Curso de Álgebra Lineal
(2022-I)

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

Notación

$\mathbf{u}, \mathbf{v}, \mathbf{w}, \ldots$	vectores (elementos de un conjunto vectorial V)
$a, b, c, \ldots$	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]
$(x_1, \ldots, x_n)$	coordenada como n-tupla
$\begin{pmatrix} x_1 & \ldots & x_n \end{pmatrix}$	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 $\mathbf{v}$
$\overline{a}$	complejo conjugado de a
$u_i v_i \equiv \sum_{i=1}^{n} u_i v_i$	*notación de Einstein* para la suma sobre un índice i
$\vert\vert\mathbf{u}\vert\vert$	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
$l.i.$	Conjunto linealmente independiente
$l.d.$	Conjunto linealmente dependiente
$\dim(V)$	Dimensión del espacio vectorial V

[a] Algunos 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 algebráicas, 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 algebráicas durante el curso. Antes, discutiremos brevemente qué constituye una estructura algebráica y repasaremos la definición de la estructura de campo —necesaria para definir la de espacio vectorial.

Estructuras algebráicas

¿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 algebráicas.

Definición de estructura algebráica y propiedades de sus operaciones

Antes de definir una estructura algebráica, 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

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

para simplificar la notación.

Def. Una *estructura algebráica* 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 algebráica formada por un conjunto A, operaciones $\star, *$ y una relación $\sim$, como $(A, \star, *, \sim)$. 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 $\star : A \times A \to A$. 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 algebráica.

Ejemplos de estructuras algebráicas

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

Ejercicio 1. Consideremos la estructura algebráica $(\mathscr{P}(A), \cup, \cap, \subseteq)$, 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 algebráica $(\mathbb{N}, +, \leq)$. 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 algebráica $(\mathbb{N}, +, \leq)$, 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 algebráica $(\mathbb{Z}, +, \cdot)$. 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 algebráica 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 algebráicas no tienen relaciones.

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

Proposición 1.2 Sean $(A, \star)$ una estructura algebráica 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 algebráica 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*:

$\forall\ a,b,c \in K$	$a+(b+c)=(a+b)+c \ \ \&\ \ a\cdot(b\cdot c)=(a\cdot b)\cdot c$	Asociatividad
$\forall\ a,b \in K$	$a+b=b+a \ \ \&\ \ a\cdot b=b\cdot a$	Conmutatividad
$\exists\ 0,1 \in K$ t.q., $\forall\ a \in K$,	$a+0=a \quad \& \quad 1\cdot a=a$	Identidades (neutros)[a]
$\forall\ a \in K$	$\exists\ -a \in K$ t.q. $a+(-a)=0$	Inversos aditivos[b]
$\forall\ a \neq 0 \in K$	$\exists\ a^{-1} \in K$ t.q. $a\cdot a^{-1}=1$	Inversos multiplicativos[c]
$\forall\ a,b,c \in K$	$a\cdot(b+c)=a\cdot b+a\cdot c$	Distributividad[d].

[a] Por el inciso (a) de la Proposición 1.2, sabemos que los elementos identidad de operaciones binarias en estructuras algebráicas 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.

[b] Por 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 $-a$.

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

[d] Observemos que podemos escribir esta propiedad de forma resumida, pues $+$ y $\cdot$ son operaciones conmutativas.

Observación 1.3 En otras palabras, una estructura algebráica $(K,+,\cdot)$ 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**;

[1] En 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 sólamente 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.

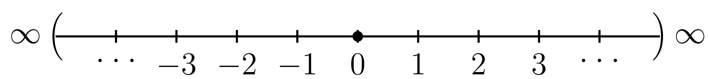

Figura 1: Representación del campo real con la recta real.

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

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

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

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

[a] Gauss 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=T647CGsuOVU.

Apoyándonos en las operaciones del campo real, podemos definir la suma y multiplicación entre números complejos como[2].

$$(a + ib) + (q + ir) := (a + q) + i(b + r),$$
$$(a + ib)(q + ir) := (aq - br) + i(ar + bq).$$

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

[2] Nótese 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 algebráicas* —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*:

$\forall\ \mathbf{u}, \mathbf{v} \in V \ \ \exists\ \mathbf{u} + \mathbf{v} \in V$	Cerradura de la adición
$\forall\ \mathbf{v} \in V, a \in K \ \ \exists\ a\mathbf{v} \in V$	Cerradura del producto de un vector por un escalar
$\forall\ \mathbf{u}, \mathbf{v}, \mathbf{w} \in V \ \ \mathbf{u} + (\mathbf{v} + \mathbf{w}) = (\mathbf{u} + \mathbf{v}) + \mathbf{w}$	Asociatividad de la adición
$\forall\ \mathbf{u}, \mathbf{v} \in V \ \ \mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$	Conmutatividad de la adición
$\exists\ \mathbf{0} \in V$ t.q. $\mathbf{v} + \mathbf{0} = \mathbf{v} \ \ \forall\ \mathbf{v} \in V$	Elemento identidad de la adición (neutro aditivo)
$\forall\ \mathbf{v} \in V \ \ \exists\ -\mathbf{v} \in V$ t.q. $\mathbf{v} + (-\mathbf{v}) = \mathbf{0}$	Elemento inverso de la adición (inverso aditivo)
$a(b\mathbf{v}) = (ab)\mathbf{v} \ \ \forall a, b \in K, \mathbf{v} \in V$	Compatibilidad del producto de un vector por un escalar con el producto entre escalares
$\exists\ 1 \in K$ t.q. $1\mathbf{v} = \mathbf{v} \ \ \forall\ \mathbf{v} \in V$	Elemento identidad del producto de un vector por un escalar
$a(\mathbf{v}+\mathbf{w}) = a\mathbf{v}+a\mathbf{w} \ \ \forall\ \mathbf{v}, \mathbf{w} \in V, a \in K$	Distributividad del producto de un vector por un escalar con respecto a la adición vectorial
$(a + b)\mathbf{v} = a\mathbf{v} + b\mathbf{v} \ \ \forall\ a, b \in K, \mathbf{v} \in V$	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 $+ : V \times V \to V$ (suma vectorial) y $\cdot : K \times V \to V$ (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 sólamente 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 $+ : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}^2$ como

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

y una operación $\cdot : \mathbb{R} \times \mathbb{R}^2 \to \mathbb{R}^2$ 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 $+ : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n$ y $\cdot : \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}^n$ 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 $f(x) = c_1x^1 + c_2x^2 + ... + c_nx^n,\ c_i \in \mathbb{R},\ i \in \{1, 2, ..., n\}$) 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 vectorial[3]. 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.

[3] En 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 $a\mathbf{0} + a0 = a\mathbf{0} + \mathbf{0}$. 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 interpretacción 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 reales[4] 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.

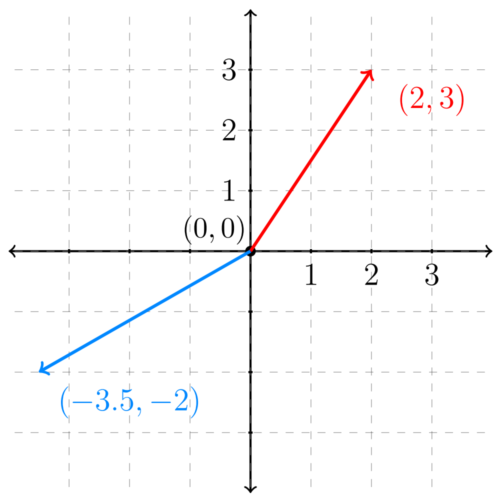

Figura 2: Ejemplo de representación de vectores en el plano cartesiano. Los vectores $(-3.5, -2)$ y $(2, 3)$ son representados por flechas que tienen su cola en el origen del plano cartesiano y su punta en las coordenadas correspondientes.

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

[4] Es decir, con $a, b \in \mathbb{R}$.

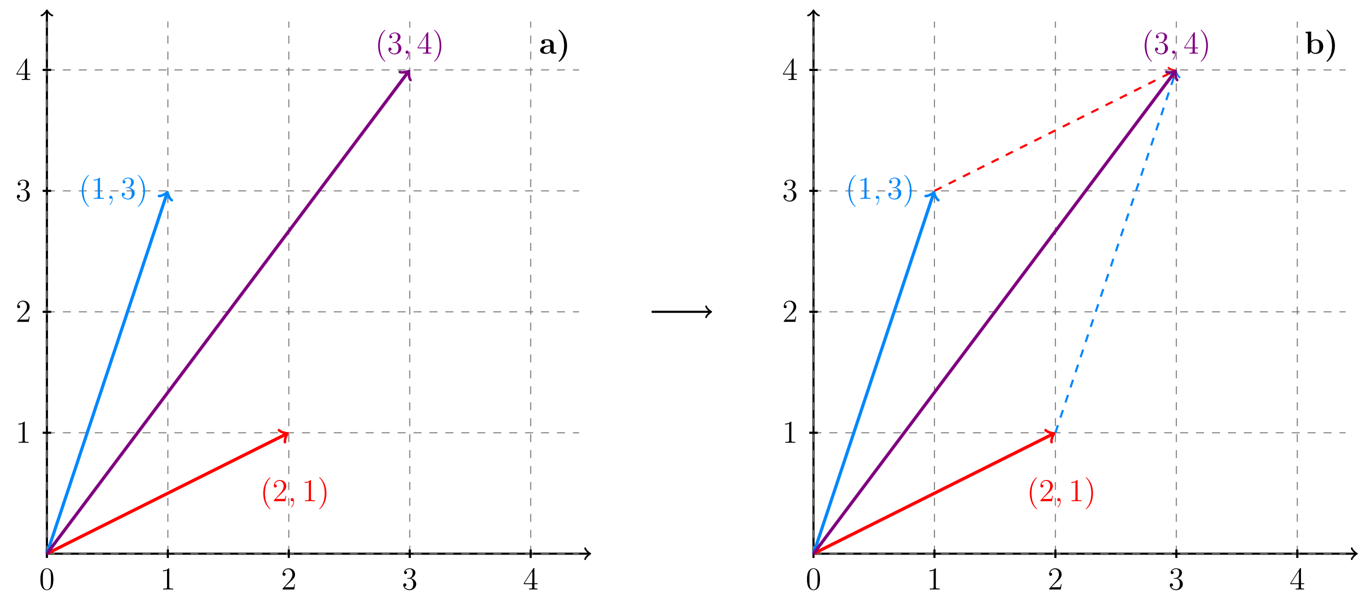

Figura 3: Interpretación geométrica de la suma vectorial en el espacio vectorial real $\mathbb{R}^2$. En la subfigura **a)** se observan los vectores $(1, 3)$ y $(2, 1)$, así como el vector resultante de la suma de los dos anteriores, $(3, 4)$. En la subfigura **b)** observamos la llamada *Ley del paralelogramo*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

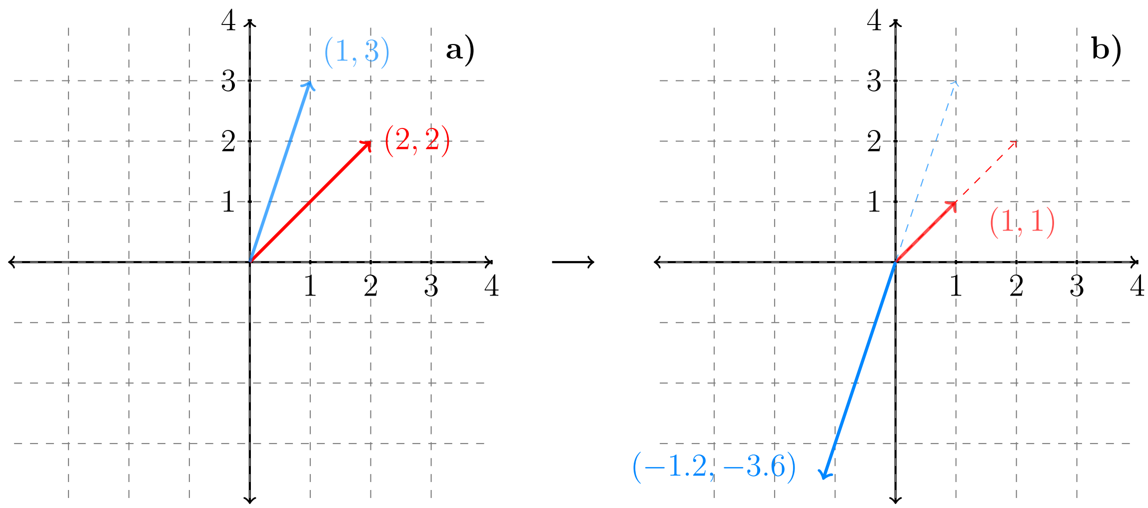

Figura 4: Interpretación geométrica del producto de un vector por un escalar en el espacio vectorial real $\mathbb{R}^2$. Comparando las subfiguras **a)** y **b)** observamos que, en caso de que se multiplique a un vector de $\mathbb{R}^2$ por un escalar de $\mathbb{R}$, es posible que la longtitud del 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 algebráicamente 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 sentido[5]. 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 geométricamente 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 $\big(a + ib\big)$ 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 $\big(a + ib\big) + \big(c + id\big) := \big((a + c) + (b + d)i\big)$ 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 $\big(1 + 2i\big) + \big(3 + 2i\big) = \big(4 + 4i\big)$, 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 $(q+ir)\,\big(s + it\big) := \big((qs - rt) + i(qt + rs)\big)$. 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 $(q)\,\big(s + it\big) := \big((qs) + (qt)i\big)$, por lo cual geométricamente 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 $\big(2 + 2i\big)$. Al hacer el producto de este vector

[5] Alternativamente, 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.

[6] Los 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.

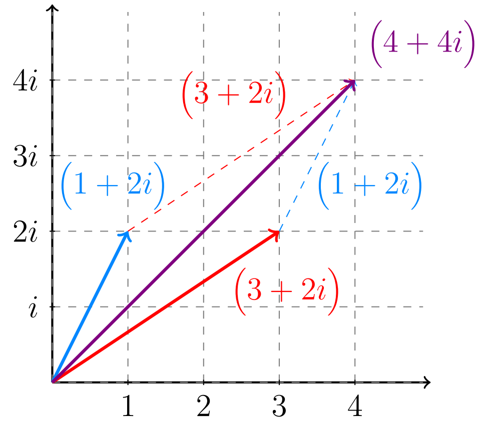

Figura 5: Interpretación geométrica de la suma vectorial en el espacio vectorial complejo $\mathbb{C}$. Observamos que, al igual que en el caso del espacio vectorial real $\mathbb{R}^2$, se cumple la *Ley del paralelogramo*.

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

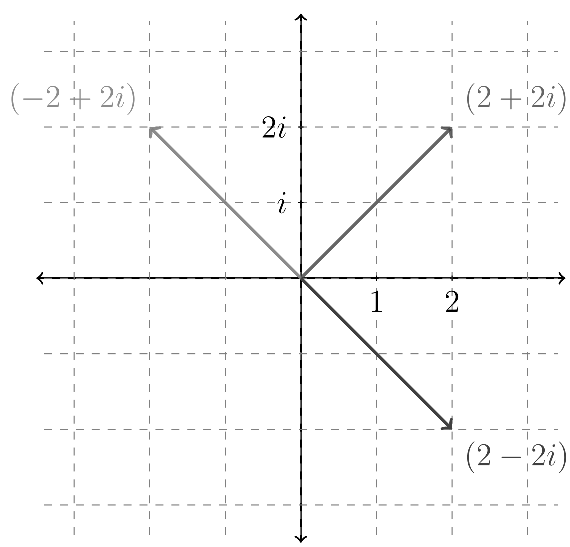

Figura 6: Interpretación geométrica del producto de un vector complejo por los números imaginarios i y $-i$. En este caso, nuestro vector base es $(2+2i)$. El producto de este vector por el escalar i resulta en el vector $(-2+2i)$, lo cual puede ser interpretado geométricamente como una rotación discreta de $\frac{\pi}{2}$ radianes. Así observamos que, en cambio, el producto de nuestro vector base $(2+2i)$ por $-i$ se puede interpretar geométricamente 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 $ai\left(b + ic\right) = \left(-ac + i(ab)\right) = a\left(-c + ib\right) = a(i\left(b + ic\right))$ —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 $r = +\sqrt{x^2 + y^2} = +\sqrt{(x + iy)(x - iy)} = +\sqrt{z\overline{z}} = |z|$, 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,

$$\begin{aligned} z_1z_2 &= |z_1|(\cos\theta_1 + i\sin\theta_1)|z_2|(\cos\theta_2 + i\sin\theta_2) \\ &= |z_1||z_2|(\cos\theta_1\cos\theta_2 + i\cos\theta_1\sin\theta_2 + i\sin\theta_1\cos\theta_2 - \sin\theta_1\sin\theta_2) \\ &= |z_1||z_2|\Big(\cos\theta_1\cos\theta_2 - \sin\theta_1\sin\theta_2 + i(\cos\theta_1\sin\theta_2 + \cos\theta_2\sin\theta_1)\Big). \end{aligned}$$

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

$$z_1z_2 = |z_1||z_2|(\cos(\theta_1 + \theta_2) + i\sin(\theta_1 + \theta_2)).$$

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 sólamente 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 vectorial[7] 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. $\square$

[7] De 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

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

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 $L = \{\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_n\} \subseteq V$ 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 $i \in \{1, 2, ..., n\}$ tales que

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

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 $L = \{\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_n\} \subseteq V$ un conjunto finito de vectores de V. Demuestra que el conjunto de todas las combinaciones lineales de L

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

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

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

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 $\{\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_n\}$ simplemente como $\langle \mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_n \rangle$ en vez de $\langle \{\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_n\} \rangle$, 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

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

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

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

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

Para dar algunos ejemplos, si elegimos cualquier vector $\mathbf{v} \in \mathbb{R}^2$, entonces el subespacio generado correspondiente $\langle \mathbf{v} \rangle = \{c\mathbf{v} \,|\, c \in \mathbb{R}\}$ 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 $N = \{\begin{pmatrix}1 & 0 & 0\end{pmatrix}, \begin{pmatrix}0 & 1 & 0\end{pmatrix}, \begin{pmatrix}0 & 0 & 1\end{pmatrix}\}$ entonces vemos que

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

pero esto es equivalente a la definición $\mathbb{R}^3 = \{\begin{pmatrix}c_1 & c_2 & c_3\end{pmatrix} \,|\, c_1, c_2, c_3 \in \mathbb{R}\}$; 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, $\langle S \rangle = \langle S' \rangle$. □

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.

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

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 $\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_n \in V$. Decimos que los vectores $\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_n$ 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,

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

En cambio, decimos que $\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_n$ 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].

[a] Observemos 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 $\mathbf{v} = c_1\mathbf{u}_1 + ... + c_n\mathbf{u}_n$ en donde los vectores son dados pero los coeficientes son desconocidos, éstos útlimos se vuelven las *incógnitas* del *sistema de ecuaciones algebráicas* dado por la ecuación $\mathbf{v} = c_1\mathbf{u}_1 + ... + c_n\mathbf{u}_2$. 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 $\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_m \in V$, 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 $\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_m$ y obtener, como resultado final, a $\mathbf{v}_n$. Adicionalmente, en el espacio vectorial complejo $\mathbb{C}$, también se podrían estar rotando los vectores $\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_m$ —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 $\big(n_1 + i(n_2)\big)$ y $\big(n_3 + i(n_4)\big)$ 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 $\begin{pmatrix} n_1 & n_4 & n_7 \end{pmatrix}$, $\begin{pmatrix} n_2 & n_5 & n_8 \end{pmatrix}$ y $\begin{pmatrix} n_3 & n_6 & n_9 \end{pmatrix} \in \mathbb{R}^3$. ¿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 $\mathbf{v_1}, \mathbf{v_2}, ..., \mathbf{v}_n$ vectores linealmente independientes de V. Sean $c_1, c_2, ..., c_n \in K$ y $d_1, d_2, ..., d_n \in K$ tales que

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

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

Demostración.

$$c_1\mathbf{v_1} + c_2\mathbf{v}_2 + ... + c_n\mathbf{v}_n = d_1\mathbf{v}_1 + d_2\mathbf{v}_2 + ... + d_n\mathbf{v}_n \iff (c_1 - d_1)\mathbf{v}_1 + (c_2 - d_2)\mathbf{v}_2 + ... + (c_n - d_n)\mathbf{v}_n = \mathbf{0}.$$

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

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

$\square$

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 $\mathbf{u} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + ... + c_n\mathbf{v}_n$ y $\{\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_n\}$ es $l.i.$ entonces no existe otra combinación de escalaras 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 $S_1 = \{\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_k\}$ y $S_2 = \{\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_n\}$ con $k \leq n$. Ya que por hipótesis S_2 es *l.i.*, por definición se cumple que

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

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

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

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 $L = \{\mathbf{u}_1, ..., \mathbf{u}_n\}$. Supongamos que $\mathbf{v} \in \langle L \rangle$, entonces existen coeficientes $c_i \in K$ tales que $\mathbf{v} = c_1\mathbf{u}_1 + ... + c_n\mathbf{u}_n$. Despejando esta ecuación, obtenemos que $c_1\mathbf{u}_1 + ... + c_n\mathbf{u}_n + (-1)\mathbf{v} = \mathbf{0}$, 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., $c_1\mathbf{u}_2 + ... + c_n\mathbf{u}_n + b\mathbf{v} = \mathbf{0}$ 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 $c_1\mathbf{u}_1 + ... + c_n\mathbf{u}_n = \mathbf{0}$; 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 $c_1\mathbf{u}_1 + ... + c_n\mathbf{u}_n + b\mathbf{v} = \mathbf{0}$ y obtener que $\mathbf{v} = \frac{-c_1}{b}\mathbf{u}_1 + ... + \frac{-c_n}{b}\mathbf{u}_n$, 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 espacio[8]. 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,

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

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 espacio[9].

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 $\{\begin{pmatrix}2 & 0\end{pmatrix}, \begin{pmatrix}0, & -3\end{pmatrix}\}, \{\begin{pmatrix}3 & 3\end{pmatrix}, \begin{pmatrix}-3 & 3\end{pmatrix}\}$ y $\{\begin{pmatrix}1 & 0\end{pmatrix}, \begin{pmatrix}0 & 1\end{pmatrix}\}$ son bases de $\mathbb{R}^2$.

Cualquier conjunto de la forma $\{c_n x^n \,|\, n \in \mathbb{N} \cup \{0\}, c_n \in \mathbb{R}\}$ 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.

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

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

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

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 $\{\mathbf{v}_1, ..., \mathbf{v}_m\}$ es l.i.). Suponiendo, por ejemplo, que $b_1 \neq 0$, tenemos que

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

por lo cual $\mathbf{u}_1$ puede ser expresado como combinación lineal de los vectores $\mathbf{v}_1, ..., \mathbf{v}_m, \mathbf{u}_2, ..., \mathbf{u}_{n-m}$. Sea $H = \{\mathbf{u}_2, ..., \mathbf{u}_{n-m}\}$, entonces $L \cup H = \{\mathbf{v}_1, ..., \mathbf{v}_{m+1}, \mathbf{u}_2, ..., \mathbf{u}_{n-m}\}$ y trivialmente tenemos que $\mathbf{v}_1, ..., \mathbf{v}_m, \mathbf{u}_2, ..., \mathbf{u}_{n-m} \in \langle L \cup H \rangle$ —lo cual también implica que $\mathbf{u}_1 \in \langle L \cup H \rangle$. Por lo tanto, tenemos que $\{\mathbf{v}_1, ..., \mathbf{v}_m, \mathbf{u}_1, ..., \mathbf{u}_{n-m}\} \subseteq \langle L \cup H \rangle$. Recordando que por hipótesis de inducción $\{\mathbf{v}_1, ..., \mathbf{v}_m, \mathbf{u}_1, ..., \mathbf{u}_{n-m}\}$ 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 linelmente dependiente?

[10]Recuerda 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 $B = \{\mathbf{b}_1, \mathbf{b}_2, ..., \mathbf{b}_n\}$ y $B' = \{\mathbf{b'}_1, \mathbf{b'}_2, ..., \mathbf{b'}_{n'}\}$ 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,

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

$$...$$

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

donde $c_{ij} \in K$.

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

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

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

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

$$...$$

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

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 $(d_1, d_2, ..., d_{n'})$. 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].

[a] En 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 $\mathbf{e}_1 = \begin{pmatrix} 1 & 0 & 0 & 0 & ... & 0 \end{pmatrix}$, $\mathbf{e}_2 = \begin{pmatrix} 0 & 1 & 0 & 0 & ... & 0 \end{pmatrix}$, ..., $\mathbf{e}_n = \begin{pmatrix} 0 & 0 & 0 & 0 & ... & 1 \end{pmatrix}$. Demuestra que $\{\mathbf{e}_1, \mathbf{e}_2, ..., \mathbf{e}_n\}$ es una base para (K^n, K) y que, por lo tanto, $\dim(K^n, K) = n$.

Observemos que esta definición *algebráica* 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 (algebráica) 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 $\mathbf{x}_2, ..., \mathbf{x}_k$ de W al conjunto $\{\mathbf{x}_1\}$ de tal forma que $\{\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_k\}$ 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 $\dim(W) = k \leq n$. □

En los teoremas anteriores demostramos que si $\dim(V) = n$ entonces cualquer 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 $B = \{\mathbf{b}_1, ..., \mathbf{b}_n\} \subset V$ 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 $B' \equiv \{\mathbf{b}_1, ..., \mathbf{b}_n, \mathbf{b}_{n+1}\}$, 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. □

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. □

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. □

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. □

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 algebráicas 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 algebráicas son aquellas que preservan la estructura en cuestión. En el caso particular de los espacios vectoriales, una función que preserva 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.

Transformaciones lineales

Def. Sean K un campo, S un subcampo de K[a] y $(V,S),(W,K)$ espacios vectoriales. Entonces, decimos que una función $T:V\to W$ 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*.

[a] En 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 $T:V\to W$ 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 transformacines 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 $T:V\to W$ 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 $T:\mathbb{R}^2\to\mathbb{R}^2$ 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 $\bar{\cdot} : \mathbb{C} \to \mathbb{C}$ abre sumas pero no es una transformación lineal.

Sea $C^\infty(\mathbb{R},\mathbb{R})$ el espacio de funciones reales de variable real continuas con todas sus derivadas continuas. Entonces la derivada $\frac{d}{dx} : C^\infty(\mathbb{R},\mathbb{R}) \to C^\infty(\mathbb{R},\mathbb{R})$ 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.

Para cualquier espacio vectorial V, la función identidad $\mathrm{Id}_V : V \to V$ es una transformación lineal.

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 $T : V \to V$ 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.

Para cualesquiera espacios vectoriales V y W, la transformación nula $T_0 : V \to W, T_0(\vec{u}) = \vec{0}_W$ para todo $\vec{u} \in V$ es una transformación lineal.

Las rotaciones en $\mathbb{R}^2, \mathbb{R}^3$ son transformaciones lineales[11].

La función $T : \mathbb{R}^2 \to \mathbb{R}^3$ dada por $T(x,y) = (x,y,0)$ es una transformación lineal. Similarmente, la función $U : \mathbb{R}^3 \to \mathbb{R}^2$ dada por $U(x,y,z) = (x,y)$ es una transformacióm 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 $\frac{d}{dx} : C^n(\mathbb{R},\mathbb{R}) \to C^{n-1}(\mathbb{R},\mathbb{R})$ 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.

[11] Puedes convencerte de esto haciendo un simple dibujo.

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

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

y la *imagen* de T como[a]

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

[a]Observemos que esto es simplemente la definición usual de la imagen de una función.

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

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

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

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

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

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

En particular, si W tiene dimensión finita, del Teorema 3.6 se sigue que $\mathrm{Im}(T)$ 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 $T : V \to W$ una transformación lineal. Entonces, definimos la *nulidad* de T como

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

y el *rango* de T como

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

Algunos ejemplos de esto están dados por los siguientes resultados.

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

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 $\vec{v} \in \mathrm{Ker}(T)$, se sigue que $\mathrm{Ker}(T) = \{\vec{0}_V\}$. Como por definición $\varnothing$ es una base de $\{\vec{0}_V\}$, se sigue que $\mathrm{nulidad}(T) = 0$.

($\Leftarrow$) Supongamos que 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 $T : V \to W$ una transformación lineal. Entonces T es una función suprayectiva si, y sólo si, 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 $T : V \to W$ una transformación lineal, con V un espacio de dimensión finita n y $B = \{\vec{b}_1, \vec{b}_2, ..., \vec{b}_n\} \subseteq V$ una base de V. Entonces, tenemos que

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

es decir, el conjunto $T(B) := \{T(\vec{b}_1), T(\vec{b}_2), ..., T(\vec{b}_n)\}$ 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 $i \in \{1, 2, ..., n\}$ 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, $\mathrm{Im}(T) \subseteq \langle T(B)\rangle$. Por lo tanto, concluimos que $\langle T(B)\rangle = \mathrm{Im}(T)$. $\square$

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

Demostración. Se sigue del Ejercicio 20 y de los Teoremas 4.4 y 3.9. $\square$

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 $T : V \to W$ una transformación lineal. Entonces,

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

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

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

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

por lo que $T(B \setminus N)$ genera a $\mathrm{Im}(T)$. Ahora, supongamos que existen coeficientes $c_{k+1}, c_{k+2}, ..., c_n \in K$ 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 $\sum_{i=k+1}^{n} c_i\vec{b}_i \in \mathrm{Ker}(T)$. Como N es una base de $\mathrm{Ker}(T)$, deben existir coeficientes $c_1, c_2, ..., c_k \in K$ 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 $i \in \{1, 2, ..., n\}$. En particular, $c_i = 0$ para todo $i \in \{k+1, k+2, ..., n\}$, de donde se sigue que $T(B \setminus N)$ es linealmente independiente. Por lo tanto, $T(B \setminus N)$ es una base de $\mathrm{Im}(T)$. Por ende,

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

de donde concluimos que

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

□

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 n y $T : V \to W$ 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 Proposiciones 4.2 y 4.3, así como el Teorema 4.6 (Fórmula de la dimensión).

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

$(b) \Rightarrow (c)$ Supongamos que T es suprayectiva. Entonces, $\mathrm{rango}(T) = n$. Luego, por la fórmula de la dimensión, tenemos que $\mathrm{nulidad}(T) = 0$, 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. □

Interacciones entre transformaciones lineales

Suma y reescalamiento de transformaciones lineales

Def. Sean $(V,S),(W,K)$ espacios vectoriales, con S un subcampo de K. Definimos el conjunto de transformaciones lineales de V a W como

$$\mathcal{L}(V,W) := \{T : V \to W \mid T(\vec{u}+a\vec{v}) = T(\vec{u}) + aT(\vec{v})\ \forall\ \vec{u},\vec{v} \in V, a \in K\}.$$

Sean $T_1, T_2 \in \mathcal{L}(V,W)$ y $a \in K$. Definimos la *suma de transformaciones lineales* T_1 y T_2 a través de su regla de correspondencia como

$$(T_1+T_2)(\vec{u}) = T_1(\vec{u}) + T_2(\vec{u}) \quad \forall\ \vec{u} \in V;$$

análogamente, definimos el *producto de la transformación lineal* (o *reescalamiento de*) T_1 *por el escalar* a como

$$(aT_1)(\vec{u}) = a\Big(T(\vec{u})\Big) \quad \forall\ \vec{u} \in V.$$

Observación 4.8 Como podrás sospechar a partir de las definiciones anteriores, el conjunto de transformaciones lineales entre dos espacios vectoriales tiene una estructura que debería resultarles familiar.

Ejercicio 21. Sean $(V,S),(W,K)$ espacios vectoriales, con S un subcampo de K. Demuestra que $\mathcal{L}(V,W)$, con las operaciones de suma de transformaciones lineales y producto de una transformación lineal por un escalar, forma un espacio vecorial sobre el campo K.

Por ende, dado que por definición las transformaciones lineales son compatibles con las operaciones esenciales de los espacios vectoriales (i.e., preservan la estructura de espacio vectorial) y que sus contradominios tienen la estructura de espacio vectorial, ¡el conjunto de todas las transformaciones lineales de un espacio a otro *hereda* la estructura de espacio vectorial!

Composición de transformaciones lineales

Def. Sean $T : V \to W$ y $U : W \to Z$ transformaciones lineales. Definimos la *composición de transformaciones lineales* $U \circ T : V \to Z$ a través de su regla de correspondencia como

$$(U \circ T)(\vec{u}) = U(T(\vec{u})) \quad \forall\ \vec{u} \in V.$$

La expresión de la composición $U \circ T$ se puede leer como "U *después de* T" para recordar cuál de las funciones se debe aplicar primero.

Observación 4.9

(1) La composición de transformaciones lineales, siempre que tenga sentido[12], es una transformación lineal:

Ejercicio 22. Sean V,W,Z espacios vectoriales y $T \in \mathcal{L}(V,W), U \in \mathcal{L}(W,Z)$. Demuestra que $U \circ T \in \mathcal{L}(V,Z)$.

[12] Esto es, siempre que el contradominio de la que se aplica primero sea un subespacio vectorial del dominio de la que se aplica después.

Sin embargo, la operación de composición de transformaciones lineales se distingue de la de suma dado que, en general, la transformación resultante de aplicar la operación de composición tiene un dominio y/o contradominio *distinto* al de las transformaciones originales.

(2) En el caso particular en que todos los dominios y contradominios de las transformaciones lineales que se componen son el mismo espacio vectorial V, entonces el resultado de la composición es un operador lineal en V. Es decir, la operación de composición de transformaciones lineales en $\mathcal{L}(V, V)$ es cerrada; equivalentemente, *la composición de operadores lineales sobre un mismo espacio vectorial es una operación binaria*.

Ejercicio 23. Sean (V, K) y (W, K) espacios vectoriales y $T : V \to W$ una transformación lineal. Supongamos que existe $T^{-1} : W \to V$. Demuestra que T^{-1} es una transformación lineal.

Def. Sea $T : V \to W$ una transformación lineal. Si T es invertible[a], decimos que es un *isomorfismo de espacios vectoriales*, o simplemente un *isomorfismo*.

[a] Equivalentemente, podemos pedir que T sea inyectiva.

5. Transformaciones lineales y sus representaciones matriciales

6. Composición de transformaciones lineales y multiplicación de matrices

7. Invertibilidad e isomorfismos

8. Matriz de cambio de base

9. Eigenvectores y eigenvalores

10. Diagonalizabilidad

11. Introducción a sistemas lineales de ecuaciones diferenciales

Esta pequeña sección de las notas tiene como finalidad introducir algunos de los conceptos que utilizaremos durante la segunda mitad del tercer módulo del curso, para el cual seguiremos principalmente el libro *Linear Algebra: A Modern Introduction* de Poole (págs. 340-348).

Sistemas lineales de ecuaciones algebráicas

Cuando hablamos de un *sistema de ecuaciones*, generalmente nos referimos a un grupo de ecuaciones que deben cumplirse simultáneamente, por ejemplo,

$$x_1 + x_2 = 1,$$

$$x_1 - x_2 = 5.$$

Para este sistema de ecuaciones, decimos que tenemos dos *incógnitas*, x_1 y x_2, que queremos encontrar.

Como sabemos, podemos formar sistemas con más ecuaciones y/o incógnitas. En general, podemos expresar un sistema de m ecuaciones y n incógnitas como

$$f_1(x_1, x_2, ..., x_n) = 0$$

$$f_2(x_1, x_2, ..., x_n) = 0$$

$$...$$

$$f_m(x_1, x_2, ..., x_n) = 0,$$

donde f_j, con $1 \leq j \leq m$, es una función arbitraria de las incógnitas x_i, donde $1 \leq i \leq n$. Decimos que tenemos una *solución* al sistema cuando encontramos los valores de las incógnitas x_i para los cuales todas las ecuaciones se verifican simultáneamente. Si reordenamos las ecuaciones del sistema que aparece al inicio de esta sección, podemos ver que en ese caso $f_1(x_1, x_2) = x_1 + x_2 - 1$ y $f_2(x_1, x_2) = x_1 - x_2 - 5$, y que la solución al sistema es $x_1 = 3, x_2 = -2$. Observemos que $m = 2$ y $n = 2$ en este caso; precisamente por esto decimos que es un sistema de dos ecuaciones con dos incógnitas.

Por lo regular, trabajamos con sistemas de ecuaciones en donde las funciones f_j son *polinomios* en las variables x_i; a este tipo de ecuaciones se les llama *algebráicas* y, por lo tanto, en esos casos decimos que tenemos un *sistema de ecuaciones algebráicas*, como el ejemplo del inicio de esta sección.

Cuando en una ecuación algebráica las incógnitas no aparecen elevadas a una potencia mayor a uno, decimos que la ecuación es *lineal*. Análogamente, cuando en un sistema de ecuaciones algebráicas las incógnitas no aparecen elevadas a una potencia mayor a uno —o, equivalentemente, cuando todos los polinomios f_j son de grado uno— tenemos un sistema de ecuaciones algebráicas lineales; generalmente, en estos casos decimos que tenemos un *sistema lineal de ecuaciones algebráicas*.

Lo que caracteriza a los sistemas de ecuaciones algebráicas (de los cuales los sistemas lineales de ecuaciones algebráicas son un caso particular) es que sus soluciones pueden ser números enteros, racionales, irracionales o hasta complejos, pero siempre son números; es decir, las *incógnitas* de este tipo de sistemas de ecuaciones —y, por tanto, sus soluciones— son *números*. Esta es la principal diferencia entre este tipo de sistemas y los sistemas de ecuaciones *diferenciales*, que veremos más adelante. Antes de eso, tenemos que entender la diferencia entre las ecuaciones algebráicas y las ecuaciones diferenciales, lo cual haremos en la siguiente sección.

Introducción a ecuaciones diferenciales (ordinarias)

A diferencia de una ecuación algebráica, en una ecuación *diferencial* la incógnita de la ecuación —y, por lo tanto, su solución—es una *función*. Por ejemplo, la ecuación

$$\dot{x} + x = 0,$$

donde $\dot{x}$ representa la derivada de la función x con respecto a alguna variable independiente —usualmente llamada t, aunque esto es sólo una etiqueta[13]—, tiene como solución a la función $x(t) = e^{-t}$, y dicha solución es válida para todo $t \in \mathbb{R}$ ya que en todo este intervalo se verifica que[14]

$$\frac{d}{dt}e^{-t} + e^{-t} = \left(\frac{d}{d(-t)}e^{-t}\right)\left(\frac{d}{dt}(-t)\right) + e^{-t} = \left(e^{-t}\right)\left(-1\right) + e^{-t} = -e^{-t} + e^{-t} = 0.$$

En general, las incógnitas de las ecuaciones diferenciales pueden ser funciones de más de una variable, y en la ecuación pueden aparecer dos o más de sus derivadas parciales; a este tipo de ecuaciones diferenciales les llamamos ecuaciones diferenciales *parciales*. En cambio, cuando nuestra función incógnita depende de una sola variable, decimos que tenemos una ecuación diferencial *ordinaria*; éstas últimas son con las cuales trabajaremos. A continuación, formalizamos algunas definiciones que utilizaremos.

> Def. Una *ecuación diferencial ordinaria* es una relación entre una función incógnita $x(t)$, sus derivadas, y una variable independiente t. En general, podemos expresar una ecuación diferencial ordinaria como
>
> $$f\Big(t, x(t), \dot{x}(t), \ddot{x}(t), \ldots, x^{(n)}(t)\Big) = 0,$$
>
> donde f es una función arbitraria y $x^{(n)}(t)$ denota la n-ésima derivada de la función incógnita x con respecto a su variable independiente t. Decimos que una función $x(t)$ es una *solución* a la ecuación diferencial anterior en un intervalo I si tanto $x(t)$ como sus derivadas cumplen dicha ecuación para todo $t \in I$.

Podemos clasificar algunos tipos de ecuaciones diferenciales ordinarias, como sigue[15]:

[13] Con *etiqueta* nos referimos a que, de haber postulado que la función x dependía de una variable independiente s, entonces la solución a la ecuación diferencial sería $x(s) = e^{-s}$, lo cual no cambia en absoluto el sentido de la ecuación ni de su solución. A este tipo de variables "etiqueta" acostumbramos llamarles *variables mudas*.

[14] Observemos que pudimos haber planteado esta ecuación diferencial de manera equivalente como $\dot{x} = -x$, y la solución sería la misma, ya que $\frac{d}{dt}e^{-t} = -e^{-t}$ para todo $t \in \mathbb{R}$.

[15] Observen la analogía con la forma en que clasificamos ecuaciones algebráicas (polinomiales), pero también las diferencias. Las definiciones presentadas serán relevantes para nuestros fines.

Def. Decimos que una ecuación diferencial ordinaria es:

- *de orden n* si la derivada de mayor grado que aparece en la ecuación es de orden n;
- *lineal* si se puede escribir de la forma

$$a_n(t)x^{(n)}(t) + a_{n-1}(t)x^{(n-1)}(t) + ... + a_1(t)\dot{x}(t) + a_0(t)x(t) + b(t) = 0$$

para funciones $a_1(t), a_2(t), ..., a_n(t), b(t)$ dadas (que pueden, en particular, ser constantes)[a];

- *autónoma* si en la ecuación no aparece la variable independiente t.

[a] En particular, si $b(t)$ es la constante cero, decimos que la ecuación es lineal *homogénea*.

Ecuaciones diferenciales (ordinarias) con condiciones inciales y sistemas lineales de ecuaciones diferenciales

Con regularidad, nos encontramos con ecuaciones diferenciales que tienen más de una solución, e inclusive pueden tener una infinidad de ellas. Por ejemplo, regresando a la ecuación

$$\dot{x} + x = 0,$$

podemos verificar que la función $x(t) = Ce^{-t}$ es una solución en todo el intervalo $\mathbb{R}$ para cualquier $C \in \mathbb{R}$. Decimos entonces que $x(t) = Ce^{-t}$ es una *solución general* a la ecuación diferencial anterior[16]. En esta solución general, el valor C representa un *grado de libertad* de la solución, el cual puede removerse agregando una *restricción* a la ecuación diferencial, la cual viene en forma de una *condición inicial*.

Supongamos que nos piden resolver la ecuación diferencial anterior bajo la condición de que al evaluar la función en $t = 0$ obtenemos un cierto valor $C_0 \in \mathbb{R}$. Es decir que la solución $x(t)$, además de cumplir la ecuación diferencial, debe cumplir que $x(t)|_{t=0} = C_0$. Partiendo de nuestra solución general, resolvemos ahora la última ecuación mostrada como sigue

$$x(t)|_{t=0} = C_0 \implies Ce^{-t}|_{t=0} = C_0 \implies Ce^0 = C_0 \implies C(1) = C_0 \implies C = C_0.$$

Por lo tanto, la solución a la ecuación diferencial $\dot{x} + x = 0$ con condición inicial $x(0) = C_0$, donde $C_0 \in \mathbb{R}$ es un valor determinado, es $x(t) = C_0 e^{-t}$. Ya que en este caso $x(t)$ sigue cumpliendo la ecuación diferencial original pero ya no tiene grados de libertad, decimos que es una *solución particular* a la ecuación diferencial $\dot{x} + x = 0$ sujeta a la *condición inicial* $x(0) = C_0$, con $C_0 \in \mathbb{R}$.

> Def. Decimos que una función $x(t)$ es una *solución particular* a una ecuación diferencial ordinaria de grado uno si, además de cumplir con dicha ecuación, cumple con la *condición inicial*
>
> $$x(t_0) = C_0,$$
>
> donde C_0 es un valor fijo.

De forma análoga a como podemos formar sistemas de ecuaciones a partir de ecuaciones algebráicas, también podemos formarlos a partir de ecuaciones diferenciales. Por ejemplo, el sistema de ecuaciones diferenciales

$$\dot{x_1} - x_2 = 0,$$

$$x_1 + \dot{x_2} = 0,$$

tiene como solución general a $x_1(t) = A\sin(t), x_2(t) = A\cos(t)$, válida en todo el intervalo $\mathbb{R}$ para cualquier $A \in \mathbb{R}$. Observemos que, en este sistema, ninguna de las incógnitas apareció elevada a una potencia mayor a uno. Por lo tanto, este sistema es un ejemplo de un *sistema lineal de ecuaciones diferenciales ordinarias*.

> Def. Decimos que un sistema de ecuaciones diferenciales es un *sistema lineal de ecuaciones diferenciales ordinarias* si todas las ecuaciones diferenciales ordinarias del sistema son lineales en todas sus incógnitas. Usualmente, se les abrevia como *sistemas lineales de ecuaciones diferenciales*.

[16] Por dar otro ejemplo, podemos decir que la ecuación $\ddot{x} + x = 0$ tiene como solución general a $x(t) = A\sin(t) + B\cos(t)$ en todo el intervalo $\mathbb{R}$ para cualesquiera valores $A, B \in \mathbb{R}$ (¡compruébalo!).

Después de esta introducción, puedes continuar leyendo las págs. 273-274 del libro *Linear Algebra* de Friedberg y 340-342 del libro *Linear Algebra: A Modern Introduction* de Poole para ver cómo podemos aplicar herramientas de álgebra lineal (en particular, diagonalización) para resolver este tipo de sistemas. Si deseas repasar el material presente antes de continuar, te sugiero resolver los ejercicios al final de esta sección. Después, deberás resolver los ejercicios $59, 61$ y 63 de la sección 4.6 del Poole.

Ejercicios de repaso

Sistemas lineales de ecuaciones algebráicas

1. Realiza un cuadro comparativo con las definiciones de sistema de ecuaciones, sistema de ecuaciones algebráicas y sistema lineal de ecuaciones algebráicas.

2. Escribe o argumenta por qué no puedes escribir un ejemplo de:

   a) una ecuación *trascendental* (i.e., no algebráica);

   b) un sistema no lineal de ecuaciones algebráicas;

   c) un sistema lineal de ecuaciones no algebráicas.

3. Demuestra que en una ecuación lineal homogénea con más de una incógnita el conjunto de soluciones forma un espacio vectorial. Después, generaliza el resultado para demostrar que en un sistema lineal de ecuaciones algebráicas homogéneas con menos ecuaciones que incógnitas el conjunto de soluciones forma un espacio vectorial.

Introducción a ecuaciones diferenciales (ordinarias)

1. Encuentra la solución general a la ecuación $\dot{x} - x = 0$ en todo el intervalo $\mathbb{R}$.

2. Sea $C^\infty(\mathbb{R})$ el conjunto de todas las funciones reales de variable real con derivadas de cualquier orden y sea $\frac{d}{dt} : C^\infty(\mathbb{R}) \to C^\infty(\mathbb{R})$ el operador lineal *derivada*. Supongamos que $x(t)$ es un eigenvector de $\frac{d}{dt}$ con eigenvalor λ, ¿entonces quién es $x(t)$?

3. Muestra que tanto la ecuación diferencial del ejemplo dado al inicio de la sección 11 como su solución general son un caso particular del ejercicio anterior.

Ecuaciones diferenciales (ordinarias) con condiciones iniciales

1. Encuentra la solución particular a la ecuación $\dot{x} + x = 0$ con la condición inicial $x(0) = 5$.

2. Encuentra la solución particular a la ecuación $\dot{x} + x = 0$ con la condición inicial $x(1) = 5$.

3. Muestra que cualquier solución particular (es decir, sin grados de libertad) de la ecuación $\ddot{x} = g$ debe tener dos condiciones iniciales: una para la función $x(t)$ y otra para $\dot{x}(t)$. ¿De qué ecuación se trata?

Sistemas lineales de ecuaciones diferenciales (ordinarias)

1. Di si el siguiente sistema de ecuaciones diferenciales ordinarias es lineal, autónomo, ambos o ninguno, y explica por qué:

$$\dot{x_1} + 3x_2 = 0,$$

$$\dot{x_2} - x_3 = 0,$$

$$x_1 - 6\dot{x_3} = 0.$$

12. Modelación con sistemas lineales de ecuaciones diferenciales

Para esta sección, seguiremos el libro *Linear Algebra: A Modern Introduction* de Poole, págs. 342-348. Deberás realizar los ejercicios 65, 66, 67, 69, 71, 73 y 75 de la sección 4.6 del Poole, así como el ejercicio 15 de la sección 5.2 del Friedberg.

13. Funcionales y espacio dual, complemento ortogonal y proyecciones ortogonales

Durante este último módulo del curso, nuestro principal objeto de estudio serán los espacios vectoriales con productos escalares, así como los operadores lineales que actúan sobre dichos espacios. Descubriremos que, con ayuda del producto escalar, podemos llegar a tener una comprensión profundamente geométrica sobre cómo actúan ciertos operadores en estos espacios —más allá de una simple descripción algrebráica dada por reglas de correspondencia. Este resultado es uno de los más importantes de todo el curso, y se le conoce como el *Teorema espectral*. Dado que haremos uso frecuente del producto escalar $\langle \cdot, \cdot \rangle$ —principalmente, a partir de la sección 13— en lo que resta del curso utilizaremos la llamada *notación bra-ket*, también conocida como *notación de Dirac*, por lo que se presenta a continuación.

Notación de bra-ket (o de Dirac) para espacios vectoriales con producto escalar:

$\lvert v\rangle$	vector o *ket*
$\langle u\vert v\rangle$	producto escalar o *bra-ket*
$\langle u\rvert$	vector dual (funcional) o *bra*

Los vectores duales (también conocidos como *funcionales* serán definidos formalmente más adelante, con ayuda de esta notación. Una diferencia muy importante con respecto a la notación utilizada anteriormente es que en esta nueva notación, por razones que veremos más adelante, la entrada *derecha* del producto escalar será lineal, mientras que la entrada *izquierda* será antilineal. Es decir, que si $|u\rangle, |v\rangle \in V$ y $\alpha \in K$, donde (V, K) es un espacio vectorial, entonces en general

$$\langle u| \Big(|\alpha v\rangle \Big) = \langle u| \Big(\alpha |v\rangle \Big) = \alpha \langle u|v\rangle$$

y

$$\Big(\langle \alpha u| \Big) |v\rangle = \Big(\overline{\alpha} \langle u| \Big) |v\rangle = \overline{\alpha} \langle u|v\rangle .$$

De las ecuaciones anteriores podemos hacer las identificaciones $|\alpha v\rangle = \alpha |v\rangle$ y $\langle \alpha u| = \overline{\alpha} \langle u|$, las cuales utilizaremos ampliamente durante este módulo.

Con esta notación podemos, por ejemplo, reescribir algunos resultados que ya conocemos sobre bases ortonormales de una manera mucho más sencilla, como se muestra a continuación[17]: sea V un espacio vectorial de dimensión finita n con base ortonormal $\beta = \{|v_1\rangle, |v_2\rangle, ..., |v_n\rangle\}$, entonces para todo $|v\rangle \in V$

$$P_{|v_i\rangle}(|v\rangle) = \langle v_i|v\rangle ,$$

$$|v\rangle = \sum_{i=1}^{n} \langle v_i|v\rangle |v_i\rangle ,$$

y para todo operador lineal $T : V \to V$ tenemos que si

$$A = [T]_\beta \implies A_{ij} = \langle T(v_j)|v_i\rangle .$$

[17] Asegúrate de entender bien esta notación y a qué resultados nos referimos antes de continuar leyendo el resto de las notas.

Funcionales y espacio dual

Como hemos visto anteriormente, si K es un campo y V y W son dos espacios vectoriales sobre K, entonces $\mathcal{L}(V, W)$ —el conjunto de todas las transformaciones lineales de V a W— es un espacio vectorial sobre K. En particular, como vimos a inicios del curso, (K, K) también es un espacio vectorial, por lo cual podemos fijar nuestra atención en el espacio vectorial $\mathcal{L}(V, K)$. En esta sección veremos el importante papel que juega este espacio vectorial; principalmente, nos enfocaremos en los espacios vectoriales con producto escalar.

Def. Sea (V, K) un espacio vectorial. Decimos que $\mathbf{f}$ es un *funcional lineal*, o simplemente un *funcional*, si $\mathbf{f} \in \mathcal{L}(V, K)$.

Def. Al espacio vectorial $\mathcal{L}(V, K)$ lo llamamos el *espacio dual* de V, denotado por V^*. Por ello, a sus elementos también se les conoce como *vectores duales*.

Ejemplos de funcionales

Sea V un espacio vectorial de dimensión finita, $\beta = (\mathbf{b}_1, \mathbf{b}_2, ..., \mathbf{b}_n)$ una base ordenada de V y $[\cdot]_\beta$ el *mapeo de coordenadas en la representación de la base ordenada* β dado por

$$[\mathbf{v}]_\beta = \begin{pmatrix} c_1 & c_2 & ... & c_n \end{pmatrix}$$

para todo $\mathbf{v} \in V$, donde los coeficientes c_i son tales que $\mathbf{v} = c_i \mathbf{b}_i$ en notación de Einstein. Si definimos a $\mathbf{f}_i(\mathbf{v}) = c_i$ para toda $1 \leq i \leq n$, entonces $\mathbf{f}_i$ es un funcional sobre V conocido como la *función de la i-ésima coordenada en la representación de la base ordenada* β.

En los espacios vectoriales de matrices cuadradas $M_{n\times n}(K)$, donde K es un campo, la traza $tr(\cdot) : M_{n\times n} \to K$ y el determinante $det(\cdot) : M_{n\times n} \to K$ son funcionales.

Sea (V, K) un espacio vectorial con producto escalar y $|v\rangle \in V$ un vector arbitrario, entonces la proyección escalar sobre $|v\rangle$ dada (en notación de Dirac) por

$$P_{|v\rangle}(|u\rangle) = \frac{\langle v|u\rangle}{||\,|v\rangle\,||}$$

para todo $|u\rangle \in V$ es un funcional sobre V.

Correspondencia entre bras y kets

Sea (V, K) un espacio vectorial arbitrario con producto escalar $\langle\cdot|\cdot\rangle : V \times V \to K$ donde, siguiendo la notación bra-ket, la primera entrada es antilineal. Si elegimos a un vector arbitrario $|v\rangle \in V$ y lo fijamos en la entrada antilineal del producto escalar, podemos definir un funcional $\langle v|\cdot\rangle : V \to K$, el cual se puede escribir más sencillamente como $\langle v|$. Ya que en notación bra-ket a los vectores $|u\rangle$ se les llama *kets* y a los productos escalares $\langle y|x\rangle$, *bra-kets* (proveniente del inglés *brackets*), a este tipo de funcionales $\langle v|$ se les conoce como *bras*.

Dado que a partir de cualquier vector $|v\rangle \in V$ se puede definir un funcional $\langle v| \in V^*$, existe una correspondencia natural entre *bras* y *kets*. Sin embargo, para entender bien esta correspondencia debemos observar un detalle crucial: si hacemos el producto de un ket $|v\rangle$ por un escalar $\alpha \in K$, obteniendo así el ket $\alpha |v\rangle$, entonces el bra correspondiente será $\overline{\alpha} \langle v|$[18]. Es decir, la correspondencia entre bras (que son un tipo de vectores duales) y kets está dada por

$$\alpha |v\rangle \leftrightarrow \overline{\alpha} \langle v| \, .$$

De esta manera, si queremos hacer el producto escalar de un vector arbitrario $\alpha_1 |u\rangle$ con otro vector $\alpha_2 |v\rangle$, podemos simplemente unir el ket $\alpha_1 |u\rangle$ con el bra correspondiente al otro vector, que sería $\overline{\alpha_2} |v\rangle$, obteniendo

$$\Big(\overline{\alpha_2} \langle v| \Big)\Big(\alpha_1 |u\rangle \Big) = \overline{\alpha_2}\alpha_1 \langle v|u\rangle \, .$$

Una ventaja de esta notación es que —siempre y cuando recordemos bien la correspondencia entre bras y kets mostrada en el párrafo anterior— ya no tenemos que preocuparnos por cómo sacar escalares del producto escalar.

Más generalmente, si $\{|b_1\rangle, |b_2\rangle, ..., |b_n\rangle\}$ es una base de V, entonces la correspondencia entre bras de V^* y kets de V puede expresarse como

$$\sum_{i=1}^{n} c_i |b_i\rangle \leftrightarrow \sum_{i=1}^{n} \overline{c_i} \langle b_i| \, .$$

[18] En la notación que usábamos antes, esto se explica como que si tomamos al vector $\alpha\mathbf{v}$ y lo fijamos en la entrada antilineal del producto escalar, obtenemos el funcional $\langle\,\cdot\,, \alpha\mathbf{v}\rangle : V \to K$ que, como sabemos, es igual a $\overline{\alpha}\langle\,\cdot\,, \mathbf{v}\rangle$.

Complemento ortogonal

<u>Def.</u> Sea V un espacio vectorial con producto escalar y $W \subseteq V$ un subespacio vectorial de V. Definimos al *complemento ortogonal* de W como

$$W^{\perp} = \{|x\rangle \in V \mid \langle x|w\rangle = 0 \ \forall \ |w\rangle \in W\}.$$

De la definición anterior podemos hacer varias observaciones:

- Ya que para cualquier espacio vectorial V con producto escalar el vector nulo es ortogonal a todos los vectores del espacio, tenemos que $\{|0\rangle\}^{\perp} = V$ y $V^{\perp} = \{|0\rangle\}$.
- Por la observación anterior de la ortogonalidad del vector nulo y las propiedades lineales del producto escalar, para todo subespacio vectorial $W \subseteq V$, $W^{\perp}$ también es un subespacio vectorial de V.
- Para cualquier subespacio vectorial $W \subseteq$ de V, $W \cap W^{\perp} = \{|0\rangle\}$.

Lema 13.1 13.3.1 Sea V un espacio vectorial con producto esclar sobre un campo K, W un subespacio vectorial de V y $\beta = \{|v_1\rangle, |v_2\rangle, ..., |v_k\rangle\}$ una base de W, entonces $|z\rangle \in W^{\perp}$ si y sólo si $\langle z|v_i\rangle = 0$ para toda $|v_i\rangle \in \beta$.

Demostración. Sea $|w\rangle \in W$ un vector arbitrario. Como β es base de W, existen coeficientes $c_i \in K$ tales que

$$|w\rangle = \sum_{i=1}^{k} c_i |v_i\rangle .$$

Sea $|z\rangle \in W^{\perp}$, entonces tenemos que $\langle z|w\rangle = 0$, pero esto ocurre si y sólo si

$$\langle z| \Big(\sum_{i=1}^{k} c_i |v_i\rangle \Big) \iff \sum_{i=1}^{k} c_i \langle z|v_i\rangle = 0 \iff \langle z|v_i\rangle = 0$$

para todo $|v_i\rangle \in \beta$, ya que los coeficientes c_i son arbitrarios. □

Teorema 13.2 13.3.2 Sea V un espacio vectorial con producto escalar y $W \subseteq V$ un subespacio vectorial de dimensión finita k, entonces para todo $|y\rangle \in V$ existen vectores únicos $|u\rangle \in W$ y $|z\rangle \in W^{\perp}$ tales que $|y\rangle = |u\rangle + |z\rangle$. Además, si $\beta = \{|v_1\rangle, |v_2\rangle, ..., |v_k\rangle\}$ es una base ortonormal de W, entonces

$$|u\rangle = \sum_{i=1}^{k} \langle v_i|y\rangle |v_i\rangle .$$

Demostración. Sea $\beta = \{|v_1\rangle, |v_2\rangle, ..., |v_k\rangle\}$ una base ortonormal de W. Definimos

$$|u\rangle = \sum_{i=1}^{k} \langle v_i|y\rangle |v_i\rangle \quad \text{y} \quad |z\rangle = |y\rangle - |u\rangle .$$

Claramente $|u\rangle \in W$ y $|y\rangle = |u\rangle + |z\rangle$. Además, observemos que

$$\langle z|u\rangle = \Big(\langle y| - \langle u| \Big) |u\rangle = \langle y|u\rangle - \langle u|u\rangle = \langle y| \left(\sum_{i=1}^{k} \langle v_i|y\rangle |v_i\rangle \right) - \left(\sum_{i=1}^{k} \langle y|v_j\rangle \langle v_i| \right) \left(\sum_{j=1}^{k} \langle v_j|y\rangle |v_j\rangle \right)$$

$$= \sum_{i=1}^{k} \langle v_i|y\rangle \langle y|v_i\rangle - \sum_{i=1}^{k} \sum_{j=1}^{k} \langle y|v_i\rangle \langle v_j|y\rangle \langle v_i|v_j\rangle = \sum_{i=1}^{k} \langle v_i|y\rangle \langle y|v_i\rangle - \sum_{i=1}^{k} \sum_{j=1}^{k} \langle y|v_i\rangle \langle v_j|y\rangle \delta_{ij}$$

$$= \sum_{i=1}^{k} \langle v_i|y\rangle \langle y|v_i\rangle - \sum_{i=1}^{k} \langle y|v_i\rangle \langle v_i|y\rangle = 0,$$

por lo que $|z\rangle \in W^{\perp}$.

Para demostrar la unicidad de $|u\rangle$ y $|z\rangle$, supongamos que existen $|u'\rangle \in W$ y $|z'\rangle \in W^{\perp}$ tales que $|y\rangle = |u'\rangle + |z'\rangle$. Entonces, tenemos que

$$|y\rangle = |u'\rangle + |z'\rangle = |u\rangle + |z\rangle \implies |u\rangle - |u'\rangle = |z\rangle - |z'\rangle .$$

Ya que el vector del lado izquierdo de la última ecuación es un elemento de W y el del lado derecho, de $W^{\perp}$, tenemos que

$$|u\rangle - |u'\rangle = |z\rangle - |z'\rangle \in W \cap W^{\perp} \implies |u\rangle - |u'\rangle = |0\rangle = |z\rangle - |z'\rangle \implies |u'\rangle = |u\rangle \quad \text{y} \quad |z'\rangle = |z\rangle .$$

□

Teorema 13.3 13.3.3 Sea $S = \{|v_1\rangle, |v_2\rangle, ..., |v_k\rangle\}$ un conjunto ortonormal de vectores de un espacio vectorial V de dimensión finita n con producto escalar. Entonces

a) S se puede extender a una base ortonormal $\{|v_1\rangle, |v_2\rangle, ..., |v_k\rangle, |v_{k+1}\rangle, ..., |v_n\rangle\}$ de V.

b) Usando la notación anterior, si $W = \langle S\rangle$, entonces $S_1 = \{|v_{k+1}\rangle, ..., |v_n\rangle\}$ es base ortonormal de $W^{\perp}$.

c) Para cualquier subespacio $W \subseteq V, \dim(V) = \dim(W) + \dim(W^{\perp})$.

Demostración. a) Por el Teorema de reemplazamiento (ver sec. 3), sabemos que podemos extender a S para formar una base $S' = \{|v_1\rangle, |v_2\rangle, ..., |v_k\rangle, |w_{k+1}\rangle, ..., |w_n\rangle\}$ de V. Aplicando el Teorema de Gram-Schmidt, podemos obtener una base $\{|v_1\rangle, |v_2\rangle, ..., |v_k\rangle, |v_{k+1}\rangle, ..., |v_n\rangle\}$ de V.

b) Observemos que, por la construcción de la base $\{|v_1\rangle, |v_2\rangle, ..., |v_k\rangle, |v_{k+1}\rangle, ..., |v_n\rangle\}$ de V, $\langle S_1\rangle \subseteq W^{\perp}$ y S_1 es un conjunto linalmente independiente. Por último, observemos que para cualquier $|v\rangle \in V$,

$$|v\rangle = \sum_{i=1}^{n} \langle v_i|v\rangle |v_i\rangle .$$

En particular, para todo $|v\rangle \in W^{\perp}$ tenemos que $\langle v_i|v\rangle = 0$ para $1 \leq i \leq k$. En este caso,

$$|v\rangle = \sum_{i=k+1}^{n} \langle v_i|v\rangle |v_i\rangle = \sum_{i=k+1}^{n} c_i |v_i\rangle \implies |v\rangle \in \langle S_1\rangle \implies W^{\perp} \subseteq \langle S_1\rangle.$$

Por otro lado, como $S_1 \subseteq W^{\perp}$ tenemos que $\langle S_1\rangle \subseteq W^{\perp}$, por lo que $\langle S_1\rangle = W^{\perp}$. Por lo tanto, S_1 es base ortonormal de $W^{\perp}$.

c) Sea W un subespacio vectorial de V. Como V es de dimensión finita, entonces W también lo es. Por ende, W tiene una base ortonormal $\{|v_1\rangle, |v_2\rangle, ..., |v_k\rangle\}$. Por los incisos a) y b), tenemos que

$$\dim(V) = n = k + (n - k) = \dim(W) + \dim(W^{\perp}).$$

□

Proyecciones ortogonales

En la sección ?? vimos que, en un espacio vectorial con producto escalar, podemos definir funciones de *proyección vectorial* que toman un vector del espacio y devuelven la componente de ese vector a lo largo de un subespacio vectorial particular. Además, en el Teorema 4.3.1.1 (ver sec. ??) vimos cómo podemos descomponer un vector en sus componentes ortogonales de manera sencilla a partir de una base ortogonal u ortonormal. Ambos resultados reflejan la importancia que tienen los llamados *operadores de proyección ortogonal* en este tipo de espacios vectoriales.

> Def. Sea V un espacio vectorial arbitrario. Decimos que un operador lineal $P : V \to V$ es un *operador de proyección* si
> $$P(P(|v\rangle)) = P(|v\rangle) \;\; \forall \;\; |v\rangle \in V.$$
> Definiendo $P^2 := P \circ P$ podemos reescribir esta condición simplemente como $P^2 = P$. En particular, si V tiene producto escalar y P es tal que
> $$\mathrm{Im}(P)^\perp = \mathrm{Ker}(P) \;\; \text{y} \;\; \mathrm{Ker}(P)^\perp = \mathrm{Im}(P),$$
> decimos que P es un *operador de proyección ortogonal*.

Observemos que la condición $P^2 = P$ para un operador de proyección P en un espacio vectorial V codifica algebráicamente la noción intuitiva de que, después de proyectar un vector arbitrario $|v\rangle$ en un subespacio de V y obtener la componente de $|v\rangle$ en dicho subespacio, proyectar esa componente *en el mismo subespacio* no le hará nada —ya que se encuentra totalmente dentro de ese subespacio.

Por otro lado, a pesar de que no sea fácil de ver directamente de la definición, cualquier operador de proyección actuando sobre un espacio vectorial de dimensión finita lo divide en dos subespacios vectoriales "ajenos" (excepto por el vector nulo, por supuesto) a través de su imagen (el subespacio sobre el cual proyectamos) y su núcleo (el conjunto de vectores que no tienen componentes en dicho subespacio). Dicho de otra forma, cualquier operador de proyección actuando sobre un espacio vectorial de dimensión finita descompone al espacio como suma directa de su imagen y su núcleo (¿Se te ocurre por qué[19]?).

Para cualquier espacio vectorial V, el operador nulo y el operador identidad son ejemplos triviales de operadores de proyección. Si V tiene producto escalar, entonces la función de proyección vectorial sobre $|v\rangle$ es un operador de proyección ortogonal para todo $|v\rangle \in V$ no nulo; en términos más precisos, este operador proyecta sobre el subespacio $\langle |v\rangle \rangle$ de dimensión uno. Observemos que podemos escribir a este operador en términos de bras y kets simplemente como $\frac{|v\rangle\langle v|}{\langle v|v\rangle} : V \to V$, de donde resulta fácil demostrar que efectivamente es un operador de proyección, ya que

$$\left(\frac{|v\rangle\langle v|}{\langle v|v\rangle}\right)^2 = \left(\frac{1}{\langle v|v\rangle}|v\rangle\langle v|\right)^2 = \frac{1}{\langle v|v\rangle^2}\Big(|v\rangle\langle v|\Big)^2 = \frac{1}{\langle v|v\rangle^2}|v\rangle\langle v|v\rangle\langle v| = \frac{\langle v|v\rangle}{\langle v|v\rangle^2}|v\rangle\langle v| = \frac{|v\rangle\langle v|}{\langle v|v\rangle}.$$

Además, si en particular $|v\rangle$ es un vector normal, la expresión se simplifica a $|v\rangle\langle v|$, de donde se ve directamente que $(|v\rangle\langle v|)^2 = |v\rangle\langle v|v\rangle\langle v| = |v\rangle\langle v|$. Por lo tanto, todo vector normal $|v\rangle$ en un espacio vectorial con producto escalar define un operador de proyección ortogonal, dado por $|v\rangle\langle v|$, y dicho operador proyecta a los vectores de todo el espacio sobre el subespacio generado por $|v\rangle$.

[19] Pista: demuestra que, si V es un espacio vectorial de dimensión finita y $T : V \to V$ es un operador lineal, entonces $\mathrm{Ker}(T) + \mathrm{Im}(T) = V \iff \mathrm{Ker}(T) \cap \mathrm{Im}(T) = \{|0\rangle\}$.

Nota aclaratoria: *Sobre bras, kets, y bra-kets...*

Como hemos visto en esta sección, en un espacio vectorial (V, K) los *kets* $|x\rangle, |y\rangle, |z\rangle$, etc. son simplemente vectores de V, mientras que los *bras* $\langle x|, \langle y|, \langle z|$, etc. son funcionales sobre V, es decir, transformaciones lineales que van de V a K.

A pesar de que exista la correspondencia entre kets $\alpha |x\rangle \in V$ y bras $\overline{\alpha} \langle x| \in V^*$ descrita en la sección 13, debemos remarcar una diferencia sutil pero importante entre la operación de aplicarle un funcional a un vector y la operación de obtener el producto escalar entre dos vectores antes de pasar a la siguiente sección.

Si $|x\rangle \in V$ y $\langle y| \in V^*$ entonces, como $\langle y| : V \to K$, tenemos que

$$\langle y| (|x\rangle) = \langle y|x\rangle ,$$

donde $\langle y|$ es el vector dual correspondiente al vector $|y\rangle$. La observación crucial es esta: del lado izquierdo de la ecuación estamos aplicándole el funcional $\langle y|$ al vector $|x\rangle$, mientras que del lado derecho estamos obteniendo el producto escalar entre el vector $|x\rangle$ (en la entrada lineal) y el vector $|y\rangle$ (en la entrada antilineal)[a]; ambas operaciones involucran objetos distintos —la primera se realiza entre un funcional ($\langle y|$) y un ket ($|x\rangle$), mientras que la segunda se realiza entre dos kets ($|x\rangle$ y $|y\rangle$)— y, sin embargo, dan el mismo escalar como resultado.

[a]Inclusive l@s más atrevid@s entre ustedes se podrían animar a agregar una equivalencia más a la ecuación anterior: $\langle y|x\rangle = (\langle y|) |x\rangle$; es decir, podrían entretener la idea de que $|x\rangle$ sea ahora un *funcional* que actúa sobre los vectores duales $\langle y| \in V^*$ por la derecha; esto está relacionado con el hecho de que, si V es un espacio de dimensión finita, entonces $(V^*)^*$ es isomorfo a V —aunque no lo demostraremos en este curso.

Ejercicios de repaso

Funcionales y espacio dual

1. Demuestra que si V es un espacio vectorial de dimensión finita, entonces $\dim(V^*) = \dim(V)$.

2. Consideremos el espacio de funciones reales de variable real continuas definidas en el intervalo $[0, 2\pi]$, que denotaremos como $C([0, 2\pi], \mathbb{R})$. Sean

$$F_{1,n}(f) := \int_0^{2\pi} \sin(nx) f(x)\ dx, \quad F_{2,n}(f) := \int_0^{2\pi} \cos(nx) f(x)\ dx \quad \forall\ f \in C([0, 2\pi], \mathbb{R}), n \in \mathbb{N}.$$

Demuestra que $F_{1,n}$ y $F_{2,n}$ son funcionales de $C([0, 2\pi], \mathbb{R})$ para todo $n \geq 0$. Los valores de estos funcionales para una función g en particular se conocen como los *n-ésimos coeficientes de Fourier*[20] de la función g.

3. Calcula los primeros 3 coeficientes de Fourier para las funciones $\sin(x)$, $f(x) = x$ y $g(x) = x^2$.

[20]Estos coeficientes están relacionados con una rama de las matemáticas llamada *Análisis de Fourier*, la cual tiene diversas aplicaciones en la solución de ecuaciones diferenciales parciales, muchas de las cuales son utilizadas para modelar procesos físicos (como la ecuación de calor o la ecuación de onda) o biológicos.

Correspondencia entre bras y kets

1. Sea (V, K) un espacio vectorial con producto escalar. Demuestra que el conjunto de todos los bras correspondientes a los vectores de V forma un subespacio vectorial de V^*.

2. Demuestra que el subespacio de V^* obtenido en el ejercicio anterior es isomorfo a V. En otras palabras, demuestra que la correspondencia entre bras y kets dada en la sección 13 es un isomorfismo.

Complemento ortogonal

1. Demuestra que para todo subespacio vectorial W de un espacio vectorial V de dimensión finita n y con producto escalar, V es igual a una suma directa de W y $W^{\perp}$. ¿Por qué es importante que V tenga dimensión finita? (Nota: tendrás que demostrar las observaciones hechas al principio de la sec. 13.)

Proyecciones ortogonales

1. Sea V un espacio vectorial. Demuestra que un operador lineal $P : V \to V$ es un operador de proyección si y sólo si $V = \mathrm{Ker}(P) \oplus \mathrm{Im}(P)$.

2. Sea V un espacio vectorial con producto escalar. Demuestra que un operador lineal $P : V \to V$ es un operador de proyección ortogonal si y sólo si $V = \mathrm{Ker}(P)^{\perp} \oplus \mathrm{Im}(P)^{\perp}$.

3. Sea V un espacio vectorial de dimensión finta n con producto escalar y sea $\{|v_1\rangle, |v_2\rangle, \dots, |v_k\rangle\}$ un conjunto ortogonal de vectores de V con $k \leq n$. Demuestra que $\sum_{i=1}^{k} \frac{|v_i\rangle\langle v_i|}{\langle v_i|v_i\rangle} : V \to V$ es un operador de proyección ortogonal.

14. Descomposición espectral y operadores adjuntos

Supongamos que tenemos un espacio vectorial (V, K) de dimensión finita n con producto escalar. Entonces, siguiendo del último ejercicio de la sección 13, tendríamos que si $\{|b_1\rangle, |b_2\rangle, \dots, |b_n\rangle\}$ es una base ortogonal de V, entonces $\sum_{i=1}^{n} \frac{|b_i\rangle\langle b_i|}{\langle b_i|b_i\rangle} : V \to V$ es un operador de proyección ortogonal[21]. En particular, para cualquier base ortonormal $\{|u_1\rangle, |u_2\rangle, \dots, |u_n\rangle\}$ de V tenemos que $\sum_{i=1}^{n} |u_i\rangle\langle u_i|$ es un operador de proyección ortogonal (equivalente al anterior, para cualquier base ortonormal arbitraria).

Ambos operadores anteriores no son más que el operador identidad en V, lo cual podemos ver en abstracto aplicándolos a un vector arbitrario $|v\rangle \in V$ o, de forma matricial, representándolos en las mismas bases que utilizamos para construirlos. A continuación, veremos cómo esta forma de entender al operador identidad (como suma de operadores de proyección más pequeños) en espacios vectoriales de dimensión finita con producto escalar nos ayuda a entender cómo actúan ciertos tipos de operadores lineales de forma geométrica.

Descomposición espectral (introducción)

Sean V un espacio vectorial de dimensión finita n con producto escalar y $T : V \to V$ un operador lineal diagonalizable tal que sus eigenvectores $|g_1\rangle, |g_2\rangle, \dots, |g_n\rangle$ forman una base[22] *ortogonal* de V; es decir, que se puede elegir a una base ortogonal de V compuesta por eigenvectores de T. Supongamos, adicionalmente, que normalizamos a cada uno de los vectores de la eigenbase y los redefinimos de tal forma que $\{|g_1\rangle, |g_2\rangle, \dots, |g_n\rangle\}$ sea una base *ortonormal* de V compuesta por eigenvectores de T. En ese caso, tendríamos que

$$I_V = \sum_{i=1}^{n} |g_i\rangle\langle g_i|.$$

Sean $\lambda_1, \lambda_2, \dots, \lambda_n$ los eigenvalores (no necesariamente todos distintos entre sí) de T correspondientes a los eigenvectores normales $|g_1\rangle, |g_2\rangle, \dots, |g_n\rangle$. Consideremos ahora al operador $T' := \sum_{i=1}^{n} \lambda_i |g_i\rangle\langle g_i|$. En particular, observemos que para todo elemento $|g_j\rangle$ de la eigenbase tenemos que

$$T'(|g_j\rangle) = \sum_{i=1}^{n} \lambda_i |g_i\rangle\langle g_i| (|g_j\rangle) = \sum_{i=1}^{n} \lambda_i |g_i\rangle\langle g_i|g_j\rangle = \sum_{i=1}^{n} \lambda_i |g_i\rangle \delta_{ij} = \lambda_j |g_j\rangle,$$

por lo que T' tiene los mismos vectores y valores *característicos* que T. Ya que T es diagonalizable y, por lo tanto, podemos formar una base de V compuesta de sus eigenvectores (por ejemplo, $\{|g_1\rangle, |g_2\rangle, \dots, |g_n\rangle\}$), se sigue que $T' = T$ y que, por ende,

$$T = \sum_{i=1}^{n} \lambda_i |g_i\rangle\langle g_i|,$$

donde $\{|g_1\rangle, |g_2\rangle, \dots, |g_n\rangle\}$ es una base ortogonnal de V compuesta de eigenvectores de T normalizados y $\lambda_1, \lambda_2, \dots, \lambda_n$ son los eigenvalores correspondientes (no necesariamente todos distintos entre sí). Esta expresión se conoce como la *descomposición espectral* de T.

[21] Este operador no depende de la base ortogonal elegida... ¿puedes adivinar de qué operador se trata?

[22] Recuerdemos que un operador $T : V \to V$ es diagonalizable si y sólo si se puede formar una base de V compuesta de eigenvectores de T. Si V es dimensión n, entonces cualquier base de V tiene exactamente n elementos, por lo que T debe tener n eigenvectores linealmente independientes; sin embargo, esto no necesariamente implica que T tenga n eigenvalores distintos, como hemos visto. Por lo tanto, si T tiene eigenvalores $\lambda_1, \lambda_2, \dots, \lambda_k$ con $k < n$, necesariamente habrá elementos distintos de la eigenbase que tengan el mismo eigenvalor y, en este caso, la correspondencia entre los subíndices de los k eigenvalores distintos y los n elementos de la eigenbase no será uno a uno, por lo que hay que tener cuidado y no dejarse llevar por la notación.

Observemos que la descomposición espectral de T nos da información completa de cómo actúa T en V geométricamente: la expresión anterior nos dice que T proyecta a un vector arbitrario $|v\rangle \in V$ en cada una de sus componentes a lo largo de los diferentes eigenespacios de T (a través de las proyecciones ortogonales $|g_i\rangle \langle g_i|$) —los cuales son ortogonales entre sí— , luego multiplica cada una de esas componentes por el eigenvalor correspondiente al eigenespacio en el que se encuentran, y luego reconstruye el vector resultante a partir de las componentes reescaladas[23].

En otras palabras, hemos visto que si T es un operador lineal sobre un espacio vectorial V de dimensión finita con producto escalar y cumple que se pueda formar una base ortogonal u ortonormal de V a partir de sus eigenvectores o, equivalentemente, que

$$V = E_{\lambda_1} \oplus E_{\lambda_2} \oplus ... \oplus E_{\lambda_k} \quad \text{y} \quad E_{\lambda_i}^{\perp} = \oplus_{j \neq i} E_{\lambda_j} \quad \forall\, i \in \{1, 2, \dots, n\},$$

donde $\lambda_1, \lambda_2, \dots, \lambda_k$ son los eigenvalores de T, entonces podemos entender a T de una forma extremadamente sencilla: a través de proyecciones (en los eigenespacios de T, que son ortogonales entre sí) y reescalamientos (por los eigenvalores de T). Un ejemplo de esto se muestra en la Figura 7.

Las condiciones descritas en el párrafo anterior en general no son fáciles de verificar, por lo que nos gustaría tener algún criterio equivalente que sea más sencillo de aplicar; de cierta manera, de eso se tratará el resto del curso. Empezando nuestra búsqueda por hipótesis más precisas que deban cumplir operadores lineales sobre espacios vectoriales de dimensión finita con producto escalar para que les podamos aplicar la descomposición espectral, nuestra primera pista se encuentra en los operadores adjuntos, que definiremos y estudiaremos en la siguiente sección.

Operadores adjuntos

Como vimos en la sección 13, en cualquier espacio vectorial (V, K) con producto escalar podemos tomar cualquier vector $|y\rangle \in V$ y definir un funcional $\mathbf{g} : V \to K$ como

$$\mathbf{g}(|x\rangle) = \langle y|x\rangle\,,$$

que podemos escribir simplemente como el bra $\langle y|$.

Lo interesante de este tipo de funcionales es que, en espacios vectoriales de dimensión finita, todas las transformaciones lineales de V a K son de esta forma; es decir, podemos hacer una identificación entre las transformaciones lineales $\mathbf{g} \in \mathcal{L}(V, K)$ y los bras $\langle y|$, como demostraremos a continuación.

Teorema 14.1 14.1 Sea (V, K) un espacio vectorial de dimensión finita con producto escalar y sea $\mathbf{g} : V \to K$ un funcional. Entonces existe un único $\langle y| \in V^*$ tal que $\mathbf{g}(|x\rangle) = \langle y|x\rangle$ para todo $|x\rangle \in V$.

Demostración. Sea $\beta = \{|b_1\rangle, |b_2\rangle, ..., |b_n\rangle\}$ una base ortonormal de V y sea

$$|y\rangle = \sum_{i=1}^{n} \overline{\mathbf{g}(|b_i\rangle)}\, |b_i\rangle\,.$$

Entonces el vector dual correspondiente a $|y\rangle \in V$ es el bra $\langle y| \in V^*$ dado por

$$\langle y| = \sum_{i=1}^{n} \mathbf{g}(|b_i\rangle)\, \langle b_i|\,.$$

[23] Aquí utilizamos la palabra “reescalamiento” como sinónimo de ”producto de un vector por un reescalar”; recordemos que, geométricamente, en general esta operación puede involucrar un reescalamiento y una rotación.

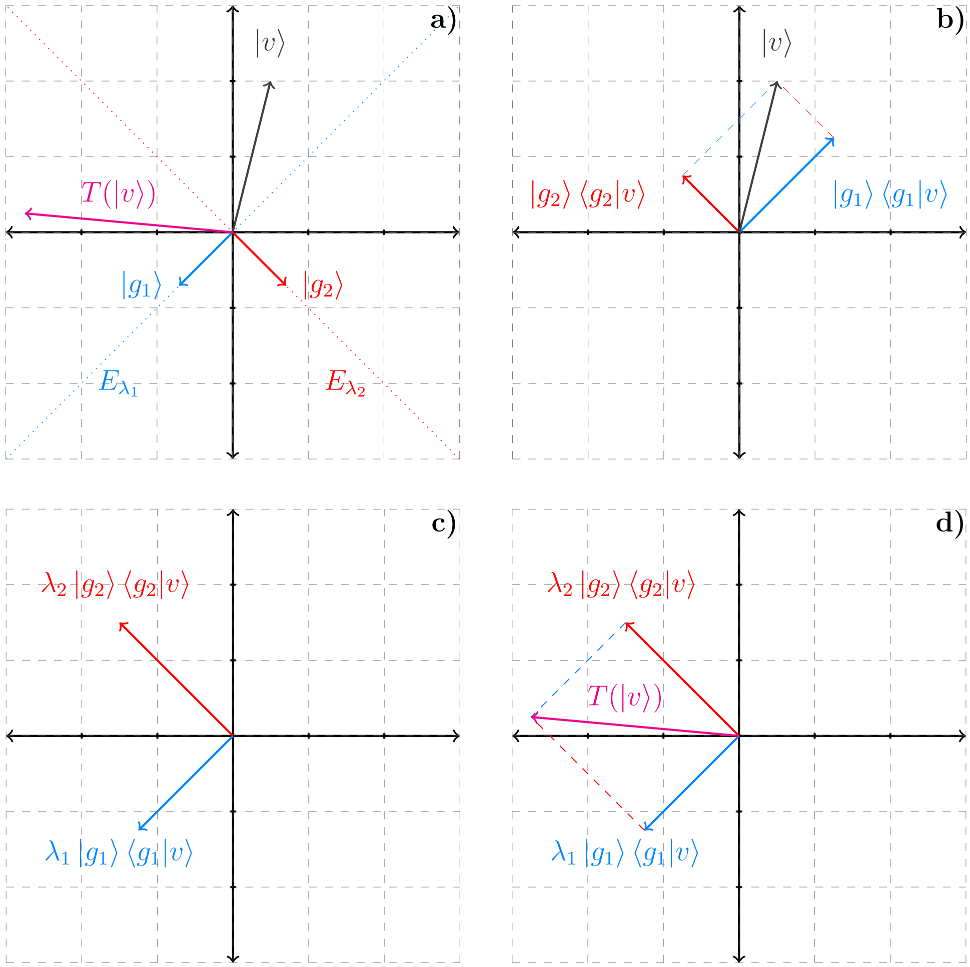

Figura 7: Aplicación de un operador $T : \mathbb{R}^2 \to \mathbb{R}^2$ a un vector $|v\rangle$ mediante descomposición espectral. En la subfigura **a)** se muestra a $|v\rangle$ y a su imagen $T(|v\rangle)$, junto con los eigenvectores $|g_1\rangle$ y $|g_2\rangle$ de T, que forman una base ortonormal; los eigenespacios correspondientes se representan con líneas punteadas. En **b)** se muestran las proyecciones ortogonales de $|v\rangle$ en cada eigenespacio de T. En **c)** se muestran las componentes anteriores reescaladas por los eigenvalores correspondientes al eigenespacio en que se encuentran. Finalmente, en **d)** vemos que la suma de estas componentes reescaladas es igual a $T(|v\rangle)$.

Por otro lado, definamos a $\mathbf{h} : V \to K$ como $\mathbf{h}(|x\rangle) = \langle y|x\rangle$ para todo $|x\rangle \in V$. Claramente, $\mathbf{h}$ es una transformación lineal. Observemos que para todos los elementos de la base β, sus imágenes bajo $\mathbf{g}$ y $\mathbf{h}$ coinciden; es decir, para todo $1 \leq j \leq n$ tenemos que

$$\mathbf{h}(|b_j\rangle) = \langle y|b_j\rangle = \left(\sum_{i=1}^{n} \mathbf{g}(|b_i\rangle) \langle b_i| \right) |b_j\rangle = \sum_{i=1}^{n} \mathbf{g}(|b_i\rangle) \langle b_i|b_j\rangle = \sum_{i=1}^{n} \mathbf{g}(|b_i\rangle)\delta_{ij} = \mathbf{g}(|b_j\rangle).$$

Dado que $\mathbf{g}$ y $\mathbf{h}$ coinciden para todos los elementos de la base y son transformaciones lineales, entonces coinciden para todos los vectores $|v\rangle \in V$. Por ende, $\mathbf{g} = \mathbf{h}$.

Para demostrar la unicidad de $\langle y|$, supongamos que $\mathbf{g}(|x\rangle) = \langle y'|x\rangle$ para todo $|x\rangle \in V$. Entonces tenemos que

$$\langle y|x\rangle = \langle y'|x\rangle \implies \langle y|x\rangle - \langle y'|x\rangle = 0 \implies \Big(\langle y| - \langle y'| \Big) |x\rangle = 0$$

para todo $|x\rangle \in V$, pero esto se cumple si y sólo si

$$\langle y| - \langle y'| = \langle 0| \implies \langle y| = \langle y'| .$$

□

Antes de continuar, recordamos la siguiente definición.

Def. Decimos que una transformación lineal es un *operador lineal* cuando su dominio y contradominio son el mismo espacio vectorial. Por lo tanto, si decimos que T es un operador lineal —o, simplemente, operador— sobre V, nos referimos a una transformación lineal $T : V \to V$.

Teorema 14.2 14.2 Sea (V, K) un espacio vectorial de dimensión finita con producto escalar y sea $T : V \to V$ un operador lineal. Entonces existe una única función $T^* : V \to V$ tal que $\langle y|T(x)\rangle = \langle T^*(y)|x\rangle$ para toda $|x\rangle, |y\rangle \in V$. Además, T^* es lineal.

Demostración. Sea $|y\rangle \in V$. Definimos a $\mathbf{g} : V \to K$ como $\mathbf{g}(|x\rangle) = \langle y|T(x)\rangle$ para todo $|x\rangle \in V$. Primero, verificamos que $\mathbf{g}$ es lineal: sean $|x_1\rangle, |x_2\rangle \in V$ y $c \in K$, entonces

$$\begin{aligned}\mathbf{g}(c\,|x_1\rangle + |x_2\rangle) &= \langle y|T(cx_1 + x_2)\rangle = \langle y|cT(x_1) + T(x_2)\rangle = \langle y|cT(x_1)\rangle + \langle y|T(x_2)\rangle \\ &= c\,\langle y|T(x_1)\rangle + \langle y|T(x_2)\rangle = c\mathbf{g}(|x_1\rangle) + \mathbf{g}(|x_2\rangle),\end{aligned}$$

por lo que T es lineal.

Luego, aplicamos el Teorema 14.1 para obtener un vector único $|y'\rangle \in V$ tal que $\mathbf{g}(|x\rangle) = \langle y'|x\rangle$ para todo $|x\rangle \in V$. Observemos que, por definición de $\mathbf{g}$, $\langle y|T(x)\rangle = \langle y'|x\rangle$ para todo $|x\rangle \in V$. Definiendo a $T^* : V \to V$ como $T^*(|y\rangle) = |y'\rangle$ tenemos que $\langle y|T(x)\rangle = \langle T^*(y)|x\rangle$, como se deseaba. Ahora, debemos demostrar que T^* es lineal: sean $|y_1\rangle, |y_2\rangle \in V$ y $c \in K$, entonces para todo $|x\rangle \in V$ se cumple que

$$\begin{aligned}\langle T^*(cy_1 + y_2)|x\rangle &= \langle cy_1 + y_2|T(x)\rangle = \overline{c}\,\langle y_1|T(x)\rangle + \langle y_2|T(x)\rangle \\ &= \overline{c}\,\langle T^*(y_1)|x\rangle + \langle T^*(y_2)|x\rangle = \langle cT^*(y_1) + T^*(y_2)|x\rangle \\ &\implies \langle T^*(cy_1 + y_2)|x\rangle - \langle cT^*(y_1) + T(y_2)|x\rangle = 0 \\ &\implies \langle T^*(cy_1 + y_2) - (cT^*(y_1) + T^*(y_2))|x\rangle = 0.\end{aligned}$$

Como esto vale para todo $|x\rangle \in V$, entonces esto implica que

$$\begin{aligned}&|T^*(cy_1 + y_2) - (cT^*(y_1) + T(y_2))\rangle = |0\rangle \\ &\implies T^*(|cy_1 + y_2\rangle) = cT^*(|y_1\rangle) + T^*(|y_2\rangle).\end{aligned}$$

Finalmente, para ver que T^* es única: sea $U : V \to V$ un operador lineal tal que $\langle y|T(x)\rangle = \langle U(y)|x\rangle$ para todo $|x\rangle, |y\rangle \in V$, entonces $\langle T^*(y)|x\rangle = \langle U(y)|x\rangle$ para todo $|x\rangle, |y\rangle \in V$, por lo que $U = T^*$. □

Como veremos dentro de poco, el operador T^* descrito en el Teorema 14.2 es de gran importancia en los espacios vectoriales con producto escalar, por lo cual le damos un nombre especial.

Def. Sea V un espacio vectorial de dimensión finita y $T^* : V \to V$ el operador lineal único que para todo $|x\rangle, |y\rangle \in V$ satisface la relación

$$\langle y|T(x)\rangle = \langle T^*(y)|x\rangle\,,$$

donde $T : V \to V$ es un operador lineal. Entonces, decimos que T^* es el *operador adjunto* de T.

En el caso en que V sea un espacio vectorial de dimensión infinita y para un operador lineal $T : V \to V$ exista otro operador lineal $T^* : V \to V$ tal que la relación anterior se cumpla, entonces nuevamente diremos que T^* es el operador adjunto de T —sin embargo, en este tipo de espacios vectoriales la existencia de T^* no está asegurada para todo operador lineal T.

Ya que varios de los teoremas que demostraremos a continuación aplican tanto para espacios vectoriales de dimensión finita como aquellos de dimensión infinita, de ahora en adelante, cuando hagamos referencia a un operador adjunto en un espacio vectorial de dimensión infinita, asumiremos implícitamente su existencia.

A continuación, demostraremos que en espacios vectoriales de dimensión finita podemos obtener la representación matricial de un operador adjunto a partir de la representación matricial del operador original de una forma muy sencilla, siempre y cuando utilicemos una base ortonormal para representar ambas matrices.

Teorema 14.3 14.3 Sea (V, K) un espacio vectorial de dimensión finita n con producto escalar, $\beta = \{|b_1\rangle, |b_2\rangle, ..., |b_n\rangle\}$ una base ortonormal de V y $T : V \to V$ un operador lineal. Entonces

$$[T^*]_\beta = \left([T]_\beta\right)^*,$$

donde A^* se define mediante la relación $A^*_{ij} = \overline{A_{ji}}$ para toda $A \in M_{n\times n}(K)$.

Demostración. Sean $A = [T]_\beta$ y $B = [T^*]_\beta$. Si partimos del sistema de ecuaciones que define a la representación matricial A y hacemos producto escalar por la izquierda en ambos lados de cada una de las ecuaciones por un elemento $|b_i\rangle$ de la base y aplicamos ortonormalidad, obtenemos la relación

$$A_{ij} = \langle b_i | T(b_j)\rangle .$$

Haciendo un procedimiento análogo con la representación matricial B, obtenemos que

$$B_{ij} = \langle b_i | T^*(b_j)\rangle ,$$

de donde se sigue que

$$B_{ij} = \overline{\langle T^*(b_j) | b_i\rangle} = \overline{\langle b_j | T(b_i)\rangle} = \overline{A_{ji}},$$

lo cual implica que $B = A^*$, es decir, que $[T^*]_\beta = \left([T]_\beta\right)^*$. □

A la luz del Teorema 14.3, damos una nueva definición, que quizá ya esperes.

Def. Sea K un campo y $A \in M_{n\times n}(K)$. Decimos que $A^* \in M_{n\times n}(K)$, definida a través de la relación

$$A^*_{ij} = \overline{A_{ji}},$$

es la *matriz adjunta* de A.

Recordando que para matrices cuadradas A la matriz transpuesta A^T se define mediante la relación

$$A^T_{ij} = A_{ji},$$

podemos ver que la matriz adjunta A^* se obtiene mediante la transposición de la matriz A y la conjugación de sus entradas, sin importar el orden en que se realicen estas dos operaciones. Si entendemos por $\overline{A}$ a la matriz que se obtiene al conjugar todas las entradas de A, podemos escribir esto simplemente como $A^* = \overline{A^T} = \overline{A}^T$.

Ejercicios de repaso

Descomposición espectral (introducción)

1. Sean V un espacio vectorial de dimensión finita, $T : V \to V$ un operador lineal y $\{\lambda_1, \lambda_2, \ldots , \lambda_k\}$ el espectro de T. Demuestra que T es diagonalizable si y sólo si $V = E_{\lambda_1} \oplus E_{\lambda_2} \oplus \ldots \oplus E_{\lambda_k}$.

2. Sean V un espacio de dimensión finita con producto escalar, $T : V \to V$ un operador lineal y $\Lambda = \{\lambda_1, \lambda_2, \ldots , \lambda_k\}$ el espectro de T. Demuestra que T se puede descomponer espectralmente si y sólo si $V = E_{\lambda_i} \oplus (E_{\lambda_i})^\perp$ para toda $\lambda_i \in \Lambda$.

3. Sean V un espacio de dimensión finita n con producto escalar, $T : V \to V$ un operador lineal, $\{|g_1\rangle , |g_2\rangle , \ldots , |g_n\rangle\}$ un conjunto de eigenvectores normales de T y $\Lambda = \{\lambda_1, \lambda_2, \ldots , \lambda_n\}$ un conjunto de eigenvalores de T (no necesariamente todos distintos) tales que $T(|g_i\rangle) = \lambda_i |g_i\rangle \ \forall\ i \in \{1, 2, \ldots , n\}$. Demuestra que $T = \sum_{i=1}^{n} \lambda_i |g_i\rangle \langle g_i|$ si y sólo si $I = \sum_{i=1}^{n} |g_i\rangle \langle g_i|$.

Operadores adjuntos

1. Sea V un espacio vectorial de dimensión finita n con producto escalar y $\beta = \{|b_1\rangle , |b_2\rangle , ..., |b_n\rangle\}$ una base ortonormal de V. Demuestra que si $T : V \to V$ es lineal, entonces $([T]_\beta)_{ij} = \langle b_i|T(b_j)\rangle$.

2. Sea K un campo y $A \in M_{n\times n}(K)$. Demuestra que $A^* = \overline{A^T} = (\overline{A})^T$ y, por lo tanto, que $\det(A^*) = \overline{\det(A)}$.

3. Sea (V, K) un espacio vectorial con producto escalar y sean T y U operadores lineales sobre V para los cuales existen operadores adjuntos T^* y U^*, respectivamente. Demuestra las siguientes propiedades:

   a) $(T + U)^* = T^* + U^*$;
   b) $(cT)^* = \overline{c}T^*$ para todo $c \in K$;
   c) $(TU)^* = U^*T^*$;
   d) $(T^*)^* = T$;
   e) $I^* = I$.

4. Sea K un campo y sean $A, B \in M_{n\times n}(K)$. Demuestra las siguientes propiedades:

   a) $(A + B)^* = A^* + B^*$;
   b) $(cA)^* = \overline{c}A^*$ para todo $c \in K$;
   c) $(AB)^* = B^*A^*$;
   d) $(A^*)^* = A$;
   e) $I^* = I$.

Una vez que hayas resuelto los ejercicios anteriores, observa de qué formas el álgebra de los operadores (que tienen operadores adjuntos) y las matrices cuadradas (que siempre tienen matrices adjuntas) junto con la operación de “sacar el adjunto” es similar al álgebra de los números complejos con la operación de *conjugar*. Luego, observa de qué formas difieren, ¿a qué se deberán estas diferencias?

15. Operadores normales y autoadjuntos y teorema espectral

Operadores normales

Para determinar si un operador tiene una base ortogonal de eigenvectores y, por lo tanto, si se puede descomponer espectralmente, una herramienta muy útil es entender cómo se relaciona con su operador adjunto. En un primer acercamiento, podemos analizar cómo se relacionan los eigenvalores de uno y otro en el caso de operadores lineales en espacios vectoriales de dimensión finita.

Lema 15.1 15.1.1 Sea T un operador lineal en un espacio vectorial V de dimensión finita con producto escalar. Si T tiene un eigenvector con eigenvalor λ, entonces T^* tiene un eigenvector con eigenvalor $\overline{\lambda}$.

Demostración. Sea $|v\rangle \in V$ un eigenvector de T con eigenvalor λ. Entonces, por el ejercicio 3 de la sección 14, para toda $|u\rangle \in V$ tenemos que

$$0 = \langle 0|u\rangle = \langle (T - \lambda I_V)(v)|u\rangle = \langle v|(T - \lambda I_V)^*(u)\rangle = \langle v|(T^* - \overline{\lambda} I_V)(u)\rangle .$$

Por ende, $|v\rangle \in V$ es un vector no nulo que es ortogonal a todos los vectores de $\operatorname{Im}(T^* - \overline{\lambda} I_V)$. Esto implica que el operador lineal $T^* - \overline{\lambda} I_V$ no es suprayectivo y, por la fórmula de la dimensión, que tiene un núcleo no trivial. Por lo tanto, todo vector no nulo en $\operatorname{Ker}(T^* - \lambda I_V)$ es un eigenvector de T^* con eigenvalor $\overline{\lambda}$. El hecho de que el núcleo sea no trivial nos asegura que existe al menos un eigenvector de T^* con eigenvalor $\overline{\lambda}$. □

El lema anterior nos dice que para cualquier operador T con espectro Λ y operador adjunto T^*, el espectro de T^* se compondrá de los complejos conjugados de los elementos de Λ; regresaremos a él más adelante.

Para continuar con nuestra búsqueda, haremos una observación interesante apoyándonos en el isomorfismo de representación matricial de operadores lineales. Supongamos que T es un operador lineal en un espacio vectorial de dimensión finita V con producto escalar tal que γ es una base ortonormal de V compuesta por eigenvectores de T. Entonces, $[T]_\gamma$ es diagonal y, por el Teorema 14.3, se sigue que

$$[T^*]_\gamma = ([T]_\gamma)^* = \overline{([T]_\gamma^T)} = \overline{([T]_\gamma)},$$

por lo que $[T^*]$ también es diagonal. Ya que las matrices diagonales conmutan, se sigue que $[T]_\gamma[T^*]_\gamma = [T^*]_\gamma[T]_\gamma$. En particular, como este isomorfismo es compatible con la multiplicación de matrices y la composición de transformaciones lineales, tenemos que $[TT^*]_\gamma = [T^*T]_\gamma$, de donde se sigue que $TT^* = T^*T$. Es decir que, si la matriz $[T]_\gamma$ es diagonal, donde γ es una base ortonormal de un espacio vectorial V de dimensión finita, entonces $TT^* = T^*T$, lo nos lleva a las siguientes definiciones.

Def. Sea V un espacio vectorial con producto escalar. Decimos que un operador lineal $T : V \to V$ es *normal* si $TT^* = T^*T$, suponiendo que T^* exista.

Def. Sea $A \in M_{n\times n}(K)$ con $K = \mathbb{R}$ ó $K = \mathbb{C}$. Decimos que A es una matriz *normal* si $AA^* = A^*A$.

Con estas nuevas definiciones, podemos ver directamente del Teorema 14.3 que $[T]_\gamma$ es normal si y sólo si T es normal, donde γ es una base *ortonormal* de un espacio vectorial V de dimensión finita sobre el que actúa T. Más aún, del mismo teorema se sigue que, si dicha base ortonormal γ *diagonaliza* a T, entonces también diagonaliza a T^*, como observamos anteriormente. Es decir que, en este caso, T y T^* son simultáneamente diagonalizables. En particular, esto implica que ambos operadores tienen los mismos eigenvectores —aunque no necesariamente con los mismos eigenvalores correspondientes, como hemos visto en el Lema Lema: 15.1.1 A continuación, veremos varias propiedades que caracterizan a los operadores normales, incluyendo la forma en que se relacionan sus eigenvalores con los de su operador adjunto.

Teorema 15.2 15.1.2 Sean (V, K) un espacio vectorial con producto escalar y T un operador normal en V. Entonces T verifica las siguientes propiedades:

(a) $||T(|x\rangle)|| = ||T^*(|x\rangle)||$;

(b) $T - cI$ es un operador normal para todo $c \in K$;

(c) para todo $|v\rangle \in V$ tal que $T(|v\rangle) = \lambda |v\rangle$, tenemos que $T^*(|v\rangle) = \overline{\lambda} |v\rangle$;

(d) Si λ_1 y λ_2 son eigenvalores distintos de T con eigenvectores correspondientes $|g_1\rangle$ y $|g_2\rangle$, respectivamente, entonces $|g_1\rangle$ y $|g_2\rangle$ son ortogonales.

Demostración. Consultar en el Friedberg y reescribir en notación de Dirac (ejercicio).

□

Después de demostrar el Lema 15.1.1, vimos que si V es un espacio vectorial de dimensión finita y $T : V \to V$ es un operador lineal con una base ortonormal γ de V tal que $[T]_\gamma$ es *diagonal* (i.e., una base ortonormal de V compuesta por eigenvectores de T), entonces T es normal. Se puede demostrar que ambas condiciones son equivalentes, es decir, que existe una base ortonormal de V compuesta por eigenvectores de T si y sólo si T es un operador normal; la implicación contraria se deja como ejercicio. Por ende, tenemos que para operadores lineales T que actúan en un espacio vectorial V de dimensión finita, las siguientes condiciones son equivalentes:

- T es un operador normal (es decir, conmuta con su operador adjunto).
- $[T]_\gamma$ es normal para toda base ortonormal γ de V.
- Existe una base ortonormal γ de V compuesta por eigenvectores de T.
- Existe una base ortonormal γ de V que diagonaliza a T.

En particular, las últimas dos propiedades junto con el inciso (a) del Teorema 14.4 nos dan una pista de por qué a este tipo de operadores se les llama *normales*.

Parece que estamos muy cerca de nuestra meta de encontrar condiciones que aseguren cuándo podemos descomponer espectralmente a un operador lineal; sin embargo, observemos lo siguiente: sean T un operador normal en un espacio vectorial (V, K) de dimensión finita con producto escalar y $|g_i\rangle, |g_j\rangle$ eigenvectores de T con eigenvalores respectivos λ_i, λ_j tales que $\lambda_i \neq \lambda_j$. Entonces, tenemos que

$$\langle g_i | TT^*(g_j)\rangle = \langle T^*(g_i)|(T^*(g_j))\rangle = \langle \overline{\lambda_i} g_i | \overline{\lambda_j} g_j\rangle = \lambda_i \overline{\lambda_j} \langle g_i | g_j\rangle ,$$

mientras que, por otro lado,

$$\langle g_i | TT^*(g_j)\rangle = \langle g_i | T^*T(g_j)\rangle = \langle T(g_i)|T(g_j)\rangle = \langle \lambda_i g_i | \lambda_j g_j\rangle = \overline{\lambda_i} \lambda_j \langle g_i | g_j\rangle = \overline{(\lambda_i \overline{\lambda_j})} \langle g_i | g_j\rangle .$$

Por lo tanto, llegamos a la igualdad

$$\left(\lambda_i \overline{\lambda_j} - \overline{(\lambda_i \overline{\lambda_j})} \right) \langle g_i | g_j\rangle = 0.$$

Notemos que, si $K = \mathbb{C}$, entonces forzosamente $|g_i\rangle$ y $|g_j\rangle$ son ortogonales, de donde se sigue que $E_{\lambda_i}^\perp = \oplus_{j \neq i} E_{\lambda_j}$ para todo eigenvalor λ_i de T —y, por ende, que T se puede descomponer espectralmente, de acuerdo al ejercicio 2 de la sección 14. Sin embargo, si $K = \mathbb{R}$, entonces $\lambda_i \overline{\lambda_j} = \overline{(\lambda_i \overline{\lambda_j})}$ trivialmente, por lo que la igualdad del párrafo anterior siempre se cumple, sin que necesariamente $\langle g_i | g_j\rangle = 0$. Por lo tanto, en un espacio vectorial de dimensión finita y con producto escalar, la condición de que T sea normal es suficiente para que se pueda descomponer espectralmente *siempre y cuando el espacio sea complejo* mientras que, si el espacio es real, *esta condición no es suficiente*. Esto nos lleva a estudiar un tipo particular de operadores normales, que veremos a continuación.

Operadores autoadjuntos

Como demostraremos más adelante, la condición suficiente para que un operador lineal que actúa en un espacio vectorial *real* de dimensión finita y con producto escalar se pueda descomponer espectralmente es que sea *igual* a su operador adjunto. Por ende, damos las siguientes definiciones.

Def. Sea V un espacio vectorial con producto escalar. Decimos que un operador lineal $T : V \to V$ es *autoadjunto* o *hermitiano*[a] si $T = T^*$, suponiendo que T^* exista.

Def. Sea $A \in M_{n\times n}(K)$ con $K = \mathbb{R}$ ó $K = \mathbb{C}$. Decimos que A es una matriz *autoadjunta* o *hermitiana* si $A = A^*$.

[a]Generalmente, el término *autoadjunto* es el predominante en matemáticas, mientras que el término *hermitiano* es preferido en física. A pesar de eso, son equivalentes, y en estas notas utilizaremos ambos términos de manera intercambiable.

Nótese que, para el caso de matrices cuadradas con entradas reales, la condición de que una matriz sea autoadjunta se reduce a $A = A^T$, es decir, que sea simétrica. Además, por un desarrollo similar al de la sección 15, podemos ver que T es un operador hermitiano si y sólo si $[T]_\gamma$ es una matriz hermitiana, con γ una base ortonormal arbitraria de V.

Lema 15.3 15.2.1 Sea T un operador autoadjunto en un espacio vectorial de dimensión finita con producto escalar V. Entonces:

(a) Todos los eigenvalores de T son reales.

(b) Si V es un espacio vectorial real, entonces el polinomio característico de T es separable.

Demostración. (a) Sean λ un eigenvalor de T y $|v\rangle$ un eigenvector correspondiente a λ. Entonces, por el Teorema 15.1.2.(c), tenemos que

$$\lambda\,|v\rangle = T(|v\rangle) = T^*(|v\rangle) = \overline{\lambda}\,|v\rangle \implies \lambda = \overline{\lambda},$$

de donde se sigue que $\lambda \in \mathbb{R}$. Por lo tanto, todos los eigenvalores del operador autoadjunto T son reales.

(b) Consultar en el Friedberg y reescribir en notación de Dirac. □

Con el lema anterior en mente, podemos ver el siguiente resultado.

Teorema 15.4 15.2.2 Sean V un espacio vectorial real de dimensión finita n con producto escalar y $T : V \to V$ un operador lineal. Entonces, T es autoadjunto si y sólo si existe una base ortonormal γ de V compuesta por eigenvectores de T.

Demostración. ($\Rightarrow$) Ejercicio. (Pista: utiliza el lema anterior.).

($\Leftarrow$) Supongamos que existe una base ortonormal $\gamma = \{|g_1\rangle, |g_2\rangle, \dots, |g_n\rangle\}$ de V compuesta por eigenvectores de T. Sea $\Lambda = \{\lambda_1, \lambda_2, \dots, \lambda_n\}$ el espectro de T tal que $T(|g_i\rangle) = \lambda_i\,|g_i\rangle\ \forall\, i \in \{1, 2, \dots, n\}$.

Ya que V es un espacio vectorial real, entonces $\overline{\lambda}_i = \lambda_i\ \ \forall\ \lambda_i \in \Lambda$. Ahora, observemos que

$$\forall\, i \in \{1,2,\dots,n\},\ \ \langle g_i|T(g_i)\rangle = \langle g_i|\lambda_i g_i\rangle = \langle g_i|\,(\lambda_i\,|g_i\rangle) = (\langle g_i|\,\lambda_i)\,|g_i\rangle = \langle\overline{\lambda}_i g_i|g_i\rangle = \langle\lambda_i g_i|g_i\rangle = \langle T(g_i)|g_i\rangle$$
$$\text{y, } \forall\, j \neq i,\ \ \langle g_j|T(g_i)\rangle = \langle g_j|\lambda_i g_i\rangle = \lambda_i\,\langle g_j|g_i\rangle = \lambda_j\,\langle g_j|g_i\rangle = \langle\overline{\lambda}_j g_j|g_i\rangle = \langle\lambda_j g_j|g_i\rangle = \langle T(g_j)|g_i\rangle\,.$$

Por lo tanto, el operador lineal T es tal que $\langle g_j|g_i\rangle = \langle g_j|g_i\rangle$ para cualesquiera elementos $|g_i\rangle, |g_j\rangle$ de la base γ de V, de donde se sigue que

$$\langle y|T(x)\rangle = \langle T(y)|x\rangle\ \ \ \forall\ \ |x\rangle, |y\rangle \in V.$$

Como V es de dimensión finita, por el Teorema 14.2 se sigue que $T : V \to V$ es el único operador con esta propiedad. Por ende, concluimos que T es autoadjunto. □

Teorema espectral

Concluida nuestra búsqueda por las hipótesis precisas bajo las cuales podemos llevar a cabo la descomposición espectral, enunciamos a continuación un teorema que recopila y sintetiza muchos resultados que hemos demostrado a lo largo de este módulo, y que utiliza una inmensa cantidad de conceptos que hemos aprendido durante todo el curso. La demostración de este teorema se deja como ejercicio.

Teorema 15.5 15.3.1 (Teorema espectral) Sea (V, K) un espacio vectorial de dimensión finita con producto escalar y $T : V \to V$ un operador lineal con eigenvalores distintos $\lambda_1, \lambda_2, ..., \lambda_k$. Supongamos que T es normal si $K = \mathbb{C}$ y que T es autoadjunto si $K = \mathbb{R}$. Sea W_i el eigenespacio de T correspondiente al eigenvalor λ_i, con $1 \leq i \leq k$, y sea T_i la proyección ortogonal de V sobre W_i. Entonces, se cumple que:

a) $V = W_1 \oplus W_2 \oplus ... \oplus W_k$;

b) si $W_i' = \oplus_{j \neq i} W_j$, entonces $W_i^{\perp} = W_i'$;

c) $T_i T_j = \delta_{ij} T_i$ para $1 \leq i, j \leq k$;

d) $I = T_1 + T_2 + ... + T_k$;

e) $T = \lambda_1 T_1 + \lambda_2 T_2 + ... + \lambda_k T_k$.

Demostración. Ejercicio.

□

Ejercicios de repaso

Operadores normales

1. Demuestra que cualesquiera dos matrices diagonales con entradas en un campo K conmutan y que, por lo tanto, los operadores lineales que representan dichas matrices también deben conmutar.

2. Demuestra el Teorema 15.1.2.

Operadores autoadjuntos

1. Completa la demostración del Lema 15.2.1.

2. Completa la demostración del Teorema 15.2.2.

Teorema espectral

1. Demuestra el teorema espectral.

Sobre la descomposición espectral de matrices...

A estas alturas del curso, nos encontramos en una excelente posición para volver a hacer énfasis en un punto de fundamental importancia en el álgebra lineal que muchas veces es pasado por alto: a simple vista, *las matrices no cuentan toda la historia*[a].

Pensemos un momento en la teoría de descomposición espectral: para encontrar a un operador que se pueda descomponer espectralmente —es decir, que proyecte vectores del dominio sobre eigenespacios ortogonales, reescale cada componente por el eigenvalor correspondiente al eigenespacio en que se encuentra, y construya a partir de estas proyecciones ortogonales reescaladas a los vectores imagen— requerimos poder formar una base ortonormal del espacio a partir de eigenvectores del operador.

Supongamos que tenemos un operador que cumple las características deseadas y actúa sobre un espacio vectorial de dimensión finita[b]. ¿Cómo se ve la representación matricial de nuestro operador en su base de eigenvectores? Pues, resulta que es una matriz diagonal, como cualquier otra...

¿Cómo puede ser? ¡Esto debe ser mentira! En efecto, lo es: realmente no es como *cualquier* matriz diagonal arbitraria, ya que la nuestra, en el fondo, *está representada en una base ortonormal* del espacio— pero eso no se ve a simple vista.

En realidad, el punto sobre el que hago énfasis es el de las *representaciones*. Dado el poder que tienen las herramientas de cambio de representación en el álgebra lineal (permitiéndonos diagonalizar matrices, por ejemplo), cuando veamos n-tuplas y matrices, sería erróneo de nuestra parte siempre asumir que se encuentran en una base ortonormal —como suele manejarse implícitamente cuando trabajamos en espacios euclideanos. En cambio, al ver n-tuplas o matrices, debemos saber que lo que estamos viendo son apenas *representaciones* de vectores y que, ya que dichas representaciones dependen totalmente de la base elegida (o *las bases elegidas*, en el caso de las matrices, las cuales representan transformaciones lineales que, como sabemos, son vectores), el tener presentes las características de la(s) base(s) en la(s) que se esté representando pueden darnos información crucial acerca de los vectores representados; ¡no lo olviden!

[a] Esto es similar a cómo, en primera instancia, las n-tuplas que podemos utilizar para representar vectores abstractos —a través de isomorfismos— tampoco nos cuentan la historia completa.

[b] Aquí ya está implícito que el espacio en cuestión debe tener producto escalar. ¿Por qué?

16. Operadores unitarios y ortogonales

A lo largo del curso, hemos estudiado funciones que preservan la estructura algebráica del espacio subyaciente, empezando por las transformaciones lineales —que preservan las operaciones esenciales de los espacios vectoriales (suma vectorial y producto de un vector por un escalar)— y los isomorfismos —que, al ser transformaciones lineales *invertibles*, preservan la estructura de todo el espacio vectorial, estableciendo una correspondencia biunívoca entre dos espacios vectoriales que, en *esencia*, son lo mismo. Para cerrar el curso, dado que durante el último módulo nos hemos enfocado en espacios vectoriales con producto escalar, estudiaremos operadores lineales invertibles que preservan el producto escalar. Como veremos, en espacios vectoriales de dimensión finita, esto es equivalente a que dichos operadores preserven la norma inducida por el producto escalar. Además, veremos que algunos operadores de este tipo pueden descomponerse espectralmente[24], aplicando lo aprendido en la sección anterior.

Def. Sea T un operador lineal sobre un espacio vectorial (V, K) de dimensión finita con producto escalar. Si T es tal que

$$||T(|v\rangle)|| = ||\,|v\rangle\,|| \quad \forall\ |v\rangle \in V,$$

entonces diremos que T es un operador *unitario* si $K = \mathbb{C}$ o un operador *ortogonal* si $K = \mathbb{R}$[a].

[a]Para el caso de dimensión infinita, T debe de preservar la norma *y ser invertible* para que se lo considere *unitario* si $K = \mathbb{C}$ u *ortogonal* si $K = \mathbb{R}$; si sólamente preserva la norma, se le conoce como una *isometría*. Claramente, esto indica que los operadores unitarios u ortogonales en espacios de dimensión finita son invertibles.

En $\mathbb{R}^3$ (y, más generalmente, en $\mathbb{R}^n$), las rotaciones alrededor de y las reflexiones con respecto a cualquier eje que pase por el origen son ejemplos de operadores ortogonales[25]. En el espacio vectorial complejo $\mathbb{C}$ (más generalmente, $\mathbb{C}^n$), cualquier operador que reescale por un número complejo de norma (valor absoluto) uno —o, en otras palabras, por un número complejo *unitario*— es un ejemplo de un operador unitario[26].

A continuación, demostraremos un lema que nos ayudará a encontrar una serie de criterios equivalentes a la definición anterior de un operador unitario u ortogonal. Las demostraciones del lema y del teorema posterior quedan como ejercicio.

Lema 16.1 16.1 Sea U un operador autoadjunto en un espacio vectorial V de dimensión finita con producto escalar. Entonces, si $\langle v|U(v)\rangle = 0\ \ \forall\ |v\rangle \in V$, entonces $U = T_0$, la transformación nula.

Demostración. Ejercicio. □

Teorema 16.2 16.2 Sea T un operador lineal en un espacio vectorial de dimensión finita con producto escalar. Entonces, la siguientes afirmaciones son equivalentes.

(a) $||T(|v\rangle)|| = ||\,|v\rangle\,||\ \ \forall\ |v\rangle \in V$.

(b) $\langle T(v)|T(v)\rangle = \langle v|v\rangle\ \ \forall\ |v\rangle \in V$.

(c) $TT^* = T^*T = I_V$.

(d) $\langle T(u)|T(v)\rangle = \langle u|v\rangle\ \ \forall\ |u\rangle, |v\rangle \in V$.

(e) Si β es una base ortonormal de V, entonces $T(\beta)$ también lo es.

[24]¿Cómo imaginas que serán sus espsectros? Pista: cambia dependiendo de si el campo es real o complejo.
[25]¿Recuerdas haber visto un operador de este tipo en tu segundo examen?
[26]¡Recuerda tu primer examen!

(f) Existe una base ortonormal β de V tal que $T(\beta)$ también es una base ortonormal de V.

Demostración. Ejercicio. □

Observemos de la primera igualdad del inciso (c) que todo operador unitario u ortogonal es normal y, por ende, los operadores unitarios pueden descomponerse espectralmente; además, a partir de este inciso, podemos hacer una analogía entre los operadores unitarios/ortogonales en un espacio vectorial con producto escalar y los elementos $z \in \mathbb{C}$ tales que $z\overline{z} = \overline{z}z = 1$.

Por otro lado, ya que la norma inducida por el producto escalar es escalable de forma absoluta, entonces se sigue que todos los eigenvalores de un operador unitario/ortogonal tienen valor absoluto 1 directamente de la definición. Es decir, sus eigenvalores son precisamente aquellos escalares $\lambda \in \mathbb{C}$ ó $\lambda \in \mathbb{R}$ tales que $\lambda\overline{\lambda} = \overline{\lambda}\lambda = 1$. Finalmente, aplicamos el Teorema Espectral a los operadores unitarios y ortogonales en los siguientes corolarios.

Corolario 16.3 16.3 Sean $T : V \to V$ un operador unitario sobre un espacio vectorial complejo de dimensión finita n con producto escalar y $\Lambda = \{\lambda_1, \lambda_2, ..., \lambda_k\}$ el espectro de T. Entonces, tenemos que $T = \lambda_1 T_1 + \lambda_2 T_2 + ... + \lambda_k T_k$, con T_i el operador de proyección ortogonal sobre el eigenespacio E_{λ_i}, y $\Lambda \subset \{z \in \mathbb{C} \mid z\overline{z} = \overline{z}z = 1\}$.

Demostración. Ejercicio. □

Corolario 16.4 16.4 Sean $T : V \to V$ un operador ortogonal sobre un espacio vectorial real de dimensión finita n con producto escalar y $\Lambda = \{\lambda_1, \lambda_2, ..., \lambda_k\}$ el espectro de T. Entonces, si T es autoadjunto, tenemos que $T = \lambda_1 T_1 + \lambda_2 T_2 + ... + \lambda_k T_k$, con T_i el operador de proyección ortogonal sobre el eigenespacio E_{λ_i}, y $\Lambda \subseteq \{\pm 1\}$.

Demostración. Ejercicio. □

Ejercicios de repaso

Operadores lineales ortogonales y unitarios

1. Sean (V, K) un espacio vectorial de dimensión finita con producto escalar y $T : V \to V$ un operador lineal tal que T es unitario si $K = \mathbb{C}$ y T es ortogonal si $K = \mathbb{R}$. Demuestra que T es invertible.

2. Da un ejemplo de un operador ortogonal $T : \mathbb{R}^2 \to \mathbb{R}^2$ que no tenga eigenvectores y que, por lo tanto, no se pueda formar una base ortogonal de $\mathbb{R}^2$ compuesta por eigenvectores de T. Esto ilustra la importante diferencia entre operadores ortogonales y autoadjuntos.

3. Da un ejemplo de un operador normal $U : \mathbb{C} \to \mathbb{C}$ que no sea unitario.

4. Demuestra el Lema 16.1.

5. Demuestra el Teorema 16.2.

6. Demuestra el Corolario 16.3.

7. Demuestra el Corolario 16.4.