

Curso de Álgebra Lineal (2021-I)

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

Notación

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

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

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

1.1. 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.

1.1.1. 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*^a. 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 .

^aNótese que la propiedad de distributividad es la única que *combina* ambas operaciones definidas en el campo.

1.1.2. 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 1.1.1, 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 $a \in \mathbb{R}$, el elemento inverso de la suma es $-a$, mientras que para todo $b \neq 0 \in \mathbb{R}$, el elemento inverso de la multiplicación es $b^{-1} = \frac{1}{b} \in \mathbb{R}$. Al conjunto \mathbb{R} junto con estas dos operaciones se le conoce como el *campo de los números reales* o, simplemente, el *campo real*.

1.1.3. El campo complejo

El conjunto de los números complejos se define como $\mathbb{C} \equiv \{a + ib \mid a, b \in \mathbb{R}\}$, 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 como¹ $(a + ib)(q + ir) \equiv (aq - br) + i(ar + bq)$ 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*². 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 1.1.1.

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

$$(a + ib) + ((q + ir) + (s + it)) = (a + ib) + ((q + s) + i(r + t)) = (a + q + s) + i(b + r + t) =$$

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

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

¹Nótese que esta definición es simplemente el resultado de desarrollar $(a + ib)(c + id)$ como un producto de binomios.

²El 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=T647CGsuOVU>.

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!

1.2. 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.

1.2.1. Definición de espacio vectorial

Def. Un *espacio vectorial* sobre un campo^a 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

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

Elemento identidad de la adición (neutro aditivo)

$$\forall \mathbf{v} \in V \quad \exists -\mathbf{v} \in V \text{ t.q. } \mathbf{v} + (-\mathbf{v}) = \mathbf{0}$$

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

$$\exists 1 \in K \text{ t.q. } 1\mathbf{v} = \mathbf{v} \quad \forall \mathbf{v} \in V$$

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*.

^aVer sec. 1.1.1.

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)³. 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.

1.2.2. 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 1.2.1. 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 notación⁴ $\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in \mathbb{R}^2$, $x_1, x_2 \in \mathbb{R}$ (lo cual también puede escribirse como $\begin{pmatrix} x_1 & x_2 \end{pmatrix}^T$ ó, simplemente, $\begin{pmatrix} x_1 & x_2 \end{pmatrix}$). 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 $\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} x_1 + y_1 \\ x_2 + y_2 \end{pmatrix}$. El elemento identidad de la suma (neutro aditivo) es el *vector origen* $\mathbf{0} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ y el inverso aditivo de un vector $\mathbf{u} = \begin{pmatrix} u_1 \\ u_2 \end{pmatrix}$ es $\begin{pmatrix} -u_1 \\ -u_2 \end{pmatrix}$, denotado como $-\mathbf{u}$. En este caso, el producto de un vector $\mathbf{v} = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} \in \mathbb{R}^2$ por un escalar $a \in \mathbb{R}$ se define como $a\mathbf{v} = a \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} \equiv \begin{pmatrix} av_1 \\ av_2 \end{pmatrix}$ (ó $a \begin{pmatrix} v_1 & v_2 \end{pmatrix} \equiv \begin{pmatrix} av_1 & av_2 \end{pmatrix}$) 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.

³Más 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.

⁴A esto se le conoce como *escribir un vector como una columna de matriz* 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 $\mathbf{v} = (v_1 \ v_2 \ \dots \ v_n)$, $v_i \in \mathbb{R}$, $i \in \{1, 2, \dots, n\}$ y son conocidos como *n-tuplas*⁵. 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 vectorial⁶. 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 $(a + ib \ c + id)$ 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 $\mathbf{0} = (0 + i0 \ 0 + i0)$ 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*⁷. 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.

⁵Nótese 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.

⁶En general, el conjunto de funciones de clase C^n sobre el campo \mathbb{R} forma un espacio vectorial.

⁷Es posible visualizar vectores más abstractos de otras formas: <https://www.youtube.com/watch?v=zwAD6dRSvYI>.

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

1.2.3. 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, $\mathbf{0} = (0)\mathbf{v} = (a + (-a))\mathbf{v} = a\mathbf{v} + (-a)\mathbf{v}$ ⁸, 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

⁸La existencia de $-a \in K$ está asegurada ya que K es un campo (ver sec. 1.1.1).

1.3. Interpretación geométrica de las operaciones esenciales de los espacios vectoriales

Como se mencionó en una nota al final de la sección 1.2.2 —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 $(a, b) \neq (b, a) \iff b \neq a^a$.

^aObservemos 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.

1.3.1. 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 reales⁹ 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 $\begin{pmatrix} a & b \end{pmatrix} \in \mathbb{R}^2$ 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 1.

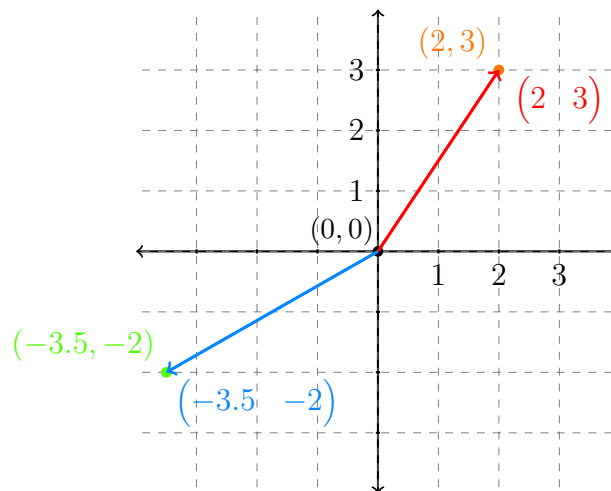


Figura 1: Ejemplo de representación de pares ordenados y vectores en el plano cartesiano. Los pares ordenados $(-3.5, -2)$ y $(2, 3)$ se representan mediante puntos que corresponden precisamente a las coordenadas cartesianas $(-3.5, -2)$ y $(2, 3)$, respectivamente, mientras que los vectores $\begin{pmatrix} -3.5 & -2 \end{pmatrix}$ y $\begin{pmatrix} 2 & 3 \end{pmatrix}$ 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.

⁹Es decir, con $a, b \in \mathbb{R}$.

Suma vectorial

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

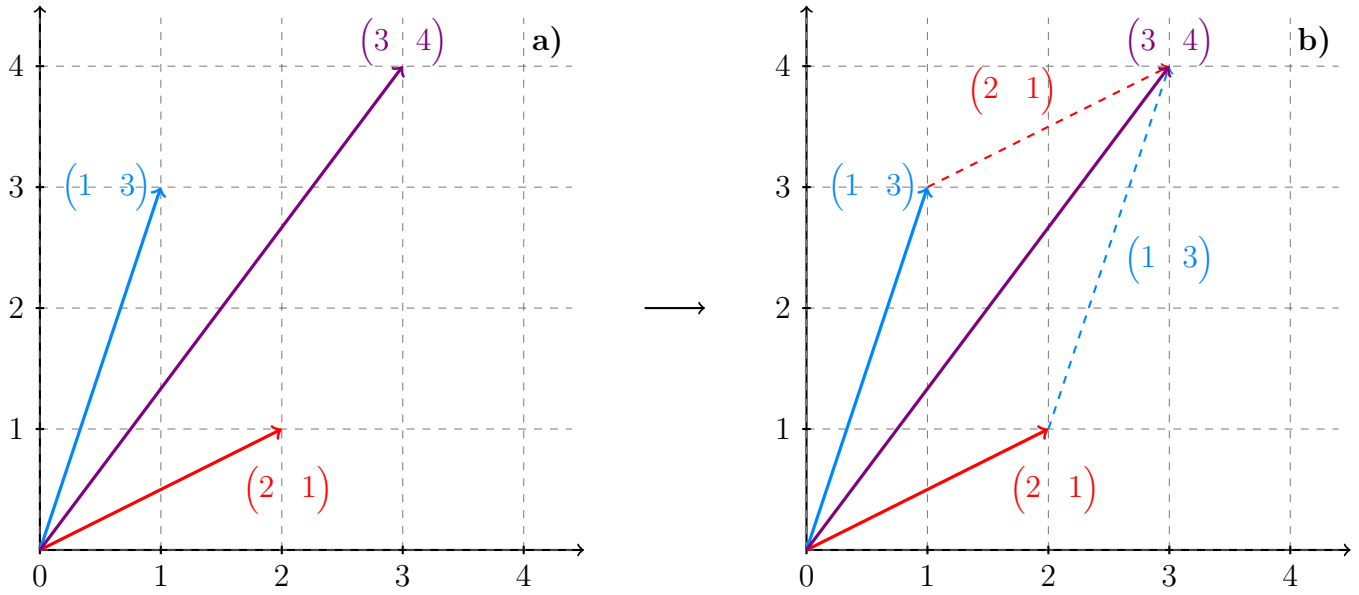


Figura 2: Interpretación geométrica de la suma vectorial en el espacio vectorial real \mathbb{R}^2 . En la subfigura a) se observan los vectores $\begin{pmatrix} 1 & 3 \end{pmatrix}$ y $\begin{pmatrix} 2 & 1 \end{pmatrix}$, así como el vector resultante de la suma de los dos anteriores, $\begin{pmatrix} 3 & 4 \end{pmatrix}$. En la subfigura 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 $c \begin{pmatrix} a & b \end{pmatrix} \equiv \begin{pmatrix} ca & cb \end{pmatrix}$. Podemos calcular, por ejemplo, los productos $\left(\frac{1}{2}\right) \begin{pmatrix} 2 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 1 \end{pmatrix}$ y $(-1.2) \begin{pmatrix} 1 & 3 \end{pmatrix} = \begin{pmatrix} -1.2 & -3.6 \end{pmatrix}$. 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 vector automáticamente se convierte en el vector nulo $\begin{pmatrix} 0 & 0 \end{pmatrix}$, como se demostró algebraicamente en el

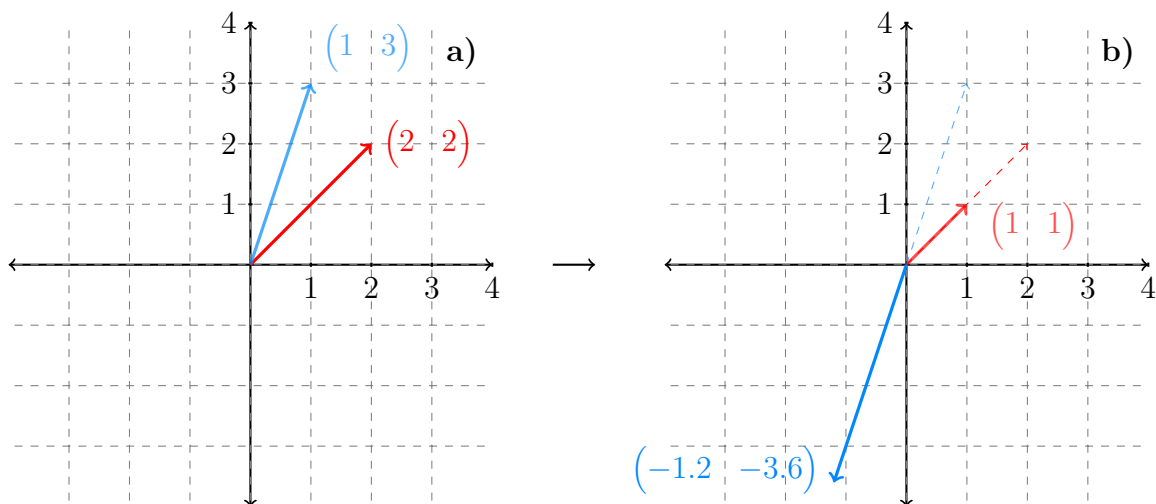


Figura 3: Interpretación geométrica del producto de un vector por un escalar en el espacio vectorial real \mathbb{R}^2 . Comparando las subfiguras **a)** y **b)** observamos que, en caso de que se multiplique a un vector de \mathbb{R}^2 por un escalar de \mathbb{R} , es posible que la longitud del vector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

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.

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

Nota: El vector nulo $\mathbf{0} = (0 \ 0)$ (también llamado *vector origen*) no tiene longitud, ya que es el único donde la cola y la punta de su flecha coinciden. Además, se dice que tampoco tiene dirección ni sentido¹⁰. 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.

1.3.2. 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.

1.3.3. 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 eje perpendicular a él de números *imaginarios*¹¹— 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

¹⁰Alternativamente, 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.

¹¹Los 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.

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

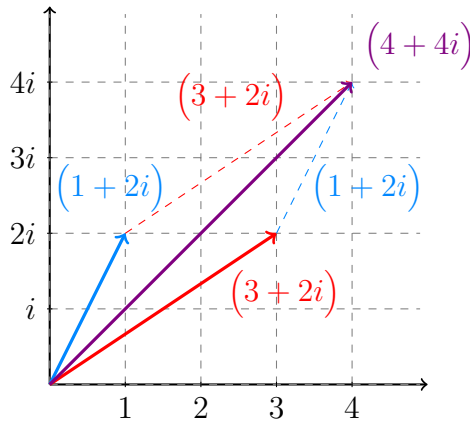


Figura 4: Interpretación geométrica de la suma vectorial en el espacio vectorial complejo \mathbb{C} . Observamos que, al igual que en el caso del espacio vectorial real \mathbb{R}^2 , se cumple la *Ley del paralelogramo*.

Producto de un vector por un escalar

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

Notemos que, en particular, si la parte imaginaria del escalar es nula (i.e., si $r = 0$), entonces el escalar es un número real y el producto resultante es $(q)(s + it) \equiv ((qs) + (qt)i)$, por lo cual geométricamente sólo se produce un reescalamiento totalmente análogo al discutido en la sección 1.3.1.

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 $(2 + 2i)$. Al hacer el producto de este vector por i obtenemos $(-2 + 2i)$. Si, en cambio, hacemos el producto de este mismo vector por el escalar $-i$, obtenemos como resultado $(-i)(2 + 2i) = (2 - 2i)$. Ambas operaciones se muestran de manera visual en la Figura 5.

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

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

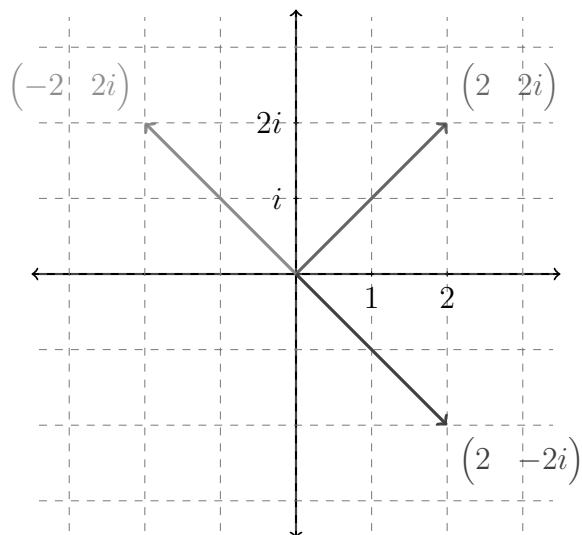


Figura 5: Interpretación geométrica del producto de un vector complejo por los números imaginarios i y $-i$. En este caso, nuestro vector base es $(2 \ 2i)$. El producto de este vector por el escalar i resulta en el vector $(-2 \ 2i)$, lo cual puede ser interpretado geoméricamente como una rotación discreta de $\frac{\pi}{2}$ radianes. Así observamos que, en cambio, el producto de nuestro vector base $(2 \ 2i)$ por $-i$ se puede interpretar geoméricamente como una rotación discreta de $-\frac{\pi}{2}$ radianes.

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 $(s + it)$ 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*¹².

¹²El asunto de las *magnitudes específicas* de estos reescalamientos y rotaciones —el cual se complica para escalares complejos en general— será precisado más adelante.

1.4. 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.

1.4.1. 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 subsecuentes¹³.
- 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 ¹⁴.

1.4.2. Ejemplos de subespacios vectoriales

El conjunto de todos los pares ordenados $\{(x_1 \ x_2) \mid x_1, x_2 \in \mathbb{R} \wedge x_1 = x_2\}$ 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 vectorial¹⁵ del espacio vectorial formado por el conjunto de polinomios de grado k .

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

¹³Esto es más o menos similar al hecho de que cualquier subconjunto puede tener subconjuntos subsecuentes.

¹⁴Aquí 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.

¹⁵De 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.

1.4.3. Algunos teoremas de subespacios vectoriales

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

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 1.4.1, 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 $B \subset C \implies A \subset C$.

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

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 $S_1 \oplus S_2 = V$.

^aAquí 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.

1.5. Ejercicios de repaso

1.5.1. Campos

1. Demuestra que el conjunto \mathbb{C} junto con las operaciones de suma y multiplicación definidas en la sección 1.1.3 cumplen las últimas tres propiedades de la definición de campo de la sección 1.1.1 para finalmente demostrar que estas tres cosas juntas (el conjunto \mathbb{C} y las dos operaciones mencionadas) forman un campo.

1.5.2. 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).
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} .
3. Sea K un campo arbitrario. Demuestra que K^n sobre K es un espacio vectorial, con $n \in \mathbb{N}$.

1.5.3. 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 $\begin{pmatrix} a & b \end{pmatrix} + c \begin{pmatrix} d & e \end{pmatrix}$ y muestra gráficamente al vector resultante aplicando la Ley del paralelogramo.
2. Como vimos en la sección 1.3, 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 $\mathbf{v} - \mathbf{u}$.
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 $\frac{1}{2}(\vec{u} + \vec{v})$.
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 $(q + ir) \in \mathbb{C}$ 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 $-s - it$.

1.5.4. 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.
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 ?

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 $n \geq m$.
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^m con $i > n$ sobre el mismo campo \mathbb{R} ? Argumenta.
5. Demuestra el Teorema 1.4.3.2 (ver sec. 1.4.3).

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

2.1. Producto escalar (punto)

Algunos espacios vectoriales están dotados de una operación entre vectores que da como resultado un escalar conocida como *producto escalar*. En estos espacios, dicha operación tiene una gran riqueza geométrica, la cual nos permite tener nociones como ortogonalidad y magnitud —y generalizarlas a espacios vectoriales más abstractos, como veremos a lo largo de esta sección¹⁶.

2.1.1. 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* $\langle \cdot, \cdot \rangle : V \times V \rightarrow K$ le asocia a cualquier par ordenado de vectores en V un escalar en K , y cumple las tres propiedades siguientes:

$$\langle \mathbf{u} + \mathbf{w}, \mathbf{v} \rangle = \langle \mathbf{u}, \mathbf{v} \rangle + \langle \mathbf{w}, \mathbf{v} \rangle,$$

$$\langle a\mathbf{u}, \mathbf{v} \rangle = a\langle \mathbf{u}, \mathbf{v} \rangle,$$

$$\langle \mathbf{u}, \mathbf{v} \rangle = \overline{\langle \mathbf{v}, \mathbf{u} \rangle},$$

donde la barra $\overline{\langle \mathbf{v}, \mathbf{u} \rangle}$ denota el complejo conjugado de $\langle \mathbf{v}, \mathbf{u} \rangle$ ^a. Si, además, se cumple la propiedad

$$\forall \mathbf{v} \in V, \langle \mathbf{v}, \mathbf{v} \rangle \geq 0; \langle \mathbf{v}, \mathbf{v} \rangle = 0 \iff \mathbf{v} = \mathbf{0},$$

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 (\mathbf{u}, \mathbf{v}) ^b.

^aRecordemos 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 $c \in \mathbb{C}$.

^bLa notación principal que utilizaremos para el producto escalar $\langle \cdot, \cdot \rangle$ 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 $\langle \mathbf{u} + a\mathbf{w}, \mathbf{v} \rangle = \langle \mathbf{u}, \mathbf{v} \rangle + a\langle \mathbf{w}, \mathbf{v} \rangle$. Para referirnos a esta propiedad específica en lenguaje matemático, decimos que el producto escalar es una operación *lineal*¹⁷ 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 $K = \mathbb{R} \implies \langle \mathbf{u}, \mathbf{v} \rangle = \langle \mathbf{v}, \mathbf{u} \rangle$. En cambio, si el campo es complejo, en

¹⁶En el último módulo del curso regresaremos a este tipo de espacios para descubrir más secretos del producto escalar.

¹⁷Como 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.

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 $\langle \mathbf{u}, \mathbf{v} + a\mathbf{w} \rangle = \langle \mathbf{u}, \mathbf{v} \rangle + \bar{a}\langle \mathbf{u}, \mathbf{w} \rangle$, donde \bar{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*¹⁸ en la segunda entrada. En cambio, si el campo es real, entonces el producto escalar es lineal en ambas entradas.
- 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 1.2.1 y 1.4.1).

2.1.2. Ejemplos de producto escalar en espacios vectoriales

En \mathbb{R}^2 el producto escalar se define como $\begin{pmatrix} u_1 \\ u_2 \end{pmatrix} \cdot \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} \equiv u_1v_1 + u_2v_2$, 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

$$\mathbf{u} \cdot \mathbf{v} = \begin{pmatrix} u_1 & u_2 & \dots & u_n \end{pmatrix} \cdot \begin{pmatrix} v_1 & v_2 & \dots & v_n \end{pmatrix} \equiv u_1v_1 + u_2v_2 + \dots + u_nv_n = \sum_{i=1}^n u_iv_i^{19}.$$

Para cumplir *todas* las propiedades descritas en la sección 2.1.1, 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

$$\langle \mathbf{a}, \mathbf{b} \rangle \equiv \mathbf{a} \cdot \mathbf{b} \equiv \sum_{i=1}^n a_i \bar{b}_i,$$

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

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

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 -tuplas²⁰. 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), *Linear Algebra: A Modern Introduction* de Poole²¹ (págs. 531-534), etc.

¹⁸Este concepto también se verá con mayor detenimiento más adelante.

¹⁹Para simplificar la notación, en varias áreas de la física se elimina el signo de suma (Σ) 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*.

²⁰Veremos otros tipos de espacios vectoriales donde se pueden definir productos escalares más adelante.

²¹Ten 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.

2.1.3. Propiedades del producto escalar

Como resumen, a partir de la definición dada en la sección 2.1.1, podemos ver que las principales propiedades del producto escalar son:

$$\begin{array}{ll} \langle \mathbf{u}, \mathbf{v} \rangle = \overline{\langle \mathbf{v}, \mathbf{u} \rangle} & \text{Conmutar vectores resulta en la conjugación del escalar} \\ \langle a\mathbf{u} + \mathbf{v}, \mathbf{w} \rangle = a\langle \mathbf{u}, \mathbf{w} \rangle + \langle \mathbf{v}, \mathbf{w} \rangle & \text{Linealidad en la primera entrada} \\ \langle \mathbf{u}, \mathbf{w} + b\mathbf{z} \rangle = \langle \mathbf{u}, \mathbf{w} \rangle + \bar{b}\langle \mathbf{u}, \mathbf{z} \rangle & \text{Antilinealidad en la segunda entrada} \end{array}$$

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

$$\langle a_1\mathbf{u}_1 + \dots + a_n\mathbf{u}_n, \mathbf{v} \rangle = a_1\langle \mathbf{u}_1, \mathbf{v} \rangle + \dots + a_n\langle \mathbf{u}_n, \mathbf{v} \rangle$$

y

$$\langle \mathbf{u}, b_1\mathbf{v}_1 + \dots + b_n\mathbf{v}_n \rangle = b_1\langle \mathbf{u}, \mathbf{v}_1 \rangle + \dots + b_n\langle \mathbf{u}, \mathbf{v}_n \rangle.$$

2.2. Norma

Otro tipo de operación que se puede definir sobre un espacio vectorial es la *norma*, que toma un sólo vector del espacio y devuelve un escalar del campo. La norma, en general, dota a los vectores de un espacio de un cierto sentido de *magnitud* mediante el cual se puede comparar a distintos vectores. Como veremos en esta sección, si un espacio vectorial tiene un producto escalar positivo definido (dotando al espacio de las nociones de proyecciones y ortogonalidad que ya vimos en la sec. 2.3), entonces se puede definir una norma sobre el espacio a partir de dicho producto escalar; sin embargo, no es necesario que una norma provenga de un producto escalar, y es posible que existan normas (y, por tanto, ciertas nociones de magnitud) en espacios vectoriales *sin* producto escalar.

2.2.1. 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 podemos demostrar que la norma es positivo definida; es decir que, además de que el único vector con norma igual a 0 es el vector nulo, todos los vectores no nulos tienen norma positiva.

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. □

2.2.2. 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 2.2.1 (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

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

A ésta se le conoce como la *norma del valor absoluto*²². En el caso real, esta norma se interpreta geoméricamente como la distancia entre el origen de la recta real y el punto correspondiente al valor r 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 euclideana 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 comunmente utilizada se define como

$$||\mathbf{u}|| = +\sqrt{\langle \mathbf{u}, \mathbf{u} \rangle} = +\sqrt{\mathbf{u} \cdot \mathbf{u}},$$

de acuerdo a las definiciones de productos escalares para n -tuplas dadas en la sección 2.1.2. 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 euclideana entre estos dos puntos. Por ende, a esta norma se le conoce como *norma euclideana*²³. En particular, a los espacios vectoriales reales \mathbb{R}^n con esta norma se les conoce como *espacios vectoriales euclidianos* (o *euclídeos*)²⁴. 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

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

siguiendo el último ejemplo de la sección 2.1.2. Es fácil demostrar que esta definición cumple con las propiedades de norma (ver sec. 2.2.1; se siguen de las propiedades de la integral vistas en Cálculo II).

2.2.3. Desigualdad de Cauchy-Schwarz

Como mencionamos al inicio de esta sección, el proceso anterior de obtener una norma a partir de un producto escalar se puede generalizar para cualquier espacio vectorial con un producto escalar positivo definido, sin importar si dichos espacios vectoriales son reales o complejos. En estos casos, se dice que la norma es *inducida* por un producto escalar positivo definido. Las normas de este tipo serán las más recurrentes durante este curso. El primer paso necesario para demostrar este resultado general es la llamada *desigualdad de Cauchy-Schwarz*, que enunciamos a continuación como teorema. El resto de la demostración se encuentra en los ejercicios de repaso al final de esta sección (ver sec. 2.6.2).

Teorema 2.2.3.1. Sea $\langle \cdot, \cdot \rangle : V \times V \rightarrow K$ un producto escalar positivo definido y sea $||\cdot|| : V \rightarrow K$ una función definida tal que $||\cdot|| = +\sqrt{\langle \cdot, \cdot \rangle}$ ²⁵, entonces para todo $\mathbf{u}, \mathbf{v} \in V$ se cumple que $|\langle \mathbf{u}, \mathbf{v} \rangle| \leq ||\mathbf{u}|| ||\mathbf{v}||$.

²²En 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.

²³Observemos que la norma del valor absoluto es simplemente un caso particular de la norma euclideana para los espacios vectoriales \mathbb{R} y \mathbb{C} , tanto en su definición algebraica como en su interpretación geométrica.

²⁴Estos son los espacios vectoriales básicos que se utilizan en geometría analítica y cálculo diferencial e integral con funciones de una o más variables reales.

²⁵Con la ayuda de este teorema, se puede demostrar que esta función $||\cdot||$ cumple con todas las propiedades de una norma, dadas en la sección 2.2.1. Ésta es la llamada *norma euclideana generalizada*.

Demostración. Sean $\mathbf{u}, \mathbf{v} \in V$. Si \mathbf{v} o \mathbf{u} (o ambos) son iguales al vector nulo, la demostración se cumple trivialmente (ver el primer ejercicio de la subsec. 2.6.1). Supongamos que $\mathbf{u}, \mathbf{v} \neq \mathbf{0}$.

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

$$0 \leq \|\mathbf{u} - a\mathbf{v}\|^2 = \langle \mathbf{u} - a\mathbf{v}, \mathbf{u} - a\mathbf{v} \rangle = \langle \mathbf{u}, \mathbf{u} - a\mathbf{v} \rangle - a\langle \mathbf{v}, \mathbf{u} - a\mathbf{v} \rangle =$$

$$\langle \mathbf{u}, \mathbf{u} \rangle - \bar{a}\langle \mathbf{u}, \mathbf{v} \rangle - a(\langle \mathbf{v}, \mathbf{u} \rangle - \bar{a}\langle \mathbf{v}, \mathbf{v} \rangle) = \langle \mathbf{u}, \mathbf{u} \rangle - \bar{a}\langle \mathbf{u}, \mathbf{v} \rangle - a\langle \mathbf{v}, \mathbf{u} \rangle + a\bar{a}\langle \mathbf{v}, \mathbf{v} \rangle.$$

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

$$0 \leq \langle \mathbf{u}, \mathbf{u} \rangle - \frac{\langle \mathbf{v}, \mathbf{u} \rangle}{\langle \mathbf{v}, \mathbf{v} \rangle} \langle \mathbf{u}, \mathbf{v} \rangle - \frac{\langle \mathbf{u}, \mathbf{v} \rangle}{\langle \mathbf{v}, \mathbf{v} \rangle} \langle \mathbf{v}, \mathbf{u} \rangle + \frac{\langle \mathbf{u}, \mathbf{v} \rangle}{\langle \mathbf{v}, \mathbf{v} \rangle} \frac{\langle \mathbf{v}, \mathbf{u} \rangle}{\langle \mathbf{v}, \mathbf{v} \rangle} \langle \mathbf{v}, \mathbf{v} \rangle =$$

$$\langle \mathbf{u}, \mathbf{u} \rangle - 2 \frac{\langle \mathbf{u}, \mathbf{v} \rangle \langle \mathbf{v}, \mathbf{u} \rangle}{\langle \mathbf{v}, \mathbf{v} \rangle} + \frac{\langle \mathbf{u}, \mathbf{v} \rangle \langle \mathbf{v}, \mathbf{u} \rangle}{\langle \mathbf{v}, \mathbf{v} \rangle} = \|\mathbf{u}\|^2 - \frac{|\langle \mathbf{u}, \mathbf{v} \rangle|^2}{\|\mathbf{v}\|^2} \implies$$

$$0 \leq \|\mathbf{u}\|^2 - \frac{|\langle \mathbf{u}, \mathbf{v} \rangle|^2}{\|\mathbf{v}\|^2} \implies \frac{|\langle \mathbf{u}, \mathbf{v} \rangle|^2}{\|\mathbf{v}\|^2} \leq \|\mathbf{u}\|^2 \implies |\langle \mathbf{u}, \mathbf{v} \rangle|^2 \leq \|\mathbf{u}\|^2 \|\mathbf{v}\|^2 \implies |\langle \mathbf{u}, \mathbf{v} \rangle| \leq \|\mathbf{u}\| \|\mathbf{v}\|.$$

□

2.3. Interpretación geométrica del producto escalar: proyecciones y ortogonalidad

2.3.1. 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

$$\langle \mathbf{u}, \mathbf{v} \rangle \equiv \|\mathbf{u}\| \|\mathbf{v}\| \cos \theta = \|\mathbf{v}\| \|\mathbf{u}\| \cos \theta,$$

donde θ 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 2.1.2. El producto $\|\mathbf{u}\| \cos \theta$ es igual a la *proyección escalar* del vector \mathbf{u} sobre el vector \mathbf{v} . Si escribimos dicha proyección escalar como $P_{\mathbf{v}}(\mathbf{u})$, tenemos que

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

Este es el escalar por el cual habría que multiplicar un vector de norma 1 en la dirección y sentido de \mathbf{v} para obtener el vector componente de \mathbf{u} colineal a \mathbf{v} , también conocido como *proyección vectorial*. Empleando esta notación, la proyección vectorial de \mathbf{u} sobre \mathbf{v} se puede escribir como $P_{\mathbf{v}}(\mathbf{u}) \frac{\mathbf{v}}{\|\mathbf{v}\|}$. En la Figura 6 se muestran ejemplos de proyecciones vectoriales entre vectores de \mathbb{R}^2 .

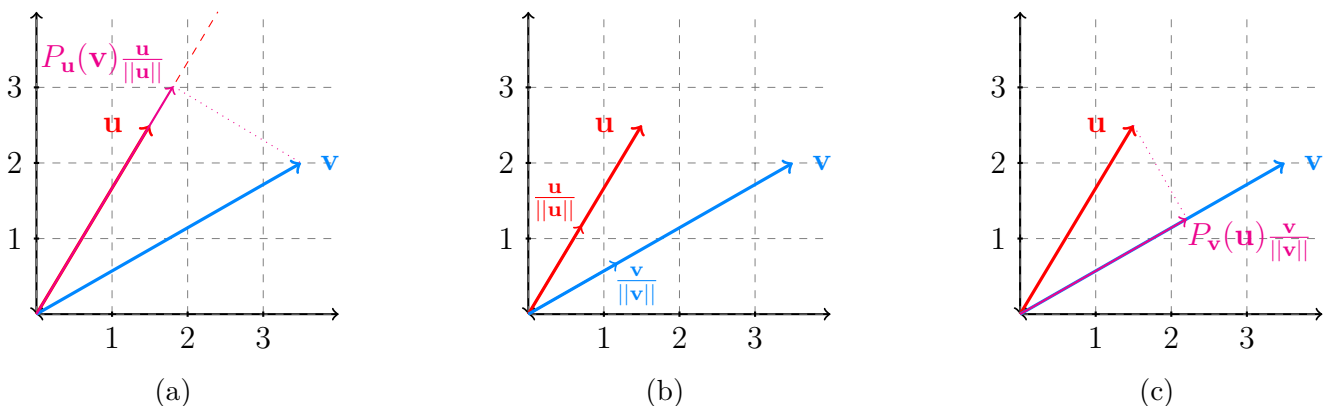


Figura 6: Proyecciones vectoriales entre dos vectores \mathbf{u} y \mathbf{v} de \mathbb{R}^2 . En la subfigura (a) tenemos la proyección vectorial del vector \mathbf{v} sobre el vector \mathbf{u} , es decir, la componente de \mathbf{v} en el eje en el cual se encuentra \mathbf{u} . En (b) se muestran los vectores de norma 1 que tienen la misma dirección y sentido que \mathbf{u} y \mathbf{v} , respectivamente. En (c) se muestra la proyección vectorial de \mathbf{u} sobre \mathbf{v} . Nótese en cada caso que la proyección vectorial se obtiene reescalando el vector sobre el que se proyecta para que tenga norma 1 y después multiplicando dicho vector por la proyección escalar correspondiente.

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

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

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

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

Geoméricamente, podemos ver esto como que $\langle \mathbf{u}', \mathbf{v} \rangle = (2\|\mathbf{u}\|)P_{\mathbf{u}}(\mathbf{v})$ o, equivalentemente, que $\langle \mathbf{u}', \mathbf{v} \rangle = \|\mathbf{u}\|(2P_{\mathbf{u}}(\mathbf{v}))$ ²⁶. Observemos que la elección del vector reescalado fue arbitraria —de haber reescalado \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 $\langle 2\mathbf{u}, \mathbf{v} \rangle = 2\langle \mathbf{u}, \mathbf{v} \rangle = \langle \mathbf{u}, 2\mathbf{v} \rangle$ en espacios vectoriales reales.

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

Sin embargo, al generalizar esta operación 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{v} es distinto al vector nulo. Definimos a la operación $P_{\mathbf{v}}(\cdot) : V \rightarrow K$, conocida como la *proyección escalar del vector \mathbf{u} sobre el vector \mathbf{v}* , como

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

Además, definimos la *proyección vectorial* de \mathbf{u} sobre \mathbf{v} como $P_{\mathbf{v}}(\mathbf{u}) \frac{\mathbf{v}}{\|\mathbf{v}\|}$.

2.3.2. 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 $\mathbf{u} \perp \mathbf{v} \iff \langle \mathbf{u}, \mathbf{v} \rangle = 0$ 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 $\forall \mathbf{u}_i, \mathbf{u}_j \in S, \mathbf{u}_i \neq \mathbf{u}_j \implies \langle \mathbf{u}_i, \mathbf{u}_j \rangle = 0$.

^aObservemos que esta definición también aplica para espacios vectoriales complejos, ya que $\bar{0} = 0$.

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

²⁶Este 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.

2.3.3. Vectores unitarios

Nótese por la definición dada en 2.3.2 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. $\hat{i}, \hat{j}, \hat{k}$, 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 $\langle \mathbf{v}_i, \mathbf{v}_j \rangle = \delta_{ij} \quad \forall \mathbf{v}_i, \mathbf{v}_j \in S$, donde δ_{ij} es la delta de Kronecker de dos entradas.

Por ejemplo, el vector $\left(\frac{1}{\sqrt{2}} \quad \frac{1}{\sqrt{2}}\right)$ es unitario en \mathbb{R}^2 (ó \mathbb{C}^2), mientras que los vectores $\left(\frac{-1}{\sqrt{2}} \quad \frac{i}{\sqrt{2}}\right)$ y $\left(\frac{-i}{\sqrt{2}} \quad \frac{1}{\sqrt{2}}\right)$ 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

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

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.

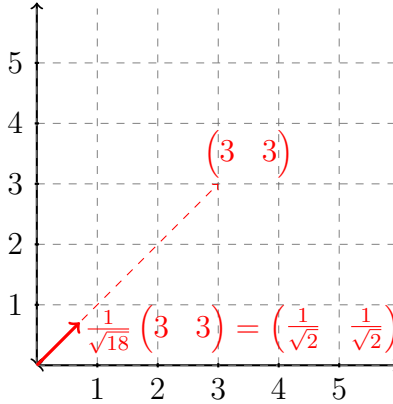


Figura 7: Ejemplo de *normalización* del vector $\begin{pmatrix} 3 & 3 \end{pmatrix} \in \mathbb{R}^2$. Ya que la norma de este vector es igual a $\sqrt{3^2 + 3^2} = \sqrt{18}$, hacemos su producto por el escalar $\frac{1}{\sqrt{18}}$ —el inverso multiplicativo de $\sqrt{18}$ en el campo real— y obtenemos al vector resultante $\begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix}$, que tiene el mismo sentido y dirección que $\begin{pmatrix} 3 & 3 \end{pmatrix}$, pero con norma igual a uno.

Observemos que podemos utilizar la notación usual de vector unitario para reescribir la proyección *vectorial* del vector \mathbf{u} sobre el vector \mathbf{v} como

$$P_{\mathbf{v}}(\mathbf{u}) \hat{\mathbf{v}},$$

donde $P_{\mathbf{v}}(\mathbf{u})$ es la proyección *escalar* de \mathbf{u} sobre \mathbf{v} .

2.4. 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.

2.4.1. 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 3×3 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 θ 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 $\mathbf{0}$.

2.4.2. 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 2.4.1 (¡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.

2.5. 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.

2.5.1. 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 $\mathbf{r} \cdot \mathbf{r} \times \mathbf{s} = 0 = \mathbf{r} \cdot \mathbf{s} \times \mathbf{r}$.

2.5.2. 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 .

2.6. Ejercicios de repaso

2.6.1. 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 (sin contar productos escalares de la forma $\langle f, f \rangle$) utilizando la definición vista en la sección 2.1.2.

2. Sea V un espacio vectorial con producto escalar. Demuestra que para todo $\mathbf{v} \in V$, $\langle \mathbf{v}, \mathbf{0} \rangle = \langle \mathbf{0}, \mathbf{v} \rangle = 0$. Además, demuestra que si $\langle \mathbf{u}, \mathbf{v} \rangle = \langle \mathbf{u}, \mathbf{w} \rangle$ para todo $\mathbf{u} \in V$, entonces $\mathbf{v} = \mathbf{w}$.
3. Explica por qué no se puede definir un producto escalar de la forma $\langle f, g \rangle = \int_a^b f(x)g(x)dx$ para funciones integrables f y g con integrales con valor infinito en $[a, b]$.
4. 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 2.1.3 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).

2.6.2. Norma

1. Calcula la norma de las funciones dadas en el primer ejercicio de la sección 2.6.1. (Nota: utiliza la norma inducida por ese mismo producto escalar.)
2. ¿El conjunto $\{(x_1, x_2, \dots, x_n) \mid x_i \in \mathbb{R} \wedge \|(x_1, x_2, \dots, x_n)\| \leq 1\}$ con las operaciones entre n -tuplas vistas en la sección 1.2.2 puede formar un espacio vectorial sobre el campo real? Argumenta.
3. Argumenta e ilustra la interpretación geométrica de la desigualdad de Cauchy-Schwarz para dos vectores cualesquiera de \mathbb{R}^2 (Nota: ver sec. 2.2.3).
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 $\|\mathbf{a}\| \equiv +\sqrt{\langle \mathbf{a}, \mathbf{a} \rangle} \quad \forall \mathbf{a} \in V$ (pista: usa la desigualdad de Cauchy-Schwarz)²⁷. Argumenta e ilustra su interpretación geométrica para dos vectores cualesquiera de \mathbb{R}^2 .

2.6.3. Interpretación geométrica del producto escalar: proyecciones y ortogonalidad

1. Sea $\mathbf{v} \in \mathbb{R}^2$ no-nulo. Prueba que el conjunto $\{\mathbf{u} \in \mathbb{R}^2 \mid \langle \mathbf{u}, \mathbf{v} \rangle = 0\}$ es una recta que pasa por el origen y llévala a una expresión de la forma $ax + by = c$.
2. Sean $\mathbf{v} = \begin{pmatrix} v_1 & v_2 \end{pmatrix}$, $\mathbf{u} = \begin{pmatrix} u_1 & u_2 \end{pmatrix}$ 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 θ 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 $\langle \mathbf{u}, \mathbf{u} - k\mathbf{v} \rangle = 0$. ¿Por qué es especial este escalar?)
3. Define a los vectores $\mathbf{u} = \begin{pmatrix} a_1 & a_2 \end{pmatrix}$, $\mathbf{v} = \begin{pmatrix} a_3 & a_4 \end{pmatrix}$ y $\mathbf{w} = \begin{pmatrix} a_5 & a_6 \end{pmatrix} \in \mathbb{R}^2$, donde a_i corresponde al i -ésimo dígito de tu número de cuenta. Explica geoméricamente el hecho de que $\langle \mathbf{u} + \mathbf{v}, \mathbf{w} \rangle = \langle \mathbf{u}, \mathbf{w} \rangle + \langle \mathbf{v}, \mathbf{w} \rangle$, en términos de las proyecciones discutidas en la sección 2.3 (Nota: recuerda la Ley del paralelogramo).

2.6.4. Producto vectorial (cruz)*

1. Demuestra las propiedades del producto vectorial de la sección 2.4.2. 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.)

²⁷Con esto habrás demostrado que a partir de cualquier producto escalar positivo definido $\langle \mathbf{a}, \mathbf{b} \rangle$ se puede definir una norma como $\|\mathbf{a}\| = +\sqrt{\langle \mathbf{a}, \mathbf{a} \rangle}$.

2.6.5. Triple producto escalar*

1. Demuestra las propiedades del triple producto escalar de la sección 2.5.2.

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

Como seguramente habrás notado al leer los libros recomendados, 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 en 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.

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

3.1. Combinaciones lineales

Como vimos en la sección 1.2.1, 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 1.2.1, 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

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

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

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

Siguiendo la discusión de la sección 1.3, 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*.

3.2. 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 ^a. Por completez, definimos $\langle \emptyset \rangle = \{\mathbf{0}\}$. Si G consiste de un sólo vector \mathbf{g} , podemos escribir el generado de G como $\langle \mathbf{g} \rangle$.

^aNótese 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 $N = \left\{ \begin{pmatrix} 1 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 1 \end{pmatrix} \right\}$ entonces vemos que

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

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

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

Teorema 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.

3.3. Dependencia e independencia lineal

Como se vio en la sección 3.1, 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 , $7\begin{pmatrix} 1 & 1 \end{pmatrix} + 5\begin{pmatrix} -1 & 1 \end{pmatrix} + 2\begin{pmatrix} -1 & -6 \end{pmatrix} = \begin{pmatrix} 0 & 0 \end{pmatrix} = \mathbf{0}$. Una consecuencia de este hecho es que podamos despejar a cualquiera de los vectores y expresarlo como combinación lineal de los demás; por ejemplo, $\begin{pmatrix} 1 & 1 \end{pmatrix} = -\frac{5}{7}\begin{pmatrix} -1 & 1 \end{pmatrix} - \frac{2}{7}\begin{pmatrix} -1 & -6 \end{pmatrix}$, ó $\begin{pmatrix} -1 & -6 \end{pmatrix} = -\frac{7}{2}\begin{pmatrix} 1 & 1 \end{pmatrix} - \frac{5}{2}\begin{pmatrix} -1 & 1 \end{pmatrix}$, etc. Este importante hecho motiva la siguiente definición.

3.3.1. 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^a.

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

^aEl 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 $\mathbf{u}_1 = \begin{pmatrix} 1 & 5 \end{pmatrix}$ y $\mathbf{u}_2 = \begin{pmatrix} -3 & -15 \end{pmatrix}$ son linealmente dependientes, ya que $\mathbf{u}_2 = -3\mathbf{u}_1$, por lo cual $3\mathbf{u}_1 + \mathbf{u}_2 = \mathbf{0}$; por otro lado, los vectores $\mathbf{v}_1 = \begin{pmatrix} 1 & 2 \end{pmatrix}$ y $\mathbf{v}_2 = \begin{pmatrix} 1 & 3 \end{pmatrix}$ son linealmente independientes, ya que no existe un número $c \in \mathbb{R}$ tal que $\mathbf{v}_1 = c\mathbf{v}_2$. Notemos que, como nuestros vectores en este caso son 2-tuplas, las ecuaciones $\mathbf{u}_2 = -3\mathbf{u}_1$ y $\mathbf{v}_1 = c\mathbf{v}_2$ son en realidad la notación compactada 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

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

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

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

3.3.2. 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.

3.3.3. 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 sec. 1.2.3), $0\mathbf{v}_{k+1} + \dots + 0\mathbf{v}_n = \mathbf{0}$ 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.

3.4. Ejercicios de repaso

3.4.1. 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 sec. 2.)
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.)

3.4.2. 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?
2. Sea $G = \left\{ \begin{pmatrix} 3 & 0 & 3 \end{pmatrix}, \begin{pmatrix} -\frac{1}{2} & 0 & \frac{1}{2} \end{pmatrix} \right\} \subset \mathbb{R}^3$. Escribe a $\langle G \rangle$ algebraicamente y descríbelo geoméricamente. ¿A qué espacio vectorial real que conoces se parece?
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.).

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, $\langle F \rangle$?

3.4.3. Dependencia e independencia lineal

1. Sea n_i el i -ésimo dígito de tu número de cuenta. Determina si los vectores $(n_1 + i(n_2))$ y $(n_3 + i(n_4)) \in \mathbb{C}$ son linealmente dependientes o independientes y muéstralo gráficamente en el plano complejo.
2. Repite el mismo ejercicio para los vectores reales $(n_1 \ n_4 \ n_7)$, $(n_2 \ n_5 \ n_8)$ y $(n_3 \ n_6 \ n_9) \in \mathbb{R}^3$. ¿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?
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.
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?
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.
6. Describe con tus palabras y con los conceptos vistos en clase el procedimiento que seguiste para resolver el ejercicio 3.4.3.5.

4. Bases, dimensión, ortogonalización y ortonormalización

4.1. 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 espacio²⁸. 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. 3.3.3), 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.

4.1.1. 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 espacio²⁹.

4.1.2. 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 $\{(2 \ 0), (0 \ -3)\}$, $\{(3 \ 3), (-3 \ 3)\}$ y $\{(1 \ 0), (0 \ 1)\}$ son bases de \mathbb{R}^2 .

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

4.1.3. Teorema de reemplazamiento

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

²⁸De lo contrario, se violarían las propiedades de cerradura vistas en la sec. 1.2.1.

²⁹El 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 ³⁰?

³⁰Recuerda para qué dijimos que nos serviría este teorema.

4.2. 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*^a.

^aEn 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 4.1.2. 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 3.2 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 3.3.3, es linealmente independiente. Si L' no genera a V , podemos repetir el proceso hasta llegar a un conjunto linealmente independiente que genera a V , i.e., una base para V . \square

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

4.3. Ortogonalización y ortonormalización

Recordemos de la sección 2.3 que dos vectores \mathbf{u} y \mathbf{v} son ortogonales si $\langle \mathbf{u}, \mathbf{v} \rangle = 0$. Supongamos que este no es el caso, i.e., que $\langle \mathbf{u}, \mathbf{v} \rangle \neq 0$ ³¹. Existe una manera de modificar cualquiera de los vectores de tal forma que se vuelva ortogonal al otro: simplemente definimos $\mathbf{u}' = \mathbf{u} - \langle \mathbf{u}, \mathbf{v} \rangle \frac{\mathbf{v}}{\|\mathbf{v}\|^2}$ y comprobamos que

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

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 $\mathbf{u}' = \mathbf{u} - \langle \mathbf{u}, \mathbf{v} \rangle \mathbf{v}$, 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..

4.3.1. 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 $\langle \mathbf{o}_i, \mathbf{o}_j \rangle = 0 \quad \forall \mathbf{o}_i, \mathbf{o}_j \in O, i \neq j$.

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 $\langle \mathbf{n}_i, \mathbf{n}_j \rangle = \delta_{ij} \quad \forall \mathbf{n}_i, \mathbf{n}_j \in N$, donde δ_{ij} es la *delta de Kronecker* de dos índices^a.

^aEn 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.

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 $\mathbf{u} \in \langle S \rangle \implies$

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

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 $1 \leq j \leq n$

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

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

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

Sustituyendo, tenemos que

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

como se quería demostrar originalmente. En particular, observamos que $c_i = \frac{P_{\mathbf{v}_i}(\mathbf{u})}{\|\mathbf{v}_i\|}$. □

³¹Aquí 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 sec. 2.1.1).

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 $c_j = \langle \mathbf{u}, \mathbf{v}_j \rangle$ en la demostración anterior y el resultado que obtuvimos se reduce a

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

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 $c_i = \frac{\langle \mathbf{0}, \mathbf{v}_i \rangle}{\|\mathbf{v}_i\|^2} = 0 \quad \forall c_i \implies S$ 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.

4.3.2. 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

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

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

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

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 forzadamente $\mathbf{v}_k \neq \mathbf{0}$.

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

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

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. \square

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.

4.4. Ejercicios de repaso

4.4.1. Bases

1. Sea K un campo arbitrario. En el ejercicio 1.5.2.4 de la sección 1.5.2 demostraste que $K \times K \times \dots \times K = K^n, n \in \mathbb{N}$ era un espacio vectorial sobre K . Sea $\mathbf{e}_1 = (1 \ 0 \ 0 \ 0 \ \dots \ 0), \mathbf{e}_2 = (0 \ 1 \ 0 \ 0 \ \dots \ 0), \dots, \mathbf{e}_n = (0 \ 0 \ 0 \ 0 \ \dots \ 1)$. 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 .
2. Sea d_i el i -ésimo dígito de tu número de cuenta. Considera el conjunto $\{(d_1 \ d_2 \ d_3), (d_4 \ d_5 \ d_6), (d_7 \ d_8 \ d_9)\}$. ¿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.
3. Sea B la base que obtuviste en el ejercicio anterior. Expresa a los vectores $(d_9 \ d_8 \ d_7), (d_6 \ d_5 \ d_4), (d_3 \ d_2 \ d_1) \in \mathbb{R}^3$ como combinaciones lineales de los vectores de B .

4.4.2. 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.
2. Demuestra que cualquier base del espacio vectorial \mathbb{C} sobre \mathbb{R} debe tener más de un elemento y que, por lo tanto, $\dim((\mathbb{C}, \mathbb{C})) \neq \dim((\mathbb{C}, \mathbb{R}))$.
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.

4.4.3. 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 .
2. Aplica el Teorema de Gram-Schmidt *modificado* y construye una base ortonormal a partir de B .
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 sec. 2.1.2).

5. Transformaciones lineales y sus representaciones matriciales
6. Composición de transformaciones lineales y multiplicación de matrices
7. Invertibilidad e isomorfismos
8. Matriz de cambio de base
9. Eigenvectores y eigenvalores
10. Diagonalizabilidad

11. Introducción a sistemas lineales de ecuaciones diferenciales

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

11.1. Sistemas lineales de ecuaciones algebraicas

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

$$x_1 + x_2 = 1,$$

$$x_1 - x_2 = 5.$$

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

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

$$f_1(x_1, x_2, \dots, x_n) = 0$$

$$f_2(x_1, x_2, \dots, x_n) = 0$$

...

$$f_m(x_1, x_2, \dots, x_n) = 0,$$

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

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

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

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

11.2. Introducción a ecuaciones diferenciales (ordinarias)

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

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

donde \dot{x} representa la derivada de la función x con respecto a alguna variable independiente —usualmente llamada t , aunque esto es sólo una etiqueta³²—, tiene como solución a la función $x(t) = e^{-t}$, y dicha solución es válida para todo $t \in \mathbb{R}$ ya que en todo este intervalo se verifica que³³

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

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

Def. Una *ecuación diferencial ordinaria* es una relación entre una función incógnita $x(t)$, sus derivadas, y una variable independiente t . En general, podemos expresar una ecuación diferencial ordinaria como

$$f(t, x(t), \dot{x}(t), \ddot{x}(t), \dots, x^{(n)}(t)) = 0,$$

donde f es una función arbitraria y $x^{(n)}(t)$ denota la n -ésima derivada de la función incógnita x con respecto a su variable independiente t . Decimos que una función $x(t)$ es una *solución* a la ecuación diferencial anterior en un intervalo I si tanto $x(t)$ como sus derivadas cumplen dicha ecuación para todo $t \in I$.

Podemos clasificar algunos tipos de ecuaciones diferenciales ordinarias, como sigue³⁴:

³²Con *etiqueta* nos referimos a que, de haber postulado que la función x dependía de una variable independiente s , entonces la solución a la ecuación diferencial sería $x(s) = e^{-s}$, lo cual no cambia en absoluto el sentido de la ecuación ni de su solución. A este tipo de variables “etiqueta” acostumbramos llamarles *variables mudas*.

³³Observemos que pudimos haber planteado esta ecuación diferencial de manera equivalente como $\dot{x} = -x$, y la solución sería la misma, ya que $\frac{d}{dt}e^{-t} = -e^{-t}$ para todo $t \in \mathbb{R}$.

³⁴Observen la analogía con la forma en que clasificamos ecuaciones algebraicas (polinomiales), pero también las diferencias. Las definiciones presentadas serán relevantes para nuestros fines.

Def. Decimos que una ecuación diferencial ordinaria es:

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

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

para funciones $a_1(t), a_2(t), \dots, a_n(t), b(t)$ dadas (que pueden, en particular, ser constantes)^a;

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

^aEn particular, si $b(t)$ es la constante cero, decimos que la ecuación es lineal *homogénea*.

11.3. Ecuaciones diferenciales (ordinarias) con condiciones iniciales y sistemas lineales de ecuaciones diferenciales

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

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

podemos verificar que la función $x(t) = Ce^{-t}$ es una solución en todo el intervalo \mathbb{R} para cualquier $C \in \mathbb{R}$. Decimos entonces que $x(t) = Ce^{-t}$ es una *solución general* a la ecuación diferencial anterior³⁵. En esta solución general, el valor C representa un *grado de libertad* de la solución, el cual puede removerse agregando una *restricción* a la ecuación diferencial, la cual viene en forma de una *condición inicial*.

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

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

Por lo tanto, la solución a la ecuación diferencial $\dot{x} + x = 0$ con condición inicial $x(0) = C_0$, donde $C_0 \in \mathbb{R}$ es un valor determinado, es $x(t) = C_0e^{-t}$. Ya que en este caso $x(t)$ sigue cumpliendo la ecuación diferencial original pero ya no tiene grados de libertad, decimos que es una *solución particular* a la ecuación diferencial $\dot{x} + x = 0$ sujeta a la *condición inicial* $x(0) = C_0$, con $C_0 \in \mathbb{R}$.

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

$$x(t_0) = C_0,$$

donde C_0 es un valor fijo.

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

$$\dot{x}_1 - x_2 = 0,$$

$$x_1 + \dot{x}_2 = 0,$$

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

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

³⁵Por dar otro ejemplo, podemos decir que la ecuación $\ddot{x} + x = 0$ tiene como solución general a $x(t) = A \sin(t) + B \cos(t)$ en todo el intervalo \mathbb{R} para cualesquiera valores $A, B \in \mathbb{R}$ (¡compruébalo!).

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

11.4. Ejercicios de repaso

11.4.1. Sistemas lineales de ecuaciones algebraicas

1. Realiza un cuadro comparativo con las definiciones de sistema de ecuaciones, sistema de ecuaciones algebraicas y sistema lineal de ecuaciones algebraicas.
2. Escribe o argumenta por qué no puedes escribir un ejemplo de:
 - a) una ecuación *trascendental* (i.e., no algebraica);
 - b) un sistema no lineal de ecuaciones algebraicas;
 - c) un sistema lineal de ecuaciones no algebraicas.
3. Demuestra que en una ecuación lineal homogénea con más de una incógnita el conjunto de soluciones forma un espacio vectorial. Después, generaliza el resultado para demostrar que en un sistema lineal de ecuaciones algebraicas homogéneas con menos ecuaciones que incógnitas el conjunto de soluciones forma un espacio vectorial.

11.4.2. Introducción a ecuaciones diferenciales (ordinarias)

1. Encuentra la solución general a la ecuación $\dot{x} - x = 0$ en todo el intervalo \mathbb{R} .
2. Sea $C^\infty(\mathbb{R})$ el conjunto de todas las funciones reales de variable real con derivadas de cualquier orden y sea $\frac{d}{dt} : C^\infty(\mathbb{R}) \rightarrow C^\infty(\mathbb{R})$ el operador lineal *derivada*. Supongamos que $x(t)$ es un eigenvector de $\frac{d}{dt}$ con eigenvalor λ , ¿entonces quién es $x(t)$?
3. Muestra que tanto la ecuación diferencial del ejemplo dado al inicio de la sección 11.3 como su solución general son un caso particular del ejercicio anterior.

11.4.3. Ecuaciones diferenciales (ordinarias) con condiciones iniciales

1. Encuentra la solución particular a la ecuación $\dot{x} + x = 0$ con la condición inicial $x(0) = 5$.
2. Encuentra la solución particular a la ecuación $\dot{x} + x = 0$ con la condición inicial $x(1) = 5$.
3. Muestra que cualquier solución particular (es decir, sin grados de libertad) de la ecuación $\ddot{x} = g$ debe tener dos condiciones iniciales: una para la función $x(t)$ y otra para $\dot{x}(t)$. ¿De qué ecuación se trata?

11.4.4. Sistemas lineales de ecuaciones diferenciales (ordinarias)

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

$$\begin{aligned} \dot{x}_1 + 3x_2 &= 0, \\ \dot{x}_2 - x_3 &= 0, \\ x_1 - 6\dot{x}_3 &= 0. \end{aligned}$$

12. Modelación con sistemas lineales de ecuaciones diferenciales

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

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

Durante este último módulo del curso, nuestro principal objeto de estudio serán los espacios vectoriales con productos escalares, así como los operadores lineales que actúan sobre dichos espacios. Descubriremos que, con ayuda del producto escalar, podemos llegar a tener una comprensión profundamente geométrica sobre cómo actúan ciertos operadores en estos espacios —más allá de una simple descripción algebraica dada por reglas de correspondencia. Este resultado es uno de los más importantes de todo el curso, y se le conoce como el *Teorema espectral*. Dado que haremos uso frecuente del producto escalar $\langle \cdot, \cdot \rangle$ —principalmente, a partir de la sección 13.2— en lo que resta del curso utilizaremos la llamada *notación de bra-ket*, también conocida como *notación de Dirac*, por lo que se presenta a continuación.

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

$$\begin{aligned} |v\rangle & \quad \text{vector o } ket \\ \langle u|v\rangle & \quad \text{producto escalar o } bra-ket \\ \langle u| & \quad \text{vector dual (funcional) o } bra \end{aligned}$$

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

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

y

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

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

Con esta notación podemos, por ejemplo, reescribir algunos resultados que ya conocemos sobre bases ortonormales de una manera mucho más sencilla, como se muestra a continuación³⁶: sea V un espacio vectorial de dimensión finita n con base ortonormal $\beta = \{|v_1\rangle, |v_2\rangle, \dots, |v_n\rangle\}$, entonces para todo $|v\rangle \in V$

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

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

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

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

³⁶Asegúrate de entender bien esta notación y a qué resultados nos referimos antes de continuar leyendo el resto de las notas.

13.1. Funcionales y espacio dual

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

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

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

Ejemplos de funcionales

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

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

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

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

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

$$P_{|v\rangle}(|u\rangle) = \frac{\langle v|u\rangle}{\| |v\rangle \|^2} |v\rangle$$

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

13.2. Correspondencia entre bras y kets

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

Dado que a partir de cualquier vector $|v\rangle \in V$ se puede definir un funcional $\langle v | \in V^*$, existe una correspondencia natural entre *bras* y *kets*. Sin embargo, para entender bien esta correspondencia debemos observar un detalle crucial: si hacemos el producto de un ket $|v\rangle$ por un escalar $\alpha \in K$, obteniendo así el ket $\alpha |v\rangle$, entonces el bra correspondiente será $\overline{\alpha} \langle v |$ ³⁷. Es decir, la correspondencia entre bras (que son un tipo de vectores duales) y kets está dada por

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

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

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

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

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

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

³⁷En la notación que usábamos antes, esto se explica como que si tomamos al vector $\alpha \mathbf{v}$ y lo fijamos en la entrada antilineal del producto escalar, obtenemos el funcional $\langle \cdot, \alpha \mathbf{v} \rangle : V \rightarrow K$ que, como sabemos, es igual a $\overline{\alpha} \langle \cdot, \mathbf{v} \rangle$.

13.3. Complemento ortogonal

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

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

De la definición anterior podemos hacer varias observaciones:

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

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

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

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

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

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

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

Teorema 13.1.1. Sea V un espacio vectorial con producto escalar y $W \subseteq V$ un subespacio vectorial de dimensión finita k , entonces para todo $|y\rangle \in V$ existen vectores únicos $|u\rangle \in W$ y $|z\rangle \in W^\perp$ tales que $|y\rangle = |u\rangle + |z\rangle$. Además, si $\beta = \{|v_1\rangle, |v_2\rangle, \dots, |v_k\rangle\}$ es una base ortonormal de W , entonces

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

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

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

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

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

$$\begin{aligned}
&= \sum_{i=1}^k \langle v_i | y \rangle \langle y | v_i \rangle - \sum_{i=1}^k \sum_{j=1}^k \langle y | v_i \rangle \langle v_j | y \rangle \langle v_i | v_j \rangle = \sum_{i=1}^k \langle v_i | y \rangle \langle y | v_i \rangle - \sum_{i=1}^k \sum_{j=1}^k \langle y | v_i \rangle \langle v_j | y \rangle \delta_{ij} \\
&= \sum_{i=1}^k \langle v_i | y \rangle \langle y | v_i \rangle - \sum_{i=1}^k \langle y | v_i \rangle \langle v_i | y \rangle = 0,
\end{aligned}$$

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

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

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

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

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

□

Teorema 13.1.2. Sea $S = \{|v_1\rangle, |v_2\rangle, \dots, |v_k\rangle\}$ un conjunto ortonormal de vectores de un espacio vectorial V de dimensión finita n con producto escalar. Entonces

- a) S se puede extender a una base ortonormal $\{|v_1\rangle, |v_2\rangle, \dots, |v_k\rangle, |v_{k+1}\rangle, \dots, |v_n\rangle\}$ de V .
- b) Usando la notación anterior, si $W = \langle S \rangle$, entonces $S_1 = \{|v_{k+1}\rangle, \dots, |v_n\rangle\}$ es base ortonormal de W^\perp .
- c) Para cualquier subespacio $W \subseteq V$, $\dim(V) = \dim(W) + \dim(W^\perp)$.

Demostración. a) Por el Teorema de reemplazamiento (ver sec. 4.1.3), sabemos que podemos extender a S para formar una base $S' = \{|v_1\rangle, |v_2\rangle, \dots, |v_k\rangle, |w_{k+1}\rangle, \dots, |w_n\rangle\}$ de V . Aplicando el Teorema de Gram-Schmidt, podemos obtener una base $\{|v_1\rangle, |v_2\rangle, \dots, |v_k\rangle, |v_{k+1}\rangle, \dots, |v_n\rangle\}$ de V .

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

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

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

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

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

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

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

□

13.4. Proyecciones ortogonales

En la sección 2.3 vimos que, en un espacio vectorial con producto escalar, podemos definir funciones de *proyección vectorial* que toman un vector del espacio y devuelven la componente de ese vector a lo largo de un subespacio vectorial particular. Además, en el Teorema 4.3.1.1 (ver sec. 4.3) vimos cómo podemos descomponer un vector en sus componentes ortogonales de manera sencilla a partir de una base ortogonal u ortonormal. Ambos resultados reflejan la importancia que tienen los llamados *operadores de proyección ortogonal* en este tipo de espacios vectoriales.

Def. Sea V un espacio vectorial arbitrario. Decimos que un operador lineal $P : V \rightarrow V$ es un *operador de proyección* si

$$P(P(|v\rangle)) = P(|v\rangle) \quad \forall |v\rangle \in V.$$

Definiendo $P^2 := P \circ P$ podemos reescribir esta condición simplemente como $P^2 = P$. En particular, si V tiene producto escalar y P es tal que

$$\text{Im}(P)^\perp = \text{Ker}(P) \quad \text{y} \quad \text{Ker}(P)^\perp = \text{Im}(P),$$

decimos que P es un *operador de proyección ortogonal*.

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

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

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

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

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

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

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

A pesar de que exista la correspondencia entre kets $\alpha|x\rangle \in V$ y bras $\bar{\alpha}\langle x| \in V^*$ descrita en la sección 13.2, debemos remarcar una diferencia sutil pero importante entre la operación de aplicarle un funcional a un vector y la operación de obtener el producto escalar entre dos vectores antes de pasar a la siguiente sección.

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

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

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

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

13.5. Ejercicios de repaso

13.5.1. Funcionales y espacio dual

1.

13.5.2. Correspondencia entre bras y kets

1. Sea (V, K) un espacio vectorial con producto escalar. Demuestra que el conjunto de todos los bras correspondientes a los vectores de V forma un subespacio vectorial de V^* .
2. Demuestra que el subespacio de V^* obtenido en el ejercicio anterior es isomorfo a V . En otras palabras, demuestra que la correspondencia entre bras y kets dada en la sección 13.2 es un isomorfismo.

13.5.3. Complemento ortogonal

1. Demuestra que para todo subespacio vectorial W de un espacio vectorial V de dimensión finita n y con producto escalar, V es igual a una suma directa de W y W^\perp . ¿Por qué es importante que V tenga dimensión finita? (Nota: tendrás que demostrar las observaciones hechas al principio de la sec. 13.3.)

13.5.4. Proyecciones ortogonales

1. Sea V un espacio vectorial. Demuestra que un operador lineal $P : V \rightarrow V$ es un operador de proyección si y sólo si $V = \text{Ker}(P) \oplus \text{Im}(P)$.
2. Sea V un espacio vectorial con producto escalar. Demuestra que un operador lineal $P : V \rightarrow V$ es un operador de proyección ortogonal si y sólo si $V = \text{Ker}(P)^\perp \oplus \text{Im}(P)^\perp$.
3. Sea V un espacio vectorial de dimensión finita n con producto escalar y sea $\{|v_1\rangle, |v_2\rangle, \dots, |v_k\rangle\}$ un conjunto ortogonal de vectores de V con $k \leq n$. Demuestra que $\sum_{i=1}^k \frac{|v_i\rangle\langle v_i|}{\langle v_i|v_i\rangle} : V \rightarrow V$ es un operador de proyección ortogonal.

14. Descomposición espectral y operadores adjuntos

Supongamos que tenemos un espacio vectorial (V, K) de dimensión finita n con producto escalar. Entonces, siguiendo del último ejercicio de la sección 13, tendríamos que si $\{|b_1\rangle, |b_2\rangle, \dots, |b_n\rangle\}$ es una base ortogonal de V , entonces $\sum_{i=1}^n \frac{|b_i\rangle\langle b_i|}{\langle b_i|b_i\rangle} : V \rightarrow V$ es un operador de proyección ortogonal³⁸. En particular, para cualquier base ortonormal $\{|u_1\rangle, |u_2\rangle, \dots, |u_n\rangle\}$ de V tenemos que $\sum_{i=1}^n |u_i\rangle\langle u_i|$ es un operador de proyección ortogonal (equivalente al anterior, para cualquier base ortonormal arbitraria).

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

14.1. Descomposición espectral (introducción)

Sean V un espacio vectorial de dimensión finita n con producto escalar y $T : V \rightarrow V$ un operador lineal diagonalizable tal que sus eigenvectores $|g_1\rangle, |g_2\rangle, \dots, |g_n\rangle$ forman una base³⁹ *ortogonal* de V ; es decir, que se puede elegir a una base ortogonal de V compuesta por eigenvectores de T . Supongamos, adicionalmente, que normalizamos a cada uno de los vectores de la eigenbase y los redefinimos de tal forma que $\{|g_1\rangle, |g_2\rangle, \dots, |g_n\rangle\}$ sea una base *ortonormal* de V compuesta por eigenvectores de T . En ese caso, tendríamos que

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

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

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

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

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

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

³⁸Este operador no depende de la base ortogonal elegida... ¿puedes adivinar de qué operador se trata?

³⁹Recordemos que un operador $T : V \rightarrow V$ es diagonalizable si y sólo si se puede formar una base de V compuesta de eigenvectores de T . Si V es dimensión n , entonces cualquier base de V tiene exactamente n elementos, por lo que T debe tener n eigenvectores linealmente independientes; sin embargo, esto no necesariamente implica que T tenga n eigenvalores distintos, como hemos visto. Por lo tanto, si T tiene eigenvalores $\lambda_1, \lambda_2, \dots, \lambda_k$ con $k < n$, necesariamente habrá elementos distintos de la eigenbase que tengan el mismo eigenvalor y, en este caso, la correspondencia entre los subíndices de los k eigenvalores distintos y los n elementos de la eigenbase no será uno a uno, por lo que hay que tener cuidado y no dejarse llevar por la notación.

Observemos que la descomposición espectral de T nos da información completa de cómo actúa T en V geoméricamente: la expresión anterior nos dice que T proyecta a un vector arbitrario $|v\rangle \in V$ en cada una de sus componentes a lo largo de los diferentes eigenspacios de T (a través de las proyecciones ortogonales $|g_i\rangle\langle g_i|$) —los cuales son ortogonales entre sí—, luego multiplica cada una de esas componentes por el eigenvalor correspondiente al eigenspacio en el que se encuentran, y luego reconstruye el vector resultante a partir de las componentes reescaladas⁴⁰.

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

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

donde $\lambda_1, \lambda_2, \dots, \lambda_k$ son los eigenvalores de T , entonces podemos entender a T de una forma extremadamente sencilla: a través de proyecciones (en los eigenspacios de T , que son ortogonales entre sí) y reescalamientos (por los eigenvalores de T). Un ejemplo de esto se muestra en la Figura 8.

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

14.2. Operadores adjuntos

Como vimos en la sección 13.2, en cualquier espacio vectorial (V, K) con producto escalar podemos tomar cualquier vector $|y\rangle \in V$ y definir un funcional $\mathbf{g} : V \rightarrow K$ como

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

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

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

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

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

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

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

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

⁴⁰Aquí utilizamos la palabra “reescalamiento” como sinónimo de “producto de un vector por un reescalar”; recordemos que, geoméricamente, en general esta operación puede involucrar un reescalamiento y una rotación.

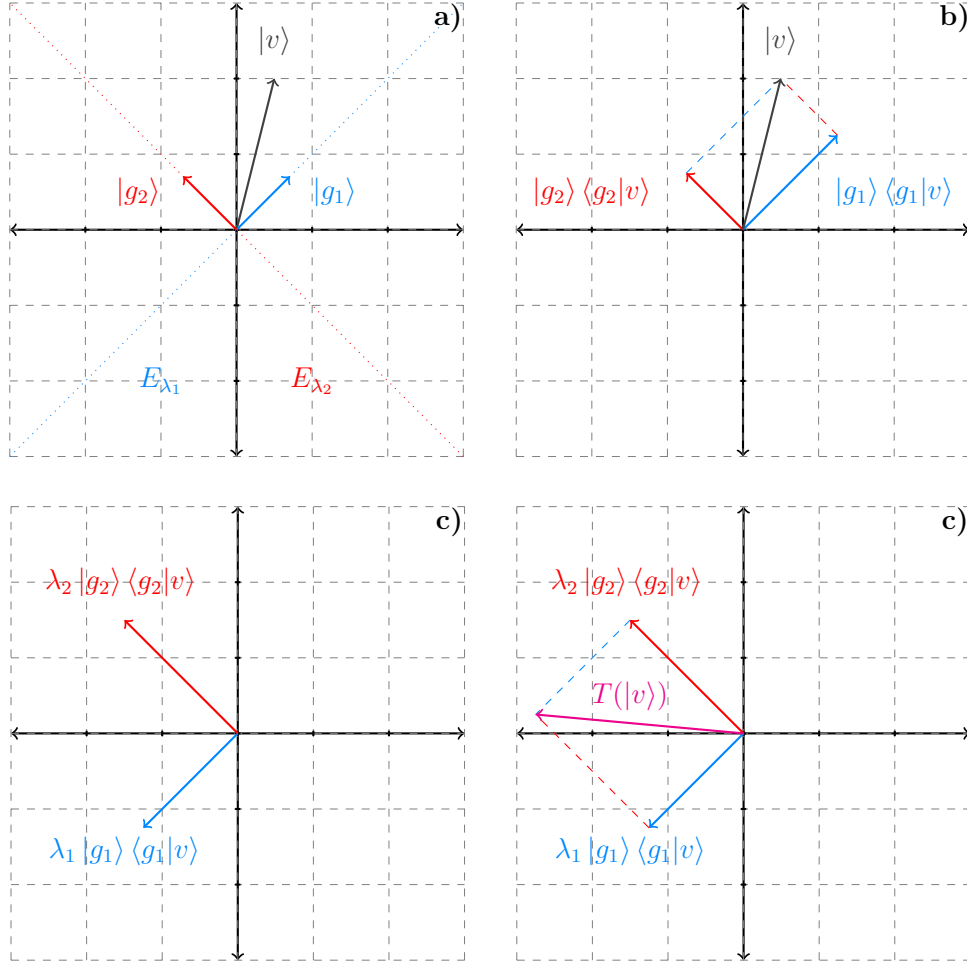


Figura 8: Aplicación de un operador $T : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ a un vector $|v\rangle$ a través de descomposición espectral. En la subfigura **a)** se muestra a $|v\rangle$ y a los eigenvectores $|g_1\rangle, |g_2\rangle$ de T , que forman una base ortonormal; los eigespacios correspondientes se representan con líneas punteadas. En **b)** se muestran las proyecciones ortogonales de $|v\rangle$ en cada uno de los eigespacios de T . En **c)** se muestran las componentes anteriores reescaladas por los eigenvalores correspondiente al eigespacio en que se encuentran. Finalmente, en **d)** vemos que la suma de estas componentes reescaladas es igual a la imagen de $|v\rangle$ bajo el operador T .

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

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

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

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

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

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

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

□

Antes de continuar, recordamos la siguiente definición.

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

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

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

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

por lo que T es lineal.

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

de donde se sigue que

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

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

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

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

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

es la *matriz adjunta* de A .

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

$$A_{ij}^T = A_{ji},$$

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

14.3. Ejercicios de repaso

14.3.1. Descomposición espectral (introducción)

1. Sean V un espacio vectorial de dimensión finita, $T : V \rightarrow V$ un operador lineal y $\{\lambda_1, \lambda_2, \dots, \lambda_k\}$ el espectro de T . Demuestra que T es diagonalizable si y sólo si $V = E_{\lambda_1} \oplus E_{\lambda_2} \oplus \dots \oplus E_{\lambda_k}$.
2. Sean V un espacio de dimensión finita con producto escalar, $T : V \rightarrow V$ un operador lineal y $\Lambda = \{\lambda_1, \lambda_2, \dots, \lambda_k\}$ el espectro de T . Demuestra que T es diagonalizable si y sólo si $V = E_{\lambda_i} \oplus (E_{\lambda_i})^\perp$ para toda $\lambda_i \in \Lambda$.
3. Sean V un espacio de dimensión finita n con producto escalar, $T : V \rightarrow V$ un operador lineal, $\{|g_1\rangle, |g_2\rangle, \dots, |g_n\rangle\}$ un conjunto de eigenvectores normales de T y $\Lambda = \{\lambda_1, \lambda_2, \dots, \lambda_n\}$ un conjunto de eigenvalores de T (no necesariamente todos distintos) tales que $T(|g_i\rangle) = \lambda_i |g_i\rangle \quad \forall i \in \{1, 2, \dots, n\}$. Demuestra que $T = \sum_{i=1}^n \lambda_i |g_i\rangle \langle g_i|$ si y sólo si $I = \sum_{i=1}^n |g_i\rangle \langle g_i|$.

14.3.2. Operadores adjuntos

1. Sea V un espacio vectorial de dimensión finita n con producto escalar y $\beta = \{|b_1\rangle, |b_2\rangle, \dots, |b_n\rangle\}$ una base ortonormal de V . Demuestra que si $T : V \rightarrow V$ es lineal, entonces $([T]_\beta)_{ij} = \langle b_i | T(b_j) \rangle$.
2. Sea K un campo y $A \in M_{n \times n}(K)$. Demuestra que $A^* = \overline{A^T} = (\overline{A})^T$ y, por lo tanto, que $\det(A^*) = \overline{\det(A)}$.
3. Sea (V, K) un espacio vectorial con producto escalar y sean T y U operadores lineales sobre V para los cuales existen operadores adjuntos T^* y U^* , respectivamente. Demuestra las siguientes propiedades:
 - a) $(T + U)^* = T^* + U^*$;
 - b) $(cT)^* = \overline{c}T^*$ para todo $c \in K$;
 - c) $(TU)^* = U^*T^*$;
 - d) $(T^*)^* = T$;
 - e) $I^* = I$.
4. Sea K un campo y sean $A, B \in M_{n \times n}(K)$. Demuestra las siguientes propiedades:
 - a) $(A + B)^* = A^* + B^*$;
 - b) $(cA)^* = \overline{c}A^*$ para todo $c \in K$;
 - c) $(AB)^* = B^*A^*$;
 - d) $(A^*)^* = A$;
 - e) $I^* = I$.

15. Operadores normales y autoadjuntos y teorema espectral

15.1. Operadores normales y autoadjuntos

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

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

Demostración. Sea $|v\rangle \in V$ un eigenvector de T con eigenvalor λ . Entonces, por el ejercicio 3 de la sección 14.3.2, para toda $|u\rangle \in V$ tenemos que

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

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

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

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

por lo que $[T^*]$ también es diagonal. Ya que las matrices diagonales conmutan, se sigue que $[T]_\gamma [T^*]_\gamma = [T^*]_\gamma [T]_\gamma$. En particular, como este isomorfismo es compatible con la multiplicación de matrices y la composición de transformaciones lineales, concluimos que $TT^* = T^*T$.

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

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

^aNótese que del Teorema 14.3 se sigue que, si V es de dimensión finita, entonces T es un operador normal si y sólo si $[T^*]$ es una matriz normal.

Nótese del párrafo anterior que, bajo las condiciones que impusimos al operador lineal T , tenemos que T y T^* tienen los mismos eigenvectores —aunque no necesariamente con los mismos eigenvalores correspondientes⁴¹. Ahora veremos cómo se relacionan los eigenvalores bajo estas hipótesis.

Teorema 14.4. Sean V un espacio vectorial de dimensión finita n con producto escalar y T un operador lineal normal en V . Entonces, para todo $|v\rangle \in V$ tal que $T(|v\rangle) = \lambda |v\rangle$, tenemos que $T^*(|v\rangle) = \bar{\lambda} |v\rangle$.

Demostración. Ya que T es normal, tenemos que $TT^* = T^*T$. Sea $|v\rangle \in V$ tal que $T(|v\rangle) = \lambda |v\rangle$. Entonces, se sigue que COMPLETAR \square

⁴¹En particular, en este caso T y T^* son simultáneamente diagonalizables.

Parece que ya estamos muy cerca de nuestra meta; sin embargo, observemos lo siguiente: si T es un operador normal en un espacio vectorial (V, K) de dimensión finita con producto escalar y $|g_i\rangle, |g_j\rangle$ son eigenvectores de T con eigenvalores λ_i, λ_j tales que $\lambda_i \neq \lambda_j$, entonces por un lado tenemos que

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

mientras que, por otro lado,

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

Por lo tanto, llegamos a la igualdad

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

Observemos que si $K = \mathbb{C}$, entonces forzosamente $|g_i\rangle$ y $|g_j\rangle$ son ortogonales, de donde se sigue que $E_{\lambda_i}^\perp = \bigoplus_{j \neq i} E_{\lambda_j}$ para todo eigenvalor λ_i de T y, por ende, que T se puede descomponer espectralmente.

15.2. Teorema espectral

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

Teorema 15.algo (Teorema espectral). Sea (V, K) un espacio vectorial de dimensión finita con producto escalar y $T : V \rightarrow V$ un operador lineal con eigenvalores distintos $\lambda_1, \lambda_2, \dots, \lambda_k$. Supongamos que T es normal si $K = \mathbb{C}$ y que T es autoadjunto si $K = \mathbb{R}$. Sea W_i el eigensubespacio de T correspondiente al eigenvalor λ_i , con $1 \leq i \leq k$, y sea T_i la proyección ortogonal de V sobre W_i . Entonces, se cumple que:

- a) $V = W_1 \oplus W_2 \oplus \dots \oplus W_k$;
- b) si $W_i' = \bigoplus_{j \neq i} W_j$, entonces $W_i^\perp = W_i'$;
- c) $T_i T_j = \delta_{ij} T_i$ para $1 \leq i, j \leq k$;
- d) $I = T_1 + T_2 + \dots + T_k$;
- e) $T = \lambda_1 T_1 + \lambda_2 T_2 + \dots + \lambda_k T_k$.

15.3. Ejercicios de repaso

15.3.1. Operadores normales y autoadjuntos

1. Demuestra que las matrices diagonales con entradas en un campo K conmutan.

15.3.2. Teorema espectral

1. Demuestra el teorema espectral.

16. Operadores ortogonales y unitarios***

16.1. Operadores ortogonales y unitarios

16.2. Ejercicios de repaso

16.2.1. Operadores lineales ortogonales y unitarios