

# Curso de Álgebra Lineal (2020-II)

Diego Alberto Barceló Nieves  
Licenciatura en Física Biomédica  
Facultad de Ciencias  
Universidad Nacional Autónoma de México

## 1. Introducción

La presente planeación del curso de Álgebra Lineal para la Licenciatura en Física Biomédica fue realizada tomando en cuenta el programa de la asignatura existente<sup>1</sup>, así como la retroalimentación del mismo por parte de estudiantes de la licenciatura.

Durante el curso, se sentarán bases teóricas sólidas de álgebra lineal para diversas aplicaciones que serán útiles en materias posteriores del plan de estudios, tales como Ecuaciones Diferenciales, Matemáticas Avanzadas, Introducción a la Física Cuántica y Mecánica Cuántica, Algoritmos Computacionales y Física Computacional, entre otras.

### 1.1. Metodología de enseñanza

Este curso está diseñado para desarrollar las habilidades autodidactas de cada estudiante, a la vez que se les da acompañamiento en los temas del curso tanto dentro como fuera del salón de clase. Desde el principio del curso, los estudiantes tendrán acceso a las notas de todo el curso, que incluyen ejercicios de repaso para reforzar los aprendizajes: esto permitirá que cada quien avance a su ritmo, y que puedan estudiar los temas que veremos en clase en caso de no asistir. Habrá cinco clases semanales en las cuales se presentarán los temas y se fomentará la discusión en grupo, así como horarios fijos establecidos fuera de clase para atender dudas de la materia de forma más personalizada (posiblemente, de manera virtual).

### 1.2. Forma de evaluación

La evaluación se hará a través de 5 exámenes parciales. Al final del curso, quienes tengan un promedio reprobatorio en la materia deberán presentar un examen final, que sustituirá sus calificaciones anteriores. Quienes tengan un promedio aprobatorio podrán realizar como máximo un examen de reposición de alguno de los parciales: una vez devuelto un examen parcial, para tener derecho a elegir ese parcial como su examen a reponer al término del curso, deberán entregar una corrección completa de dicho examen como máximo una semana después de que se les haya entregado retroalimentado y calificado. Las personas con promedio aprobatorio también podrán optar por hacer el examen final para subir calificación; en ese caso, su calificación en el final reemplazará las anteriores.

Para definir la calificación final del curso, en los casos donde la calificación obtenida sea muy cercana a la calificación siguiente, se tomará en cuenta el esfuerzo extra realizado por cada estudiante durante el semestre mediante la entrega de tareas “morales” (opcionales) —que se dejarán para reforzar lo visto en clase— y/o correcciones de exámenes para definir un posible cambio.

---

<sup>1</sup>El cual puede ser consultado en la página <http://www.fciencias.unam.mx/asignaturas/1330.pdf>.

## 2. Cronograma tentativo y bibliografía

### 2.1. Cronograma tentativo del curso

Semana	Temas
1	Campos, espacios vectoriales, interpretación geométrica de las operaciones esenciales de los espacios vectoriales y subespacios vectoriales
2	Producto escalar (punto), norma, interpretación geométrica del producto escalar: proyecciones y ortogonalidad, producto vectorial (cruz)* y triple producto escalar*
3	Combinaciones lineales, subespacio generado y conjunto generador, dependencia e independencia lineal
4	Bases, dimensión, ortogonalización y ortonormalización
5	Definición de transformación lineal y espacio de transformaciones lineales, núcleo e imagen de una transformación lineal, nulidad y rango
6	Composición de transformaciones lineales, transformación lineal inversa y espacios vectoriales isomorfos y representación matricial de una transformación lineal
7	Espacios vectoriales de matrices, multiplicación de matrices, representación de un vector en una base ordenada, matriz de cambio de base y matrices similares
8	Definición y propiedades del determinante y la traza de una matriz, interpretación geométrica del determinante
9	Matriz transpuesta, matriz inversa e invertibilidad de matrices
10	Matriz diagonal, eigenvalores, eigenvectores y polinomio característico
11	Diagonalizabilidad, eigenespacios y subespacios invariantes
12	Definición de operador adjunto, funcional (vector dual) y espacio dual
13	Definición y propiedades de operadores hermitianos
14	Descomposición espectral de transformaciones simétricas
15	Sistemas de ecuaciones lineales algebraicas y diferenciales, espacio de soluciones y subespacios invariantes
16	Espacios de Hilbert, conmutadores y producto tensorial

Tabla 1: Cronograma semanal tentativo del curso. Los temas marcados con asterisco (\*) no se verán en clase, pero sí se mencionarán en las notas y aparecerán ejercicios relacionados a ellos en las tareas, los cuales tendrán valor de *puntos extra*.

### 2.2. Bibliografía recomendada para el curso

- S. H. Friedberg, *Linear Algebra*, 4ta ed. - es el texto básico para este tipo de cursos.
- S. Lang, *Introduction to Linear Algebra*, 2da ed. (Springer, 1986, EUA) - otro texto básico que cubre la muchos de los temas del curso y complementa el Friedberg.
- S. Lang, *Linear Algebra*, 3a ed. (Springer, 1987, EUA) - no es tan básico como *Introduction to Linear Algebra*, pero sirve para *repasar* temas básicos y aprender algunos más avanzados.
- D. Poole, *Linear Algebra: A Modern Introduction*, 4ta ed. (Cengage Learning, 2015, EUA) - este texto básico ve los temas a través de diversos ejemplos y aplicaciones; es útil para las personas que quieran ver algunas aplicaciones de los conceptos al mismo tiempo que los aprenden.

Les sugiero que hojeen *todos* los libros recomendados al inicio del curso, y que consulten los de su agrado constantemente durante el mismo, o bien, busquen otros que les sirvan mejor para aprender.

## 2.3. Otros recursos educativos

No sólo se aprende de libros; hay que aprovechar todo el contenido que ofrece el internet para nuestra educación. A lo largo de los apuntes pondré hipervínculos a algunas páginas con recursos relevantes para el tema en cuestión; sin embargo, aquí enlistaré algunos recursos útiles para aprender álgebra lineal. Por supuesto, les invito a que busquen más recursos por su propia cuenta; de encontrarlos, les agradecería que me notificaran para revisarlos.

- [https://www.youtube.com/watch?v=fNk\\_zzaMoSs&list=PLZHQObOWTQDPD3MizzM2xVFitgF8hE\\_ab](https://www.youtube.com/watch?v=fNk_zzaMoSs&list=PLZHQObOWTQDPD3MizzM2xVFitgF8hE_ab) - Lista de reproducción *Essence of Linear Algebra* del canal de YouTube 3Blue1Brown.
- <http://immersivemath.com/ila/learnmore.html> y <https://textbooks.math.gatech.edu/ila/index2.html> - Libros de texto *interactivos* que sirven para generar intuición acerca de algunos conceptos básicos del álgebra lineal.
- <https://www.lem.ma/books/AIApowDnj1DDQrp-u0ZVow/landing> y <https://www.khanacademy.org/math/linear-algebra> - Series de videos sobre álgebra lineal con interfaces para resolver ejercicios al final de cada lección.

## Notación

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

---

<sup>a</sup>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.

### 3. Campos, espacios vectoriales, interpretación geométrica de las operaciones esenciales de los espacios vectoriales y subespacios vectoriales

#### 3.1. Campos

Uno de los conceptos más fundamentales del álgebra lineal es el de espacio vectorial; sin embargo, para definir formalmente un espacio vectorial se requiere de una estructura algebraica conocida como campo —el ejemplo más común siendo el *campo de los números reales*. Esta estructura quizá la viste 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— *para los números reales*); sin embargo, aquí mencionaremos su definición y los dos ejemplos de campos que más utilizaremos durante este curso (el campo de los números reales y el de los complejos) antes de definir los espacios vectoriales.

##### 3.1.1. Definición de campo

Def. Un *campo* es un conjunto  $K$  con dos operaciones (llamadas *adición* o *suma* y *multiplicación*) que cumplen las propiedades siguientes:

$$\forall a, b \in K \quad a + b = b + a; \quad ab = ba \quad \text{Conmutatividad}$$

$$\forall a, b, c \in K \quad a + (b + c) = (a + b) + c; \quad a(bc) = (ab)c \quad \text{Asociatividad}$$

$$\exists 0, 1 \in K \text{ t.q. } a + 0 = a, \quad 1a = a \quad \forall a \in K \quad \text{Elementos identidad (neutros)}$$

$$\forall a \in K \quad \exists -a \in K \text{ t.q. } a + (-a) = 0 \quad \text{Elemento inverso de la suma}$$

$$\forall a \neq 0 \in K \quad \exists a^{-1} \in K \text{ t.q. } a(a^{-1}) = 1 \quad \text{Elemento inverso de la multiplicación}$$

$$\forall a, b, c \in K \quad a(b + c) = ab + ac \quad \text{Distributividad.}$$

A estas propiedades también se les conoce como *axiomas de campo*<sup>a</sup>. Por otro lado, a pesar de que la definición de campo incluya un conjunto  $K$  y dos operaciones que cumplen las propiedades anteriormente mencionadas, por simplicidad, se suele denotar a todo el campo como  $K$ .

<sup>a</sup>Nótese que la última de ellas (distributividad) es la única que *combina* ambas operaciones definidas en el campo

##### 3.1.2. El campo real

El conjunto de los números reales  $\mathbb{R}$  junto con las operaciones de suma y multiplicación (que aprendimos de forma intuitiva durante nuestra educación básica) cumplen todas las propiedades enlistadas en la sección 3.1.1, ya que dichas operaciones son conmutativas, asociativas y cumplen la propiedad de distributividad. El elemento identidad (neutro) de la suma es  $0 \in \mathbb{R}$  y el de la multiplicación es  $1 \in \mathbb{R}$ . Para todo número  $a \neq 0 \in \mathbb{R}$ , el elemento inverso de la multiplicación es  $a^{-1} = \frac{1}{a} \in \mathbb{R}$ . Al conjunto  $\mathbb{R}$  junto con estas dos operaciones se le conoce como el *campo de los números reales* o, simplemente, el *campo real*.

### 3.1.3. El campo complejo

El conjunto de los números complejos se define como  $\mathbb{C} \equiv \{a + ib \mid a, b \in \mathbb{R}\}$ <sup>2</sup>, donde  $i \equiv +\sqrt{-1}$ . Definiendo la suma entre números complejos como  $(a + ib) + (q + ir) \equiv (a + q) + i(b + r)$ , la multiplicación entre números complejos como<sup>3</sup>  $(a + ib)(q + ir) \equiv (aq - br) + i(ar + bq)$  y apoyándonos en el campo real, podemos comprobar que estas dos operaciones junto con el conjunto  $\mathbb{C}$  forman un campo. A los números complejos de la forma  $0 + ia, a \in \mathbb{R}$  se les conoce como números *imaginarios* o *laterales*<sup>4</sup>. A continuación, demostraremos que el conjunto  $\mathbb{C}$  con las dos operaciones definidas previamente cumplen las primeras tres propiedades de los campos enlistadas en la sección 3.1.1.

Sean  $a, b, q, r, s, t$  elementos del campo  $\mathbb{R}$  tal que  $a + ib, q + ir, s + it \in \mathbb{C}$ .

Ya que  $a, b, q, r$  pertenecen al campo  $\mathbb{R}$ , entonces las sumas y multiplicaciones entre estos elementos son conmutativas. Por lo tanto, tenemos que

$$(a + ib) + (q + ir) = (a + q) + i(b + r) = (q + a) + i(r + b) = (q + ir) + (a + ib),$$

por lo cual la suma en  $\mathbb{C}$  que definimos en el párrafo anterior es conmutativa, y que

$$(a + ib)(s + it) = (as - bt) + i(bs + at) = (sa - tb) + i(sb + ta) = (s + it)(a + ib),$$

por lo cual la multiplicación en  $\mathbb{C}$  que definimos también es conmutativa.

Nuevamente, ya que  $a, b, r, q, s, t$  pertenecen al campo  $\mathbb{R}$ , entonces las sumas y multiplicaciones entre estos elementos son asociativas. Por ende, tenemos que

$$\begin{aligned} (a + ib) + ((q + ir) + (s + it)) &= (a + ib) + ((q + s) + i(r + t)) = (a + q + s) + i(b + r + t) = \\ &= ((a + q) + i(b + r)) + (s + it) = ((a + ib) + (q + ir)) + (s + it), \end{aligned}$$

por lo cual la suma en  $\mathbb{C}$  es asociativa, y que

$$\begin{aligned} (a + ib)((q + ir)(s + it)) &= (a + ib)((qs - rt) + i(qt + rs)) = ((a(qs - rt) - b(qt + rs)) + i(a(qt + rs) + b(qs - rt))) = \\ &= (aqs - art - bqt - brs) + i(aqt + ars + bqs - brt) = (aqs - brs - art - bqt) + i(aqt - brt + ars + bqs) = \\ &= ((aq - br)s - (ar + bq)t) + i((aq - br)t + (ar + bq)s) = ((aq - br) + i(ar + bq))(s + it) = ((a + ib)(q + ir))(s + it), \end{aligned}$$

por lo cual concluimos que la multiplicación también es asociativa.

Por otro lado, recordando la definición de  $\mathbb{C} = \{a + ib \mid a, b \in \mathbb{R}\}$  vemos que, en particular,  $0 + i0 \in \mathbb{C}$  y  $1 + i0 \in \mathbb{C}$ . Observemos que

$$(a + ib) + (0 + i0) = ((a + 0) + i(b + 0)) = (a + ib),$$

por lo cual existe un elemento identidad de la suma (neutro aditivo) en este campo, y que

$$(1 + i0)(q + ir) = (1q - 0r + i(1r + 0q)) = ((q - 0) + i(r + 0)) = (q + ir),$$

por lo cual también existe un elemento identidad de la multiplicación (neutro multiplicativo); por simplicidad, escribimos  $0 + i0$  como  $0$  y  $1 + i0$  como  $1$ : se pueden identificar con los elementos neutros del campo real. Las últimas tres propiedades de la definición de campo de la sección 3.1.1 se demuestran

<sup>2</sup>Notemos que, por definición, el conjunto  $\mathbb{R}$  es subconjunto de  $\mathbb{C}$ , i.e.,  $\mathbb{R} \subset \mathbb{C}$ .

<sup>3</sup>Nótese que esta definición es simplemente el resultado de desarrollar  $(a + ib)(c + id)$  como un producto de dos binomios.

<sup>4</sup>El mismísimo Gauss hubiera preferido este segundo nombre, ya que creía que era mucho más intuitivo, y que llamarlos *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: <https://www.youtube.com/watch?v=T647CGsu0VU>.

de manera similar, y se dejan como ejercicio. A este campo se le conoce como el *campo de los números complejos* o *campo complejo*.

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

## 3.2. Espacios vectoriales

Un espacio vectorial es una estructura algebraica abstracta que cumple una serie de propiedades específicas que veremos en el siguiente apartado. Dicha estructura tiene una gran variedad de aplicaciones en muchas áreas de las matemáticas, la física, la computación y la biomedicina, por lo cual, para arrancar el curso, es vital su comprensión desde un punto de vista teórico.

### 3.2.1. Definición de espacio vectorial

Def. Un *espacio vectorial* sobre un campo<sup>a</sup>  $K$  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:

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

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

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

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

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

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

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

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

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

$(a + b)\mathbf{v} = a\mathbf{v} + b\mathbf{v} \quad \forall a, b \in K, \mathbf{v} \in V$  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* y los elementos  $\mathbf{u}, \mathbf{v}, \mathbf{w} \in V$  que cumplen todas las propiedades anteriores son llamados *vectores*. A las propiedades anteriores también se les conoce como *axiomas de espacio vectorial*.

---

<sup>a</sup>Ver sección 3.1.1.

Partiendo de esta definición, podemos hacer varias observaciones:

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



definición de vector como *elemento con magnitud, dirección y sentido* (y, más precisamente, que además es invariante bajo rotaciones propias e impropias) utilizada en algunas áreas de la física, siendo la primera definición más general.

- La definición de *espacio vectorial* incluye dos operaciones *nuevas* (con respecto a las operaciones de campo) con una importante diferencia entre ellas: una es solamente entre los elementos del conjunto  $V$  (adición o suma vectorial) y, la otra, entre los elementos del conjunto  $V$  y el campo  $K$  (producto de un vector por un escalar)<sup>5</sup>. Sin embargo, *ambas dan como resultado un vector en  $V$* .
- 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. Por simplicidad, al campo se le denota como  $K$  y al espacio vectorial, como  $V$ .

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](https://www.youtube.com/watch?v=fNk_zzaMoSs).

### 3.2.2. Ejemplos de espacios vectoriales

El producto cartesiano del conjunto de los números reales consigo mismo  $\mathbb{R} \times \mathbb{R}$  (o  $\mathbb{R}^2$ ) sobre el campo  $\mathbb{R}$  es un espacio vectorial, ya que los elementos de  $\mathbb{R}^2$  (conocidos como *pares ordenados* o *2-tuplas*) junto con los de  $\mathbb{R}$  satisfacen la definición de la sección 3.2.1. Usualmente, en geometría analítica, estos pares ordenados se escriben en forma de *coordenadas*  $(x, y)$  con  $x, y \in \mathbb{R}$  y tienen una correspondencia uno a uno con *puntos* en el plano cartesiano; sin embargo, en álgebra lineal, cuando hablemos de pares ordenados como *vectores* de  $\mathbb{R}^2$ , es preferible emplear la notación<sup>6</sup>  $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \in \mathbb{R}^2$ ,  $x_1, x_2 \in$

$\mathbb{R}$  (o, equivalentemente,  $\begin{bmatrix} x_1 & x_2 \end{bmatrix}^T$ ). Estos vectores se pueden representar visualmente mediante una correspondencia uno a uno con *flechas* en el plano cartesiano, las cuales *tienen su cola en el origen y su punta en el punto correspondiente a las coordenadas  $(x_1, x_2)$* . La suma vectorial se define naturalmente como  $\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} x_1 + y_1 \\ x_2 + y_2 \end{bmatrix}$ . El elemento identidad de la suma (neutro aditivo) es el *vector origen*

$\mathbf{0} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$  y el inverso aditivo de un vector  $\mathbf{u} = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$  es  $\begin{bmatrix} -u_1 \\ -u_2 \end{bmatrix}$ , denotado como  $-\mathbf{u}$ . En este caso, el producto de un vector  $\mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} \in \mathbb{R}^2$  por un escalar  $a \in \mathbb{R}$  se define como  $a\mathbf{v} = a \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} \equiv \begin{bmatrix} av_1 \\ av_2 \end{bmatrix}$  (o  $a \begin{bmatrix} v_1 & v_2 \end{bmatrix}^T \equiv \begin{bmatrix} av_1 & av_2 \end{bmatrix}^T$ ) y el elemento de identidad de esta operación es el escalar  $1 \in \mathbb{R}$ .

El espacio vectorial de  $\mathbb{R}^3$  sobre  $\mathbb{R}$  se define de manera análoga. Los vectores de  $\mathbb{R}^3$  también se pueden representar como flechas que parten del origen de un espacio tridimensional, en una correspondencia uno a uno con las coordenadas del espacio tridimensional. Este tipo de espacios vectoriales se pueden generalizar de manera abstracta (i.e., no visual), como veremos en el siguiente ejemplo.

<sup>5</sup>Más adelante veremos otras operaciones que se pueden definir entre vectores y escalares, pero las dos que hemos visto hasta ahora son las únicas necesarias para *definir* los espacios vectoriales. Por ende, más adelante podremos referirnos a ellas como las operaciones *esenciales* de los espacios vectoriales.

<sup>6</sup>A esto se le conoce como *escribir un vector como una columna de matriz* (ver el recuadro 2.3) y facilita la notación al momento de hacer productos de matrices por la izquierda con vectores, como haremos más adelante en el curso.

El conjunto obtenido al realizar un producto cartesiano de un número entero positivo  $n$  de conjuntos  $\mathbb{R}$  ( $\mathbb{R} \times \mathbb{R} \times \dots \times \mathbb{R} = \mathbb{R}^n$ ) sobre el campo  $\mathbb{R}$  también es un espacio vectorial. Sus elementos vectoriales son de la forma  $\mathbf{v} = [v_1 \ v_2 \ \dots \ v_n]^T$ ,  $v_i \in \mathbb{R}$ ,  $i \in \{1, 2, \dots, n\}$  y son conocidos como *n-tuplas*<sup>7</sup>. Las operaciones entre los vectores de  $\mathbb{R}^n$  y con los escalares en  $\mathbb{R}$  se definen de manera análoga al ejemplo de  $\mathbb{R}^2$ . A este tipo de espacios vectoriales les llamamos *espacios vectoriales reales*, de acuerdo con la definición siguiente:

Def. Un *espacio vectorial real* es aquel definido sobre el campo  $\mathbb{R}$  (campo real) o, equivalentemente, aquel donde los escalares son números reales.

El conjunto de todas las funciones polinomiales de una variable real de grado  $n$  (i.e., con regla de correspondencia de la forma  $f(x) = c_1x^1 + c_2x^2 + \dots + c_nx^n$ ,  $c_i \in \mathbb{R}$ ,  $i \in \{1, 2, \dots, n\}$ ) y con un mismo dominio  $\mathcal{D}$  forma un espacio vectorial sobre el campo  $\mathbb{R}$ . 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 ( $(f + g)(x) \equiv f(x) + g(x)$ ) y del producto de una función arbitraria  $f(x)$  por una función constante  $a$ , respectivamente, vistas en cálculo —las cuales aplican para la 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  $\mathbb{R}$  forma un espacio vectorial<sup>8</sup>. Esto probablemente lo viste de manera implícita en tu curso de cálculo diferencial de una variable, cuando viste los teoremas de derivada 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.

El conjunto  $\mathbb{C} \times \mathbb{C}$  ( $\mathbb{C}^2$ ) sobre el campo  $\mathbb{C}$  también es un espacio vectorial. Sus vectores son de la forma  $[a + ib \ c + id]^T$  con  $a, b, c, d \in \mathbb{R}$  e  $i \equiv +\sqrt{-1}$  (ya que esto implica que  $a + ib, c + id \in \mathbb{C}$ , es decir, que sus entradas son complejas). El elemento identidad de la suma vectorial es  $\mathbf{0} = [0 + i0 \ 0 + i0]^T$  y el del producto de un vector por un escalar es  $1 + i0 \in \mathbb{C}$ ; las operaciones en  $\mathbb{C}^2$  y  $\mathbb{C}$  se definen como las del ejemplo de  $\mathbb{R}^2$  y  $\mathbb{R}$ . Análogamente, el conjunto  $\mathbb{C}^n$  sobre el campo  $\mathbb{C}$  es un espacio vectorial: sus vectores tienen  $n$  entradas complejas y sus escalares también son complejos. A este tipo de espacio vectorial le llamamos *espacio vectorial complejo*, de acuerdo a la siguiente definición:

Def. Un *espacio vectorial complejo* es aquel definido sobre el campo  $\mathbb{C}$  (campo complejo) o, equivalentemente, aquel donde los escalares son números complejos.

Nota: no podemos visualizar los vectores de  $\mathbb{R}^n$  con  $n > 3$ , los de  $\mathbb{C}^m$  con  $m > 1$ , ni los del conjunto de funciones de clase  $C^1$ , etc., como *flechas que parten de un mismo origen*<sup>9</sup>. Sin embargo, sí podemos hacer operaciones entre estos vectores de manera análoga a como lo haríamos con vectores de  $\mathbb{R}^2$  o  $\mathbb{R}^3$ , por lo cual trabajar en estos espacios *visualizables* puede ayudarnos a generar intuición sobre espacios vectoriales más abstractos.

Para ver más ejemplos de espacios vectoriales pueden revisar, por ejemplo, *Linear Algebra* de Fried-

<sup>7</sup>Nótese que inclusive en el caso  $n = 1$  el conjunto  $\mathbb{R}$  sobre sí mismo forma un espacio vectorial. Es decir, en este caso,  $\mathbb{R}$  funciona como conjunto vectorial y como campo.

<sup>8</sup>En general, el conjunto de funciones de clase  $C^n$  sobre el campo  $\mathbb{R}$  forma un espacio vectorial.

<sup>9</sup>Es posible visualizar vectores más abstractos de otras formas: <https://www.youtube.com/watch?v=zwAD6dRSVYI>.

berg (págs. 8-11.), *Introduction to Linear Algebra* de Lang (págs. 89-90). o *Linear Algebra: A Modern Introduction* de Poole (págs. 430-432) entre otros.

### 3.2.3. Algunos teoremas de espacios vectoriales

**Teorema 1.2.3.1.** Sean  $\mathbf{x}, \mathbf{y}, \mathbf{z}$  vectores de  $V$  tales que  $\mathbf{x} + \mathbf{z} = \mathbf{y} + \mathbf{z}$ , entonces  $\mathbf{x} = \mathbf{y}$ .

*Demostración.* Ya que  $\mathbf{z}$  es un vector de  $V \implies \exists -\mathbf{z} \in V$  tal que  $\mathbf{z} + (-\mathbf{z}) = \mathbf{0}$ , por la propiedad de existencia de inversos aditivos de los espacios vectoriales. Sumando este elemento  $-\mathbf{z}$  a cada lado de la igualdad inicial, tenemos que

$$\begin{aligned}\mathbf{x} + \mathbf{z} = \mathbf{y} + \mathbf{z} &\iff \mathbf{x} + \mathbf{z} + (-\mathbf{z}) = \mathbf{y} + \mathbf{z} + (-\mathbf{z}) \iff \mathbf{x} + (\mathbf{z} + (-\mathbf{z})) = \mathbf{y} + (\mathbf{z} + (-\mathbf{z})) \iff \\ &\mathbf{x} + \mathbf{0} = \mathbf{y} + \mathbf{0} \iff \mathbf{x} = \mathbf{y},\end{aligned}$$

donde en la tercera igualdad se aplicó la propiedad asociativa de la suma vectorial, en la cuarta igualdad se aplicó la propiedad de existencia de los inversos aditivos y, en la última, se aplicó la propiedad de existencia de neutro aditivo.  $\square$

A este teorema se le conoce como *Ley de cancelación para la suma vectorial* y con él se puede demostrar la unicidad del neutro aditivo y de los inversos aditivos.

**Teorema 1.2.3.2.** Sea  $V$  sobre  $K$  un espacio vectorial arbitrario con  $\mathbf{v} \in V$  y  $a \in K$ , entonces se verifica que:

1.  $0\mathbf{v} = \mathbf{0}$
2.  $a\mathbf{0} = \mathbf{0}$
3.  $(-a)\mathbf{v} = -(a\mathbf{v}) = a(-\mathbf{v})$

*Demostración.* 1.  $0\mathbf{v} + 0\mathbf{v} = (0 + 0)\mathbf{v} = 0\mathbf{v} = 0\mathbf{v} + \mathbf{0} \iff 0\mathbf{v} = \mathbf{0}$ , donde se aplicaron las propiedades de distributividad del producto de un vector por un escalar con respecto a la suma *escalar*, existencia del neutro aditivo y la ley de cancelación de la suma vectorial (Teorema 1.2.3.1).

2.  $a\mathbf{0} + a\mathbf{0} = a(\mathbf{0} + \mathbf{0}) = a\mathbf{0} = a\mathbf{0} + \mathbf{0} \iff a\mathbf{0} = \mathbf{0}$ , donde se aplicaron las propiedades de distributividad del producto de un vector por un escalar con respecto a la adición *vectorial*, existencia del neutro aditivo y la ley de cancelación de la suma vectorial.

3. Por el primer inciso y por distributividad,  $\mathbf{0} = (0)\mathbf{v} = (a + (-a))\mathbf{v} = a\mathbf{v} + (-a)\mathbf{v}$ <sup>10</sup>, por lo cual  $a\mathbf{v} + (-a)\mathbf{v} = \mathbf{0}$ . Por otro lado, por la propiedad de cerradura del producto de un vector por un escalar  $a\mathbf{v} \in V$  y, por la existencia de inversos aditivos,  $\exists -(a\mathbf{v}) \in V$  tal que  $a\mathbf{v} + [-(a\mathbf{v})] = \mathbf{0}$ . Por lo tanto, tenemos que  $\mathbf{0} = \mathbf{0} \iff a\mathbf{v} + (-a)\mathbf{v} = a\mathbf{v} + [-(a\mathbf{v})] \iff (-a)\mathbf{v} = -(a\mathbf{v})$  por la ley de cancelación. En particular,  $(-1)\mathbf{v} = -\mathbf{v}$ ; por la propiedad de compatibilidad del producto de un vector por un escalar con el producto entre escalares, se sigue que  $a(-\mathbf{v}) = a[(-1)\mathbf{v}] = [a(-1)]\mathbf{v} = (-a)\mathbf{v} = -(a\mathbf{v})$ .  $\square$

---

<sup>10</sup>La existencia de  $-a \in K$  está asegurada ya que  $K$  es un campo (ver sección 3.1.1).

### 3.3. Interpretación geométrica de las operaciones esenciales de los espacios vectoriales

Como se mencionó en una nota al final de la sección 3.2.2 —de la cual retomaremos muchas ideas a continuación—, podemos desarrollar nuestra intuición sobre muchos temas del álgebra lineal trabajando en espacios vectoriales *visualizables*, para luego extenderla a espacios vectoriales más generales. Por ende, ahora haremos hincapié en la interpretación geométrica de las operaciones de suma vectorial y producto de un vector por un escalar en los espacios vectoriales reales  $\mathbb{R}^2$  y  $\mathbb{R}^3$ , así como en el espacio vectorial complejo  $\mathbb{C}$ . Antes de empezar, necesitamos recordar la siguiente definición.

**Def.** Un *par ordenado* (o *2-tupla*) es un par de números  $(a, b)$  en donde el orden de los números importa; matemáticamente, decimos que el par ordenado  $(a, b) \neq (b, a) \iff b \neq a^a$ .

<sup>a</sup>Observemos que esto implicaría que  $(a, b) = (b, a) \iff b = a$ , lo cual tiene sentido ya que, si ambos números son el mismo, es imposible distinguir el orden. Esta es una definición equivalente de par ordenado.

#### 3.3.1. En 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 los pares ordenados  $(a, b)$  de entradas reales<sup>11</sup> y los 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. De manera análoga, en álgebra lineal, cada vector  $\begin{bmatrix} a & b \end{bmatrix}^T \in \mathbb{R}^2$  tiene una correspondencia biunívoca con una flecha en el plano cartesiano que tiene cola en el origen y punta en la coordenada cartesiana  $(a, b)$  correspondiente. Ambas correspondencias se muestran en la Figura 6.

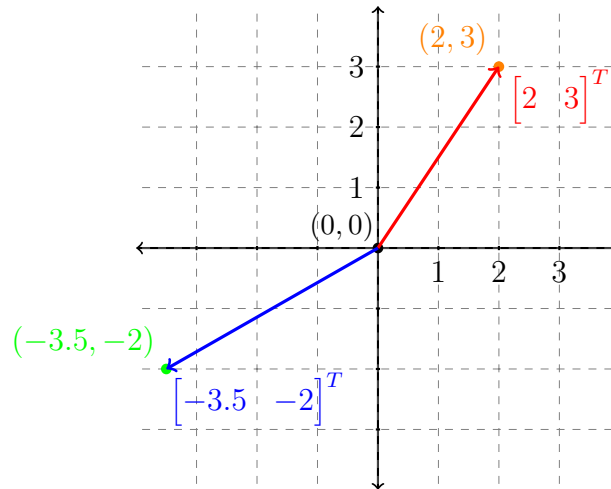


Figura 1: Ejemplo de representación de pares ordenados y vectores en el plano cartesiano. Los pares ordenados  $(-3.5, -2)$  y  $(2, 3)$  se representan mediante puntos que corresponden precisamente a las coordenadas cartesianas  $(-3.5, -2)$  y  $(2, 3)$ , respectivamente, mientras que los vectores  $\begin{bmatrix} -3.5 & -2 \end{bmatrix}^T$  y  $\begin{bmatrix} 2 & 3 \end{bmatrix}^T$  son representados por flechas que tienen su cola en el origen del plano cartesiano y su punta en las coordenadas  $(-3.5, -2)$  y  $(2, 3)$ , respectivamente.

<sup>11</sup>Es decir, con  $a, b \in \mathbb{R}$ .

## Suma vectorial

En este espacio, la suma vectorial se define como  $\begin{bmatrix} a & b \end{bmatrix}^T + \begin{bmatrix} c & d \end{bmatrix}^T \equiv \begin{bmatrix} a+c & b+d \end{bmatrix}^T$ . Podemos calcular, por ejemplo, la suma  $\begin{bmatrix} 2 & 1 \end{bmatrix}^T + \begin{bmatrix} 1 & 3 \end{bmatrix}^T = \begin{bmatrix} 3 & 4 \end{bmatrix}^T$ . Los tres vectores mencionados se muestran en la Figura 2.

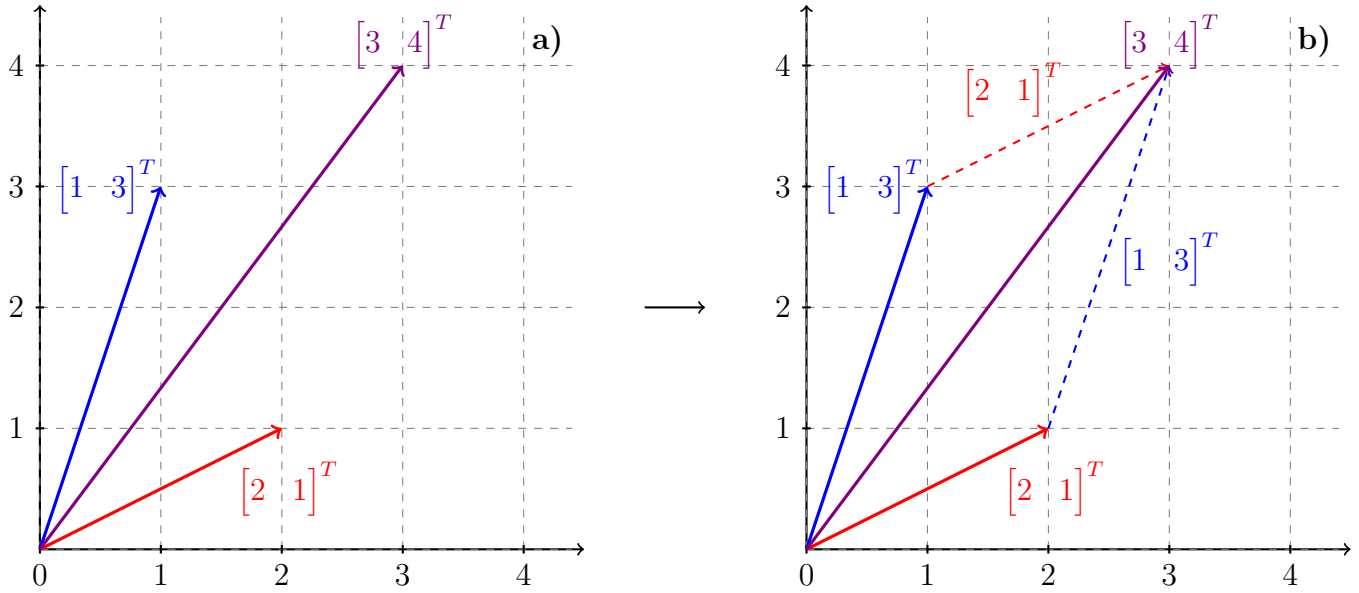


Figura 2: Interpretación geométrica de la suma vectorial en el espacio vectorial real  $\mathbb{R}^2$ . En la figura **a)** se observan los vectores  $\begin{bmatrix} 1 & 3 \end{bmatrix}^T$  y  $\begin{bmatrix} 2 & 1 \end{bmatrix}^T$ , así como el vector resultante de la suma de los dos anteriores, 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

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*. En particular,  $\forall \mathbf{v} \in \mathbb{R}^2$ ,  $\mathbf{0} + \mathbf{v} = \mathbf{v}$ , lo cual concuerda con el hecho de que el vector  $\mathbf{v}$  corresponda a una flecha con cola en el origen.

## Producto de un vector por un escalar

En este espacio, el producto de un vector por un escalar se define como  $c \begin{bmatrix} a & b \end{bmatrix}^T \equiv \begin{bmatrix} ca & cb \end{bmatrix}^T$ . Podemos calcular, por ejemplo, los productos  $(\frac{1}{2}) \begin{bmatrix} 2 & 2 \end{bmatrix}^T = \begin{bmatrix} 1 & 1 \end{bmatrix}^T$  y  $(-1.2) \begin{bmatrix} 1 & 3 \end{bmatrix}^T = \begin{bmatrix} -1.2 & -3.6 \end{bmatrix}^T$ . La representación gráfica de estas operaciones se muestra en la Figura 3.

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

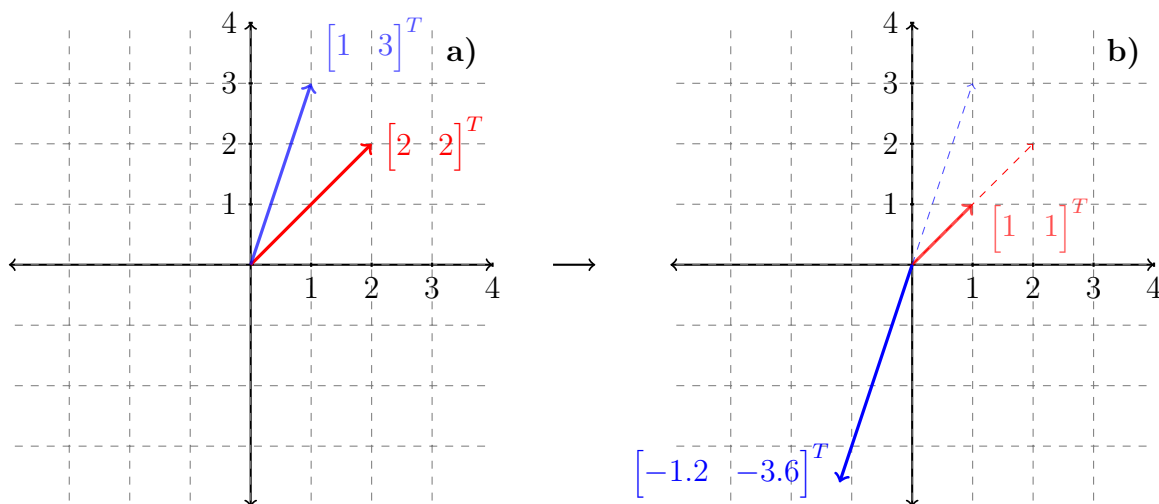


Figura 3: Interpretación geométrica del producto de un vector por un escalar en el espacio vectorial real  $\mathbb{R}^2$ . Comparando las figuras **a)** y **b)** observamos que, en caso de que se multiplique a un vector de  $\mathbb{R}^2$  por un escalar de  $\mathbb{R}$ , es posible que la longitud del vector cambie y que su sentido se invierta, pero su dirección no cambia.

vector automáticamente se convierte en el vector nulo  $\begin{bmatrix} 0 & 0 \end{bmatrix}^T$ , como se demostró algebraicamente en el primer inciso del Teorema 1.2.3.2. 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](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  $\mathbf{0} = \begin{bmatrix} 0 & 0 \end{bmatrix}^T$  (también llamado *vector origen*) no tiene longitud, ya que es el único donde la cola y la punta de su flecha coinciden. Además, se dice que tampoco tiene dirección ni sentido<sup>12</sup>. 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 segundo inciso del Teorema 1.2.3.2.

### 3.3.2. En el espacio vectorial real $\mathbb{R}^3$

La suma vectorial y el producto de un vector por un escalar (o *reescalamiento*) en el espacio vectorial real  $\mathbb{R}^3$  tienen la misma interpretación geométrica que en  $\mathbb{R}^2$ , con una dimensión extra añadida. Esto es de esperarse, ya que las definiciones de estas operaciones y las correspondencias entre vectores y flechas que salen del origen a una coordenada específica son análogas en ambos espacios vectoriales.

### 3.3.3. 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

<sup>12</sup>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.

eje perpendicular a él de números *imaginarios*<sup>13</sup>— nos sirve para representar vectores con una entrada compleja. Así, cada vector de una entrada compleja  $[a + ib]$  con  $a, b \in \mathbb{R}$  tiene una correspondencia uno a uno con una flecha con cola en el origen del plano y flecha en la coordenada  $(a, ib)$  del plano complejo, la cual corresponde a, desde el origen, moverse  $a$  unidades sobre el eje real y  $b$  unidades sobre el eje imaginario.

### Suma vectorial

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

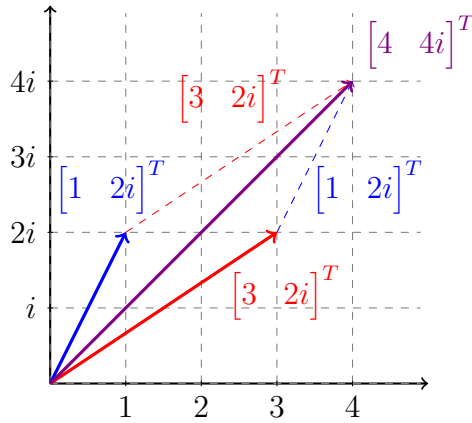


Figura 4: Interpretación geométrica de la suma vectorial en el espacio vectorial complejo  $\mathbb{C}$ . Observamos que, al igual que en el caso del espacio vectorial real  $\mathbb{R}^2$ , se cumple la *Ley del paralelogramo*.

### Producto de un vector por un escalar

Por definición, el producto de un vector por un escalar es  $(q + ir) [s + it]^T \equiv [(qs - rt) + i(qt + rs)]^T$ .

Notemos que, en particular, si la parte imaginaria del escalar es nula (i.e., si  $r = 0$ ), entonces el escalar es un número real y el producto resultante es  $(q) [s + it]^T \equiv [(qs) + (qt)i]^T$ , por lo cual geométricamente sólo se produce un reescalamiento totalmente análogo al discutido en la sección 3.3.1.

En cambio, ahora observemos qué sucede si la parte real del escalar es nula y la parte imaginaria es igual a 1 (i.e., si multiplicamos por el escalar  $i$ ). Tomemos, por ejemplo, al vector  $[2 + 2i]^T$ . Al hacer el producto de este vector por  $i$  obtenemos  $[-2 + 2i]^T$ . Si, en cambio, hacemos el producto de este mismo vector por el escalar  $-i$ , obtenemos como resultado  $(-i) [2 + 2i]^T = [2 - 2i]^T$ . Ambas operaciones se muestran de manera visual en la Figura 5.

Aquí vemos que hacer el producto de un vector por el escalar  $i$  equivale a hacer una rotación de  $90^\circ$  ó  $\frac{\pi}{2}$  radianes. Análogamente, el producto de un vector por el escalar  $-i$  equivale a hacer una rotación de  $-90^\circ$  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.

<sup>13</sup>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.



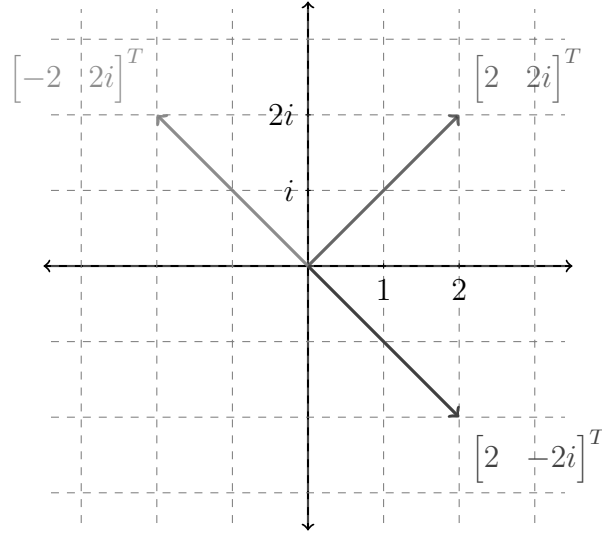


Figura 5: Interpretación geométrica del producto de un vector complejo por los números imaginarios  $i$  y  $-i$ . En este caso, nuestro vector base es  $\begin{bmatrix} 2 & 2i \end{bmatrix}^T$ . El producto de este vector por el escalar  $i$  resulta en el vector  $\begin{bmatrix} -2 & 2i \end{bmatrix}^T$ , 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  $\begin{bmatrix} 2 & 2i \end{bmatrix}^T$  por  $-i$  se puede interpretar geométricamente como una rotación discreta de  $-\frac{\pi}{2}$  radianes.

¿Y si multiplicamos un vector de  $\mathbb{C}$  por un escalar  $ai$  con  $a \neq 0, 1$ ? Ya que  $ai \begin{bmatrix} b + ic \end{bmatrix}^T = \begin{bmatrix} -ac + i(ab) \end{bmatrix}^T = a \begin{bmatrix} -c + ib \end{bmatrix}^T = a(i \begin{bmatrix} b + ic \end{bmatrix}^T)$  —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 donde  $a$  es negativo, 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  $a$ : ambos razonamientos son equivalentes.

Dicho lo anterior, estamos listos para el caso más general: multiplicar un vector complejo  $\begin{bmatrix} s + it \end{bmatrix}^T$  por un escalar complejo  $q + ir$  con  $q, r \neq 0$  *reescalará* el vector en el plano complejo y lo *rotará* en el sentido correspondiente al signo de  $r$ . Es decir que, en los espacios vectoriales complejos, los escalares no sólo pueden *reescalar* vectores, sino que también los pueden *rotar*<sup>14</sup>.

### 3.4. Subespacios vectoriales

En ciertas formas *específicas*, las cuales iremos detallando, los subespacios vectoriales son a los espacios vectoriales lo que los subconjuntos a los conjuntos. Por ejemplo: así como cualquier subconjunto es, en sí mismo, un conjunto, cualquier subespacio vectorial es, en sí mismo, un espacio vectorial.

<sup>14</sup>El asunto de las *magnitudes específicas* de estos reescalamientos y rotaciones —el cual se complica para escalares complejos en general— será precisado más adelante.



### 3.4.1. Definición de subespacio vectorial

Def. Sea un conjunto  $V$  sobre un campo  $K$  un espacio vectorial. Un *subespacio vectorial* de  $V$  es un subconjunto  $W \subset V$  sobre el campo  $K$  con las operaciones de suma vectorial y producto de un vector por un escalar que cumple las propiedades siguientes:

$$\forall \mathbf{w}, \mathbf{x} \in W \quad \exists \mathbf{w} + \mathbf{x} \in W$$

Cerradura de la adición

$$\forall \mathbf{w} \in W, a \in K \quad \exists a\mathbf{w} \in W$$

Cerradura del producto de un vector por un escalar

$$\exists \mathbf{0} \in W \text{ t.q. } \mathbf{w} + \mathbf{0} = \mathbf{w} \quad \forall \mathbf{w} \in W$$

Elemento identidad de la adición (neutro aditivo).

En este caso, a  $W$  se le conoce como un *subespacio vectorial* de  $V$  (nuevamente, simplificando la notación); sin embargo,  $W$  también es, en sí mismo, un espacio vectorial.

Observemos que:

- Ya que cualquier subespacio vectorial es un espacio vectorial, entonces cualquier subespacio vectorial puede tener subespacios vectoriales subsecuentes<sup>15</sup>.
- La definición de subespacio vectorial sólo incluye tres propiedades. Esto nos indica que, ya que  $W$  es subconjunto de  $V$  y ambos espacios vectoriales están definidos sobre el mismo campo  $K$ , si  $W$  cumple explícitamente las tres propiedades mencionadas, las demás propiedades de un espacio vectorial se siguen trivialmente.
- Para todo espacio vectorial  $V$ ,  $V$  y  $\{\mathbf{0}\}$  son subespacios vectoriales de  $V$ <sup>16</sup>.

### 3.4.2. Ejemplos de subespacios vectoriales

El conjunto de todos los pares ordenados  $\left\{ \begin{bmatrix} x_1 & x_2 \end{bmatrix}^T \mid x_1, x_2 \in \mathbb{R} \wedge x_1 = x_2 \right\}$  es un subespacio vectorial en el espacio vectorial real  $\mathbb{R}^2$  (o  $\mathbb{R} \times \mathbb{R}$ ).

El conjunto  $\mathbb{R}$  sobre el campo  $\mathbb{R}$  es un subespacio vectorial del conjunto  $\mathbb{C}$  sobre el mismo campo  $\mathbb{R}$ .

Sean  $j, k \in \mathbb{N}$  t.q.  $j < k$ . El conjunto de polinomios de grado  $j$  es un subespacio vectorial<sup>17</sup> del espacio vectorial formado por el conjunto de polinomios de grado  $k$ .

El conjunto de todas las funciones reales de clase  $C^\infty$  es un subespacio vectorial del espacio vectorial formado por el conjunto de todas las funciones reales de clase  $C^n$  (con  $n \in \mathbb{N}$ ).

### 3.4.3. Algunos teoremas de subespacios vectoriales

**Teorema 1.4.3.1.** Cualquier intersección de dos subespacios vectoriales de  $V$  es un subespacio vectorial de  $V$ .

<sup>15</sup>Esto es más o menos similar al hecho de que cualquier subconjunto puede tener subconjuntos subsecuentes.

<sup>16</sup>Aquí se sobreentiende que los conjuntos  $V$  y  $\{\mathbf{0}\}$  están definidos como espacios vectoriales sobre el mismo campo  $K$  y que  $\{\mathbf{0}\}$  representa al espacio vectorial que sólo tiene al vector nulo de  $V$  como vector.

<sup>17</sup>De aquí en adelante, asumiremos que cualquier espacio vectorial  $V$  está definido por un conjunto vectorial  $V$  sobre el campo  $\mathbb{R}$  (espacio vectorial real), a menos que se indique lo contrario.

*Demostración.* Sea  $V$  sobre  $K$  un espacio vectorial y sea  $C$  una colección de subespacios vectoriales de  $V$  (definidos sobre el mismo campo  $K$ ). Sea  $W$  la intersección de los conjuntos vectoriales de  $C$ . Entonces, ya que cada subespacio vectorial en  $C$  contiene al neutro aditivo de  $V$ ,  $\mathbf{0} \in W$ . Además, sea  $a \in K$  y sean  $\mathbf{u}, \mathbf{v} \in W$ , entonces  $\mathbf{u}, \mathbf{v}$  están en todos los subespacios de la colección  $C$ , cada uno de los cuales es cerrado por la adición vectorial y por el producto de un vector por un escalar, de donde se sigue que  $a\mathbf{u}, a\mathbf{v}$  y  $\mathbf{u} + \mathbf{v}$  están en todos los subespacios, por lo cual también están en  $W$ . Por lo tanto, por la definición de la sección 3.4.1,  $W$  es un subespacio vectorial de  $V$ .  $\square$

**Teorema 1.4.3.2.** Sea  $Z$  un subespacio vectorial de  $W$  y sea  $W$ , a su vez, subespacio vectorial de  $V$ . Entonces  $Z$  es un subespacio vectorial de  $V$ .

La demostración del Teorema 1.4.3.2 se deja como ejercicio. Este último teorema nos muestra otra analogía válida entre subconjuntos y subespacios vectoriales, ya que si  $A \subset B$  y  $B \subset C \implies A \subset C$ .

### 3.4.4. Suma y suma directa de espacios vectoriales

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 el espacio vectorial definido por  $S_1 + S_2 = \{\mathbf{x} + \mathbf{y} \mid \mathbf{x} \in S_1, \mathbf{y} \in S_2\}$ <sup>a</sup>.

Def. Si, además, se cumple que  $S_1 + S_2 = V$  y  $S_1 \cap S_2 = \{\mathbf{0}\}$ , decimos que el espacio vectorial  $V$  es la *suma directa* de  $S_1$  y  $S_2$ , lo cual denotamos como  $S_1 \oplus S_2 = V$ .

<sup>a</sup>Aquí se sobreentiende que  $S_1, S_2$  y  $S_1 + S_2$  están definidos sobre el mismo campo.

La operación de suma entre subespacios vectoriales en realidad es una suma entre sus *conjuntos vectoriales* —así como la intersección de dos espacios vectoriales es en realidad una intersección de los conjuntos vectoriales—; el conjunto resultante de la suma forma un espacio vectorial sobre el mismo campo que define a los subespacios. Observemos que la definición de suma vectorial pide que  $S_1$  y  $S_2$  sean *subespacios* de un espacio vectorial  $V$ , y no sólo espacios vectoriales arbitrarios: esto asegura que su suma  $S_1 + S_2$  también sea un espacio vectorial.

Por otro lado, una suma directa de la forma  $V_1 \oplus V_2 \oplus \dots \oplus V_n = W$  nos da la sensación de que, en cierto sentido, el espacio vectorial  $W$  se puede *descomponer* en sus subespacios  $V_1, V_2, \dots, V_n$ , dado que el único elemento común entre cualesquiera de estos dos subespacios es el neutro aditivo (vector nulo).

Para dar un ejemplo: sea  $C$  el espacio vectorial de todas las funciones constantes  $f(x) = c$  con  $c \in \mathbb{R}$ , y sea  $D$  el de todas las funciones de la forma  $f(x) = dx$  para algún  $d \in \mathbb{R}$ . Sea  $P^n$  el espacio vectorial de todos los polinomios de grado  $n$ , es decir, de todas las funciones con regla de correspondencia  $f(x) = c_0x^1 + c_1x^1 + \dots + c_nx^n$  con  $c_i \in \mathbb{R}$ , entonces  $C \oplus D = P^1$  (de hecho, nótese que  $P^0 = C$  por lo cual pudimos haber escrito  $P^0 \oplus D = P^1$  de manera equivalente).

Volveremos a esta idea de *descomponer un espacio vectorial como una suma directa de sus subespacios vectoriales* más adelante en el curso. Antes de eso, debemos ver otro tipo de operación de los espacios vectoriales, la cual se realiza entre dos vectores y da como resultado un escalar.

## 3.5. Ejercicios de repaso

### 3.5.1. Campos

1. Demuestra que el conjunto  $\mathbb{C}$  junto con las operaciones de suma y multiplicación definidas en la sección 3.1.3 cumplen las últimas tres propiedades de la definición de campo de la sección 3.1.1 para finalmente demostrar que estas tres cosas juntas (el conjunto  $\mathbb{C}$  y las dos operaciones mencionadas) forman un campo. (0.5 ptos. extra)\*

### 3.5.2. Espacios vectoriales

1. Sea  $K$  un campo arbitrario. ¿Siempre puede definirse un espacio vectorial de  $K$  (como conjunto vectorial) sobre sí mismo (como campo)? Si sí, demuéstalo. Si no, da un contraejemplo (Nota: repasa los axiomas de campo y ten cuidado con tu notación). (1 pto.)
2. Explica por qué **no** puede existir un espacio vectorial con vectores reales en  $\mathbb{R}^n$  sobre el campo  $\mathbb{C}$ , pero sí puede haber un espacio vectorial con vectores complejos en  $\mathbb{C}^n$  sobre el campo  $\mathbb{R}$ . (1 pto.)
3. Sea  $K$  un campo arbitrario. Demuestra que  $K \times K \times \dots \times K = K^n, n \in \mathbb{N}$  sobre  $K$  es un espacio vectorial. (0.5 ptos. extra)\*

### 3.5.3. Interpretación geométrica de los espacios vectoriales

1. Define las variables  $a, b, c, d, e \in \mathbb{R}$  de acuerdo a los últimos 5 dígitos de tu número de cuenta. Calcula algebraicamente al vector  $\begin{bmatrix} a & b \end{bmatrix}^T + c \begin{bmatrix} d & e \end{bmatrix}^T$  y muestra gráficamente al vector resultante aplicando la Ley del paralelogramo. (1 pto.)
2. Como vimos en la sección 3.3, la suma de dos vectores  $\mathbf{v}, \mathbf{u} \in \mathbb{R}^2$  sobre  $\mathbb{R}$  se puede representar geoméricamente con ayuda de un paralelogramo en el plano cartesiano, en donde la diagonal del paralelogramo que sale del origen es igual al vector resultante de la suma  $\mathbf{u} + \mathbf{v}$  ó  $\mathbf{v} + \mathbf{u}$  (ver, e.g., la Figura 2). Muestra intuitivamente que si transportamos a la *otra* diagonal del paralelogramo hasta el origen (lo cual se puede hacer de dos maneras distintas), obtenemos la representación del vector resultante de la resta  $\mathbf{u} - \mathbf{v}$  ó  $\mathbf{v} - \mathbf{u}$ . (1 pto.)
3. Muestra intuitivamente que el punto de intersección de las diagonales del paralelogramo mencionado en el ejercicio anterior es igual a la punta del vector  $\frac{1}{2}(\vec{u} + \vec{v})$ . (1 pto.)
4. Elige los dos últimos dígitos de tu número de cuenta y llámalos  $q$  y  $r$ , respectivamente. Al segundo y tercer dígito de tu número de cuenta, llámalos  $s$  y  $t$ . Dibuja el vector complejo  $\begin{bmatrix} q + ir \end{bmatrix}^T \in \mathbb{C}$  en el plano complejo. Explica qué le pasaría a este vector si realizáramos su producto por los escalares complejos  $s + it$ ,  $-s + it$ ,  $s - it$  y  $-s - it$ . (1 pto.)

### 3.5.4. Subespacios vectoriales

1. En la página 17 del libro *Linear Algebra* de Friedberg, la definición de un subespacio vectorial incluye una cuarta propiedad que deben cumplir (la de existencia de inversos aditivos), la cual no incluí en las notas del curso, ya que es redundante. Explica por qué se puede desechar esta cuarta propiedad sin afectar la definición de subespacio vectorial (Nota: ver, por ejemplo, la definición de subespacios vectoriales dada en la página 91 de *Introduction to Linear Algebra* de Lang, que no incluye dicha propiedad). (1 pto.)
2. Demuestra que el conjunto  $\{(a, 2a, 3a) \in \mathbb{R}^3 \mid a \in \mathbb{R}\}$  sobre el campo  $\mathbb{R}$  es un subespacio vectorial de  $\mathbb{R}^3$ . ¿Qué representa este conjunto visualmente en  $\mathbb{R}^3$ ? (1 pto.)
3. Demuestra que el espacio vectorial de todas las funciones polinomiales de una variable real de grado  $n$  es un subespacio vectorial del espacio vectorial de todas las funciones reales de una variable de clase  $C^m$  para toda  $n \geq m$ . (1 pto.)
4. Demuestra que el espacio vectorial de todas las funciones de una variable real que son derivables y tienen derivada continua (de clase  $C^1$ ) sobre  $\mathbb{R}$  es un subespacio vectorial del espacio vectorial de todas las funciones reales continuas (de clase  $C^0$ ) de una variable sobre  $\mathbb{R}$ . ¿Cómo generalizarías

este resultado a que las de clase  $C^i$  formen un subespacio de las de clase  $C^n$  con  $i > n$  sobre el mismo campo  $\mathbb{R}$ ? Argumenta. (0.5 ptos. extra)\*

5. Demuestra el Teorema 1.4.3.2 (ver sección 3.4.3). (1 pto.)

## 4. Producto escalar (punto), norma, interpretación geométrica del producto escalar: proyecciones y ortogonalidad, producto vectorial (cruz)\* y triple producto escalar\*

### 4.1. Producto escalar (punto)

#### 4.1.1. Definición de producto escalar (punto)

Def. Sea  $V$  sobre  $K$  un espacio vectorial, con  $\mathbf{u}, \mathbf{v}, \mathbf{w} \in V$  y  $a \in K$ . El *producto escalar*  $(\cdot \cdot \cdot) : V \times V \rightarrow K$  le asocia a cualquier par ordenado de vectores en  $V$  un escalar en  $K$ , y cumple las tres propiedades siguientes:

$$(\mathbf{u} + \mathbf{w}, \mathbf{v}) = (\mathbf{u}, \mathbf{v}) + (\mathbf{w}, \mathbf{v}),$$

$$(a\mathbf{u}, \mathbf{v}) = a(\mathbf{u}, \mathbf{v}),$$

$$(\mathbf{u}, \mathbf{v}) = \overline{(\mathbf{v}, \mathbf{u})},$$

donde la barra  $\overline{(\mathbf{v}, \mathbf{u})}$  denota el complejo conjugado de  $(\mathbf{v}, \mathbf{u})$ <sup>a</sup>. Si, además, se cumple la propiedad

$$\forall \mathbf{v} \in V, (\mathbf{v}, \mathbf{v}) \geq 0; (\mathbf{v}, \mathbf{v}) = 0 \iff \mathbf{v} = \mathbf{0},$$

se dice que el producto escalar es *positivo definido*.

Para vectores que son  $n$ -tuplas (como aquellos de  $\mathbb{R}^n$  o  $\mathbb{C}^n$ ), es común que el producto escalar de  $\mathbf{u}$  y  $\mathbf{v}$  se denote por  $\mathbf{u} \cdot \mathbf{v}$ , por lo cual también se le conoce como *producto punto*; sin embargo, también existen notaciones alternativas como  $\langle \mathbf{u}, \mathbf{v} \rangle$ <sup>b</sup>.

<sup>a</sup>Recordemos que, en general, nuestro campo  $K$  puede ser complejo; en ese caso, el producto escalar en  $V$  le asignará a cada par ordenado de dos vectores de  $V$  un escalar complejo  $c \in \mathbb{C}$ .

<sup>b</sup>Esta notación para el producto escalar es similar a la utilizada en Mecánica Cuántica  $\langle \phi | \psi \rangle$  —en donde el campo siempre es complejo—, y esconde un profundo significado, el cual veremos más adelante.

A partir de la definición, observemos lo siguiente:

- Las primeras dos propiedades juntas nos dicen que  $(\mathbf{u} + a\mathbf{w}, \mathbf{v}) = (\mathbf{u}, \mathbf{v}) + a(\mathbf{w}, \mathbf{v})$ . Para referirnos a esta propiedad específica en lenguaje matemático, decimos que el producto escalar es una operación *lineal*<sup>18</sup> en la primera entrada.
- La tercera propiedad nos dice que, si el campo es real, el producto escalar es una operación conmutativa, es decir, que si  $K = \mathbb{R} \implies (\mathbf{u}, \mathbf{v}) = (\mathbf{v}, \mathbf{u})$ . En cambio, si el campo es complejo, en general, el producto escalar es no conmutativo.
- Las primeras tres propiedades juntas nos dicen que, si el espacio vectorial está definido sobre un campo complejo, entonces  $(\mathbf{u}, \mathbf{v} + a\mathbf{w}) = (\mathbf{u}, \mathbf{v}) + \bar{a}(\mathbf{u}, \mathbf{w})$ , donde  $\bar{a}$  (ó  $a^*$ ) denota el complejo conjugado del escalar complejo  $a$ . Esto nos dice que, si el campo es complejo, el producto escalar es una operación *antilineal*<sup>19</sup> en la segunda entrada. En cambio, si el campo es real, entonces el producto escalar es lineal en ambas entradas.

<sup>18</sup>Como debes sospechar, el concepto de operación *lineal* es fundamental para el álgebra *lineal*, y lo veremos con más detalle en las secciones de transformaciones y operadores lineales.

<sup>19</sup>Este concepto también se verá con mayor detenimiento más adelante.

- Para un producto escalar positivo definido, el único vector que puede tener como resultado el escalar 0 al hacer producto escalar consigo mismo es el vector nulo  $\mathbf{0}$  (el neutro aditivo que vimos en las secciones 3.2.1 y 3.4.1).

#### 4.1.2. Ejemplos de producto escalar en espacios vectoriales

En  $\mathbb{R}^2$  el producto escalar se define como  $\begin{bmatrix} u_1 \\ u_2 \end{bmatrix} \cdot \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} \equiv u_1v_1 + u_2v_2$ , donde hemos utilizado la notación de punto, ya que los vectores son 2-tuplas. En general, en el espacio vectorial real  $\mathbb{R}^n$  el producto escalar se define como

$$\mathbf{u} \cdot \mathbf{v} = \begin{bmatrix} u_1 & u_2 & \dots & u_n \end{bmatrix}^T \cdot \begin{bmatrix} v_1 & v_2 & \dots & v_n \end{bmatrix}^T \equiv u_1v_1 + u_2v_2 + \dots + u_nv_n = \sum_{i=1}^n u_iv_i^{20}.$$

Para cumplir *todas* las propiedades descritas en la sección 4.1.1, el producto escalar en el espacio vectorial complejo  $\mathbb{C}^n$  se debe definir de una forma ligeramente distinta. Sean  $\mathbf{a}, \mathbf{b} \in \mathbb{C}^n$  vectores con  $n$  entradas complejas, entonces el producto escalar se define como

$$(\mathbf{a}, \mathbf{b}) \equiv \mathbf{a} \cdot \mathbf{b} \equiv \sum_{i=1}^n a_ib_i^*,$$

es decir, se realiza un producto entre la  $i$ -ésima entrada de  $\mathbf{a}$  y el *complejo conjugado* de la  $i$ -ésima entrada de  $\mathbf{b}$  para cada  $i$ , y luego se suman dichos productos.

Sea  $C^0([a, b])$  el conjunto de todas las funciones de variable real continuas en el intervalo cerrado  $[a, b]$  con integral finita, entonces podemos definir un producto escalar en el espacio vectorial del conjunto  $C^0([a, b])$  sobre el campo  $\mathbb{R}$  como

$$(f, g) = \int_a^b f(x)g(x)dx.$$

Observemos que, como muestra el último ejemplo, se puede definir un producto escalar en muchos tipos de espacios vectoriales diferentes, y no sólo en aquellos que tienen como vectores a  $n$ -tuplas<sup>21</sup>. Nótese además que, en cada caso, el resultado del producto escalar es un escalar del campo sobre el cual está definido el espacio vectorial.

Para ver más ejemplos de productos escalares pueden revisar *Linear Algebra* de Friedberg (págs. 330-331), *Introduction to Linear Algebra* de Lang (págs. 172-173), *Linear Algebra: A Modern Introduction* de Poole<sup>22</sup> (págs. 531-534), etc.

#### 4.1.3. Propiedades del producto escalar

Las principales propiedades del producto escalar son:

$(\mathbf{u}, \mathbf{v}) = \overline{(\mathbf{v}, \mathbf{u})}$	Conmutar vectores resulta en la conjugación del escalar
$(a\mathbf{u} + \mathbf{v}, \mathbf{w}) = a(\mathbf{u}, \mathbf{w}) + (\mathbf{v}, \mathbf{w})$	Linealidad en la primera entrada
$(\mathbf{u}, \mathbf{w} + b\mathbf{z}) = (\mathbf{u}, \mathbf{w}) + b^*(\mathbf{u}, \mathbf{z})$	Antilinealidad en la segunda entrada

<sup>20</sup>Para simplificar la notación, en varias áreas de la física se elimina el signo de suma ( $\Sigma$ ) cuando aparecen índices repetidos, siguiendo la convención de que éstos implícitamente indican una suma sobre el índice. Así,  $\sum_{i=1}^n u_iv_i$  se puede escribir simplemente como  $u_iv_i$ . A esto se le conoce como *notación de Einstein* o *convención de suma de Einstein*.

<sup>21</sup>Veremos otros tipos de espacios vectoriales donde se pueden definir productos escalares más adelante.

<sup>22</sup>Ten en cuenta que algunos libros introducen el producto escalar (al cual pueden llamar producto punto) únicamente con espacios vectoriales reales, y lo generalizan a espacios vectoriales complejos en secciones posteriores.

En particular, el producto escalar definido sobre un espacio vectorial real es una operación lineal en ambas entradas (o *bilíneal*), es decir, que

$$(a_1 \mathbf{u}_1 + \dots + a_n \mathbf{u}_n, \mathbf{v}) = a_1(\mathbf{u}_1, \mathbf{v}) + \dots + a_n(\mathbf{u}_n, \mathbf{v}); \quad (\mathbf{u}, b_1 \mathbf{v}_1 + \dots + b_n \mathbf{v}_n) = b_1(\mathbf{u}, \mathbf{v}_1) + \dots + b_n(\mathbf{u}, \mathbf{v}_n).$$

## 4.2. Norma

### 4.2.1. Definición de norma

Def. Una *norma* es una operación  $\|\cdot\| : V \rightarrow K$  que toma sólo un vector y devuelve un escalar, y que cumple las siguientes propiedades:

$$\|\mathbf{u} + \mathbf{v}\| \leq \|\mathbf{u}\| + \|\mathbf{v}\| \quad \text{Satisface la desigualdad del triángulo}$$

$$\|a\mathbf{u}\| = |a| \|\mathbf{u}\| \quad \text{Es escalable de forma absoluta}$$

$$\|\mathbf{u}\| = 0 \iff \mathbf{u} = \mathbf{0} \quad \text{Distingue al vector nulo.}$$

A partir de la definición anterior se puede demostrar que, al igual que el producto escalar, la norma también es positivo definida.

**Corolario** Sea  $V$  un espacio vectorial con una norma  $\|\cdot\|$ , entonces  $\forall \mathbf{v} \in V$  con  $\mathbf{v} \neq \mathbf{0}$  se cumple que  $\|\mathbf{v}\| > 0$ .

*Demostración.* Por definición de norma tenemos que  $\|\mathbf{0}\| = 0$ . Sea  $\mathbf{v} \in V$  tal que  $\mathbf{v} \neq \mathbf{0}$ ; ya que la norma distingue al vector nulo, entonces  $\|\mathbf{v}\| \neq 0$ . Por definición de espacio vectorial,  $\exists -\mathbf{v} \in V$  tal que  $\mathbf{v} + (-\mathbf{v}) = \mathbf{0} \implies \|\mathbf{v} + (-\mathbf{v})\| = \|\mathbf{0}\| = 0$ . Ya que la norma cumple la desigualdad del triángulo por definición, tenemos que

$$0 = \|\mathbf{0}\| = \|\mathbf{v} + (-\mathbf{v})\| \leq \|\mathbf{v}\| + \|-\mathbf{v}\| \implies 0 \leq \|\mathbf{v}\| + \|-\mathbf{v}\|.$$

Además, por definición, la norma es escalable de forma absoluta, por lo cual

$$0 \leq \|\mathbf{v}\| + \|-\mathbf{v}\| = \|\mathbf{v}\| + |-1| \|\mathbf{v}\| = \|\mathbf{v}\| + \|\mathbf{v}\| = 2\|\mathbf{v}\|.$$

Ya que  $\|\mathbf{v}\| \neq 0$ , el resultado anterior  $2\|\mathbf{v}\| \geq 0 \implies \|\mathbf{v}\| > 0$ , como se quería demostrar. □

### 4.2.2. Ejemplos de norma

Tanto en el espacio vectorial real  $\mathbb{R}$  como en el espacio vectorial complejo  $\mathbb{C}$  se puede definir una norma como

$$\|x\| = |x|$$

para cualquier vector  $x$  de dichos espacios, ya que el valor absoluto cumple trivialmente las propiedades de la sección 4.2.1 (lo cual quizá demostraste en tu curso de Cálculo I, por lo menos en el caso real). Como recordatorio, las definiciones del valor absoluto son

$$|r| = +\sqrt{r^2} \quad \forall r \in \mathbb{R}; \quad |c| = +\sqrt{cc^*} \quad \forall c \in \mathbb{C}.$$

A ésta se le conoce como la *norma del valor absoluto*<sup>23</sup>. En el caso real, esta norma se interpreta geométricamente como la distancia entre el origen de la recta real y el punto correspondiente al valor  $r$

<sup>23</sup>En el caso de los números reales, el valor absoluto equivale a cambiar los signos de los números negativos por signos positivos y no hacerle nada a los números no negativos.

o, equivalentemente, como la longitud de la flecha que tiene cola en 0 y punta en  $r$ ; en el caso complejo, se interpreta como la distancia mínima (o *euclídeana*) entre el origen del plano complejo y el punto correspondiente al valor complejo  $c$  o, equivalentemente, como la longitud de la flecha que tiene cola en el origen del plano complejo y punta en  $c$ .

Para los vectores que son  $n$ -tuplas, la norma más comúnmente utilizada se define como

$$\|\mathbf{u}\| = +\sqrt{(\mathbf{u}, \mathbf{u})} = +\sqrt{\mathbf{u} \cdot \mathbf{u}},$$

de acuerdo a las definiciones de productos escalares para  $n$ -tuplas dadas en la sección 4.1.2. Esta norma se interpreta geoméricamente como la longitud de la flecha que tiene cola en el origen del espacio vectorial y punta en la coordenada dada por las entradas del vector  $\mathbf{u}$ , lo cual es equivalente a la distancia euclídeana entre estos dos puntos. Por ende, a esta norma se le conoce como *norma euclídeana*<sup>24</sup>. En particular, a los espacios vectoriales reales  $\mathbb{R}^n$  con esta norma se les conoce como *espacios vectoriales euclídeanos* (o *euclídeos*)<sup>25</sup>. Cuando la norma está asociada geoméricamente a la longitud de la flecha que representa un vector, también se dice que está relacionada con la *magnitud* de ese vector.

También, en un espacio vectorial real de funciones reales con integral finita en el intervalo  $[a, b]$ , podemos definir una norma a partir del producto escalar como

$$\|f\| = +\sqrt{(f, f)} = +\sqrt{\int_a^b f(x)f(x)dx},$$

siguiendo el último ejemplo de la sección 4.1.2. Es fácil demostrar que esta definición cumple con todas las propiedades de norma (ver sección 4.2.1: se siguen de las propiedades de la integral que se ven en Cálculo II).

### 4.2.3. Desigualdad de Cauchy-Schwarz

El proceso anterior de obtener una norma a partir de un producto escalar se puede generalizar para cualquier producto escalar positivo definido (ver la sección 4.6.2), tanto para espacios vectoriales reales como complejos; entonces, se dice que este tipo de normas son *inducidas* por un producto escalar positivo definido. En general, las normas de este tipo serán las más recurrentes durante este curso, aunque también veremos otros tipos de normas no inducidas por productos escalares más adelante. El primer paso necesario para demostrar este resultado general es la llamada *desigualdad de Cauchy-Schwarz*, que enunciamos a continuación como teorema.

**Teorema 2.2.3.1.** Sea  $(\cdot, \cdot) : V \times V \rightarrow K$  un producto escalar positivo definido y sea  $\|\cdot\| : V \rightarrow K$  una función definida tal que  $\|\mathbf{u}\| = +\sqrt{(\mathbf{u}, \mathbf{u})}$ <sup>26</sup>, entonces se cumple que  $|(\mathbf{u}, \mathbf{v})| \leq \|\mathbf{u}\| \|\mathbf{v}\|$ .

*Demostración.* Si  $\mathbf{v} = \mathbf{0}$ , la demostración se cumple trivialmente. Supongamos que  $\mathbf{v} \neq \mathbf{0}$ .

Sabemos que para cualquier  $a \in K$  se cumple que

$$\begin{aligned} 0 \leq \|\mathbf{u} - a\mathbf{v}\|^2 &= (\mathbf{u} - a\mathbf{v}, \mathbf{u} - a\mathbf{v}) = (\mathbf{u}, \mathbf{u} - a\mathbf{v}) - a(\mathbf{v}, \mathbf{u} - a\mathbf{v}) = \\ &= (\mathbf{u}, \mathbf{u}) - a^*(\mathbf{u}, \mathbf{v}) - a((\mathbf{v}, \mathbf{u}) - a^*(\mathbf{v}, \mathbf{v})) = (\mathbf{u}, \mathbf{u}) - a^*(\mathbf{u}, \mathbf{v}) - a(\mathbf{v}, \mathbf{u}) + aa^*(\mathbf{v}, \mathbf{v}). \end{aligned}$$

<sup>24</sup>Observemos que la norma del valor absoluto es simplemente un caso particular de la norma euclídeana para los espacios vectoriales  $\mathbb{R}$  y  $\mathbb{C}$ , tanto en su definición algebraica como en su interpretación geométrica.

<sup>25</sup>Estos son los espacios básicos que se utilizan en geometría analítica y cálculo diferencial e integral con funciones de una o más variables reales.

<sup>26</sup>Con la ayuda de este teorema, se puede demostrar que esta función  $\|\cdot\|$  cumple con todas las propiedades de una norma, dadas en la sección 4.2.1. Ésta es la llamada *norma euclídeana* generalizada.



En particular, tomando  $a = \frac{(\mathbf{u}, \mathbf{v})}{(\mathbf{v}, \mathbf{v})}$  y sustituyendo en la desigualdad anterior, obtenemos que

$$\begin{aligned}
 0 &\leq (\mathbf{u}, \mathbf{u}) - \frac{(\mathbf{v}, \mathbf{u})}{(\mathbf{v}, \mathbf{v})}(\mathbf{u}, \mathbf{v}) - \frac{(\mathbf{u}, \mathbf{v})}{(\mathbf{v}, \mathbf{v})}(\mathbf{v}, \mathbf{u}) + \frac{(\mathbf{u}, \mathbf{v})}{(\mathbf{v}, \mathbf{v})} \frac{(\mathbf{v}, \mathbf{u})}{(\mathbf{v}, \mathbf{v})}(\mathbf{v}, \mathbf{v}) = \\
 &(\mathbf{u}, \mathbf{u}) - 2 \frac{(\mathbf{u}, \mathbf{v})(\mathbf{v}, \mathbf{u})}{(\mathbf{v}, \mathbf{v})} + \frac{(\mathbf{u}, \mathbf{v})(\mathbf{v}, \mathbf{u})}{(\mathbf{v}, \mathbf{v})} = \|\mathbf{u}\|^2 - \frac{|(\mathbf{u}, \mathbf{v})|^2}{\|\mathbf{v}\|^2} \implies \\
 0 &\leq \|\mathbf{u}\|^2 - \frac{|(\mathbf{u}, \mathbf{v})|^2}{\|\mathbf{v}\|^2} \implies \frac{|(\mathbf{u}, \mathbf{v})|^2}{\|\mathbf{v}\|^2} \leq \|\mathbf{u}\|^2 \implies |(\mathbf{u}, \mathbf{v})|^2 \leq \|\mathbf{u}\|^2 \|\mathbf{v}\|^2 \implies |(\mathbf{u}, \mathbf{v})| \leq \|\mathbf{u}\| \|\mathbf{v}\|.
 \end{aligned}$$

□

### 4.3. Interpretación geométrica del producto escalar: proyecciones y ortogonalidad

#### 4.3.1. Producto escalar y proyecciones

Como quizás aprendiste en tus cursos de Geometría Analítica y/o Mecánica Vectorial, el producto escalar entre vectores de  $\mathbb{R}^2$  tiene una definición geométrica dada por

$$(\mathbf{u}, \mathbf{v}) \equiv \|\mathbf{u}\| \|\mathbf{v}\| \cos \theta,$$

donde  $\theta$  es el ángulo positivo más chico entre  $\mathbf{u}$  y  $\mathbf{v}$ ; esta definición es equivalente a la algebraica dada al principio de la sección 4.1.2. El producto  $\|\mathbf{v}\| \cos \theta$  es igual a la proyección del vector  $\mathbf{v}$  sobre el vector  $\mathbf{u}$ , lo cual puede escribirse como  $P_{\mathbf{u}}(\mathbf{v})$ , por lo cual podemos abreviar la expresión anterior como

$$(\mathbf{u}, \mathbf{v}) = \|\mathbf{u}\| P_{\mathbf{u}}(\mathbf{v}) \implies P_{\mathbf{u}}(\mathbf{v}) = \frac{(\mathbf{u}, \mathbf{v})}{\|\mathbf{u}\|}.$$

Observemos que, ya que tanto  $\cos \theta$  como las normas  $\|\mathbf{u}\|$  y  $\|\mathbf{v}\|$  son números reales, entonces

$$(\mathbf{u}, \mathbf{v}) = \|\mathbf{u}\| P_{\mathbf{u}}(\mathbf{v}) = \|\mathbf{u}\| \|\mathbf{v}\| \cos \theta = \|\mathbf{v}\| \|\mathbf{u}\| \cos \theta = \|\mathbf{v}\| P_{\mathbf{v}}(\mathbf{u}) = (\mathbf{v}, \mathbf{u}).$$

Esto no es más que la propiedad conmutativa del producto escalar en un espacio vectorial real —visto desde un punto de vista geométrico.

Ahora imaginemos que tenemos dos vectores  $\mathbf{u}, \mathbf{v} \in \mathbb{R}^2$  de la misma magnitud (i.e.,  $\|\mathbf{u}\| = \|\mathbf{v}\|$ ) trazados en el plano cartesiano. En este caso, la bisectriz del ángulo que los separa traza un eje de simetría. Si trazamos la proyección de cada uno de los vectores sobre el otro, veremos de forma más intuitiva que

$$\|\mathbf{u}\| P_{\mathbf{u}}(\mathbf{v}) = \|\mathbf{v}\| P_{\mathbf{v}}(\mathbf{u}),$$

con lo cual reafirmamos que el producto escalar conmuta. Supongamos ahora que reescalamos alguno de los vectores, digamos  $\mathbf{u}$ , al doble de su longitud —efectivamente duplicando así su norma—, y llamamos a este nuevo vector  $\mathbf{u}'$ . Entonces, algebraicamente vemos que

$$(\mathbf{u}', \mathbf{v}) = \|\mathbf{u}'\| P_{\mathbf{u}}(\mathbf{v}) = 2\|\mathbf{u}\| P_{\mathbf{u}}(\mathbf{v}) = 2(\mathbf{u}, \mathbf{v}).$$

Geométricamente, podemos ver esto como que  $(\mathbf{u}', \mathbf{v}) = (2\|\mathbf{u}\|) P_{\mathbf{u}}(\mathbf{v})$  o, equivalentemente, que  $(\mathbf{u}', \mathbf{v}) = \|\mathbf{u}\| (2P_{\mathbf{u}}(\mathbf{v}))$ <sup>27</sup>. Observemos que la elección del vector reescalado fue arbitraria —de haber reescalado

<sup>27</sup>Este mismo argumento se expone de forma animada en el siguiente video: [https://www.youtube.com/watch?v=LyGKycYT2v0&list=PL\\_w8oSr1JpVCZ5pKXHKz6PkjGCBpBSBYv&index=11&t=255s](https://www.youtube.com/watch?v=LyGKycYT2v0&list=PL_w8oSr1JpVCZ5pKXHKz6PkjGCBpBSBYv&index=11&t=255s).

$\mathbf{v}$  por 2 en vez de  $\mathbf{u}$ , hubiéramos obtenido el mismo resultado. La discusión anterior es simplemente la interpretación geométrica del hecho de que  $(2\mathbf{u}, \mathbf{v}) = 2(\mathbf{u}, \mathbf{v}) = (\mathbf{u}, 2\mathbf{v})$  en espacios vectoriales reales.

Similarmente, la interpretación geométrica del hecho de que  $(\mathbf{u} + \mathbf{v}, \mathbf{w}) = (\mathbf{u}, \mathbf{w}) + (\mathbf{v}, \mathbf{w})$  ó  $(\mathbf{w}, \mathbf{u} + \mathbf{v}) = (\mathbf{w}, \mathbf{u}) + (\mathbf{w}, \mathbf{v})$  se puede ver directamente de la Ley del paralelogramo.

Sin embargo, al generalizar esta operación de *proyectar un vector sobre otro (no nulo)* a cualquier tipo de espacios vectoriales nos encontramos con un problema: en general, el producto escalar no es conmutativo, por lo cual debemos decidir de qué manera precisa definiremos a esta operación. Por razones que veremos más adelante, la definición que resulta más conveniente es la siguiente:

Def. Sean  $\mathbf{u}, \mathbf{v}$  vectores de un espacio vectorial  $V$  arbitrario tal que  $\mathbf{u}$  es distinto al vector nulo. Definimos a la operación  $P_{\mathbf{u}}(\cdot) : V \rightarrow K$ , conocida como la *proyección del vector  $\mathbf{v}$  sobre el vector  $\mathbf{u}$* , como

$$P_{\mathbf{u}}(\mathbf{v}) \equiv \frac{(\mathbf{v}, \mathbf{u})}{\|\mathbf{u}\|}.$$

### 4.3.2. Ortogonalidad

Por otro lado, ya que  $\cos(\frac{\pi}{2}) = 0$ , podemos ver directamente que, si dos vectores en  $\mathbb{R}^2$  son perpendiculares entre sí en el plano, su producto escalar será cero. Esto significaría que la proyección de cualquiera de estos vectores sobre el otro sería nula —lo cual, en el plano cartesiano, es sinónimo de que los vectores sean perpendiculares—; más generalmente, a esta propiedad se le conoce como *ortogonalidad*, y siempre se asocia con un producto escalar nulo por razones que veremos más adelante.

Def. Se dice que dos vectores  $\mathbf{u}, \mathbf{v} \in V$  son *ortogonales* si su producto escalar es nulo (i.e., si la proyección de cualquiera de los vectores sobre el otro es cero). Matemáticamente, esto se escribe como  $\mathbf{u} \perp \mathbf{v} \iff (\mathbf{u}, \mathbf{v}) = 0$  ó, equivalentemente,  $\mathbf{u} \perp \mathbf{v} \iff P_{\mathbf{u}}(\mathbf{v}) = 0 = P_{\mathbf{v}}(\mathbf{u})^a$ .

Def. Se dice que un subconjunto  $S \subset V$  es *ortogonal* si cualquier vector de  $S$  es ortogonal al resto de los vectores del conjunto, es decir, si  $\forall \mathbf{u}_i, \mathbf{u}_j \in S, \mathbf{u}_i \neq \mathbf{u}_j \implies (\mathbf{u}_i, \mathbf{u}_j) = 0$ .

<sup>a</sup>Observemos que esta definición también aplica para espacios vectoriales complejos, ya que  $0^* = 0$ .

Por ejemplo, en  $\mathbb{R}^3$ ,  $\begin{bmatrix} 2 & 3 & 1 \end{bmatrix}^T \perp \begin{bmatrix} -4 & 2 & 2 \end{bmatrix}^T$ , ó  $\begin{bmatrix} 2 & 3 & 1 \end{bmatrix}^T \perp \begin{bmatrix} -8 & 4 & 4 \end{bmatrix}^T$ ; además, el subconjunto  $\{\begin{bmatrix} 5 & 0 & 0 \end{bmatrix}^T, \begin{bmatrix} 0 & 7 & 0 \end{bmatrix}^T, \begin{bmatrix} 0 & 0 & -4 \end{bmatrix}^T\}$  es ortogonal. En cambio, en el espacio de funciones polinomiales reales  $P^2$  restringido al intervalo  $[-1, 1]$ , donde definimos el producto escalar como  $(f, g) \equiv \int_{-1}^1 f(x)g(x) dx$ , el conjunto  $\{1, x, \frac{1}{2}(3x^2 - 1)\}$  es ortogonal (¿podrás comprobarlo?).

### 4.3.3. Vectores unitarios

Nótese por la definición dada en 4.3.2 y las propiedades del producto escalar que, si dos vectores son ortogonales, entonces reescalar cualquiera de ellos no afecta a su ortogonalidad. Además, notemos que la definición geométrica del producto escalar para vectores en  $\mathbb{R}^2$  se simplifica cuando la norma de los vectores es uno. De hecho, como veremos más adelante, este tipo de vectores en general pueden simplificar muchas expresiones de operaciones realizadas en espacios vectoriales, por lo cual les damos un nombre especial.

Def. Decimos que un vector  $\mathbf{v} \in V$  es *unitario* cuando su norma es igual a uno, i.e., si y sólo si  $\|\mathbf{v}\| = 1$ . Los vectores unitarios suelen denotarse con una cuña encima de la letra, e.g.  $\hat{i}, \hat{j}, \hat{k}, \hat{v}, \hat{u}$ , etc.

Def. Decimos que un subconjunto  $S \subset V$  es *ortonormal* si es un conjunto ortogonal donde todos los vectores son unitarios. Es decir,  $S$  es *ortonormal* si y sólo si  $(\mathbf{v}_i, \mathbf{v}_j) = \delta_{ij} \quad \forall \mathbf{v}_i, \mathbf{v}_j \in S$ , donde  $\delta_{ij}$  es la delta de Kronecker de dos entradas.

Por ejemplo, el vector  $\begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}^T$  es unitario en  $\mathbb{R}^2$  (ó  $\mathbb{C}^2$ ), mientras que los vectores  $\begin{bmatrix} -1 & i \end{bmatrix}^T$  y  $\begin{bmatrix} -i & 1 \end{bmatrix}^T$  son unitarios en  $\mathbb{C}^2$ .

Por último, notemos que si  $\mathbf{v} \in V$  es un vector arbitrario no unitario distinto al vector nulo, entonces el vector

$$\mathbf{v}' = \frac{\mathbf{v}}{\|\mathbf{v}\|}$$

es unitario, i.e., tiene norma igual a uno. Por ello, a este proceso de multiplicar un vector no nulo y no unitario por el inverso multiplicativo de su norma se le conoce como *normalización*. Además, ya que por la definición de norma  $\|c\mathbf{v}\| = |c|\|\mathbf{v}\|$  entonces, si hacemos el producto de un vector unitario por cualquier escalar de valor absoluto igual a uno (i.e., con  $|c| = 1$ ), obtendremos nuevamente un vector unitario.

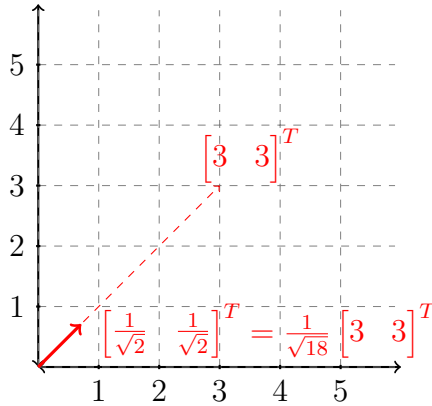


Figura 6: Ejemplo de *normalización* del vector  $\begin{bmatrix} 3 & 3 \end{bmatrix}^T \in \mathbb{R}^2$ . Ya que la norma de este vector es igual a  $\sqrt{3^2 + 3^2} = \sqrt{18}$ , hacemos su producto por el escalar  $\frac{1}{\sqrt{18}}$  —el inverso multiplicativo de  $\sqrt{18}$  en el campo real— y obtenemos al vector resultante  $\begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}^T$ , que tiene el mismo sentido y dirección que  $\begin{bmatrix} 3 & 3 \end{bmatrix}^T$ , pero con norma igual a uno.

#### 4.4. Producto vectorial (cruz)\*

El producto vectorial es una operación de gran utilidad, principalmente en algunas ramas de la física clásica (en las cuales se trabaja con vectores en  $\mathbb{R}^3$ ), y tiene una interpretación geométrica bastante rica, por lo cual se menciona brevemente en estas notas. En esta sección también se mencionan por primera vez los temas de matrices, determinantes y vectores canónicos, los cuales asumo que ya debes conocer, aunque más adelante en el curso los repasaremos.

#### 4.4.1. Definición del producto vectorial (cruz)

En  $\mathbb{R}^3$  se puede definir una operación entre dos vectores que da como resultado un tercer vector, el cual es ortogonal a los dos anteriores. A esta operación se le conoce como *producto vectorial*.

Def. En  $\mathbb{R}^3$ , el producto vectorial  $\times : V \times V \rightarrow V$  entre dos vectores  $\mathbf{r} = (r_1, r_2, r_3)$  y  $\mathbf{s} = (s_1, s_2, s_3)$  se define como

$$\mathbf{r} \times \mathbf{s} \equiv (r_2 s_3 - r_3 s_2, r_3 s_1 - r_1 s_3, r_1 s_2 - r_2 s_1).$$

Dado el símbolo ( $\times$ ) utilizado para denotar esta operación, también se le conoce como *producto cruz*.

Algebráicamente, el producto cruz  $\mathbf{r} \times \mathbf{s}$  también puede ser visto como el determinante de una matriz de  $3 \times 3$  como sigue:

$$\mathbf{r} \times \mathbf{s} = \det \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ r_1 & r_2 & r_3 \\ s_1 & s_2 & s_3 \end{vmatrix} = \mathbf{i}(r_2 s_3 - r_3 s_2) + \mathbf{j}(r_3 s_1 - r_1 s_3) + \mathbf{k}(r_1 s_2 - r_2 s_1),$$

donde  $\mathbf{i}, \mathbf{j}$  y  $\mathbf{k}$  son los vectores canónicos de  $\mathbb{R}^3$   $(1, 0, 0)$ ,  $(0, 1, 0)$  y  $(0, 0, 1)$ , respectivamente.

Geométricamente, si nombramos como  $\theta$  el menor ángulo de separación positivo entre dos vectores  $\mathbf{r}$  y  $\mathbf{s} \in \mathbb{R}^3$ , entonces podemos hacer la definición equivalente  $\mathbf{r} \times \mathbf{s} = \|\mathbf{r}\| \|\mathbf{s}\| \sin \theta$ . En este caso, la magnitud del vector resultante de hacer el producto vectorial  $\mathbf{r} \times \mathbf{s}$  se interpreta como la magnitud del área del paralelogramo formado por los vectores  $\mathbf{r}$  y  $\mathbf{s}$ . De aquí se sigue que el producto vectorial de dos vectores ortogonales sea igual al producto de sus normas, mientras que el producto cruz de dos vectores colineales sea 0 (en particular, el producto cruz de cualquier vector consigo mismo es igual a 0). Además, de ambas definiciones (algebraica y geométrica) se sigue que el producto vectorial de cualquier vector en  $\mathbb{R}^3$  con el vector nulo ( $\mathbf{0}$ ) sea  $\mathbf{0}$ .

#### 4.4.2. Propiedades del producto vectorial (cruz)

Las siguientes propiedades son fáciles de demostrar para cualesquiera  $\mathbf{r}, \mathbf{s}, \mathbf{t} \in \mathbb{R}^3$  a partir de la definición de la sección 4.4.1 (¡inténtalo!):

$$\mathbf{s} \times \mathbf{r} = -\mathbf{r} \times \mathbf{s}$$

Anticonmutatividad del producto vectorial

$$(\mathbf{r} \times \mathbf{s}) \times \mathbf{t} \neq \mathbf{r} \times (\mathbf{s} \times \mathbf{t})$$

No asociatividad del producto vectorial

$$\mathbf{r} \times (\mathbf{s} + \mathbf{t}) = \mathbf{r} \times \mathbf{s} + \mathbf{r} \times \mathbf{t}$$

Distributividad bajo la suma vectorial

$$(a\mathbf{r} \times \mathbf{s}) = a(\mathbf{r} \times \mathbf{s}) = (\mathbf{r} \times a\mathbf{s})$$

Compatibilidad con el producto de un vector por un escalar

$$\mathbf{r} \times (\mathbf{s} \times \mathbf{t}) + \mathbf{s} \times (\mathbf{t} \times \mathbf{r}) + \mathbf{t} \times (\mathbf{r} \times \mathbf{s}) = \mathbf{0}$$

Identidad de Jacobi para el producto vectorial.

#### 4.5. Triple producto escalar\*

Al unir el producto escalar con el producto vectorial, se obtiene una operación llamada *triple producto escalar* entre tres vectores, que da como resultado un escalar. La magnitud del triple producto escalar se interpreta geométricamente como la magnitud del volumen del paralelepípedo formado por los tres vectores.

#### 4.5.1. Definición de triple producto escalar

Def. El *triple producto escalar* entre tres vectores  $\mathbf{r}, \mathbf{s}, \mathbf{t} \in \mathbb{R}^3$  se define como

$$\mathbf{r} \cdot \mathbf{s} \times \mathbf{t} \equiv r_1(s_2t_3 - s_3t_2) + r_2(s_3t_1 - s_1t_3) + r_3(s_1t_2 - s_2t_1),$$

es decir, primero se realiza el producto vectorial  $\mathbf{s} \times \mathbf{t}$  y luego se realiza el producto escalar entre el vector resultante de esa operación y  $\mathbf{r}$ .

Recordemos que el producto escalar se realiza entre dos vectores y da como resultado un escalar, mientras que el producto vectorial se realiza entre dos vectores pero da como resultado un vector. Por lo tanto, no hay ambigüedad en la expresión  $\mathbf{r} \cdot \mathbf{s} \times \mathbf{t}$ , ya que la única forma lógica de juntar estas dos operaciones es realizando primero el producto vectorial y después el producto escalar; por ende, escribir esta operación como  $\mathbf{r} \cdot (\mathbf{s} \times \mathbf{t})$  sería redundante.

Al igual que el producto vectorial, el triple producto escalar también puede ser visto algebraicamente como el determinante de una matriz:

$$\mathbf{r} \cdot \mathbf{s} \times \mathbf{t} = \begin{vmatrix} r_1 & r_2 & r_3 \\ s_1 & s_2 & s_3 \\ t_1 & t_2 & t_3 \end{vmatrix}.$$

Además, recordando que el producto escalar de dos vectores ortogonales es igual a 0 y que el vector resultante de la operación  $\mathbf{r} \times \mathbf{s}$  es ortogonal tanto a  $\mathbf{r}$  como a  $\mathbf{s}$ , se sigue directamente que  $\mathbf{r} \cdot \mathbf{r} \times \mathbf{s} = 0 = \mathbf{r} \cdot \mathbf{s} \times \mathbf{r}$ .

#### 4.5.2. Propiedades del triple producto escalar

$$\mathbf{r} \cdot \mathbf{s} \times \mathbf{t} = \mathbf{s} \cdot \mathbf{t} \times \mathbf{r} = \mathbf{t} \cdot \mathbf{r} \times \mathbf{s}$$

Conmutatividad bajo permutaciones cíclicas

$$\mathbf{r} \cdot \mathbf{s} \times \mathbf{t} = \mathbf{r} \times \mathbf{s} \cdot \mathbf{t}$$

Invariancia bajo intercambio de operadores

$$\mathbf{r} \cdot \mathbf{s} \times \mathbf{t} = -\mathbf{r} \cdot \mathbf{t} \times \mathbf{s} = -\mathbf{s} \cdot \mathbf{r} \times \mathbf{t} = -\mathbf{t} \cdot \mathbf{s} \times \mathbf{r}$$

Anticonmutatividad bajo intercambio de dos vectores

Nota: también se puede definir un *triple producto vectorial* entre tres vectores que da como resultado un vector como  $\mathbf{r} \times \mathbf{s} \times \mathbf{t}$ . Éste es de gran utilidad para hacer demostraciones de algunas identidades de cálculo vectorial en  $\mathbb{R}^3$ .

### 4.6. Ejercicios de repaso

#### 4.6.1. Producto escalar

1. Sean las funciones  $f(x) = \frac{x}{40.5}$ ,  $g(x) = \frac{x^2}{243}$  y  $h(x) = \frac{1}{9}$  definidas con dominio  $[0, 9]$ . Calcula el producto escalar entre estas funciones en ese dominio (sin contar productos escalares de la forma  $(f, f)$ ) utilizando la definición vista en la sección 4.1.2. (1 pto.)
2. Explica por qué no se puede definir un producto escalar de la forma  $(f, g) = \int_a^b f(x)g(x)dx$  para funciones integrables  $f$  y  $g$  con integrales con valor infinito en  $[a, b]$ . (1 pto.)
3. Sea  $\mathbb{C}^2$  un espacio vectorial complejo con  $\mathbf{q}, \mathbf{r} \in \mathbb{C}^2$ . Explica cuál(es) de las propiedades de la sección 4.1.3 no se cumpliría(n) si definiéramos el producto escalar en este espacio vectorial ingenuamente como  $\mathbf{q} \cdot \mathbf{r} = \sum_{i=1}^n q_i r_i$  con  $n = 2$  (nótese que esta misma definición de producto escalar nos generaría problemas al intentar usarla para definir una norma en este espacio). (1 pto.)

#### 4.6.2. Norma

1. Calcula la norma de las funciones dadas en el primer ejercicio de la sección 4.6.1. (Nota: utiliza la norma inducida por ese mismo producto escalar.) (1 pto.)
2. ¿El conjunto  $\{(x_1, x_2, \dots, x_n) \mid x_i \in \mathbb{R} \wedge \|(x_1, x_2, \dots, x_n)\| \leq 1\}$  puede formar un espacio vectorial sobre el campo real? Argumenta. (1 pto.)
3. Argumenta e ilustra la interpretación geométrica de la desigualdad de Cauchy-Schwarz para dos vectores cualesquiera de  $\mathbb{R}^2$  (Nota: ver sección 4.2.3). (1 pto.)
4. Demuestra la desigualdad del triángulo  $\|\mathbf{a} + \mathbf{b}\| \leq \|\mathbf{a}\| + \|\mathbf{b}\|$  para cualesquiera dos vectores en un espacio  $V$  con producto escalar positivo definido donde la norma  $\|\mathbf{a}\| \equiv +\sqrt{(\mathbf{a}, \mathbf{a})} \quad \forall \mathbf{a} \in V$  (pista: usa la desigualdad de Cauchy-Schwarz)<sup>28</sup>. Argumenta e ilustra su interpretación geométrica para dos vectores cualesquiera de  $\mathbb{R}^2$ . (1 pto.)

#### 4.6.3. Interpretación geométrica del producto escalar: proyecciones y ortogonalidad

1. Sea  $\mathbf{v} \in \mathbb{R}^2$  no-nulo. Prueba que el conjunto  $\{\mathbf{u} \in \mathbb{R}^2 \mid (\mathbf{u}, \mathbf{v}) = 0\}$  es una recta que pasa por el origen y llévala a una expresión de la forma  $ax + by = c$ . (1 pto.)
2. Sean  $\mathbf{v} = \begin{bmatrix} v_1 & v_2 \end{bmatrix}^T$ ,  $\mathbf{u} = \begin{bmatrix} u_1 & u_2 \end{bmatrix}^T$  vectores del espacio real  $\mathbb{R}^2$ . Deduce que  $\mathbf{u} \cdot \mathbf{v} = u_1 v_1 + u_2 v_2 = \|\mathbf{u}\| \|\mathbf{v}\| \cos(\theta)$ , donde  $\theta$  es el ángulo mínimo entre ambos vectores en el plano cartesiano. (Nota: supón que ambos vectores son no nulos y encuentra al escalar  $k \in \mathbb{R}$  tal que  $(\mathbf{u}, \mathbf{u} - k\mathbf{v}) = 0$ . ¿Por qué es especial este escalar?) (1 pto.)
3. Define a los vectores  $\mathbf{u} = \begin{bmatrix} a_1 & a_2 \end{bmatrix}^T$ ,  $\mathbf{v} = \begin{bmatrix} a_3 & a_4 \end{bmatrix}^T$  y  $\mathbf{w} = \begin{bmatrix} a_5 & a_6 \end{bmatrix}^T \in \mathbb{R}^2$ , donde  $a_i$  corresponde al  $i$ -ésimo dígito de tu número de cuenta. Explica geoméricamente el hecho de que  $(\mathbf{u} + \mathbf{v}, \mathbf{w}) = (\mathbf{u}, \mathbf{w}) + (\mathbf{v}, \mathbf{w})$ , en términos de las proyecciones discutidas en la sección 4.3 (Nota: recuerda la Ley del paralelogramo). (1 pto.)

#### 4.6.4. Producto vectorial (cruz)\*

1. Demuestra las propiedades del producto vectorial de la sección 4.4.2. Además, da una interpretación geométrica para la primera, tercera y cuarta propiedad enlistadas. (Nota: si quieres hacer estas demostraciones utilizando una notación condensada, te recomiendo investigar acerca del símbolo de Levi-Civita el cual, junto con la delta de Kronecker, facilitan la escritura de muchas demostraciones de cálculo vectorial, entre otras áreas de las matemáticas.) (1.5 pts. extra)\*

#### 4.6.5. Triple producto escalar\*

1. Demuestra las propiedades del triple producto escalar de la sección 4.5.2. (0.5 pts. extra)\*

---

<sup>28</sup>Con esto habrás demostrado que a partir de cualquier producto escalar positivo definido  $(\mathbf{a}, \mathbf{b})$  se puede definir una norma como  $\|\mathbf{a}\| = +\sqrt{(\mathbf{a}, \mathbf{a})}$

### Nota aclaratoria: *Sobre nombres y traducciones...*

Como seguramente habrás notado al leer los libros recomendados en la sección 2.2, al producto escalar en inglés se le conoce como *inner product*, y a los espacios vectoriales dotados de un producto escalar se les conoce como *inner product spaces*. En español, al producto escalar también se le conoce como *producto interior*; sin embargo, existe otro tipo de producto diferente al producto escalar al cual en inglés, desafortunadamente, le llaman *interior product*.

Esto significa que, en inglés, la convención es que *scalar product* e *interior product* sean operaciones diferentes mientras que, en español, la convención es que *producto escalar* y *producto interior* se refieran a la misma operación. Algunos textos en español utilizan *producto interno* (en vez de producto interior) como sinónimo de *producto escalar* para homologar los nombres con los utilizados en inglés pero, por ahora, las convenciones preponderantes en español e inglés no permiten una traducción directa.

Por lo anterior, en estas notas decidí usar únicamente el nombre de *producto escalar* para la operación entre dos vectores que da como resultado un escalar (la cual acostumbramos llamar *producto punto* cuando esos vectores son  $n$ -tuplas), pero es importante que sepan que esta operación es equivalente al *inner product* de los textos en inglés.

Para empeorar la situación, algunos textos en inglés se refieren a la operación de producto de un vector por un escalar como *scalar multiplication* —por lo cual algunos textos en español pueden referirse a esta operación como *multiplicación escalar* o, inclusive, *producto escalar*— mientras que otros textos utilizan el término *scalar multiplication* para referirse a la multiplicación entre dos elementos del campo (escalares) que da como resultado otro elemento del campo (escalar). Por lo tanto, debemos recordar que el significado de estos términos depende del contexto en que se utilicen.

Finalmente, remarcamos que la convención de este texto:

- Los términos *producto de un vector por un escalar* y *reescalamiento* se emplean para referirnos a la operación realizada entre un vector del conjunto vectorial y un escalar del campo que da como resultado otro vector.
- El término *producto escalar* se reserva para la operación realizada entre dos vectores que da como resultado un escalar.
- El término *producto entre escalares* se utiliza para referirnos a la multiplicación entre dos escalares del campo que da como resultado otro escalar del campo.

## 5. Combinaciones lineales, subespacio generado y conjunto generador, dependencia e independencia lineal

### 5.1. Combinaciones lineales

Como vimos en la sección 3.2.1, las dos operaciones *necesarias* para definir un espacio vectorial son la suma vectorial (realizada entre elementos del conjunto vectorial) y el producto de un vector por un escalar (realizada entre un elemento del conjunto vectorial y un elemento del campo), ambas dando como resultado un vector en el mismo espacio. La operación más general que podemos realizar a partir de dichas operaciones se define a continuación.

Def. Sea  $V$  sobre  $K$  un espacio vectorial y  $L = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\} \subset V$  un conjunto con un número finito de vectores de ese espacio. Decimos que  $\mathbf{u}$  es una *combinación lineal* de los vectores de  $L$  si y sólo si

$$\mathbf{u} = c_1 \mathbf{v}_1 + c_2 \mathbf{v}_2 + \dots + c_n \mathbf{v}_n = \sum_{i=1}^n c_i \mathbf{v}_i, \quad c_i \in K, \mathbf{v}_i \in L, n \in \mathbb{N}.$$

Observemos que, dadas las propiedades de cerradura de ambas operaciones vistas en la sección 3.2.1, claramente  $\mathbf{u} \in V$ , es decir, *el resultado de la combinación lineal de vectores de un espacio vectorial  $V$  siempre es un vector del mismo espacio.*

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

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

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

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

Siguiendo la discusión de la sección 3.3, donde se plantearon por separado las interpretaciones geométricas de las operaciones de suma entre vectores y producto de un vector por un escalar, podemos pensar que, si combinamos ambas operaciones, la operación resultante se puede interpretar como una combinación de flechas rectas reescaladas (y/o rotadas, si el espacio vectorial es complejo y el escalar tiene una parte imaginaria no nula), cada una con su propia longitud, dirección y sentido, a las cuales aplicamos la Ley del paralelogramo para obtener una nueva flecha (o línea) como resultado. Precisamente por ello es que a esta operación general se le conoce como *combinación lineal*.

### 5.2. Subespacio generado y conjunto generador

Sea  $V$  sobre  $K$  un espacio vectorial y  $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n \in V$  vectores. Si tomamos un conjunto de estos vectores, digamos,  $G = \{\mathbf{v}_1, \mathbf{v}_2\}$ , podemos también definir el conjunto de todos los vectores que se pueden generar a través de combinaciones lineales de los vectores de  $G$  como  $\{a\mathbf{v}_1 + b\mathbf{v}_2 \mid a, b \in K\}$ . Este nuevo conjunto cumple con todas las propiedades de un espacio vectorial, y este hecho es generalizable a cualquier conjunto  $G$  con un número finito de elementos, lo cual motiva la definición siguiente.



Def. Sea  $V$  sobre  $K$  un espacio vectorial y  $G \subset V$ . Entonces, definimos

$$\langle G \rangle \equiv \{c_1 \mathbf{v}_1 + c_2 \mathbf{v}_2 + \dots + c_n \mathbf{v}_n \mid c_i \in K, \mathbf{v}_i \in G\} \subseteq V.$$

A  $G$  se le conoce como el *conjunto generador* y a  $\langle G \rangle$ , como el *subespacio generado* por  $G$  ya que, por definición, cumple con todas las propiedades de un espacio vectorial y además, por cerradura de  $V$ , cualquier vector generado por  $G$  es elemento de  $V$ <sup>a</sup>. Por completez, definimos  $\langle \emptyset \rangle = \{\mathbf{0}\}$ . Si  $G$  consiste de un sólo vector  $\mathbf{g}$ , podemos escribir el generado de  $G$  como  $\langle \mathbf{g} \rangle$ .

<sup>a</sup>Nótese por la definición que, dependiendo de cómo sea el conjunto  $G$ , es posible que  $\langle G \rangle = V$ .

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

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

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

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

**Teorema 3.2.1.** Sea  $V$  un espacio vectorial,  $S \subset V$  un conjunto de vectores de  $V$  y  $\mathbf{v} \in V$  un vector arbitrario. Si  $S' = S \cup \{\mathbf{v}\}$ , entonces  $\langle S \rangle = \langle S' \rangle \iff \mathbf{v} \in \langle S \rangle$ .

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

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

### 5.3. Dependencia e independencia lineal

Como se vio en la sección 5.1, un vector puede ser expresado como diferentes combinaciones lineales de otros vectores del mismo espacio. En particular, el vector nulo  $\mathbf{0}$  de cualquier espacio vectorial  $V$  puede ser obtenido a través de la *combinación lineal trivial* de cualesquiera  $n$  vectores del espacio: sólo basta con que todos los coeficientes sean cero, i.e.

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

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

### 5.3.1. Definición de dependencia e independencia lineal

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

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

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

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

<sup>a</sup>El vector nulo se excluye de estas definiciones ya que, de lo contrario, cumpliría ambas trivialmente para cualquier conjunto arbitrario de vectores dado.

Por ejemplo, en  $\mathbb{R}^2$  los vectores  $\mathbf{u}_1 = \begin{bmatrix} 1 & 5 \end{bmatrix}^T$  y  $\mathbf{u}_2 = \begin{bmatrix} -3 & -15 \end{bmatrix}^T$  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  $\mathbf{v}_1 = \begin{bmatrix} 1 & 2 \end{bmatrix}^T$  y  $\mathbf{v}_2 = \begin{bmatrix} 1 & 3 \end{bmatrix}^T$  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 de 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

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

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

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

### 5.3.2. Interpretación geométrica de la dependencia e independencia lineal

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

### 5.3.3. Algunos teoremas sobre dependencia e independencia lineal

**Teorema 3.3.3.1.** Sea  $V$  un espacio vectorial y  $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n$  vectores linealmente independientes de  $V$ . Sean  $c_1, c_2, \dots, c_n \in K$  y  $d_1, d_2, \dots, d_n \in K$  tales que

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

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

*Demostración.*

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

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

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

□

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

**Teorema 3.3.3.2.** Sea  $V$  un espacio vectorial y  $S_1, S_2$  subespacios tales que  $S_1 \subseteq S_2 \subseteq V$ . Si  $S_2$  es linealmente independiente, entonces  $S_1$  también es linealmente independiente.

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

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

Si  $k = n$  entonces  $S_1$  también es *l.i.* trivialmente. Supongamos que  $k < n$ . Entonces, por el Teorema 1.2.3.2 (ver sección 3.2.3),  $0\mathbf{v}_{k+1} + \dots + 0\mathbf{v}_n = \mathbf{0}$  por lo cual lo anterior implica que

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

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

□

Este teorema nos dice que si removemos un vector de un conjunto linealmente independiente, el conjunto resultante también es linealmente independiente.

**Teorema 3.3.3.3.** Sea  $V$  un espacio vectorial sobre un campo  $K$  y  $L \subset V$  un conjunto con  $n$  elementos linealmente independientes entre sí. Entonces, para cualquier  $\mathbf{v} \in V$ , el conjunto  $L' \equiv L \cup \{\mathbf{v}\}$  es *l.i.*  $\iff \mathbf{v} \notin \langle L \rangle$ .

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

Por otro lado, supongamos que  $L'$  es linealmente dependiente. Entonces, existe una combinación lineal no trivial de los vectores de  $L'$  que resulta en el vector nulo, i.e.,  $c_1\mathbf{u}_1 + \dots + c_n\mathbf{u}_n + b\mathbf{v} = \mathbf{0}$  con al menos un coeficiente distinto de cero. En este caso, el coeficiente  $b \neq 0$ : si  $b$  fuera igual a cero, la

ecuación restante sería  $c_1 \mathbf{u}_1 + \dots + c_n \mathbf{u}_n = \mathbf{0}$ ; ya que  $L$  es linealmente independiente, entonces todos los vectores  $c_i$  deben ser iguales a cero pero, ya que estamos suponiendo que también  $b = 0$ , entonces el conjunto  $L'$  también sería linealmente independiente, contradiciendo la hipótesis. Así, sabiendo que  $b \neq 0$  podemos despejar la ecuación  $c_1 \mathbf{u}_1 + \dots + c_n \mathbf{u}_n + b \mathbf{v} = \mathbf{0}$  y obtener que  $\mathbf{v} = \frac{-c_1}{b} \mathbf{u}_1 + \dots + \frac{-c_n}{b} \mathbf{u}_n$ , lo cual implica que  $\mathbf{v} \in \langle L \rangle$ .

Por contraposición, tenemos que  $L' \equiv L \cup \{\mathbf{v}\}$  es *l.i.*  $\iff \mathbf{v} \notin \langle L \rangle$ .

□

La demostración de este teorema básicamente 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.

## 5.4. Ejercicios de repaso

### 5.4.1. Combinaciones lineales

1. Demuestra que si tenemos tres vectores  $\mathbf{u}$ ,  $\mathbf{v}$  y  $\mathbf{w}$  no nulos y ortogonales entre sí, entonces no puede haber una combinación lineal de dos de ellos que dé como resultado el tercero (Nota: repasa la definición de ortogonalidad y las propiedades del producto escalar vistas en la sección 4.) (1 pto.)
2. Da un ejemplo de la demostración anterior en  $\mathbb{R}^3$ , e interpreta este resultado geoméricamente (Nota: recuerda la relación entre el producto escalar y las proyecciones.) (1 pto.)

### 5.4.2. Subespacio generado y conjunto generador

1. Sea  $a \in \mathbb{R}$  un vector arbitrario del espacio vectorial real  $\mathbb{R}$ . Demuestra que este espacio vectorial sólo tiene subespacios vectoriales triviales. ¿Quiénes son los valores  $a \in \mathbb{R}$  que generan a estos subespacios, en cada caso? (1 pto.)
2. Sea  $G = \left\{ \begin{bmatrix} 3 & 0 & 3 \end{bmatrix}^T, \begin{bmatrix} -\frac{1}{2} & 0 & \frac{1}{2} \end{bmatrix}^T \right\} \subset \mathbb{R}^3$ . Escribe a  $\langle G \rangle$  algebraicamente y descríbelo geoméricamente. ¿A qué espacio vectorial real que conoces se parece? (1 pto.)
3. Sea  $\mathbf{c}$  un vector arbitrario del espacio vectorial complejo  $\mathbb{C}$ . Da una interpretación geométrica para  $\langle \mathbf{c} \rangle$  (Nota: recuerda que, en este caso,  $K = \mathbb{C}$ ; además, te sugiero que primero escribas a  $\langle \mathbf{c} \rangle$  algebraicamente y, a partir de ahí, busques interpretarlo geoméricamente.). (1 pto.)
4. Sea  $F = \{x^0, x^1, x^2, \dots, x^n\}$  con  $n \in \mathbb{N}$  un conjunto de funciones reales de una variable real. ¿Quién es, entonces,  $\langle F \rangle$ ? (1 pto.)

### 5.4.3. Dependencia e independencia lineal

1. Sea  $n_i$  el  $i$ -ésimo dígito de tu número de cuenta. Determina si los vectores  $\begin{bmatrix} n_1 + i(n_2) \end{bmatrix}^T$  y  $\begin{bmatrix} n_3 + i(n_4) \end{bmatrix}^T \in \mathbb{C}$  son linealmente dependientes o independientes y muéstralo gráficamente en el plano complejo. (1 pto.)
2. Repite el mismo ejercicio para los vectores reales  $\begin{bmatrix} n_1 & n_4 & n_7 \end{bmatrix}^T, \begin{bmatrix} n_2 & n_5 & n_8 \end{bmatrix}^T$  y  $\begin{bmatrix} n_3 & n_6 & n_9 \end{bmatrix}^T \in \mathbb{R}^3$ . ¿Cuál es el conjunto linealmente dependiente más grande que puedes armar con estos vectores, sin considerar los subespacios generados por éstos? ¿Qué hay del conjunto linealmente independiente más grande? (1 pto.)

3. Sea  $V$  un espacio vectorial y  $S_1 \subseteq S_2 \subseteq V$ . Demuestra que si  $S_1$  es linealmente dependiente, entonces  $S_2$  es linealmente dependiente. (1 pto.)
4. Sean las funciones  $p_1(x) = 3x^2 - x$ ,  $p_2(x) = x^3 + 5$ ,  $p_3(x) = 4x + 1$ ,  $p_4(x) = 4x^3 + 4x + 16$  vectores de  $P^4$ , el espacio vectorial real de funciones polinomiales de grado  $n \leq 4$ . ¿Cuáles son los tres conjuntos de tres vectores linealmente independientes que podemos formar a partir de estos cuatro vectores? (1 pto.)
5. Da un ejemplo de un conjunto de tres vectores en  $\mathbb{R}^3$  linealmente independientes y ortogonales entre sí, donde todas las entradas de dichos vectores sean distintas de cero. (0.5 ptos. extra)\*
6. Describe con tus palabras y con los conceptos vistos en clase el procedimiento que seguiste para resolver el ejercicio 3.4.3.5 (1 pto. extra)\*

## 6. Bases, dimensión, ortogonalización y ortonormalización

### 6.1. Bases

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<sup>29</sup>. Cuando trabajamos en un espacio vectorial  $V$ , resulta conveniente tener un conjunto de vectores  $B \subset V$  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 3.3.3.1 (sec. 5.3.3), 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.

#### 6.1.1. Definición de base

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

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

Nótese por la definición que, ya que muchos conjuntos de vectores distintos pueden ser linealmente independientes y generar a un mismo espacio vectorial, un espacio vectorial puede tener muchas bases distintas. Esto implica que cualquier vector arbitrario de un espacio vectorial puede ser expresado a través de diferentes combinaciones lineales (correspondientes a distintas bases del espacio, y únicas para cada base). Dicho de otra forma, dado un espacio vectorial con más de una base, cualquier vector de ese espacio puede ser *representado en las distintas bases* de ese espacio<sup>30</sup>.

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

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

#### 6.1.3. Teorema de reemplazamiento

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

<sup>29</sup>De lo contrario, se violarían las propiedades de cerradura vistas en la sección 3.2.1.

<sup>30</sup>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.

**Teorema 4.1.1.** 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  $V$ .

*Demostración.* Esta demostración se hará por inducción.

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

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

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

Sea  $L = \{\mathbf{v}_1, \dots, \mathbf{v}_{m+1}\}$  un subconjunto linealmente independiente de  $V$ . Por el Teorema 3.3.3.2, el conjunto  $\{\mathbf{v}_1, \dots, \mathbf{v}_m\}$  es l.i., por lo cual podemos aplicar la hipótesis de inducción y concluir que  $m \leq n$  y que existe un subconjunto  $\{\mathbf{u}_1, \dots, \mathbf{u}_{n-m}\} \subset G$  tal que  $\{\mathbf{v}_1, \dots, \mathbf{v}_m\} \cup \{\mathbf{u}_1, \dots, \mathbf{u}_{n-m}\}$  genera a  $V$ . Por lo tanto, existen escalares  $a_1, \dots, a_m, b_1, \dots, b_{n-m}$  tales que

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

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

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

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

Recordando que por hipótesis de inducción  $\{\mathbf{v}_1, \dots, \mathbf{v}_m, \mathbf{u}_1, \dots, \mathbf{u}_{n-m}\}$  genera a  $V$ , entonces el hecho de que esté contenido en  $L \cup H$  implica necesariamente que  $\langle L \cup H \rangle = V$ . Finalmente, ya que  $H$  es un subconjunto de  $G$  con  $(n - m) - 1 = n - (m + 1)$  vectores, el teorema se cumple para  $m + 1$ , terminando así nuestra demostración. □

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  $V$ <sup>31</sup>?

<sup>31</sup>Recuerda para qué dijimos que nos serviría este teorema.

## 6.2. 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 4.2.1.** Sean  $B = \{\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_n\}$  y  $B' = \{\mathbf{b}'_1, \mathbf{b}'_2, \dots, \mathbf{b}'_{n'}\}$  bases de  $V$ , entonces  $n = n'$ .

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

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

...

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

donde  $c_{ij} \in K$ .

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

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

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

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

...

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

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

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

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

<sup>a</sup>En el resto de estas notas, supondremos que los espacios vectoriales mencionados tienen dimensión finita, a menos que se indique lo contrario.

Observemos que esta definición *algebraica* de dimensión difiere de las definiciones geométricas y físicas usuales de dimensión. Por ejemplo, a pesar de que el espacio vectorial complejo  $\mathbb{C}$  se represente en el plano cartesiano —el cual tiene dimensión geométrica 2—, este espacio vectorial es de dimensión (algebraica) 1, como vimos en los ejemplos de la sección 6.1.2. 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 (por ejemplo, el espacio vectorial  $\mathbb{C}$  sobre  $\mathbb{R}$  no puede tener dimensión 1: ¿podrías demostrarlo?).



**Teorema 4.2.2.** Sea  $V$  un espacio vectorial de dimensión finita y  $W$  un subespacio de  $V$ , entonces  $W$  tiene dimensión finita y  $\dim(W) \leq \dim(V)$ .

*Demostración.* Sea  $\dim(V) = n$ . Si  $W = \{\mathbf{0}\} \implies \dim(W) = 0 \leq n$ . Consideremos ahora que  $W$  contiene a un vector no nulo  $\mathbf{x}_1$ , entonces el conjunto  $\{\mathbf{x}_1\}$  es linealmente independiente. Supongamos que seguimos agregando más vectores  $\mathbf{x}_2, \dots, \mathbf{x}_k$  de  $W$  al conjunto  $\{\mathbf{x}_1\}$  de tal forma que  $\{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_k\}$  sea linealmente independiente. Ya que  $\dim(V) = n$ , entonces cualquier base de  $V$  tiene  $n$  elementos. Esto implica que ningún subconjunto de  $V$  linealmente independiente puede tener más de  $n$  elementos, por lo cual el proceso anterior debe detenerse para algún  $k \leq n$ . De acuerdo al Teorema 3.3.3.3, este conjunto genera a  $W$ , por lo cual forma una base de  $W$ , de donde concluimos que  $\dim(W) = k \leq n$ .  $\square$

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

**Teorema 4.2.3.** Sea  $V$  un espacio vectorial. Si  $\dim(V) = n$  entonces cualquier conjunto de  $n$  vectores linealmente independientes de  $V$  es una base de  $V$ .

*Demostración.* Esta prueba se hará por contradicción.

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

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

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

**Corolario** De los dos teoremas anteriores podemos concluir que si  $W$  es un subespacio vectorial de  $V$  y  $\dim(W) = \dim(V) \implies W = V$ .

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 4.2.4.** 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 3.2.1 de la sección 5.2 sabemos que  $\langle D' \rangle = \langle D \rangle = V$ . Si el conjunto generador  $D'$  no es linealmente independiente, podemos seguir retirando vectores de la misma forma sin afectar su espacio generado hasta obtener un conjunto linealmente independiente que genera a  $V$ , es decir, una base para  $V$ .  $\square$

**Teorema 4.2.5.** 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 4.2.1. 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 3.3.3.3 de la sección 5.3.3, es linealmente independiente. Si  $L'$  no genera a  $V$ , podemos repetir el proceso hasta llegar a un conjunto linealmente independiente que genera a  $V$ , i.e., una base para  $V$ .  $\square$

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

### 6.3. Ortogonalización y ortonormalización

Recordemos de la sección 4.3 que dos vectores  $\mathbf{u}$  y  $\mathbf{v}$  son ortogonales si  $(\mathbf{u}, \mathbf{v}) = 0$ . Supongamos que este no es el caso, i.e., que  $(\mathbf{u}, \mathbf{v}) \neq 0$ <sup>32</sup>. Existe una manera de modificar cualquiera de los vectores de tal forma que se vuelva ortogonal al otro: simplemente definimos  $\mathbf{u}' = \mathbf{u} - (\mathbf{u}, \mathbf{v}) \frac{\mathbf{v}}{\|\mathbf{v}\|^2}$  y comprobamos que

$$(\mathbf{u}', \mathbf{v}) = \left( \mathbf{u} - (\mathbf{u}, \mathbf{v}) \frac{\mathbf{v}}{\|\mathbf{v}\|^2}, \mathbf{v} \right) = (\mathbf{u}, \mathbf{v}) - \left( \frac{(\mathbf{u}, \mathbf{v})}{\|\mathbf{v}\|^2} \mathbf{v}, \mathbf{v} \right) = (\mathbf{u}, \mathbf{v}) - \frac{(\mathbf{u}, \mathbf{v})}{(\mathbf{v}, \mathbf{v})} (\mathbf{v}, \mathbf{v}) = (\mathbf{u}, \mathbf{v}) - (\mathbf{u}, \mathbf{v}) \cdot 1 = 0.$$

A esto se le conoce como un proceso de *ortogonalización*. Observemos que, si  $\mathbf{v}$  es un vector unitario (i.e., si  $\|\mathbf{v}\| = 1$ ) entonces la definición de  $\mathbf{u}'$  se reduce a  $\mathbf{u}' = \mathbf{u} - (\mathbf{u}, \mathbf{v})\mathbf{v}$ , simplificando el proceso. Por razones que irán quedando más claras con la experiencia, a menudo es conveniente trabajar con bases de vectores que sean ortogonales entre sí y, en casos específicos, que además sean unitarios, por lo cual damos las siguientes definiciones..

#### 6.3.1. Definiciones

Def. Sea  $O = \{\mathbf{o}_1, \mathbf{o}_2, \dots, \mathbf{o}_n\}$  una base de un espacio vectorial  $V$ . Decimos que  $O$  es una *base ortogonal* si cada uno de sus vectores es ortogonal a todos los demás, i.e., si  $(\mathbf{o}_i, \mathbf{o}_j) = 0 \quad \forall \mathbf{o}_i, \mathbf{o}_j \in O, i \neq j$ .

Def. Sea  $N = \{\mathbf{n}_1, \mathbf{n}_2, \dots, \mathbf{n}_n\}$  una base de un espacio vectorial  $V$ . Decimos que  $N$  es una *base ortonormal* si es una base ortogonal y, además, todos sus vectores son unitarios, i.e., si  $(\mathbf{n}_i, \mathbf{n}_j) = \delta_{ij} \quad \forall \mathbf{n}_i, \mathbf{n}_j \in N$ , donde  $\delta_{ij}$  es la *delta de Kronecker* de dos índices<sup>a</sup>.

<sup>a</sup>En general, la delta de Kronecker de  $n$  índices se define como  $\delta_{ab\dots n} = 1$  si  $a = b = \dots = n$  y  $\delta_{ab\dots n} = 0$  en cualquier otro caso.

Antes de ver cómo podemos construir bases ortogonales y ortonormales, el siguiente teorema y corolario nos ayudarán a comenzar a entender su utilidad.

<sup>32</sup>Aquí implícitamente estamos asumiendo que  $\mathbf{u}$  y  $\mathbf{v}$  son vectores no nulos ya que, por definición, el producto escalar *distingue* al vector nulo (ver sección 4.1.1).

**Teorema 4.3.1.1.** Sea  $V$  sobre  $K$  un espacio vectorial con producto escalar y  $S = \{\mathbf{v}_1, \dots, \mathbf{v}_n\}$  un subconjunto ortogonal de  $V$  con  $\mathbf{v}_i \neq \mathbf{0} \quad \forall 1 \leq i \leq n$ . Si  $\mathbf{u} \in \langle S \rangle \implies$

$$\mathbf{u} = \sum_{i=1}^n \frac{(\mathbf{u}, \mathbf{v}_i)}{\|\mathbf{v}_i\|^2} \mathbf{v}_i.$$

*Demostración.* Ya que  $\mathbf{u} \in \langle S \rangle \implies \mathbf{u} = c_1 \mathbf{v}_1 + \dots + c_n \mathbf{v}_n = \sum_{i=1}^n c_i \mathbf{v}_i$  para algunos coeficientes  $c_i \in K$ . Para ver precisamente quiénes son esos coeficientes  $c_i$ , observemos que para  $1 \leq j \leq n$

$$(\mathbf{u}, \mathbf{v}_j) = \left( \sum_{i=1}^n c_i \mathbf{v}_i, \mathbf{v}_j \right).$$

Aplicando las propiedades del producto escalar y la definición de conjunto ortogonal, tenemos que

$$(\mathbf{u}, \mathbf{v}_j) = \sum_{i=1}^n (c_i \mathbf{v}_i, \mathbf{v}_j) = \sum_{i=1}^n c_i (\mathbf{v}_i, \mathbf{v}_j) = c_j (\mathbf{v}_j, \mathbf{v}_j) = c_j \|\mathbf{v}_j\|^2 \iff c_j \|\mathbf{v}_j\|^2 = (\mathbf{u}, \mathbf{v}_j) \iff c_j = \frac{(\mathbf{u}, \mathbf{v}_j)}{\|\mathbf{v}_j\|^2}.$$

Sustituyendo, tenemos que

$$\mathbf{u} = \sum_{i=1}^n c_i \mathbf{v}_i = \sum_{i=1}^n \frac{(\mathbf{u}, \mathbf{v}_i)}{\|\mathbf{v}_i\|^2} \mathbf{v}_i,$$

como se quería demostrar originalmente.  $\square$

Habíamos visto con anterioridad que cualquier vector puede ser expresado como una combinación lineal única de elementos de su base; sin embargo, no habíamos entrado en detalles sobre cómo obtener los coeficientes necesarios para esto más allá de plantear y resolver un sistema de ecuaciones. Si aplicamos el teorema anterior a una base ortogonal  $O$  de un espacio vectorial  $V$ , entonces  $\forall \mathbf{u} \in V$  tenemos una *receta* para obtener directamente los coeficientes necesarios para expresar a  $\mathbf{v}$  como combinación lineal de los vectores de  $O$ : de ahí viene la utilidad de las bases ortogonales.

Nótese que, en particular, si el conjunto  $S$  es *ortonormal*, entonces  $c_j = (\mathbf{u}, \mathbf{v}_j) \mathbf{v}_j$  en la demostración anterior y el resultado que obtuvimos se reduce a

$$\mathbf{u} = \sum_{i=1}^n (\mathbf{u}, \mathbf{v}_i) \mathbf{v}_i.$$

Por lo tanto, si aplicamos el Teorema 4.3.1.1 a una base ortonormal  $N$ , los coeficientes mencionados en el párrafo anterior son simplemente el producto escalar del vector  $\mathbf{u}$  con cada vector de la base ortonormal.

**Corolario** Sea  $V$  sobre  $K$  un espacio vectorial con producto interior y  $S$  un subconjunto ortogonal con vectores no nulos, entonces  $S$  es linealmente independiente.

*Demostración.* Supongamos que  $S = \{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ . Fijamos nuestra atención en la combinación lineal  $\sum_{i=1}^n c_i \mathbf{v}_i = \mathbf{0}$ , con  $c_i \in K$ . Aplicando el Teorema 4.3.1.1 con  $\mathbf{u} = \mathbf{0}$  tenemos que  $c_i = \frac{(\mathbf{0}, \mathbf{v}_i)}{\|\mathbf{v}_i\|^2} = 0 \quad \forall c_i \implies S$  es linealmente independiente.  $\square$

Observemos que este corolario es una generalización del primer ejercicio de repaso de la sección 3. Además, este corolario y el teorema del cual se desprende nos dicen que los conjuntos ortogonales son buenos candidatos para bases, ya que son linealmente independientes y el cálculo de los coeficientes  $c_i$  es sumamente sencillo. En particular, aplicando este corolario a los conjuntos ortogonales vemos que, si  $V$  es un espacio vectorial de dimensión  $n$ , entonces cualquier conjunto ortogonal de  $n$  vectores es una base de  $V$  (ver teorema 4.2.3).

Así como a partir de cualquier conjunto linealmente independiente de un espacio vectorial  $V$  se puede construir una base para  $V$  se puede, además, realizar un proceso de *ortogonalización* u *ortonormalización* de esa misma base. La demostración de este hecho—que también nos deletrea el proceso a seguir para lograrlo— se conoce como el Teorema de Gram-Schmidt.

### 6.3.2. Teorema de Gram-Schmidt

**Teorema 4.3.2.1 (Gram-Schmidt).** Sea  $V$  sobre  $K$  un espacio vectorial y  $S = \{\mathbf{u}_1, \dots, \mathbf{u}_n\}$  un subconjunto linealmente independiente de  $V$ . Si definimos al conjunto  $S' = \{\mathbf{v}_1, \dots, \mathbf{v}_n\}$  de tal forma que  $\mathbf{v}_1 = \mathbf{u}_1$  y

$$\mathbf{v}_k = \mathbf{u}_k - \sum_{j=1}^{k-1} \frac{(\mathbf{u}_k, \mathbf{v}_j)}{\|\mathbf{v}_j\|^2} \mathbf{v}_j, \quad 2 \leq k \leq n,$$

entonces  $S'$  es un subconjunto ortogonal de  $V$  tal que  $\langle S' \rangle = \langle S \rangle$ .

*Demostración.* Esta demostración se hará por inducción sobre  $n$  y, para realizarla, definimos a  $S_k \equiv \{\mathbf{u}_1, \dots, \mathbf{u}_k\}$  para  $k = 1, 2, \dots, n$ .

**Base inductiva** Si  $n = 1$ , entonces el teorema se demuestra trivialmente, ya que  $S' = S$ , por lo cual  $\langle S' \rangle = \langle S \rangle$  y, además,  $S'$  sería un subconjunto ortogonal de  $V$  por vacuidad.

**Hipótesis de inducción** Supongamos que el conjunto  $S'_{k-1} = \{\mathbf{v}_1, \dots, \mathbf{v}_{k-1}\}$  ha sido construido siguiendo el proceso descrito en el planteamiento del teorema y que cumple las propiedades deseadas, es decir, que es un conjunto ortogonal tal que  $\langle S'_{k-1} \rangle = \langle S_{k-1} \rangle$ , donde  $S_{k-1}$  es linealmente independiente.

**Paso inductivo** Sea  $S_k$  un conjunto linealmente independiente y  $S'_k = \{\mathbf{v}_1, \dots, \mathbf{v}_{k-1}, \mathbf{v}_k\}$  tal que

$$\mathbf{v}_k = \mathbf{u}_k - \sum_{j=1}^{k-1} \frac{(\mathbf{u}_k, \mathbf{v}_j)}{\|\mathbf{v}_j\|^2} \mathbf{v}_j.$$

Si  $\mathbf{v}_k = \mathbf{0}$  entonces la ecuación anterior implicaría que  $\mathbf{u}_k \in \langle S'_{k-1} \rangle = \langle S_{k-1} \rangle$ , lo cual contradice el hecho de que  $S_k$  es un conjunto linealmente independiente, por lo cual forzosamente  $\mathbf{v}_k \neq \mathbf{0}$ .

Para  $1 \leq i \leq k-1$  se sigue que

$$(\mathbf{v}_k, \mathbf{v}_i) = \left( \mathbf{u}_k - \sum_{j=1}^{k-1} \frac{(\mathbf{u}_k, \mathbf{v}_j)}{\|\mathbf{v}_j\|^2} \mathbf{v}_j, \mathbf{v}_i \right) = (\mathbf{u}_k, \mathbf{v}_i) - \sum_{j=1}^{k-1} \frac{(\mathbf{u}_k, \mathbf{v}_j)}{\|\mathbf{v}_j\|^2} (\mathbf{v}_j, \mathbf{v}_i) = (\mathbf{u}_k, \mathbf{v}_i) - \frac{(\mathbf{u}_k, \mathbf{v}_i)}{\|\mathbf{v}_i\|^2} \|\mathbf{v}_i\|^2 = 0,$$

ya que en nuestra hipótesis inductiva supusimos que  $S'_{k-1}$  es ortogonal. Por ende,  $S'_k$  es un conjunto ortogonal de vectores no nulos.

Finalmente, por construcción  $S'_k \subseteq \langle S_k \rangle$ , lo cual también implica que  $\langle S'_k \rangle \subseteq \langle S_k \rangle$ ; análogamente, por la forma en que construimos  $S'_k$ ,  $S_k \subseteq \langle S'_k \rangle$ , lo cual implica que  $\langle S_k \rangle \subseteq \langle S'_k \rangle$ . Juntando ambos resultados, concluimos que  $\langle S_k \rangle = \langle S'_k \rangle$ , como queríamos demostrar. □

Aplicando este teorema a bases, nos dice que a partir de una base arbitraria  $B$  de un espacio vectorial  $V$ , se puede construir una base ortogonal  $O$  para ese mismo espacio  $V$ : sólo hace falta escoger algún vector  $\mathbf{b}_1 \in B$  con el cual definir  $\mathbf{o}_1 \equiv \mathbf{b}_1$  y después seguir el procedimiento descrito en el teorema.

Además, observemos que se podría modificar ligeramente el procedimiento seguido en el teorema anterior definiendo al primer vector de la nueva base como un vector unitario. Es más, supongamos que definimos un nuevo conjunto generador  $N$  con  $\mathbf{n}_1 = \frac{\mathbf{b}_1}{\|\mathbf{b}_1\|}$  y después, cada vez que obtenemos un nuevo vector ortogonal a los anteriores siguiendo el proceso del Teorema 4.3.2.1, lo normalizamos antes de añadirlo a  $N$ : en este caso, el conjunto generador resultante será una base *ortonormal* de  $V$ . Es decir, podríamos hacer un Teorema de Gram-Schmidt *modificado* de tal forma que a partir de cualquier subconjunto linealmente independiente podamos generar un subconjunto ortonormal que genere al mismo subespacio que el anterior. Aplicado a bases, estaríamos asegurando que a partir de cualquier base  $B$  de un espacio vectorial  $V$ , se puede construir una base ortonormal para  $V$ , además de detallar el proceso mediante el cual se construye dicha base ortonormal.

## 6.4. Ejercicios de repaso

### 6.4.1. Bases

1. Sea  $K$  un campo arbitrario. En el ejercicio 1.5.2.4 de la sección 3.5.2 demostraste que  $K \times K \times \dots \times K = K^n, n \in \mathbb{N}$  era un espacio vectorial sobre  $K$ . Sea  $\mathbf{e}_1 = \begin{bmatrix} 1 & 0 & 0 & 0 & \dots & 0 \end{bmatrix}^T, \mathbf{e}_2 = \begin{bmatrix} 0 & 1 & 0 & 0 & \dots & 0 \end{bmatrix}^T, \dots, \mathbf{e}_n = \begin{bmatrix} 0 & 0 & 0 & 0 & \dots & 1 \end{bmatrix}^T$ . Demuestra que  $\{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n\}$  es una base para  $K^n$  sobre  $K$ . En particular, demuestra que es una base ortonormal. A esta base se le conoce como la *base canónica* del espacio vectorial  $K^n$  sobre  $K$ . (1.25 pts.)
2. Sea  $d_i$  el  $i$ -ésimo dígito de tu número de cuenta. Considera el conjunto  $\left\{ \begin{bmatrix} d_1 & d_2 & d_3 \end{bmatrix}^T, \begin{bmatrix} d_4 & d_5 & d_6 \end{bmatrix}^T, \begin{bmatrix} d_7 & d_8 & d_9 \end{bmatrix}^T \right\}$ . ¿Forma una base para  $\mathbb{R}^3$ ? Si sí lo es, demuéstalo. Si no, demuestra por qué no es base de  $\mathbb{R}^3$  y realiza las modificaciones necesarias para obtener una base de  $\mathbb{R}$ . (1.25 pts.)
3. Sea  $B$  la base que obtuviste en el ejercicio anterior. Expresa a los vectores  $\begin{bmatrix} d_9 & d_8 & d_7 \end{bmatrix}^T, \begin{bmatrix} d_6 & d_5 & d_4 \end{bmatrix}^T, \begin{bmatrix} d_3 & d_2 & d_1 \end{bmatrix}^T \in \mathbb{R}^3$  como combinaciones lineales de los vectores de  $B$ . (1.25 pts.)

### 6.4.2. Dimensión

1. 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. (1.25 pts.)
2. Demuestra que cualquier base del espacio vectorial  $\mathbb{C}$  sobre  $\mathbb{R}$  debe tener más de un elemento y que, por lo tanto, la  $\dim((\mathbb{C}, \mathbb{C})) \neq \dim((\mathbb{C}, \mathbb{R}))$ . (1.25 pts.)
3. Da un ejemplo distinto a  $\mathbb{C}$  en donde la dimensión (algebraica) de un espacio vectorial no corresponda con la dimensión (geométrica) de su representación geométrica. Además, da un ejemplo distinto al del ejercicio anterior en donde la dimensión de un espacio vectorial cambie según el campo sobre el cual está definido. (1.25 pts.)

### 6.4.3. Ortogonalización y ortonormalización

1. Sea  $B$  la base que obtuviste del ejercicio 4.4.1.2. Aplica el Teorema de Gram-Schmidt y construye una base ortogonal de  $\mathbb{R}^3$  a partir de  $B$ . (1.25 pts.)
2. Aplica el Teorema de Gram-Schmidt *modificado* y construye una base ortonormal a partir de  $B$ . (1.25 pts.)
3. Sea  $P_0(x) = 1, P_1(x) = x, P_2(x) = \frac{3x^2-1}{2}, P_3(x) = \frac{5x^3-3x}{2}$ . Demuestra que el conjunto  $\{P_0, P_1, P_2, P_3\}$  es una base del espacio vectorial  $P^3$  y di si es ortogonal y/o ortonormal en el intervalo  $[-1, 1]$  (nota: tendrás que utilizar la definición de producto escalar para funciones vista en la sección 4.1.2). (1.5 pts. extra)\*

## 7. Definición de transformación lineal y espacio de transformaciones lineales, núcleo e imagen de una transformación lineal, nulidad y rango

### 7.1. Definición de transformación lineal y espacio de transformaciones lineales

Recordemos de la sección 3.2.1 que las dos operaciones que definen a los espacios vectoriales son la suma vectorial y el producto de un vector por un escalar. Cuando una operación es compatible con las dos operaciones mencionadas anteriormente, se dice que tiene *propiedades lineales*, o bien, que es una *operación lineal* (e.g., la operación de producto escalar es *lineal en la primera entrada*, como vimos en la sección 4.1.1).

En general, se pueden definir funciones arbitrarias que vayan de un espacio vectorial  $V$  a un espacio vectorial  $W$  (admitiendo la posibilidad de que  $V = W$ ). A aquellas de estas funciones que, además, sean compatibles con las operaciones de suma vectorial y producto de un vector por un escalar, se les conoce como *transformaciones lineales*.

#### 7.1.1. Definición de transformación lineal

Def. Sean  $V$  y  $W$  espacios vectoriales sobre un mismo campo  $K$ , entonces una *transformación lineal* es una función  $T : V \rightarrow W$  tal que

$$T(\mathbf{v}_1 + \mathbf{v}_2) = T(\mathbf{v}_1) + T(\mathbf{v}_2) \text{ y } T(c\mathbf{v}_1) = cT(\mathbf{v}_1) \quad \forall \mathbf{v}_1, \mathbf{v}_2 \in V, c \in K.$$

Si juntamos ambas propiedades, podemos observar que las transformaciones lineales son compatibles con las combinaciones lineales:

$$T\left(\sum_{i=1}^n c_i \mathbf{v}_i\right) = T(c_1 \mathbf{v}_1 + \dots + c_n \mathbf{v}_n) = T(c_1 \mathbf{v}_1) + \dots + T(c_n \mathbf{v}_n) = c_1 T(\mathbf{v}_1) + \dots + c_n T(\mathbf{v}_n) = \sum_{i=1}^n c_i T(\mathbf{v}_i).$$

#### 7.1.2. Ejemplos de transformaciones lineales

##### Identidad

En cualquier espacio vectorial  $V$ , la transformación identidad  $I_V : V \rightarrow V$  dada por  $I_V(\mathbf{x}) = \mathbf{x} \quad \forall \mathbf{x} \in V$  es una transformación lineal.

##### Producto de un vector por un escalar

Sea  $V$  sobre  $K$  un espacio vectorial con  $\mathbf{v} \in V$  y  $a \in K$ . Definimos a la transformación lineal  $T_a : V \rightarrow V$  como  $T_a(\mathbf{v}) = a\mathbf{v}$  y observamos que  $\forall \mathbf{v}, \mathbf{u} \in V, \forall b \in K$  tenemos que

$$T_a(b\mathbf{u} + \mathbf{v}) = a(b\mathbf{u} + \mathbf{v}) = a(b\mathbf{u}) + a\mathbf{v} = b(a\mathbf{u}) + a\mathbf{v} = b(T_a(\mathbf{u})) + T_a(\mathbf{v}),$$

por lo cual  $T_a$  es una transformación lineal  $\forall a \in K$ . En particular,  $T_0$  es una transformación lineal que envía a todos los vectores de  $V$  al vector nulo  $\mathbf{0}_V$ , conocida como la *transformación nula* de  $V$  a  $V$ .

## Rotaciones en dos dimensiones

Sea  $\mathbf{v} = \begin{bmatrix} v_1 & v_2 \end{bmatrix}^T$  un vector de  $\mathbb{R}^2$  que forma un ángulo  $\alpha$  con el eje horizontal del plano cartesiano. Entonces, la transformación  $T_\theta : \mathbb{R}^2 \rightarrow \mathbb{R}^2$  dada por

$$T_\theta(\mathbf{v}) = \begin{bmatrix} \sqrt{v_1^2 + v_2^2} \cos(\alpha + \theta) & \sqrt{v_1^2 + v_2^2} \sin(\alpha + \theta) \end{bmatrix}^T = \begin{bmatrix} v_1 \cos \theta - v_2 \sin \theta & v_1 \sin \theta + v_2 \cos \theta \end{bmatrix}^T,$$

para cualquier ángulo  $\theta$ , es una transformación lineal. Geométricamente, lo que hace esta transformación lineal es rotar  $\theta$  grados a cualquier vector  $\mathbf{v} \in \mathbb{R}^2$ .

## Reflexiones

En  $\mathbb{R}^3$ , las reflexiones con respecto a los ejes  $\hat{x}$ ,  $\hat{y}$  y  $\hat{z}$ , los planos  $\hat{x}\hat{y}$ ,  $\hat{y}\hat{z}$ ,  $\hat{x}\hat{z}$  y el origen  $O$  del espacio cartesiano son transformaciones lineales. Podemos definir las como sigue:

$$T_{\hat{x}}(\mathbf{v}) = \begin{bmatrix} v_1 & -v_2 & -v_3 \end{bmatrix}^T; \quad T_{\hat{y}}(\mathbf{v}) = \begin{bmatrix} -v_1 & v_2 & -v_3 \end{bmatrix}^T; \quad T_{\hat{z}}(\mathbf{v}) = \begin{bmatrix} -v_1 & -v_2 & v_3 \end{bmatrix}^T;$$

$$T_{\hat{x}\hat{y}}(\mathbf{v}) = \begin{bmatrix} v_1 & v_2 & -v_3 \end{bmatrix}^T; \quad T_{\hat{y}\hat{z}}(\mathbf{v}) = \begin{bmatrix} -v_1 & v_2 & v_3 \end{bmatrix}^T; \quad T_{\hat{x}\hat{z}}(\mathbf{v}) = \begin{bmatrix} v_1 & -v_2 & v_3 \end{bmatrix}^T; \quad T_O(\mathbf{v}) = \begin{bmatrix} -v_1 & -v_2 & -v_3 \end{bmatrix}^T.$$

En general, las reflexiones son transformaciones lineales en  $\mathbb{R}^n$ ,  $n \in \mathbb{N}$  definiéndolas de manera análoga —aunque para  $n \geq 4$  no sea fácil de visualizar. Además, también se cumple para los espacios vectoriales complejos  $\mathbb{C}^n$ ,  $n \in \mathbb{N}$ .

## Producto escalar

Sea  $V$  un espacio vectorial real con producto escalar. Entonces, por lo observado en la sección 4.1.1, en este caso el producto escalar es una operación lineal *en ambas entradas*, también conocida como *transformación bilineal* o, simplemente, *forma bilineal*. En cambio, si  $V$  es un espacio vectorial complejo, entonces  $T_{\mathbf{u}} : V \rightarrow \mathbb{C}$  dada por  $T_{\mathbf{u}}(\mathbf{v}) = (\mathbf{v}, \mathbf{u})$  es una transformación lineal.

## Proyecciones

Recordando la definición algebraica de la proyección sobre  $\mathbf{u}$  de un vector  $\mathbf{v}$  dada en la sección 4.3.1

$$P_{\mathbf{u}}(\mathbf{v}) = \frac{(\mathbf{v}, \mathbf{u})}{\|\mathbf{u}\|^2} \mathbf{u},$$

donde  $\mathbf{u} \neq \mathbf{0}$ , observemos que  $\forall a \in K$ ,  $\forall \mathbf{v}, \mathbf{w} \in V$  se cumple que

$$P_{\mathbf{u}}(a\mathbf{v} + \mathbf{w}) = \frac{(a\mathbf{v} + \mathbf{w}, \mathbf{u})}{\|\mathbf{u}\|^2} \mathbf{u} = \frac{a(\mathbf{v}, \mathbf{u}) + (\mathbf{w}, \mathbf{u})}{\|\mathbf{u}\|^2} \mathbf{u} = a \frac{(\mathbf{v}, \mathbf{u})}{\|\mathbf{u}\|^2} \mathbf{u} + \frac{(\mathbf{w}, \mathbf{u})}{\|\mathbf{u}\|^2} \mathbf{u} = aP_{\mathbf{u}}(\mathbf{v}) + P_{\mathbf{u}}(\mathbf{w});$$

es decir que, en general, la proyección de un vector sobre otro es una transformación lineal.

## Derivadas e integrales

En  $P^n$ , el espacio vectorial real de funciones polinomiales reales de grado  $\leq n$ , la derivada  $\frac{d}{dx} : P^n \rightarrow P^{n-1}$  es una transformación lineal. Por otro lado, integral  $\int dx : P^n \rightarrow P^{n+1}$  también es una transformación lineal. Ambos hechos se desprenden de las propiedades

$$\begin{aligned} \frac{d}{dx}(cf(x)) &= c \frac{d}{dx}f(x), \quad \frac{d}{dx}(f(x) + g(x)) = \frac{d}{dx}f(x) + \frac{d}{dx}g(x), \\ \int cf(x)dx &= c \int f(x)dx \quad \text{y} \quad \int f(x) + g(x)dx = \int f(x)dx + \int g(x)dx, \end{aligned}$$

vistas en cursos de cálculo, las cuales son llamadas, apropiadamente, *propiedades lineales de la derivada y la integral*.

### 7.1.3. Espacio de transformaciones lineales

Sean  $V$  y  $W$  espacios vectoriales sobre el mismo campo  $K$ . Definimos al conjunto de transformaciones lineales de  $V$  a  $W$  como

$$\mathcal{L}(V, W) = \{T : V \rightarrow W \mid T(\mathbf{u} + \lambda \mathbf{v}) = T(\mathbf{u}) + \lambda T(\mathbf{v}) \quad \forall \mathbf{u}, \mathbf{v} \in V, \lambda \in K\}.$$

Sean  $T_1, T_2 \in \mathcal{L}(V, W)$ . Ya que  $T_1$  y  $T_2$  son simplemente un tipo particular de *funciones*, resulta natural definir la suma de las transformaciones lineales  $T_1$  y  $T_2$  como  $(T_1 + T_2)(\mathbf{x}) = T_1(\mathbf{x}) + T_2(\mathbf{x})$ . Observemos entonces que  $\forall \mathbf{u}, \mathbf{v} \in V, \lambda \in K$ ,

$$(T_1 + T_2)(\mathbf{u} + \lambda \mathbf{v}) = T_1(\mathbf{u} + \lambda \mathbf{v}) + T_2(\mathbf{u} + \lambda \mathbf{v}) = T_1(\mathbf{u}) + \lambda T_1(\mathbf{v}) + T_2(\mathbf{u}) + \lambda T_2(\mathbf{v}).$$

Ya que las imágenes  $T_1(\mathbf{u}), T_1(\mathbf{v}), T_1(\mathbf{u})$  y  $T_2(\mathbf{u})$  son vectores de  $W$  y la suma de vectores en  $W$  conmuta y se distribuye con respecto al producto por escalares del campo  $K$ , tenemos que

$$(T_1 + T_2)(\mathbf{u} + \lambda \mathbf{v}) = T_1(\mathbf{u}) + T_2(\mathbf{u}) + \lambda(T_1(\mathbf{v}) + T_2(\mathbf{v})) = (T_1 + T_2)(\mathbf{u}) + \lambda(T_1 + T_2)(\mathbf{v}),$$

por lo cual  $T_1 + T_2 \in \mathcal{L}(V, W)$ . Además, ya que la suma de funciones es asociativa y conmutativa, entonces la suma de transformaciones lineales también lo es.

Ahora, sea  $\alpha \in K$  y  $T \in \mathcal{L}(V, W)$ . Nuevamente, ya que  $T$  es una función, resulta natural definir el producto de la transformación  $T$  por  $\alpha$  como  $(\alpha T)(\mathbf{x}) = \alpha(T(\mathbf{x}))$ <sup>33</sup>. A partir de esta definición, podemos observar que

$$(\alpha T)(\mathbf{u} + \lambda \mathbf{v}) = \alpha(T(\mathbf{u} + \lambda \mathbf{v})) = \alpha(T(\mathbf{u}) + \lambda T(\mathbf{v})) = \alpha T(\mathbf{u}) + \alpha(\lambda T(\mathbf{v})).$$

Luego, por compatibilidad del producto de un vector por un escalar con el producto entre escalares, tenemos que

$$(\alpha T)(\mathbf{u} + \lambda \mathbf{v}) = \alpha T(\mathbf{u}) + (\alpha \lambda)T(\mathbf{v}) = \alpha T(\mathbf{u}) + (\lambda \alpha)T(\mathbf{v}) = (\alpha T)(\mathbf{u}) + \lambda((\alpha T)(\mathbf{v})),$$

por lo cual  $\alpha T \in \mathcal{L}(V, W)$ .

Dadas las dos operaciones que acabamos de definir sobre el conjunto  $\mathcal{L}(V, W)$ , podemos verificar las demás propiedades de los espacios vectoriales y finalmente demostrar que este conjunto, junto con las dos operaciones ya mencionadas, forma un espacio vectorial sobre  $K$  (la verificación de las propiedades restantes se deja como ejercicio). En este caso, el neutro aditivo de  $\mathcal{L}(V, W)$  es la transformación nula  $T_0(\mathbf{v}) = \mathbf{0}_W \quad \forall \mathbf{v} \in V$  (verifica que es lineal) y el inverso aditivo de cualquier transformación  $T \in \mathcal{L}(V, W)$  es  $-T = (-1)T \in \mathcal{L}(V, W)$  (¿por qué?).

En resumen, debido a la naturaleza de las transformaciones lineales (es decir, que por definición sean compatibles con las operaciones esenciales de los espacios vectoriales), siempre que tengamos dos espacios vectoriales, ¡el conjunto de todas las transformaciones lineales de uno a otro *hereda* la estructura de espacio vectorial!

## 7.2. Núcleo e imagen de una transformación lineal

Para entender mejor cómo actúa una transformación lineal dada sobre un espacio vectorial, enfocamos ahora nuestra atención en dos subconjuntos especiales.

<sup>33</sup>En analogía al producto de una función por una función constante, podríamos pensar que esta definición de  $\alpha T$  es, naturalmente, *el producto de la transformación lineal  $T$  por la transformación lineal constante  $\alpha$* ; sin embargo, esta afirmación sería errónea. ¿Por qué?



Def. Sean  $V, W$  espacios vectoriales sobre  $K$  y  $T : V \rightarrow W$  una transformación lineal. Definimos al *núcleo* o *espacio nulo* de  $T$  como el conjunto de todos los vectores de  $V$  que son enviados por  $T$  al vector nulo de  $W$ . Matemáticamente<sup>a</sup>,

$$\ker(T) = \{\mathbf{v} \in V \mid T(\mathbf{v}) = \mathbf{0}_W\}.$$

Def. Definimos a la *imagen* de  $T$  como el conjunto de todos los vectores en  $W$  obtenidos al aplicar la transformación lineal  $T$  a los vectores de  $V$ , es decir,

$$\operatorname{im}(T) = \{\mathbf{w} \in W \mid T(\mathbf{v}) = \mathbf{w}, \mathbf{v} \in V\}.$$

---

<sup>a</sup>La notación  $\ker(T)$  empleada es la estándar y proviene del inglés, en donde al núcleo de una transformación lineal se le conoce como su *kernel*.

Observemos de la definición que el núcleo de una transformación lineal  $\ker(T)$  es, en cierta forma, análogo al conjunto de las raíces de una función polinomial  $p(x)$ , mientras que la definición de la imagen  $\operatorname{im}(T)$  es totalmente análoga a la definición de la imagen de una función real. Además, por definición de transformación lineal, si  $T : V \rightarrow W$  es una transformación lineal, entonces  $T(\mathbf{0}_V) = \mathbf{0}_W$ .

El siguiente teorema nos muestra un hecho interesante que se desprende a partir de las definiciones de  $\ker(T)$  e  $\operatorname{im}(T)$ .

**Teorema 5.2.1.** Sean  $V, W$  espacios vectoriales y  $T : V \rightarrow W$  una transformación lineal, entonces  $\ker(T)$  e  $\operatorname{im}(T)$  son subespacios vectoriales de  $V$  y  $W$ , respectivamente.

*Demostración.* Sean  $\mathbf{0}_V, \mathbf{0}_W$  los vectores nulos de  $V$  y  $W$ , respectivamente. Ya que  $T(\mathbf{0}_V) = \mathbf{0}_W$ , tenemos que  $\mathbf{0}_V \in \ker(T)$ . Sean  $\mathbf{u}, \mathbf{v} \in \ker(T)$  y  $a \in K$ , entonces

$$T(\mathbf{u} + \mathbf{v}) = T(\mathbf{u}) + T(\mathbf{v}) = \mathbf{0}_W + \mathbf{0}_W = \mathbf{0}_W, \quad T(c\mathbf{u}) = cT(\mathbf{u}) = c\mathbf{0}_W = \mathbf{0}_W,$$

lo cual implica que  $\mathbf{u} + \mathbf{v}, c\mathbf{u} \in \ker(T)$ , por lo cual  $\ker(T)$  es un subespacio vectorial de  $V$ .

Por otro lado,  $T(\mathbf{0}_V) = \mathbf{0}_W \implies \mathbf{0}_W \in \operatorname{im}(T)$ . Sean  $\mathbf{x}, \mathbf{y} \in \operatorname{im}(T)$  y  $c \in K$ . Por lo tanto, existen  $\mathbf{u}, \mathbf{v} \in V$  tal que  $T(\mathbf{u}) = \mathbf{x}$  y  $T(\mathbf{v}) = \mathbf{y}$ . Entonces, por definición de transformación lineal,

$$T(\mathbf{u} + \mathbf{v}) = T(\mathbf{u}) + T(\mathbf{v}) = \mathbf{x} + \mathbf{y}, \quad T(c\mathbf{u}) = cT(\mathbf{u}) = c\mathbf{x},$$

lo cual implica que  $\mathbf{x} + \mathbf{y}, c\mathbf{x} \in \operatorname{im}(T)$ , por lo cual  $\operatorname{im}(T)$  es un subespacio vectorial de  $W$ . □

A continuación, veremos un teorema que nos muestra cómo encontrar un conjunto generador de  $\operatorname{im}(T)$  para cualquier transformación lineal  $T$ , a partir del cual podemos construir una base para  $\operatorname{im}(T)$ .

**Teorema 5.2.2.** Sean  $V$  y  $W$  sobre  $K$  espacios vectoriales, y  $T : V \rightarrow W$  una transformación lineal. Si  $B = \{\mathbf{v}_1, \dots, \mathbf{v}_n\}$  es una base de para  $V$ , entonces

$$\langle T(B) \rangle = \langle \{T(\mathbf{v}_1), \dots, T(\mathbf{v}_n)\} \rangle = \operatorname{im}(T).$$

*Demostración.* Ya que  $\mathbf{v}_i \in V \implies T(\mathbf{v}_i) \in W \forall i$ . Además, ya que  $\operatorname{im}(T)$  es un subespacio vectorial de  $W$ , sus elementos son cerrados por la suma vectorial y el producto de vectores por escalares; es decir, cualquier combinación lineal de vectores de  $\operatorname{im}(T)$  resulta en un vector de  $\operatorname{im}(T) \implies \langle T(B) \rangle \subseteq \operatorname{im}(T)$ .

Por otro lado, sea  $\mathbf{w} \in \operatorname{im}(T) \implies \mathbf{w} = T(\mathbf{v})$  para algún  $\mathbf{v} \in V$ . Ya que  $B$  es base de  $V$ , existen coeficientes  $c_i \in K$  tales que

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

Ya que  $T$  es lineal, se sigue que

$$\mathbf{w} = T(\mathbf{v}) = \sum_{i=1}^n c_i T(\mathbf{v}_i) \in \langle T(B) \rangle \implies \text{im}(T) \subseteq \langle T(B) \rangle.$$

De lo anterior, concluimos que  $\langle T(B) \rangle = \langle \{T(\mathbf{v}_1), \dots, T(\mathbf{v}_n)\} \rangle = \text{im}(T)$ . □

A partir de la definición de  $\ker(T)$  podemos, además, demostrar un criterio interesante para ver cuándo una transformación lineal  $T : V \rightarrow W$  es inyectiva (es decir, que a cada vector de  $\mathbf{v} \in V$  le corresponda sólo un vector de  $\mathbf{w} \in W$  tal que  $T(\mathbf{v}) = \mathbf{w}$ ).

**Teorema 5.2.3.** Sean  $V$  y  $W$  espacios vectoriales, y  $T : V \rightarrow W$  una transformación lineal. Entonces,  $T$  es inyectiva si y sólo si  $\ker(T) = \{\mathbf{0}_V\}$ .

*Demostración.* Supongamos que  $T$  es inyectiva y que  $\mathbf{x} \in \ker(T) \implies T(\mathbf{x}) = \mathbf{0}_W = T(\mathbf{0}_V) \implies \mathbf{x} = \mathbf{0}_V$ , ya que estamos suponiendo que  $T$  es inyectiva. Por lo tanto,  $\ker(T) = \{\mathbf{0}_V\}$ .

Por otro lado, supongamos que  $\ker(T) = \{\mathbf{0}_V\}$  y que  $T(\mathbf{x}) = T(\mathbf{y}) \implies T(\mathbf{x}) - T(\mathbf{y}) = \mathbf{0}_W \implies T(\mathbf{x} - \mathbf{y}) = \mathbf{0}_W \implies \mathbf{x} - \mathbf{y} \in \ker(T) \implies \mathbf{x} - \mathbf{y} = \mathbf{0}_V \implies \mathbf{x} = \mathbf{y}$ , por lo cual  $T$  es inyectiva. □

### 7.3. Nulidad y rango

Def. Sean  $V, W$  espacios vectoriales,  $T : V \rightarrow W$  una transformación lineal, y  $\ker(T), \text{im}(T)$  subespacios de dimensión finita de  $V$  y  $W$ , respectivamente, entonces a la dimensión de  $\ker(T)$  la llamamos la *nulidad* de  $T$  y a la dimensión de  $\text{im}(T)$ , el *rango* de  $T$ .

El resultado más relevante de la imagen y el núcleo de una transformación lineal es la llamada *fórmula de la dimensión*, que se demuestra a continuación.

**Teorema 5.3.1.** Sean  $V, W$  espacios vectoriales de dimensión finita y  $T : V \rightarrow W$  una transformación lineal, entonces la nulidad( $T$ ) + rango( $T$ ) = dim( $V$ ).

*Demostración.* Sea  $\dim(V) = n$ ,  $\dim(\ker(T)) = \text{nulidad}(T) = k$ , y  $N = \{\mathbf{v}_1, \dots, \mathbf{v}_k\}$  una base para  $\ker(T)$ . Por el Teorema 4.2.5 de la sección 6.2, sabemos que podemos extender a  $N$  hasta formar una base  $B = \{\mathbf{v}_1, \dots, \mathbf{v}_k, \mathbf{v}_{k+1}, \dots, \mathbf{v}_n\}$  de  $V$ . Ahora demostraremos que  $S = \{T(\mathbf{v}_{k+1}), \dots, T(\mathbf{v}_n)\}$  es una base de la  $\text{im}(T)$ .

Ya que  $B$  es base de  $V$ , entonces por el Teorema 5.2.2.  $\langle \{T(\mathbf{v}_1), T(\mathbf{v}_2), \dots, T(\mathbf{v}_n)\} \rangle = \text{im}(T) \implies \text{im}(T) = \langle \{T(\mathbf{v}_1), T(\mathbf{v}_2), \dots, T(\mathbf{v}_n)\} \rangle = \langle \{\mathbf{0}_W, \dots, \mathbf{0}_W, T(\mathbf{v}_{k+1}), \dots, T(\mathbf{v}_n)\} \rangle = \langle \{T(\mathbf{v}_{k+1}), \dots, T(\mathbf{v}_n)\} \rangle = \langle S \rangle$ , por lo cual  $S$  genera a  $\text{im}(T)$ . Ahora sólo falta demostrar que es un conjunto linealmente independiente. Supongamos que existen coeficiente  $b_{k+1}, \dots, b_n \in K$  tales que

$$\sum_{i=k+1}^n b_i T(\mathbf{v}_i) = \mathbf{0}_W \implies T\left(\sum_{i=k+1}^n b_i \mathbf{v}_i\right) = \mathbf{0}_W \implies \sum_{i=k+1}^n b_i \mathbf{v}_i \in \ker(T).$$

Ya que  $N$  es una base para  $\ker(T)$ , deben existir coeficientes  $c_1, \dots, c_k \in K$  tales que

$$\sum_{i=1}^k c_i \mathbf{v}_i = \sum_{i=k+1}^n b_i \mathbf{v}_i \iff \sum_{i=1}^k (-c_i) \mathbf{v}_i + \sum_{i=k+1}^n b_i \mathbf{v}_i = \mathbf{0}_W.$$

Ya que  $B$  es base de  $V$ , tenemos que  $b_i = 0 \ \forall i$ , por lo cual concluimos que  $S$  es linealmente independiente y, en particular, que es una base de  $\text{im}(T)$ .

Finalmente, tenemos que  $\text{rango}(T) = n - k \implies \text{nulidad}(T) + \text{rango}(T) = \dim(V)$ . □

El resultado anterior es relevante porque nos ayuda a determinar fácilmente cuándo una transformación lineal es biyectiva, como se verá en el siguiente teorema.

**Teorema 5.3.2.** Sean  $V$  y  $W$  sobre  $K$  espacios vectoriales de la misma dimensión finita, y  $T : V \rightarrow W$  una transformación lineal. Entonces, los enunciados siguientes son equivalentes:

- a)  $T$  es inyectiva
- b)  $T$  es suprayectiva
- c)  $\text{rango}(T) = \dim(V)$ .

*Demostración.* Por el Teorema 5.2.3, sabemos que  $T$  es inyectiva si y sólo si  $\ker(T) = \{\mathbf{0}_V\}$ , i.e., si el núcleo de  $T$  es un espacio de dimensión cero, lo cual implica que  $\text{nulidad}(T) = 0$ . De la fórmula de la dimensión, tenemos que  $\text{nulidad}(T) + \text{rango}(T) = \dim(V) \implies \text{rango}(T) = \dim(V) \iff \dim(\text{im}(T)) = \dim(V)$ .  $\square$

## 7.4. Ejercicios de repaso

### 7.4.1. Definición de transformación lineal y espacio de transformaciones lineales

- Sea  $V$  sobre  $K$  un espacio vectorial y  $\mathbf{v} \in V$ . Definimos una función  $T_{\mathbf{v}} : V \rightarrow V$  con regla de correspondencia  $T_{\mathbf{v}}(\mathbf{u}) = \mathbf{u} + \mathbf{v}$  para todo  $\mathbf{u} \in V$ . Demuestra que, en general, esto no es una transformación lineal. ¿En qué caso sí lo es? (1 pto.)
- Demuestra que para cualquier transformación lineal  $T : V \rightarrow W$ ,  $T(\mathbf{0}_V) = \mathbf{0}_W$ . (1 pto.)
- Sea  $d_i$  el  $i$ -ésimo dígito de tu número de cuenta. Sean  $\mathbf{u} = \begin{bmatrix} d_1 & d_2 \end{bmatrix}^T$ ,  $\mathbf{v} = \begin{bmatrix} d_3 & d_4 \end{bmatrix}^T$ ,  $\mathbf{w} = \begin{bmatrix} d_5 & d_6 \end{bmatrix}^T$  vectores de  $\mathbb{R}^2$ , y  $d_7 \in \mathbb{R}$  un escalar. Muestra algebraicamente que  $P_{\mathbf{u}}(d_7\mathbf{v} + \mathbf{w}) = d_7P_{\mathbf{u}}(\mathbf{v}) + P_{\mathbf{u}}(\mathbf{w})$  y da una interpretación geométrica para este resultado (Nota: te sugiero repasar la sección 4.3.1 para la interpretación geométrica). (1 pto.)
- Demuestra que la función  $T_{\theta} : \mathbb{R}^2 \rightarrow \mathbb{R}^2$  definida en la sección 7.1.2 es una transformación lineal. Además, dado el vector  $\mathbf{x} = \begin{bmatrix} d_8 & d_9 \end{bmatrix}^T$ , calcula  $T_{\frac{\pi}{3}}(\mathbf{x})$  y dibuja a ambos vectores en el plano cartesiano (nota: tendrás que utilizar identidades trigonométricas). (1 pto.)
- Sea  $T : V \rightarrow V$  una transformación lineal. Decimos que  $\mathbf{v} \in V$  es un *punto fijo* de  $T$  si  $T(\mathbf{v}) = \mathbf{v}$ . ¿Quiénes son los puntos fijos de  $T_{\hat{x}\hat{y}} : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ ,  $T_{\hat{z}} : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ ,  $\frac{d}{dx} : C^\infty \rightarrow C^\infty$  y  $\frac{d^2}{dx^2} : C^\infty \rightarrow C^\infty$ ? (1 pto.)
- Sean  $V$  sobre  $K$  y  $W$  sobre  $F$  espacios vectoriales, y sea  $T : V \rightarrow W$  una función que lleva vectores de  $V$  a vectores de  $W$ . ¿Cómo debe ser la relación entre los campos  $W$  y  $F$  para que exista la posibilidad de que  $T$  sea una transformación lineal? (0.5 ptos. extra)\*

### 7.4.2. Núcleo e imagen de una transformación lineal

- Sea  $T : V \rightarrow W$  una transformación lineal arbitraria. Demuestra que  $\ker(T)$  es un subespacio vectorial de  $V$  y que  $\text{im}(T)$  es un subespacio vectorial de  $W$  (puedes dar por hecho que  $V$  y  $W$  son espacios vectoriales). (1 pto.)
- Sea  $D$  el espacio vectorial real de todas las funciones  $f : \mathbb{R} \rightarrow \mathbb{R}$  de clase  $C^n$ , con  $n \in \mathbb{N}$ . ¿Quiénes son  $\ker(\frac{d}{dx})$ ,  $\text{im}(\frac{d}{dx})$ ,  $\ker(\int_a^b dx)$  e  $\text{im}(\int_a^b dx)$ ? (Nota: puedes escoger arbitrariamente a  $a, b \in \mathbb{R}$  para las integrales) (1 pto.)
- Sea  $T : V \rightarrow W$  una transformación lineal biyectiva. Demuestra que  $L \subset V$  es un conjunto linealmente independiente si y sólo si  $T(L) \subset W$  también lo es. (1 pto.)

### 7.4.3. Nulidad y rango

1. Repite el ejercicio 5.4.2.2 para el espacio vectorial de funciones de  $\mathbb{R} \rightarrow \mathbb{R}$  de clase  $C^\infty$ . Además, explica, ¿se cumple aquí la fórmula de la dimensión? ¿Por qué? (1 pto.)
2. Sea  $T : P^n \rightarrow P^{n+1}$  con  $n \in \mathbb{N}$  dada por  $T(p(x)) = xp(x) + \frac{d}{dx}p(x)$ . Demuestra que  $T$  es una transformación lineal, encuentra una base para  $\ker(T)$  y  $\operatorname{im}(T)$  y verifica que se cumpla la fórmula de la dimensión. ¿ $T$  es una transformación inyectiva, suprayectiva o ambas (biyectiva)? Argumenta. (1 pto.)

## 8. Composición de transformaciones lineales, transformación lineal inversa y representación matricial de una transformación lineal

### 8.1. Composición de transformaciones lineales

Recordarás de tus cursos de cálculo que si tenemos dos funciones reales de variable real, entonces podemos definir una operación llamada *composición de funciones*, denotada por el símbolo  $\circ$ , tal que para todas las funciones reales de variable real  $f$  y  $g$  se tiene una nueva función dada por  $(f \circ g)(x) = f(g(x))$ . Además, si  $g : D_g \rightarrow Y$  y  $f : D_f \rightarrow Z$  entonces, para que la función  $f \circ g$  esté bien definida, necesitamos que  $\text{Im}(g) \subseteq D_f$ , lo cual se logra restringiendo el dominio de  $G$  a todos los números reales para los cuales  $\text{Im}(g) \in D_f$ . Esta idea fácilmente se puede generalizar a funciones vectoriales de variable vectorial.

En particular, como hemos visto, las transformaciones lineales no son más que funciones que van de espacio vectorial a espacio vectorial (i.e., funciones vectoriales de variable vectorial) que, además, son compatibles con la suma vectorial y el producto de un vector por un escalar. Por ende, dada la discusión anterior, nos debería parecer natural definir la composición de transformaciones lineales como sigue:

Def. Sean  $T_1 : V \rightarrow W$  y  $T_2 : W \rightarrow Z$  transformaciones lineales, entonces definimos la *composición de las transformaciones lineales*  $T_2$  y  $T_1$  como<sup>a</sup>

$$T_2 \circ T_1 : V \rightarrow Z, \quad (T_2 \circ T_1)(\mathbf{v}) = T_2(T_1(\mathbf{v})), \quad \mathbf{v} \in V.$$

<sup>a</sup>En esta definición se sobreentiende que los espacios vectoriales  $V, W$  y  $Z$  están definidos sobre un mismo campo  $K$ , pero vale la pena preguntarnos: ¿qué problemas podríamos tener si esto no fuera así?

La observación más crucial que podemos hacer a partir de la definición anterior es que,  $\forall a \in K$  y  $\forall \mathbf{u}, \mathbf{v} \in V$ , tenemos que

$$(T_2 \circ T_1)(a\mathbf{u} + \mathbf{v}) = T_2(T_1(a\mathbf{u} + \mathbf{v})) = T_2(aT_1(\mathbf{u}) + T_1(\mathbf{v})) = aT_2(T_1(\mathbf{u})) + T_2(T_1(\mathbf{v})) \implies$$

$$(T_2 \circ T_1)(a\mathbf{u} + \mathbf{v}) = a(T_2 \circ T_1)(\mathbf{u}) + (T_2 \circ T_1)(\mathbf{v}),$$

es decir, que la composición de transformaciones lineales *es nuevamente una transformación lineal*.

Por ejemplo, sean  $\mathbf{v} \in V$  y  $a, b \in K$ , y sean  $T_a$  y  $T_b$  transformaciones lineales de  $V$  a  $V$  definidas como en la sección 7.1.2, entonces podemos componerlas como

$$(T_a \circ T_b)(\mathbf{v}) = T_a(T_b(\mathbf{v})) = T_a(b\mathbf{v}) = a(b\mathbf{v}).$$

En este caso particular, observemos que  $T_a \circ T_b = T_b \circ T_a$  (¿a qué propiedades de los espacios vectoriales se debe esto?).

En cambio, si pensamos en las reflexiones de  $\mathbb{R}^3$  definidas en la sección 7.1.2, observamos que

$$T_{\hat{x}} = T_{\hat{x}\hat{y}} \circ T_{\hat{x}\hat{z}} = T_{\hat{x}\hat{z}} \circ T_{\hat{x}\hat{y}}, \quad T_{\hat{y}} = T_{\hat{x}\hat{y}} \circ T_{\hat{y}\hat{z}} = T_{\hat{y}\hat{z}} \circ T_{\hat{x}\hat{y}}, \quad T_{\hat{z}} = T_{\hat{x}\hat{z}} \circ T_{\hat{y}\hat{z}} = T_{\hat{y}\hat{z}} \circ T_{\hat{x}\hat{z}},$$

$$T_O = T_{\hat{y}\hat{z}} \circ T_{\hat{x}\hat{z}} \circ T_{\hat{x}\hat{y}}, \quad T_{\hat{x}} \circ T_{\hat{x}} = T_{\hat{y}} \circ T_{\hat{y}} = T_{\hat{z}} \circ T_{\hat{z}} = T_O \circ T_O = \dots = I_{\mathbb{R}^3}.$$

De la última equivalencia notamos algo interesante: cualquier reflexión compuesta consigo misma da como resultado la identidad. ¿Será posible esto mismo para otros tipos de transformaciones lineales? Más generalmente, ¿para toda transformación lineal existirá una transformación lineal tal que al componerlas se obtenga la identidad, sin importar en qué orden se componen? La búsqueda de las respuestas a estas preguntas motivan la nuestra sección.

## 8.2. Transformación lineal inversa

Siguiendo con la analogía de transformaciones lineales con funciones (dado que en realidad son funciones de un cierto tipo específico), definimos la transformación lineal inversa:

Def. Sea  $T : V \rightarrow W$  una transformación lineal, entonces decimos que  $T^{-1} : W \rightarrow V$  es su transformación lineal inversa si  $\forall \mathbf{v} \in V, \mathbf{w} \in W$  se cumple que

$$(T^{-1} \circ T)(\mathbf{v}) = \mathbf{v} \quad \text{y} \quad (T \circ T^{-1})(\mathbf{w}) = \mathbf{w}.$$

Dicho de otro modo,  $T : V \rightarrow W$  y  $T^{-1} : W \rightarrow V$  son transformaciones lineales inversas si y sólo si

$$T^{-1} \circ T = I_V \quad \text{y} \quad T \circ T^{-1} = I_W,$$

donde  $I_V$  y  $I_W$  son las funciones identidad en los espacios vectoriales  $V$  y  $W$ , respectivamente.

Por ejemplo, recordando la transformación lineal  $T_a(\mathbf{v}) = a\mathbf{v}$  de la sección 7.1.2 vemos que si  $a \neq 0$ , entonces la transformación lineal  $T_{a^{-1}}$  es su transformación lineal inversa<sup>34</sup>. Claramente, la transformación nula  $T_0$  no tiene ninguna inversa, por lo cual no todas las transformaciones lineales tienen inversas.

Si pensamos en las reflexiones en  $\mathbb{R}^3$ , entonces cualquiera de ellas *es su propia inversa*, como observamos al final de la sección anterior. Esto tiene sentido geoméricamente ya que, si reflejamos a un vector con respecto a un cierto eje/plano/origen dos veces seguidas, volvemos a nuestro vector original. En general, las reflexiones en  $\mathbb{R}^n$  son transformaciones lineales que, compuestas consigo mismas en cualquier orden, dan la identidad. Adoptando la notación  $T \circ T = T^2$ , donde  $T : V \rightarrow V$ , entonces las transformaciones lineales que cumplen  $T^2 = I_V$ , donde  $I_V$  es la identidad del espacio vectorial  $V$  en cuestión, se llaman transformaciones lineales *involutivas*. En particular, para cualquier espacio vectorial  $V$ , la transformación identidad en ese espacio es una transformación lineal involutiva.

## 8.3. Espacios vectoriales isomorfos

Como tal vez recordarás, habíamos dicho en clase que podíamos identificar el subespacio vectorial  $(\{a + i0 \mid a \in \mathbb{R}\}, \mathbb{R})$  del espacio vectorial complejo  $\mathbb{C}$  con el espacio vectorial real  $\mathbb{R}$ . La forma precisa de hacer esta identificación es mediante una transformación lineal biyectiva. Observemos que, una vez establecida una biyección entre estos dos espacios vectoriales, a cualquier operación esencial realizada en uno de ellos le corresponderá una operación equivalente en el otro espacio vectorial (por ejemplo, a la operación  $(a + i0) + (b + i0)$  del primer espacio vectorial le corresponderá la operación  $a + b$  en el segundo, y vice versa). Cuando esto ocurre, decimos que son espacios *isomorfos*.

Def. Decimos que dos espacios vectoriales  $V$  y  $W$  son *isomorfos* si existe una transformación lineal biyectiva entre ellos. Matemáticamente,

$$V \cong W \iff \exists T : V \rightarrow W, T^{-1} : W \rightarrow V \quad \text{t.q.} \quad \forall \mathbf{u}, \mathbf{v} \in V, \lambda \in K$$

$$T(\mathbf{u} + \lambda\mathbf{v}) = T(\mathbf{u}) + \lambda T(\mathbf{v}) \quad \text{y} \quad T \circ T^{-1} = I_W, \quad T^{-1} \circ T = I_V.$$

Por ejemplo, recordemos el ejercicio 3.4.2.2 donde vimos que  $\langle G \rangle = \langle \left\{ \begin{bmatrix} 3 & 0 & 3 \end{bmatrix}^T, \begin{bmatrix} -\frac{1}{2} & 0 & \frac{1}{2} \end{bmatrix}^T \right\} \rangle$  se parecía a  $\mathbb{R}^2$  (en algún sentido que no habíamos precisado). Sea  $T : \langle G \rangle \rightarrow \mathbb{R}^2$  dada por

$$T\left(\begin{bmatrix} a & 0 & b \end{bmatrix}^T\right) = \begin{bmatrix} a & b \end{bmatrix}^T,$$

<sup>34</sup>Aquí vale la pena preguntarnos: ¿cómo se relaciona esto con los axiomas de campo  $y$  de los espacios vectoriales?.

entonces podemos verificar que es lineal para todos los vectores de  $\langle G \rangle$  y escalares de  $\mathbb{R}$  y que, además, su inversa  $T^{-1} : \mathbb{R}^2 \rightarrow \langle G \rangle$  está dada por

$$T^{-1}\begin{bmatrix} c & d \end{bmatrix}^T = \begin{bmatrix} c & 0 & d \end{bmatrix}^T,$$

la cual también es lineal. Siguiendo la definición, entonces  $\langle G \rangle$  y  $\mathbb{R}^2$  son espacios vectoriales isomorfos. Esto significa que cualquier operación realizada con vectores de  $\langle G \rangle$  se puede identificar con una operación análoga realizada en  $\mathbb{R}^2$ , y vice versa. Geométricamente, asumiendo que estamos trabajando en la base canónica, podemos interpretar a  $\langle G \rangle$  como un espacio vectorial de tres dimensiones geométricas que contiene a una copia de  $\mathbb{R}^2$  en su plano  $\hat{x}\hat{z}$ , y no contiene más vectores fuera de este plano.

Más generalmente, en álgebra decimos que dos estructuras algebraicas del mismo tipo (como campos, espacios vectoriales, grupos, etc.) son *isomorfas* si existe un mapeo biyectivo entre ellas<sup>35</sup>. Dos estructuras algebraicas isomorfas tienen las mismas propiedades algebraicas, por lo cual pueden ser identificadas entre sí. De ahí que la palabra empleada, *isomorfismo*, venga del griego *isos* (misma) y *morphí* (forma).

## 8.4. Representación de un vector en una base ordenada

Al trabajar en un espacio vectorial  $V$ , vimos anteriormente que resulta conveniente tener un conjunto de vectores con los cuales podamos expresar a todos los vectores de  $V$  mediante una combinación lineal única; a este conjunto le llamamos una *base* de  $V$ . También hablamos de cómo es posible que un espacio vectorial tenga diferentes bases, por lo cual los vectores del espacio pueden ser *representados* de diferentes formas. Antes de meternos más a fondo en el tema de las representaciones, para mayor comodidad, les exigiremos una propiedad adicional a las bases: un orden.

**Def.** Sea  $V$  un espacio vectorial, entonces una *base ordenada* de  $V$  es un conjunto generador de  $V$  linealmente independiente donde sus elementos están dotados de un orden. Denotamos a una base ordenada de  $n$  elementos como una  $n$ -tupla ordenada. Mientras que las bases

$$\{\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_{n-1}, \mathbf{b}_n\} = \{\mathbf{b}_2, \mathbf{b}_1, \dots, \mathbf{b}_n, \mathbf{b}_{n-1}\} = \dots,$$

etcétera, tenemos que las *bases ordenadas*

$$(\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_{n-1}, \mathbf{b}_n) = (\mathbf{b}_2, \mathbf{b}_1, \dots, \mathbf{b}_n, \mathbf{b}_{n-1}) \iff \mathbf{b}_1 = \mathbf{b}_2, \dots, \mathbf{b}_{n-1} = \mathbf{b}_n.$$

En particular, si tenemos una base ordenada  $(\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_n)$  de un espacio vectorial  $V$  y un vector  $\mathbf{v} \in V$  tal que

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

con  $c_i \in K$ , entonces podemos representar al vector  $\mathbf{v}$  en esta base ordenada simplemente como

$$\mathbf{v} \doteq (c_1, c_2, \dots, c_n) \quad \text{ó} \quad \mathbf{v} \doteq \begin{bmatrix} c_1 & c_2 & \dots & c_n \end{bmatrix}^T,$$

en donde utilizamos el símbolo de igualdad  $\doteq$  para denotar explícitamente que esto es solamente una *representación* del vector  $\mathbf{v}$  en una base particular.

¿Qué significa geométricamente que una base tenga un orden? Veamos un ejemplo en  $\mathbb{R}^2$ : pensemos en el vector de la Figura 7, inciso a). Ya que hemos visto que los vectores ortonormales son excelentes

<sup>35</sup>Claramente, para poder ser biyectivo, este mapeo debe ser compatible con las operaciones de las estructuras algebraicas en cuestión. Por ejemplo, cualquier biyección entre dos espacios vectoriales necesariamente debe ser compatible con la suma de vectores y el producto de un vector por un escalar, i.e., debe ser una transformación lineal (¿podrías demostrar esto?).

candidatos para formar bases, nos gustaría expresar algebraicamente al vector mostrado en **a)** como combinación lineal de dos vectores ortogonales de norma uno. Dado nuestro trabajo previo en el plano cartesiano (y que sus ejes son perpendiculares), lo más natural resulta tomar a los vectores normales que apuntan en el sentido positivo en cada uno de los ejes. Sin embargo, si queremos representar a este vector como un par *ordenado* de números reales (correspondientes a los coeficientes de cada elemento de la base mediante los cuales se obtiene a dicho vector como combinación lineal) surge una ambigüedad: ¿a cuál de los elementos de la base corresponderá el primer número y a cuál elemento, el segundo? Siguiendo las convenciones usuales de geometría analítica, nos inclinamos por tomar a nuestra base ordenada como  $(\mathbf{e}_1, \mathbf{e}_2)$ , donde  $\mathbf{e}_1$  denota al vector ortonormal sobre el eje horizontal y en sentido positivo. De hecho, ¡hemos estado utilizando esta base implícitamente desde la sección 3.3! El resultado de esta elección de base ordenada se muestra en la Figura 7, inciso b).

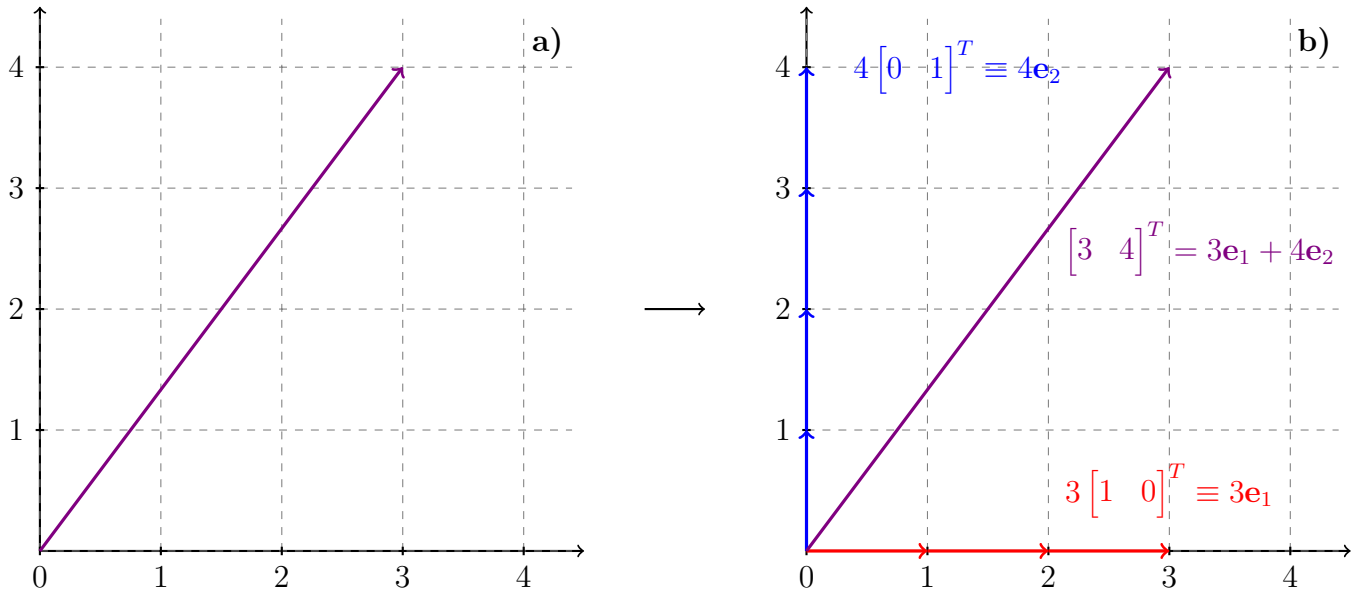


Figura 7: Interpretación geométrica de una base ordenada. En el inciso **a)** se muestra a un vector de  $\mathbb{R}^2$  que puede ser representado en una infinidad de bases distintas de  $\mathbb{R}^2$ . En el inciso **b)**, se elige a la base canónica *ordenada*  $(\mathbf{e}_1, \mathbf{e}_2)$ . Utilizando la notación de 2-tuplas para denotar vectores, este vector puede representarse en esta base ordenada como  $(3, 4)$ .

El hecho de que podamos representar a cualquier vector de  $\mathbb{R}^2$  como una combinación lineal de la base ordenada  $(\mathbf{e}_1, \mathbf{e}_2)$  lo podemos interpretar geoméricamente como que podemos *descomponer* cualquier vector *en sus componentes sobre cada eje* del plano cartesiano —cada uno de los cuáles es un múltiplo de un elemento de la base antes mencionada.

Extendiendo esta idea, en  $\mathbb{R}^3$  la *base canónica*  $(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$  es una base ordenada y ortonormal<sup>36</sup>. Más generalmente, para los espacios vectoriales de  $n$ -tuplas reales o complejas  $K^n$ , la  $n$ -tupla  $(\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n)$  es la base ordenada usual, ya que el hecho de que sea ortonormal resulta de gran utilidad. Cualquier  $\mathbf{e}_i$  de esta base ordenada se representa *en esta misma base* como una  $n$ -tupla que tiene un 1 en la  $i$ -ésima entrada y un 0 en las demás entradas (¿por qué?).

<sup>36</sup>En libros de texto de Física, a veces esta misma base canónica ordenada de  $\mathbb{R}^3$  se denota como  $(\hat{i}, \hat{j}, \hat{k})$ .



## 8.5. Representación matricial de una transformación lineal

Supongamos que tenemos un vector  $\mathbf{v} = \begin{bmatrix} v_1 & v_2 & \dots & v_n \end{bmatrix}^T \in K^n$  representado en una base ordenada arbitraria  $(\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_n)$  y sea  $T : K^n \rightarrow K^n$  una transformación lineal tal que

$$T\left(\begin{bmatrix} v_1 & v_2 & \dots & v_n \end{bmatrix}^T\right) = \begin{bmatrix} w_1 & w_2 & \dots & w_n \end{bmatrix}^T,$$

donde el segundo vector está representado en la misma base. Ya que  $T$  es lineal, tenemos que la ecuación anterior se cumple si y sólo si

$$T(v_1\mathbf{b}_1 + v_2\mathbf{b}_2 + \dots + v_n\mathbf{b}_n) = v_1T(\mathbf{b}_1) + v_2T(\mathbf{b}_2) + \dots + v_nT(\mathbf{b}_n) = w_1\mathbf{b}_1 + w_2\mathbf{b}_2 + \dots + w_n\mathbf{b}_n.$$

Es decir que, si conocemos cómo  $T$  actúa *sobre los elementos de la base*, podemos conocer cómo actúa *sobre cualquier vector arbitrario representado en esa misma base*.

Por ejemplo, sean  $T_{\hat{x}\hat{y}}, T_{\hat{y}\hat{z}}, T_{\hat{x}\hat{z}}$  las reflexiones por los planos  $\hat{x}\hat{y}, \hat{y}\hat{z}$  y  $\hat{x}\hat{z}$  de  $\mathbb{R}^3$ , respectivamente. Utilicemos la base canónica  $(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$  de  $\mathbb{R}^3$ ; por conveniencia, vamos a reetiquetarla como  $(\hat{x}, \hat{y}, \hat{z})$ . Una reflexión por el plano  $\hat{x}\hat{y}$  no le hace nada a los vectores  $\hat{x}$  y  $\hat{y}$ , pero manda a  $\hat{z}$  a su inverso aditivo. Más generalmente, podemos ver que

$$T_{\hat{a}\hat{b}}(\hat{a}) = \hat{a}, \quad T_{\hat{a}\hat{b}}(\hat{b}) = \hat{b}, \quad T_{\hat{a}\hat{b}}(\hat{c}) = -\hat{c}, \quad \forall \hat{a}, \hat{b}, \hat{c} \in \{\hat{x}, \hat{y}, \hat{z}\}, \quad \hat{a} \neq \hat{b} \neq \hat{c}$$

Entonces, para cualquier vector arbitrario  $\mathbf{v} = c_1\hat{a} + c_2\hat{b} + c_3\hat{c} \in \mathbb{R}^3$  con  $\hat{a}, \hat{b}, \hat{c} \in \{\hat{x}, \hat{y}, \hat{z}\}$ ,  $\hat{a} \neq \hat{b} \neq \hat{c}$  tenemos que

$$T_{\hat{a}\hat{b}}(\mathbf{v}) = T_{\hat{a}\hat{b}}(c_1\hat{a} + c_2\hat{b} + c_3\hat{c}) = c_1T_{\hat{a}\hat{b}}(\hat{a}) + c_2T_{\hat{a}\hat{b}}(\hat{b}) + c_3T_{\hat{a}\hat{b}}(\hat{c}) = c_1(\hat{a}) + c_2(\hat{b}) + c_3(-\hat{c}) = c_1\hat{a} + c_2\hat{b} - c_3\hat{c},$$

lo cual coincide con lo visto en la sección 7.1.2. Esta misma idea se puede aplicar al resto de las reflexiones de  $\mathbb{R}^3$  (con respecto al origen y cada uno de los ejes).

Ahora, sea  $\mathbf{v} = \begin{bmatrix} v_1 & v_2 & v_3 \end{bmatrix}^T \in \mathbb{R}^3$  representado en la base canónica y sea  $M_{\hat{x}\hat{y}} \in M_{3 \times 3}(\mathbb{R})$  una matriz de tres por tres con entradas reales dada por

$$\begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix}.$$

Queremos que al multiplicar a  $\mathbf{v}$  por  $M_{\hat{x}\hat{y}}$  obtengamos al vector  $\begin{bmatrix} v_1 & v_2 & -v_3 \end{bmatrix}^T$ , es decir, que esta matriz refleje a  $\mathbf{v}$  con respecto al plano  $\hat{x}\hat{y}$ . Resolviendo el sistema de ecuaciones correspondiente, vemos que

$$M_{\hat{x}\hat{y}}\mathbf{v} = T_{\hat{x}\hat{y}}(\mathbf{v}) \iff \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = \begin{bmatrix} v_1 \\ v_2 \\ -v_3 \end{bmatrix} \iff M_{\hat{x}\hat{y}} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}.$$

Evidentemente,  $M_{\hat{x}\hat{y}}$  no es más que la representación matricial de  $T_{\hat{x}\hat{y}}$  en la base canónica. De manera similar podemos obtener las representaciones matriciales de las demás reflexiones en  $\mathbb{R}^3$ .

Ahora viene lo realmente interesante: sea  $M_{ab}$  la entrada en el renglón  $a$  y la columna  $b$  de la matriz  $M_{\hat{x}\hat{y}}$ , entonces podemos observar que la ecuación

$$M_{ij} = (T_{\hat{x}\hat{y}}(\mathbf{e}_j), \mathbf{e}_i)$$

se verifica para todas las entradas de la matriz  $M_{\hat{x}\hat{y}}$ <sup>37</sup>.

<sup>37</sup>¡Asegúrate de que esto es cierto! Este hecho no es una mera coincidencia, como veremos a continuación, pero quedará más claro si lo verificas primero.

Ahora, vamos a generalizar esta idea: sea  $V$  sobre  $K$  un espacio vectorial con una base ortogonal ordenada<sup>38</sup>  $(\mathbf{o}_1, \mathbf{o}_2, \dots, \mathbf{o}_n)$  y  $T$  una transformación lineal de  $V$  a  $V$ . Entonces las entradas  $A_{ij} \in K$  de la matriz de  $n \times n$  asociada a  $T$  representada en esta misma base están dadas por

$$A_{ij} = \frac{(T(\mathbf{o}_j), \mathbf{o}_i)}{\|\mathbf{o}_i\|^2}.$$

Veamos de dónde viene esto: ya que  $(\mathbf{o}_1, \mathbf{o}_2, \dots, \mathbf{o}_n)$  es una base ordenada de  $V$ , entonces  $V$  tiene dimensión  $n$ . Sea  $\mathbf{v} = \begin{bmatrix} v_1 & v_2 & \dots & v_n \end{bmatrix}^T \in V$  un vector de  $n$  entradas y  $A \in M_{n \times n}(K)$  una matriz de  $n \times n$  entradas escalares  $A_{ij}$ . Pensemos ahora en la ecuación  $A\mathbf{v} = T(\mathbf{v})$ . Por un lado, sabemos que podemos aplicar las propiedades lineales de  $T$  y reescribir  $T(\mathbf{v})$  como  $v_1T(\mathbf{o}_1) + v_2T(\mathbf{o}_2) + \dots + v_nT(\mathbf{o}_n)$ . Por el otro lado sabemos que, al multiplicar al vector  $\mathbf{v}$  por la matriz  $A$ , la primera entrada del vector resultante será igual a hacer un producto entrada por entrada entre el vector columna  $\mathbf{v}$  y el primer renglón de la matriz  $A$ , y después sumar todos los términos. Más generalmente, la  $i$ -ésima entrada del vector resultante  $A\mathbf{v}$  será igual a la suma  $\sum_{j=1}^n A_{ij}v_j$ , por lo cual podemos escribir explícitamente al vector resultante en términos de la base ortogonal como  $\sum_{i=1}^n \sum_{j=1}^n A_{ij}v_j\mathbf{o}_i$ . Hasta ahora, tenemos que

$$A\mathbf{v} = T(\mathbf{v}) \iff \sum_{i=1}^n \sum_{j=1}^n A_{ij}v_j\mathbf{o}_i = \sum_{j=1}^n v_jT(\mathbf{o}_j).$$

¿Qué podemos hacer ahora para igualar las expresiones de ambos lados de manera astuta? En general  $T$  podría actuar sobre elementos de la base  $\mathbf{o}_i$  cambiándolos a otros elementos de la base  $\mathbf{o}_j, j \neq i$  (e.g., mediante rotaciones), así que no sabemos cuáles términos de la suma del lado derecho de la ecuación terminarán acompañados del elemento de la base  $\mathbf{o}_i$ , de tal manera que podamos igualar la ecuación término a término apoyándonos en el índice  $i$ . Aquí viene la ventaja de trabajar con bases ortogonales para representar transformaciones lineales como matrices: podemos simplemente tomar el producto escalar con  $\mathbf{o}_i$  de ambos lados de la ecuación<sup>39</sup>. Observemos qué sucede en este caso:

$$\begin{aligned} A\mathbf{v} = T(\mathbf{v}) &\iff \left( \sum_{i=1}^n \sum_{j=1}^n A_{ij}v_j\mathbf{o}_i, \mathbf{o}_i \right) = \left( \sum_{j=1}^n v_jT(\mathbf{o}_j), \mathbf{o}_i \right) \iff \sum_{i=1}^n \sum_{j=1}^n A_{ij}v_j(\mathbf{o}_i, \mathbf{o}_i) = \sum_{i=j}^n v_j(T(\mathbf{o}_j), \mathbf{o}_i) \\ &\iff \sum_{i=1}^n \sum_{j=1}^n A_{ij}v_j\|\mathbf{o}_i\|^2 = \sum_{j=1}^n v_j(T(\mathbf{o}_j), \mathbf{o}_i). \end{aligned}$$

Ahora sólo hace falta observar que

$$\sum_{i=1}^n \sum_{j=1}^n A_{ij}v_j\|\mathbf{o}_i\|^2 = \sum_{i=1}^n \sum_{j=1}^n A_{ij}(v_1 + v_2 + \dots + v_n)\|\mathbf{o}_i\|^2 = (v_1 + v_2 + \dots + v_n) \sum_{i=1}^n \sum_{j=1}^n A_{ij}\|\mathbf{o}_i\|^2$$

y

$$\sum_{j=1}^n v_j(T(\mathbf{o}_j), \mathbf{o}_i) = \sum_{j=1}^n (v_1 + v_2 + \dots + v_n)(T(\mathbf{o}_j), \mathbf{o}_i) = (v_1 + v_2 + \dots + v_n) \sum_{j=1}^n (T(\mathbf{o}_j), \mathbf{o}_i),$$

por lo cual, asumiendo que  $\mathbf{v}$  es no nulo<sup>40</sup>, podemos dividir cada lado entre  $v_1 + v_2 + \dots + v_n$  y obtener

$$A\mathbf{v} = T(\mathbf{v}) \iff \sum_{i=1}^n \sum_{j=1}^n A_{ij}\|\mathbf{o}_i\|^2 = \sum_{j=1}^n (T(\mathbf{o}_j), \mathbf{o}_i).$$

Finalmente, si fijamos  $i$  y  $j$  en la expresión anterior, obtenemos la ecuación inicial.

<sup>38</sup>Recordemos que, ya que esta base es ortogonal, entonces  $(\mathbf{o}_i, \mathbf{o}_j) = 0$  si  $i \neq j$ .

<sup>39</sup>Al hacer esto, estamos fijando nuestra atención sólo en la  $i$ -ésima entrada del vector resultante  $A\mathbf{v}$  (ó  $T(\mathbf{v})$ ), por lo cual el resto del desarrollo se debe cumplir para cada  $i \in \{1, 2, \dots, n\}$ .

<sup>40</sup>Si  $\mathbf{v}$  es el vector nulo, entonces la igualdad se cumple trivialmente, por lo cual ese caso ya está cubierto.

## 8.6. Ejercicios de repaso

### 8.6.1. Composición de transformaciones lineales

1. Sean  $T_{\theta_1}, T_{\theta_2}$  rotaciones en  $\mathbb{R}^2$ . Demuestra que  $T_{\theta_1} \circ T_{\theta_2} = T_{\theta_1+\theta_2} = T_{\theta_2+\theta_1} = T_{\theta_2} \circ T_{\theta_1}$ , es decir, que la composición de dos rotaciones por ángulos  $\theta_1$  y  $\theta_2$  es igual a una sola rotación por un ángulo  $\theta_1 + \theta_2$ , sin importar el orden en que compongas las rotaciones originales (Nota: repasa identidades trigonométricas). (1.25 pts.)
2. De la sección 7.1.2 podemos ver que componer tres reflexiones de  $\mathbb{R}^3$  con respecto a los tres planos distintos resulta en una reflexión por el origen. ¿Habrá alguna manera de componer las reflexiones  $T_{\hat{x}}, T_{\hat{y}}$  y  $T_{\hat{z}}$  para obtener a  $T_O$ ? (1.25 pts.)

### 8.6.2. Transformación lineal inversa

1. Sea  $T : V \rightarrow W$  una transformación lineal y  $L : W \rightarrow V$  una función vectorial tal que  $L \circ T = I_V$  y  $T \circ L = I_W$ . Demuestra que  $L$  es forzosamente una transformación lineal. (1.25 pts.)
2. Sea  $T_\theta : \mathbb{R}^2 \rightarrow \mathbb{R}^2$  una rotación por un ángulo  $\theta \in (-\infty, \infty)$ . Demuestra que existen una infinidad de transformaciones lineales para  $T_\theta$ . ¿Cómo podríamos deshacernos de esta redundancia? (1.25 pts.)

### 8.6.3. Espacios vectoriales isomorfos

1. Demuestra que los espacios vectoriales  $\mathbb{R}^n$  y  $P^{n-1}$ ,  $n \in \mathbb{N}$  son isomorfos. (1.25 pts. extra)
2. Demuestra que la relación de isomorfismo entre espacios vectoriales sobre un campo  $K$  es una relación de equivalencia (i.e., que es reflexiva, simétrica y transitiva) (1 pto. extra)\*

### 8.6.4. Representación de un vector en una base ordenada

1. Pensando en la Figura 7, imaginemos que tomamos cualesquiera otros dos vectores ortonormales de  $\mathbb{R}^2$  que llamaremos  $\mathbf{f}_1, \mathbf{f}_2$ , distintos de  $\mathbf{e}_1$  y  $\mathbf{e}_2$ , respectivamente. Ahora, tomemos un vector no nulo  $\mathbf{v} \in \mathbb{R}^2$ . ¿Cómo se relacionarían las representaciones del vector  $\mathbf{v}$  en las bases ordenadas  $(\mathbf{e}_1, \mathbf{e}_2)$  y  $(\mathbf{f}_1, \mathbf{f}_2)$ ? ¿Con qué transformación lineal podríamos describir esta relación? (1.25 pts.)
2. Siguiendo el ejercicio anterior, representa a la matriz asociada a esta transformación lineal en ambas bases. (1 pto. extra)\*

### 8.6.5. Representación matricial de una transformación lineal

1. Supón que tenemos una transformación lineal incógnita  $X : V \rightarrow W$  de la cual no conocemos su regla de correspondencia general, pero que sabemos que actúa sobre la base ortogonal ordenada  $(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3)$  de  $V$  de la manera siguiente:

$$X(\mathbf{v}_1) = \mathbf{v}_1, X(\mathbf{v}_2) = \mathbf{v}_2, X(\mathbf{v}_3) = -\mathbf{v}_3.$$

Además, supón que  $(\mathbf{v}_1, \mathbf{v}_2) = a$ ,  $(\mathbf{v}_2, \mathbf{v}_3) = b$  y  $(\mathbf{v}_1, \mathbf{v}_3) = c$ , con  $a, b, c \in \mathbb{R}$ . Encuentra la representación de  $X$  en las bases ordenadas  $(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3)$ ,  $(\mathbf{v}_2, \mathbf{v}_1, \mathbf{v}_3)$  y  $(\mathbf{v}_3, \mathbf{v}_2, \mathbf{v}_1)$ . (1.25 pts.)

2. Repite el ejercicio anterior, pero ahora suponiendo que  $X$  va de  $\mathbb{R}^3$  a  $\mathbb{R}^3$  y que la base ordenada dada es la base canónica  $(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$ . Viendo la representación matricial de esta transformación lineal en la base canónica de  $\mathbb{R}^3$ , ¿cómo puedes interpretar geoméricamente esta transformación lineal? ¿Cómo sería la representación matricial de la transformación lineal inversa? (1.25 pts.)