

Curso de Álgebra Lineal (2022-I)

Diego Alberto Barceló Nieves
Facultad de Ciencias
Universidad Nacional Autónoma de México

Notación

$\mathbf{u}, \mathbf{v}, \mathbf{w}, \dots$	vectores (elementos de un conjunto vectorial V)
a, b, c, \dots	escalares (elementos de un campo K que define un espacio vectorial)
ab	producto entre los escalares a y b
$a\mathbf{u}$	producto del vector \mathbf{u} por el escalar a^a
(x_1, \dots, x_n)	coordenada como n-tupla
$\begin{pmatrix} x_1 & \dots & x_n \end{pmatrix}$	vector como n-tupla
$V + W$	suma de los espacios vectoriales V y W
$V \oplus W$	suma directa de los espacios vectoriales V y W
$\langle \mathbf{u}, \mathbf{v} \rangle$	producto escalar entre los vectores \mathbf{u} y \mathbf{v}
\bar{a}	complejo conjugado de a
$u_i v_i \equiv \sum_{i=1}^n u_i v_i$	<i>notación de Einstein</i> para la suma sobre un índice i
$\ \mathbf{u}\ $	norma del vector \mathbf{u}
$P_{\mathbf{u}}(\mathbf{v})$	proyección del vector \mathbf{v} sobre el vector \mathbf{u}
$\mathbf{u} \perp \mathbf{v}$	ortogonalidad de los vectores \mathbf{u} y \mathbf{v}
$\mathbf{a} \times \mathbf{b}$	producto vectorial (cruz) de dos vectores $\mathbf{a}, \mathbf{b} \in \mathbb{R}^3$
$\mathbf{a} \cdot \mathbf{b} \times \mathbf{c}$	triple producto escalar entre tres vectores $\mathbf{a}, \mathbf{b}, \mathbf{c} \in \mathbb{R}^3$.
$\langle G \rangle$	Espacio vectorial generado por G
<i>l.i.</i>	Conjunto linealmente independiente
<i>l.d.</i>	Conjunto linealmente dependiente
$\dim(V)$	Dimensión del espacio vectorial V

^aAlgunos textos se refieren a esta operación —realizada entre un vector y un escalar, y que da como resultado un vector— como *multiplicación escalar* (o *scalar multiplication*, en inglés); sin embargo, es fácil que esta operación se confunda con la de *producto escalar*, que da como resultado un escalar. Debemos tener esto en mente cuando leamos otros textos de álgebra lineal, tanto en español como en inglés.

1. Estructuras algebraicas, campos y espacios vectoriales

El álgebra lineal se puede definir como el estudio de los espacios vectoriales. En esta sección definiremos qué son, así como algunas nociones básicas que guiarán nuestro estudio de este tipo de estructuras algebraicas durante el curso. Antes, discutiremos brevemente qué constituye una estructura algebraica y repasaremos la definición de la estructura de campo —necesaria para definir la de espacio vectorial.

Estructuras algebraicas

¿Qué es el álgebra, y qué es lo que estudia? A menudo en los primeros cursos de álgebra esta cuestión no queda clara. El hecho de que esta área de las matemáticas tenga una historia de evolución que comenzó hace miles de años y continúa hasta hoy en día complica aún más la situación. A pesar de no ser la visión más general que existe, en este curso tomaremos por respuesta que el álgebra es el estudio de estructuras algebraicas.

Definición de estructura algebraica y propiedades de sus operaciones

Antes de definir una estructura algebraica, recordaremos una definición importante.

Def. Un *par ordenado* de elementos de dos conjuntos A y B es un par de elementos (a, b) , con $a \in A$ y $b \in B$, en donde el orden de los elementos importa. Dos pares ordenados $(a, b), (a', b')$ son iguales si, y sólo si, tienen los mismos elementos en el mismo orden; matemáticamente

$$(a, b) = (a', b') \iff a = a' \wedge b = b'.$$

El conjunto de pares ordenados de elementos de dos conjuntos A y B es el producto cartesiano

$$A \times B := \{(a, b) \mid a \in A, b \in B\}.$$

En particular, el conjunto $A \times A$ de pares ordenados de elementos de A se denota por A^2 . Inductivamente, para todo $n \geq 2$ podemos definir el conjunto de n -tuplas de elementos de A como $A^n := A^{n-1} \times A$, el cual identificamos con el conjunto

$$\{(a_1, a_2, \dots, a_n) \mid a_1, a_2, \dots, a_n \in A\}$$

para simplificar la notación.

Def. Una *estructura algebraica* es un conjunto no vacío A con (al menos) una operación en A y una colección (posiblemente vacía) de relaciones en A . Denotamos a una estructura algebraica formada por un conjunto A , operaciones \star, \circ y una relación \sim , como (A, \star, \circ, \sim) . Decimos que \star es una operación *binaria* si toma pares ordenados de elementos de A y devuelve elementos de A , i.e., si su dominio es $A \times A$ y su contradominio es A , lo que denotamos como $\star : A \times A \rightarrow A$. Si \star es una operación binaria, decimos que:

- \star es *asociativa* si, para cualesquiera $a, b, c \in A$, se tiene que

$$(a \star b) \star c = a \star (b \star c);$$

- \star es *conmutativa* si, para cualesquiera $a, b \in A$, se tiene que

$$a \star b = b \star a;$$

- \star tiene un elemento *identidad* (o *neutro*) si existe $e \in A$ tal que, para todo $a \in A$,

$$e \star a = a = a \star e;$$

- $b \in A$ es el elemento *inverso* de $a \in A$ (bajo \star) si \star tiene un elemento identidad $e \in A$ y

$$a \star b = e = b \star a;$$

- $B \subseteq A$ es *cerrado bajo la operación binaria* \star si la restricción de \star a $B \times B$ es binaria, i.e., si para cualesquiera $x, y \in B$,

$$x \star y \in B;$$

- \star se *distribuye con respecto a* $*$ si $*$ también es binaria y, para cualesquiera $a, b, c \in A$,

$$a \star (b * c) = (a \star b) * (a \star c) \quad \& \quad (b * c) \star a = (b \star a) * (c \star a).$$

Observación 1.1 Por definición, la relación de “ser inverso bajo una operación” es simétrica. Es decir, a es inverso de b bajo la operación \star si, y sólo si, b es inverso de a bajo \star . Más aún, por definición, el elemento identidad de una operación, si existe, siempre es su propio inverso bajo esa operación. Frecuentemente diremos simplemente *identidad* o *neutro* para referirnos al elemento identidad de una operación; así mismo, utilizaremos sólo la palabra *estructura* para referirnos a una estructura algebraica.

Ejemplos de estructuras algebraicas

Para cualquier conjunto arbitrario A , su conjunto potencia $\mathcal{P}(A)$ junto con las operaciones de unión e intersección de conjuntos (\cup y \cap , respectivamente) y la relación “contención” (\subseteq) forma una estructura algebraica, que podemos escribir explícitamente como $(\mathcal{P}(A), \cup, \cap, \subseteq)$. Observemos que ambas operaciones son binarias, asociativas y conmutativas, como seguramente viste en tu curso de Álgebra.

Ejercicio 1. Consideremos la estructura algebraica $(\mathcal{P}(A), \cup, \cap, \subseteq)$, donde A es un conjunto arbitrario. ¿Quiénes son los elementos neutros de \cup y \cap ? Demuéstralo.

El conjunto de los números naturales (\mathbb{N}) junto con la operación de suma ($+$) y la relación “menor o igual que” (\leq) forma una estructura algebraica $(\mathbb{N}, +, \leq)$. Observemos que, en este caso, $+$ es una operación binaria, asociativa y conmutativa. Si incluimos al número 0 en el conjunto \mathbb{N} , entonces 0 es el elemento identidad de la suma, y es el único elemento de \mathbb{N} que tiene inverso (¿Por qué?).

Ejercicio 2. Consideremos la estructura algebraica $(\mathbb{N}, +, \leq)$, donde $0 \in \mathbb{N}$. Demuestra que $\{0\}$ es el único subconjunto finito \mathbb{N} cerrado bajo la suma.

El conjunto de los números enteros junto con las operaciones de suma y multiplicación forma una estructura algebraica $(\mathbb{Z}, +, \cdot)$. Similarmente, los conjuntos de los números racionales y los reales con las mismas operaciones —entre números racionales y reales, respectivamente— forman las estructuras $(\mathbb{Q}, +, \cdot)$ y $(\mathbb{R}, +, \cdot)$. Claramente, estas tres estructuras son distintas, pues los conjuntos que los forman son distintos. Sin embargo, $(\mathbb{Q}, +, \cdot)$ y $(\mathbb{R}, +, \cdot)$ forman el mismo *tipo* de estructura algebraica puesto que, como veremos más adelante, sus operaciones cumplen las mismas propiedades. Estudiaremos este tipo de estructura, conocida como *campo*, en la sección 1 como un primer paso hacia la definición de otro tipo de estructura conocida como *espacio vectorial*, que definiremos en 1. Además, veremos que no se necesitan relaciones para definir a estos tipos de estructuras. Por lo tanto, de ahora en adelante asumiremos que nuestras estructuras algebraicas no tienen relaciones.

Antes de definir estos tipos de estructuras algebraicas, veremos algunas nociones generales de estructuras que usaremos tanto en campos como en espacios vectoriales.

Proposición 1.2 Sean (A, \star) una estructura algebraica con una operación binaria y $e \in A$ un elemento identidad de \star . Entonces

- (a) e es único;
- (b) si la operación \star es asociativa y $a \in A$ tiene un inverso b bajo \star , entonces b es único.
- (c) si la operación \star es asociativa y todos los elementos de A tienen inversos bajo \star entonces, para cualesquiera $a, b, c \in A$, se tiene que $a \star c = b \star c$ si, y sólo si, $a = b$. (Ley de Cancelación)

Demostración.

- (a) Supongamos que $e' \in A$ es un elemento identidad de \star . Entonces $e = e \star e' = e'$. Por lo tanto, el elemento identidad e de \star es único.
- (b) Supongamos que \star es asociativa, $a \in A$ y que existe $b \in A$ tal que b es inverso de a bajo \star . Adicionalmente, supongamos que $b' \in A$ es un inverso de a bajo \star . Entonces

$$\begin{aligned}
 b &= b \star e && (e \text{ es neutro}) \\
 &= b \star (a \star b') && (b' \text{ es inverso de } a) \\
 &= (b \star a) \star b' && (\star \text{ es asociativa}) \\
 &= e \star b' && (b \text{ es inverso de } a) \\
 &= b'.
 \end{aligned}$$

Por lo tanto, los elementos inversos bajo \star , si existen, son únicos.

- (c) Supongamos que \star es asociativa y que todos los elementos de A tienen inversos bajo \star . Sean $a, b, c \in A$ tales que $a \star c = b \star c$ y sea $c' \in A$ el elemento inverso de c bajo \star . Entonces,

$$\begin{aligned}
 a \star c = b \star c &\iff (a \star c) \star c' = (b \star c) \star c' \\
 &\iff a \star (c \star c') = b \star (c \star c') \\
 &\iff a \star e = b \star e \\
 &\iff a = b.
 \end{aligned}$$

Más aún, haciendo un procedimiento análogo, podemos demostrar que $c \star a = c \star b$ si, y sólo si, $a = b$.

□

Campos

Un campo es un tipo de estructura algebraica que formaliza varias de las nociones intuitivas que adquirimos durante nuestra educación básica sobre la aritmética en los números reales; esto es, que la suma y la multiplicación son operaciones binarias, asociativas y conmutativas, que la multiplicación se distribuye con respecto a la suma, y que el 0 y el 1 son números “especiales” en cierto sentido, el cual precisaremos más adelante.

Es probable que hayas visto este tipo de estructura explícitamente en tu curso de Álgebra y/o implícitamente en tu curso de Cálculo I (a través de los *axiomas de campo* —o *de cuerpo*¹— para los números reales); sin embargo, a continuación mencionaremos su definición y estudiaremos los dos ejemplos de campos que más utilizaremos durante este curso (el campo real y el complejo).

Definición de campo

Def. Un *campo* es un conjunto, que suele denotarse por K , con dos operaciones binarias llamadas *suma* y *multiplicación*, denotadas por $+$ y \cdot , respectivamente, tales que cumplen las siguientes propiedades, conocidas como los *axiomas de campo* o *de cuerpo*:

$$\forall a, b, c \in K \quad a + (b + c) = (a + b) + c \quad \& \quad a \cdot (b \cdot c) = (a \cdot b) \cdot c \quad \text{Asociatividad}$$

$$\forall a, b \in K \quad a + b = b + a \quad \& \quad a \cdot b = b \cdot a \quad \text{Conmutatividad}$$

$$\exists 0, 1 \in K \text{ t.q.}, \forall a \in K, \quad a + 0 = a \quad \& \quad 1 \cdot a = a \quad \text{Identidades (neutros)}^a$$

$$\forall a \in K \quad \exists -a \in K \quad \text{t.q.} \quad a + (-a) = 0 \quad \text{Inversos aditivos}^b$$

$$\forall a \neq 0 \in K \quad \exists a^{-1} \in K \quad \text{t.q.} \quad a \cdot a^{-1} = 1 \quad \text{Inversos multiplicativos}^c$$

$$\forall a, b, c \in K \quad a \cdot (b + c) = a \cdot b + a \cdot c \quad \text{Distributividad}^d.$$

^aPor el inciso (a) de la Proposición 1.2, sabemos que los elementos identidad de operaciones binarias en estructuras algebraicas son únicos. Por convención, se suele denotar al neutro aditivo de un campo como 0 y al neutro multiplicativo, como 1. Además, como en este caso las operaciones son conmutativas, podemos escribir esta propiedad de manera más sucinta.

^bPor el inciso (c) de la Proposición 1.2, sabemos que los inversos, si existen, son únicos. Por convención, si a es elemento de un campo, su inverso aditivo suele denotarse por $-a$.

^cPor convención, si a es un elemento de un campo distinto del neutro aditivo, su inverso multiplicativo suele denotarse por a^{-1} o $\frac{1}{a}$.

^dObservemos que podemos escribir esta propiedad de forma resumida, pues $+$ y \cdot son operaciones conmutativas.

Observación 1.3 En otras palabras, una estructura algebraica $(K, +, \cdot)$ es un campo si:

- (a) $+$ es una operación binaria, asociativa, conmutativa, con un elemento identidad y con inversos aditivos para **todos sus elementos**;
- (b) \cdot es una operación binaria, asociativa, conmutativa, con un elemento identidad y con inversos multiplicativos para **todos sus elementos excepto el neutro aditivo**;

¹En francés, este tipo de estructura es llamado *corps*, cuya traducción directa al español es “cuerpo”.

(c) \cdot se distribuye con respecto a $+$.

Nótese la asimetría en la propiedad de existencia de inversos aditivos en ambas operaciones: el neutro aditivo **no** requiere tener inverso multiplicativo, mientras que **todos** los elementos deben tener inversos aditivos. Frecuentemente escribiremos sólo K para referirnos al campo $(K, +, \cdot)$ y escribiremos la multiplicación implícitamente, omitiendo el símbolo “ \cdot ”.

Ejemplos de campos

El campo real

El conjunto de los números reales \mathbb{R} junto con las operaciones de suma y multiplicación (que aprendimos desde la educación básica) cumplen todas las propiedades enlistadas en la sección 1, por lo que forman un campo $(\mathbb{R}, +, \cdot)$, conocido como el *campo real*. Este campo puede ser representado geométricamente con la recta real, como en la Figura 1.

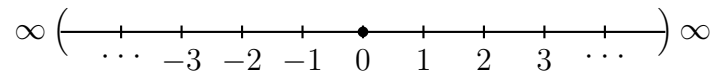


Figura 1: Representación del campo real con la recta real.

El campo complejo

Def. El conjunto de los números complejos se define como

$$\mathbb{C} := \{a + ib \mid a, b \in \mathbb{R}\}, \quad i := +\sqrt{-1}.$$

Sea $z = a + ib$ un número complejo. Decimos que a es su *parte real* y que b es su *parte imaginaria*. A los números complejos con parte real nula, i.e., aquellos de la forma $0 + ib, b \in \mathbb{R}$, se les conoce como números *imaginarios*^a. El *complejo conjugado* de un número complejo de $z = a + ib$ es

$$\bar{z} = a - ib,$$

y el valor absoluto o *módulo* de un número complejo z es

$$|z| = +\sqrt{z\bar{z}}.$$

^aGauss prefería llamarles números *laterales*, ya que creía que era un nombre más intuitivo, y que llamarles *imaginarios* les dotaba de una opacidad misteriosa e innecesaria. Sugiero ver este video introductorio (o la serie completa, llamada *Imaginary Numbers Are Real*) para perderles el miedo a los números imaginarios: <https://www.youtube.com/watch?v=T647CGsu0VU>.

Apoyándonos en las operaciones del campo real, podemos definir la suma y multiplicación entre números complejos como².

$$\begin{aligned}(a + ib) + (q + ir) &:= (a + q) + i(b + r), \\ (a + ib)(q + ir) &:= (aq - br) + i(ar + bq).\end{aligned}$$

Así mismo, apoyándonos en el campo real, podemos comprobar que el conjunto \mathbb{C} junto con estas dos operaciones forma un campo.

²Nótese que la definición de multiplicación es igual al desarrollo de $(a + ib)(c + id)$ como producto de binomios.

Ejercicio 3. Prueba que \mathbb{C} , con las operaciones de suma y multiplicación definidas anteriormente, forma un campo.

Subcampos

Def. Sea $(K, +, \cdot)$ un campo. Si $S \subseteq K$ es tal que las operaciones de suma y multiplicación en K restringidas al dominio $S \times S$ forman un campo, decimos que $(S, +, \cdot)$ es un *subcampo* de $(K, +, \cdot)$.

Observación 1.4

- (1) Todo campo K es trivialmente un subcampo de sí mismo.
- (2) Si hacemos una identificación entre el conjunto de los números reales y los números complejos de la forma $a + i0 \in \mathbb{C}, a \in \mathbb{R}$, entonces podemos considerar a $(\mathbb{R}, +, \cdot)$ como un subcampo de $(\mathbb{C}, +, \cdot)$. En efecto, si restringimos las operaciones de suma y multiplicación en el campo complejo al conjunto $\{z \in \mathbb{C} \mid z = a + i0, a \in \mathbb{R}\}$, entonces estas operaciones son idénticas a las operaciones de suma y multiplicación en el campo real.

Prácticamente todas las operaciones que realizamos cotidianamente como calcular fechas, dar cambio, aproximar áreas, repartir comida, etc., toman lugar en un campo. Es decir, las ideas intuitivas que nos formamos durante la educación básica de que la suma siempre debe ser conmutativa y asociativa —al igual que la multiplicación—, que existe la resta y la división, que el 0 y el 1 son números *especiales* en cierto sentido y que siempre se cumple la propiedad de distributividad, son un *hecho* para cualquier estructura de campo. Sin embargo, estas mismas ideas intuitivas *no siempre se cumplen en otros tipos de estructuras algebraicas* —algunas de las cuales veremos más adelante—, ¡así que no te confíes!

Espacios vectoriales

El conjunto de soluciones de ciertos tipos de problemas que aparecen frecuentemente en varias áreas de las matemáticas, la física, la computación y la biomedicina tienen estructura de espacio vectorial, por lo que el conocimiento de la teoría de los espacios vectoriales —es decir, del álgebra lineal— se puede traducir en aplicaciones directas en dichas áreas.

Definición de espacio vectorial

Def. Un *espacio vectorial* sobre un **campo** K es un conjunto V con dos operaciones (llamadas *adición* o *suma vectorial* y *producto de un vector por un escalar*) que satisfacen las siguientes propiedades, conocidas como los *axiomas de los espacios vectoriales*:

$$\forall \mathbf{u}, \mathbf{v} \in V \quad \exists \mathbf{u} + \mathbf{v} \in V \quad \text{Cerradura de la adición}$$

$$\forall \mathbf{v} \in V, a \in K \quad \exists a\mathbf{v} \in V \quad \text{Cerradura del producto de un vector por un escalar}$$

$$\forall \mathbf{u}, \mathbf{v}, \mathbf{w} \in V \quad \mathbf{u} + (\mathbf{v} + \mathbf{w}) = (\mathbf{u} + \mathbf{v}) + \mathbf{w} \quad \text{Asociatividad de la adición}$$

$$\forall \mathbf{u}, \mathbf{v} \in V \quad \mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u} \quad \text{Conmutatividad de la adición}$$

$$\exists \mathbf{0} \in V \text{ t.q. } \mathbf{v} + \mathbf{0} = \mathbf{v} \quad \forall \mathbf{v} \in V \quad \text{Elemento identidad de la adición (neutro aditivo)}$$

$$\forall \mathbf{v} \in V \quad \exists -\mathbf{v} \in V \text{ t.q. } \mathbf{v} + (-\mathbf{v}) = \mathbf{0} \quad \text{Elemento inverso de la adición (inverso aditivo)}$$

$$a(b\mathbf{v}) = (ab)\mathbf{v} \quad \forall a, b \in K, \mathbf{v} \in V \quad \text{Compatibilidad del producto de un vector por un escalar con el producto entre escalares}$$

$$\exists 1 \in K \quad \text{t.q. } 1\mathbf{v} = \mathbf{v} \quad \forall \mathbf{v} \in V \quad \text{Elemento identidad del producto de un vector por un escalar}$$

$$a(\mathbf{v} + \mathbf{w}) = a\mathbf{v} + a\mathbf{w} \quad \forall \mathbf{v}, \mathbf{w} \in V, a \in K \quad \text{Distributividad del producto de un vector por un escalar con respecto a la adición vectorial}$$

$$(a + b)\mathbf{v} = a\mathbf{v} + b\mathbf{v} \quad \forall a, b \in K, \mathbf{v} \in V \quad \text{Distributividad del producto de un vector por un escalar con respecto a la suma escalar.}$$

A los elementos $a, b \in K$ del campo utilizado para definir el espacio vectorial se les llama *escalares*, a los elementos $\mathbf{u}, \mathbf{v}, \mathbf{w} \in V$ que cumplen todas las propiedades anteriores se les llama *vectores*, y el conjunto V es llamado *conjunto vectorial*.

Observación 1.5

- (1) La definición matemática de *vectores* como *elementos cualesquiera de un conjunto V que, junto con un campo K y las operaciones $+: V \times V \rightarrow V$ (suma vectorial) y $\cdot: K \times V \rightarrow V$ (producto de un vector por un escalar), cumplen las propiedades de un espacio vectorial* es muy distinta a la definición de vector como *elemento con magnitud, dirección y sentido* (y, más precisamente, que

además es invariante bajo rotaciones propias e impropias) utilizada en algunas áreas de la física, siendo la primera definición más general.

- (2) La definición de *espacio vectorial* incluye dos operaciones *nuevas* (con respecto a las operaciones de campo) con una importante diferencia entre ellas: una es solamente entre los elementos del conjunto V (suma vectorial) y, la otra, entre los elementos del conjunto V y el campo K (producto de un vector por un escalar). Sin embargo, *ambas dan como resultado un vector en V* .
- (3) Así como la definición de *campo* incluye un conjunto K con dos operaciones (suma y producto) entre sus elementos que cumplen propiedades específicas, la definición de *espacio vectorial* incluye un conjunto V y un campo K con dos operaciones (suma vectorial y producto de un vector por un escalar) entre sus elementos que cumplen propiedades específicas.

Denotaremos a un espacio vectorial formado por un conjunto de vectores V y un campo K como (V, K) , o simplemente V cuando sea claro sobre qué campo está definido. Más adelante veremos otras operaciones en espacios vectoriales que se pueden definir entre vectores y escalares; sin embargo, la suma vectorial y el producto de un vector por un escalar son las únicas necesarias para *definir* a los espacios vectoriales, por lo que nos referiremos a ellas como las operaciones *esenciales* de los espacios vectoriales.

Corolario 1.6 (Ley de Cancelación para espacios vectoriales) Sean V un espacio vectorial y $\mathbf{x}, \mathbf{y}, \mathbf{z} \in V$ tales que $\mathbf{x} + \mathbf{z} = \mathbf{y} + \mathbf{z}$. Entonces, $\mathbf{x} = \mathbf{y}$.

Demostración. Se sigue la Ley de Cancelación (inciso (c) de la Proposición 1.2) ya que, por definición de espacio vectorial, la suma vectorial es asociativa y todos los elementos de V tienen elementos inversos bajo esta operación. Más aún, dado que la suma vectorial es conmutativa, se sigue que $\mathbf{x} + \mathbf{z} = \mathbf{z} + \mathbf{y}$ implica que $\mathbf{x} = \mathbf{y}$. \square

Para complementar la discusión al respecto de qué es un vector y apreciar cómo funcionan las operaciones de los espacios vectoriales (suma vectorial y producto de un vector por un escalar) de manera visual, sugiero ver el siguiente video: https://www.youtube.com/watch?v=fNk_zzaMoSs.

Ejemplos de espacios vectoriales

Los ejemplos de espacios vectoriales más sencillos —y, a menudo, más útiles— se siguen de una muy importante relación existente entre las definiciones de campo y espacio vectorial.

Ejercicio 4. Sea $(K, +, \cdot)$ un campo. Demuestra que (K, K) forma un espacio vectorial, con las operaciones de suma vectorial y producto de un vector por un escalar dadas por la suma y la multiplicación en el campo, respectivamente.

Por el Ejercicio 4 tenemos que cualquier campo forma un espacio vectorial sobre sí mismo. En particular, (\mathbb{R}, \mathbb{R}) y (\mathbb{C}, \mathbb{C}) son espacios vectoriales.

Consideremos el conjunto \mathbb{R}^2 . Aprovechando las operaciones de suma y multiplicación en el campo real, podemos definir una operación $+$: $\mathbb{R}^2 \times \mathbb{R}^2 \rightarrow \mathbb{R}^2$ como

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

y una operación \cdot : $\mathbb{R} \times \mathbb{R}^2 \rightarrow \mathbb{R}^2$ como

$$a(x_1, x_2) := (ax_1, ax_2).$$

Se puede verificar que, con estas operaciones, \mathbb{R}^2 forma un espacio vectorial sobre el campo real, en el cual el neutro aditivo es $(0, 0) \in \mathbb{R}^2$ y el inverso aditivo de $(x_1, x_2) \in \mathbb{R}^2$ es $(-x_1, -x_2) \in \mathbb{R}^2$. Esto se

debe a que definimos las nuevas operaciones “entrada por entrada”, por lo que la demostración de que $(\mathbb{R}^2, \mathbb{R})$ es un espacio vectorial es análoga a la del Ejercicio 4. Más aún, para cualquier entero positivo n podemos definir operaciones $+: \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}^n$ y $\cdot: \mathbb{R} \times \mathbb{R}^n \rightarrow \mathbb{R}^n$ de manera análoga a las anteriores, obteniendo así un espacio vectorial $(\mathbb{R}^n, \mathbb{R})$. Por ende, tenemos el siguiente corolario del Ejercicio 4.

Corolario 1.7 Sean K un campo y n un entero positivo. Entonces (K^n, K) es un espacio vectorial, definiendo las operaciones de suma vectorial y producto de un vector por un escalar entrada por entrada.

Por el Corolario 1.7, se sigue que $(\mathbb{C}^n, \mathbb{C})$ es un espacio vectorial para cualquier entero positivo n . Ahora, veamos otros ejemplos de espacios vectoriales.

El conjunto de todas las funciones polinomiales de una variable real de grado n (i.e., con regla de correspondencia de la forma $f(x) = c_1x^1 + c_2x^2 + \dots + c_nx^n$, $c_i \in \mathbb{R}$, $i \in \{1, 2, \dots, n\}$) y con un mismo dominio \mathcal{D} forma un espacio vectorial sobre el campo real. Aquí, las definiciones de suma vectorial y de producto de un vector por un escalar se siguen naturalmente de la definición de la suma de funciones $(f + g)(x) := f(x) + g(x)$ y del producto de una función arbitraria $f(x)$ por una función constante a , respectivamente, vistas en cálculo —las cuales aplican para las intersecciones de los dominios. El elemento identidad de la suma vectorial (neutro aditivo) es la función constante cero $f(x) = 0 \ \forall x \in \mathcal{D}$ y el inverso aditivo de una función $g(x)$ es $-g(x)$. Observemos que, en este caso, los *vectores* de nuestro espacio vectorial *son funciones* (en particular, en este ejemplo, son funciones polinomiales).

El conjunto de todas las funciones de una variable real derivables y con derivada continua (i.e., funciones de clase C^1) sobre el campo real forma un espacio vectorial³. Esto probablemente lo viste de manera implícita en tu curso de cálculo diferencial de una variable, cuando viste los teoremas de derivadas de una suma/multiplicación/división de funciones (también conocido como *álgebra de derivadas*) para funciones de este tipo. Las operaciones en este espacio vectorial, así como los elementos identidad (neutros) e inversos, se definen de la misma forma que en el ejemplo de las funciones polinomiales.

Para futura referencia, dejamos la siguiente definición.

Def. Un *espacio vectorial real (complejo)* es aquel definido sobre el campo real (complejo) o, equivalentemente, aquel donde los escalares son números reales (complejos).

Para ver más ejemplos de espacios vectoriales pueden consultar, por ejemplo, *Linear Algebra* de Friedberg (págs. 8-11.) o *Linear Algebra: A Modern Introduction* de Poole (págs. 430-432), entre otros.

Teorema 1.8 Sea (V, K) un espacio vectorial arbitrario con $\mathbf{v} \in V$ y $a \in K$. Entonces, tenemos que:

- (a) $0\mathbf{v} = \mathbf{0}$;
- (b) $a\mathbf{0} = \mathbf{0}$;
- (c) $(-a)\mathbf{v} = -(a\mathbf{v}) = a(-\mathbf{v})$.

Demostración.

³En general, el conjunto de funciones de clase C^n sobre el campo \mathbb{R} forma un espacio vectorial; aunque, estrictamente hablando, C^1 y C^n son *clases* y no conjuntos, en este curso ignoraremos este tecnicismo.

- (a) Por la propiedad de cerradura del producto de un vector por un escalar, sabemos que $0\mathbf{v} \in V$. Ahora, observemos que

$$\begin{aligned} 0\mathbf{v} + 0\mathbf{v} &= (0 + 0)\mathbf{v} && \text{(distributividad)} \\ &= 0\mathbf{v} \\ &= 0\mathbf{v} + \mathbf{0}, && \text{(neutro aditivo)} \end{aligned}$$

donde la primera igualdad se sigue de la distributividad de la suma de escalares con respecto al producto de un vector por un escalar. Por ende, tenemos que $0\mathbf{v} + 0\mathbf{v} = 0\mathbf{v} + \mathbf{0}$. Luego, de la Ley de Cancelación para espacios vectoriales (Corolario 1.6) se sigue que $0\mathbf{v} = \mathbf{0}$.

- (b) Observemos que

$$\begin{aligned} a\mathbf{0} + a\mathbf{0} &= a(\mathbf{0} + \mathbf{0}) \\ &= a\mathbf{0} \\ &= a\mathbf{0} + \mathbf{0}, \end{aligned}$$

donde la primera igualdad se sigue de la distributividad de la suma vectorial con respecto al producto de un vector por un escalar. Por ende, tenemos que $a\mathbf{0} + a\mathbf{0} = a\mathbf{0} + \mathbf{0}$. Luego, de la Ley de Cancelación para espacios vectoriales (Corolario 1.6) se sigue que $a\mathbf{0} = \mathbf{0}$.

- (c) Observemos que

$$\begin{aligned} \mathbf{0} &= 0\mathbf{v} \\ &= (a + (-a))\mathbf{v} \\ &= a\mathbf{v} + (-a)\mathbf{v}, \end{aligned}$$

donde se utilizó el inciso (a), la propiedad de existencia de inversos aditivos en un campo y la distributividad de la suma de escalares con respecto al producto de un vector por un escalar. De la ecuación $a\mathbf{v} + (-a)\mathbf{v} = \mathbf{0}$ y la conmutatividad de la suma vectorial se sigue que $(-a)\mathbf{v} = -(a\mathbf{v})$. Análogamente, tenemos que

$$\begin{aligned} \mathbf{0} &= a\mathbf{0} \\ &= a(\mathbf{v} + (-\mathbf{v})) \\ &= a\mathbf{v} + a(-\mathbf{v}), \end{aligned}$$

donde se utilizó el inciso (b), la propiedad de existencia de inversos aditivos en un espacio vectorial y la distributividad de el producto de un vector por un escalar con respecto a la suma vectorial. De la ecuación $a\mathbf{v} + a(-\mathbf{v}) = \mathbf{0}$ y la conmutatividad de la suma vectorial se sigue que $a(-\mathbf{v}) = -(a\mathbf{v})$. Por ende, $(-a)\mathbf{v} = -(a\mathbf{v}) = a(-\mathbf{v})$.

□

Interpretación geométrica de las operaciones esenciales de los espacios vectoriales

Como se mencionó en una nota al final de la sección 1 —de la cual retomaremos muchas ideas a continuación—, podemos desarrollar nuestra intuición sobre muchos temas del álgebra lineal trabajando en espacios vectoriales *visualizables*, para luego extenderla a espacios vectoriales más generales. Por ende, ahora haremos hincapié en la interpretación geométrica de las operaciones de suma vectorial y producto de un vector por un escalar en el espacio vectorial real \mathbb{R}^2 , así como en el espacio vectorial complejo \mathbb{C} .

El espacio vectorial real \mathbb{R}^2

En geometría analítica aprendimos que, con la ayuda de un sistema de coordenadas, podemos formar una correspondencia uno a uno (o *biunívoca*) entre pares ordenados (a, b) con entradas reales⁴ y puntos del plano cartesiano $\mathbb{R} \times \mathbb{R}$. En particular, si tomamos el sistema de coordenadas cartesianas, entonces a cualquier par ordenado de entradas reales (a, b) le corresponde un punto en el plano cartesiano $\mathbb{R} \times \mathbb{R}$ con coordenadas cartesianas (a, b) , y vice versa. En álgebra lineal, es preferible considerar a cada vector $(a, b) \in \mathbb{R}^2$ en una correspondencia biunívoca con la flecha en el plano cartesiano que tiene cola en el origen y punta en la coordenada cartesiana correspondiente al punto (a, b) , como se muestra en la Figura 2.

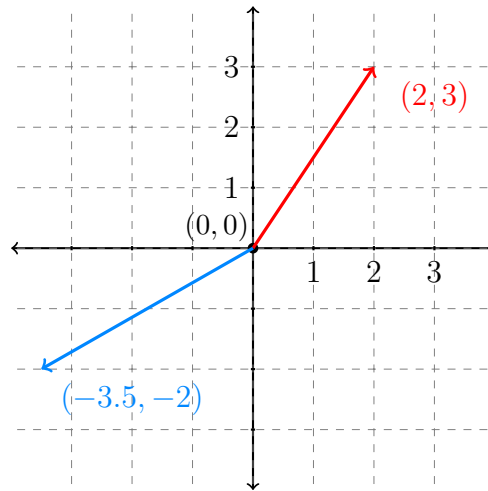


Figura 2: Ejemplo de representación de vectores en el plano cartesiano. Los vectores $(-3.5, -2)$ y $(2, 3)$ son representados por flechas que tienen su cola en el origen del plano cartesiano y su punta en las coordenadas correspondientes.

Suma vectorial

Recordemos que, en este espacio, la suma vectorial se define como $(a, b) + (c, d) = (a + c, b + d)$, i.e., entrada por entrada. Podemos calcular, por ejemplo, la suma $(2, 1) + (1, 3) = (3, 4)$. Los tres vectores mencionados se muestran en la Figura 3.

Observemos que, visualmente, esto corresponde a trazar uno de los vectores en el plano cartesiano y luego trazar el otro colocando la cola en la punta del vector anterior, como si ése fuese su origen. Nótese que no importa cuál vector trazamos primero y cuál después, lo cual concuerda con la conmutatividad de la suma vectorial (esta misma interpretación geométrica es válida para la suma de tres o más vectores de \mathbb{R}^2 : basta irlos sumando de dos en dos vectores); a esto se le conoce como la *Ley del paralelogramo*. Es decir, la suma de dos vectores $\mathbf{u}, \mathbf{v} \in \mathbb{R}^2$ es igual a la diagonal principal del paralelogramo formado por las flechas que los representan.

Ejercicio 5. Muestra intuitivamente que si transportamos a la *otra* diagonal del paralelogramo hasta el origen (lo cual se puede hacer de dos maneras distintas), obtenemos a las representaciones de los vectores $\mathbf{u} - \mathbf{v}$ y $\mathbf{v} - \mathbf{u}$.

Producto de un vector por un escalar

⁴Es decir, con $a, b \in \mathbb{R}$.

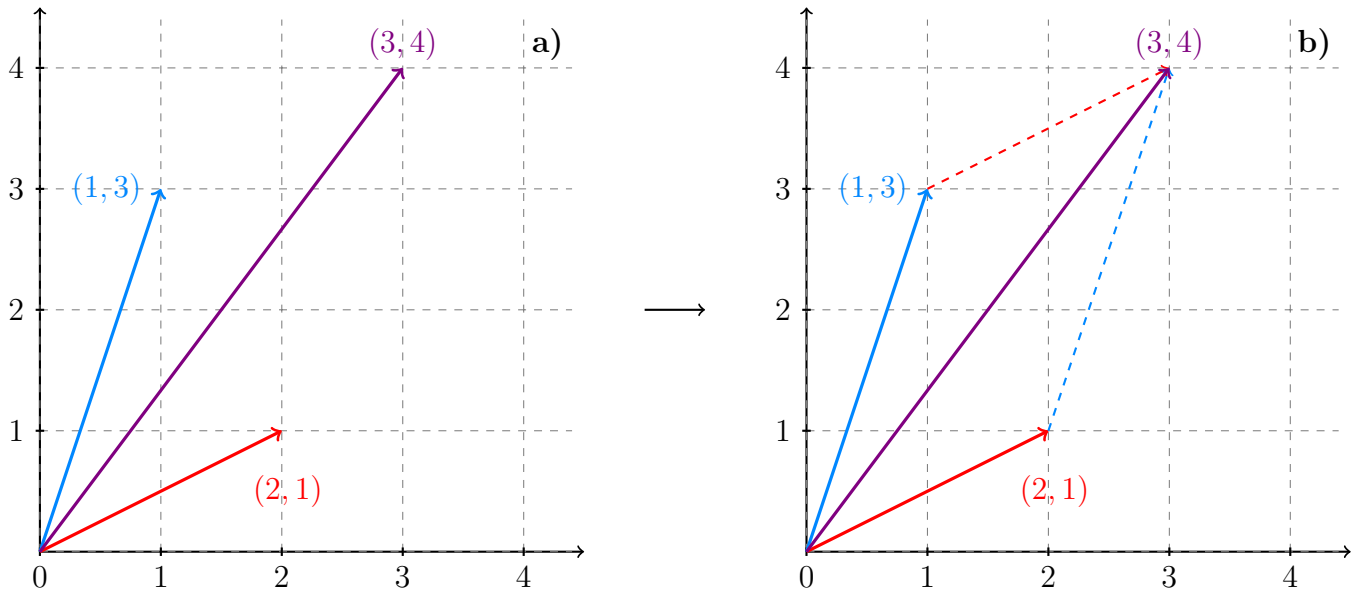


Figura 3: Interpretación geométrica de la suma vectorial en el espacio vectorial real \mathbb{R}^2 . En la subfigura **a)** se observan los vectores $(1, 3)$ y $(2, 1)$, así como el vector resultante de la suma de los dos anteriores, $(3, 4)$. En la subfigura **b)** observamos la llamada *Ley del paralelogramo* para la suma de dos vectores.

Ahora, recordemos que en este espacio el producto de un vector por un escalar se define como $c(a, b) = (ca, cb)$ (entrada por entrada). Podemos calcular, por ejemplo, los productos $\left(\frac{1}{2}\right)(2, 2) = (1, 1)$ y $(-1.2)(1, 3) = (-1.2, -3.6)$. La representación gráfica de estas operaciones se muestra en la Figura 4.

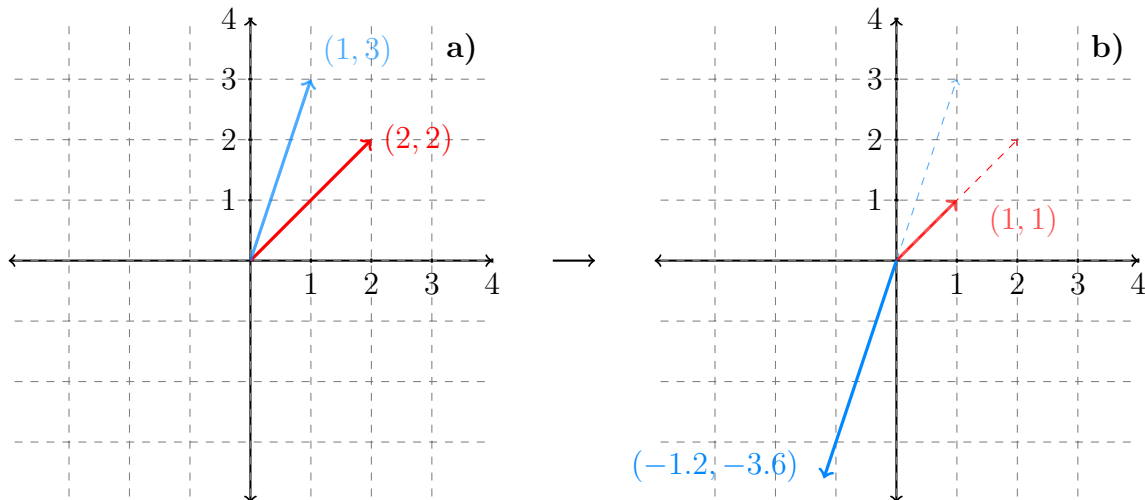


Figura 4: Interpretación geométrica del producto de un vector por un escalar en el espacio vectorial real \mathbb{R}^2 . Comparando las subfiguras **a)** y **b)** observamos que, en caso de que se multiplique a un vector de \mathbb{R}^2 por un escalar de \mathbb{R} , es posible que la longitud del vector cambie y que su sentido se invierta, pero su dirección no cambia.

Como podemos observar, el primer producto redujo la longitud del vector sin cambiar su sentido, mientras que el segundo producto aumentó la longitud del vector, a la vez que invirtió su sentido; sin embargo, en ambos casos, el producto de un vector por un escalar no cambió la *dirección* de los vectores—es decir, los mantuvo en la misma *línea*. En general, si el escalar $c \in \mathbb{R}$ que multiplica al vector tiene $|c| > 1$, lo *alarga*; si tiene $|c| < 1$, lo *acorta*; finalmente, si tiene $|c| = 1$, no cambia su longitud. Por este cambio de longitud es que al producto de un vector por un escalar también se le

conoce por el nombre *reescalamiento*. Además, si $c > 0$, el vector mantiene su misma dirección y sentido (sigue en la misma línea y apunta hacia el mismo lado) mientras que, si $c < 0$, el vector conserva su dirección pero se invierte su sentido (sigue en la misma línea pero apunta hacia el lado opuesto); si $c = 0$ entonces el vector automáticamente se convierte en el vector nulo $(0, 0)$, como se demostró algebraicamente en el inciso (a) del Teorema 1.8. Para visualizar las operaciones de adición vectorial y producto de un vector por un escalar de forma interactiva, recomiendo la sección **Vector Algebra and Geometry** de <https://textbooks.math.gatech.edu/ila/vectors.html>, así como la ilustración interactiva http://immersivemath.com/ila/ch02_vectors/ch02.html#fig_vec_scaling.

Así, en general, si combinamos las operaciones de suma vectorial y producto de un vector por un escalar, visualmente lo que estaremos haciendo será *combinar líneas* con diferentes longitudes, direcciones y sentidos en el plano cartesiano.

Nota: El vector nulo $\mathbf{0} = (0, 0)$ (también llamado *vector origen*) no tiene longitud, ya que es el único donde la cola y la punta de su flecha coinciden. Además, tampoco tiene dirección ni sentido⁵. Si asumimos que este vector no tiene longitud, dirección ni sentido, entonces queda claro por qué cualquier reescalamiento de este vector no lo modifica, como se demostró en el inciso (b) del Teorema 1.8.

Ejercicio 6. Interpreta geoméricamente las operaciones esenciales del espacio vectorial \mathbb{R}^3 .

En el espacio vectorial complejo \mathbb{C}

Como hemos visto, el plano cartesiano nos sirve para representar vectores con dos entradas reales. De manera similar, el *plano complejo* —con un eje de números *reales* (por convención, el horizontal) y otro eje perpendicular a él de números *imaginarios*⁶— nos sirve para representar vectores con una entrada compleja. Así, cada vector de una entrada compleja $(a + ib)$ con $a, b \in \mathbb{R}$ tiene una correspondencia uno a uno con una flecha con cola en el origen del plano y flecha en la coordenada (a, b) del plano complejo, la cual corresponde a, desde el origen, moverse a unidades sobre el eje real y b unidades sobre el eje imaginario.

Suma vectorial

De la definición de suma vectorial $(a + ib) + (c + id) := ((a + c) + (b + d)i)$ se deduce que la suma vectorial entre vectores de \mathbb{C} tiene la misma interpretación geométrica que aquella entre vectores de \mathbb{R}^2 . Por ejemplo, si calculamos $(1 + 2i) + (3 + 2i) = (4 + 4i)$, podemos representarlo visualmente en la Figura 5.

Producto de un vector por un escalar

Por definición, el producto de un vector por un escalar es $(q + ir)(s + it) := ((qs - rt) + i(qt + rs))$. Notemos que, en particular, si la parte imaginaria del escalar es nula (i.e., si $r = 0$), entonces el escalar es un número real y el producto resultante es $(q)(s + it) := ((qs) + (qt)i)$, por lo cual geoméricamente sólo se produce un reescalamiento totalmente análogo al discutido en el caso de \mathbb{R}^2 . En cambio, ahora observemos qué sucede si la parte real del escalar es nula y la parte imaginaria es igual a 1 (i.e., si multiplicamos por el escalar i). Tomemos, por ejemplo, al vector $(2 + 2i)$. Al hacer el producto de este vector

⁵Alternativamente, se dice que tiene *todas las direcciones y todos los sentidos simultáneamente*: en la práctica, ambas interpretaciones son equivalentes, pero la primera puede ser más fácil de asimilar.

⁶Los números imaginarios son aquellos números complejos con la parte real igual a cero, i.e. $0 + ib = ib \in \mathbb{C}$, donde b es un número real. En otras palabras, son el resultado de multiplicar el número imaginario i por cualquier número real.

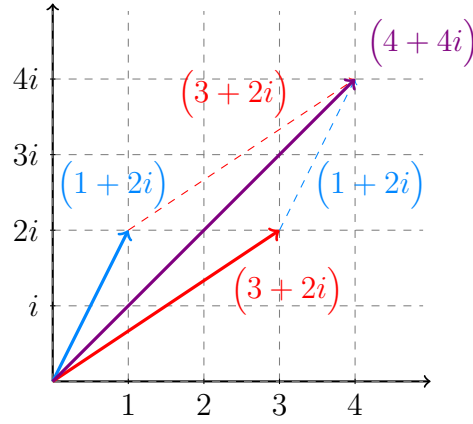


Figura 5: Interpretación geométrica de la suma vectorial en el espacio vectorial complejo \mathbb{C} . Observamos que, al igual que en el caso del espacio vectorial real \mathbb{R}^2 , se cumple la *Ley del paralelogramo*.

por i obtenemos $(-2 + 2i)$. Si, en cambio, hacemos el producto de este mismo vector por el escalar $-i$, obtenemos como resultado $(-i)(2 + 2i) = (2 - 2i)$. Ambas operaciones se muestran de manera visual en la Figura 6.

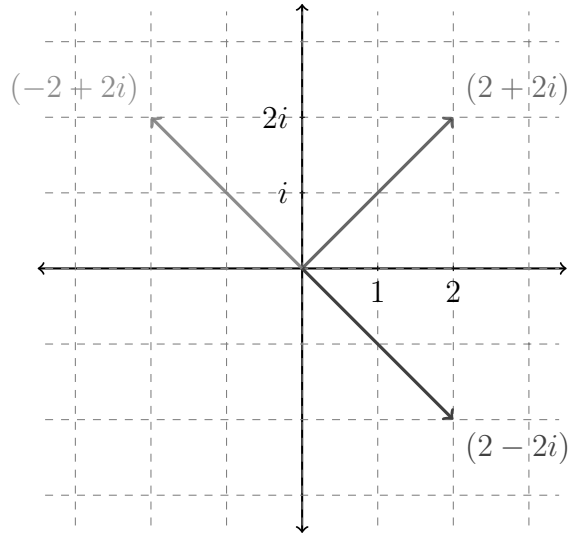


Figura 6: Interpretación geométrica del producto de un vector complejo por los números imaginarios i y $-i$. En este caso, nuestro vector base es $(2 + 2i)$. El producto de este vector por el escalar i resulta en el vector $(-2 + 2i)$, lo cual puede ser interpretado geoméricamente como una rotación discreta de $\frac{\pi}{2}$ radianes. Así observamos que, en cambio, el producto de nuestro vector base $(2 + 2i)$ por $-i$ se puede interpretar geoméricamente como una rotación discreta de $-\frac{\pi}{2}$ radianes.

Aquí vemos que hacer el producto de un vector por el escalar i equivale a hacer una rotación de 90° ó $\frac{\pi}{2}$ radianes. Análogamente, el producto de un vector por el escalar $-i$ equivale a hacer una rotación de -90° ó $-\frac{\pi}{2}$ radianes. Esto tiene sentido ya que $-i = -1(i) = i(-1)$ lo cual implica que, debido a la compatibilidad del producto de un vector por un escalar con el producto entre escalares, es lo mismo multiplicar un vector por $(-i)$ a multiplicarlo por i y después por -1 , o vice versa: el razonamiento geométrico correspondiente es que da lo mismo rotar un vector $-\frac{\pi}{2}$ radianes a rotarlo $\frac{\pi}{2}$ radianes y después invertir su sentido, o primero invertir su sentido y después rotarlo $\frac{\pi}{2}$ radianes.

¿Y si multiplicamos un vector de \mathbb{C} por un escalar ai con $a \neq 0, 1$? Ya que $ai(b + ic) = (-ac + i(ab)) = a(-c + ib) = a(i(b + ic))$ —es decir, por la compatibilidad entre productos— podemos deducir que hacer el producto de un vector complejo por un número imaginario arbitrario ai tendrá dos consecuencias: rotarlo de acuerdo a i ($\frac{\pi}{2}$ radianes a contrarreloj) y reescalarlo de acuerdo al valor de a (invirtiendo el sentido si $a < 0$). En este último caso, ya que $ai = |a|(-i) = (-i)|a| \quad \forall a < 0$, también podríamos pensar que se rota al vector complejo de acuerdo a $-i$ ($\frac{\pi}{2}$ radianes en el sentido de las manecillas) y se reescala de acuerdo al valor absoluto de a : ¡ambas interpretaciones son equivalentes!

Dicho lo anterior, estamos listos para el caso más general, el cual es fácil de entender de forma precisa recordando la representación polar de los números complejos. Sea $z = x + iy \in \mathbb{C}$. Entonces sus coordenadas cartesianas en el plano complejo son (x, y) , mientras que sus coordenadas polares son (r, θ) , donde $x = r \cos \theta, y = r \sin \theta$. Observemos que $r = +\sqrt{x^2 + y^2} = +\sqrt{(x + iy)(x - iy)} = +\sqrt{z\bar{z}} = |z|$, por lo que el *módulo* de z es igual a la longitud de la flecha que representa a z en el plano complejo. Por otro lado, θ es el ángulo que va de la parte positiva del eje real a la flecha que representa a z en el plano y se conoce como el *argumento* de z . Por ende, tenemos que

$$\begin{aligned} z &= x + iy \\ &= |z| \cos \theta + i|z| \sin \theta \\ &= |z|(\cos \theta + i \sin \theta) \end{aligned}$$

para todo $z \in \mathbb{C}$. Sean $z_1, z_2 \in \mathbb{C}$ con argumentos θ_1 y θ_2 , respectivamente. Entonces, por lo anterior,

$$\begin{aligned} z_1 z_2 &= |z_1|(\cos \theta_1 + i \sin \theta_1)|z_2|(\cos \theta_2 + i \sin \theta_2) \\ &= |z_1||z_2|(\cos \theta_1 \cos \theta_2 + i \cos \theta_1 \sin \theta_2 + i \sin \theta_1 \cos \theta_2 - \sin \theta_1 \sin \theta_2) \\ &= |z_1||z_2|(\cos \theta_1 \cos \theta_2 - \sin \theta_1 \sin \theta_2 + i(\cos \theta_1 \sin \theta_2 + \cos \theta_2 \sin \theta_1)). \end{aligned}$$

Luego, aplicando las identidades trigonométricas para el coseno y seno de una suma de ángulos

$$\begin{aligned} \cos(\alpha + \beta) &= \cos(\alpha) \cos(\beta) - \sin(\alpha) \sin(\beta), \\ \sin(\alpha + \beta) &= \cos \alpha \sin \beta + \cos \beta \sin \alpha, \end{aligned}$$

tenemos que

$$z_1 z_2 = |z_1||z_2|(\cos(\theta_1 + \theta_2) + i \sin(\theta_1 + \theta_2)).$$

Por lo tanto, el producto de dos números complejos se obtiene *multiplicando sus módulos y sumando sus argumentos*.

Ya que el producto de un vector por un escalar en (\mathbb{C}, \mathbb{C}) es igual al producto entre escalares complejos, concluimos que multiplicar un vector complejo $(s + it)$ por un escalar complejo $q + ir$ con $q, r \neq 0$ reescalará el vector $(s + it)$ en el plano complejo por el módulo de $q + ir$ y lo rotará de acuerdo al argumento de $q + ir$. En general, en los espacios vectoriales complejos los escalares no sólo pueden *reescalar* vectores, sino que también los pueden *rotar*.

2. Subespacios vectoriales, combinaciones lineales, conjunto generador y subespacio generado

Anteriormente, vimos que es posible tomar a un subconjunto de un campo de tal manera que las operaciones del campo, restringidas al subconjunto, formen un campo, y llamamos a esto un *subcampo*. Resulta que podemos hacer algo similar con espacios vectoriales, definiendo el concepto de *subespacio vectorial*. Más aún, podemos dar condiciones explícitas para que un subconjunto del conjunto vectorial de un espacio forme un subespacio vectorial de dicho espacio.

Definición de subespacio vectorial

Def. Sea (V, K) un espacio vectorial. Si S es un subcampo de K y $W \subseteq V$ es tal que las operaciones de suma vectorial y producto de un vector por un escalar en (V, K) restringidas a $W \times W$ y $S \times W$, respectivamente, forman un espacio vectorial, entonces decimos que (W, S) es un *subespacio vectorial* de (V, K) .

Observación 2.1 Si S es un subcampo de K , entonces (S, S) es un subespacio vectorial de (K, K) . En efecto: Esto se sigue del Ejercicio 4. En particular, de los ejemplos de ese ejercicio se sigue que (\mathbb{R}, \mathbb{R}) es un subespacio vectorial de (\mathbb{C}, \mathbb{C}) .

Si (V, K) es un espacio vectorial y $W \subseteq V$, podemos preguntarnos: ¿Qué condiciones necesita cumplir el subconjunto W para que forme un subespacio vectorial sobre K de (V, K) ? La respuesta está dada por el siguiente resultado, que nos da una caracterización muy útil de los subespacios vectoriales.

Proposición 2.2 (Caracterización de subespacios vectoriales) Sean (V, K) un espacio vectorial y $W \subseteq V$. Entonces las condiciones siguientes equivalen a que (W, K) sea un subespacio vectorial de (V, K) :

- (a) W es cerrado bajo la suma vectorial,
- (b) W es cerrado bajo el producto de un vector por un escalar, y
- (c) W contiene al vector nulo de V .

Demostración.

Supongamos que (W, K) es un subespacio vectorial de (V, K) . Entonces, (W, K) forma un espacio vectorial con las operaciones de suma vectorial y producto de un vector por un escalar en (V, K) restringidas a $W \times W$ y $K \times W$. Por definición de espacio vectorial, W es cerrado bajo la suma vectorial y el producto de un vector por un escalar en K , por lo que se cumplen (a) y (b), y W contiene a un vector nulo. Como $W \subseteq V$ y V , por definición de espacio vectorial, contiene a un vector nulo, entonces del inciso (c) de la Proposición 1.2, se sigue que el vector nulo de W es el mismo de V .

Por otro lado, supongamos que W cumple (a), (b) y (c). Puesto que (V, K) es un espacio vectorial y $W \subseteq V$, entonces de (a) y (b) se sigue directamente que la suma vectorial en V restringida a W es asociativa y conmutativa, el producto de un vector por un escalar restringido a $K \times W$ es compatible con el producto entre escalares, existe un elemento identidad del producto de un vector por un escalar $1 \in K$, y que el producto de un vector por un escalar restringido a W y K se distribuye con respecto a la suma vectorial y con respecto a la suma escalar. En particular, por la existencia de inversos aditivos en el campo K , la existencia del elemento $1 \in K$ y la distributividad del producto de un vector por un escalar con respecto a la suma escalar, se sigue que existen inversos aditivos para todos los elementos de W . Finalmente, por (c), existe un neutro aditivo en W . Por lo tanto, W es tal que las operaciones

de suma vectorial y producto de un vector por un escalar restringidas a $W \times W$ y $K \times W$ forman un espacio vectorial, de donde se sigue que (W, K) es un subespacio vectorial de (V, K) . \square

Observación 2.3 Las condiciones (a) y (b) de la Proposición 2.2 son equivalentes a decir que, si $\mathbf{u}, \mathbf{v} \in W$ y $a \in K$, entonces $\mathbf{u} + a\mathbf{v} \in W$.

Corolario 2.4 Sean (V, K) un espacio vectorial, $W \subseteq V$ y S es un subcampo de K . Entonces, las condiciones de la Proposición 2.2 —restringiendo el producto de un vector por un escalar a $S \times W$ — son equivalentes a que (W, S) sea un subespacio vectorial de (V, K) .

Demostración. Es totalmente análoga a la de demostración de la Proposición 2.2. \square

Observación 2.5

- (1) Como todo subespacio vectorial es, en particular, un espacio vectorial, entonces cualquier subespacio vectorial puede tener subespacios vectoriales subsecuentes, todos con el mismo vector nulo.
- (2) Para todo espacio vectorial V , V y $\{\mathbf{0}\}$ son trivialmente subespacios vectoriales de V .

Ejemplos de subespacios vectoriales

El conjunto de todos los pares ordenados $\{(x_1, x_2) \in \mathbb{R}^2 | x_1 = x_2\}$ es un subespacio vectorial del espacio vectorial real \mathbb{R}^2 .

Si $j, k \in \mathbb{N}$ son tales que $j < k$, entonces el conjunto de polinomios de grado j es un subespacio vectorial⁷ del espacio vectorial real de polinomios de grado k .

Si $n \in \mathbb{N}$, entonces el conjunto de todas las funciones reales de clase C^∞ es un subespacio vectorial de (C^n, \mathbb{R}) .

Intersección y suma de subespacios vectoriales

A continuación, presentamos algunas operaciones que podemos realizar entre subespacios vectoriales de un cierto espacio vectorial para obtener nuevos subespacios de dicho espacio. En realidad, estas operaciones se realizarán entre los *conjuntos vectoriales* de dichos subespacios, resultando en un conjunto de vectores que forma un subespacio vectorial sobre el mismo campo que el espacio vectorial original.

Teorema 2.6 Sea V un espacio vectorial. Entonces, cualquier intersección de dos subespacios vectoriales de V es un subespacio vectorial de V .

Demostración. Sea V sobre K un espacio vectorial y sean W_1, W_2 subespacios vectoriales de V . Por la Proposición 2.2, cada subespacio vectorial de V contiene al neutro aditivo de V , por lo que $\mathbf{0} \in W_1 \cap W_2$.

Sean $\mathbf{u}, \mathbf{v} \in W_1 \cap W_2$ y $a \in K$. Por la Observación 2.3, cada subespacio contiene a $\mathbf{u} + a\mathbf{v}$, por lo que $\mathbf{u} + a\mathbf{v} \in W_1 \cap W_2$. De lo anterior, concluimos que $W_1 \cap W_2$ es un subespacio vectorial de V . \square

Corolario 2.7 Cualquier intersección finita de subespacios vectoriales de un espacio vectorial V es un subespacio vectorial de V .

Demostración. Se sigue del Teorema 2.6 y de que la intersección de conjuntos es una operación asociativa. \square

⁷De aquí en adelante, asumiremos que cualquier espacio vectorial V está definido por un conjunto vectorial V sobre el campo real a menos que se indique lo contrario.

Ejercicio 7. Sea Z un subespacio vectorial de W y sea W , a su vez, subespacio vectorial de V . Demuestra que Z es un subespacio vectorial de V .

Def. Sean S_1 y S_2 subespacios de un espacio vectorial V . Definimos a la *suma de los subespacios vectoriales* S_1 y S_2 como

$$S_1 + S_2 = \{\mathbf{x} + \mathbf{y} \mid \mathbf{x} \in S_1, \mathbf{y} \in S_2\}.$$

Ejercicio 8. Demuestra que cualquier suma finita de subespacios vectoriales de un espacio vectorial V es un subespacio vectorial de V .

Combinaciones lineales, conjunto generador y subespacio generado

Ahora, veremos más formas de obtener subespacios vectoriales a partir de un espacio vectorial. Sabemos que las operaciones necesarias para definir a un espacio vectorial son la suma vectorial y el producto de un vector por un escalar. La operación más general que podemos realizar a partir de dichas operaciones se define a continuación.

Def. Sea (V, K) un espacio vectorial y $L = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\} \subseteq V$ un conjunto finito de vectores de V . Decimos que \mathbf{u} es una *combinación lineal* de los vectores de L si existen escalares $c_i \in K$ para $i \in \{1, 2, \dots, n\}$ tales que

$$\mathbf{u} = c_1 \mathbf{v}_1 + c_2 \mathbf{v}_2 + \dots + c_n \mathbf{v}_n = \sum_{i=1}^n c_i \mathbf{v}_i.$$

En este caso, decimos que los escalares c_i son los *coeficientes* de la combinación lineal $\sum_{i=1}^n c_i \mathbf{v}_i$.

Observación 2.8

- (1) Dadas las propiedades de cerradura de las operaciones esenciales de los espacios vectoriales, cualquier combinación lineal de vectores de un espacio vectorial V resultará en un vector de V .
- (2) El vector nulo de un espacio vectorial puede ser obtenido como combinación lineal de cualquier conjunto de vectores. Por el inciso (a) del Teorema 1.8, basta fijar a todos los coeficientes de la combinación lineal como el neutro aditivo del campo. A este tipo de combinación lineal se le conoce como *combinación lineal trivial*.
- (3) Siguiendo las interpretaciones geométricas de las operaciones de suma entre vectores y producto de un vector por un escalar vistas anteriormente, podemos interpretar a esta operación generalizada como la combinación de flechas (o líneas) reescaladas y posiblemente rotadas, si el espacio vectorial es complejo y el escalar tiene una parte imaginaria no nula, a las cuales aplicamos la Ley del paralelogramo para obtener una nueva flecha (o línea) como resultado. Precisamente por esta razón es que a esta operación general se le conoce como *combinación lineal*.

Ejercicio 9. Sea (V, K) un espacio vectorial y $L = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\} \subseteq V$ un conjunto finito de vectores de V . Demuestra que el conjunto de todas las combinaciones lineales de L

$$\{c_1 \mathbf{v}_1 + c_2 \mathbf{v}_2 + \dots + c_n \mathbf{v}_n \mid c_i \in K\}$$

es un subespacio vectorial de V .

En vista del Ejercicio 9, damos la siguiente definición.

Def. Sea V sobre K un espacio vectorial y $L \subseteq V$ finito. Entonces, definimos al *subespacio generado por L* como

$$\langle L \rangle := \{c_1 \mathbf{v}_1 + c_2 \mathbf{v}_2 + \dots + c_n \mathbf{v}_n \mid c_i \in K, \mathbf{v}_i \in L\}.$$

A L se le conoce como el *conjunto generador*. Por completez, definimos $\langle \emptyset \rangle = \{\mathbf{0}\}$.

Observación 2.9 A menudo denotaremos al subespacio generado por un conjunto de vectores $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\}$ simplemente como $\langle \mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n \rangle$ en vez de $\langle \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\} \rangle$, cuando esto no lleve a una confusión.

En un espacio vectorial arbitrario es posible expresar a cualquiera de sus vectores como combinación lineal de otros vectores del mismo espacio. Por ejemplo, en \mathbb{R}^2 el vector

$$\begin{aligned} \begin{pmatrix} 1 & 5 \end{pmatrix} &= \begin{pmatrix} 1 & 0 \end{pmatrix} + 5 \begin{pmatrix} 0 & 1 \end{pmatrix} = 2 \begin{pmatrix} 1 & 1.5 \end{pmatrix} + (-0.5) \begin{pmatrix} 2 & -4 \end{pmatrix} \\ &= (-4) \begin{pmatrix} 0.5 & -3 \end{pmatrix} + 3 \begin{pmatrix} 1 & 1 \end{pmatrix} + (-5) \begin{pmatrix} 0 & 2 \end{pmatrix} = \dots \end{aligned}$$

Observamos que, en cada caso, el valor de los coeficientes $c_i \in \mathbb{R}$ depende de los vectores $\mathbf{v}_i \in \mathbb{R}^2$ con los cuales se realiza la combinación lineal. Para dar otro ejemplo, en P^2 , si definimos los vectores $f(x) = 7x^2 - 5x + 2, g(x) = x^2, h(x) = 9x, i(x) = 7, j(x) = x^2 + x + 1$, podemos verificar que

$$f(x) = 7g(x) - \frac{5}{9}h(x) + \frac{2}{7}i(x) = 7j(x) - \frac{4}{3}h(x) + \frac{1}{7}i(x) = 3j(x) + 4g(x) - \frac{8}{9}h(x) - \frac{1}{7}i(x) = \dots$$

Para dar algunos ejemplos, si elegimos cualquier vector $\mathbf{v} \in \mathbb{R}^2$, entonces el subespacio generado correspondiente $\langle \mathbf{v} \rangle = \{c\mathbf{v} \mid c \in \mathbb{R}\}$ se puede interpretar geométricamente en el plano cartesiano como el conjunto de todas las flechas posibles de obtener a partir de reescalamientos de \mathbf{v} . Por otro lado, si en \mathbb{R}^3 definimos a $N = \left\{ \begin{pmatrix} 1 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 1 \end{pmatrix} \right\}$ entonces vemos que

$$\langle N \rangle = \{c_1 \begin{pmatrix} 1 & 0 & 0 \end{pmatrix} + c_2 \begin{pmatrix} 0 & 1 & 0 \end{pmatrix} + c_3 \begin{pmatrix} 0 & 0 & 1 \end{pmatrix} \mid c_1, c_2, c_3 \in \mathbb{R}\},$$

pero esto es equivalente a la definición $\mathbb{R}^3 = \left\{ \begin{pmatrix} c_1 & c_2 & c_3 \end{pmatrix} \mid c_1, c_2, c_3 \in \mathbb{R} \right\}$; es decir, en este caso *el espacio generado por los vectores de N es igual a \mathbb{R}^3* , i.e., $\langle N \rangle = \mathbb{R}^3$.

A continuación, veremos un teorema que será de gran importancia en las secciones posteriores.

Teorema 2.10 Sea V un espacio vectorial, $S \subseteq V$ un conjunto finito de vectores de V y $\mathbf{v} \in V$ un vector arbitrario. Si $S' = S \cup \{\mathbf{v}\}$, entonces $\langle S \rangle = \langle S' \rangle \iff \mathbf{v} \in \langle S \rangle$.

Demostración. Ya que $\mathbf{v} \in S'$ entonces trivialmente se cumple que $\mathbf{v} \in \langle S' \rangle$; por lo tanto, si $\mathbf{v} \notin \langle S \rangle \implies \langle S \rangle \neq \langle S' \rangle$. Por otro lado, si $\mathbf{v} \in \langle S \rangle$ entonces $S' \subset \langle S \rangle$, lo cual implica que $\langle S' \rangle \subset \langle S \rangle$. Además, ya que $S \subset S'$, entonces trivialmente se cumple que $\langle S \rangle \subset \langle S' \rangle$. En conclusión, $\langle S \rangle = \langle S' \rangle$. \square

Este teorema nos dice que agregar un vector a un conjunto generador no necesariamente cambiará el subespacio generado por ese conjunto generador. Para que este cambio realmente suceda, el vector añadido debe ser en algún sentido *ajeno* a los del conjunto generador original. En la siguiente sección, veremos algunas definiciones necesarias para precisar esta idea.

3. Dependencia e independencia lineal, bases y dimensión

Dependencia e independencia lineal

Como se vio en la sección anterior, un vector puede ser expresado como diferentes combinaciones lineales de otros vectores del mismo espacio. En particular, el vector nulo $\mathbf{0}$ de cualquier espacio vectorial V puede ser obtenido a través de la *combinación lineal trivial* de cualesquiera n vectores del espacio: sólo basta con que todos los coeficientes sean cero, i.e.

$$0\mathbf{v}_1 + 0\mathbf{v}_2 + \dots + 0\mathbf{v}_n = \mathbf{0}, \quad \forall \mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n \in V.$$

Sin embargo, también pueden existir combinaciones lineales entre n vectores de un espacio vectorial V que den como resultado $\mathbf{0}$ pero sean no triviales (es decir, tengan coeficientes distintos de cero), e.g., en \mathbb{R}^2 , $7 \begin{pmatrix} 1 & 1 \end{pmatrix} + 5 \begin{pmatrix} -1 & 1 \end{pmatrix} + 2 \begin{pmatrix} -1 & -6 \end{pmatrix} = \begin{pmatrix} 0 & 0 \end{pmatrix}$. Una consecuencia de este hecho es que podamos despejar a cualquiera de los vectores y expresarlo como combinación lineal de los demás; por ejemplo, $\begin{pmatrix} 1 & 1 \end{pmatrix} = -\frac{5}{7} \begin{pmatrix} -1 & 1 \end{pmatrix} - \frac{2}{7} \begin{pmatrix} -1 & -6 \end{pmatrix}$, ó $\begin{pmatrix} -1 & -6 \end{pmatrix} = -\frac{7}{2} \begin{pmatrix} 1 & 1 \end{pmatrix} - \frac{5}{2} \begin{pmatrix} -1 & 1 \end{pmatrix}$, etc. Este importante hecho motiva la siguiente definición.

Definición de dependencia e independencia lineal

Def. Sea V un espacio vectorial y $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n \in V$. Decimos que los vectores $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n$ son *linealmente independientes* entre sí si la única combinación lineal de ellos que da como resultado el vector nulo es la trivial (i.e., en la cual todos los coeficientes son cero). Matemáticamente,

$$\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n \text{ son l.i.} \iff (c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \dots + c_n\mathbf{v}_n = \mathbf{0} \implies c_1, c_2, \dots, c_n = 0).$$

En cambio, decimos que $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n$ son *linealmente dependientes* si existe al menos una combinación lineal no trivial que dé como resultado el vector nulo o, equivalentemente, si cualquiera de los vectores \mathbf{v}_i puede ser expresado como una combinación lineal de los demás.

Si todos los vectores de un conjunto S son linealmente dependientes (independientes) entre sí, se dice que el conjunto S es linealmente dependiente (independiente)^a.

^aObservemos que cualquier conjunto que contenga al vector nulo será linealmente dependiente.

Por ejemplo, en \mathbb{R}^2 los vectores $\mathbf{u}_1 = \begin{pmatrix} 1 & 5 \end{pmatrix}$ y $\mathbf{u}_2 = \begin{pmatrix} -3 & -15 \end{pmatrix}$ son linealmente dependientes, ya que $\mathbf{u}_2 = -3\mathbf{u}_1$, por lo cual $3\mathbf{u}_1 + \mathbf{u}_2 = \mathbf{0}$; por otro lado, los vectores $\mathbf{v}_1 = \begin{pmatrix} 1 & 2 \end{pmatrix}$ y $\mathbf{v}_2 = \begin{pmatrix} 1 & 3 \end{pmatrix}$ son linealmente independientes, ya que no existe un número $c \in \mathbb{R}$ tal que $\mathbf{v}_1 = c\mathbf{v}_2$. Notemos que, como nuestros vectores en este caso son 2-tuplas, las ecuaciones $\mathbf{u}_2 = -3\mathbf{u}_1$ y $\mathbf{v}_1 = c\mathbf{v}_2$ son en realidad la notación compactada para un *sistema de ecuaciones*, con una ecuación por cada entrada del vector. En particular, la ecuación $\mathbf{v}_1 = c\mathbf{v}_2$ puede ser reescrita como

$$\begin{pmatrix} 1 & 2 \end{pmatrix} = c \begin{pmatrix} 1 & 3 \end{pmatrix} \iff 1 = c1 \wedge 2 = c3,$$

de donde vemos directamente que no existe solución para c , por lo cual estos vectores son linealmente independientes.

En general, cuando expresamos combinaciones lineales del tipo $\mathbf{v} = c_1\mathbf{u}_1 + \dots + c_n\mathbf{u}_n$ en donde los vectores son dados pero los coeficientes son desconocidos, éstos últimos se vuelven las *incógnitas* del *sistema de ecuaciones algebraicas* dado por la ecuación $\mathbf{v} = c_1\mathbf{u}_1 + \dots + c_n\mathbf{u}_n$. El número de ecuaciones

del sistema dependerá de la naturaleza de los vectores, mientras que el número de incógnitas será igual al número de coeficientes desconocidos. Por lo tanto, la pregunta de si un vector es linealmente independiente o dependiente de otro(s) se reduce a la de si el sistema de ecuaciones asociado a la combinación lineal de ellos que da como resultado el vector nulo tiene una solución no trivial o no.

Para ver más ejemplos de conjuntos de vectores linealmente dependientes e independientes pueden consultar, e.g., el Friedberg (págs. 36-38), el Lang introductorio (págs. 104-109), etc.

Interpretación geométrica de la dependencia e independencia lineal

Como ya mencionamos, si un vector $\mathbf{v}_n \in V$ es linealmente *dependiente* de otros vectores $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_m \in V$, entonces puede ser expresado como combinación lineal de esos vectores. Geométricamente, en los espacios vectoriales reales \mathbb{R}^2 y \mathbb{R}^3 esto quiere decir que es posible reescalar y combinar (mediante la Ley del paralelogramo) las flechas de los vectores $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_m$ y obtener, como resultado final, a \mathbf{v}_n . Adicionalmente, en el espacio vectorial complejo \mathbb{C} , también se podrían estar rotando los vectores $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_m$ —además de reescalarlos y combinarlos— para formar, finalmente, a \mathbf{v}_n . Si son linealmente *independientes*, entonces lo anterior no es posible.

Ejercicio 10. Sea n_i el i -ésimo dígito de tu número de cuenta. Determina si los vectores complejos $(n_1 + i(n_2))$ y $(n_3 + i(n_4))$ son linealmente dependientes o independientes en (\mathbb{C}, \mathbb{C}) y muéstralo gráficamente en el plano complejo. Luego, repite el mismo ejercicio para los vectores reales $(n_1 \ n_4 \ n_7), (n_2 \ n_5 \ n_8)$ y $(n_3 \ n_6 \ n_9) \in \mathbb{R}^3$. ¿Cuáles son los conjuntos linealmente independientes más grandes que puedes formar usando a estos vectores?

Algunos teoremas sobre dependencia e independencia lineal

Teorema 3.1 Sea V un espacio vectorial y $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n$ vectores linealmente independientes de V . Sean $c_1, c_2, \dots, c_n \in K$ y $d_1, d_2, \dots, d_n \in K$ tales que

$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \dots + c_n\mathbf{v}_n = d_1\mathbf{v}_1 + d_2\mathbf{v}_2 + \dots + d_n\mathbf{v}_n,$$

entonces se tiene que $c_i = d_i \ \forall i \in \{1, 2, \dots, n\}$.

Demostración.

$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \dots + c_n\mathbf{v}_n = d_1\mathbf{v}_1 + d_2\mathbf{v}_2 + \dots + d_n\mathbf{v}_n \iff (c_1 - d_1)\mathbf{v}_1 + (c_2 - d_2)\mathbf{v}_2 + \dots + (c_n - d_n)\mathbf{v}_n = \mathbf{0}.$$

Pero, ya que por hipótesis estos vectores son linealmente independientes, entonces por definición

$$c_i - d_i = 0 \ \forall i \in \{1, 2, \dots, n\} \iff c_i = d_i \ \forall i \in \{1, 2, \dots, n\}.$$

□

Este teorema nos dice que si un vector es resultado de una combinación lineal de vectores linealmente independientes, entonces esa combinación lineal es *la única* que da como resultado a ese vector. Es decir, que si $\mathbf{u} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \dots + c_n\mathbf{v}_n$ y $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\}$ es *l.i.* entonces no existe otra combinación de escalares y vectores $c_i\mathbf{v}_i$ tal que la suma de todos ellos dé \mathbf{u} .

Teorema 3.2 Sean V un espacio vectorial y S_1, S_2 subconjuntos de V tales que $S_1 \subseteq S_2 \subseteq V$. Si S_2 es linealmente independiente, entonces S_1 también es linealmente independiente.

Demostración. Sean $S_1 = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k\}$ y $S_2 = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\}$ con $k \leq n$. Ya que por hipótesis S_2 es l.i., por definición se cumple que

$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \dots + c_n\mathbf{v}_n = \mathbf{0} \iff c_i = 0 \quad \forall i \in \{1, 2, \dots, n\}.$$

Si $k = n$ entonces S_1 también es l.i. trivialmente. Supongamos que $k < n$. Entonces, por un Teorema anterior, $0\mathbf{v}_{k+1} + \dots + 0\mathbf{v}_n = \mathbf{0}$ por lo cual lo anterior implica que

$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \dots + c_k\mathbf{v}_k = \mathbf{0} \iff c_i = 0 \quad \forall i \in \{1, 2, \dots, k\}.$$

Por lo tanto, por definición, S_1 también es l.i. □

Este teorema nos dice que si removemos vectores de un conjunto linealmente independiente, el conjunto resultante también es linealmente independiente. El resultado opuesto se deja como ejercicio.

Ejercicio 11. Sean V un espacio vectorial y S_1, S_2 subconjuntos de V tales que $S_1 \subseteq S_2 \subseteq V$. Demuestra que, si S_1 es linealmente dependiente, entonces S_2 es linealmente dependiente.

Teorema 3.3 Sea V un espacio vectorial sobre un campo K y $L \subset V$ un conjunto con n elementos linealmente independientes entre sí. Entonces, para cualquier $\mathbf{v} \in V$, el conjunto $L' \equiv L \cup \{\mathbf{v}\}$ es l.i. $\iff \mathbf{v} \notin \langle L \rangle$.

Demostración. Sea $L = \{\mathbf{u}_1, \dots, \mathbf{u}_n\}$. Supongamos que $\mathbf{v} \in \langle L \rangle$, entonces existen coeficientes $c_i \in K$ tales que $\mathbf{v} = c_1\mathbf{u}_1 + \dots + c_n\mathbf{u}_n$. Despejando esta ecuación, obtenemos que $c_1\mathbf{u}_1 + \dots + c_n\mathbf{u}_n + (-1)\mathbf{v} = \mathbf{0}$, es decir, que existe una combinación lineal no trivial entre los vectores de L' que dan como resultado al vector nulo, por lo cual L' es un conjunto linealmente dependiente.

Por otro lado, supongamos que L' es linealmente dependiente. Entonces, existe una combinación lineal no trivial de los vectores de L' que resulta en el vector nulo, i.e., $c_1\mathbf{u}_1 + \dots + c_n\mathbf{u}_n + b\mathbf{v} = \mathbf{0}$ con al menos un coeficiente distinto de cero. En este caso, el coeficiente $b \neq 0$: si b fuera igual a cero, la ecuación restante sería $c_1\mathbf{u}_1 + \dots + c_n\mathbf{u}_n = \mathbf{0}$; ya que L es linealmente independiente, entonces todos los vectores c_i deben ser iguales a cero pero, ya que estamos suponiendo que también $b = 0$, entonces el conjunto L' también sería linealmente independiente, contradiciendo la hipótesis. Así, sabiendo que $b \neq 0$ podemos despejar la ecuación $c_1\mathbf{u}_1 + \dots + c_n\mathbf{u}_n + b\mathbf{v} = \mathbf{0}$ y obtener que $\mathbf{v} = \frac{-c_1}{b}\mathbf{u}_1 + \dots + \frac{-c_n}{b}\mathbf{u}_n$, lo cual implica que $\mathbf{v} \in \langle L \rangle$. □

La demostración de este teorema nos dice que si tenemos un conjunto linealmente independiente y agregamos a un vector de su subespacio generado a este conjunto, entonces se vuelve linealmente dependiente. En contraposición, concluimos que en cualquier conjunto linealmente dependiente existe una especie de *redundancia* entre sus elementos, ya que se puede remover a cualquiera de ellos sin alterar el subespacio generado por este conjunto. En cambio, remover un vector de un conjunto linealmente independiente *sí* altera el subespacio generado por ese conjunto.

Bases y dimensión

Como hemos visto en secciones anteriores, cualquier vector de un espacio vectorial se puede expresar como combinación lineal de otros vectores de ese mismo espacio⁸. Cuando trabajamos en un espacio vectorial V , resulta conveniente tener un conjunto de vectores $B \subseteq V$ con el cual se pueda expresar a *cualquier vector del espacio vectorial V* de forma *única* —lo cual se logra, precisamente, a través de una combinación lineal única de los vectores del conjunto B . Tomando en cuenta el Teorema 3.1, vemos que los conjuntos linealmente independientes son buenos candidatos para lograr que las expresiones mediante combinaciones lineales sean *únicas*, por lo cual pediremos que B sea linealmente independiente; además, ya que queremos ser capaces de expresar a *cualquier* vector arbitrario de V como combinación lineal *única* de los vectores de B , sería necesario que el conjunto linealmente independiente B generara a *todos* los vectores de V . A cualquier conjunto que cumpla ambas propiedades se le conoce como una *base* para el espacio vectorial en cuestión.

Definición de base

Def. Una base de un espacio vectorial V es un conjunto de vectores linealmente independientes que generan a todo el espacio vectorial V . En lenguaje matemático,

$$B \subset V \text{ es una base de } V \iff B \text{ es l.i. y } \langle B \rangle = V.$$

Nótese por la definición que, ya que muchos conjuntos de vectores distintos pueden ser linealmente independientes y generar a un mismo espacio vectorial, un espacio vectorial puede tener muchas bases distintas. Esto implica que cualquier vector arbitrario de un espacio vectorial puede ser expresado a través de diferentes combinaciones lineales (correspondientes a distintas bases del espacio, y únicas para cada base). Dicho de otra forma, dado un espacio vectorial con más de una base, cualquier vector de ese espacio puede ser *representado en las distintas bases* de ese espacio⁹.

Ejemplos de bases

El conjunto $\{1\}$ es una base para el espacio vectorial complejo \mathbb{C} . De hecho, si cambiamos a 1 en el conjunto anterior por cualquier número complejo no nulo, también tendremos una base para el espacio complejo \mathbb{C} . ¿A qué propiedades se debe esto?.

Los conjuntos $\{(2 \ 0), (0 \ -3)\}$, $\{(3 \ 3), (-3 \ 3)\}$ y $\{(1 \ 0), (0 \ 1)\}$ son bases de \mathbb{R}^2 .

Cualquier conjunto de la forma $\{c_n x^n \mid n \in \mathbb{N} \cup \{0\}, c_n \in \mathbb{R}\}$ es una base del espacio vectorial de las funciones polinomiales de grado n .

Teorema de reemplazamiento

A continuación, veremos un importante teorema que nos ayudará a construir bases más adelante.

Teorema 3.4 Sea V un espacio vectorial generado por un conjunto G con n vectores, y sea L un subconjunto linealmente independiente de V con m vectores. Entonces $m \leq n$ y existe un subconjunto $H \subseteq G$ que contiene $n - m$ vectores tal que $L \cup H$ genera a V .

⁸De lo contrario, se violarían las propiedades de cerradura de los espacios vectoriales.

⁹El tema de las *representaciones* es de gran interés en algunas ramas de las matemáticas y sus aplicaciones son de suma importancia en varias áreas de la física. En este curso, lo veremos sobre todo en las secciones de representación matricial de una transformación lineal, representación de una matriz en distintas bases y representaciones de un operador lineal en distintos espacios vectoriales.

Demostración. Esta demostración se hará por inducción.

Base inductiva Sea $m = 0$, entonces $L = \emptyset$. Si tomamos $H = G$ obtenemos el resultado deseado.

Hipótesis de inducción Supongamos que la hipótesis del teorema se cumple para $L = \{\mathbf{v}_1, \dots, \mathbf{v}_m\}$ con $m > 0$.

Paso inductivo Ahora debemos demostrar que, bajo la hipótesis de inducción (donde el teorema se cumple para alguna $m > 0$), el teorema se debe cumplir para $m + 1$.

Sea $L = \{\mathbf{v}_1, \dots, \mathbf{v}_{m+1}\}$ un subconjunto linealmente independiente de V . Por el Teorema 3.2 el conjunto $\{\mathbf{v}_1, \dots, \mathbf{v}_m\}$ es l.i., por lo cual podemos aplicar la hipótesis de inducción y concluir que $m \leq n$ y que existe un subconjunto $\{\mathbf{u}_1, \dots, \mathbf{u}_{n-m}\} \subset G$ tal que $\{\mathbf{v}_1, \dots, \mathbf{v}_m\} \cup \{\mathbf{u}_1, \dots, \mathbf{u}_{n-m}\}$ genera a V . Por lo tanto, existen escalares $a_1, \dots, a_m, b_1, \dots, b_{n-m}$ tales que

$$a_1 \mathbf{v}_1 + \dots + a_m \mathbf{v}_m + b_1 \mathbf{u}_1 + \dots + b_{n-m} \mathbf{u}_{n-m} = \mathbf{v}_{m+1}.$$

Observemos que, ya que L es linealmente independiente, $n - m > 0 \implies n > m \implies n \geq m + 1$. Además, alguna b_i debe ser distinta de cero, por lo cual podemos despejarla (de lo contrario, estaríamos contradiciendo la hipótesis de inducción, que nos asegura que $\{\mathbf{v}_1, \dots, \mathbf{v}_m\}$ es l.i.). Suponiendo, por ejemplo, que $b_1 \neq 0$, tenemos que

$$\mathbf{u}_1 = \frac{-a_1}{b_1} \mathbf{v}_1 + \dots + \frac{-a_m}{b_1} \mathbf{v}_m + \frac{-b_2}{b_1} \mathbf{u}_2 + \dots + \frac{-b_{n-m}}{b_1} \mathbf{u}_{n-m},$$

por lo cual \mathbf{u}_1 puede ser expresado como combinación lineal de los vectores $\mathbf{v}_1, \dots, \mathbf{v}_m, \mathbf{u}_2, \dots, \mathbf{u}_{n-m}$. Sea $H = \{\mathbf{u}_2, \dots, \mathbf{u}_{n-m}\}$, entonces $L \cup H = \{\mathbf{v}_1, \dots, \mathbf{v}_{m+1}, \mathbf{u}_2, \dots, \mathbf{u}_{n-m}\}$ y trivialmente tenemos que $\mathbf{v}_1, \dots, \mathbf{v}_m, \mathbf{u}_2, \dots, \mathbf{u}_{n-m} \in \langle L \cup H \rangle$ —lo cual también implica que $\mathbf{u}_1 \in \langle L \cup H \rangle$. Por lo tanto, tenemos que $\{\mathbf{v}_1, \dots, \mathbf{v}_m, \mathbf{u}_1, \dots, \mathbf{u}_{n-m}\} \subseteq \langle L \cup H \rangle$. Recordando que por hipótesis de inducción $\{\mathbf{v}_1, \dots, \mathbf{v}_m, \mathbf{u}_1, \dots, \mathbf{u}_{n-m}\}$ genera a V , entonces el hecho de que esté contenido en $L \cup H$ implica necesariamente que $\langle L \cup H \rangle = V$. Finalmente, ya que H es un subconjunto de G con $(n - m) - 1 = n - (m + 1)$ vectores, el teorema se cumple para $m + 1$, terminando así nuestra demostración. \square

El teorema anterior se conoce como el teorema de *reemplazamiento* ya que, partiendo de un conjunto linealmente independiente L y otro conjunto H que juntos cumplen $\langle L \cup H \rangle = V$ (sin que $L \cup H$ sea necesariamente l.i.), lo que estamos haciendo con cada paso consecutivo de la inducción es reemplazar a los vectores de H por vectores que podemos añadir a L tal que este conjunto siga siendo linealmente independiente y se siga cumpliendo que la unión de ambos genere a V . De esta forma, L es un conjunto linealmente independiente que va creciendo y que cada vez necesita a menos vectores de H para poder, a través de la unión, generar a V . ¿Qué pasará cuando L sea un conjunto linealmente independiente que no necesita a ningún vector de H para generar a V ¹⁰?

Ejercicio 12. Considera al espacio vectorial \mathbb{R}^n para distintos valores de n y responde: ¿A partir de qué cardinalidad cualquier conjunto de vectores de \mathbb{R}^n es necesariamente linealmente dependiente?

¹⁰Recuerda para qué dijimos que nos serviría este teorema.

Dimensión

Como quizá notaste en los ejemplos de la sección anterior, pareciera que todas las bases de un mismo espacio vectorial tienen el mismo número de elementos. A continuación, demostraremos este hecho.

Teorema 3.5 Sean $B = \{\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_n\}$ y $B' = \{\mathbf{b}'_1, \mathbf{b}'_2, \dots, \mathbf{b}'_{n'}\}$ bases de V , entonces $n = n'$.

Demostración. Supongamos que $n' > n$. Ya que $B' \subset V$ y $\langle B \rangle = V \implies B' \subset \langle B \rangle$, por lo cual podemos expresar cualquier vector de B' como combinación lineal de los de B . Entonces,

$$\mathbf{b}'_1 = c_{11}\mathbf{b}_1 + c_{12}\mathbf{b}_2 + \dots + c_{1n}\mathbf{b}_n,$$

...

$$\mathbf{b}'_{n'} = c_{n'1}\mathbf{b}_1 + c_{n'2}\mathbf{b}_2 + \dots + c_{n'n}\mathbf{b}_n,$$

donde $c_{ij} \in K$.

Sea $\mathbf{z} \in V$ un vector arbitrario. Como B' es base de $V \implies \mathbf{z} = d_1\mathbf{b}'_1 + d_2\mathbf{b}'_2 + \dots + d_{n'}\mathbf{b}'_{n'}$. Sustituyendo con las ecuaciones, obtenemos que

$$\mathbf{z} = d_1(c_{11}\mathbf{b}_1 + \dots + c_{1n}\mathbf{b}_n) + \dots + d_{n'}(c_{n'1}\mathbf{b}_1 + \dots + c_{n'n}\mathbf{b}_n) = (d_1c_{11} + \dots + d_{n'}c_{n'1})\mathbf{b}_1 + \dots + (d_1c_{1n} + \dots + d_{n'}c_{n'n})\mathbf{b}_n.$$

En particular, si $\mathbf{z} = \mathbf{0}$, ya que por hipótesis B es linealmente independiente, obtenemos

$$d_1c_{11} + \dots + d_{n'}c_{n'1} = 0,$$

...

$$d_1c_{1n} + \dots + d_{n'}c_{n'n} = 0.$$

Sin embargo, ya que al inicio de la demostración supusimos que $n' > n$, entonces el sistema de ecuaciones anterior tiene más incógnitas que ecuaciones y, por ende, una solución no trivial para $(d_1, d_2, \dots, d_{n'})$. Esto contradice el hecho de que B sea una linealmente independiente, por lo cual tampoco podría ser una base. Análogamente, si $n > n'$ se llega a una contradicción similar. Por lo tanto, por tricotomía concluimos que, si B y B' son bases, $n = n'$. \square

El hecho de que todas las bases de un mismo espacio vectorial tengan el mismo número de elementos motiva la siguiente definición.

Def. La *dimensión* de un espacio vectorial V es igual al número de elementos (i.e., la cardinalidad) de cualquiera de sus bases. Si cualquier base de V tiene un número finito n de elementos, decimos que V es un *espacio de dimensión (finita) n* y escribimos esto como $\dim(V) = n$; de lo contrario decimos que V es un espacio de dimensión *infinita*^a.

^aEn el resto de estas notas, supondremos que los espacios vectoriales mencionados tienen dimensión finita, a menos que se indique lo contrario.

Ejercicio 13. Sean K un campo arbitrario y $\mathbf{e}_1 = \begin{pmatrix} 1 & 0 & 0 & 0 & \dots & 0 \end{pmatrix}, \mathbf{e}_2 = \begin{pmatrix} 0 & 1 & 0 & 0 & \dots & 0 \end{pmatrix}, \dots, \mathbf{e}_n = \begin{pmatrix} 0 & 0 & 0 & 0 & \dots & 1 \end{pmatrix}$. Demuestra que $\{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n\}$ es una base para (K^n, K) y que, por lo tanto, $\dim(K^n, K) = n$.

Observemos que esta definición *algebraica* de dimensión difiere de las definiciones geométricas y físicas usuales de dimensión. Por ejemplo, a pesar de que el espacio vectorial complejo \mathbb{C} se represente en el plano cartesiano —el cual tiene dimensión geométrica 2—, este espacio vectorial es de dimensión (algebraica) 1, como vimos en los ejemplos de la sección 3. Otra observación es que la dimensión de un espacio vectorial no sólo depende del conjunto vectorial, sino también del campo sobre el cual se define, como en el siguiente ejercicio.

Ejercicio 14. Demuestra que $\dim(\mathbb{C}, \mathbb{C}) = 1$ pero $\dim(\mathbb{C}, \mathbb{R}) = 2$.

Teorema 3.6 Sea V un espacio vectorial de dimensión finita y W un subespacio de V , entonces W tiene dimensión finita y $\dim(W) \leq \dim(V)$.

Demostración. Sea $\dim(V) = n$. Si $W = \{\mathbf{0}\} \implies \dim(W) = 0 \leq n$. Consideremos ahora que W contiene a un vector no nulo \mathbf{x}_1 , entonces el conjunto $\{\mathbf{x}_1\}$ es linealmente independiente. Supongamos que seguimos agregando más vectores $\mathbf{x}_2, \dots, \mathbf{x}_k$ de W al conjunto $\{\mathbf{x}_1\}$ de tal forma que $\{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_k\}$ sea linealmente independiente. Ya que $\dim(V) = n$, entonces cualquier base de V tiene n elementos. Esto implica que ningún subconjunto de V linealmente independiente puede tener más de n elementos, por lo cual el proceso anterior debe detenerse para algún $k \leq n$. De acuerdo a un Teorema anterior, este conjunto genera a W , por lo cual forma una base de W , de donde concluimos que $\dim(W) = k \leq n$. \square

En los teoremas anteriores demostramos que si $\dim(V) = n$ entonces cualquier base de V tiene precisamente n elementos, y que cualquier subespacio vectorial tiene dimensión finita. Resulta, además, que en este caso cualquier conjunto de n vectores linealmente independientes de V es también una base para V —es decir, que también genera a todo el espacio V , como veremos en el siguiente teorema.

Teorema 3.7 Sea V un espacio vectorial. Si $\dim(V) = n$ entonces cualquier conjunto de n vectores linealmente independientes de V es una base de V .

Demostración. Esta prueba se hará por contradicción.

Sea V un espacio vectorial con $\dim(V) = n$, $n \in \mathbb{N}$ y $B = \{\mathbf{b}_1, \dots, \mathbf{b}_n\} \subset V$ un conjunto de n vectores de V que son linealmente independientes entre sí.

Supongamos que $\langle B \rangle \neq V$, es decir, que $\exists \mathbf{b}_{n+1} \in V$ tal que éste no puede ser expresado como combinación lineal de los vectores de B . Por definición, entonces dicho vector es linealmente independiente de los vectores de B . Por lo tanto, podemos definir al conjunto $B' \equiv \{\mathbf{b}_1, \dots, \mathbf{b}_n, \mathbf{b}_{n+1}\}$, que tiene $n + 1$ elementos linealmente independientes entre sí. Supongamos que, ahora sí, $\langle B' \rangle = V$; en ese caso, por definición, B' sería una base de V . Sin embargo, ya que $\dim(V) = n$, por lo demostrado en el Teorema 4.2.1 hemos llegado a una contradicción, ya que cualquier base de V debe tener exactamente n elementos.

Ya que la suposición $\langle B \rangle \neq V$ fue la que nos llevó a esta contradicción, tenemos que $\langle B \rangle = V$, por lo cual B —un conjunto arbitrario de n elementos linealmente independientes de V — es una base de V . \square

Corolario 3.8 Sean V es un espacio vectorial de dimensión finita y W un subespacio vectorial de V . Entonces, $\dim(W) = \dim(V)$ implica que $W = V$.

Demostración. Se sigue del Teorema 3.7. \square

Ejercicio 15. Sea V un espacio vectorial de dimensión n y D un conjunto linealmente dependiente tal que $\langle D \rangle = V$. Demuestra que D tiene más de n elementos.

Para terminar esta sección, veremos algunas formas en las cuales se pueden construir bases de un espacio vectorial de dimensión n a partir de conjuntos linealmente independientes con menos de n elementos, o de conjuntos linealmente dependientes que generan V y tienen más de n elementos.

Teorema 3.9 Sea V un espacio vectorial de dimensión n , entonces cualquier conjunto finito linealmente dependiente que genera a V puede reducirse hasta convertirse en una base de V .

Demostración. Sea D un conjunto finito linealmente dependiente tal que $\langle D \rangle = V$. Ya que D es linealmente dependiente, entonces existe un vector \mathbf{v} que puede ser expresado como combinación lineal de los demás por lo cual, definiendo $D' = D \setminus \{\mathbf{v}\}$ tenemos que $\mathbf{v} \in \langle D' \rangle$, donde claramente D' también es finito. Por el Teorema 2.10 sabemos que $\langle D' \rangle = \langle D \rangle = V$. Si el conjunto generador D' no es linealmente independiente, podemos seguir retirando vectores de la misma forma sin afectar su espacio generado hasta obtener un conjunto linealmente independiente que genera a V , es decir, una base para V . \square

Teorema 3.10 Sea V un espacio vectorial de dimensión n , entonces cualquier conjunto linealmente independiente con menos de n elementos no genera a V , pero puede extenderse hasta convertirse en una base de V .

Demostración. Sea L un conjunto linealmente independiente con m elementos, donde $m < n$. Entonces L no genera a V ya que, de lo contrario, sería una base de V y tendríamos dos bases de V con diferente cardinalidad, lo cual contradice al Teorema 3.5. Sea S un conjunto generador de V . Ya que L no genera a V , debe haber algún vector en $\mathbf{v} \in S$ tal que $\mathbf{v} \notin \langle L \rangle$. Definimos ahora al conjunto $L' = L \cup \{\mathbf{v}\}$ el cual, por el Teorema 3.3, es linealmente independiente. Si L' no genera a V , podemos repetir el proceso hasta llegar a un conjunto linealmente independiente que genera a V , i.e., una base para V . \square

Sabiendo que un mismo espacio vectorial puede tener muchas bases diferentes —todas con el mismo número de elementos—, en la siguiente sección nos enfocaremos a ver algunos tipos de bases que resultan ser útiles comunmente, y a entender cómo podemos construirlas y usarlas.

4. Nociones básicas de transformaciones lineales

Las estructuras algebraicas no sólo se pueden estudiar a través de sus *elementos*, sino también a través de *funciones* entre ellas. En particular, un tipo de funciones muy útiles para estudiar estructuras algebraicas son aquellas que preservan la estructura en cuestión. En el caso particular de los espacios vectoriales, una función que preserve este tipo de estructura tiene espacios vectoriales como dominio y contradominio y es compatible con las operaciones de suma vectorial y producto de un vector por un escalar; a este tipo de funciones se les conoce como transformaciones lineales. Una definición más operativa es la siguiente.

Definición de transformación lineal

Def. Sean K un campo, S un subcampo de K^a y $(V, S), (W, K)$ espacios vectoriales. Entonces, decimos que una función $T : V \rightarrow W$ es una *transformación lineal* si para todo $a \in S$, $\vec{u}, \vec{v} \in V$, se tiene que

$$T(\vec{u} + \vec{v}) = T(\vec{u}) + T(\vec{v}) \quad \& \quad T(a\vec{v}) = aT(\vec{v}).$$

Informalmente, decimos que una transformación lineal T *abre sumas y saca escalares*.

^aEn particular, podemos tomar $S = K$.

Ejercicio 16. Demuestra que las transformaciones lineales mandan al vector nulo de su dominio al vector nulo de su contradominio. Es decir, si $T : V \rightarrow W$ es una transformación lineal y los vectores nulos de V y W son $\vec{0}_V$ y $\vec{0}_W$, respectivamente, entonces $T(\vec{0}_V) = \vec{0}_W$.

Una forma equivalente y más compacta de definir a las transformaciones lineales está dada por la siguiente caracterización.

Ejercicio 17 (Caracterización de transformaciones lineales). Sean (V, S) y (W, K) espacios vectoriales, con S un subcampo de K . Demuestra que $T : V \rightarrow W$ es una transformación lineal si, y sólo si, para todo $a \in S$, $\vec{u}, \vec{v} \in V$, se tiene que $T(\vec{u} + a\vec{v}) = T(\vec{u}) + aT(\vec{v})$.

Ejemplos de transformaciones lineales

La función $T : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ dada por $T(x, y) = (x, -y)$ es una transformación lineal, pues para todo $a \in \mathbb{R}$, $(x, y), (x', y') \in \mathbb{R}^2$, tenemos que

$$\begin{aligned} T(x, y) + aT(x', y') &= (x, -y) + a(x', -y') \\ &= (x, -y) + (ax', -ay') \\ &= (x + ax', -y - ay') \\ &= (x + ax', -(y + ay')) \\ &= T(x + ax, y + ay) \\ &= T((x, y) + (ax', ay')) \\ &= T((x, y) + a(x', y')), \end{aligned}$$

por lo que $T((x, y) + a(x', y')) = T(x, y) + aT(x', y')$. Podemos interpretar geométricamente a esta transformación como una reflexión en el plano cartesiano a lo largo del eje vertical. De manera análoga, se puede demostrar que la reflexión con respecto al eje horizontal y al origen y, más generalmente, que las reflexiones en \mathbb{R}^n son transformaciones lineales.

Siguiendo del ejemplo anterior en \mathbb{R}^2 , observemos que podemos considerar a una función que haga algo “semejante” en el espacio vectorial complejo \mathbb{C} : la función de conjugación compleja $\overline{a + ib} = a - ib$. Sin embargo, en este caso, esta función no preserva la estructura del espacio \mathbb{C} . ¿Por qué?

Ejercicio 18. Demuestra que la función de conjugación compleja $\bar{\cdot} : \mathbb{C} \rightarrow \mathbb{C}$ abre sumas pero no es una transformación lineal.

Sea $C^\infty(\mathbb{R}, \mathbb{R})$ el espacio de funciones reales de variable real con todas sus derivadas continuas. Entonces la derivada $\frac{d}{dx} : C^\infty(\mathbb{R}, \mathbb{R}) \rightarrow C^\infty(\mathbb{R}, \mathbb{R})$ es una transformación lineal. Si consideramos el subespacio compuesto por todas las funciones integrables, entonces la integral $\int_{-\infty}^{\infty} dx$ es una transformación lineal en este subespacio.

Todas las transformaciones lineales anteriores van de un espacio vectorial en sí mismo. A este tipo de transformaciones lineales les damos un nombre especial.

Def. Sea V un espacio vectorial. Decimos que una transformación lineal $T : V \rightarrow V$ es un *operador lineal*.

Nos enfocaremos en el estudio de operadores lineales en la última parte del curso. Ahora, veamos ejemplos de transformaciones lineales que no son operadores lineales.

Las rotaciones en $\mathbb{R}^2, \mathbb{R}^3$ son transformaciones lineales¹¹.

La función $T : \mathbb{R}^2 \rightarrow \mathbb{R}^3$ dada por $T(x, y) = (x, y, 0)$ es una transformación lineal. Similarmente, la función $U : \mathbb{R}^3 \rightarrow \mathbb{R}^2$ dada por $U(x, y, z) = (x, y)$ es una transformación lineal. No es difícil mostrar que T es inyectiva y U es suprayectiva. Diremos que T es una *inclusión* de \mathbb{R}^2 en \mathbb{R}^3 , y que U es una *proyección* de \mathbb{R}^3 sobre \mathbb{R}^2 .

Sea $C^n(\mathbb{R}, \mathbb{R})$ el espacio de funciones reales de variable real con n derivadas continuas. Entonces $\frac{d}{dx} : C^n(\mathbb{R}, \mathbb{R}) \rightarrow C^{n-1}(\mathbb{R}, \mathbb{R})$ es una transformación lineal.

Núcleo e imagen de una transformación lineal

Podemos estudiar espacios vectoriales a través de transformaciones lineales entre ellos y, a la vez, podemos estudiar transformaciones lineales a través de algunos conjuntos que se definen a partir de ellas.

Def. Sean V, W espacios vectoriales y $T : V \rightarrow W$ una transformación lineal. Entonces definimos el *núcleo* de T como

$$\text{Ker}(T) := \{\vec{v} \in V \mid T(\vec{v}) = \vec{0}_W\}$$

y la *imagen* de T como^a

$$\text{Im}(T) := \{\vec{w} \in W \mid \exists \vec{v} \in V \text{ tal que } T(\vec{v}) = \vec{w}\}.$$

^aObservemos que esto es simplemente la definición usual de la imagen de una función.

Observación 4.1 Sea $T : V \rightarrow W$ una transformación lineal.

- (1) Por definición, $\text{Ker}(T) \subseteq V$ y, por el Ejercicio 16, $\vec{0}_V \in \text{Ker}(T)$, lo cual es una de las condiciones de la caracterización de subespacios vectoriales vista en la Proposición 2.2.

¹¹Puedes convencerte de esto haciendo un simple dibujo.

Ejercicio 19. Demuestra que $\text{Ker}(T)$ es un subespacio vectorial de V .

En particular, si V tiene dimensión finita, del Teorema 3.6 se sigue que $\text{Ker}(T)$ también tiene dimensión finita.

(2) Similarmente, por definición, $\text{Im}(T) \subseteq W$ y, por el Ejercicio 16, $\vec{0}_W \in \text{Im}(T)$.

Ejercicio 20. Demuestra que $\text{Im}(T)$ es un subespacio vectorial de W .

En particular, si W tiene dimensión finita, del Teorema 3.6 se sigue que $\text{Im}(T)$ también tiene dimensión finita.

Las dimensiones de los núcleos e imágenes de transformaciones lineales nos dan información útil sobre ellas, por lo que se les da un nombre especial.

Def. Sea $T : V \rightarrow W$ una transformación lineal. Entonces, definimos la *nulidad* de T como

$$\text{nulidad}(T) := \dim(\text{Ker}(T))$$

y el *rango* de T como

$$\text{rango}(T) := \dim(\text{Im}(T)).$$

Algunos ejemplos de esto están dados por los siguientes resultados.

Proposición 4.2 [Caracterización de transformaciones lineales inyectivas] Sean V, W espacios vectoriales y $T : V \rightarrow W$ una transformación lineal. Entonces T es una función inyectiva si, y sólo si, $\text{nulidad}(T) = 0$.

Demostración.

(\Rightarrow) Supongamos que T es una función inyectiva. Entonces, para cualesquiera $\vec{u}, \vec{v} \in V$, $T(\vec{u}) = T(\vec{v})$ implica que $\vec{u} = \vec{v}$. Por el Ejercicio 16, sabemos que $T(\vec{0}_V) = \vec{0}_W$. Dado que $T(\vec{v}) = \vec{0}_W$ para todo $\vec{v} \in \text{Ker}(T)$, se sigue que $\text{Ker}(T) = \{\vec{0}_V\}$. Como por definición \emptyset es una base de $\{\vec{0}_V\}$, se sigue que $\text{nulidad}(T) = 0$.

(\Leftarrow) Supongamos que $\text{nulidad}(T) = 0$. Entonces, $\dim(\text{Ker}(T)) = 0$, es decir, $\text{Ker}(T) = \{\vec{0}_V\}$. Luego, para cualesquiera $\vec{u}, \vec{v} \in V$, tenemos que

$$\begin{aligned} T(\vec{u}) = T(\vec{v}) &\implies T(\vec{u}) - T(\vec{v}) = \vec{0}_W \\ &\implies T(\vec{u} - \vec{v}) = \vec{0}_W && (T \text{ es lineal}) \\ &\implies \vec{u} - \vec{v} \in \text{Ker}(T) \\ &\implies \vec{u} - \vec{v} = \vec{0}_V && (\text{por hipótesis}) \\ &\implies \vec{u} = \vec{v}. \end{aligned}$$

Por lo tanto, se sigue que T es una función inyectiva. □

Proposición 4.3 (Caracterización de transformaciones lineales suprayectivas con contradominios de dimensión finita) Sean V, W espacios vectoriales con $\dim(W) = n$ finita y $T : v \rightarrow W$ una transformación lineal. Entonces T es una función suprayectiva si, y sólo si, $\text{rango}(T) = n$.

Demostración.

(\Rightarrow) Supongamos que T es una función suprayectiva. Entonces, $\text{Im}(T) = W$, por lo que

$$\begin{aligned}\text{rango}(T) &= \dim(\text{Im}(T)) \\ &= \dim(W) \\ &= n.\end{aligned}$$

(\Leftarrow) Se sigue del Corolario 3.8. □

El siguiente resultado nos dice cómo encontrar un conjunto generador de la imagen de una transformación lineal que tiene como dominio a un espacio vectorial de dimensión finita a partir de una base de su dominio. Aplicando el Teorema 3.9 a dicho conjunto generador podemos obtener una base de la imagen, por lo que este resultado nos servirá para calcular el rango de transformaciones lineales de este tipo.

Teorema 4.4 Sean V, W espacios vectoriales y $T : V \rightarrow W$ una transformación lineal, con V un espacio de dimensión finita n y $B = \{\vec{b}_1, \vec{b}_2, \dots, \vec{b}_n\} \subseteq V$ una base de V . Entonces, tenemos que

$$\langle \{T(\vec{b}_1), T(\vec{b}_2), \dots, T(\vec{b}_n)\} \rangle = \text{Im}(T),$$

es decir, el conjunto $T(B) := \{T(\vec{b}_1), T(\vec{b}_2), \dots, T(\vec{b}_n)\}$ es un conjunto generador de $\text{Im}(T)$.

Demostración. Como $B \subseteq V$, por definición de imagen tenemos que $T(B) \subseteq W$. Puesto que, por el Ejercicio 20, $\text{Im}(T)$ es un subespacio vectorial de W , entonces de la Proposición 2.2 se sigue que $\text{Im}(T)$ es un conjunto cerrado por combinaciones lineales. Por ende, $\langle T(B) \rangle \subseteq \text{Im}(T)$.

Por otro lado, sea $\vec{w} \in \text{Im}(T)$. Entonces, existe $\vec{v} \in V$ tal que $T(\vec{v}) = \vec{w}$. Dado que B es una base de V , entonces existen coeficientes $c_i \in K$ para $i \in \{1, 2, \dots, n\}$ tales que

$$\vec{v} = \sum_{i=1}^n c_i \vec{b}_i.$$

Ya que T es lineal, se sigue que

$$\begin{aligned}\vec{w} &= T(\vec{v}) \\ &= T\left(\sum_{i=1}^n c_i \vec{b}_i\right) \\ &= \sum_{i=1}^n T(c_i \vec{b}_i) \\ &= \sum_{i=1}^n c_i T(\vec{b}_i).\end{aligned}$$

Por ende, $\text{Im}(T) \subseteq \langle T(B) \rangle$. Por lo tanto, concluimos que $\langle T(B) \rangle = \text{Im}(T)$. □

Corolario 4.5 Si $T : V \rightarrow W$ es una transformación lineal y V es de dimensión finita, entonces $\text{Im}(T)$ es un subespacio vectorial de W de dimensión finita.

Demostración. Se sigue del Ejercicio 20 y de los Teoremas 4.4 y 3.9. □

Uno de los resultados más importantes para transformaciones lineales entre espacios vectoriales de dimensión finita es el siguiente.

Teorema 4.6 (Fórmula de la dimensión) Sean V, W espacios vectoriales de dimensión finita y $T : V \rightarrow W$ una transformación lineal. Entonces,

$$\dim(V) = \text{rango}(T) + \text{nulidad}(T).$$

Demostración. Sea $\dim(V) = n$. Por el Ejercicio 19 y el Teorema 3.6, sabemos que $\text{Ker}(T)$ es un subespacio vectorial de V de dimensión finita menor o igual a n . Sean $\text{nulidad}(T) = \dim(\text{Ker}(T)) = k$ y $N = \{\vec{b}_1, \vec{b}_2, \dots, \vec{b}_k\}$ una base de $\text{Ker}(T)$. Como $k \leq n$, por el Teorema 3.10 sabemos que N puede extenderse hasta convertirse en una base $B = \{\vec{b}_1, \vec{b}_2, \dots, \vec{b}_k, \vec{b}_{k+1}, \dots, \vec{b}_n\}$ de V . Demostraremos que $T(B \setminus N) := \{T(\vec{b}_{k+1}), T(\vec{b}_{k+2}), \dots, T(\vec{b}_n)\}$ es una base de $\text{Im}(T)$.

Observemos que, dado que V es de dimensión finita y B es una base de V , por el Teorema 4.4, tenemos que

$$\begin{aligned} \text{Im}(T) &= \langle T(B) \rangle \\ &= \langle \{T(\vec{b}_1), T(\vec{b}_2), \dots, T(\vec{b}_k), T(\vec{b}_{k+1}), \dots, T(\vec{b}_n)\} \rangle \\ &= \langle \{\vec{0}_W, \vec{0}_W, \dots, \vec{0}_W, T(\vec{b}_{k+1}), \dots, T(\vec{b}_n)\} \rangle \quad (N \subseteq \text{Ker}(T)) \\ &= \langle \{T(\vec{b}_{k+1}), T(\vec{b}_{k+2}), \dots, T(\vec{b}_n)\} \rangle \\ &= \langle T(B \setminus N) \rangle, \end{aligned}$$

por lo que $T(B \setminus N)$ genera a $\text{Im}(T)$. Ahora, supongamos que existen coeficientes $c_{k+1}, c_{k+2}, \dots, c_n \in K$ tales que

$$\sum_{i=k+1}^n c_i T(\vec{b}_i) = \vec{0}_W.$$

Entonces, por linealidad de T , se sigue que

$$T\left(\sum_{i=k+1}^n c_i \vec{b}_i\right) = \vec{0}_W,$$

por lo que $\sum_{i=k+1}^n c_i \vec{b}_i \in \text{Ker}(T)$. Como N es una base de $\text{Ker}(T)$, deben existir coeficientes $c_1, c_2, \dots, c_k \in K$ tales que

$$\sum_{i=1}^k c_i \vec{b}_i = \sum_{j=k+1}^n c_j \vec{b}_j,$$

lo que implica que

$$\sum_{i=1}^k c_i \vec{b}_i - \sum_{j=k+1}^n c_j \vec{b}_j = \sum_{i=1}^k c_i \vec{b}_i + \sum_{j=k+1}^n (-c_j) \vec{b}_j = \vec{0}_W.$$

Dado que B es una base de V , tenemos que $c_i = 0$ para todo $i \in \{1, 2, \dots, n\}$. En particular, $c_i = 0$ para todo $i \in \{k+1, k+2, \dots, n\}$, de donde se sigue que $T(B \setminus N)$ es linealmente independiente. Por lo tanto, $T(B \setminus N)$ es una base de $\text{Im}(T)$. Por ende,

$$\begin{aligned} \text{rango}(T) &= \dim(\text{Im}(T)) \\ &= |B \setminus N| \\ &= n - k, \end{aligned}$$

de donde concluimos que

$$\dim(V) = \text{rango}(T) + \text{nulidad}(T).$$

□

Con la ayuda de la fórmula de la dimensión, podemos obtener una caracterización para saber cuándo una transformación lineal entre espacios vectoriales de la misma dimensión es biyectiva, lo cual será de suma importancia más adelante.

Teorema 4.7 (Caracterización de transformaciones lineales biyectivas) Sean V, W espacios vectoriales de la misma dimensión finita y $T : V \rightarrow W$ una transformación lineal. Entonces, las siguientes condiciones son equivalentes.

- (a) T es inyectiva.
- (b) T es suprayectiva.
- (c) T es biyectiva.

Demostración.

En esta demostración, utilizaremos las Propositiones 4.2 y 4.3, así como el Teorema 4.6 (Fórmula de la dimensión).

(a) \Rightarrow (b) Supongamos que T es inyectiva. Entonces, $\text{nulidad}(T) = 0$. Luego, por la fórmula de la dimensión, tenemos que $\text{rango}(T) = n$, de donde se sigue que T es suprayectiva.

(b) \Rightarrow (c) Supongamos que T es suprayectiva. Entonces, $\text{rango}(T) = n$. Luego, por la fórmula de la dimensión, tenemos que $\text{nulidad}(T) = 0$, de donde se sigue que T es inyectiva. Como T es inyectiva y suprayectiva, entonces biyectiva.

(c) \Rightarrow (a) Se sigue de la definición de función biyectiva. □