a_{ij} = 0$ para cada $i = 1,\dots,m$ y cada $j = 1,\dots,n$. 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.

Ejemplo 2.90. En $\mathbb{K}[x]$ la familia $S = \{p_i(x) \,|\, i \in \mathbb{N}\cup\{0\}\}$ definida en el Ejemplo 2.66 es l.i. Hay que comprobar que cada familia finita de S es l.i. Tomamos cualquier familia finita $S' = \{p_{i_1}(x),\dots,p_{i_m}(x)\}$ de 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 $i = 1,\dots,m$.

Finalmente, veamos que el papel de K es importante cuando hablamos de independencia lineal.

Ejemplo 2.91. 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 $v_2 = a \cdot v_1$).

2.3.3. Bases y dimensión

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 familias l.i. que sean lo más grandes posibles, y planteamos dos cuestiones (la 2.69 y 2.82) 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* comp para demostrar que dos bases de un espacio vectorial f.g. tienen el mismo número de vectores.

Teorema 2.92 (Teorema de Steinitz)**.** *Sea V un e.v. sobre un cuerpo K. Supongamos que tenemos dos subconjuntos finitos $S = \{u_1, u_2, \dots, u_m\}$, que es un s.d.g. y $S' = \{v_1, v_2, \dots, v_k\}$, que es l.i. Entonces, se cumple que $m \geq k$.*

Demostración. Razonaremos por reducción al absurdo. Suponemos que $m+1 \leq k$ y pretendemos llegar a contradicción. La idea consiste en sustituir los vectores de S por vectores de S' hasta obtener, en m pasos, que $S_m := \{v_1,\dots,v_m\}$ es un s.d.g. de V. Como $m+1 \leq k$, existe el vector v_{m+1} de S', el cual se escribirá como c.l. de S_m. Esto contradice que S' es l.i.

Paso 1. Como S es un s.d.g. de V, la Proposición 2.70 (i) nos dice que $\overline{S}_1 = \{v_1, u_1, u_2, \dots, u_m\}$ 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 $a_1, a_2, \dots, a_m \in K$ 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' una familia l.i. Así, en la igualdad de arriba no todos los coeficientes del miembro derecho son nulos. Renombrando si fuese necesario los vectores de S, podemos suponer $a_1 \neq 0$. 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 $\{v_1, u_2, \dots, u_m\}$, es decir, $u_1 \in L(\overline{S}_1 - \{u_1\})$. Aplicando la Proposición 2.70 (ii) deducimos que $S_1 = \overline{S}_1 - \{u_1\} = \{v_1, u_2, \dots, u_m\}$ es un s.d.g. de V.

Ahora distinguimos dos casos. Si $m = 1$ entonces $S_1 = \{v_1\}$ y como $2 = m+1 \leq k$, existe el vector $v_2 \in S'$, que será c.l. de S_1 (al ser S_1 un s.d.g. de V). Esto es una contradicción porque S' era una familia 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 2.70 (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:

$$v_2 = b_1 \cdot v_1 + b_2 \cdot u_2 + \ldots + b_m \cdot u_m.$$

La ecuación anterior se puede escribir como:

$$v_2 - b_1 \cdot v_1 = b_2 \cdot u_2 + \ldots + b_m \cdot u_m.$$

Por otro lado, el conjunto $\{v_1, v_2\}$ es l.i. al serlo S', por lo que el miembro izquierdo es $\neq 0$. En consecuencia, en el derecho tendrá que existir algún $b_i \neq 0$ con $i \in \{2, \ldots, m\}$. Tras renombrar los vectores se puede suponer que $b_2 \neq 0$. Despejando el vector u_2 en la ecuación de arriba, obtenemos:

$$u_2 = b_2^{-1} \cdot v_2 + (-b_2^{-1} \cdot b_1) \cdot v_1 + \ldots + (-b_2^{-1} \cdot b_m) \cdot u_m,$$

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 2.70 (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 como $k \geq m + 1 = 3$, tenemos el vector $v_3 \in S'$, que será c.l. de S_2 (al ser S_2 un s.d.g. de V). Esto contradice que S' es una familia 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 entonces que el conjunto $S_m \cup \{v_{m+1}\}$ es l.i, por estar incluido en S' (el vector v_{m+1} debe existir porque $m + 1 \leq k$). 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:

$$v_{m'+1} = c_1 \cdot v_1 + \cdots + c_{m'} \cdot u_{m'} + c_{m'+1} \cdot u_{m'+1} + \ldots + c_m \cdot u_m.$$

La ecuación anterior se puede escribir como:

$$v_{m'+1} + (-c_1) \cdot v_1 + \cdots + (-c_{m'}) \cdot u_{m'} = c_{m'+1} \cdot u_{m'+1} \ldots + c_m \cdot u_m.$$

Como el conjunto $\{v_1, ..., v_{m'+1}\}$ es l.i. (al estar incluido en S'), el miembro izquerdo es distinto de 0. Por tanto, en el miembro derecho algún $c_i \neq 0$ con $i \in \{m'+1, \ldots, m\}$. 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 2.70 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 2.93. *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 familia l.i. más allá del teorema de Steinitz).*

La idea de la demostración anterior sugiere la relevancia de conjuntos que sean a la vez s.d.g. y l.i.

Definición 2.94. *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.*

La definición previa 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. Nos interesarán los e.v. que tengan bases finitas.

Ejemplo 2.95. Dado un sistema de referencia cartesiano en el espacio de vectores libres $\overrightarrow{\mathbb{R}^2}$ con vectores asociados $\{\vec{e}_1, \vec{e}_2\}$, entonces $\mathcal{B} = \{\vec{e}_1, \vec{e}_2\}$ es una base de $\overrightarrow{\mathbb{R}^2}$. Dado un sistema de referencia cartesiano en $\overrightarrow{\mathbb{R}^3}$ con vectores $\{\vec{e}_1, \vec{e}_2, \vec{e}_3\}$, entonces $\mathcal{B} = \{\vec{e}_1, \vec{e}_2, \vec{e}_3\}$ es una base de $\overrightarrow{\mathbb{R}^3}$.

Ejemplo 2.96. En K^n la familia $\mathcal{B}_u = \{e_1, \dots, e_n\}$ definida en el Ejemplo 2.64 es una base, que llamaremos *base usual o canónica* de K^n.

Ejemplo 2.97. En $M_{m\times n}(K)$ la familia $\mathcal{B}_u = \{E_{ij} \,|\, i = 1, \dots, m,\ j = 1, \dots, n\}$ definida en el Ejemplo 2.65 es una base, que llamaremos *base usual o canónica* de $M_{m\times n}(K)$.

Ejemplo 2.98. En $\mathbb{K}[x]$ la familia $\mathcal{B}_u = \{p_i(x) \,|\, i \in \mathbb{N} \cup \{0\}\}$ 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 2.99 (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 $S = \{v_1, \dots, v_m\}$, 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 $i \in \{1, \dots, m\}$ tal que $v_i \in L(S - \{v_i\})$. Gracias a la Proposición 2.70 (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 2.100. 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 ($s \in X$, 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 2.83 y, por tanto, M no sería maximal), por lo que $v \in L(M)$ (contrarrecíproco de la proposición 2.83).

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. A continuación mostraremos que, aunque haya 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 2.101 (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 una familia l.i. entonces cualquier familia finita en $\mathcal{B}_2$ también lo es. Tomando una familia $\mathcal{B}_2'$ 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 2.99 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 familia 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 2.102. *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 2.103. Tenemos $\dim_{\mathbb{R}}(\overrightarrow{\mathbb{R}^2}) = 2$, $\dim_{\mathbb{R}}(\overrightarrow{\mathbb{R}^3}) = 3$, $\dim_K(K^n) = n$, $\dim_K(M_{m\times n}(K)) = m\cdot n$, $\dim_K(M_n(K)) = n^2$ 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 2.66, o bien porque ya sabemos que $\mathbb{K}[x]$ tiene una base infinita (Ejemplo 2.98), 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 2.104. 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 2.105. *Sea $V(K)$ un espacio vectorial f.g. Dados $S = \{v_1,\dots,v_m\}$, un s.d.g. de V, y $S' = \{u_1,\dots,u_k\}$, una familia 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 familia 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 familia l.i. deducimos que $k \leq n$, concluyéndose la demostración. ∎

Observación 2.106 (Respuesta a las cuestiones 2.69 y 2.82)**.** 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 2.69. Asimismo $\dim_K(V)$ es el mayor número de vectores que una familia l.i. en V puede tener. Esto resuelve la cuestión 2.82. 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 2.99 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 una familia l.i.

Teorema 2.107 (ampliación de la base)**.** *Sea V un e.v. con $\dim_K(V) = n \in \mathbb{N}$. Si $S = \{v_1,\dots,v_m\}$ es un l.i. y $m \leq n$, entonces existen $n-m$ vectores, $v_{m+1},\dots,v_n \in V$ tales que $\mathcal{B} = \{v_1,\dots,v_m, v_{m+1},\dots,v_n\}$ es una base de V.*

Demostración. En el caso $m = n$, bsta 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 2.83) con $k = n+1$ vectores, en contradicción con el corolario anterior.

En el caso $m < n$, S no es un s.d.g. de V (de lo contrario S sería una base, en contradicción con el teorema 2.101), por lo que existe $v_{m+1} \in V$ con $v_{m+1} \notin L(S)$. En consecuencia, el conjunto $S \cup \{v_{m+1}\} = \{v_1,\dots,v_m,v_{m+1}\}$ es l.i. (de nuevo por la proposición 2.83). Si $m+1 = n$ este conjunto será una base, aplicando el caso $m = n$ antes estudiado. En caso contrario, se repite el argumento un total de $n-m$ veces hasta conseguir una familia $\mathcal{B} = \{v_1,\dots,v_m,v_{m+1},\dots,v_n\}$ que es l.i. y tiene exactamente n vectores, la cual será una base aplicando de nuevo el caso $m = n$. ∎

Observación 2.108. 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 una familia l.i. no es única en general.

Ejemplo 2.109. 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 2.110. El teorema 2.107 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 2.100.

Ejercicio 2.111. *Sea V un e.v. no trivial sobre un cuerpo K. Supongamos que V es f.g. y que S, S' son dos familias finitas 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 2.105 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 2.112. *Sea V un e.v. sobre un cuerpo K con $\dim_K(V) = n \geq 1$. Dada una familia $\mathcal{B} = \{v_1, \ldots, v_n\}$ 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 una familia l.i. de 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 2.99 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 2.101. Por último, (iii) implica (i) se sigue directamente del teorema 2.107 con $n = m$. ∎

Ejemplo 2.113. Consideremos la familia $\mathcal{B} = \{u, v\}$ del Ejemplo 2.109. 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 (su apartado (3)) también tiene gran utilidad práctica.

Proposición 2.114. *Sea $V(K)$ un e.v. y $S = \{v_1, \ldots, v_m\} \subset V$ y $v \in V$ tal que $v = a_1 v_1 + \ldots a_m v_m$ 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 2.70 (ii)).

(2) La demostración puede hacerse directamente (como se propone en el ejercicio 2.115 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).[9]

(3) Inmediato de los casos anteriores. ∎

Ejercicio 2.115. *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 2.116. 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 2.117. 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 2.118. Si $S = \{v_1, \dots, v_m\}$ es una familia 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 2.105.

Ejemplo 2.119. En $\mathbb{K}_n[x]$ la familia $\mathcal{B}_u = \{p_i(x) \,|\, i \in \{0,1,\dots,n\}\} = \{1, x, \dots, x^n\}$ 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, $\dim_{\mathbb{K}}(\mathbb{K}_n[x]) = n+1$.

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 2.120. *Calcúlese una base de cada subespacio de matrices $S_n(K)$ y $A_n(K)$ (con característica de $K \neq 2$), y compruébese $\dim_K S_n(K) = n(n+1)/2$, $\dim_K A_n(K) = n(n-1)/2$.*

[9] 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.

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 2.121. *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 2.107 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 2.112), y $U = L(\mathcal{B}_U) = V$.

(ii) La primera afirmación es una consecuencia inmediata del Teorema 2.107. 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 $L(\mathcal{B}) \subset L(U \cup W)$ y, por ser $\mathcal{B}$ un s.d.g, $V \subset U + W$.

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 2.122. 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.

Ejemplo 2.123. 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 2.124. *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 2.125. Es obvio que $U = V$ si y sólo si $\operatorname{codim}_K(U) = 0$. 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 2.126 (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 2.127. 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 2.121 (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 2.121 (iv) sabemos que $U_i = (U_1 \cap U_2) \oplus W_i$, para cada $i = 1,2$. Consideremos ahora la familia 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 $U_1 + U_2$.

Comprobemos primero que $\mathcal{B}$ un s.d.g. de $U_1 + U_2$, esto es, $L(\mathcal{B}) = U_1 + U_2$. Claramente $\mathcal{B} \subset U_1 \cup U_2 \subset U_1 + U_2$, por lo que $L(\mathcal{B}) \subset U_1 + U_2$. Recíprocamente, nótese que $\mathcal{B}_i \subset \mathcal{B}$ para cada $i = 1,2$. Como $U_i = L(\mathcal{B}_i)$ se sigue que $U_1, U_2 \subset L(\mathcal{B})$ y, por tanto, $U_1 + U_2 \subset L(\mathcal{B})$.

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:

$$\sum_{i=1}^{m} a_i \cdot v_i + \sum_{i=m+1}^{n_1} b_i \cdot v^1_i = - \sum_{i=m+1}^{n_2} c_i \cdot v^2_i, \tag{2.9}$$

donde el miembro izquierdo de la igualdad es un vector de U_1 y el derecho es un vector de W_2. Ahora bien, observemos que $U_1 \cap W_2 = \{0\}$. En efecto, por construcción $(U_1 \cap U_2) \cap W_2 = \{0\}$ y, trivialmente,

$W_2 = U_2 \cap W_2$, por lo que $\{0\} = (U_1 \cap U_2) \cap W_2 = U_1 \cap (U_2 \cap W_2) = U_1 \cap W_2$. Así, los dos miembros de la igualdad (2.9) 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 una familia l.i. y $\{v_{m+1}^2, \ldots, v_{n_2}^2\} \subset \mathcal{B}_2$, que es también l.i. Esto completa la demostración. ∎

Como consecuencia inmediata tenemos este resultado, cuya prueba queda como ejercicio.

Corolario 2.128. *Sea V un espacio f.g. sobre un cuerpo K. 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 2.129. *Sea $V(K)$ un espacio y $\{U_1, U_2, U\}$ subespacios de V, U f.g., tales que $U_1, U_2 \subset U$. Demostrar que son equivalentes las siguientes afirmaciones:*

(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 2.130 (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 2.81(ii) y la definición 2.56).

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.[10]

[10]Para el desarrollo de esta sección es preciso tener algunos conocimientos elementales sobre el producto de matrices (esencialmente su definición, asociatividad y, en el caso de matrices cuadradas, las propiedades asociadas a su estructura de anillo unitario), que se unen a las propiedades ya vistas de espacio vectorial respecto a la suma y el producto por escalares, y se repasarán aparte. También conviene, desde el punto de vista práctico, tener un procedimiento algorítmico que permita calcular la matriz inversa de una matriz regular. Aquí también se verá aparte un procedimiento sencillo basado en la resolución de SEL por el método de Gauss; más adelante se verá otro usando determinantes, que probablemente conozca ya el alumno.

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 2.131 (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 2.79 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,

$$\begin{aligned} v &= a_1 \cdot v_1 + \ldots + a_n \cdot v_n, \\ v &= b_1 \cdot v_1 + \ldots + b_n \cdot v_n. \end{aligned}$$

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. de la familia $\mathcal{B}$. Como $\mathcal{B}$ es l.i. deducimos que $a_i - b_i = 0$ para cada $i = 1, \ldots, n$, como se quería demostrar. ∎

Observación 2.132. 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 2.133. *Diremos que una n-úpla $(v_1, \ldots, v_n) \in V \times \ldots \times^{(n)} V$ 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 2.131.

Definición 2.134. *Sea $V(K)$ un e.v. y $\mathcal{B} = (v_1, \ldots, v_n)$ 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 $\mathcal{B}$. *En tal caso escribiremos:*

$$v_{\mathcal{B}} = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}.$$

Ejemplo 2.135. Las bases ordenadas $\mathcal{B} = (v_1, v_2, \ldots, v_n)$ y $\mathcal{B}' = (v_2, v_1, \ldots, v_n)$ dan lugar a la misma base conjuntista. Dado $v \in V$, nótese si

$$v_{\mathcal{B}} = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix}$$

son las coordenadas de v respecto de $\mathcal{B}$, entonces:

$$v_{\mathcal{B}'} = \begin{pmatrix} a_2 \\ a_1 \\ \vdots \\ a_n \end{pmatrix}$$

son las coordenadas de v respecto de $\mathcal{B}'$; obviamente $v_{\mathcal{B}} \neq v_{\mathcal{B}'}$.

Convenio 2.136. 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 (aunque no se diga explícitamente que está ordenada) se habla de coordenadas en la base $\mathcal{B}$. Así, en el ejemplo anterior se puede decir simplemente que v_B son las coordenadas de v en la base $\{v_1, \ldots, v_n\}$.

Observación 2.137. Dada una base ordenada $\mathcal{B} = (v_1, \ldots, v_n)$ en V, las coordenadas $v_{\mathcal{B}}$ de un vector $v \in V$ proporcionan un elemento de K^n, si identificamos este espacio con $M_{n\times 1}(K)$. Esto nos da una conexión entre V y K^n a través de la aplicación $f_{\mathcal{B}} : V \to K^n$ definida por $f_{\mathcal{B}}(v) = v_{\mathcal{B}}$. Es sencillo comprobar que $f_{\mathcal{B}}$ es biyectiva. De hecho, anticipando el lenguaje del próximo tema $f_{\mathcal{B}}$ es un *isomorfismo de espacios vectoriales*: además de la biyectividad, cumple que $f_{\mathcal{B}}(a \cdot u + b \cdot v) = a \cdot f_{\mathcal{B}}(u) + b \cdot f_{\mathcal{B}}(v)$, para todo $a, b \in K$ y todo $u, v \in V$. Esto equivale a la identidad $(a \cdot u + b \cdot v)_{\mathcal{B}} = a \cdot u_{\mathcal{B}} + b \cdot v_{\mathcal{B}}$, que se puede demostrar como ejercicio.

Este isomorfismo $f_{\mathcal{B}}$ establece una "identificación" (dependiente de $\mathcal{B}$) entre V y K^n como espacios vectoriales sobre K. Así, 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[11] directamente que $S = \{u_1, \ldots, u_m\}$ es un s.d.g. (resp. una familia l.i.) de V si y sólo si $S_{\mathcal{B}} = \{(u_1)_{\mathcal{B}}, \ldots, (u_m)_{\mathcal{B}}\}$ es un s.d.g. (resp. una familia l.i.) de K^n.

[11] Demostraremos un resultado más general válido para todo isomorfismo de e.v. en el tema siguiente.

Ejercicio 2.138. *Sea $V(K)$ un e.v. y $\mathcal{B} = (v_1, \ldots, v_n)$ una base ordenada de V. ¿Qué coordenadas tiene el vector nulo respecto de $\mathcal{B}$? ¿Y cada uno de los vectores v_i?*

Ahora veremos cómo se calculan las coordenadas en algunas de las bases estudiadas con anterioridad.

Ejemplo 2.139. Si en K^n tomamos la base (ordenada) usual $\mathcal{B}_u = (e_1, \ldots, e_n)$ 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:

$$v_{\mathcal{B}_u} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix},$$

es decir, las coordenadas coinciden con las componentes del vector. Esta es una propiedad exclusiva de $\mathcal{B}_u$ que facilita mucho los cálculos.

Ejemplo 2.140. Si en $M_{m\times n}(K)$ tomamos $\mathcal{B}_u = \{E_{ij} \,|\, i = 1, \ldots, m,\ j = 1, \ldots, n\}$, 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 $\mathcal{B}$ como una base ordenada (consistentemente con el convenio 2.136, ordenaremos $\mathcal{B}$ 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 $\mathcal{B}_u$ (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 2.141. Si en $\mathbb{K}_n[x]$ tomamos la base usual $\mathcal{B}_u = (1, x, \ldots, x^n)$ 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 $\mathcal{B}_u$ 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 2.142. En $\mathbb{R}^2$ consideramos el conjunto $\mathcal{B} = \{v_1, v_2\}$ con $v_1 = (1, 0)$ y $v_2 = (-1, 1)$. Es obvio que $\mathcal{B}$ es una base de $\mathbb{R}^2$ pues $\dim_{\mathbb{R}}(\mathbb{R}^2) = 2$ y $\mathcal{B}$ es l.i. (v_1 y v_2 no son proporcionales). Calculemos las coordenadas de $v = (x, y) \in \mathbb{R}^2$ respecto de $\mathcal{B}$ (con su ordenación natural). Éstas serán los únicos números $(a, b) \in \mathbb{R}^2$ 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:

$$v_{\mathcal{B}} = \begin{pmatrix} x+y \\ y \end{pmatrix}.$$

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 $\mathcal{B}_u$ 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 $\mathcal{B} = (v_1, \ldots, v_n)$ y $\mathcal{B}' = (v'_1, \ldots, v'_n)$ de V. Dado $v \in V$, la cuestión es: ¿qué relación hay entre $v_{\mathcal{B}}$ y $v_{\mathcal{B}'}$? (p. ej., supongamos que conocemos $v_{\mathcal{B}}$ y queremos determinar $v_{\mathcal{B}'}$ a partir de $v_{\mathcal{B}}$). Para ello necesitamos conocer cómo los vectores de $\mathcal{B}$ se expresan como c.l. de los de $\mathcal{B}'$.

Definición 2.143. *En las condiciones anteriores, la* matriz de cambio de base de $\mathcal{B}$ a $\mathcal{B}'$ *es la matriz que denotaremos* $M(I,\mathcal{B}'\leftarrow\mathcal{B})$ *en* $M_n(K)$ *cuya columna* j*-ésima contiene las coordenadas del vector* v_j *de* $\mathcal{B}$ *respecto de* $\mathcal{B}'$*. Lo simbolizamos así:*

$$M(I,\mathcal{B}'\leftarrow\mathcal{B})=\big((v_1)_{\mathcal{B}'}\,|\,(v_2)_{\mathcal{B}'}\,|\,\dots\,|\,(v_n)_{\mathcal{B}'}\big).$$

Ahora ya podemos resolver el problema que nos habíamos planteado.

Proposición 2.144 (cambio de base)**.** *En las condiciones anteriores, se tiene que:*

$$v_{\mathcal{B}'}=M(I,\mathcal{B}'\leftarrow\mathcal{B})\cdot v_{\mathcal{B}}.$$

Llamaremos a la ecuación anterior la expresión matricial del cambio de base de $\mathcal{B}$ a $\mathcal{B}'$.

Demostración. Supongamos que:

$$v_j=a_{1j}\cdot v_1'+\dots+a_{nj}\cdot v_n',\ \text{para cada } j=1,\dots,n.$$

De esta forma, la j-ésima columna de $M(I,\mathcal{B}'\leftarrow\mathcal{B})$ contiene exactamente a los escalares a_{ij} con $i=1,\dots,n$. Supongamos también que:

$$v=x_1\cdot v_1+\dots+x_n\cdot v_n,$$

es decir, $v_{\mathcal{B}}$ contiene a los escalares x_i con $i=1,\dots,n$. Buscamos expresar v como c.l. de $\mathcal{B}'$. Sustituyendo la primera ecuación en la segunda tenemos la igualdad:

$$\begin{aligned} v&=x_1\cdot(a_{11}\cdot v_1'+\dots+a_{n1}\cdot v_n')+\dots+x_n\cdot(a_{1n}\cdot v_1'+\dots+a_{nn}\cdot v_n')\\ &=(a_{11}\cdot x_1+\dots+a_{1n}\cdot x_n)\cdot v_1'+\dots+(a_{n1}\cdot x_1+\dots+a_{nn}\cdot x_n)\cdot v_n', \end{aligned}$$

que nos expresa v como c.l. de $\mathcal{B}'$. Esto significa que los coeficientes de la c.l. anterior son las coordenadas de v respecto de $\mathcal{B}'$. Por definición de producto de matrices, la entrada i-ésima de $v_{\mathcal{B}'}$ es el producto de la fila i-ésima de $M(I,\mathcal{B}'\leftarrow\mathcal{B})$ con el vector columna $v_{\mathcal{B}}$. ∎

Ejemplo 2.145. En K^n, la matriz $M(I,\mathcal{B}_u\leftarrow\mathcal{B})=\big((v_1)_{\mathcal{B}_u}\,|\,(v_2)_{\mathcal{B}_u}\,|\,\dots\,|\,(v_n)_{\mathcal{B}_u}\big)$ es aquella cuyas columnas contienen las componentes de los vectores de $\mathcal{B}$.

Ejemplo 2.146. En $\mathbb{R}^2$ consideramos la base $\mathcal{B}=(v_1,v_2)$ con $v_1=(1,0)$ y $v_2=(-1,1)$ del último ejemplo. Ya vimos que, para cada $v=(x,y)\in\mathbb{R}^2$ se tiene:

$$v_{\mathcal{B}}=\begin{pmatrix} x+y\\ y\end{pmatrix},$$

con un cálculo directo de las coordenadas. Vamos a llegar ahora al mismo resultado usando un cambio de base. Tomando $\mathcal{B}_u$ como base de referencia, la Proposición 2.144 nos dice que:

$$v_{\mathcal{B}}=M(I,\mathcal{B}\leftarrow\mathcal{B}_u)\cdot v_{\mathcal{B}_u},\ \text{donde } M(I,\mathcal{B}\leftarrow\mathcal{B}_u)=((e_1)_{\mathcal{B}},(e_2)_{\mathcal{B}}).$$

Necesitamos las coordenadas de e_1 y e_2 con respecto a $\mathcal{B}$. Tras unos cálculos sencillos:

$$(e_1)_B=\begin{pmatrix}1\\0\end{pmatrix}\ \text{y}\ (e_2)_B=\begin{pmatrix}1\\1\end{pmatrix},\ \text{por lo que } M(I,\mathcal{B}\leftarrow\mathcal{B}_u)=\begin{pmatrix}1&1\\0&1\end{pmatrix}.$$

Por último, teniendo en cuenta que en $\mathcal{B}_u$ las coordenadas de un vector están dadas por sus componentes, deducimos que:

$$v_{\mathcal{B}} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x+y \\ y \end{pmatrix},$$

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.

Definición 2.147. *Sea A una matriz cuadrada, $A \in M_n(\mathbb{K})$. Diremos que A es 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 $j,k = 1,\dots,n$.*

Proposición 2.148. *Sea $V(K)$ un e.v. de dimensión $n \in \mathbb{N}$, y sean $\mathcal{B}, \mathcal{B}', \mathcal{B}''$ tres bases ordenadas de $V(K)$. Se verifica:*

(i) $M(I, \mathcal{B} \leftarrow \mathcal{B}'') = M(I, \mathcal{B} \leftarrow \mathcal{B}') \cdot M(I, \mathcal{B}' \leftarrow \mathcal{B}'')$.

(ii) $M(I, \mathcal{B} \leftarrow \mathcal{B})$ es la matriz identidad I_n.

(iii) $M(I, \mathcal{B}' \leftarrow \mathcal{B})$ es una matriz regular*, esto es, que admite inversa para el producto de matrices, verificando ésta además: $M(I, \mathcal{B}' \leftarrow \mathcal{B})^{-1} = M(I, \mathcal{B} \leftarrow \mathcal{B}')$*

Demostración. Para (i), sea v cualquier vector de V. La Proposición 2.144 nos dice:

$$v_{\mathcal{B}''} = M(I, \mathcal{B}'' \leftarrow \mathcal{B}') \cdot v_{\mathcal{B}'} \qquad v_{\mathcal{B}'} = M(I, \mathcal{B}' \leftarrow \mathcal{B}) \cdot v_{\mathcal{B}}.$$

Sustituyendo la segunda igualdad en la primera, y usando la asociatividad del producto de matrices:

$$v_{\mathcal{B}''} = \big(M(I, \mathcal{B}'' \leftarrow \mathcal{B}') \cdot M(I, \mathcal{B}' \leftarrow \mathcal{B})\big) \cdot v_{\mathcal{B}}.$$

Por otra parte, también sabemos:

$$v_{\mathcal{B}''} = M(I, \mathcal{B}'' \leftarrow \mathcal{B}) \cdot v_{\mathcal{B}}$$

Restando las dos igualdades anteriores se tiene entonces:

$$\big(M(I, \mathcal{B}'' \leftarrow \mathcal{B}) - M(I, \mathcal{B} \leftarrow \mathcal{B}') \cdot M(I, \mathcal{B}' \leftarrow \mathcal{B})\big) \cdot v_{\mathcal{B}} = 0_{n\times 1}.$$

Tomando $v_{\mathcal{B}} = (v_j)_{\mathcal{B}}$ donde v_j es el $j-$ésimo vector de la base $\mathcal{B}$ (esto es $(v_j)_{\mathcal{B}} (= (e_j)_{B_u})$ 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, esto es, $M(I, \mathcal{B}'' \leftarrow \mathcal{B}) - M(I, \mathcal{B} \leftarrow \mathcal{B}') \cdot M(I, \mathcal{B}' \leftarrow \mathcal{B}) = 0_n$ (matriz nula $n \times n$), de donde se deduce el resultado.

La afirmación (ii) es inmediata. Para (iii), aplicando primero la parte (i) y luego la (ii):

$$M(I, \mathcal{B} \leftarrow \mathcal{B}') \cdot M(I, \mathcal{B}' \leftarrow \mathcal{B}) = M(I, \mathcal{B} \leftarrow \mathcal{B}) = I_n,$$

y para la igualdad $M(I, \mathcal{B}' \leftarrow \mathcal{B}) \cdot M(I, \mathcal{B} \leftarrow \mathcal{B}') = I_n$ basta con intercambiar el papel de las bases. ∎

Ejemplo 2.149. Una de las aplicaciones usuales del resultado anterior es realizar con facilidad cambios de base en K^n a través de la base usual $\mathcal{B}_u$, siempre que se sepa calcular la matriz inversa.

Por ejemplo, sea $\mathcal{B}$ una base de K^n y supongamos que queremos calcular $M(I, \mathcal{B} \leftarrow \mathcal{B}_u)$, el cambio de base de coordenadas de $\mathcal{B}_u$ a $\mathcal{B}$. Por el corolario previo tenemos $M(I, \mathcal{B} \leftarrow \mathcal{B}_u) = M(I, \mathcal{B}_u \leftarrow \mathcal{B})^{-1}$. La ventaja de $M(I, \mathcal{B}_u \leftarrow \mathcal{B})$ frente a $M(I, \mathcal{B} \leftarrow \mathcal{B}_u)$ es que su cálculo suele ser inmediato, pues basta escribir por columnas las componentes de los vectores de $\mathcal{B}$ (eso sí, para terminar de obtener $M(I,\mathcal{B} \leftarrow \mathcal{B}_u)$ necesitamos saber calcular $M(I,\mathcal{B}_u \leftarrow \mathcal{B})^{-1}$). Así, el resultado del ejemplo 2.146 se reobtendría:

$$M(I,\mathcal{B} \leftarrow \mathcal{B}_u) = M(I,\mathcal{B}_u \leftarrow \mathcal{B})^{-1} = \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix}^{-1} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix},$$

Más aún, dadas cualesquiera dos bases $\mathcal{B},\mathcal{B}'$ de K^n se puede escribir:

$$M(I,\mathcal{B}' \leftarrow \mathcal{B}) = M(I,\mathcal{B} \leftarrow \mathcal{B}_u) \cdot M(I,\mathcal{B}_u \leftarrow \mathcal{B}),$$

donde $M(I,\mathcal{B} \leftarrow \mathcal{B}_u)$ se calcula por el procedimiento anterior, y $M(I,\mathcal{B}_u \leftarrow \mathcal{B})$ es inmediata.

2.4. Aplicaciones: rango de una matriz y ecuaciones de un subespacio

Terminaremos este tema aprovechando la teoría de espacios vectoriales que hemos estudiado para obtener algunas consecuencias interesantes en la teoría de matrices y la discusión de SEL. Deduciremos también algunos principios que serán de utilidad en el cálculo de bases y dimensiones de subespacios vectoriales.

2.4.1. Bases y dimensión de $U = L(S)$

Sea V un e.v. sobre K y $U = L(S)$ un s.v. no trivial generado por $S = \{v_1,\dots,v_m\}$. Queremos calcular una base (ordenada) de U y $\dim_K(U)$.

Sabemos que $\dim_K(U)$ es el número máximo de vectores l.i. de S. Así, $\dim_K(U) = m$ si y sólo si S es l.i. (lo que equivale a que S sea una base de U). En general, $\dim_K(U) \leq m$ pues S no tiene por qué ser l.i. En tal caso, el Teorema 2.99 dice que hay una base $\mathcal{B}$ de U con $\mathcal{B} \subset S$. Para construir $\mathcal{B}$ debemos ir eliminando de S sucesivamente un vector que se exprese como c.l. del resto de generadores. Este proceso puede ser muy tedioso si m es grande. Para realizarlo más fácilmente construiremos un s.d.g. de U en el que la independencia lineal sea sencilla de estudiar. La idea es modificar S a través del siguiente resultado.

Proposición 2.150. *Sea V un e.v. sobre K y $U = L(S)$ con $S = \{v_1,\dots,v_i,\dots,v_j,\dots,v_m\} \subset V$. Entonces, se tiene:*

1. *Si se considera S como un conjunto ordenado (como en la nota 2.78), entonces cualquier reordenación S' de S es también un s.d.g.*
2. *Si $a \in K$ y $a \neq 0$, entonces $S'' = \{v_1,\dots,a\cdot v_i,\dots,v_j,\dots,v_m\}$ es un s.d.g. de U.*
3. *Si $a \in K$ entonces $S''' = \{v_1,\dots,v_i,\dots,v_j + a\cdot v_i,\dots,v_m\}$ es un s.d.g. de U.*

Demostración. El lema es consecuencia tanto de la Proposición 2.70 como de la 2.114. Usaremos la primera de ellas, aunque resulta más inmedito de la segunda.

Es obvio que, desde el punto de vista conjuntista, $S = S'$ y, por tanto[12], $L(S') = L(S) = U$. 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 2.70(2) resulta aplicable). ∎

[12] 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.

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 2.151. Calculemos una base y la dimensión del s.v. de $\mathbb{R}^4$ dado por $U = L(S)$, donde $S = \{(1,-2,1,0),(2,3,-2,1),(4,-1,0,1)\}$. La matriz asociada M es:

$$M = \begin{pmatrix} 1 & -2 & 1 & 0 \\ 2 & 3 & -2 & 1 \\ 4 & -1 & 0 & 1 \end{pmatrix}$$

Aplicamos transformaciones elementales por filas hasta conseguir la matriz de un SEL escalonado. Enseguida llegamos a:

$$M \sim \begin{pmatrix} 1 & -2 & 1 & 0 \\ 0 & 7 & -4 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Como el vector nulo se puede eliminar de cualquier s.d.g., se sigue que $\mathcal{B} = \{(1,-2,1,0),(0,7,-4,1)\}$ sigue siendo un s.d.g. Además, $\mathcal{B}$ es claramente l.i. (esto es una propiedad general para las filas no nulas de cualquier matriz escalonada), por lo que $\mathcal{B}$ resulta ser una base y $\dim_{\mathbb{R}}(U) = 2$.

Observación 2.152. Es de remarcar que este procedimiento resulta extensible a cualquier s.d.g. finito S de un s.v. U de cualquier e.v. finitamente generado $V(K)$. Basta con fijar una base $\mathcal{B}$ 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 $\mathcal{B}$ de la base requerida de $L(S)$ (recuérdese la observación 2.137).

2.4.2. Rango de una matriz

Asociaremos a cada matriz A un número $\mathrm{rg}(A) \in \mathbb{N} \cup \{0\}$ que desempeñará un papel importante en lo sucesivo. Para ello nos basaremos en la teoría de espacios vectoriales estudiada hasta ahora (existen otros enfoques para definir el rango).

Definición 2.153. *Sea K un cuerpo y $m, n \in \mathbb{N}$. Dada una matriz $A = (a_{ij})$ en $M_{m\times n}(K)$, definimos $v_j = (a_{1j}, \ldots, a_{mj})$, 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 como:*

$$rg(A) = dim_K(U), \text{ con } U = L(v_1, \ldots, v_n).$$

Así, $rg(A)$ es el número máximo de columnas l.i. en A cuando se interpretan como vectores de K^m.

Observación 2.154. 1. En principio la definición anterior sería el *rango por columnas* de A, y habría una definición análoga para el *rango por filas* (que sería igual al rango por columnas de A^t). No obstante, como veremos este rango por filas coincidirá con el que hemos definido usando columnas.

2. Es claro que $\mathrm{rg}(A) \in \mathbb{N} \cup \{0\}$ y que $\mathrm{rg}(A) \leq \min\{n, m\}$. Además, $\mathrm{rg}(A) = 0$ si y sólo si $A = 0_{m\times n}$.

Gracias a la proposición 2.150 y a la definición de rango deducimos lo siguiente.

Corolario 2.155. *El rango de A es invariante por transformaciones elementales por columnas de A.*

Observación 2.156. Las transformaciones elementales por columnas de A pueden verse como transformaciones elementales por filas de A^t. Como se dijo en la observación anterior, veremos 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[13].

Ejercicio 2.157. *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.

Lema 2.158. *Sea $S = \{v_1, \dots, v_n\}$ una familia de vectores en K^m. Consideremos la matriz en $M_{m\times n}(K)$ dada por $A = \big((v_1)_{\mathcal{B}_u} \,|\, (v_2)_{\mathcal{B}_u} \,|\, \dots \,|\, (v_n)_{\mathcal{B}_u}\big)$. Entonces:*

(i) S es un s.d.g. de K^m (necesariamente entonces $n \geq m$) si y sólo si $\operatorname{rg}(A) = m$.

(ii) S es una familia l.i. en K^m (necesariamente entonces $n \leq m$) si y sólo si $\operatorname{rg}(A) = n$.

(iii) S es una base de K^m si y sólo si $n = m$ y $\operatorname{rg}(A) = m$.

Demostración. Sabemos que $\operatorname{rg}(A) = \dim_K(U)$ con $U = L(S)$. Si S es s.d.g. de K^m entonces $U = L(S) = K^m$ y $\operatorname{rg}(A) = m$. Recíprocamente, si $\operatorname{rg}(A) = m$, 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, $\operatorname{rg}(A) = \dim_K(U) = n$. Recíprocamente, si $\operatorname{rg}(A) = n$, entonces $\dim_K(U) = n$. Como S es un s.d.g. de U con exactamente n vectores, el corolario 2.112 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). ∎

Ejercicio 2.159. *Estudiar si la familia $S = \{(1,0,-1), (1,2,5), (-2,1,3)\}$ es una base de $\mathbb{R}^3$.*

2.4.3. El teorema de Rouché-Frobenius

Hasta ahora hemos usado los SEL para resolver problemas de geometría lineal. Ahora veremos cómo la teoría de espacios vectoriales se puede emplear para discutir eficazmente un SEL sin necesidad de resolverlo.

Teorema 2.160 (Rouché-Frobenius)**.** *Consideremos un* SEL *con ecuación matricial asociada $A \cdot x = b$, donde $A \in M_{m\times n}(K)$ y $b \in M_{m\times 1}(K)$. Entonces:*

(i) El sistema es compatible si y sólo si $\operatorname{rg}(A) = \operatorname{rg}(A|b)$.

(ii) El sistema es compatible determinado si y sólo si $\operatorname{rg}(A) = \operatorname{rg}(A|b) = n$.

[13] Esta libertad para el cálculo del rango contrasta con la necesidad de hacer las transformaciones elementales sólo por filas en el procedimiento de extraer una base visto entre la proposición 2.150 y el ejercicio 2.151; de de hecho, si allí se llevaran a cabo transformaciones elementales por columnas cambiaríamos las coordenadas de los correspondientes vectores.

(iii) El sistema es compatible indeterminado si y sólo si $\text{rg}(A) = \text{rg}(A|b) = r < n$*. En tal caso, las soluciones quedan expresadas en función de* $n-r$ *parámetros de* K*.*

Demostración. Sea v_j (resp. v_0) el vector de K^m asociado a la columna j-ésima de A (resp. a la columna b). Por la definición del producto de matrices:

$$(x_1,\ldots,x_n) \text{ es solución del SEL si y sólo si } x_1 \cdot v_1 + \ldots + x_n \cdot v_n = v_0. \tag{2.10}$$

Por tanto, el SEL es compatible si y sólo si $v_0 \in L(v_1,\ldots,v_n)$. En tal caso, las soluciones del SEL coinciden con los coeficientes que permiten expresar v_0 como c.l. de $\{v_1,\ldots,v_n\}$.

Vamos a demostrar (i). Se tiene que:

$$\begin{aligned}\text{SEL compatible} &\Leftrightarrow v_0 \in L(v_1,\ldots,v_n) \Leftrightarrow L(v_1,\ldots,v_n,v_0) = L(v_1,\ldots,v_n)\\ &\Leftrightarrow \dim_K(L(v_1,\ldots,v_n,v_0)) = \dim_K(L(v_1,\ldots,v_n)) \Leftrightarrow \text{rg}(A|b) = \text{rg}(A),\end{aligned}$$

donde en la segunda y en la tercera equivalencia hemos empleado la Proposición 2.70 y el segundo apartado de la Proposición 2.121.

Demostremos (ii). Supongamos que el SEL es compatible determinado. Gracias a (2.10) existen escalares únicos $x_1,\ldots,x_n \in K$ tales que $x_1 \cdot v_1 + \ldots + x_n \cdot v_n = v_0$. Queremos comprobar que $\dim_K(U) = n$ con $U = L(v_1,\ldots,v_n)$. Para ver que $\{v_1,\ldots,v_n\}$ es l.i. tomamos una c.l.

$$a_1 \cdot v_1 + \ldots + a_n \cdot v_n = 0,$$

con $a_1,\ldots,a_n \in K$. Sumando el vector v_0 a ambos lados de la igualdad:

$$a_1 \cdot v_1 + \ldots + a_n \cdot v_n + v_0 = v_0.$$

Teniendo en cuenta la expresión de v_0 de arriba:

$$(a_1 + x_1) \cdot v_1 + \ldots + (a_n + x_n) \cdot v_n = v_0.$$

Por la unicidad de los escalares x_i se tiene que $a_i + x_i = x_i$ y, por tanto, $a_i = 0$ para cada $i = 1,\ldots,n$. Recíprocamente, si $\text{rg}(A) = n$, entonces $\dim_K(U) = n$ y $\{v_1,\ldots,v_n\}$ es una base de U por el corolario 2.112. Como $\text{rg}(A) = \text{rg}(A|b)$ el SEL es compatible, es decir, $v_0 \in U$. Por el Teorema 2.131 existen escalares únicos $x_1,\ldots,x_n \in K$ tales que $v_0 = x_1 \cdot v_1 + \ldots + x_n \cdot v_n$. Gracias a (2.10) el SEL es compatible determinado.

Vamos a demostrar (iii). Por (i) y (ii) el SEL es compatible indeterminado si y sólo si $\text{rg}(A) = \text{rg}(A|b) = r < n$ (nótese que $\text{rg}(A) \leq m,n$). Veamos que las soluciones se expresan en función de $n-r$ parámetros de K. Como $\dim_K(U) = \text{rg}(A) = r$, tenemos r vectores l.i. en $\{v_1,\ldots,v_n\}$. Tras renombrar los vectores si fuera preciso, podemos suponer que la familia $\{v_1,\ldots,v_r\}$ es l.i y, por tanto, una base de U. Gracias a (2.10) se tiene que:

$$(x_1,\ldots,x_n) \text{ es solución si y sólo si } x_1 \cdot v_1 + \ldots + x_r \cdot v_r = v_0 - x_{r+1} \cdot v_{r+1} - \ldots - x_n \cdot v_n.$$

Usando el Teorema 2.131, cada vez que tomamos $\lambda_1,\ldots,\lambda_{n-r} \in K$, existen $x_1,\ldots,x_r \in K$ únicos tales que $(x_1,\ldots,x_r,\lambda_1,\ldots,\lambda_{n-r})$ es solución del SEL. A partir de aquí se sigue que

$$\{(x_1,\ldots,x_r,\lambda_1,\ldots,\lambda_{n-r}) \,|\, \lambda_1,\ldots,\lambda_{n-r} \in K\}$$

es el conjunto de soluciones del SEL. ∎

Corolario 2.161. *Sea W el conjunto de soluciones del* SEL *homogéneo con ecuación matricial $A \cdot x = 0$, donde $A \in M_{m\times n}(K)$. Entonces, W es un s.v. de K^n con* $\dim_K(W) = n - \mathrm{rg}(A)$.

Si se considera ahora un SEL no homogéneo $Ax = b$ que sea compatible, y x_0 es una solución suya, entonces el conjunto de todas las soluciones es $x_0 + W := \{x_0 + w \mid w \in W\}$.

Demostración. Como el SEL es homogéneo sabemos que W es un s.v. de K^n. Gracias a los argumentos en la demostración del teorema anterior tenemos, sabemos que W admite una representación con $n-r$ parámetros del tipo:

$$\{(x_1,\dots,x_r,\lambda_1,\dots,\lambda_{n-r}) \mid \lambda_1,\dots,\lambda_{n-r} \in K\},$$

donde, para cada $\lambda_1,\dots,\lambda_{n-r} \in K$, los escalares $x_1,\dots,x_r$ son los únicos tales que:

$$x_1 \cdot v_1 + \ldots + x_r \cdot v_r = -\lambda_1 \cdot v_{r+1} - \ldots - \lambda_{n-r} \cdot v_n. \tag{2.11}$$

Sea $w_i \in W$ la única solución del SEL cuando $\lambda_i = 1$ y $\lambda_j = 0$ si $i \neq j$. Basta con comprobar que $\mathcal{B}_W = \{w_1,\dots,w_{n-r}\}$ es una base de W. Es sencillo demostrar que $\mathcal{B}_W$ es l.i. Además, $L(\mathcal{B}_W) \subset W$ al ser W un s.v. que contiene a $\mathcal{B}_W$. Sea $w = (x_1,\dots,x_n) \in W$. Sabemos que $x_1,\dots,x_r$ son los únicos escalares en K que cumplen (2.11) con $\lambda_i = x_{r+i}$ para cada $i = 1,\dots,n-r$. Afirmamos que $w = w'$ con $w' = x_{r+1} \cdot w_1 + \ldots + x_n \cdot w_{n-r}$. Nótese que $w' \in W$. Si ponemos $w' = (y_1,\dots,y_n)$ entonces $y_{r+i} = x_{r+i}$, para cada $i = 1,\dots,n-r$. Por la unicidad en las soluciones de (2.11) concluimos que $y_i = x_i$ para cada $i = 1,\dots,r$. Esto muestra que $w = w' \in L(\mathcal{B}_W)$, lo que termina la demostración para W.

Para la última afirmación, obsérvese que si x_0' es otra solución del SEL no homogéneo entonces $A \cdot (x_0' - x_0) = b - b = 0$, esto es, $x_0' - x_0 \in W$ y $x_0' \in x_0 + W$. Recíprocamente, si $x_0' \in x_0 + W$ entonces $x_0' = x_0 + w$ para algún $w \in W$ y

$$A \cdot x_0' = A \cdot (x_0 + w) = A \cdot x_0 + A \cdot w = b + 0 = b,$$

esto es, x_0' también es una solución del SEL no homogéneo. ∎

Observación 2.162. Desde el punto de vista práctico, conviene tener en cuenta que una base de W se obtiene mediante las $n-r$ eleccionea de los parámetros $(\lambda_1,\dots,\lambda_r)$ iguales a $(1,0\dots,0,0)$, ..., $(0,0,\dots,0,1)$.

Ejemplo 2.163. Vamos a calcular una base de:

$$W = \{(x,y,z,t) \in \mathbb{R}^4 : x + y = 0, z + t = 0\}.$$

Una forma de hacerlo es resolver el SEL homogéneo. Otra se basa en el corolario previo. La matriz de coeficientes del SEL es:

$$A = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{pmatrix}.$$

Es obvio que $\mathrm{rg}(A) = 2$, ya que el s.v. que generan sus columnas es $\mathbb{R}^2$ (hay al menos dos columnas l.i.). Por tanto, $\dim_{\mathbb{R}}(W) = 2$. Para obtener una base de W necesitamos dos soluciones l.i. del SEL. Como $w_1 = (1,-1,0,0)$ y $w_2 = (0,0,1,-1)$ cumplen esto, entonces $\mathcal{B}_W = \{w_1, w_2\}$ es una base de W.

Como una aplicación, caractericemos de modo sencillo a las matrices regulares.

Corolario 2.164. *Una matriz cuadrada* $A \in M_n(K)$ *es regular si y sólo si* $\mathrm{rg}(A) = n$.

Demostración. Si A es regular, cada SEL con ecuación matricial asociada $A \cdot x = b$ es compatible determinado (multiplicando por A^{-1} a la izquierda de ambos miembros y usando la asociatividad del producto se obtiene la única solución $x = A^{-1} \cdot b$). Aplicando Rouché-Frobenius, $\mathrm{rg}(A) = n$.

Recíprocamente, si $\mathrm{rg}(A) = n$ y $\mathcal{B} = (v_1, \ldots, v_n)$, donde v_j se identifica con la j-ésima columna de A, el Lema 2.158 nos dice que $\mathcal{B}$ es una base de K^n. Como $A = M(I, \mathcal{B}_u \leftarrow \mathcal{B})$, la proposición 2.148 implica que A es regular. ∎

Proposición 2.165. *Sea* $\mathrm{GL}(n, \mathbb{K})$ *el conjunto formado por todas las matrices regulares de* $M_n(K)$. *La operación producto de matrices*

$$\mathrm{GL}(n, \mathbb{K}) \times \mathrm{GL}(n, \mathbb{K}) \to \mathrm{GL}(n, \mathbb{K}), \qquad (A, C) \mapsto A \cdot C$$

está bien definida en $\mathrm{GL}(n, \mathbb{K})$*, y* $\mathrm{GL}(n, K)$*, con esta operación, tiene estructura de grupo. Así, llamaremos a* $\mathrm{GL}(n, \mathbb{K})$ grupo[14] lineal general *(de orden n sobre el cuerpo K).*

Demostración. Para comprobar que está bien definida, debemos demostrar que si $A, C \in \mathrm{GL}(n, K)$ entonces $A \cdot C \in \mathrm{GL}(n, K)$. Para ello, fijemos la base usual $\mathcal{B}_u$ de K^n. Por ser C regular existe una única base $\mathcal{B}'$ tal que $C = M(I, \mathcal{B}' \leftarrow \mathcal{B})$ y, obtenida $\mathcal{B}'$, por ser A regular existe una única base $\mathcal{B}''$ tal que $A = M(I, \mathcal{B}'' \leftarrow \mathcal{B}')$. Usando la proposición 2.148

$$A \cdot C = M(I, \mathcal{B}'' \leftarrow \mathcal{B}') \cdot M(I, \mathcal{B}' \leftarrow \mathcal{B}) = M(I, \mathcal{B}' \leftarrow \mathcal{B}),$$

la cual es una matriz de cambio de base y, por tanto, regular.

La asociatividad se sigue por ser ésta una propiedad general del producto de matrices (sean cuadradas o no, siempre que se puedan multiplicar). Por computación directa se tiene que $I_n \in \mathrm{GL}(n, K)$ verifica ser el elemento neutro para el producto. El elemento simétrico existe por la propia definición de $\mathrm{GL}(n, K)$ como el conjunto de las matrices regulares. ∎

Observación 2.166. En la demostración anterior, la base $\mathcal{B}_u$ 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 $\mathcal{B}$, y los pasos de la demostración funcionarían igual reemplazándolos por $K^n(K)$ y $\mathcal{B}_u$.

2.4.4. El rango de la matriz traspuesta

Nuestro objetivo a continuación es demostrar que el rango de una matriz $A \in M_{n \times m}(K)$ coincide con el de su traspuesta $A^t \in M_{m \times n}(K)$, esto es, que el número máximo de columnas independientes de A (vistas como elementos de K^m coincide con el número máximo de filas independientes (vistas como elementos de K^n). La demostración se basa en el siguiente lema, inspirado en técnicas del espacio dual que se desarrollarán en el Tema 4 (de hecho, en ese lema se dará una nueva demostración).

Observemos en primer lugar que cada ecuación de un SEL será del tipo $a_1 \cdot x_1 + \ldots + a_n \cdot x_n = 0$. Si w representa al vector columna en $M_{n \times 1}(K)$ de entradas $a_1, \ldots, a_n$, entonces la ecuación se escribe como $v \cdot w = 0$, con $v = (x_1\ x_2\ \ldots\ x_n)$.

[14]Obsérvese que la operación de este grupo es el producto matricial (por lo que no es conmutativo en general) y que la suma matricial *no* se puede restringir como operación a $\mathrm{GL}(n, \mathbb{K})$.

Lema 2.167. *Sea U un s.v. de K^n con $\dim_K(U) = m$. Definimos el conjunto:*

$$U_0 = \{w \in M_{n\times 1}(K) \mid v \cdot w = 0, \text{ para todo } v \in U\}.$$

Entonces U_0 es un s.v. de K^n con $\dim_K(U_0) = n - m$.

Demostración. Para comprobar que U_0 es un s.v. de K^n basta usar la Proposición 2.22 y propiedades del producto de matrices.

Sobre la dimensión, en el caso $m = 0$ entonces $U = \{0\}$ y $U_0 = M_{n\times 1}(K)$, por lo que $\dim_K(U_0) = n$. Si $m = n$ entonces $U = K^n$, y tomando $v = e_i$ (i-ésimo vector de la base usual en K^n) para cada $i = 1,\ldots,n$ se sigue que $U_0 = \{0_{n\times 1}\}$ y, por tanto, $\dim_K(U_0) = 0$.

Supongamos el caso no trivial $1 \leq m < n$. Sea $\mathcal{B}_U = \{v_1,\ldots,v_m\}$ una base de U. Por el teorema de ampliación de la base encontramos una base de K^n del tipo $\mathcal{B} = \{v_1,\ldots,v_m,v_{m+1},\ldots,v_n\}$. Consideremos la matriz en $M_n(K)$ dada por $A = \big((v_1)_{\mathcal{B}_u} \,|\, (v_2)_{\mathcal{B}_u} \,|\, \ldots \,|\, (v_n)_{\mathcal{B}_u}\big)$. Como $\mathcal{B}$ es una base de K^n y $A = M(I,\mathcal{B}_u \leftarrow \mathcal{B})$ se sigue por la proposición 2.148 que A es regular. Esto implica que A^t también es regular; de hecho, $A^t \cdot (A^{-1})^t = (A^{-1} \cdot A)^t = I_n^t = I_n$, y la inversa de A^t resulta ser $(A^{-1})^t$. Escribiendo $C = A^t$ la ecuación matricial $C \cdot x = e_i$ tiene solución única $w_i = C^{-1} \cdot e_i \in M_{n\times 1}(K)$. Afirmamos que $\mathcal{B}' = \{w_1,\ldots,w_n\}$ es una base de $M_{n\times 1}(K)$. Por el corolario 2.112 basta comprobar que $\mathcal{B}'$ es l.i. Tomemos una c.l. del tipo $a_1 \cdot w_1 + \ldots + a_n \cdot w_n = 0_{n\times 1}$. Multiplicando por la izquierda por C obtenemos $a_1 \cdot e_1 + \ldots + a_n \cdot e_n = 0_{n\times 1}$. Como la familia $\{e_1,\ldots,e_n\}$ es l.i. en $M_{n\times 1}(K)$ se sigue que $a_i = 0$, para cada $i = 1,\ldots,n$.

Para terminar la demostración veamos que $\mathcal{B}_0 = \{w_{m+1},\ldots,w_n\}$ es una base de U_0. Nótese que las filas de C se identifican con los vectores de $\mathcal{B}$. Por definición de producto de matrices se tiene que la entrada j-ésima de $C \cdot w_i$ coincide con $v_j \cdot w_i$. Como $C \cdot w_i = e_i$, entonces $v_j \cdot w_i = 0$ para cada $j \neq i$. En particular, $v_j \cdot w_i = 0$ para cada $j = 1,\ldots,m$ y cada $i = m+1,\ldots,n$. Esto demuestra que $\mathcal{B}_0 \subset U_0$ (cada vector de U es c.l. de $\{v_1,\ldots,v_m\} = \mathcal{B}_U$). Es claro que $\mathcal{B}_0$ es l.i. ya que $\mathcal{B}_0 \subset \mathcal{B}'$ y $\mathcal{B}'$ es l.i. Veamos por último que $L(\mathcal{B}_0) = U_0$. Sea $w \in U_0$. Como $\mathcal{B}'$ es una base de $M_{n\times 1}(K)$ existen escalares únicos $a_1,\ldots,a_n \in K$ tales que $w = a_1 \cdot w_1 + \ldots + a_n \cdot w_n$. Multiplicando por la matriz C obtenemos $C \cdot w = a$, donde $a \in M_{n\times 1}(K)$ tiene como componentes $a_1,\ldots,a_n$. La componente i-ésima de $C \cdot w$ es $v_i \cdot w$. Como $w \in U_0$ deducimos que $a_i = 0$ para cada $i = 1,\ldots,m$. Sustituyendo esta información llegamos a $w = a_{m+1} \cdot w_{m+1} + \ldots + a_n \cdot w_n$. Así, $w \in L(\mathcal{B}_0)$ y se concluye. ∎

Podemos ahora establecer el siguiente resultado fundamental.

Teorema 2.168. *Para cada $A \in M_{m\times n}(K)$ se cumple que $\mathrm{rg}(A^t) = \mathrm{rg}(A)$.*

Demostración. Sea v_j el vector de K^m que se identifica con la columna j-ésima de A. Llamemos $U = L(v_1,\ldots,v_n)$. Sea W el subespacio de $M_{m\times 1}(K)$ de las soluciones de la ecuación matricial $A^t \cdot y = 0_{n\times 1}$. Escribiendo las matrices A y A^t se demuestra enseguida que $W = U_0$. Teniendo en cuenta el Corolario 2.161 y el Lema 2.167, deducimos que:

$$\mathrm{rg}(A^t) = m - \dim_K(W) = m - \dim_K(U_0) = m - (m - \dim_K(U)) = \dim_K(U) = \mathrm{rg}(A),$$

como se quería demostrar. ∎

Corolario 2.169. *El rango de A es invariante por transformaciones elementales por filas o por columnas de A.*

De este modo, podemos calcular $\mathrm{rg}(A)$ haciendo transformaciones elementales por filas de A hasta obtener la matriz de un SEL escalonado. Entonces $\mathrm{rg}(A)$ coincidirá con el número de filas no nulas de dicha matriz.

2.4.5. Ecuaciones paramétricas e implícitas de un subespacio

Sea U un s.v. de K^n con $\dim_K(U) = m \geq 1$. Tomemos una base $\mathcal{B}_U = (v_1, v_2, \ldots, v_m)$ de U. Como $U = L(\mathcal{B}_U)$, los vectores de U son exactamente de la forma $v = \lambda_1 \cdot v_1 + \lambda_2 \cdot v_2 + \ldots + \lambda_m \cdot v_m$ con $\lambda_i \in K$, para cada $i = 1, \ldots, m$. Si ponemos $v = (x_1, x_2, \ldots, x_n)$ y $v_i = (x_1^i, x_2^i, \ldots, x_n^i)$ entonces, al tomar componentes en la igualdad $v = \lambda_1 \cdot v_1 + \ldots + \lambda_m \cdot v_m$ obtenemos:

$$\begin{aligned} x_1 &= x_1^1 \cdot \lambda_1 + x_1^2 \cdot \lambda_2 + \ldots + x_1^m \cdot \lambda_m, \\ x_2 &= x_2^1 \cdot \lambda_1 + x_2^2 \cdot \lambda_2 + \ldots + x_2^m \cdot \lambda_m, \\ &\vdots \\ x_n &= x_n^1 \cdot \lambda_1 + x_n^2 \cdot \lambda_2 + \ldots + x_n^m \cdot \lambda_m, \end{aligned}$$

que son unas *ecuaciones paramétricas* para U. Nótese que hay n ecuaciones paramétricas con m parámetros $\lambda_1, \ldots, \lambda_m$. Los coeficientes que acompañan a cada λ_i son las coordenadas de v_i en la base usual $\mathcal{B}_u$. 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 $\mathcal{B}_U$ se recupera de las ecuaciones paramétricas sin más que tomar $\lambda_1 = \cdots = \lambda_{i-1} = 0$, $\lambda_i = 1$, $\lambda_{i+1} = \cdots = \lambda_m = 0$.

Ahora veremos que cada subespacio U de K^n es el conjunto de soluciones de un SEL homogéneo. De acuerdo con el Corolario 2.161 la matriz A de este SEL cumplirá $\mathrm{rg}(A) = n - m$ si $\dim_K(U) = m$. Usaremos en la siguiente demostración elementos del lema 2.167.

Proposición 2.170. *Sea U un s.v. de K^n con $\dim_K(U) = m$ y $1 \leq m < n$. Existe un* SEL *homogéneo con matriz de coeficientes $A \in M_{(n-m)\times n}(K)$ y $\mathrm{rg}(A) = n - m$, cuyo subespacio de soluciones es U.*

Demostración. Por el lema 2.167 sabemos que $\dim_K(U_0) = n - m$. Sea $\{w_1, \ldots, w_{n-m}\}$ una base de U_0. Si las entradas de w_i son $(w_i)_1, \ldots, (w_i)_n$ entonces tenemos una ecuación cartesiana $(w_i)_1 \cdot x_1 + \ldots + (w_i)_n \cdot x_n = 0$. Esto produce un SEL de $n - m$ ecuaciones con expresión matricial $A \cdot x = 0_{(n-m)\times 1}$, donde A es la matriz en $M_{(n-m)\times n}(K)$ cuya i-ésima fila se identifica con w_i. Por otro lado:

$$\mathrm{rg}(A) = \mathrm{rg}(A^t) = \dim_K\big(L(w_1, \ldots, w_{n-m})\big) = \dim_K(U_0) = n - m.$$

Sea W el subespacio de soluciones del SEL asociado a $A \cdot x = 0_{(n-m)\times 1}$. Sabemos que $\dim_K(W) = m$ por el Corolario 2.161. Como $U \subset W$ se concluye que $U = W$ por el segundo apartado de la Proposición 2.121. ∎

Definición 2.171. *En las hipótesis de la proposición anterior, diremos que las $n - m$ ecuaciones del SEL homogéneo $A \cdot x = 0_{(n-m)\times 1}$ son unas* ecuaciones implícitas o cartesianas *de U.*

Observación 2.172. Obsérvese que en esta definición se asume tácitamente que las $n - m$ ecuaciones son independientes (esto es, el rango de A es $n - m$).

La técnica empleada para demostrar la proposición 2.170 sirve directamente para obtener unas ecuaciones cartesianas de U. Veamos a continuación otro método alternativo que esquiva esa proposición (usa sólo que el rango de una matriz es igual al de su traspuesta) y puede servir para entender mejor el proceso; ambos métodos se aplican después a un ejemplo concreto.

Sea $\mathcal{B}_U = \{v_1,\dots,v_m\}$ una base de U. Consideremos la matriz en $M_{n\times m}(K)$ dada por

$$C = \big((v_1)_{\mathcal{B}_u}\,|\,(v_2)_{\mathcal{B}_u}\,|\,\dots\,|\,(v_m)_{\mathcal{B}_u}\big).$$

Sea $v = (x_1,\dots,x_n) \in K^n$. Como $\mathcal{B}_U$ es una base de U, deducimos (a partir de (2.10)) que $v \in U$ si y sólo si el SEL con matriz ampliada $(C|x)$ es compatible, siendo $x = (v)_{\mathcal{B}_u}$ el vector columna que se identifica con v. Por el teorema de Rouché-Frobenius, esto equivale a que $\mathrm{rg}(C|x) = \mathrm{rg}(C) = m$. Realizamos transformaciones elementales por filas en $(C|x)$ hasta conseguir la matriz de un SEL escalonado $(\tilde{C}|\tilde{x})$. Esta matriz tiene el mismo rango por filas (y, por tanto, por columnas) que $(C|x)$; por tanto, su rango m si y sólo si $v \in U$. No obstante por ser escalonado el SEL de $(\tilde{C}|\tilde{x})$, las filas $m+1,\dots,n$ de $\tilde{C}$ deben ser todas nulas, y el rango de $(\tilde{C}|\tilde{x})$ será m si y sólo si las últimas $n-m$ componentes del término independiente $\tilde{x}$ son 0. Por tanto, al igualar a cero estas componentes, se obtienen las $n-m$ ecuaciones lineales homogéneas (necesariamente independientes) buscadas.

Ejemplo 2.173. Calculamos unas ecuaciones cartesianas para la recta $U = L((1,3,-2))$ en $\mathbb{R}^3$. Claramente $\mathcal{B}_U = \{(1,3,-2)\}$ es una base de U. Así, sabemos que U tiene asociadas dos ecuaciones cartesianas. Las obtendremos de dos maneras distintas.

Método 1: Cada ecuación cartesiana es del tipo $ax+by+cz=0$. 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_0 tiene dimensión 2. Una base de U_0 viene dada por $\mathcal{B}_0 = \{(-3,1,0),(2,0,1)\}$. Así, unas ecuaciones cartesianas para U son $-3x+y=0$ y $2x+z=0$.

Método 2: Un vector $v=(x,y,z)$ de $\mathbb{R}^3$ está en U si es proporcional a $(1,3,-2)$. Esto equivale a que la matriz dada por:

$$(C|v) = \left(\begin{array}{r|c} 1 & x \\ 3 & y \\ -2 & z \end{array}\right)$$

cumpla que $\mathrm{rg}(C|v) = 1$. Al realizar transformaciones por filas se llega a:

$$(C|v) \sim \left(\begin{array}{c|c} 1 & x \\ 0 & -3x+y \\ 0 & 2x+z \end{array}\right)$$

Igualando a 0 las filas segunda y tercera obtenemos para U las mismas ecuaciones de antes.

Observación 2.174. Cuando estudiamos la posibilidad de obtener una base del s.v. $L(S)$ generado por un conjunto finito de vectores $S \subset K^n$, vimos que el procedimiento se extendía para cuando S está incluido en cualquier e.v. finitamente generado $V(K)$ (observación 2.152): bastaba con fijar una base $\mathcal{B}$ de $V(K)$ y trabajar con las coordenadas de los vectores en K^n. Análogamente, ahora los conceptos y modo de construcción de ecuaciones paramétricas e implícitas se extienden para cualquier s.v U de un e.v. f.g $V(K)$ sin más que fijar una tal base[15] $\mathcal{B}$. Como antes, la clave está en el isomorfismo $f_{\mathcal{B}} : V \to K^n$, que establece una biyección entre los subespacios vectoriales (resp. s.d.g.; conjuntos l.i.) de $V(K)$ y de $K^n(K)$ (observación 2.137).

[15]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 $\mathcal{B}_u$, como hemos hecho en esta sección.

2.4.6. Apéndice: cálculo de la matriz inversa por el método de Gauss-Jordan

Aunque más adelante se verá otro método de cálculo de la matriz inversa, conviene por su simplicidad conceptual estudiar el problema por el procedimiento de Gauss-Jordan.

Sea $A \in M_n(K)$ una matriz regular. Calcular su inversa equivale a solucionar la ecuación de n^2 incógnitas (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 columnsas 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:

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

Este sistema es compatible y determinado (por ser A regular tiene rango n y, por tanto, la matriz del sistema tiene rango n^2). 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 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:

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