

Notas de Álgebra Lineal

Amélie Bernès

Primavera, 2025

BUAP
FCFM

Contents

1	Información general del curso	3
1.1	Forma de evaluación	3
1.2	Algunos conceptos que debes repasar	4
1.3	Bibliografía	4
2	Espacios vectoriales	5
2.1	Algunas estructuras algebraicas	6
2.2	Definición de espacio vectorial	8
2.3	Ejemplos de espacios vectoriales	8
2.3.1	El F -espacio vectorial F^n	8
2.3.2	El F -espacio vectorial F^X de funciones de X en F	9
2.3.3	Espacio de funciones de soporte finito $F^{(X)}$	11
2.4	Propiedades básicas de espacios vectoriales	13
2.5	Subespacios vectoriales	14
2.6	El subespacio generado por un subconjunto y suma de subespacios .	16
2.6.1	Suma directa de subespacios	20
2.7	Caso práctico: Representación de información nutricional	21
2.8	Ejercicios I	24
2.9	Dependencia e independencia lineal	27
2.10	Bases y dimensión	31
2.11	Cómo construir bases de un espacio vectorial	33
2.12	Caso práctico: Interpolación con polinomios de Lagrange	39
2.13	Algunos resultados de dimensión	41
2.14	Ejercicios II	45
2.15	Simulacro de examen, primer parcial	46
3	Transformaciones lineales	48
3.1	Transformaciones lineales	49
3.1.1	Ejemplos de transformaciones lineales	49
3.2	El teorema fundamental de las transformaciones lineales	52
3.2.1	No ejemplos de transformaciones lineales	53
3.3	Inyectividad y suprayectividad de transformaciones lineales	54
3.4	Isomorfismos	56
3.5	La propiedad universal de las bases	58
3.6	Caracterización de transformaciones lineales de \mathbb{R}^n en \mathbb{R}^m	60
3.7	Ejercicios III	63

4	Representaciones matriciales de transformaciones lineales	65
4.1	El espacio $\mathcal{L}(V, W)$	66
4.2	El isomorfismo $[\cdot]_{\beta}$	66
4.3	Representación matricial de una transformación lineal	68
4.4	El isomorfismo Φ_{β}^{γ}	69
4.5	Operaciones entre transformaciones lineales usando sus representaciones matriciales	70
4.6	Transformaciones lineales de la forma L_A	73
4.7	Ejemplos	75
4.8	Cambio de sistema coordenado	75
4.8.1	Matrices similares	79
4.9	Cálculo del rango de transformaciones usando representaciones matriciales	80
4.10	Ejercicios IV	83
4.11	Simulacro de examen, segundo parcial	85
5	Diagonalización	87
5.1	Operadores lineales diagonalizables	88
5.2	El determinante de un operador y su polinomio característico	92
5.3	Ejemplos de diagonalización	96
5.4	Criterios para determinar cuándo un operador es diagonalizable . . .	98

Chapter 1

Información general del curso

Intentaremos acordar al menos una hora a la semana en la que podamos tener asesorías grupales, para poder discutir con los alumnos que así lo deseen las dudas que tengan respecto al curso.

1.1 Forma de evaluación

- **Examen: 70%.** Se realizarán dos exámenes parciales a lo largo del semestre. El promedio de la calificación de estos será el 70% de la calificación final del alumno (i.e. un promedio perfecto de 10 le da 7 puntos de 10 en su calificación final). Por posibles cuestiones de salud no se prohíben las salidas breves al sanitario durante el examen, pero se exhorta a los alumnos a evitar este tipo de situaciones y permanecer en el salón mientras responden su examen. No se permitirá el uso de celulares durante el examen.
- **Tareas: 20%.** Antes de cada parcial, se enviará por correo un ejercicio que el alumno deberá resolver pulcramente para entregar en una fecha acordada. Antes del examen se les regresará su ejercicio calificado. El alumno obtendrá 2 puntos de su calificación final si entrega ambos ejercicios y 1 si entrega uno solo. Se les pide encarecidamente no copiar o usar modelos de lenguaje para hacer sus tareas, ya que el objetivo principal de esta actividad es, de mi parte, dar la mayor retroalimentación posible para ayudarlos a identificar posibles errores en su razonamiento y redacción, en resumen, hacerles comentarios que puedan tomar en cuenta para cuando presenten su examen. Si la tarea contiene errores muy graves de comprensión o calidad esta se anulará.
- **Participación: 10%.** El alumno deberá participar voluntariamente al menos una vez durante el semestre para conseguir 1 punto correspondiente al rubro de participación. Se tomarán en cuenta las participaciones realizadas fuera del horario de clase durante las tutorías.

Las calificaciones reprobatorias (i.e. menores a 6) no se redondean. Si el alumno ha asistido al menos al 60% de las sesiones, tiene derecho a presentar un examen extraordinario al final del semestre. La calificación obtenida en este será la final y definitiva del curso. Si el alumno había aprobado el curso, pero quiere presentar el examen extraordinario para subir su calificación, podrá presentarse para resolverlo, sin embargo, si lo entrega para su revisión, acepta que renuncia

a su calificación aprobatoria inicial, y que su nota final será la que obtenga en el examen extraordinario (aún cuando esta sea reprobatoria).

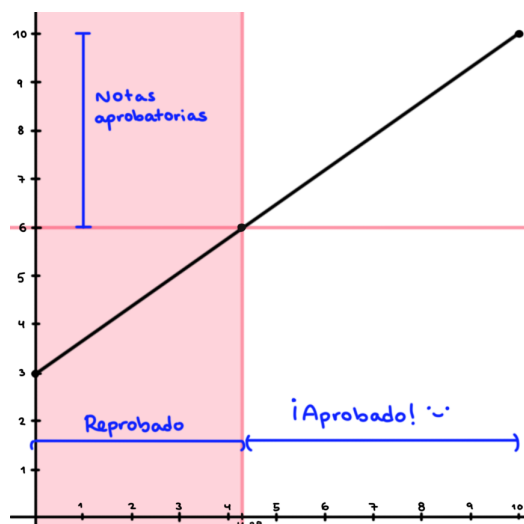


Figure 1.1: En el eje x se marca el promedio de ambos exámenes parciales, y en el eje y la calificación final del curso: la recta muestra la calificación final que obtendrías suponiendo que participaste y entregaste tus dos tareas.

1.2 Algunos conceptos que debes repasar

Se usarán constantemente definiciones que estudiaste en cursos pasados, en particular

- el concepto de función, dominio, codominio, imágenes e imágenes inversas,
- manipulaciones algebraicas de matrices (suma, multiplicación, producto por un número real), matrices simétricas, cuadradas, la traza de una matriz y el rango de esta,
- resolución de sistemas de ecuaciones lineales,
- la definición de unión e intersección de una familia de conjuntos

1.3 Bibliografía

Los textos principales del curso serán [FSH03], [Mej11] y [Axl], pero para algunos temas particulares me apoyaré de los libros [Bro], [Gol] y [God].

Chapter 2

Espacios vectoriales

TODO: usar notación de span para subespacios. No quiero que se confunda eso con el producto punto.

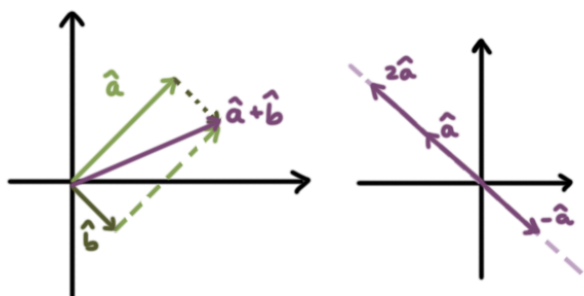
¿Qué entiendes por un *vector*?

En clases previas de física, usabas estos objetos para representar *fuerzas* pues, para hablar de una fuerza, se necesita especificar tanto su *magnitud* como su *dirección*.

Algunos aspectos implícitos eran que

- los vectores eran elementos de \mathbb{R}^2 o de \mathbb{R}^3 ,
- podías “sumar vectores” usando la ley del paralelogramo (por ejemplo, si varias fuerzas actúan en un objeto, las sumabas para obtener la *fuerza neta* ejercida en el objeto),
- además, podías “multiplicar vectores por números reales” para cambiar su magnitud y, dependiendo del signo del número real, mantener la dirección o invertirla (considera, por ejemplo, la segunda Ley de Newton: $\hat{F} = m\hat{a}$).

Por ejemplo, la fuerza de gravedad que ejerce el planeta Tierra en los objetos sobre ella es de $9.8m/s^2$, pero ¿qué dirección tiene?



Estas manipulaciones geométricas además satisfacen ciertas propiedades algebraicas; por ejemplo, dados tres vectores $\hat{a}, \hat{b}, \hat{c} \in \mathbb{R}^2$ y dos escalares $m, n \in \mathbb{R}$, puedes comprobar geométricamente que

$$\begin{aligned}\hat{a} + (\hat{b} + \hat{c}) &= (\hat{a} + \hat{b}) + \hat{c}, \\ m(\hat{a} + \hat{b}) &= m\hat{a} + m\hat{b},\end{aligned}\tag{2.1}$$

o que

$$m(n\hat{a}) = (mn)\hat{a}.$$

Abstrayendo este escenario, tenemos

- un conjunto V de “vectores”, que pueden ser sumados entre sí,
- otro conjunto F de “escalares”, cuyos elementos pueden multiplicarse por elementos de V . En F además se pueden hacer sumas y productos.
- Además, tenemos ciertas relaciones entre las operaciones de F y V (c.f. (2.1)).

2.1 Algunas estructuras algebraicas

Ya comentamos antes que sabíamos sumar vectores en el plano o en el espacio; empecemos pues por abstraer el concepto de operación. Estamos acostumbrados a sumar números reales: dados $a, b \in \mathbb{R}$, conocemos una *regla* para asignarle a la pareja (a, b) un número real llamado su suma, que solemos denotar por $a + b$. También hablamos de la resta de los números a, b . Observe que estas operaciones tienen distintas propiedades, por ejemplo, la suma es asociativa y conmutativa, pero la resta no.

Definición 2.1.1. Dado $X \neq \emptyset$ un conjunto no vacío, una **operación binaria** en X es una función con dominio $X \times X$ y codominio X .

O sea, una operación binaria $*$: $X \times X \longrightarrow X$ es una regla que a cada par ordenado (x_1, x_2) le asigna un único elemento x_3 de X . Por ser este elemento único, puede dársele un nombre; es usual denotar a x_3 por $x_1 * x_2$.

Es decir, con $x_1 * x_2$ estamos abreviando a $*(x_1, x_2)$.

Definición 2.1.2. Dado $X \neq \emptyset$ y $+: X \times X \longrightarrow X$ una operación binaria en X , la pareja $(X, +)$ es un **grupo** si

GR-1) (asociatividad) Para cualesquiera $x_1, x_2, x_3 \in X$,

$$(x_1 + x_2) + x_3 = x_1 + (x_2 + x_3). \quad (2.2)$$

La asociatividad de la operación $*$ significa en la práctica que los paréntesis en (2.2) pueden omitirse.

GR-2) (existencia de elemento neutro) Existe $e \in X$ tal que

$$\forall x \in X : e + x = x = x + e. \quad (2.3)$$

GR-3) (existencia de inversos) Se cumple que

$$\forall x \in X \exists -x \in X : x + (-x) = e = (-x) + x. \quad (2.4)$$

Es fácil demostrar la unicidad del elemento neutro y de los inversos - por eso es legítimo darles un nombre particular a estos en la definición de grupo. Si la operación binaria $+$ se entiende por el contexto, al grupo $(X, +)$ se le denota por X .

Definición 2.1.3. A todo grupo $(X, +)$ que satisfaga que

GR-4) (conmutatividad)

$$\forall x_1, x_2 \in X : x_1 + x_2 = x_2 + x_1 \quad (2.5)$$

se le llamará **abeliano**.

La motivación del concepto de grupo es tener un conjunto con una operación binaria en el que se puedan hacer cancelaciones; si en un grupo $(G, +)$ se tiene una ecuación de la forma

$$a + b = a + c,$$

sumando a ambos lados de la igualdad el inverso de a y usando el que la operación sea asociativa, se tiene que

$$e + b = e + c,$$

luego, que

$$b = c.$$

Ejemplo 2.1.4. Considera al conjunto \mathbb{Z} de los números enteros.

- Si $+$ denota a la suma usual de números enteros, comprueba que $(\mathbb{Z}, +)$ es un grupo.
- Muestra que, si \cdot denota al producto usual de enteros, (\mathbb{Z}, \cdot) no es un grupo.
- Si ahora $+$ y \cdot denotan, respectivamente, la suma y producto de números racionales, demuestra que tanto $(\mathbb{Q}, +)$ como $(\mathbb{Q} - \{0\}, \cdot)$ son grupos abelianos.

Definición 2.1.5. Dado $F \neq \emptyset$ un conjunto no vacío y dos operaciones binarias $+: F \times F \rightarrow F$ y $\cdot: F \times F \rightarrow F$, $(F, +, 0, \cdot, 1)$ con $0 \neq 1$ es un **campo** si

- $(F, +)$ es un grupo abeliano con neutro 0 , es decir, si
 - $\forall a, b, c \in F: a + (b + c) = (a + b) + c,$
 - $\forall a \in F: a + 0 = a = 0 + a,$
 - $(\forall a \in F) (\exists -a \in F): a + (-a) = 0 = (-a) + a,$
 - $\forall a, b \in F: a + b = b + a,$
- $(F - \{0\}, \cdot)$ es un grupo abeliano con neutro 1 , es decir,
 - $\forall a, b, c \in F - \{0\}: a \cdot (b \cdot c) = (a \cdot b) \cdot c$
 - $\forall a \in F - \{0\}: a \cdot 1 = a = 1 \cdot a,$
 - $(\forall a \in F - \{0\}) (\exists a^{-1} \in F - \{0\}): a \cdot a^{-1} = 1 = a^{-1} \cdot a,$
 - $\forall a, b \in F - \{0\}: a \cdot b = b \cdot a.$

- (distributividad) Para cualesquiera $a, b, c \in F$:

$$a \cdot (b + c) = a \cdot b + a \cdot c. \quad (2.6)$$

La distributividad es la propiedad que relaciona las estructuras $(X, +)$ y (X, \cdot) que conforman un campo. Por costumbre, al neutro de $(X, +)$ se le suele llamar el **neutro aditivo** del campo, y al neutro de $(X - \{0\}, \cdot)$ se le llama el **neutro multiplicativo**.

Por tener la ley distributiva, si $x \in X$ cualquiera,

$$x \cdot 0 = x \cdot (0 + 0) = x \cdot 0 + x \cdot 0,$$

luego,

$$x \cdot 0 = 0.$$

Así, no existe $x \in X$ tal que $x \cdot 0 = 1$; por eso se debe quitar el cero para hablar del grupo multiplicativo $(X - \{0\}, \cdot)$.

Este último punto muestra que frases del tipo “el conjunto X es un grupo” son ambiguas. Puede ser posible definir dos operaciones binarias $+$ y \cdot en X de tal forma que tanto $(X, +)$ como (X, \cdot) sean grupos. Los conjuntos subyacentes de estos grupos son los mismos, pero estos no son el mismo grupo.

2.2 Definición de espacio vectorial

Ya tenemos los elementos necesarios para dar una de las definiciones centrales del curso.

Definición 2.2.1. Sean $(V, +)$ un grupo abeliano y F un campo. Si $\cdot : F \times V \longrightarrow V$ es tal que

$$EV-5) \quad (\forall x \in V) : 1_F \cdot x = x,$$

$$EV-6) \quad (\forall a, b \in F)(\forall x \in V) : (ab) \cdot x = a \cdot (b \cdot x),$$

$$EV-7) \quad (\forall a \in F)(\forall x, y \in V) : a \cdot (x + y) = a \cdot x + a \cdot y,$$

$$EV-8) \quad (\forall a, b \in F)(\forall x \in V) : (a + b) \cdot x = a \cdot x + b \cdot x,$$

entonces $(V, +, F, \cdot)$ es un **F -espacio vectorial**.

Recuerda que, el que $(V, +)$ sea grupo abeliano, significa que

$$EV-1) \quad (\forall x, y, z \in V) : x + (y + z) = (x + y) + z,$$

$$EV-2) \quad (\exists 0_V \in V)(\forall x \in V) : 0_V + x = x = x + 0_V,$$

$$EV-3) \quad (\forall x \in V)(\exists -x \in V) : x + (-x) = 0_V = (-x) + x,$$

$$EV-4) \quad (\forall x, y \in V) : x + y = y + x.$$

Por tradición, a los elementos de V se les llama “vectores” y a los de F “escalares”. Nota que los vectores entonces son los elementos del grupo abeliano (cuyos elementos no necesariamente pertenecen a \mathbb{R}^2 o \mathbb{R}^3).

Se permite en la definición de espacio vectorial que F sea cualquier campo pero, a menos que se indique lo contrario, en lo que sigue entenderemos que $F \in \{\mathbb{R}, \mathbb{C}\}$. Si hay riesgo de confusión, al vector cero se le denotará por $\hat{0}$ y al escalar cero se le denota por 0_F .

La función \cdot (que no es una operación binaria, ¿verdad?) es la que conecta las estructuras de grupo abeliano y de campo que conforman a un espacio vectorial.

2.3 Ejemplos de espacios vectoriales

Vamos a dar las construcciones de espacios vectoriales concretos con los que trabajaremos ampliamente en el curso. En la lista de ejercicios 2.8 se dan más ejemplos.

2.3.1 El F -espacio vectorial F^n

Sea $n \geq 1$ entero. Tanto \mathbb{R}^n como \mathbb{C}^n con la suma definida “entrada a entrada” son grupos abelianos:

$$\forall x = (x_k)_{k=1}^n, y = (y_k)_{k=1}^n \in \mathbb{R}^n : x + y := (x_k + y_k)_{k=1}^n. \quad (2.7)$$

De forma similar, fijado un campo, podemos definir multiplicación escalar “entrada a entrada”: si $\cdot : \mathbb{R} \times \mathbb{R}^n \longrightarrow \mathbb{R}^n$ se define como

$$\forall x = (x_k)_{k=1}^n \in \mathbb{R}^n, a \in \mathbb{R} : ax := (ax_k)_{k=1}^n, \quad (2.8)$$

puede comprobarse que \mathbb{R}^n con esta multiplicación escalar es un \mathbb{R} -espacio vectorial. De forma similar se puede definir una \mathbb{C} -multiplicación escalar en \mathbb{C}^n y también una \mathbb{R} -multiplicación escalar en \mathbb{C} . ¿Por qué no tiene sentido cambiar al campo de escalares en (2.8) por \mathbb{C} ?

En general, si F es un campo y se considera el grupo abeliano F^n de n -tuplas con entradas en F y suma definida como en (2.7), con la multiplicación escalar definida como en (2.8), obtenemos un F -espacio vectorial. Haciendo $n = 1$, se tiene que F es también un F -espacio vectorial (la multiplicación por escalares en este contexto es la multiplicación en el campo F).

Ejemplo 2.3.1. *El grupo abeliano \mathbb{R}^n con la suma entrada a entrada con la multiplicación por escalares entrada a entrada es un \mathbb{R} -espacio vectorial. \mathbb{C}^n puede ser visto como \mathbb{C} -espacio vectorial o como \mathbb{R} -espacio vectorial, dependiendo de en qué campo se tomen los escalares.*

2.3.2 El F -espacio vectorial F^X de funciones de X en F

Vamos ahora a definir espacios vectoriales cuyos vectores son funciones; lo hacemos de esta forma porque esta construcción general engloba varios ejemplos de interés práctico, por ejemplo, a los espacios de sucesiones o espacios de funciones continuas - ambos importantes en el análisis - o el espacio de polinomios con coeficientes en un campo.

Sean $X \neq \emptyset$ un conjunto no vacío cualquiera y $(F, +, 0, \cdot, 1)$ un campo. Consideremos al conjunto

$$F^X := \{f : X \longrightarrow F : f \text{ es función}\}$$

de funciones de X en F ; aprovechando la estructura de campo en F , vamos a dotar al conjunto F^X de estructura de F -espacio vectorial.

- **Definición de suma de funciones:** Definimos a la operación binaria suma

$$\hat{+} : F^X \times F^X \longrightarrow F^X \quad (2.9)$$

como sigue: dadas $f, g \in F^X$ cualesquiera, la función $f \hat{+} g \in F^X$ se define puntualmente como

$$(f \hat{+} g)(x) = f(x) + g(x), \quad x \in X. \quad (2.10)$$

Nota que, en la ecuación anterior, el signo $\hat{+}$ que aparece a la izquierda es la operación binaria que estamos definiendo en (2.9), mientras que el signo $+$ de la derecha es la suma del campo F : a partir de la operación suma en F estamos diciendo cómo sumar elementos de F^X .

- **Definición de producto por escalares:** Definimos a un producto por escalares

$$\star : F \times F^X \longrightarrow F^X \quad (2.11)$$

como sigue: si $a \in F$ y $f \in F^X$, la función $a \star f \in F^X$ se define puntualmente como

$$(a \star f)(x) = af(x), \quad x \in X. \quad (2.12)$$

Nota cómo estamos usando la operación producto del campo para definir el producto por escalares.

- **Neutro aditivo:** Proponemos como neutro para la operación suma de funciones (c.f. (2.9)) a la función $\hat{0} : X \longrightarrow F$ definida como sigue:

$$\hat{0}(x) = 0, \quad x \in X, \quad (2.13)$$

es decir, a la función que mapea todo punto de X al neutro aditivo del campo F .

Demostremos que $(F^X, \hat{+}, \star)$ es un F -espacio vectorial.

En lo que sigue, vamos a demostrar igualdades que tienen lugar en F^X , es decir, igualdades entre funciones. Recuerda que dos funciones son iguales si tienen el mismo dominio, codominio y misma regla de correspondencia: puesto que todos los elementos de F^X tienen dominio X y codominio F , para demostrar que $f, g \in F^X$ son iguales basta ver que

$$f(x) = g(x) \quad \text{para toda } x \in X.$$

1. **Asociatividad de $\hat{+}$:** sean $f, g, h \in F^X$. Mostremos que

$$(f \hat{+} g) \hat{+} h = f \hat{+} (g \hat{+} h).$$

Sea $x \in X$. Usando la definición de suma de funciones dada en (2.10), tenemos que

$$\begin{aligned} ((f \hat{+} g) \hat{+} h)(x) &= (f \hat{+} g)(x) + h(x) = (f(x) + g(x)) + h(x) \\ &= f(x) + (g(x) + h(x)) = f(x) + (g \hat{+} h)(x) \\ &= (f \hat{+} (g \hat{+} h))(x). \end{aligned}$$

La tercera igualdad se da porque la suma en F es asociativa.

2. **Elemento neutro para la suma:** Si $f \in F^X$ cualquiera, veamos que

$$\hat{0} \hat{+} f = f = f \hat{+} \hat{0}.$$

Probemos la primera igualdad; la otra se tiene de forma análoga. Usaremos la definición (2.13) de la función cero $\hat{0}$ y la definición de suma de funciones dada en (2.10). Para toda $x \in X$,

$$(\hat{0} \hat{+} f)(x) = \hat{0}(x) + f(x) = 0 + f(x) = f(x).$$

3. **Existencia de inversos aditivos:** **Ejercicio:** Dada una función $f \in F^X$, propón una función $g \in F^X$ que sea el neutro aditivo de f , es decir, una función tal que

$$f \hat{+} g = \hat{0} = g \hat{+} f.$$

Sugerencia: Defínela puntualmente, usa la existencia de inversos aditivos en el campo F . Naturalmente, a tal función g la denotaremos por $-f$.

4. **Conmutatividad de $\hat{+}$:** **Ejercicio:** Demuestra que la suma de funciones es conmutativa, es decir, que dadas cualesquiera $f, g \in F^X$,

$$f \hat{+} g = g \hat{+} f.$$

5. **(EV-5)** Sea $f \in F^X$ cualquiera; es fácil ver que ocurre

$$1 \star f = f$$

pues, según la definición del producto por escalares dada en (2.12), para toda $x \in X$ se tiene que

$$(1 \star f)(x) = 1f(x) = f(x).$$

6. (EV-6) Si $a, b \in F$ y $f \in F^X$, demostremos que

$$(ab) \star f = a \star (b \star f).$$

Sea $x \in X$.

$$((ab) \star f)(x) = abf(x) = a(bf(x)) = a((b \star f)(x)) = (a \star (b \star f))(x).$$

7. (EV-7) **Ejercicio:** Demuestra que, para todo $a \in F$ y cualesquiera $f, g \in F^X$,

$$a \star (f \hat{+} g) = a \star f \hat{+} a \star g.$$

8. (EV-8) **Ejercicio:** Demuestra que, para toda $f \in F^X$ y cualesquiera $a, b \in F$,

$$(a + b) \star f = a \star f \hat{+} b \star f.$$

Ejemplo 2.3.2. Si $\mathcal{C}(\mathbb{R}) = \{f : \mathbb{R} \rightarrow \mathbb{R} : f \text{ es continua}\}$, es claro que $\mathcal{C}(\mathbb{R}) \subseteq \mathbb{R}^{\mathbb{R}}$.

2.3.3 Espacio de funciones de soporte finito $F^{(X)}$

Definición 2.3.3. Sean $X \neq \emptyset$ un conjunto, F un campo. Dada $f : X \rightarrow F$ función (i.e. un elemento de F^X), definimos su **soporte** como el conjunto de los puntos del dominio para los que f no se anula:

$$\text{sop}(f) := \{x \in X \mid f(x) \neq 0\}.$$

Ejemplo 2.3.4. Tomemos a $X = \mathbb{N}$. Acostumbramos llamar a los elementos de $F^{\mathbb{N}}$ (i.e. a las funciones de \mathbb{N} en F) **sucesiones en F** . De hecho, dada una función $f \in F^{\mathbb{N}}$, es usual identificarla con el conjunto de sus imágenes:

$$f = (f_n)_{n \in \mathbb{N}}, \quad \text{donde } f_n := f(n).$$

Ejercicio: Demuestre que, si $f = (f_n)_{n \in \mathbb{N}} \in F^{\mathbb{N}}$ tiene soporte finito, entonces se debe tener $\lim_{n \rightarrow \infty} f(n) = 0$. ¿Se vale la otra implicación? \diamond

Definición 2.3.5. Denotaremos por $F^{(X)}$ al conjunto de las funciones de X en F que tienen soporte finito, es decir,

$$F^{(X)} := \{f \in F^X : |\text{sop}(f)| < \infty\}.$$

Para lo que sigue, necesitamos que el lector recuerde (o al menos esté de acuerdo en aceptar como ciertos) los siguientes hechos:

- La cardinalidad del conjunto vacío es 0. El conjunto vacío es el único conjunto de cardinalidad 0.
- Si A y B son dos conjuntos finitos, entonces $A \cup B$ es también finito.
- Si A está contenido en B y B es finito, entonces A también lo es.

Proposición 2.3.6. Si $X \neq \emptyset$ y F es un campo, entonces $F^{(X)}$ es un subespacio de F^X .

Demostración. Usemos los criterios dados en la proposición (\star -sub) para mostrar que $F^{(X)} \leq F^X$.

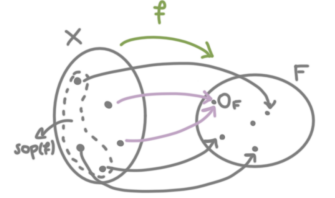


Figure 2.1: Representación gráfica del soporte de una función.

1. Si $\hat{0} : X \rightarrow F$ es la función cero,

$$\text{sop}(\hat{0}) = \{x \in X \mid \hat{0}(x) \neq 0\} = \emptyset,$$

luego, $\hat{0}$ tiene soporte finito, o sea, $\hat{0} \in F^{(X)}$.

2. Sean $f, g \in F^{(X)}$; veamos que $f + g \in F^{(X)}$. Es decir, usando que f y g tienen soporte finito, mostremos que $f + g$ también tiene soporte finito. Si $\text{sop}(f + g) = \emptyset$, acabamos. Supongamos ahora que $\text{sop}(f + g) \neq \emptyset$. Note que, si $x \in \text{sop}(f + g)$, entonces

$$f(x) + g(x) = (f + g)(x) \neq 0,$$

luego, no ocurre $f(x) = 0 = g(x)$, es decir,

$$x \notin (\text{sop}(f))^c \cap (\text{sop}(g))^c,$$

o sea,

$$x \in \text{sop}(f) \cup \text{sop}(g).$$

Con esto demostramos la contención

$$\text{sop}(f + g) \subseteq \text{sop}(f) \cup \text{sop}(g).$$

Como $\text{sop}(f)$ y $\text{sop}(g)$ son, por hipótesis, ambos finitos, su unión también lo es, luego $\text{sop}(f + g)$, por ser subconjunto de un conjunto finito, es finito.

3. Sean $f \in F^{(x)}$ y $a \in F$; probemos que $af \in F^{(X)}$. Si $a = 0$, $af = \hat{0} \in F^{(X)}$; supongamos ahora $a \neq 0$. Entonces,

$$\begin{aligned} x \in \text{sop}(af) &\Leftrightarrow af(x) \neq 0 \\ &\Leftrightarrow a \neq 0 \text{ y } f(x) \neq 0 \\ &\Leftrightarrow f(x) \neq 0 \Leftrightarrow x \in \text{sop}(f). \end{aligned}$$

Así, $\text{sop}(af) = \text{sop}(f)$, luego, af tiene, al igual que f , soporte finito.

□

Nota 2.3.7. Por un polinomio, por lo general se entiende una función $f : F \rightarrow F$ de la forma

$$f(x) = a_0 + a_1x + \dots + a_nx^n, \quad n \in \mathbb{N}, a_k \in F, \quad (2.14)$$

con los “coeficientes” $a_k \in F$ y la “incógnita” x , que puede tomar valores en el campo F .

La construcción formal usual de polinomios con coeficientes en un campo F consiste justamente en considerar al espacio vectorial $F^{(\mathbb{N})}$ de sucesiones en el campo F de soporte finito, pues, la propiedad que se intenta capturar es que dos expresiones de la forma $a_0 + a_1x + \dots + a_nx^n$ y $b_0 + b_1x + \dots + b_mx^m$ sean iguales si y sólo si $a_i = b_i$ para toda i (si $n < m$, claro que se supone $a_i = 0$ para $i > n$). Esto claro está se consigue si se identifica al polinomio en x (2.14) con la sucesión

$$f = (a_0, a_1, \dots, a_n, 0, 0, 0, \dots) \in F^{(\mathbb{N})}.$$

Los detalles de la construcción pueden consultarse en [Jac85], [Rot15].

Recuerda que, en un campo, el producto de dos elementos es cero si y sólo si al menos uno de los dos es cero.

En realidad, basta con partir de un anillo R - que no necesariamente sea campo - para construir al anillo de polinomios con coeficientes en R .

2.4 Propiedades básicas de espacios vectoriales

Deducimos a continuación resultados inmediatos de la definición 2.2.1.

Proposición 2.4.1. *Sea V un F -espacio vectorial.*

- Para todo $v \in V$, $0_F \cdot v = \hat{0}$.
- Para todo $a \in F$, $a \cdot \hat{0} = \hat{0}$.
- Para todo $v \in V$, $(-1) \cdot v = -v$.

Demostración.

- Sólo observe que

$$0_F \cdot v + 0_F \cdot v = (0_F + 0_F) \cdot v = 0_F \cdot v = 0_F \cdot v + \hat{0},$$

note que esta es una ecuación en el grupo abeliano V , luego, podemos cancelar el elemento $0_F \cdot v$ a ambos lados de la igualdad, para llegar a que $0_F \cdot v = \hat{0}$.

- **Ejercicio.**
- El elemento $(-1) \cdot v \in V$ es tal que

$$(-1) \cdot v + v = (-1) \cdot v + 1 \cdot v = (-1 + 1) \cdot v = 0_F \cdot v = \hat{0},$$

luego, por unicidad del inverso aditivo en V , concluimos que

$$(-1) \cdot v = -v.$$

□

¿Verdadero o falso?

- “Todo espacio vectorial contiene un vector cero”
- “Un espacio vectorial puede contener más de un vector cero”
- “En un F -espacio vectorial V , si $a, b \in F$ y $x \in V$, entonces la ecuación $ax = bx$ implica $a = b$ ”
- “En un F -espacio vectorial V , si $a \in F$ y $x, y \in V$, entonces la ecuación $ax = ay$ implica $x = y$ ”

Proposición 2.4.2. *Sea V un F -espacio vectorial.*

- Si $x \in V - \{0\}$, entonces, para cualesquiera escalares $a, b \in F$, $ax = bx$ implica $a = b$.
- Si $a \in F - \{0\}$, entonces, para cualesquiera vectores $x, y \in V$, $ax = ay$ implica $x = y$.

Demostración.

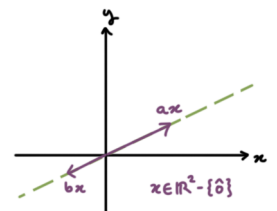


Figure 2.2: Ilustración del primer punto de la Proposición 2.4.2 para el caso $V = \mathbb{R}^2$.

- Sea x un vector no cero, y supongamos que existen escalares a, b distintos entre sí tales que $ax = bx$. Entonces, si $c := a - b \neq 0$,

$$cx = (a - b)x = ax - bx = \hat{0}.$$

Como c es un elemento no cero del campo F , tiene inverso multiplicativo c^{-1} . Así,

$$x = 1_F \cdot x = (c^{-1}c) \cdot x = c^{-1} \cdot (c \cdot x) = c^{-1} \cdot \hat{0} = \hat{0},$$

contradiciendo la hipótesis.

- **Ejercicio.**

□

La versión de la Proposición 2.4.2 para el \mathbb{R} -espacio vectorial \mathbb{R}^2 sería

Si $x = (x_1, x_2) \in \mathbb{R}^2 - \{(0, 0)\}$ y $a, b \in \mathbb{R}$ tales que $ax = bx$, entonces $a = b$.

Podemos demostrar esta afirmación directamente; como x no es el vector cero, sin pérdida de generalidad $x_1 \neq 0$. De la hipótesis

$$(ax_1, ax_2) = ax = bx = (bx_1, bx_2)$$

se tiene la igualdad en \mathbb{R} $ax_1 = bx_1$. Como $x_1 \neq 0$, existe su inverso multiplicativo; multiplicando por este a ambos lados de la igualdad, se llega a que $a = b$, como se quería.

Nota que la diferencia entre este argumento y el de la demostración de la Proposición 2.4.2 es que este primero hace referencia a la forma de los elementos del espacio vectorial \mathbb{R}^2 , mientras que el segundo sólo usa los axiomas de campo y espacio vectorial.

Ejemplo 2.4.3. El espacio vectorial trivial: considere al grupo abeliano $V = \{0\}$ - con la suma definida de la única forma posible. Sea F un campo cualquiera. La única forma de definir una función $\cdot : F \times V \longrightarrow V$ es por

$$\forall a \in F : a \cdot 0 = 0.$$

Puede comprobar que, con este producto escalar, V es un F -espacio vectorial, al que solemos denominar el **espacio vectorial cero** o el **espacio trivial**.

2.5 Subespacios vectoriales

En muchas ocasiones, dado V un F -espacio vectorial, nos interesará trabajar con subestructuras de V que “con la misma suma y multiplicación por escalares”, sean en sí mismas espacios vectoriales.

Definición 2.5.1. Sea $(V, +, 0_V, F, \cdot : F \times V \longrightarrow V)$ un espacio vectorial. Un subconjunto no vacío $\emptyset \neq W \subseteq V$ es un **subespacio vectorial** de V si

$$(W, +|_{W \times W}, 0_V, F, \cdot|_{F \times W} : F \times W \longrightarrow W)$$

es un espacio vectorial. En este caso, escribimos $W \leq V$.

Si $f : A \longrightarrow B$ es una función y C es un subconjunto de A , por $f|_C$ denotamos a la **restricción de f a C** .

En la definición de subespacio, nótese que se está suponiendo que las restricciones $+_{W \times W}$ y $\cdot_{F \times W}$ tengan ambas como codominio a W - pues se están usando para dotar a W de estructura de espacio vectorial.

Si quisieramos usar la Definición 2.5.1 para ver si un subconjunto W de V es subespacio, nota que las propiedades $EV5 - EV8$ se cumplen todas “por herencia de V ” - si son ciertas para todo elemento de V , por supuesto que lo son para todo elemento del subconjunto W de V . Entonces, W es subespacio de V si y sólo si los codominios de $+_{W \times W}$ y $\cdot_{F \times W}$ son W y $(W, +_{W \times W})$ es un grupo abeliano, es decir, si

- Para cualesquiera $x, y \in W$: $x + y \in W$
- Para cualesquiera $x \in W$, $a \in F$: $ax \in W$
- $0_V \in W$
- Para todo $x \in W$, también $-x \in W$.

De hecho, el cumplir estas primeras tres condiciones es necesario y suficiente para ser subespacio.

Teorema 2.5.2. *Sea V un F -espacio vectorial, W un subconjunto de V . Entonces, W es subespacio de V si y sólo si*

- $\hat{0} \in W$
- (Cerradura bajo la suma) Para cualesquiera $x, y \in W$: $x + y \in W$
- (Cerradura bajo multiplicación escalar) Para cualesquiera $x \in W$, $a \in F$: $ax \in W$

Demostración.

\Rightarrow) Trivial

\Leftarrow Por la discusión de arriba, basta ver que, para todo $x \in W$, también $-x \in W$, pero esto por la cerradura bajo multiplicación por escalares y la Proposición 2.4.1 se cumple, pues

$$-x = (-1_F) \cdot x \in W.$$

□

Ejemplo 2.5.3. *Dados $a < b$ números reales, considere al \mathbb{R} -espacio vectorial de funciones $\mathbb{R}^{[a,b]}$. Sea*

$$\mathcal{C}(a, b) := \{f \in \mathbb{R}^{[a,b]} : f \text{ es continua en } [a, b]\}.$$

Se demuestra que

- la función constante cero $\hat{0}_{[a,b]}$ es una función continua,
- la suma de funciones continuas es continua, y
- la multiplicación de una función continua por un escalar resulta en una función continua,

El Teorema 2.5.2 es el que en la práctica se usa para determinar si $\emptyset \neq W \subseteq V$ es subespacio o no.

luego, $\mathcal{C}(a, b)$ es un subespacio de $\mathbb{R}^{[a, b]}$.

Muchas veces, si se quiere probar que W es espacio vectorial, resulta más sencillo verlo como subconjunto de un espacio vectorial y mostrar que es subespacio de este, pues lo primero requiere comprobar tres condiciones (c.f. Teorema 2.5.2), mientras que lo segundo requiere comprobar ocho condiciones (c.f. Definición 2.2.1).

Ejercicio: considere al conjunto

$$U = \{(x, x, 0, y) \in \mathbb{R}^4 : x, y \in \mathbb{R}\}.$$

Si la suma y multiplicación por escalares se definen entrada a entrada, demuestre que U es un \mathbb{R} -espacio vectorial.

Ejercicio: sea ahora

$$U' = \{(x, x, 1, y) \in \mathbb{R}^4 : x, y \in \mathbb{R}\}.$$

¿Es U' subespacio de \mathbb{R}^4 ?

2.6 El subespacio generado por un subconjunto y suma de subespacios

Dado V un F -espacio vectorial, sea $\{W_\alpha\}_{\alpha \in I}$ una familia no vacía de subespacios de V . A partir de esta familia, ¿cómo podemos construir otro subespacio de V ? Consideremos a la unión e intersección de la familia, o sea, a los conjuntos

$$\bigcap_{\alpha \in I} W_\alpha := \{x \in V \mid x \in W_\alpha \text{ para toda } \alpha \in I\} \subseteq V, \quad (2.15)$$

y

$$\bigcup_{\alpha \in I} W_\alpha := \{x \in V \mid x \in W_\alpha \text{ para alguna } \alpha \in I\} \subseteq V. \quad (2.16)$$

¿Son estos subespacios de V ? Tenemos una respuesta afirmativa y otra negativa.

Proposición 2.6.1. *La intersección (2.15) de la familia de subespacios $\{W_\alpha\}_{\alpha \in I}$ de V es un subespacio de V .*

Demostración.

- i) Puesto que el cero es elemento de *toda* subespacio de V , claro que $0 \in W_\alpha$ para toda $\alpha \in I$, luego, $0 \in \bigcap_{\alpha \in I} W_\alpha$
- ii, iii) Sean $x, y \in \bigcap_{\alpha \in I} W_\alpha$, $a \in F$; claro que $x + y, ax \in \bigcap_{\alpha \in I} W_\alpha$ pues, dado $\alpha \in I$ un índice cualquiera, $x, y \in W_\alpha$, luego, como W_α es subespacio de V , $x + y, ax \in W_\alpha$.

□

La demostración anterior fue muy sencilla pues, para que un vector sea elemento de la intersección $\bigcap_{\alpha \in I} W_\alpha$, debe estar en *todos* los elementos de la familia a la vez, y todos los W_α cumplen los tres puntos de la proposición (*-sub). La condición que debe cumplir un vector para estar en la unión de la familia es mucho más laxa; es necesario y suficiente que sea elemento de un solo miembro de la familia $\{W_\alpha\}_{\alpha \in I}$. Dados $x, y \in \bigcup_{\alpha \in I} W_\alpha$, existen $\alpha_1, \alpha_2 \in I$ tales que $x \in W_{\alpha_1}$ y $y \in W_{\alpha_2}$. Nota que *no podemos decir que x y y son elementos de un mismo subespacio*, luego, esta información no parece implicar que la suma sea elemento de la unión.

Ejemplo 2.6.2. (Que muestra que la unión de subespacios puede no ser un subespacio). Sea $(a, b) \subseteq \mathbb{R}$ un intervalo abierto cualquiera. Considere al \mathbb{R} -espacio vectorial

$$\mathbb{R}^{(a,b)} = \{f : (a, b) \longrightarrow \mathbb{R} \mid f \text{ es función.}\}$$

Sean

$$\mathcal{B}_{(a,b)} = \{f \in \mathbb{R}^{(a,b)} \mid f \text{ es acotada}\}, \quad \mathcal{C}_{(a,b)} = \{f \in \mathbb{R}^{(a,b)} \mid f \text{ es continua}\}.$$

Como sabes por tus cursos de cálculo, estos son subespacios de $\mathbb{R}^{(a,b)}$. Claro que ninguno de ellos está contenido en el otro (hay funciones definidas en (a, b) que son continuas pero no acotadas, así como acotadas pero no continuas). Nota que

- $\hat{0} \in \mathcal{B}_{(a,b)} \cup \mathcal{C}_{(a,b)}$, y que
- si $\lambda \in \mathbb{R}$ y $f \in \mathcal{B}_{(a,b)} \cup \mathcal{C}_{(a,b)}$, entonces también $\lambda f \in \mathcal{B}_{(a,b)} \cup \mathcal{C}_{(a,b)}$,

sin embargo, $\mathcal{B}_{(a,b)} \cup \mathcal{C}_{(a,b)}$ **no** es un subespacio de $\mathbb{R}^{(a,b)}$, pues no es cerrado bajo la suma. En la imagen se muestran la gráfica de una función continua f y una acotada g en (a, b) para las que la suma no es ni continua ni acotada (i.e. tales que $f, g \in \mathcal{B}_{(a,b)} \cup \mathcal{C}_{(a,b)}$ pero $f + g \notin \mathcal{B}_{(a,b)} \cup \mathcal{C}_{(a,b)}$).

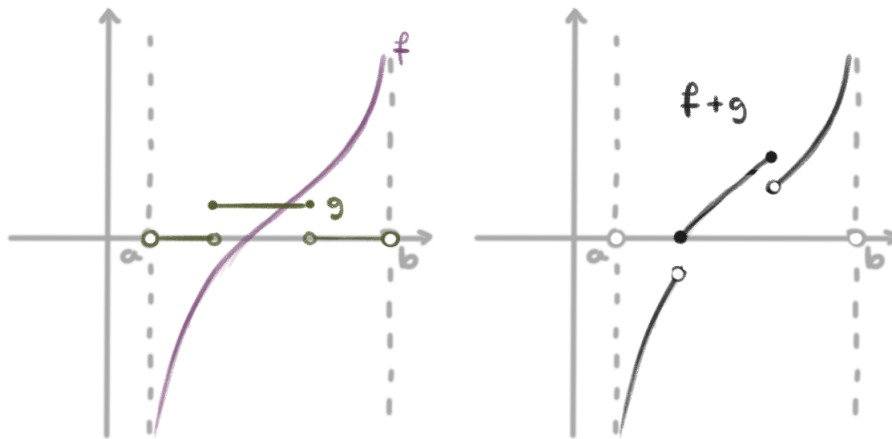


Figure 2.3: Con el sumando continuo rompemos la acotación, y con el sumando acotado, la continuidad.

◇

Ejemplo 2.6.3. Ahora considera a los espacios de funciones $\mathbb{R}^{[a,b]}$, $\mathcal{B}_{[a,b]}$, $\mathcal{C}_{[a,b]}$. Demostraste en tu curso de cálculo de una variable que toda función con dominio **cerrado y acotado** que sea continua, es también acotada (c.f. *cita Spivak, p. 178*). Entonces,

$$\mathcal{C}_{[a,b]} \subseteq \mathcal{B}_{[a,b]},$$

luego, la unión

$$\mathcal{C}_{[a,b]} \cup \mathcal{B}_{[a,b]} = \mathcal{B}_{[a,b]}$$

trivialmente es un subespacio de $\mathbb{R}^{[a,b]}$.

Ejercicio: Si $\{W_\alpha\}_{\alpha \in I}$ es una colección de subespacios de V tal que

$$(\forall i, j \in I)(\exists k \in I) : W_i, W_j \subseteq W_k,$$

entonces $\bigcup_{\alpha \in I} W_\alpha$ es un subespacio de V .

Los Ejemplos 2.6.2 y 2.6.3 nos hacen ver que aspectos propios del dominio X común a las funciones que conforman el espacio vectorial (en particular, su topología), afectan la forma que tengan estos espacios.

Ahora, dado V un F -espacio vectorial, a veces nos interesará limitarnos a trabajar no con todo el espacio V , sino con un subconjunto X de este. Como también queremos aprovechar la estructura algebraica, nos interesará que tal colección X sea, más que subconjunto, subespacio de V . Esto puede ocurrir o no: lo que siempre podemos hacer es considerar a la familia $\{W \leq V : X \subseteq W\}$ de subespacios más pequeños (en el sentido que están propiamente contenidos en V) que contienen a nuestro conjunto de interés de X , y “condensar” a esta familia via su intersección.

Definición 2.6.4. Sean V un F -espacio vectorial, $X \subseteq V$. El subespacio

$$\text{span}(X) := \bigcap \{W \leq V : X \subseteq W\}, \quad (2.17)$$

o sea, la intersección de la familia de todos los subespacios de V que contienen a X , es llamado el **subespacio de V generado por X** .

Proposición 2.6.5. $\text{span}(X)$ es el menor subespacio de V que contiene a X , es decir,

$$\forall W \leq V : X \subseteq W \Rightarrow \text{span}(X) \leq W.$$

Demostración. Es clara pues, $\text{span}(X)$, al ser por definición la intersección de la familia de subespacios que contienen a X , está contenido en cada uno de sus integrantes.

Estamos considerando a la contención de conjuntos como orden; el que un conjunto A sea menor que un conjunto B significa que A está contenido en B .

□

En resumen: si X no es subespacio, siempre podemos considerar a $\text{span}(X)$, el menor subespacio de V que contiene al conjunto de interés.

Ejercicio: Demuestre que, si $X \leq V$, entonces $X = \text{span}(X)$.

Observación 2.6.6. Por definición, $\text{span}(\emptyset)$ es la intersección de la familia de subespacios que contienen a \emptyset - o sea, es la intersección de todos los subespacios de V . Como $\{0\}$ es el menor subespacio de V , concluimos que

$$\text{span}(\emptyset) = \{0\}.$$

En (2.17) decimos cómo construir teóricamente a tal $\text{span}(X)$; el objetivo ahora es dar una descripción completa de sus elementos, es decir, dar un criterio concreto en base al cual determinar cuándo un vector x del espacio pertenece a $\text{span}(X)$.

Según el Teorema 2.5.2, X puede no ser subespacio por cumplirse al menos una de las tres razones siguientes:

- $0 \notin X$; corregimos esto agregando al vector cero.
- Existen $x, y \in X$ tales que $x + y \notin X$; esto se puede arreglar agregando sumas de elementos de X .
- Existen $a \in F$ y $x \in X$ para los que $ax \notin X$. Para que esto no ocurra, podemos agregar los múltiplos escalares de los elementos de X .

Parece que el problema de no ser subespacio se arregla si agregamos sumas finitas de múltiplos escalares de elementos de X (pues así estamos forzando a que los tres puntos del Teorema 2.5.2 ocurran). Demos un nombre a tales sumas de elementos de X .

Definición 2.6.7. Sea V un F -espacio vectorial, $X \subseteq V$. Una **combinación lineal** en V de elementos de X es cualquier vector de la forma

$$\sum_{i=1}^n a_i x_i, \quad \text{con } n \geq 1 \text{ entero, } x_i \in X, a_i \in F \text{ para toda } 1 \leq i \leq n.$$

Según la discusión anterior, parece que para “extender” a X a un subespacio de la forma mínima, hay que agregar todas las combinaciones lineales de elementos de X , pues con esto parece quedar asegurada la cerradura bajo suma y multiplicación por escalares. Confirmamos nuestras sospechas con el siguiente resultado.

Proposición 2.6.8. Si $\emptyset \neq X \subseteq V$, entonces $\text{span}(X)$ consta de todas las combinaciones lineales de elementos de X , es decir,

$$\text{span}(X) = \left\{ \sum_{i=1}^n a_i x_i \mid n \in \mathbb{N}, x_i \in X, a_i \in F \text{ para toda } 1 \leq i \leq n \right\} \quad (2.18)$$

Demostración. Sea

$$\mathcal{A} = \left\{ \sum_{i=1}^n a_i x_i \mid n \in \mathbb{N}, x_i \in X, a_i \in F \text{ para toda } 1 \leq i \leq n \right\};$$

veamos que $\mathcal{A} = \langle X \rangle$.

⊇] Claro que \mathcal{A} es subespacio de V (**Ejercicio:** compruebe los detalles!). Además, \mathcal{A} contiene a X (pues $x = \sum_{i=1}^1 1_F x \in \mathcal{A}$). De esto se deduce que $\text{span}(X) \subseteq \mathcal{A}$.

⊆] Sea ahora $a_1 x_1 + \dots + a_n x_n$ un elemento genérico de \mathcal{A} ; mostremos que este es elemento de *cualquier* subespacio W de V que contenga a X - de esto podremos concluir la contención deseada, pues $\text{span}(X)$ es la intersección de tales subespacios W . Sea pues $W \leq V$ con $X \subseteq W$. Como cada x_i es elemento de X , todos serán elementos de W , luego, como W es subespacio de V , es cerrado bajo multiplicaciones escalares y sumas, por lo tanto, $a_1 x_1 + \dots + a_n x_n \in W$.

□

Definición 2.6.9. Si $\{W_\alpha\}_{\alpha \in I}$ es una familia de subespacios de V , definimos a su suma como el subespacio generado por su unión, es decir,

$$\begin{aligned} \sum_{\alpha \in I} W_\alpha &:= \text{span} \left(\bigcup_{\alpha \in I} W_\alpha \right) \\ &= \left\{ a_1 w_1 + \dots + a_n w_n \mid n \geq 1, a_i \in F, w_i \in \bigcup_{\alpha \in I} W_\alpha \right\} \\ &= \{ a_1 w_1 + \dots + a_n w_n \mid n \geq 1, a_i \in F, w_i \in W_\alpha \text{ para alguna } \alpha \in I \}. \end{aligned} \quad (2.19)$$

Ejercicio: Demuestre que una forma equivalente de escribir al subespacio (2.19) es como

$$\sum_{\alpha \in I} W_\alpha = \{ w_1 + \dots + w_n : n \geq 1, w_i \in W_i \forall 1 \leq i \leq n \}. \quad (2.20)$$

Recapitulemos: dados $W_1, W_2 \leq V$,

- $W_1 \cap W_2$ es un subespacio de V ,

- $W_1 \cup W_2$ es subespacio si y sólo si $W_1 \subseteq W_2$ o bien $W_2 \subseteq W_1$. Nota que el resultado de esta operación no da lugar a un nuevo subespacio de V .
- El menor subespacio de V que contiene a W_1 y W_2 es su suma, o sea,

$$W_1 + W_2 = \{w_1 + w_2 \mid w_1 \in W_1, w_2 \in W_2\}.$$

Definición 2.6.10. Si $X \subseteq V$ es tal que $\text{span}(X) = V$, decimos que X **genera al espacio** V .

Es decir, X genera a V si no hay subespacios propios de V que contengan a X ; esto, según la proposición 2.6.8, significa que *todo elemento de V puede expresarse como combinación lineal finita de elementos de X .*

2.6.1 Suma directa de subespacios

Sean V un F -espacio vectorial, $\{W_i\}_{i=1}^n$ una familia finita de subespacios de V . Por la proposición 2.6.8 sabemos que *todo elemento de su suma es de la forma $w_1 + \dots + w_n$, con $W_i \in W_i$ para $1 \leq i \leq n$.*

Ejemplo 2.6.11. Sea el \mathbb{R} -espacio vectorial \mathbb{R}^3 . Considere a los subespacios

$$W_1 = \{(a, b, 0) \mid a, b \in \mathbb{R}\} \leq \mathbb{R}^3,$$

$$W_2 = \{(0, c, d) \mid c, d \in \mathbb{R}\} \leq \mathbb{R}^3.$$

Es claro que $\mathbb{R}^3 = W_1 + W_2$; de hecho, dado $(x, y, z) \in \mathbb{R}^3$ un vector cualquiera del espacio, para toda $\alpha \in \mathbb{R}$ se tiene que

$$(x, y, z) = (x, \alpha y, 0) + (0, (1 - \alpha)y, z). \quad (2.21)$$

Esto muestra que hay una infinidad de formas de expresar a un vector $(x, y, z) \in \mathbb{R}^3$ como suma de elementos de $W_1 \cup W_2$.

Considere ahora a los subespacios

$$V_1 = \{(a, 0, 0) \mid a \in \mathbb{R}\},$$

$$V_2 = \{(0, b, 0) \mid b \in \mathbb{R}\},$$

$$V_3 = \{(0, 0, c) \mid c \in \mathbb{R}\}.$$

Claro que $\mathbb{R}^3 = V_1 + V_2 + V_3$; dado $(x, y, z) \in \mathbb{R}^3$ cualquiera,

$$(x, y, z) = (x, 0, 0) + (0, y, 0) + (0, 0, z).$$

Una representación de esta forma es única. Por ejemplo, el que las primeras entradas de los últimos dos sumandos deban ser cero (por definición de V_1 y V_2) fuerza que la primera entrada del primer sumando sea x . \diamond

Definición 2.6.12. Sea $\{W_i\}_{i=1}^n$ una familia finita de subespacios de V . El subespacio suma $\sum_{i=1}^n W_i$ es llamada una **suma directa** si cada uno de sus elementos tiene una única representación de la forma

$$w_1 + \dots + w_n, \quad \text{con } w_i \in W_i.$$

En este caso, al subespacio $\sum_{i=1}^n W_i$ se le denota por

$$W_1 \oplus \dots \oplus W_n.$$

Las entradas cero de los elementos de W_1 y W_2 determinan las entradas x y z en la ecuación (2.21), pero en las entradas centrales tenemos libertad de elección -esto se ve reflejado por el parámetro α .

El símbolo \oplus es usado para representar una propiedad del espacio suma $\sum_{i=1}^n W_i$, a saber, la unicidad de las representaciones de sus elementos como sumas de vectores en $\bigcup_{i=1}^n W_i$.

Proposición 2.6.13. Sean U, W subespacios de V . La suma $U + W$ es directa si y sólo si $U \cap W = \{\hat{0}\}$.

Demostración.

\Rightarrow] Si existe $x \in U \cap W$ con $x \neq \hat{0}$, entonces,

$$x = \hat{0} + x, \quad \text{con } \hat{0} \in U, x \in W,$$

y

$$x = x + \hat{0}, \quad \text{con } x \in U, \hat{0} \in W,$$

luego, la suma $U + W$ no es directa.

\Leftarrow] Sea $x \in U + W$; sean $u_1, u_2 \in U$ y $w_1, w_2 \in W$ tales que

$$u_1 + w_1 = x = u_2 + w_2.$$

Entonces,

$$u_1 - u_2 = w_2 - w_1 \in U \cap W = \{0\},$$

o sea, $u_1 = u_2$ y $w_1 = w_2$.

□

Note lo siguiente: según el Teorema 2.5.2, todos los subespacios tienen en común al vector cero $\hat{0}$ del espacio, luego, no tiene sentido buscar subespacios disjuntos. Lo más disimiles que pueden llegar a ser $U, W \leq V$ es que su intersección sea $\{0\}$. Esto, según la Proposición 2.6.13, equivale a que la suma $U + W$ sea directa.

Suma de subespacios \sim Unión de subconjuntos

Suma directa de subespacios \sim Unión disjunta de subconjuntos

Note que en la Proposición 2.6.13 estamos lidiando sólo con dos subespacios. ¿Podemos generalizar este resultado para la suma de n -subespacios? ¿Será cierto que la suma $\sum_{i=1}^n W_i$ es directa si y sólo si la intersección $\bigcap_{i=1}^n W_i$ es $\{\hat{0}\}$? La respuesta es **no** (c.f. ejercicio 10). **Poner la teoría general de suma directa? La necesito para el tema de diagonalización.**

2.7 Caso práctico: Representación de información nutricional

Veamos cómo los conceptos introducidos pueden ser de utilidad para modelar y resolver una problemática de la vida real.

Supongamos que a una nutrióloga le interesa la ingesta de

- calorías
- agua
- hidratos de carbono
- fibra
- grasas, y

- proteínas

que un paciente puede obtener al consumir frutas de cierto grupo de interés (por ejemplo, las frutas que puede encontrar a precio accesible en su entorno). Sus datos iniciales son los valores de dichos componentes presentes en las frutas de interés, datos presentados en forma de tabla como se muestra a continuación.

ENERGÍA Y COMPONENTES MAYORITARIOS DE ALGUNAS FRUTAS DE CONSUMO FRECUENTE EN ESPAÑA						
Fruta	Energía (kcal)	Agua (g)	Hidratos de carbono (g)	Fibra dietética (g)	Grasas (g)	Proteínas (g)
Aguacate	134	79	1,3	2,4	13,8	1,3
Albaricoque	40	86	9,5	2,1	0,1	0,8
Cerezas	58	82	13,5	1,5	0,5	0,8
Ciruela	45	84	11	2,1	0,15	0,6
Fresas	34	88	7	2,2	0,5	0,7
Kiwi	53	83	12,1	1,5	0,44	1
Limón	39	87	9	1	0,3	0,7
Mandarina	39	86	9	1,9	0,19	0,8
Manzana (Golden)	41	85	10,5	2,3	Trazas	0,3
Melocotón	37	86	9	2,3	0,1	0,6
Melón	37	88	8,4	0,8	0,28	0,9
Naranja	38	86	8,6	2,3	0,3	0,8
Pera	48	84	11,7	2,2	0,3	0,4
Piña	46	85	11,5	1,2	0,1	0,5
Plátano	85	73	20,8	2,5	0,27	1,2
Sandía	22	93	4,5	0,3	0,3	0,5
Uvas (blancas)	63	81	16,1	0,9	Trazas	0,6
Uvas (negras)	67	81	15,5	0,4	0,7	0,6

Figure 2.4: Fuente: gruporem-ucam.com

Algunos de sus pacientes son alérgicos a las frutas de su lista; digamos que su nuevo paciente no puede comer manzanas. La nutrióloga se pregunta si es posible sustituir los nutrientes aportados por una manzana con otras frutas de la lista y, si la respuesta es afirmativa, quisiera dar explícitamente combinaciones de frutas que aporten lo mismo que aportaba una cierta cantidad de manzanas (que no pueden ser incluidas en la dieta del paciente). ¿Cómo podemos modelar esta situación?

Podemos representar nuestros datos iniciales como vectores de \mathbb{R}^6 ; por ejemplo, el vector nutricional de una manzana sería

$$x_{manzana} = (41, 85, 10.5, 2.3, 0, 0.3) \in \mathbb{R}^6.$$

Definamos a \mathcal{A} como el subconjunto de \mathbb{R}^6 que consta de todos los vectores fruta,

$$\mathcal{A} = \{x_{aguacate}, x_{albaricoque}, \dots, x_{uvaB}, x_{uvaN}\}.$$

Nota que, en este contexto, las combinaciones lineales con coeficientes naturales tienen una interpretación física; por ejemplo, el vector

$$2x_{pera} + 5x_{fresa} + 3x_{cereza}$$

representa la información nutrimental de 2 peras, 5 fresas y 3 cerezas (luego, la cantidad de calorías y componentes que alguien estaría obteniendo al consumir esta cantidad de frutas).

El problema de la nutrióloga queda expresado en términos de álgebra lineal como sigue: ¿es $x_{manzana}$ elemento del subespacio de \mathbb{R}^6 generado por los vectores $x_{aguacate}, x_{albaricoque}, \dots, x_{uvaB}, x_{uvaN}$? Nota que, implícitamente, el problema nos permite limitarnos al conjunto de las combinaciones lineales finitas de los vectores fruta (menos el vector manzana) y no considerar a todo \mathbb{R}^6 (para este problema, no nos interesan vectores que no sean combinaciones lineales de vectores fruta). La teoría estudiada antes nos asegura que

- Tal conjunto de combinaciones de vectores fruta es un subespacio de \mathbb{R}^6 y,
- de hecho, es el menor subespacio de \mathbb{R}^6 que contiene a los vectores fruta menos el manzana.

La pregunta de la nutrióloga queda entonces expresada en ver si ocurre

$$x_{\text{manzana}} \in \text{span}(\mathcal{A} - \{x_{\text{manzana}}\})$$

o no.

Observa que **los vectores son mucho más que una forma conveniente de almacenar información: podemos hacer operaciones de ellos e interpretarlas en el contexto del problema físico planteado.**

Con las definiciones e ideas que planteamos en los siguientes capítulos, podremos

- determinar cuándo existen combinaciones de frutas que aporten lo mismo que cierta cantidad de manzanas,
- decir si hay una única forma o varias de expresar manzanas en términos de combinaciones lineales de frutas específicas,
- usar argumentos de dimensionalidad para determinar cuándo subconjuntos de frutas son suficientes para sustituir a todas las demás,

entre otras cosas. Es fácil pensar requerimientos razonables que complicarían aún más la situación; por ejemplo, considera que, en un escenario realista,

- un paciente podría tener no sólo posibles alergias como limitantes para definir una dieta, sino también un presupuesto mensual que no pueda sobrepasar, presupuesto que podría reducir aún más las opciones o cantidades de fruta que puede consumir (por ejemplo, esto podría implicar que combinaciones lineales que contengan al vector x_{aguacate} sólo sean consideradas cuando su coeficiente -i.e. el escalar por el que se multiplica a x_{aguacate} - no sea mayor a cierto valor).


No nos interesan todos los elementos de \mathbb{R}^6 , sino sólo aquellos que sean combinaciones lineales de vectores fruta; estamos reemplazando a nuestro espacio de trabajo \mathbb{R}^6 por $\langle \mathcal{A} \rangle$.



Figure 2.5: Cada paciente tiene sus necesidades personales, por ejemplo, las frutas a las que es alérgico (o que no le gustan, y quiere evitar en su dieta) y su presupuesto mensual. Esto cambia el espacio de combinaciones lineales a considerar para cada uno, por lo tanto, las combinaciones lineales aceptables para construirle una dieta.

Incluso en esta situación tan simplificada, vemos la utilidad de usar el marco teórico ofrecido por el álgebra lineal para modelar la situación; con los conocimientos de los próximos capítulos seremos capaces de resolver los problemas aquí planteados.

2.8 Ejercicios I

En lo que sigue, a menos que se indique lo contrario, V es un F -espacio vectorial. Los ejercicios marcados con el símbolo  son obligatorios para los estudiantes de matemáticas, pero opcionales para los de actuaría.

Ejercicio 1. Demuestre que

$$W_1 = \{(a_1, \dots, a_n) \in \mathbb{R}^n : a_1 + \dots + a_n = 0\}$$

es un subespacio de \mathbb{R}^n , pero que

$$W_1 = \{(a_1, \dots, a_n) \in \mathbb{R}^n : a_1 + \dots + a_n = 1\}$$

no lo es.

Ejercicio 2. Demuestre que un subconjunto W de un espacio vectorial V es un subespacio de V si y sólo si

1. $0 \in W$, y
2. para todo $a \in F$ y $x, y \in W$, se tiene que $ax + y \in W$.

Ejercicio 3. **El F -espacio vectorial de polinomios con coeficientes en un campo F :** Sea F un campo. Un **polinomio con coeficientes en el campo F** es una función $f: F \rightarrow F$ de la forma

$$\begin{aligned} f(x) &= a_n x^n + a_{n-1} x^{n-1} + \dots + a_1 x + a_0 \\ &= \sum_{i=0}^n a_i x^i, \end{aligned} \quad (2.22)$$

donde $n \geq 1$ entero y $a_0, a_1, \dots, a_n \in F$. A los a_i se les conoce como **coeficientes del polinomio**. Si todos los coeficientes son cero, decimos que f es el **polinomio cero**, y definimos a su **grado** como -1 . Si $a_n \neq 0$, decimos que el grado del polinomio es n y a_n es su **coeficiente principal**. Todo polinomio cuyo coeficiente principal sea 1 será llamado **mónico**.

Dos polinomios $f(x)$ y $g(x)$ son iguales si y sólo si tienen el mismo grado y sus correspondientes coeficientes son iguales.

Sea

$$F[x] := \{f \in F^F \mid f \text{ es un polinomio}\}$$

el conjunto de todos los polinomios con coeficientes en F . Definimos

1. La suma de dos polinomios

$$f(x) = \sum_{i=0}^n a_i x^i, \quad g(x) = \sum_{i=0}^m b_i x^i, \quad a_n, b_m \neq 0$$

como sigue; sin pérdida de generalidad, supongamos que $n \leq m$. La suma de f con g es

$$(f + g)(x) = \sum_{i=0}^m (a_i + b_i) x^i,$$

donde, para $n \leq i \leq m$, definimos $a_i = 0$.

Haciendo $F = \mathbb{R}$, el polinomio $f(x) = x^2 - 1$ tiene grado dos y es mónico, $f(x) = 4$ tiene grado cero.

Puedes encontrar esta definición del espacio vectorial de polinomios en el Friedberg, p.10. Te recomiendo que estudies también cómo $F[x]$ es esencialmente igual al espacio $F^{(\mathbb{N})}$ de sucesiones en F con soporte finito.

2. La multiplicación escalar de un polinomio cualquiera (2.22) por un escalar $\alpha \in F$ como el polinomio

$$(\alpha f)(x) = \sum_{i=0}^n (\alpha a_i) x^i$$

- Demuestre que $F[x]$ es, con la suma y la multiplicación escalar definida arriba, un F -espacio vectorial.
- Sea

$$M[x] = \{f \in \mathbb{R}[x] : f \text{ es mónico}\}.$$

¿Es $M[x] \leq \mathbb{R}[x]$?

- ¿Es $\mathbb{R}[x] - M[x] \leq \mathbb{R}[x]$?
- Para $n \geq 0$, Sea

$$\mathbb{R}_n[x] = \{f \in \mathbb{R}[x] \mid \text{el grado de } f \text{ es menor o igual a } n\}.$$

Demuestre que $\mathbb{R}_n[x] \leq \mathbb{R}[x]$. Muestre además que $\{1, x, x^2, \dots, x^n\}$ genera a $\mathbb{R}_n[x]$.

Ejercicio 4. El F -espacio vectorial de matrices de dimensión $m \times n$ con coeficientes en F :

Dado un campo F , una **matriz** $m \times n$ **dimensional con coeficientes en F** es una función de la forma

$$A: \{1, \dots, m\} \times \{1, \dots, n\} \longrightarrow F.$$

Normalmente abreviamos

$$A(i, j) = a_{i,j}, \quad 1 \leq i \leq m, \quad 1 \leq j \leq n,$$

y representamos a la matriz A como un arreglo rectangular

$$A = (a_{ij}) = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}.$$

Definimos tanto la suma de matrices como la multiplicación escalar entrada a entrada: para cualesquiera $A = (a_{ij}), B = (b_{ij}), \alpha \in F$,

$$A + B := (a_{ij} + b_{ij}), \quad \alpha A := (\alpha a_{ij}).$$

- Demuestre que el conjunto $M_{m \times n}(F)$ de las matrices $m \times n$ dimensionales con entradas en F y las operaciones definidas arriba es un F -espacio vectorial.
- Demuestre que las matrices

$$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$$

generan a $M_{2 \times 2}(F)$.

- Demuestre que

$$D_n := \{A = (a_{ij}) \in M_{n \times n}(F) \mid i \neq j \Rightarrow a_{ij} = 0\},$$

el conjunto de las matrices diagonales n -dimensionales, es un subespacio de $M_{n \times n}(F)$.

- Si

$$TS_{m,n} := \{A = (a_{ij}) \in M_{m \times n}(F) \mid i > j \Rightarrow a_{ij} = 0\}$$

es el conjunto de las matrices triangulares superiores de dimensión $m \times n$, demuestre que $TS_{m,n} \leq M_{m \times n}(F)$.

Ejercicio 5. Demuestre que un subconjunto W de V es subespacio de V si y sólo si coincide con el subespacio que genera en V , es decir, que

$$W \leq V \Leftrightarrow W = \langle W \rangle$$

Ejercicio 6. Si X y Y son subespacios de V , demuestre que

$$\langle X \cup Y \rangle = \langle X \rangle + \langle Y \rangle = \langle \langle X \rangle \cup \langle Y \rangle \rangle$$

Pista: para la segunda igualdad, simplemente use la definición de suma de subespacios.

Ejercicio 7. (♣♣♣) De un ejemplo de un subconjunto no vacío U de \mathbb{R}^2 que sea cerrado bajo multiplicación escalar, pero que no sea subespacio de \mathbb{R}^2 .

Ejercicio 8. (♣♣♣) Si W es un subespacio de V , describa al subespacio $W + W$.

Ejercicio 9. Si $\{W_i\}_{i=1}^n$ es una familia de subespacios de V , demuestre que las siguientes proposiciones son equivalentes:

- La suma $\sum_{i=1}^n W_i$ es directa.
- Si los vectores $w_i \in W_i$ ($1 \leq i \leq n$) son tales que $w_1 + \dots + w_n = 0$, entonces toda w_i es cero.

Es decir, demuestre que determinar si representaciones de la forma $w_1 + \dots + w_n$ con $w_i \in W_i$ son únicas equivale a ver la unicidad de la representación sólo del vector cero como suma de elementos de $\cup W_i$.

Ejercicio 10. Use a los subespacios

$$W_1 = \{(x, y, 0) \in \mathbb{R}^3 : x, y \in \mathbb{R}\},$$

$$W_2 = \{(0, 0, x) \in \mathbb{R}^3 : x \in \mathbb{R}\},$$

$$W_3 = \{(0, x, x) \in \mathbb{R}^3 : x \in \mathbb{R}\}$$

de \mathbb{R}^3 para dar un ejemplo de una familia de subespacios para la que, a pesar de que la intersección es $\{0\}$, la suma no es directa. *Pista:* para facilitarle el trabajo, use el ejercicio 9.

Ejercicio 11. Si

$$W_1 = \{(a_1, \dots, a_n) \in F^n \mid a_1 = 0\},$$

$$W_2 = \{(a_1, \dots, a_n) \in F^n \mid a_2 = \dots = a_n = 0\},$$

demuestre que $F^n = W_1 \oplus W_2$.

Ejercicio 12. Sean $X, Y \subseteq V$. Demuestre que

- Si $X \subseteq Y$ entonces $\langle X \rangle \leq \langle Y \rangle$.
- Si $X \subseteq Y$ y $\langle X \rangle = V$, entonces $\langle Y \rangle = V$.
- $\langle X \cap Y \rangle \subseteq \langle X \rangle \cap \langle Y \rangle$. Busque un ejemplo en el que se de la igualdad y otro en el que la contención sea propia.

Ejercicio 13. ($\clubsuit\spadesuit$) Sea $b \in \mathbb{R}$. Defina a la familia

$$\mathcal{I}_b := \left\{ f \in \mathbb{R}^{[0,1]} : \int_0^1 f(x) dx = b \right\}.$$

Demuestre que \mathcal{I}_b es subespacio de $\mathbb{R}^{[0,1]}$ si y sólo si $b = 0$.

Ejercicio 14. ($\clubsuit\spadesuit$) Sea $A \neq \emptyset$ un conjunto cualquiera, $\mathbb{Z}_2 = \{0, 1\}$ el campo de los enteros módulo 2.

Podemos identificar a todo subconjunto $B \subseteq A$ con su **función característica**, es decir, la función $\chi_B : A \rightarrow \mathbb{Z}_2$ definida como

$$\chi_B(x) = \begin{cases} 1 & \text{si } x \in B, \\ 0 & \text{si } x \notin B; \end{cases}$$

recíprocamente, toda función $f : A \rightarrow \mathbb{Z}_2$ induce un subconjunto de A como

$$B = \{x \in A : f(x) = 1\}.$$

Note que estos procesos son uno el inverso del otro, y que ellos nos permiten establecer una biyección entre $\mathcal{P}(A)$ -el conjunto potencia de A - y el conjunto \mathbb{Z}_2^A de funciones de A en \mathbb{Z}_2 . En la sección ?? vimos cómo dotar a \mathbb{Z}_2^A de estructura de \mathbb{Z}_2 -espacio vectorial.

Dados $\alpha \in \mathbb{Z}_2$, $f, g \in \mathbb{Z}_2^A$, si B y C son los subconjuntos de A que estas funciones inducen,

- ¿Qué subconjunto de A induce la función suma $f + g$?
- ¿Cuál es el subconjunto de A inducido por la función cero $\hat{0} : A \rightarrow \mathbb{Z}_2$?
- ¿Qué subconjunto de A induce $-f$?
- ¿Cuál induce αf ?
- Si $\{B_i\}_{i \in I}$ es una familia de subespacios de A tal que $\{\chi_{B_i} \mid i \in I\}$ genera al \mathbb{Z}_2 -espacio vectorial \mathbb{Z}_2^A , demuestre que $A = \bigcup_{i \in I} B_i$.

2.9 Dependencia e independencia lineal

Definición 2.9.1. Sea V un F -espacio vectorial. Decimos que $S \subseteq V$ es **linealmente dependiente** si existe un número finito de vectores distintos entre si $x_1, \dots, x_n \in S$ y escalares $a_1, \dots, a_n \in F$ no todos cero tales que

$$\sum_{i=1}^n a_i x_i = \hat{0}.$$

Abreviamos los términos **linealmente dependiente** (resp. **independiente**) como **l.d.** (resp. **l.i.**).

Si S no es linealmente dependiente, diremos que es **linealmente independiente**.

En la definición 2.9.1 pedimos

- que los vectores x_i sea distintos entre sí, y
- que al menos uno de los escalares a_i sea no cero,

pues las igualdades

$$1_F x + (-1_F)x = x - x = \hat{0}$$

y

$$0_F x = \hat{0},$$

son ciertas *para todo* $x \in V$, ¡no las vamos a usar para definir un nuevo concepto!

Si un conjunto finito $S = \{x_1, \dots, x_n\}$ es l.d.(resp. l.i.), acostumbramos decir que los vectores x_1, \dots, x_n son l.d. (resp. l.i.).

Ejercicio: Niegue la definición de dependencia lineal dada en 2.9.1 para obtener la definición de independencia lineal.

Ejemplo 2.9.2. • *Por vacuidad, el conjunto vacío es linealmente independiente*

- *Todo subconjunto de S que contenga al cero es linealmente dependiente*

◇

Ejercicio: demuestre a detalle los puntos del ejemplo anterior.

Ejercicio: demuestre que, dados $S \subseteq T \subseteq V$,

- si T es l.i., entonces S es l.i.
- si S es l.d., entonces T es l.d.

Ejercicio: demuestre que si x_1, \dots, x_n distintos entre sí son l.i. y $a_1 x_1 + \dots + a_n x_n = \hat{0}$, entonces todos los escalares a_i son iguales a cero.

Proposición 2.9.3. $S \subseteq V$ es linealmente dependiente si y sólo si existe $x \in S$ tal que $x \in \text{span}(S - \{x\})$.

El enunciado dado en la proposición 2.9.3 es la definición de dependencia lineal manejada en [Mej11].

Demostración.

⇒] Sean $x_1, \dots, x_n \in S$ distintos entre sí y a_1, \dots, a_n escalares no todos cero tales que $\sum_{i=1}^n a_i x_i = \hat{0}$. Sin pérdida de generalidad, supongamos que $a_1 \neq 0$. Entonces,

$$a_1 x_1 = - \sum_{i=2}^n a_i x_i \in \text{span}(S - \{x_1\}),$$

luego, $x_1 \in \text{span}(S - \{x_1\})$.

⇐] Sea $x \in S$ tal que $x \in \text{span}(S - \{x\})$. Sean pues $x_1, \dots, x_n \in S - \{x\}$ distintos entre si tales que $x = \sum_{i=1}^n a_i x_i$. Restando x de ambos lados de la igualdad llegamos a que

$$\hat{0} = \sum_{i=1}^n a_i x_i - x = \sum_{i=1}^n a_i x_i + (-1_F)x.$$

(c.f. prop REF). Como $-1_F \neq 0_F$, esta última igualdad muestra que S es linealmente dependiente.

□

Proposición 2.9.4. *Un singulete $\{x\} \subseteq V$ es linealmente independiente en V si y sólo si $x \neq \hat{0}$.*

Demostración. En efecto, según la Proposición 2.9.3, $\{x\}$ es linealmente dependiente si y sólo si $x \in \text{span}(\{x\} - \{x\}) = \text{span}(\emptyset) = \{0\}$, o sea, si y sólo si x es el vector cero. \square

Proposición 2.9.5. *$S \subseteq V$ es linealmente dependiente si y sólo si existe $x \in S$ tal que $\text{span}(S) = \text{span}(S - \{x\})$.*

Demostración. Usaremos la equivalencia probada en la Proposición 2.9.3.

\Rightarrow] Sea $x \in S$ tal que $x \in \text{span}(S - \{x\})$. Como $S - \{x\}$ está contenido en S , trivialmente $\text{span}(S - \{x\})$ está contenido en $\text{span}(S)$. La otra contención se da también pues, por poder x expresarse como combinación lineal de elementos de $S - \{x\}$, en cualquier combinación lineal de elementos de S se puede reemplazar a x con múltiplos escalares de $S - \{x\}$, es decir, todo elemento de $\text{span}(S)$ es elemento de $\text{span}(S - \{x\})$.

\Leftarrow] Si existe $x \in S$ tal que $\text{span}(S) = \text{span}(S - \{x\})$, entonces

$$x \in S \subseteq \text{span}(S) = \text{span}(S - \{x\}).$$

\square

Este resultado nos indica que, cuando un conjunto S es l.d., hay al menos un vector que es **redundante**, en el sentido en que hay otros vectores en S que pueden reemplazar a x para generar al span de S ; el quitar a x de S no implica una pérdida de información.

Proposición 2.9.6. *Dado $S \subseteq V$, son equivalentes las siguientes proposiciones:*

- *S es linealmente independiente*
- *Todo subconjunto finito de S es linealmente independiente.*

*La proposición 2.9.6 dice que ser l.i. es una **propiedad de carácter finito** (ref wiki).*

Demostración.

\Rightarrow] En general, todo subconjunto de un conjunto linealmente independiente es también linealmente independiente.

\Leftarrow] Por contraposición. Sea S linealmente dependiente; sean pues $x_1, \dots, x_n \in S$ distintos entre sí y $a_1, \dots, a_n \in F$ no todos cero tales que

$$\sum_{i=1}^n a_i x_i = \hat{0};$$

esta ecuación muestra que el subconjunto finito $\{x_1, \dots, x_n\}$ de S es linealmente dependiente.

\square

Ejemplo 2.9.7. *En el \mathbb{R} -espacio vectorial \mathbb{R}^4 , considere a los vectores*

$$x_1 = (2, -1, 4), x_2 = (1, -1, 3), x_3 = (1, 1, -1), x_4 = (1, -2, -1).$$

¿Son linealmente dependientes en \mathbb{R}^4 ? Para que lo fuesen, deben existir escalares $a_1, a_2, a_3, a_4 \in \mathbb{R}$ no todos cero tales que

$$a_1x_1 + a_2x_2 + a_3x_3 + a_4x_4 = \hat{0}.$$

Podemos desglosar esta igualdad en tres (una para cada entrada de los vectores), y llegar a las tres condiciones equivalentes

$$\begin{cases} 2a_1 + a_2 + a_3 + a_4 &= 0, \\ -a_1 - a_2 + a_3 - 2a_4 &= 0, \\ 4a_1 + 3a_2 - a_3 - a_4 &= 0. \end{cases}$$

Aquí estamos reescribiendo nuestro problema original - que consiste de una ecuación en \mathbb{R}^3 - en términos de tres ecuaciones en \mathbb{R} .

Usando operaciones elementales, podemos encontrar un sistema de ecuaciones diagonal y equivalente (en el sentido que tenga exactamente las mismas soluciones que el sistema inicial);

$$\left(\begin{array}{cccc|c} 1 & 1 & -1 & 2 & 0 \\ 1 & 2 & -1 & -1 & 0 \\ 5 & 7 & -4 & 9 & 0 \end{array} \right) \sim \left(\begin{array}{cccc|c} 1 & 1 & -1 & 2 & 0 \\ 0 & -1 & 3 & -3 & 0 \\ 0 & -1 & 3 & -9 & 0 \end{array} \right) \sim \left(\begin{array}{cccc|c} 1 & 1 & -1 & 2 & 0 \\ 0 & -1 & 3 & -3 & 0 \\ 0 & 0 & 0 & -6 & 0 \end{array} \right).$$

Buscamos entonces soluciones no cero del sistema de ecuaciones

$$\begin{cases} a_1 + a_2 - a_3 + 2a_4 &= 0, \\ -a_2 + 3a_3 - 3a_4 &= 0, \\ 6a_4 &= 0. \end{cases}$$

La última ecuación implica $a_4 = 0$. La segunda entonces da la relación $3a_3 = a_2$; para asegurarnos de encontrar una solución no trivial, hagamos a_2 no cero, digamos, $a_2 = 3$. Entonces, $a_3 = 1$. La primera ecuación se reescribe entonces como $a_1 + 3 - 1 = 0$, luego, $a_1 = -2$.

Note que, en efecto, ocurre

$$-2x_1 + 3x_2 + x_3 = \hat{0}.$$

Con esto concluimos que $\{x_1, x_2, x_3, x_4\}$ es linealmente dependiente. \diamond

Ejemplo 2.9.8. Sean los vectores

$$\hat{x} = (2, 3, 1), \hat{y} = (1, -1, 2), \hat{z} = (7, 3, \alpha) \in \mathbb{R}^3.$$

Determinemos para qué valores de α esta colección es linealmente independiente. Sean $a, b, c \in \mathbb{R}$ tales que

$$a\hat{x} + b\hat{y} + c\hat{z} = \hat{0}.$$

a, b, c son soluciones del sistema

$$\begin{cases} 2a + b + 7c &= 0, \\ 3a - b + 3c &= 0, \\ a + 2b + \alpha c &= 0. \end{cases}$$

Busquemos un sistema de ecuaciones diagonal equivalente a este;

$$\left(\begin{array}{ccc|c} 1 & 2 & \alpha & 0 \\ 2 & 1 & 7 & 0 \\ 3 & -1 & 3 & 0 \end{array} \right) \sim \left(\begin{array}{ccc|c} 1 & 2 & \alpha & 0 \\ 0 & -3 & 7-2\alpha & 0 \\ 0 & -7 & 3-3\alpha & 0 \end{array} \right) \sim \left(\begin{array}{ccc|c} 1 & 2 & \alpha & 0 \\ 0 & 1 & -\frac{7}{3} + \frac{2}{3}\alpha & 0 \\ 0 & 0 & -\frac{40}{3} + \frac{5}{3}\alpha & 0 \end{array} \right).$$

Entonces, el sistema a resolver equivale a

$$\begin{cases} a + 2b + \alpha c &= 0, \\ b + \left(-\frac{7}{3} + \frac{2}{3}\alpha\right)c &= 0, \\ \left(-\frac{40}{3} + \frac{5}{3}\alpha\right)c &= 0. \end{cases}$$

Esta forma nos permite ver que el que la solución trivial $a = b = c = 0$ sea la única solución del sistema -es decir, el que la colección de vectores dada al inicio sea linealmente independiente - equivale a que $-\frac{40}{3} + \frac{5}{3}\alpha$ sea cero, o sea, a que $\alpha = 8$. (¿por qué?). \diamond

Ejemplo 2.9.9. ¿Verdadero o falso?

- “Si S es un conjunto l.d., entonces cada elemento de S es combinación lineal de otros elementos de S .” Falso. **Ejercicio:** construye un ejemplo.
- “Subconjuntos de conjuntos l.d. son l.d.” Falso. Toma un conjunto l.d. que contenga al menos un vector x no cero; $\{x\}$ es un subconjunto de este que es l.i..

2.10 Bases y dimensión

Definición 2.10.1. Sea V un F -espacio vectorial. Un subconjunto de V que sea linealmente independiente y que genere a V será llamado una **base** del espacio.

Ejemplo 2.10.2. **Ejercicio:** demuestre las siguientes afirmaciones.

- El vacío es la única base del F -espacio vectorial $\{0\}$.
- En F^n , si e_i es el vector de F^n cuyas entradas son cero menos la i -ésima, que es uno, entonces

$$\{e_1, \dots, e_n\}$$

es una base de F^n . A esta se le conoce como la **base canónica de F^n** .

- En $F[x]$ - el F -espacio vectorial de polinomios con coeficientes en F - el conjunto

$$\{x^k : k \in \mathbb{N}\}, \quad \mathbb{N} = \{0, 1, 2, \dots\}$$

es una base de este espacio.

- El conjunto

$$\{1, x, x^2, \dots, x^n\}$$

es una base de $\mathbb{R}_n[x]$.

El siguiente Teorema aclara la importancia del concepto de base.

Teorema 2.10.3. Sea V un F -espacio vectorial distinto de $\{0\}$. Las siguientes proposiciones son equivalentes:

- $\emptyset \neq \mathcal{B} \subseteq V$ una base del espacio.
- Todo vector $\hat{x} \in V$ puede ser expresado de manera única como una combinación lineal de vectores de \mathcal{B} .

En este sentido los conjuntos generadores linealmente independientes son los elementos constructivos de los espacios vectoriales.

Demostración. Antes de comenzar, recuerde que, según la proposición 2.6.8,

$$\text{span}(\mathcal{B}) = \left\{ \sum_{i=1}^n a_i v_i \mid n \in \mathbb{N}, v_i \in \mathcal{B}, a_i \in F \text{ para toda } 1 \leq i \leq n \right\}$$

Además, el que un vector $\hat{x} \in V$ pueda expresarse de forma única como combinación lineal de elementos de \mathcal{B} significa que cualesquiera dos combinaciones lineales de \mathcal{B} iguales a \hat{x} tendrán los mismos coeficientes.

\Rightarrow) Supongamos que \mathcal{B} es base de V . Sea $\hat{x} \in V$. Puesto que $V = \text{span}(\mathcal{B})$, sabemos que \hat{x} es combinación lineal finita de elementos de \mathcal{B} .

Sean dos representaciones (véase la nota [★])

$$\sum_{i=1}^n a_i v_i = \hat{x} = \sum_{i=1}^n b_i v_i, \quad (2.23)$$

con $v_i \in \mathcal{B}$ para toda $1 \leq i \leq n$. Entonces,

$$\hat{0} = \hat{x} - \hat{x} = \sum_{i=1}^n (a_i - b_i) v_i;$$

por ser \mathcal{B} linealmente independiente, esto último implica que, para toda i , $a_i - b_i = 0$, luego, que $a_i = b_i$, y que las representaciones (2.23) son iguales.

\Leftarrow) El que todo vector del espacio pueda expresarse como combinación lineal de elementos de \mathcal{B} ya significa que \mathcal{B} genera a todo el espacio V . Para demostrar que \mathcal{B} es también l.i., tomemos un subconjunto finito arbitrario $\{v_1, \dots, v_n\}$ de él y mostremos que es l.i. (c.f. proposición 2.9.6). Si los escalares a_1, \dots, a_n son tales que

$$a_1 x_1 + \dots + a_n x_n = \hat{0},$$

puesto que también se tiene

$$0x_1 + \dots + 0x_n = \hat{0},$$

por unicidad concluimos que $a_i = 0$ para toda $i = 1, \dots, n$.

□

Nota [★:] En realidad, debíamos de partir de dos representaciones de \hat{x} de la forma

$$\hat{x} = \sum_{j=1}^l c_j w_j, \quad \hat{x} = \sum_{k=1}^m d_k u_k, \quad c_i, d_k \in F, w_i, u_k \in \mathcal{B} \quad (2.24)$$

y demostrar que $l = m$, que los w_i son una permutación de los u_k (i.e. un reordenamiento) y que los correspondientes coeficientes c_i y d_k son iguales entre sí. Sin embargo, podemos limitarnos a usar representaciones como las dadas en (2.23) - que tienen el mismo número de sumandos n y consideran a los mismos vectores v_i - sin pérdida de generalidad, pues podemos convertir dos representaciones (2.24) a unas de la forma (2.23) sumando ceros y renombrando los vectores.

Por ejemplo, si $F = \mathbb{R}$ y tenemos las representaciones

$$\hat{x} = 2w_1 + 3u_1 - 7w_3 + 4u_3 \quad (w_2 = u_1, w_4 = u_3)$$

y

$$\hat{x} = 1u_1 - 0.3u_2 + 5u_3,$$

podemos reemplazarlas por

$$\hat{x} = 3u_1 + 0u_2 + 4u_3 + 2w_1 - 7w_3$$

y

$$\hat{x} = 1u_1 - 0.3u_2 + 5u_3 + 0w_1 + 0w_3.$$

respectivamente

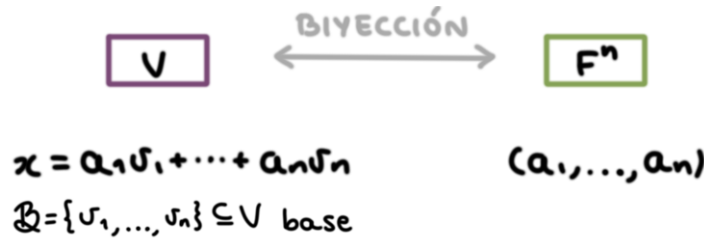


Figure 2.6: Si se tiene una base finita \mathcal{B} de V , el teorema 2.10.3 da una biyección entre V y F^n . Más adelante que esta es más que una biyección, pues también preserva la estructura algebraica de los subespacios.

2.11 Cómo construir bases de un espacio vectorial

Si $S \subseteq V$ es un subconjunto de un F -espacio vectorial V ,

- si S es linealmente independiente, la única forma en la que no puede ser base es si el span de S es un subespacio propio de V . Esto significa que, formando combinaciones lineales con los elementos de S no podemos construir todo el espacio:

$$\exists v \in V - \text{span}(S).$$

Este problema puede intentar arreglarse (cuando se pueda) agregando vectores del espacio que hagan falta para generarlo, pero ¡sin arruinar la independencia lineal!

- Si S es generador, la única forma en la que no puede ser base es si es linealmente dependiente, es decir, si hay redundancia en S (c.f. Proposición 2.9.3), es decir, si

$$\exists v \in S : v \in \text{span}(S - \{v\}). \quad (2.25)$$

Recuerda que esto significa que el elemento v de S en realidad no es necesario para construir a todas las combinaciones lineales de elementos de S . Podríamos intentar arreglar esto quitando a los elementos de S que pueden ser sustituidos con otros, en el sentido de la ecuación (2.25).

Si partimos de un subconjunto l.i. del espacio, demos un criterio necesario y suficiente para extenderlo (i.e. agregar más vectores del espacio) sin alterar la independencia lineal.

Lema 2.11.1. Sean S un subconjunto l.i. de V , $x \in V - S$. Entonces, son equivalentes

- $S \cup \{x\}$ es linealmente dependiente,
- $x \in \text{span}(S)$.

Demostración.

\Rightarrow) Suponiendo que x no está en el generado por S , mostremos que $S \cup \{x\}$ es l.i. o, equivalentemente, que todo subconjunto finito de él lo es. Sea $\{x_1, \dots, x_{n-1}\} \cup \{x\}$ un subconjunto finito de $S \cup \{x\}$ (¿por qué sólo interesa el caso en el que el subconjunto contiene a x ?). Sean a_i escalares tales que

$$a_1x_1 + \dots + a_{n-1}x_{n-1} + a_nx = \hat{0}.$$

Si a_n fuese distinto de cero, despejando tendríamos que

$$x = -\frac{a_1}{a_n}x_1 + \dots + -\frac{a_{n-1}}{a_n}x_{n-1} \in \text{span}(S) \quad \nexists$$

Así, $a_n = 0$. Luego,

$$a_1x_1 + \dots + a_{n-1}x_{n-1} = \hat{0}$$

y, como S es l.i., esto último implica la igualdad a cero de todos los coeficientes.

\Leftarrow) Puesto que

$$x \in \text{span}(S) = \text{span}((S \cup \{x\}) - \{x\}),$$

según la Proposición 2.9.3, $S \cup \{x\}$ es l.d..

□

Definición 2.11.2. Sea S un subconjunto del F -espacio vectorial V . Decimos que S es **linealmente independiente maximal** si para todo $x \in V - S$, $S \cup x$ es l.d..

Según el Lema 2.11.1, S es l.i. maximal si todo elemento de $V - S$ está en el generado de S . Puesto que, trivialmente $S \subseteq \text{span}(S)$, tenemos que el ser l.i. maximal implica ser base del espacio. Recíprocamente, si S es base de V , en particular es l.i., además, si $x \in V - S$, entonces $x \in V = \text{span}(S)$, luego, según el lema 2.11.1, $S \cup \{x\}$ es l.d., luego, S es l.i. maximal. Con esto probamos que

Observación 2.11.3. Los conceptos de base y subconjunto linealmente independiente maximal coinciden.

El nombre “l.i. maximal” puede ser más ilustrativo que el término “base”, pues da a entender que una base es un l.i. al que no se le puede agregar otro vector del espacio sin arruinar la independencia lineal.

Teorema 2.11.4. Si un espacio vectorial V es generado por un conjunto finito S , entonces un subconjunto de S es base para V .

El teorema consiste en extraer bases de conjuntos generadores finitos.

Demostración. Si $S = \emptyset, \{0\}$, entonces $V = \{0\}$, y en ambos casos puede extraerse del conjunto generador S al vacío, que es una base del espacio.

Supongamos ahora que existe $x_1 \in S - \{0\}$. El singulete $\{x_1\}$ es entonces un subconjunto l.i. de S (c.f. Proposición 2.9.4).

• Si

$$\forall z \in S - \{x_1\} : z \in \text{span}(\{x_1\}),$$

entonces tendremos que $S \subseteq \text{span}(\{x_1\})$ - pues trivialmente $x_1 \in \text{span}(\{x_1\})$, luego,

$$V = \text{span}(S) \leq \text{span}(\text{span}(\{x_1\})) = \text{span}(\{x_1\}) \leq V,$$

por lo tanto,

$$\text{span}(\{x_1\}) = V,$$

y entonces $\{x_1\} \subseteq S$ es un l.i. que genera al espacio, i.e. una base de este.

- Supongamos que ocurre lo contrario, es decir, que

$$\exists x_2 \in S - \{x_1\} : x_2 \notin \text{span}(\{x_1\}).$$

Entonces, según el Lema 2.11.1, $\{x_1, x_2\} \subseteq S$ es l.i..

en este último caso, podemos seguir el argumento de antes, y preguntarnos si existe algún $z \in S - \{x_1, x_2\}$ tal que $z \notin \text{span}(\{x_1, x_2\})$:

- si la respuesta a esta pregunta es “no”, entonces tendremos, como antes, que $S \subseteq (\{x_1, x_2\})$, de donde, formando el span a ambos lados de la contención, se seguirá que el l.i. $\{x_1, x_2\}$ genera a V , luego, es base de V ;
- si, por el contrario, la respuesta es “sí”, por el Lema 2.11.1 podemos agregar este elemento al l.i. $\{x_1, x_2\}$ sin alterar la independencia lineal.

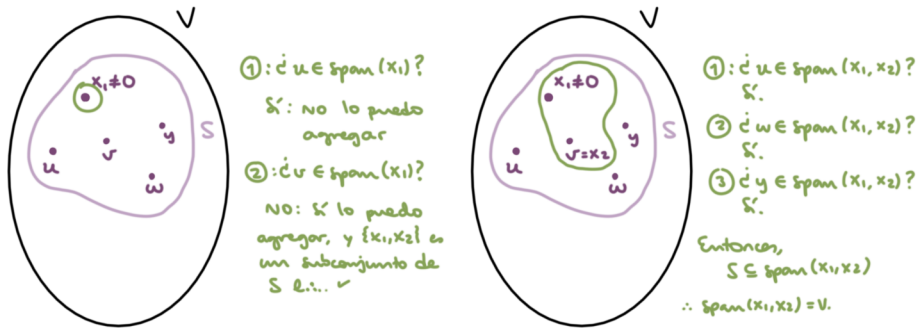


Figure 2.7: Tengo que responder una cantidad finita de preguntas para saber si puedo seguir extendiendo a mi subconjunto l.i. de S .

Siguiendo este proceso, debemos conseguir un subconjunto l.i. $\{x_1, \dots, x_n\}$ de S con la propiedad

$$\forall z \in S - \{x_1, \dots, x_n\} : z \in \text{span}(\{x_1, \dots, x_n\});$$

pues el conjunto generador S del que partimos es **finito** (luego, para llegar a tal l.i. hay que responder una cantidad finita de preguntas, el proceso no se extiende sin límite). Con el argumento de antes se seguirá que tal l.i. es una base del espacio.

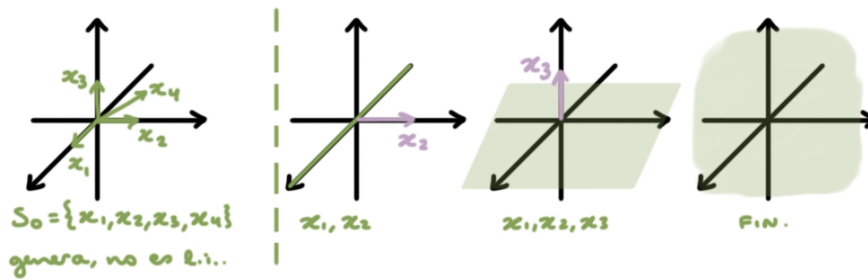


Figure 2.8: Figura que ilustra el proceso para la colección de cuatro vectores de \mathbb{R}^3 que se muestra en el diagrama.

□

Corolario 2.11.5. Si un espacio vectorial tiene un subconjunto finito que lo genera, entonces tiene base (finita).

La ventaja de la demostración del Teorema 2.11.4 es que es *constructiva*, es decir, no sólo muestra la existencia de tales bases, sino que explica cómo construirlas.

Nota que hasta el momento no hemos dicho nada sobre la existencia de bases en un espacio vectorial; nos vamos a limitar en esta sección a hacer inferencias sobre espacios vectoriales que, por hipótesis, tengan bases.

Teorema 2.11.6. Sean V un F -espacio vectorial, β una base V con n elementos, $n \in \mathbb{N}$. Sea $S = \{y_1, \dots, y_m\}$ un subconjunto linealmente independiente con $m \leq n$. Entonces, existe S_1 subconjunto de β con $n - m$ elementos tal que $\langle S \cup S_1 \rangle = V$.

Demostración. Vamos a proceder por inducción sobre $m \leq n$.



- *Base de inducción:* si $m = 0$, o sea, si $S = \emptyset$, entonces $S_1 = \beta$ funciona.
- *Paso inductivo:* supongamos el Teorema cierto para $m < n$ y demostremos que el teorema también se cumple para $m+1$. Sea pues $S = \{y_1, \dots, y_m, y_{m+1}\}$ un subconjunto l.i. de $m+1$ elementos. Como $\{y_1, \dots, y_m\}$ es un l.i. de m elementos, por hipótesis de inducción existe $\{x_1, \dots, x_{n-m}\} \subseteq \beta$ tal que

$$\text{span}(\{y_1, \dots, y_m\} \cup \{x_1, \dots, x_{n-m}\}) = V. \quad (2.26)$$

Existen pues escalares $a_i, b_j \in F$ tales que

$$y_{m+1} = a_1 y_1 + \dots + a_m y_m + b_1 x_1 + \dots + b_{n-m} x_{n-m}.$$

Como S es l.i., y_{m+1} no puede ponerse como combinación lineal de otros elementos de S , luego, al menos algún b_j debe no ser cero; sin pérdida de generalidad, digamos que $b_1 \neq 0$. Entonces, podemos despejar a x_1 de la ecuación anterior y llegar a que

$$x_1 \in \text{span}(\{y_1, \dots, y_m, y_{m+1}\} \cup \{x_2, \dots, x_{n-m}\});$$

de esto se sigue que

$$\{y_1, \dots, y_m, x_1, x_2, \dots, x_{n-m}\} \subseteq \text{span}(\{y_1, \dots, y_m, y_{m+1}, x_2, \dots, x_{n-m}\}).$$

Formando el span a ambos lados de la contención y usando (2.26), llegamos a que

$$\begin{aligned} V &= \text{span}(\{y_1, \dots, y_m, x_1, x_2, \dots, x_{n-m}\}) \\ &\subseteq \text{span}(\{y_1, \dots, y_m, y_{m+1}, x_2, \dots, x_{n-m}\}), \end{aligned}$$

luego, el subconjunto $\{x_2, \dots, x_{n-m}\}$ de $n - (m+1)$ elementos de β es tal que

$$\text{span}(\{y_1, \dots, y_{m+1}\} \cup \{x_2, \dots, x_{n-m}\}) = V.$$



□

Corolario 2.11.7. Sea V un F -espacio vectorial que tiene una base β con n elementos. Entonces, cualquier subconjunto l.i. de V con n elementos es base de V .

Demostración. Según el Teorema 2.11.6, si S es un tal subconjunto l.i., entonces existe $S_1 \subseteq \beta$ tal que $S \cup S_1$ genera a V y S_1 tiene $n - n = 0$ elementos; entonces, tal S_1 debe ser el vacío, y $S = S \cup \emptyset$, además de ser l.i., genera a V , luego, es base del espacio.

□

Ejemplo 2.11.8. En los ejercicios mostraste que \mathbb{R}^3 tiene a $\{(1, 0, 0), (0, 1, 0), (0, 0, 1)\}$ como base. Sean los vectores $x_1 = (1, -3, 2)$, $x_2 = (4, 1, 0)$ y $x_3 = (0, 2, -1)$. Es fácil ver que estos son linealmente independientes. Según el Corolario 2.11.7, esto implica que ellos conforman una base para \mathbb{R}^3 .

Este Corolario nos indica que la cardinalidad de una base finita de V acota la cardinalidad de los subconjuntos l.i. del espacio.

Corolario 2.11.9. Sea V un F -espacio vectorial. Si existe β base de V con n elementos, entonces cualquier subconjunto de V con más de n elementos es l.d..

Demostración. Sea $S \subseteq V$ con $|S| > n$. Supongamos que S es l.i.. Si S_1 es un subconjunto de S con al menos n elementos, entonces al igual que S es l.i., luego, por el corolario 2.11.7 es una base del espacio, entonces, $\text{span}(S_1) = V$. Si $x \in S - S_1$, como $x \in V = \text{span}(S_1)$, según el Lema 2.11.1, $S_1 \cup \{x\} \subseteq S$ es l.d., luego, S es l.d. \nmid

□

Corolario 2.11.10. Sea V un F -espacio vectorial. Si existe $\beta \subseteq V$ base de n elementos, entonces cualquier otra base de V tendrá n elementos.

Demostración. Sea $\gamma \subseteq V$ otra base de V . Como γ (respectivamente, β) es linealmente independiente y β (respectivamente, γ) es base del espacio, por el Corolario 2.11.9 $|\gamma| \leq |\beta|$ (respectivamente, $|\beta| \leq |\gamma|$).

□

Este último resultado nos permite introducir la noción de dimensión.

Definición 2.11.11. Un espacio vectorial se dice **finito dimensional** si tiene una base que consta de un número finito de elementos. El único número de elementos en cada base del espacio se llama la **dimensión de V** , y se denotará por $\dim(V)$.

Todo espacio vectorial que no sea finito dimensional será llamado **infinito dimensional**.

Nota que hasta el momento hemos encontrado resultados válidos para cuando el espacio vectorial tiene una base finita; aunque con lo visto hasta ahora no podemos asegurar que cualquier espacio vectorial tiene base, esto es cierto y puede demostrarse usando el Lema de Zorn. Puedes consultar los detalles en [FSH03], sección 1.7. En la práctica, casi siempre se usarán espacios de dimensión finita (de hecho, algún \mathbb{R}^n).

Ejemplo 2.11.12. Demuestre las siguientes afirmaciones:

- El espacio vectorial $\{0\}$ tiene dimensión cero.
- El espacio vectorial F^n tiene dimensión n .
- El espacio vectorial $M_{m \times n}(F)$ tiene dimensión $m \times n$.
- El espacio vectorial $P_n(F)$ tiene dimensión $n + 1$.
- El espacio vectorial $P(F)$ es infinito dimensional.

Ejemplo 2.11.13. Ilustramos a continuación el que la dimensión de un F -espacio vectorial V depende no solo del grupo abeliano V , sino también del campo F .

- El \mathbb{C} -espacio vectorial \mathbb{C} tiene dimensión 1. En efecto, $\{1\}$ es base para él, pues
 - el singulete $\{1\}$ es l.i., y,
 - dado $z \in \mathbb{C}$, $z = 1 \cdot z \in \text{span}(\{1\})$, luego, $\text{span}(\{1\}) = \mathbb{C}$.
- El \mathbb{R} -espacio vectorial \mathbb{C} tiene dimensión 2, pues una base de este espacio es $\{1, i\}$;
 - Dado $z = a + bi \in \mathbb{C}$ (donde $a, b \in \mathbb{R}$), $z = a \cdot 1 + b \cdot i \in \langle \{1, i\} \rangle$, y
 - no existe $a \in \mathbb{R}$ tal que $a \cdot 1 = i$, luego, i no es múltiplo escalar de 1, por lo que $\{1, i\}$ es l.i..

Corolario 2.11.14. Sea V un F -espacio vectorial de dimensión n . Si $S \subseteq V$ genera a V y tiene a lo más n elementos, entonces S es base de V (luego, $|S| = n$).

Demostración. Por el Teorema 2.11.4, sabemos que existe $S_1 \subseteq S$ base de V . Entonces, como V es n -dimensional, $|S_1| = n$; entonces, $|S| \leq n$ y S tiene un subconjunto S_1 de n elementos, luego S tiene n elementos y coincide con S_1 , por lo tanto, es base de V . \square

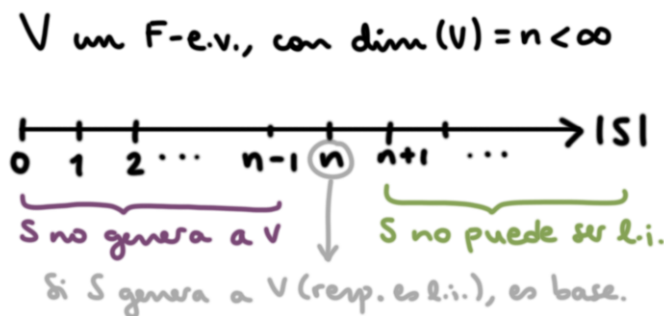


Figure 2.9: Según lo estudiado ahora, la cardinalidad de un subconjunto de un espacio vectorial n -dimensional nos indica si puede o no ser l.i. o generador.

Corolario 2.11.15. Sea β una base de un espacio vectorial V de dimensión n . Sea $S \subseteq V$ linealmente independiente. Existe $S_1 \subseteq \beta$ tal que $S \cup S_1$ es base de V .

Este corolario nos explica cómo extender l.i.'s a bases finitas.

Demostración. Por el Corolario 2.11.9, $m := |S| \leq n$, entonces, por el teorema 2.11.6, existe $S_1 \subseteq \beta$ con $|S_1| = n - m$ tal que $S \cup S_1$ genera a V . Claro que $|S \cup S_1| \leq n$; así, por el corolario 2.11.14, $S \cup S_1$ es base de V . \square

Resumimos lo deducido hasta ahora:

- Una base de un espacio vectorial es un subconjunto de este que lo genera y es linealmente independiente.
- Si V tiene una base finita, entonces cualquier base de V tiene el mismo número de vectores. A este número n se le llama la dimensión de V .
- En este caso, todo subconjunto linealmente independiente o generador de n elementos es base del espacio,
- todo subconjunto l.i. contiene a lo más n vectores (c.f. corolario 2.11.9), y en caso de tener menos de n vectores (i.e. en caso de no ser base) puede completarse a una base del espacio (c.f. corolario 2.11.15).
- Recíprocamente, si $\dim(V) = n$, entonces todo subconjunto generador contiene por lo menos n elementos y de él se puede extraer una base de V eliminando vectores que hagan al generador redundante (recuerda este proceso explicado en la demostración del teorema 2.11.4).



Agregar el proceso para extender un l.i. a una base de un espacio finito dimensional con la ayuda del Lema 2.11.1. Poner el ejemplo práctico en $M_{2 \times 2}(\mathbb{R})$.

Proposición 2.11.16. *lo pongo aquí? checa el ejemplo 3.1.5 Sean $U, W \leq V$ subespacios de un F -espacio vectorial V . Las siguientes son equivalentes:*

- $V = U \oplus W$
- Si $\gamma_1 \subseteq U$, $\gamma_2 \subseteq W$ son bases de estos subespacios entonces $\beta := \gamma_1 \cup \gamma_2$ es base de V .

2.12 Caso práctico: Interpolación con polinomios de Lagrange

Relación con el determinante de Vandermonde? paper Uncertainty principles and signal recovery, p. 909. La situación es la siguiente: se tiene un conjunto de $n + 1$ datos

$$\{(c_0, b_0), (c_1, b_1), \dots, (c_n, b_n)\},$$

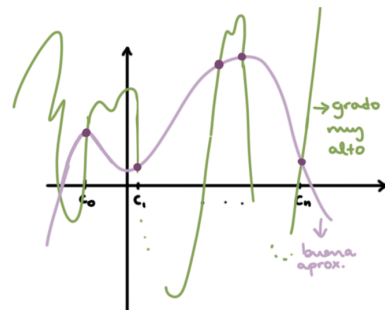
con $c_i \neq c_j$ si $i \neq j$, y se requiere hacer una interpolación de estos con un polinomio del menor grado posible.

- Nos gustaría usar un polinomio, pues es una función con la que es muy fácil trabajar, tanto teórica como prácticamente.
- Además, no queremos que el grado sea alto (relativo a la cantidad de datos) para evitar situaciones como las de la figura: el aumentar el grado del polinomio aumenta el número de ceros de este, luego, la cantidad de oscilaciones, por lo que usar un grado demasiado alto hace que, a pesar de que el polinomio coincida con los puntos b_i en los valores c_i , no modele bien su patrón de comportamiento.

Para $0 \leq i \leq n$, definamos

$$f_i(x) := \frac{(x - c_0)(x - c_1) \cdots (x - c_{i-1})(x - c_{i+1}) \cdots (x - c_n)}{(c_i - c_0)(c_i - c_1) \cdots (c_i - c_{i-1})(c_i - c_{i+1}) \cdots (c_i - c_n)} = \prod_{\substack{j=0 \\ j \neq i}}^n \frac{x - c_j}{c_i - c_j}. \quad (2.27)$$

Interpolación: obtención de nuevos puntos partiendo del conocimiento de un conjunto de puntos.



A estos $n + 1$ polinomios se les llama los **polinomios de Lagrange asociados a** c_0, c_1, \dots, c_n . Por supuesto que

$$f_i(c_j) = \begin{cases} 0 & \text{si } j \neq i, \\ 1 & \text{si } j = i. \end{cases} \quad (2.28)$$

La condición 2.28 es la que define a los polinomios de Lagrange.

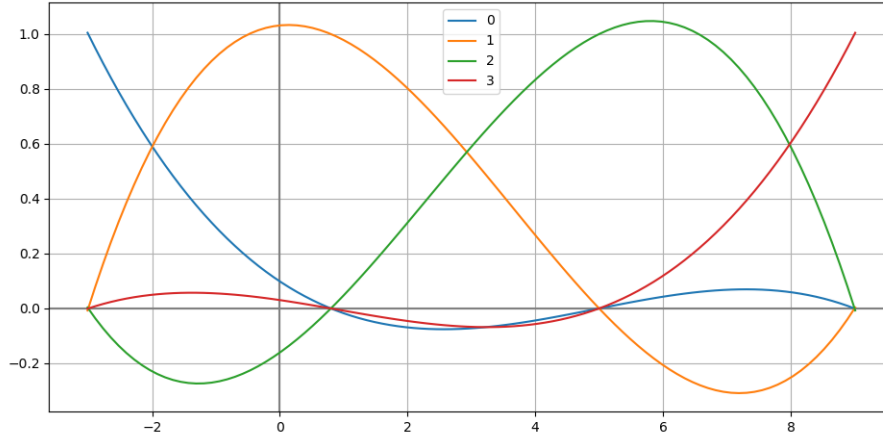


Figure 2.10: Polinomios de Lagrange asociados a la malla $[-3, 0.8, 5, 9]$.

Proposición 2.12.1. Sean $c_0 < c_1 < \dots < c_n$ números reales. Si $\beta = \{f_0, \dots, f_n\}$ es el conjunto de los polinomios de Lagrange (2.27) asociados a estos números, entonces β es una base del F espacio vectorial $P_n(F)$.

Demostración. Si demostramos que el conjunto de $n+1$ vectores β es l.i. podremos concluir que es base de $P_n(F)$ (¿por qué?). Sean $a_0, a_1, \dots, a_n \in F$ tales que $a_0 f_0 + \dots + a_n f_n$ es el polinomio ero. Entonces, para toda $0 \leq j \leq n$, usando la propiedad (2.28) se tiene que

$$\begin{aligned} 0 &= \hat{0}(c_j) = (a_0 f_0 + a_1 f_1 + \dots + a_n f_n)(c_j) \\ &= \sum_{i=0}^n a_i f_i(c_j) = a_j \cdot 1 = a_j. \end{aligned}$$

□

Así, dado $g \in P_n(F)$ cualquiera, existen únicos $b_i \in F$ tales que $g = \sum_{i=0}^n b_i f_i$; de hecho, la propiedad (2.28) nos permite dar explícitamente a tales coeficientes b_i ; para toda $0 \leq j \leq n$,

$$g(c_j) = \left(\sum_{i=0}^n b_i f_i \right)(c_j) = b_j.$$

Así,

$$g = \sum_{i=0}^n g(c_i) f_i. \quad (\text{Ecuación de interpolación de Lagrange}) \quad (2.29)$$

Regresando a la situación planteada al inicio, dada la colección de datos

$$\{(c_0, b_0), (c_1, b_1), \dots, (c_n, b_n)\},$$

el único polinomio g de grado a lo más n tal que $g(c_i) = b_i$ para toda i es (2.29).

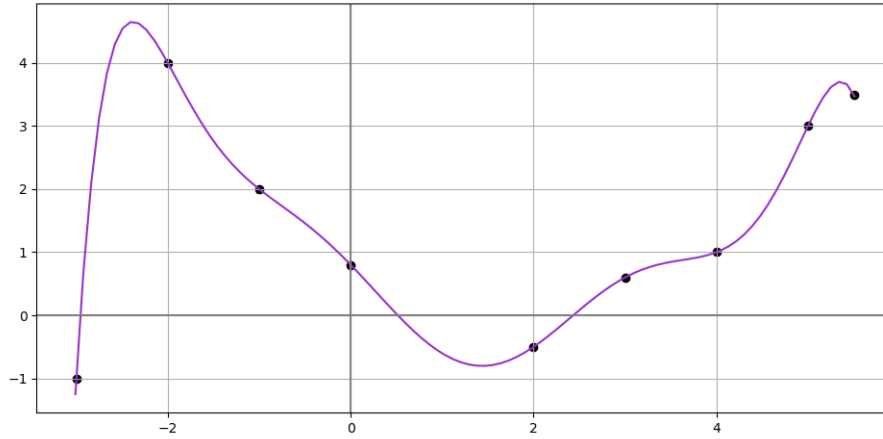


Figure 2.11: Polinomio de interpolación de Lagrange para los datos $(-3, -1)$, $(-2, 4)$, $(-1, 2)$, $(0, 0.8)$, $(2, -0.5)$, $(3, 0.6)$, $(4, 1)$, $(5, 3)$, $(5.5, 3.5)$.

2.13 Algunos resultados de dimensión

Teorema 2.13.1. *Sea V un F -espacio vectorial de dimensión n . Todo subespacio de V será también finito dimensional, y si dimensión será menor o igual a n ; si es igual a n , entonces coincide con todo el espacio V .*

Demostración. Sea $W \leq V$. Si $W = \{0\}$, entonces $\dim(W) = 0 \leq n$. Supongamos ahora que W tiene al menos un elemento no cero x_1 . El singulete $\{x_1\}$ es entonces l.i. (c.f. Proposición 2.9.4). Podemos continuar de esta forma tomando elementos x_1, \dots, x_k de W tales que $\{x_1, \dots, x_k\}$ es l.i. pero, adjuntando otro vector de W , se pierde la independencia lineal (esto porque en V no puede haber subconjuntos linealmente independientes de más de elementos, luego, de hecho ocurre $k \leq n$). Según el lema 2.11.1, esto implica que

$$\forall x \in W - \{x_1, \dots, x_k\} : x \in \langle \{x_1, \dots, x_k\} \rangle,$$

luego,

$$W = \text{span}(\{x_1, \dots, x_k\}).$$

Así, el l.i. $\{x_1, \dots, x_k\}$ genera a W , por lo tanto, es base de W . Entonces, $\dim(W) = k \leq n$.

Si ocurriese $\dim(W) = n$, entonces una base β de W es un subconjunto l.i. de V de $n = \dim(V)$ elementos, por lo tanto, es también una base de V (c.f. corolario 2.11.7), luego,

$$W = \text{span}(\beta) = V.$$

□

Corolario 2.13.2. *Si V es finito dimensional, entonces toda base β de un subespacio de W puede extenderse a una base de V .*

Demostración. Si β es base de W , en particular es un subconjunto l.i. de V , luego, según el Corolario 2.11.15, puede extenderse a una base de V . □

Teorema 2.13.3. *Si W_1, W_2 son dos subespacios de V finito dimensionales, entonces $W_1 + W_2$ es también finito dimensional, y*

$$\dim(W_1 + W_2) = \dim(W_1) + \dim(W_2) - \dim(W_1 \cap W_2).$$

Demostración. Note primero que $W_1 \cap W_2$ es subespacio de un espacio finito dimensional (por ejemplo, de W_1), luego, es también finito dimensional. Sea pues $\beta_0 = \{x_1, \dots, x_k\}$ base de $W_1 \cap W_2$, y sean $\beta_1 = \{y_1, \dots, y_r\}$, $\beta_2 = \{z_1, \dots, z_m\}$ tales que

$$\beta_0 \cup \beta_i \quad \text{es base de } W_i, \quad i = 1, 2.$$

- Veamos que $\beta_0 \cup \beta_1 \cup \beta_2$ es l.i.. Sea escalares a_i , b_j y c_l tales que

$$a_1x_1 + \dots + a_kx_k + b_1y_1 + \dots + b_ry_r + c_1z_1 + \dots + c_mz_m = \hat{0}. \quad (2.30)$$

Definiendo

$$v_0 = a_1x_1 + \dots + a_kx_k, \quad v_1 = b_1y_1 + \dots + b_ry_r, \quad v_2 = c_1z_1 + \dots + c_mz_m,$$

la ecuación (2.30) se reescribe como

$$v_0 + v_1 + v_2 = \hat{0}.$$

Por supuesto que $v_0 \in W_1 \cap W_2$, $v_1 \in W_1$, $v_2 \in W_2$. Despejando a v_2 de esta última ecuación, se tiene que $v_2 = v_0 + v_1 \in W_1$, luego, $v_2 \in W_1 \cap W_2$. Puesto que β_0 es base de este espacio, existen escalares d_i tales que

$$v_2 = d_1x_1 + \dots + d_nx_n; \quad (2.31)$$

sustituyendo en (2.30), se llega a que

$$(a_1 + d_1)x_1 + \dots + (a_k + d_k)x_k + b_1y_1 + \dots + b_ry_r = \hat{0};$$

la independencia lineal de $\beta_0 \cup \beta_1$ implica que todos los escalares de la combinación lineal anterior son cero, en particular, que $b_1 = \dots = b_r = 0$. Sustituyendo esto en (2.30), se tiene que

$$(a_1)x_1 + \dots + (a_k)x_k + c_1z_1 + \dots + c_mz_m = \hat{0}.$$

Ahora la independencia lineal de $\beta_0 \cup \beta_2$ nos permite concluir que también los escalares a_i y c_l son todos cero.

Con esto demostramos la independencia lineal de $\beta_0 \cup \beta_1 \cup \beta_2$. Nota que esto implica que β_0 , β_1 y β_2 son ajenos dos a dos, luego,

$$|\beta_0 \cup \beta_1 \cup \beta_2| = k + r + m. \quad (2.32)$$

- Mostremos que $\beta_0 \cup \beta_1 \cup \beta_2$ genera a $W_1 + W_2$;

Es fácil comprobar que, para cualesquiera subconjuntos $A, B \subseteq V$,

$$\begin{aligned} \langle \beta_0 \cup \beta_1 \cup \beta_2 \rangle &= \langle (\beta_0 \cup \beta_1) \cup (\beta_0 \cup \beta_2) \rangle \\ &= \langle \beta_0 \cup \beta_1 \rangle + \langle \beta_0 \cup \beta_2 \rangle \\ &= W_1 + W_2. \end{aligned} \quad \langle A \cup B \rangle = \langle A \rangle + \langle B \rangle.$$

Hemos demostrado así que $\beta_0 \cup \beta_1 \cup \beta_2$ es base de $W_1 + W_2$; de esto y (2.32) se concluye que

$$\begin{aligned} \dim(W_1 + W_2) &= k + r + m = (k + r) + (k + m) - k \\ &= \dim(W_1) + \dim(W_2) - \dim(W_1 \cap W_2). \end{aligned}$$

□

De este Teorema y de la Proposición 2.6.13 se sigue el siguiente

Corolario 2.13.4. Si W_1 y W_2 son subespacios de V finito dimensionales, entonces la suma $W_1 + W_2$ es directa si y sólo si

$$\dim(W_1 + W_2) = \dim(W_1) + \dim(W_2).$$

Ejemplo 2.13.5. Si U, V son dos subespacios de \mathbb{R}^9 ambos de dimensión 5, entonces su suma no puede ser directa, pues

$$9 \geq \dim(U + V) = \dim(U) + \dim(V) - \dim(U \cap V) = 10 - \dim(U \cap V),$$

luego, $\dim(U \cap V) \geq 1$, por lo que no puede ocurrir $U \cap V = \{0\}$.

Ejemplo 2.13.6. Considere al \mathbb{R} espacio vectorial $P_4(\mathbb{R})$ de los polinomios con coeficientes reales y grado no mayor a cuatro. Sea

$$U = \{f \in P_4(\mathbb{R}) \mid f(6) = 0\}$$

el conjunto de los polinomios de grado a lo más cuatro que tienen al 6 como raíz. Claro que U es un subespacio (propia) de $P_4(\mathbb{R})$, pues

- El polinomio cero tiene a 6 como raíz,
- Si $f, g \in U$ y $a \in \mathbb{R}$, entonces 6 es raíz del polinomio $af + g$, pues $(af + g)(6) = a \cdot f(6) + g(6) = a \cdot 0 + 0 = 0$.

Según el teorema 2.13.1, $0 \leq \dim(U) \leq 4$ (pues $\dim(P_4) = 5$). Observe que los cuatro polinomios

$$f_i(x) := (x - 6)^i, \quad 1 \leq i \leq 4$$

son todos elementos de U , y además son linealmente independientes (sus grados son todos distintos entre sí), luego, ellos conforman una base para el espacio, y $\dim(U) = 4$.

Una base de $P_4(\mathbb{R})$ que contiene a esta base de U es

$$\beta = \{1, f_1, f_2, f_3, f_4\}.$$

Nota que el subespacio

$$W = \{f \in P_4(\mathbb{R}) \mid f \text{ es un polinomio constante o el polinomio cero}\}$$

es tal que $P_4(\mathbb{R}) = U \oplus W$, pues $U \cap W = \{0\}$ - el único elemento de W que tiene a 6 como raíz es el polinomio cero (de hecho, todo número real es raíz del polinomio cero, no hay nada especial en el número 6 tomado para este ejemplo).

Terminemos con unos comentarios más sobre el concepto de dimensión: es muy importante no confundir la noción de dimensión con la de cardinalidad: por ejemplo, en \mathbb{R}^3 ,

- toda recta que pasa por el origen es un subespacio de dimensión uno, y su cardinalidad $|\mathbb{R}|$,
- todo plano que pasa por el origen es un subespacio de dimensión dos, y su cardinalidad es $|\mathbb{R} \times \mathbb{R}| = |\mathbb{R}|$.

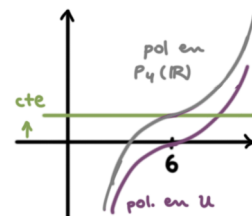


Figure 2.12: Todo polinomio en $P_4(\mathbb{R})$ se expresa de forma única como la suma de un polinomio en U y uno constante.

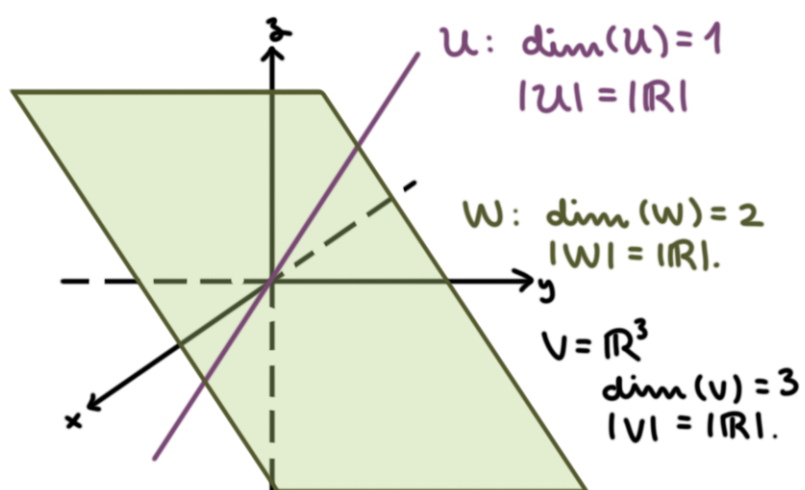


Figure 2.13: Nótese que la dimensión parece ser un mejor indicador del “tamaño” de un subespacio, no la cardinalidad de este.

2.14 Ejercicios II

Ejercicio 15. Demuestre que si $f \in P_n(F)$ es tal que $f(c_j) = 0$ para $n+1$ elementos distintos c_0, \dots, c_n del campo F , entonces f es el polinomio cero. Pista: use polinomios de interpolación de Lagrange

Ejercicio 16. Usando el Teorema 2.13.1, demuestra que

- los únicos subespacios de \mathbb{R}^2 son $\{0\}$, rectas que pasan por el origen y \mathbb{R}^2 , y
- los únicos subespacios de \mathbb{R}^3 son $\{0\}$, rectas y planos que pasan por el origen, y \mathbb{R}^3 .

2.15 Simulacro de examen, primer parcial

Problema 2.15.1. (0.5 p.) En un F -espacio vectorial, ¿el vector cero puede ser elemento de un subconjunto linealmente independiente? Argumenta tu respuesta.

Problema 2.15.2. (0.5 p.) Sean V un F -espacio vectorial, $n \geq 1$ un entero. ¿Qué significa que la dimensión de V sea n ?

Problema 2.15.3. (0.5 p.) Explica por qué el \mathbb{R} -espacio vectorial de los polinomios con coeficientes reales es infinito dimensional.

Problema 2.15.4. (0.5 p.) “Si X genera al F -espacio vectorial V , entonces cualquier vector de V puede expresarse de forma única como combinación lineal de elementos de X ”. ¿Es esto cierto o falso? Argumenta tu respuesta.

Problema 2.15.5. (1 p.) Demuestra que, en un F -espacio vectorial V , un singulete $\{x\}$ es linealmente independiente si y sólo si $x \neq 0$.

Problema 2.15.6. (1 p.) Sea V un \mathbb{Z}_2 espacio vectorial. Demuestre que todo vector coincide con su inverso aditivo, es decir, que

$$\forall v \in V : v = -v.$$

Problema 2.15.7. (1 p.) Sea $X \neq \emptyset$ un conjunto no vacío, F un campo. Considere al F -espacio vectorial F^X de las funciones de X en F . Seleccionando un punto $x_0 \in X$, define

$$W := \{f \in F^X \mid f(x_0) = 0\}.$$

Demuestre que W es un subespacio de F^X .

Problema 2.15.8. (2 p.) Sea

$$U = \{(a, b, c) \in \mathbb{R}^3 \mid a^2 = c^2\}.$$

¿Es U un subespacio del \mathbb{R} -espacio vectorial \mathbb{R}^3 ? Argumenta tu respuesta.

Problema 2.15.9. (3 p.) Sea V un F -espacio vectorial. Sea la familia

$$\mathcal{I} = \{X \subseteq V \mid X \text{ es linealmente independiente}\}.$$

Argumente si las siguientes proposiciones son verdaderas o falsas:

- La familia \mathcal{I} de subconjuntos de V es no vacía.
- Si $X \in \mathcal{I}$, entonces todo subconjunto de X es también elemento de \mathcal{I} .
- Si $X, Y \in \mathcal{I}$, entonces $X \cup Y \in \mathcal{I}$.

Problema 2.15.10. (3 p.) Considere al \mathbb{R} -espacio vectorial \mathbb{R}^3 . Si

$$X = \{v_1 = (1, 1, 1), v_2 = (0, 1, -1), v_3 = (1, 0, 2)\}.$$

Demuestra que el vector $u = (4, 2, 1) \in \mathbb{R}^3$ **no** es elemento del generado de X .

Problema 2.15.11. (3 p.) del problema (2.15.10) se infiere que el conjunto X (que tiene 3 elementos) no genera a \mathbb{R}^3 (siendo este un espacio de dimensión 3), por lo tanto, no puede ser linealmente independiente. Quitá un vector v de X de tal forma que $X - \{v\}$ sea linealmente independiente, y extiende este subconjunto linealmente independiente a una base de \mathbb{R}^3

Estos fueron los problemas del examen parcial aplicado al curso del Otoño 2024. Puedes contestar todos los que quieras en dos horas, y dependiendo de qué tan bien está tu solución, podrás ganar una fracción de los puntos que vale cada problema. Si acumulas 10 o más puntos, tu calificación es 10.

Problema 2.15.12. (3 p.) El conjunto

$$\beta = \left\{ A = \begin{pmatrix} 0 & 1 \\ 2 & 3 \end{pmatrix}, B = \begin{pmatrix} 1 & 12 \\ 14 & 8 \end{pmatrix}, C = \begin{pmatrix} 1 & 9 \\ 8 & -1 \end{pmatrix}, D = \begin{pmatrix} 1 & 7 \\ 6 & 2 \end{pmatrix}, E = \begin{pmatrix} 1 & 6 \\ 8 & 21 \end{pmatrix}, F = \begin{pmatrix} 1 & 5 \\ 4 & 7 \end{pmatrix} \right\}$$

genera al \mathbb{R} -espacio vectorial $M_{2 \times 2}(\mathbb{R})$. De este generador extrae una base para el espacio $M_{2 \times 2}(\mathbb{R})$.

Chapter 3

Transformaciones lineales

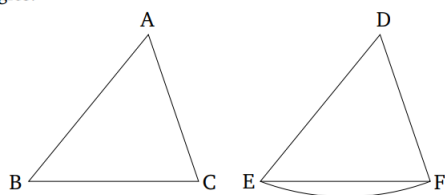
Ya definimos y estudiamos el concepto de espacio vectorial, dimos también unos ejemplos que ilustran lo útil que puede ser considerar espacios vectoriales como marcos teóricos para modelar situaciones prácticas. Lo que queremos hacer ahora es establecer relaciones entre espacios vectoriales; más formalmente, queremos trabajar con funciones de un espacio vectorial a otro que preserven la estructura algebraica de estos. Como veremos, algunos de los conceptos matemáticos más usados pueden verse de forma natural como transformaciones lineales, por ejemplo,

- las operaciones de integración y diferenciación del cálculo, o
- las proyecciones, reflexiones y proyecciones de la geometría.

Explicaremos además cómo codificar la información de una transformación lineal entre espacios vectoriales finito dimensionales en una matriz (que dependerá de las bases escogidas para el dominio y codominio), hecho que hará que la teoría desarrollada sea fácilmente llevada a la práctica - pues estaremos sustituyendo a las funciones por objetos discretos.

Proposition 4

If two triangles have two sides equal to two sides, respectively, and have the angle(s) enclosed by the equal straight-lines equal, then they will also have the base equal to the base, and the triangle will be equal to the triangle, and the remaining angles subtended by the equal sides will be equal to the corresponding remaining angles.



(That is) ABC to DEF ,

and ACB to DFE .

For if triangle ABC is applied to triangle DEF ,[†] the point A being placed on the point D , and the straight-line

Proposition 12

To draw a straight-line perpendicular to a given infinite straight-line from a given point which is not on it.

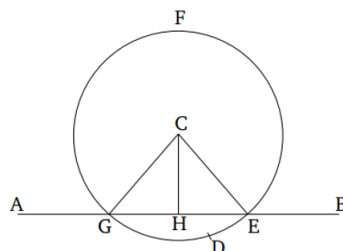


Figure 3.1: Euclides usa en sus argumentos transformaciones lineales como las traslaciones, rotaciones y proyecciones.

3.1 Transformaciones lineales

Definición 3.1.1. Sean V, W dos espacios vectoriales sobre un mismo campo F . Toda función $T: V \rightarrow W$ tal que

$$(\forall x, y \in V)(\forall a \in F): T(ax + y) = aT(x) + T(y) \quad (3.1)$$

será llamada una **transformación lineal** de V en W .

Nota que V y W deben ser espacios vectoriales sobre un campo común para que ambos lados de la ecuación 3.1 tengan sentido.

Observación 3.1.2. Si $T: V \rightarrow W$ es una transformación lineal, entonces

- $T(0_V) = T(0_W)$, y
- Para cualesquiera $n \geq 1$ entero, $a_1, \dots, a_n \in F$ y $x_1, \dots, x_n \in V$,

$$T\left(\sum_{i=1}^n a_i x_i\right) = \sum_{i=1}^n a_i T(x_i).$$

Demostración. En efecto, tenemos la siguiente ecuación en el grupo abeliano W

$$T(0_V) = T(0_V + 0_V) = T(0_V) + T(0_V),$$

luego, $T(0_V) = 0_W$. El segundo punto se demuestra por inducción, usando a la condición (3.1) como base de inducción. \square

Ejercicio: demuestra que la condición 3.1 es equivalente a las siguientes dos condiciones:

1. (**aditividad**) Para cualesquiera $x, y \in V: T(x + y) = T(x) + T(y)$,
2. (**homogeneidad**) Para cualesquiera $x \in V, a \in F, T(ax) = aT(x)$.

Estas condiciones dicen que, no importa si primero se realiza la suma o multiplicación escalar en el espacio V y luego se aplica la transformación lineal, o si primero se llevan los vectores a W via T y luego se efectúa la suma o multiplicación escalar en W , el resultado es el mismo.

3.1.1 Ejemplos de transformaciones lineales

Tenemos la siguiente lista de ejemplos canónicos de transformaciones lineales.

- Dados V y W F -espacios vectoriales, las funciones $I_V: V \rightarrow V$ y $T_0: V \rightarrow W$ definidas como

$$\forall x \in V: I_V(x) = x, T_0(x) = 0_W$$

son transformaciones lineales, llamadas respectivamente la **transformación identidad** y la **transformación cero** en V .

- La función $T: P_n(\mathbb{R}) \rightarrow P_{n-1}(\mathbb{R})$ definida como

$$\forall f \in P_n(\mathbb{R}): T(f) = f'$$

es una transformación lineal.

- La función $T: \mathcal{C}(\mathbb{R}) \rightarrow \mathbb{R}$ definida como

$$\forall f \in \mathcal{C}(\mathbb{R}): T(f) = \int_a^b f(t) dt$$

es una transformación lineal.

- Sea $0 \leq \theta < 2\pi$. La función $T_\theta : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ definida como

$$\forall (a, b) \in \mathbb{R}^2 : T_\theta(a, b) = (a \cos(\theta) - b \sin(\theta), a \sin(\theta) + b \cos(\theta))$$

es una transformación lineal llamada **rotación de θ radianes**.

- La función $T : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ definida como

$$\forall (a, b) \in \mathbb{R}^2 : T_\theta(a, b) = (a, -b)$$

es llamada la **reflexión sobre el eje x** , y es una transformación lineal tal que $T \circ T = I$.

- La función $T : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ definida como

$$\forall (a, b) \in \mathbb{R}^2 : T_\theta(a, b) = (a, 0)$$

es una transformación lineal, que llamamos **proyección sobre el eje- x** .

Profundicemos un poco más la definición de proyección. Claramente, $W_1 = \{(a, 0) : a \in \mathbb{R}\}$ y $W_2 = \{(0, b) : b \in \mathbb{R}\}$ son subespacios de \mathbb{R}^2 tales que $\mathbb{R}^2 = W_1 \oplus W_2$. Se definió arriba a la proyección sobre el eje- x a la función que, a cada $v \in \mathbb{R}^2$, le asigna su sumando correspondiente al espacio W_1 . Definamos en general el término proyección.

Definición 3.1.3. Sea V un F -espacio vectorial, $W_1 \leq V$. Si $W_2 \leq V$ es tal que $W_1 \oplus W_2 = V$, la función $T : V \longrightarrow V$ definida como

$$T(x) = x_1, \quad x = x_1 + x_2, \quad x_1 \in W_1, x_2 \in W_2 \quad (3.2)$$

es llamada una **proyección sobre W_1** .

Nota 3.1.4. Si U, V, W son subespacios de un F -espacio vectorial, la igualdad $U \oplus V = U \oplus W$ **no implica** la igualdad $V = W$.

Por ejemplo, considérese a los subespacios de \mathbb{R}^2

$$U = \text{span}(\{(1, 0)\}), \quad \text{eje } x,$$

$$V = \text{span}(\{(0, 1)\}), \quad \text{eje } y,$$

$$W = \text{span}(\{(1, 1)\}), \quad \text{gráfica de la recta } y = x.$$

Puesto que $U \cap V = \{0\} = U \cap W$, la suma de U con V y de U con W es directa (c.f. Proposición 2.6.13), y de hecho es todo el espacio:

$$U \oplus V = \mathbb{R}^2 = U \oplus W.$$

Sin embargo, claro que $V \neq W$. Observa que

$$\forall (x, y) \in \mathbb{R}^2 : x(1, 0) + y(0, 1) = (x, y) = (x - y)(1, 0) + y(1, 1).$$

◇

Observe que, como se usa una suma directa para definir una proyección, la expresión (3.2) en efecto define una función T , de hecho lineal, pues, si $x = x_1 + x_2$, $y = y_1 + y_2$ y $a \in F$, entonces,

$$T(ax + y) = T((ax_1 + y_1) + (x_2 + y_2)) = ax_1 + x_2 = aT(x_1) + T(x_2).$$

Note que $W_1 = \{x \in V \mid T(x) = x\}$, es decir, el conjunto de puntos fijos de una proyección en W_1 coincide con W_1 . En la Definición 3.1.3 hablamos de “una” proyección a W_1 ; esto es porque, como mostramos a continuación, hay tantas transformaciones lineales que satisfacen la definición de proyección a W_1 como subespacios cuya suma directa con W_1 es todo el espacio.

Proposición 3.1.5. Sea $W_1 \leq V$. Si W_2, \tilde{W}_2 son dos subespacios de V distintos entre si tales que $V = W_1 \oplus W_2 = W_1 \oplus \tilde{W}_2$, entonces las transformaciones lineales $T, \tilde{T} : V \longrightarrow V$ definidas como

$$T(x) = x_1, \quad x = x_1 + x_2, \quad x_1 \in W_1, x_2 \in W_2$$

y

$$\tilde{T}(x) = \tilde{x}_1, \quad x = \tilde{x}_1 + \tilde{x}_2, \quad \tilde{x}_1 \in W_1, \tilde{x}_2 \in \tilde{W}_2$$

son distintas entre si.

Demostración. Busquemos un punto en el que T y \tilde{T} difieren. Como $W_2 \neq \tilde{W}_2$, sin pérdida de generalidad podemos suponer que existe $y \in \tilde{W}_2 - W_2$. Se tiene que

$$x_1 + x_2 = y = 0 + y,$$

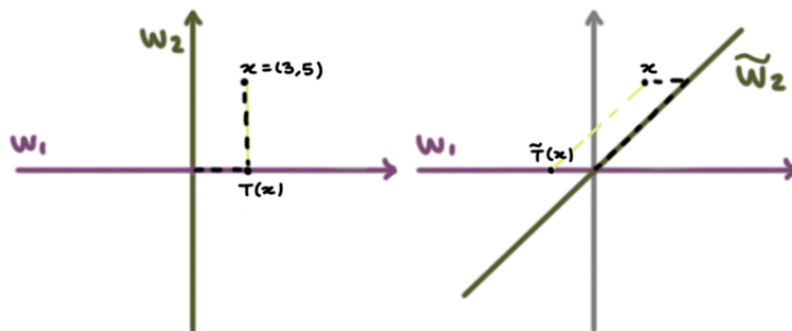
con $x_1 \in W_1, x_2 \in W_2$ y $y \in \tilde{W}_2$. Note que x_1 no es cero, de lo contrario, se tendría

$$y = x_2 \in W_2 \quad \nexists$$

Así,

$$T(y) = x_1 \neq 0 = \tilde{T}(y).$$

□



Ejemplo 3.1.6. *Poner este ejemplo mucho antes.* Sea V un F -espacio vectorial con $\dim(V) = n$. Si β es una base de V , dividámosla en dos subconjuntos α, γ ajenos cuya unión sea β . Entonces, $V = \langle \alpha \rangle \oplus \langle \gamma \rangle$, pues

- $V = \langle \alpha \cup \gamma \rangle = \langle \alpha \rangle + \langle \gamma \rangle$, y
- la suma es directa, pues

$$\begin{aligned} \dim(V) &= \dim(\langle \alpha \rangle + \langle \gamma \rangle) = \dim(\langle \alpha \rangle) + \dim(\langle \gamma \rangle) - \dim(\langle \alpha \rangle \cap \langle \gamma \rangle) \\ &= \dim(V) + \dim(\langle \alpha \rangle \cap \langle \gamma \rangle), \end{aligned}$$

luego, $\dim(\langle \alpha \rangle \cap \langle \gamma \rangle) = 0$, o sea, $\langle \alpha \rangle \cap \langle \gamma \rangle = \{0\}$.

3.2 El teorema fundamental de las transformaciones lineales

La condición (3.1) significa que la función T preserva la estructura algebraica de V . Como veremos a continuación, toda transformación lineal también preserva subespacios.

Proposición 3.2.1. Sean V y W F -espacios vectoriales, $T : V \rightarrow W$ una transformación lineal entre estos.

- La imagen de todo subespacio de V bajo T es un subespacio de W , es decir,

$$\forall X \subseteq V : X \leq V \Rightarrow T(X) \leq W.$$

- La preimagen de todo subespacio de W bajo T es un subespacio de V , es decir,

$$\forall Y \subseteq W : Y \leq W \Rightarrow T^{-1}(Y) = \{x \in V \mid T(x) \in Y\} \leq V.$$

Demostración. En efecto, si X es un subespacio de V , entonces $0_V \in X$, luego, $0_W = T(0_V) \in T(X)$. Además, si a es un escalar cualquiera y $y_1, y_2 \in T(X)$, entonces existen $x_1, x_2 \in X$ tales que $y_i = T(x_i)$, con $i = 1, 2$. Así,

$$ay_1 + y_2 = aT(x_1) + T(x_2) = T(ax_1 + x_2)$$

es elemento de $T(X)$ pues X , al ser subespacio de V , contiene a $ax_1 + x_2$. La demostración del segundo punto es dual. \square

Vamos ahora a asociar a una transformación lineal dos espacios vectoriales (uno será un subespacio del dominio, otro del codominio) de gran importancia.

Definición 3.2.2. Sean V, W F -espacios vectoriales, $T : V \rightarrow W$ una transformación lineal.

- Se define al **espacio nulo** de T o **kernel** de T como

$$\text{Ker}(T) := T^{-1}(\{0_W\}) \leq V. \quad (3.3)$$

- La **imagen** de T es

$$T(V) \leq W. \quad (3.4)$$

Si $\text{Ker}(T)$ es finito dimensional, a su dimensión se le denomina la **nulidad** de T . Si $T(V)$ es finito dimensional, su dimensión se conoce como el **rango** de T .

Observa que el “tamaño” del kernel de una transformación lineal parece ser inversamente proporcional al de la imagen de esta; en efecto, si muchos vectores pertenecen al kernel, entonces no habrá muchos vectores en la imagen de T que no sean cero, lo que achica a $T(V)$. El siguiente teorema, pilar del álgebra lineal, da forma a esta intuición.

Teorema 3.2.3. (fundamental de las transformaciones lineales), teorema de la dimensión Sean V, W F -espacios vectoriales, con V finito dimensional. Para toda transformación lineal $T : V \rightarrow W$,

$$\dim(V) = \dim(\text{Ker}(T)) + \dim(T(V)), \quad (3.5)$$

es decir, la dimensión del espacio de origen V es igual a la nulidad de T más el rango de T .

Recuerda que, si $f : A \rightarrow B$ es una función cualquiera y $X \subseteq A, Y \subseteq B$, definimos

$$f(X) := \{y \in B \mid \exists x \in X : f(x) = y\},$$

y

$$f^{-1}(B) = \{x \in A \mid f(x) \in B\}$$

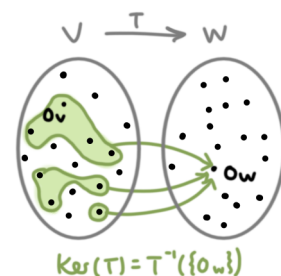
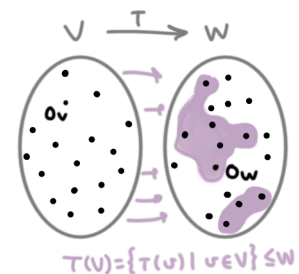


Figure 3.2: Etimológicamente, “kernel” significa semilla, centro, esencia.



Demostración. Como V es finito dimensional, el kernel de T también lo es. Sea pues $\{n_1, \dots, n_k\}$ una base de $\text{Ker}(T)$. Extendamos este subconjunto l.i. de V a una base de V ; sea $\{v_1, \dots, v_l\}$ tal que

$$\beta := \{n_1, \dots, n_k\} \cup \{v_1, \dots, v_l\}$$

es base de V . Entonces, $\dim(V) = k + l$. Afirmamos que $\{T(v_1), \dots, T(v_l)\}$ es base de la imagen $T(V)$.

El Teorema 3.2.3 es también conocido como el “Teorema de la dimensión”.

- Sean a_1, \dots, a_l escalares tales que

$$a_1 T(v_1) + \dots + a_l T(v_l) = 0_W$$

Supongamos que alguno de ellos no es cero; sin pérdida de generalidad, digamos que $a_1 \neq 0$. Por la linealidad de T , la ecuación anterior se reescribe como

$$T(a_1 v_1 + \dots + a_l v_l) = 0_W,$$

es decir,

$$a_1 v_1 + \dots + a_l v_l \in \text{Ker}(T) = \text{span}(\{n_1, \dots, n_k\}),$$

por lo tanto,

$$v_1 \in \text{span}(\{n_1, \dots, n_k\} \cup \{v_2, \dots, v_l\}).$$

Esto, según el Lema 2.11.1, contradice la independencia lineal de β .

- Claro que $\{T(v_1), \dots, T(v_l)\}$ genera a $T(V)$ pues, dado $y \in T(V)$, existe $x \in V$ tal que $y = T(x)$. Como β es base de V , existen escalares a_i, b_j tales que

$$x = a_1 n_1 + \dots + a_k n_k + b_1 v_1 + \dots + b_l v_l;$$

evaluando ambos lados de la igualdad bajo T y recordando que los vectores n_i son mapeados al cero (pues son elementos del kernel de T), concluimos que

$$y = T(x) = b_1 T(v_1) + \dots + b_l T(v_l).$$

Así,

$$\dim(V) = k + l = \dim(\text{Ker}(T)) + \dim(T(V)).$$

□

Poner una nota para que ya no confundan el rango de T con el codominio de T .

Corolario 3.2.4. Sean V y W son F -espacios vectoriales con V finito dimensional, $T : V \rightarrow W$ lineal. Para toda $\beta \subseteq V$ base de V se tiene que $T(\beta)$ genera a $T(V)$.

3.2.1 No ejemplos de transformaciones lineales

- La función $f : \mathbb{R} \rightarrow \mathbb{R}$ definida como

$$f(x) = e^x$$

no es lineal, pues $f(0) \neq 0$ (de hecho, la exponencial es una función estrictamente positiva).

- La función valor absoluto $|\cdot| : \mathbb{R} \rightarrow \mathbb{R}$ no es lineal. Por ejemplo,

$$|2 + (-1)| = |1| = 1 \neq 3 = |2| + |-1|.$$

- El polinomio $f : \mathbb{R} \rightarrow \mathbb{R}$ definido como $f(x) = x^2$ no es una transformación lineal pues, en general, no se cumple que $(x + y)^2$ coincida con $x^2 + y^2$. De hecho, lo que ocurre es $(x + y)^2 = x^2 + 2xy + y^2$.
- Sea la función $f : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ definida como

$$f(x, y) = \begin{cases} (2x, 0) & \text{si } y = 0 \\ (x, y) & \text{si } y \neq 0. \end{cases}$$

Es fácil comprobar que, a pesar de que f es homogénea, no es lineal. Por ejemplo,

$$f((1, 0) + (0, 1)) = f(1, 1) = (1, 1),$$

pero

$$f(1, 0) + f(0, 1) = (2, 0) + (0, 1) = (2, 1).$$

- La función coseno evaluada en cero vale uno, luego, no es lineal.
- Según el Teorema 3.2.3, si $T : \mathbb{R} \rightarrow \mathbb{R}$ es lineal, entonces

$$1 = \dim(\text{Ker}(T)) + \dim(T(V)),$$

luego, tenemos dos casos:

1. $\dim(T(V)) = 1$, luego, como $T(V)$ es subespacio de \mathbb{R} , con \mathbb{R} uno dimensional, tenemos que $T(V)$ es suprayectiva.
2. $\dim(T(V)) = 0$, es decir, $\dim(\text{Ker}(T)) = 1$. Puesto que $\text{Ker}(T) \leq \mathbb{R}$ y $\dim(\mathbb{R}) = 1$, se tiene que $\text{Ker}(T) = \mathbb{R}$, luego, T es la transformación lineal cero.

Puesto que la función seno

$$\text{sen}(x) : \mathbb{R} \rightarrow \mathbb{R}$$

no es ni la función cero ni suprayectiva, no puede ser lineal.

3.3 Inyectividad y suprayectividad de transformaciones lineales

Recuerda que, en general, si $f : V \rightarrow W$ es una función de un conjunto V a otro conjunto W , f se dice

- **inyectiva** si

$$\forall x, y \in V : T(x) = T(y) \Rightarrow x = y,$$

o, equivalentemente, si las imágenes de puntos distintos son distintas

- **suprayectiva** si

$$(\forall z \in W) (\exists x \in V) : T(x) = z.$$

Resulta que, si V y W son F -espacios vectoriales y $T : V \rightarrow W$ es, no sólo una función, sino una transformación lineal entre ellos, entonces podemos encontrar equivalencias de ser inyectiva o suprayectiva usando propiedades del Kernel y la preservación de la generación o inyectividad de subconjuntos de V .

Veremos que, en el contexto de transformaciones lineales, los conceptos de inyectividad e independencia lineal están íntimamente ligados, así como los de suprayectividad y generación.

Proposición 3.3.1. Sean V, W dos F -espacios vectoriales. Si $T : V \longrightarrow W$ es lineal, las siguientes son equivalentes:

1. T es inyectiva
2. $\text{Ker}(T) = \{0\}$
3. Si $X \subseteq V$ es linealmente independiente, entonces $T(X) \subseteq W$ es también linealmente independiente.

Demostración.

1) \Rightarrow 2) Si $x \in \text{Ker}(T)$ entonces, por definición del kernel, $T(x) = 0$. Además, como T es lineal, también se tiene $T(0) = 0$, luego, como T es inyectiva, tenemos que $x = 0$. Así, $\text{Ker}(T) = \{0\}$.

2) \Rightarrow 1) Sean $x, y \in V$ tales que $T(x) = T(y)$. Entonces, por la linealidad de T , $T(x - y) = T(x) - T(y) = 0$, así, $x - y \in \text{Ker}(T) = \{0\}$, es decir, $x - y = 0$, o sea, $x = y$.

2) \Rightarrow 3) Sean $v_1, \dots, v_n \in X$ cualesquiera; mostremos que $\{T(v_1), \dots, T(v_n)\}$ es linealmente independiente. Sean $a_i \in F$ escalares tales que

$$a_1 T(v_1) + \dots + a_n T(v_n) = \hat{0}_W.$$

Por ser T lineal, podemos reescribir el lado izquierdo de la igualdad anterior;

$$T(a_1 v_1 + \dots + a_n v_n) = \hat{0}_W.$$

Esto muestra que $a_1 v_1 + \dots + a_n v_n \in \text{Ker}(T) = \{0\}$, luego,

$$a_1 v_1 + \dots + a_n v_n = 0_V.$$

Como X es l.i., esto implica que todos los escalares a_i son cero.

3) \Rightarrow 2) Supongamos que existe $x \in \text{Ker}(T) - \{0\}$. Como x no es el vector cero, el singulete $\{x\}$ es l.i. y, sin embargo, $\{T(x)\} = \{0_W\}$ es l.d.. Esto contradice nuestra hipótesis.

□

Proposición 3.3.2. Sean V, W dos F -espacios vectoriales. Si $T : V \longrightarrow W$ es lineal, las siguientes son equivalentes:

1. T es suprayectiva
2. Si $X \subseteq V$ genera a V entonces $T(X)$ genera a W .

Demostración.

1) \Rightarrow 2) Sea X un generador de V , es decir, un subconjunto tal que $\text{span}(X) = V$. Esto significa que el único subespacio de V que contiene a X es el mismo V . Puesto que $X \subseteq T^{-1}(\text{span}(T(X)))$ (pues, dado $x \in X$, $T(x) \in T(X) \subseteq \text{span}(T(X))$), se tiene entonces $T^{-1}(\text{span}(T(X))) = V$. Evaluando ambos lados de la igualdad bajo la función suprayectiva T concluimos que

$$W = T(V) = T(T^{-1}(\text{span}(T(X)))) = \text{span}(T(X)),$$

o sea, que $T(X)$ genera a W .

En la implicación 2) \Rightarrow 3), estamos mostrando que un subconjunto finito arbitrario de $T(X)$ es l.i. suponiendo que X es l.i.. Recuerda que esto es necesario y suficiente para demostrar la independencia lineal de todo $T(X)$.

Se demuestra que si $f : A \longrightarrow B$ es una función suprayectiva, entonces, para todo $Y \subseteq B$, se tiene $f(f^{-1}(Y)) = Y$.

2) \Rightarrow 1) V trivialmente se genera a sí mismo, luego, por hipótesis debe ocurrir que $T(V)$ genere a W , o sea, que

$$W = \text{span}(T(V)) = T(V),$$

por lo tanto T es suprayectiva.

□

Proposición 3.3.3. Sean V y W dos F -espacios vectoriales finito dimensionales.

- Si $\dim(V) > \dim(W)$, entonces no existen transformaciones lineales de V en W inyectivas.
- Si $\dim(V) < \dim(W)$, entonces no existen transformaciones lineales de V en W suprayectivas.

Demostración. En efecto,

- si existe $T : V \rightarrow W$ inyectiva, entonces, según la proposición 3.3.1, $\text{Ker}(T) = \{0\}$, luego, la ecuación (3.5) se reescribe como

$$\dim(V) = \dim(T(V)) \leq \dim(W).$$

- Si existe $T : V \rightarrow W$ suprayectiva, entonces $T(V) = W$, luego, (3.5) se reescribe como

$$\dim(V) = \dim(\text{Ker}(T)) + \dim(W) \geq \dim(W).$$

□

3.4 Isomorfismos

Definición 3.4.1. Sean V, W dos F -espacios vectoriales. Toda transformación lineal $T : V \rightarrow W$ que sea biyectiva (i.e. inyectiva y suprayectiva) será llamada un **isomorfismo**. Si existe un isomorfismo entre dos espacios vectoriales V y W decimos que V y W son **isomorfos**.

Es decir, una transformación lineal $T : V \rightarrow W$ es un isomorfismo si $\text{Ker}(T) = \{0_V\}$ y $T(V) = W$.

Recuerda de tus cursos anteriores que una función $f : A \rightarrow B$ es biyectiva si y sólo si es invertible (i.e. si y sólo si existe una función $g : B \rightarrow A$ tal que $f \circ g = \text{Id}_B$ y $g \circ f = \text{Id}_A$). Entonces, una transformación lineal $T : V \rightarrow W$ es un isomorfismo si y sólo si existe una función $U : W \rightarrow V$ que sea su inversa. Claro que tal inversa de existir es única, y se le suele denotar por T^{-1} . Como establecemos a continuación, tal función es, al igual que T , una transformación lineal.

Proposición 3.4.2. Si $T : V \rightarrow W$ es un isomorfismo, entonces su función inversa T^{-1} es también un isomorfismo.

Demostración. Basta probar la linealidad de T^{-1} . Sean pues $y, z \in W$, $\lambda \in F$. Puesto que T es lineal y $u := T^{-1}(y)$, $v := T^{-1}(z)$ son vectores de V , se tiene que

$$T(\lambda u + v) = \lambda T(u) + T(v) = \lambda y + z,$$

luego,

$$T^{-1}(\lambda y + z) = \lambda u + v = \lambda T^{-1}(y) + T^{-1}(z).$$

□

Mostremos ahora que, si V es finito dimensional, entonces puede establecerse un isomorfismo entre V y otro F -espacio vectorial W sólo si W tiene la misma dimensión que V .

Proposición 3.4.3. *Sea $T : V \longrightarrow W$ lineal, con V de dimensión finita. Si T es un isomorfismo, entonces W también es finito dimensional y, de hecho, $\dim(W) = \dim(V)$.*

Demostración. Digamos que $\dim(V) = n$. Por ser T un isomorfismo, se tiene que

$$\text{Ker}(T) = \{0_V\} \quad \text{y} \quad T(V) = W.$$

Además, como V es finito dimensional, podemos usar el Teorema 3.2.3 para deducir que

$$\begin{aligned} n = \dim(V) &= \dim(\text{Ker}(T)) + \dim(T(V)) \\ &= \dim(\{0_V\}) + \dim(W) \\ &= 0 + \dim(W) = \dim(W). \end{aligned}$$

□

Mostremos ahora que, cuando se trata con espacios vectoriales finito dimensionales, los conceptos de inyectividad, suprayectividad y biyectividad en transformaciones lineales son equivalentes.

Teorema 3.4.4. *Sean V, W dos F -espacios vectoriales, ambos de dimensión n . Entonces, para cualquier transformación lineal $T : V \longrightarrow W$, son equivalentes*

1. T es inyectiva
2. T es suprayectiva
3. T es biyectiva (i.e. un isomorfismo).

Demostración. Basta demostrar que 1) implica 2) y que 2) implica 3).

1) \Rightarrow 2) Si T es inyectiva entonces $\text{Ker}(T) = \{0\}$, luego, por el Teorema 3.2.3,

$$n = \dim(V) = \dim(\text{Ker}(T)) + \dim(T(V)) = 0 + \dim(T(V)) = \dim(T(V)),$$

luego, como $\dim(W) = n$ y $T(V) \leq W$ tiene dimensión n , concluimos que $T(V) = W$, o sea, que T es suprayectiva.

2) \Rightarrow 3) Si $T(V) = W$, entonces

$$n = \dim(\text{Ker}(T)) + \dim(T(V)) = \dim(\text{Ker}(T)) + \dim(W) = \dim(\text{Ker}(T)) + n,$$

luego, $\dim(\text{Ker}(T)) = 0$ o, equivalentemente, $\text{Ker}(T) = \{0_V\}$, i.e. T es inyectiva.

□

Ejemplo 3.4.5. (para mostrar la importancia de la hipótesis de dimensión finita en el Teorema 3.4.4) Considere al \mathbb{R} -espacio vectorial $\mathbb{R}^{\mathbb{N}}$ de sucesiones en \mathbb{R} (c.f. Sección 2.3.2). Sean $L, R: \mathbb{R}^{\mathbb{N}} \rightarrow \mathbb{R}^{\mathbb{N}}$ las funciones definidas como

$$\forall x = (x_1, x_2, x_3, \dots) \in \mathbb{R}^{\mathbb{N}}: \quad R(x) = (0, x_1, x_2, \dots), \quad L(x) = (x_2, x_3, \dots). \quad (3.6)$$

Claro que R y L son ambas lineales, y que $L \circ R = Id_{\mathbb{R}^{\mathbb{N}}}$, luego, si R tiene inversa, tiene que ser L . Sin embargo, $R \circ L \neq Id_{\mathbb{R}^{\mathbb{N}}}$, por lo tanto, R no es invertible, luego, no es un isomorfismo. Sin embargo, sí es inyectiva. Similarmente puede notar que L no es un isomorfismo, pero que sí es suprayectiva. \diamond

Ejemplo 3.4.6. Sea la función $T: P_2(\mathbb{R}) \rightarrow P_3(\mathbb{R})$ definida como

$$T(f)(u) = 2f'(u) + \int_0^u 3f(x)dx. \quad (\text{polinomio en la variable } u).$$

Puesto que T es combinación lineal de transformaciones lineales, es también lineal. Según el Corolario 3.2.4, $\{T(1), T(x), T(x^2)\}$ genera al rango de T . Se calcula que

$$T(1)(u) = 2 \cdot 0 + \int_0^u 3dx = 3u,$$

$$T(x)(u) = 2 + \int_0^u 3xdx = 2 + \frac{3}{2}x^2 \Big|_{x=0}^{x=u} = 2 + \frac{3u^2}{2},$$

$$T(x^2)(u) = 4u + \int_0^u 3x^2dx = 4u + x^3 \Big|_{x=0}^{x=u} = 4u + u^3.$$

Tenemos entonces que

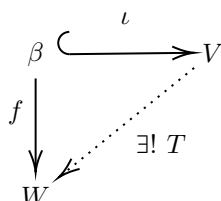
$$\{g_1(u) = 3u, g_2(u) = 2 + (3/2)u^2, g_3(u) = 4u + u^3\}$$

genera a $T(P_2(\mathbb{R}))$; puesto que los elementos de este generador son polinomios de grados distintos, de hecho son linealmente independientes, luego, esta es una base del rango de T . Así, como $\dim(P_2(\mathbb{R})) = 3$, se debe tener que T es inyectiva. \diamond

3.5 La propiedad universal de las bases

Recuerda que, dados V y W dos F -espacios vectoriales, una transformación lineal $T: V \rightarrow W$ es una función que además “abre” combinaciones lineales. Como cualquier función, T está completamente determinada por los valores que toma en su dominio V . Como veremos a continuación, en el caso de las transformaciones lineales, estas de hecho quedan determinadas por su definición en una base cualquiera del espacio V ; conocer los valores de T en una base de V nos permite saber los valores de T en cualquier punto de V .

Teorema 3.5.1. (propiedad universal de las bases) Sea V un F -espacio vectorial, con V finito-dimensional. Son equivalentes para $\beta \subseteq V$ las siguientes:



A las transformaciones (3.6) se les conoce como “right shift” y “left shift”, resp.

Al Teorema 3.5.1 también se le conoce como el Teorema fundamental de las bases. Se usará en repetidas ocasiones para desarrollar la teoría de las siguientes secciones, por lo que se recomienda entenderlo bien.

- β es base de V
- Si W es un F -espacio vectorial cualquiera, toda función $f : \beta \longrightarrow W$ se puede extender linealmente de forma única a todo V , es decir, existe una única transformación lineal $T : V \longrightarrow W$ tal que $T|_{\beta} = f$.

Nota: en realidad, este teorema es cierto aún cuando V es infinito dimensional (c.f. [Mej11] p. 84). Sin embargo, como en el curso no hemos demostrado los resultados necesarios para probar esto cuando se trabaja en un espacio infinito dimensional (no estudiamos el Lema de Zorn, por lo que no pudimos demostrar hechos fundamentales como la existencia de bases para cualquier espacio vectorial o el hecho de que un l.i. de un espacio infinito dimensional puede extenderse a una base de este), nos limitaremos a establecer y probar la propiedad universal de las bases en dimensión finita - que en realidad son los tipos de espacios que se usan siempre en las aplicaciones.

Demostración.

\Rightarrow) Sea $f : \beta \longrightarrow W$ una función de la base $\beta = \{v_1, \dots, v_n\}$ de V escogida a W . Sean $x, y \in V$ cualesquiera; digamos que

$$x = \sum_{i=1}^n a_i v_i, \quad y = \sum_{i=1}^n b_i v_i. \quad (3.7)$$

Observe que, si $T : V \longrightarrow W$ es una transformación lineal que extiende a f (i.e. tal que $T \circ \iota = f$), entonces, deberá ocurrir

$$T(x) = T\left(\sum_{i=1}^n a_i v_i\right) = \sum_{i=1}^n a_i T(v_i) = \sum_{i=1}^n a_i f(v_i),$$

es decir, T tiene que ser la función definida como

$$T\left(\sum_{i=1}^n a_i v_i\right) := \sum_{i=1}^n a_i f(v_i). \quad (3.8)$$

Esta es en efecto una transformación lineal, pues

$$\begin{aligned} T(ax + y) &= T\left(a \sum_{i=1}^n a_i v_i + \sum_{i=1}^n b_i v_i\right) = T\left(\sum_{i=1}^n (aa_i + b_i) v_i\right) \\ &= \sum_{i=1}^n (aa_i + b_i) f(v_i) = a \sum_{i=1}^n a_i f(v_i) + \sum_{i=1}^n b_i f(v_i) = aT(x) + T(y), \end{aligned}$$

y, en efecto, $T \circ \iota = f$, pues, dado $v_i \in \beta$ cualquiera,

$$(T \circ \iota)(v_i) = T(v_i) = T(0v_1 + \dots + 1v_i + \dots + 0v_n) = f(v_i).$$

\Leftarrow) Mostremos que un subconjunto $\beta = \{v_1, \dots, v_n\}$ de V con tal propiedad es base de V .

- Independencia lineal: Supongamos que $v_1 \in \langle \beta - \{v_1\} \rangle$, o sea, que existen escalares c_i tales que $v_1 = \sum_{i=2}^n c_i v_i$. Sea la función $f : \beta \longrightarrow F$ definida como

$$f(v_1) = 1, \quad f(v_i) = 0, \quad 2 \leq i \leq n.$$

Sea T la única extensión lineal de esta función. Se tiene que

$$1 = T(v_1) = T\left(\sum_{i=2}^n c_i v_i\right) = \sum_{i=2}^n c_i T(v_i) = \sum_{i=2}^n c_i 0_W = 0 \quad \nexists$$

Observe que la ecuación (3.8) en efecto define una función T de V en W , pues, por ser β base de V , la representación de x dada en (3.7) como combinaciones lineales de elementos de β son únicas

- Generación: Supongamos que β no genera a V . Como ya mostramos que β es l.i., podemos extender β a una base de V (c.f. corolario 2.11.15). Sea pues $\gamma \subseteq V$ tal que $\beta \cup \gamma$ es base de V . Si $f : \beta \rightarrow V$ se define como $f(v_i) = v_i$, observe que la función identidad $I_V : V \rightarrow V$ es tal que $I_V \circ \iota = f$, pero también la proyección $T : V = \langle \beta \rangle \oplus \langle \gamma \rangle \rightarrow V$ definida como

$$T(x) = b, \quad x = b + c, b \in \beta, c \in \gamma$$

satisface que $T \circ \iota = f$. Como $T \neq I_V$, esto contradice la unicidad de la extensión lineal que suponemos por hipótesis.

□

Corolario 3.5.2. Sea V, W F -espacios vectoriales, con V finito dimensional, $T : V \rightarrow W, U : V \rightarrow W$ dos transformaciones lineales. Si para una base $\beta = \{v_1, \dots, v_n\}$ de V se tiene que

$$T(v_i) = U(v_i), \quad 1 \leq i \leq n,$$

entonces $T = U$.

Terminemos mostrando que todo F -espacio vectorial de dimensión n es esencialmente \mathbb{R}^n .

Proposición 3.5.3. Sea V un F -espacio vectorial. V es n -dimensional si y sólo si V es isomorfo a \mathbb{R}^n .

Demostración.

\Leftarrow) Inmediata de notar que $\dim(\mathbb{R}^n) = n$ y la Proposición 3.4.3.

\Rightarrow) Sea $\beta = \{v_1, \dots, v_n\}$ base de V . Sea $f : \beta \rightarrow \mathbb{R}^n$ la función que a v_i le asigna \hat{e}_i . Según la propiedad universal de las bases (c.f. Teorema 3.5.1), podemos extender a β linealmente. Sea $T : V \rightarrow \mathbb{R}^n$ tal extensión lineal. Puesto que $T(V)$ contiene a la base canónica de \mathbb{R}^n , $T(V) = \mathbb{R}^n$, o sea, T es suprayectiva. De esto, según el Teorema 3.4.4, se deduce que T es un isomorfismo.

□

3.6 Caracterización de transformaciones lineales de \mathbb{R}^n en \mathbb{R}^m

Es gracias a la propiedad universal de las bases establecida en el Teorema 3.5.1 que vamos a poder caracterizar a las transformaciones lineales entre F -espacios vectoriales finito dimensionales. Como veremos más adelante, esto nos permitirá capturar toda la información de una transformación lineal a partir de una cantidad finita de números, que almacenaremos en un arreglo numérico rectangular - i.e. en una matriz. Antes de abordar la teoría en general, para familiarizarnos con las ideas que vamos a encontrar más adelante, estudiemos el caso de transformaciones lineales de un \mathbb{R}^n a un \mathbb{R}^m , con $m, n \geq 1$ enteros.

Para simplificar la notación, conviene introducir la noción del producto punto euclídeo en \mathbb{R}^n .

Es decir, para demostrar la igualdad entre transformaciones lineales basta comprobar que sus definiciones en una base cualquiera del dominio coinciden.

Definición 3.6.1. Para $\hat{x} = (a_i)_{i=1}^n, \hat{y} = (b_i)_{i=1}^n \in \mathbb{R}^n$, definimos su **producto punto (euclídeo)** como

$$\langle \hat{x}, \hat{y} \rangle = \sum_{i=1}^n a_i b_i.$$

Recordando la definición de producto de matrices, podemos reinterpretar al producto punto de dos vectores de \mathbb{R}^n como producto de matrices;

$$\langle \hat{x}, \hat{y} \rangle = (a_1, a_2, \dots, a_n) \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix} = \left(\sum_{i=1}^n a_i b_i \right).$$

Este es un buen momento para recordar que, por lo general, la multiplicación de matrices no es conmutativa.

En lo que sigue, vamos a identificar a una matriz de 1×1 con su única entrada. Tampoco haremos distinción de los “vectores fila” (a_1, a_2, \dots, a_n) con “vectores columna”

$$\begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix}.$$

Para ser más formales, deberíamos de escribir al vector columna anterior como $(a_1, \dots, a_n)^t$.

I) Sea $T: \mathbb{R} \rightarrow \mathbb{R}^n$ lineal. Consideremos a la base $\{1\}$ de \mathbb{R} . Según el Teorema 3.5.1, T queda completamente determinada por su valor en 1; digamos que $T(1) = (c_1, c_2, \dots, c_n)$. Entonces, para toda $a \in \mathbb{R}$ se tiene que

$$T(a) = T(a \cdot 1) = aT(1) = a(c_1, c_2, \dots, c_n) = \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{pmatrix} (a).$$

II) Sea $T: \mathbb{R}^n \rightarrow \mathbb{R}^n$ lineal. Ahora vamos a considerar a la base canónica $\{\hat{e}_1, \hat{e}_2, \dots, \hat{e}_n\}$ de \mathbb{R}^n . Nuevamente, el Teorema 3.5.1 nos asegura que, con conocer los valores

$$c_i = T(\hat{e}_i) \in \mathbb{R}^n, \quad 1 \leq i \leq n,$$

conocemos los valores de la transformación lineal T en todo punto de \mathbb{R}^n ; en efecto, si $\hat{x} = (a_1, a_2, \dots, a_n) \in \mathbb{R}^n$, entonces

$$T(\hat{x}) = T\left(\sum_{i=1}^n a_i \hat{e}_i\right) = \sum_{i=1}^n a_i T(\hat{e}_i) = \sum_{i=1}^n a_i c_i = \langle \hat{x}, \hat{c} \rangle = (c_1, c_2, \dots, c_n) \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix},$$

donde $\hat{c} = (c_1, c_2, \dots, c_n)$.

II) Ahora sí, consideremos el caso más general de una transformación lineal $T: \mathbb{R}^n \rightarrow \mathbb{R}^m$, con $m, n \geq 1$ enteros cualesquiera. De nuevo tomamos a la base canónica de \mathbb{R}^n compuesta por los vectores \hat{e}_i . Si

$$T(\hat{e}_i) = (c_{1i}, c_{2i}, \dots, c_{mi}), \quad 1 \leq i \leq n,$$

entonces, para cualquier $\hat{x} = (a_1, \dots, a_n) \in \mathbb{R}^n$,

$$\begin{aligned}
 T(\hat{x}) &= T\left(\sum_{i=1}^n a_i e_i\right) = \sum_{i=1}^n a_i T(e_i) \\
 &= \sum_{i=1}^n a_i (c_{1i}, c_{2i}, \dots, c_{mi}) \\
 &= \sum_{i=1}^n (a_i c_{1i}, a_i c_{2i}, \dots, a_i c_{mi}) \\
 &= \left(\sum_{i=1}^n a_i c_{1i}, \sum_{i=1}^n a_i c_{2i}, \dots, \sum_{i=1}^n a_i c_{mi}\right) \\
 &= (\langle \hat{x}, M_1 \rangle, \langle \hat{x}, M_2 \rangle, \dots, \langle \hat{x}, M_m \rangle),
 \end{aligned}$$

donde M_1, M_2, \dots, M_m son las m filas de la matriz

$$\begin{pmatrix} c_{11} & c_{12} & \dots & c_{1n} \\ c_{21} & c_{22} & \dots & c_{2n} \\ \dots & \dots & \cdot & \cdot \\ c_{m1} & c_{m2} & \dots & c_{mn} \end{pmatrix} \in M_{m \times n}(\mathbb{R}),$$

que se forma usando a los n vectores $T(e_i)$ de \mathbb{R}^m como columnas.

3.7 Ejercicios III

Demos un repaso de los conceptos estudiados.

I) Construyamos una base de \mathbb{R}^3 que contenga al vector $v_1 = (4, 5, 9)$.

- Para que v_2 sea tal que $\{v_1, v_2\}$ sea l.i., debe cumplirse que

$$v_2 \in \mathbb{R}^3 - \text{span}(\{v_1\}),$$

es decir, que v_2 no sea múltiplo escalar de v_1 . Pongamos pues a $v_2 = (5, 3, 7)$.

- Para que v_3 sea tal que $\{v_1, v_2, v_3\}$ sea l.i., se deberá tener

$$\begin{aligned} v_3 \in \mathbb{R}^3 - \text{span}(\{v_1, v_2\}) &= \mathbb{R}^3 - \{a(4, 5, 9) + b(5, 3, 7) : a, b \in \mathbb{R}\} \\ &= \mathbb{R}^3 - \{(4a + 5b, 5a + 3b, 9a + 7b) \mid a, b \in \mathbb{R}\}. \end{aligned}$$

Haciendo $a = 1$, $b = -1$, tenemos que el vector $(-1, 2, 2)$ es elemento de $\text{span}(\{v_1, v_2\})$; cambiando una de las tres entradas obtendremos un vector que no está en este espacio (pues el sistema planteado para expresar a este nuevo vector como combinación lineal de v_1 y v_2 no tendría solución, recuerde el Teorema de Cramer). Tomemos pues a $v_3 = (-1, 2, 0)$.

Obtuvimos así al subconjunto linealmente independiente

$$\beta = \{v_1 = (4, 5, 9), v_2 = (5, 3, 7), v_3 = (-1, 2, 0)\}$$

de \mathbb{R}^3 ; como este tiene tres elementos y la dimensión de \mathbb{R}^3 es 3, tenemos que es una base de \mathbb{R}^3 .

II) Encontremos ahora, usando el argumento de la demostración del Teorema fundamental de las bases 3.5.1, una transformación lineal $T: \mathbb{R}^3 \rightarrow \mathbb{R}^2$ tal que

$$T(v_1) = (1, 0), \quad T(v_2) = (2, 5), \quad T(v_3) = (-3, 1). \quad (3.9)$$

Tenemos primero que ver cómo expresar a cualquier vector $(x, y, z) \in \mathbb{R}^3$ como combinación lineal de elementos de la base β . Estamos buscando entonces a los únicos escalares $a, b, c \in \mathbb{R}$ tales que

$$\begin{aligned} (x, y, z) &= a(4, 5, 9) + b(5, 3, 7) + c(-1, 2, 0) \\ &= (4a + 5b - c, 5a + 3b + 2c, 9a + 7b). \end{aligned}$$

Esta igualdad en \mathbb{R}^3 es equivalente al sistema de ecuaciones que resolvemos a continuación:

$$\begin{aligned} \left(\begin{array}{ccc|c} 4 & 5 & -1 & x \\ 5 & 3 & 2 & y \\ 9 & 7 & 0 & z \end{array} \right) &\sim \left(\begin{array}{ccc|c} 1 & 5/4 & -1/4 & x/4 \\ 5 & 3 & 2 & y \\ 9 & 7 & 0 & z \end{array} \right) \sim \left(\begin{array}{ccc|c} 1 & 5/4 & -1/4 & x/4 \\ 0 & -13/4 & 13/4 & y - 5x/4 \\ 0 & -17/4 & 9/4 & z - 9x/4 \end{array} \right) \\ &\sim \left(\begin{array}{ccc|c} 1 & 5/4 & -1/4 & x/4 \\ 0 & 1 & -1 & -4y/13 + 5x/13 \\ 0 & -17 & 9 & 4z - 9x \end{array} \right) \sim \left(\begin{array}{ccc|c} 1 & 5/4 & -1/4 & x/4 \\ 0 & 1 & -1 & -4y/13 + 5x/13 \\ 0 & 0 & 8 & -4z + 68y/13 - 32x/13 \end{array} \right). \end{aligned}$$

Tenemos entonces que

$$c = -\frac{1}{8} \left(4z - \frac{68}{13}y - \frac{32}{13}x \right) = \frac{4}{13}x + \frac{17}{26}y - \frac{1}{2}z, \quad (3.10)$$

Para esta construcción estamos usando el Lema 2.11.1.

En este argumento, x , y y z son las constantes, mientras que las incógnitas son los escalares a , b y c .

Ejercicio: sustituyendo algunos valores para x , y y z compruebe que las fórmulas (3.10), (3.11) y (3.12) son correctas.

$$b = c - \frac{4}{13}y + \frac{5}{13}x = \frac{9}{13}x + \frac{9}{26}y - \frac{1}{2}z, \quad (3.11)$$

y

$$a = \frac{1}{4}(x + c - 5b) = -\frac{7}{13}x - \frac{7}{26}y + \frac{1}{2}z. \quad (3.12)$$

Usando estas tres expresiones, tenemos que la única transformación que extiende la definición (3.9) en la base β es

$$\begin{aligned} T(x, y, z) &= T(av_1 + bv_2 + cv_3) \\ &= aT(v_1) + bT(v_2) + cT(v_3) \\ &= \left(-\frac{7}{13}x - \frac{7}{26}y + \frac{1}{2}z\right)(1, 0) + \left(\frac{9}{13}x + \frac{9}{26}y - \frac{1}{2}z\right)(2, 5) + \left(\frac{4}{13}x + \frac{17}{26}y - \frac{1}{2}z\right)(-3, 1) \\ &= \left(-\frac{1}{13}x - \frac{20}{13}y + z, \frac{49}{13}x + \frac{31}{13}y - 3z\right). \end{aligned}$$

Ejercicio: compruebe que la transformación lineal propuesta cumple las condiciones (3.9).

Ejercicio 17. Sean U , V y W tres F -espacios vectoriales, $a \in F$. Demuestra que

- Si $f, g : V \rightarrow W$ son transformaciones lineales, entonces $f + g$ y af son transformaciones lineales.
- Si $f : V \rightarrow W$ y $g : W \rightarrow U$ son transformaciones lineales, la función composición $g \circ f : V \rightarrow U$ es también lineal.

Ejercicio 18. Argumente por qué existe una única transformación lineal $T : \mathbb{R}^2 \rightarrow \mathbb{R}^3$ tal que

$$T(2, 5) = (1, 0, 1), \quad T(-1, 4) = (2, 5, 7)$$

y de su fórmula explícitamente.

Ejercicio 19. Argumente por qué existe una única transformación lineal $T : P_2(\mathbb{R}) \rightarrow P_2(\mathbb{R})$ tal que

$$T(1) = x^2, \quad T(1+x) = 2x, \quad T(x^2-1) = 2.$$

y de su fórmula explícitamente.

Chapter 4

Representaciones matriciales de transformaciones lineales

The proper treatment of calculus for functions of several variables requires vector ideas; the budding statistician and the coming physicist need them; modern analysis is unthinkable without the notion of linear dependence and all that flows from it. Throughout these courses the infusion of a geometrical point of view is of paramount importance. A vector is geometrical; it is an element of a vector space, defined by suitable axioms - whether the scalars be real numbers or elements of a general field. A vector is not an n -tuple of numbers until a coordinate system has been chosen. Any teacher and any text book which starts with the idea that vectors are n -tuples is committing a crime for which the proper punishment is ridicule. The n -tuple idea is not "easier", it is harder; it is not clearer, it is more misleading. (Mac Lane, 1954)

Now imagine that the query is simply an abstract vector in the space, we would still have to define it with respect to the basis of the space, but it would be up to us, or the user, to refer the objects in the space to different bases depending on their point of view. A change of basis constitutes a change of point of view. (Keith van Rijsbergen)

Desarrollada la teoría de transformaciones lineales en el capítulo anterior, nos proponemos ahora usar la propiedad universal de las bases (c.f. Teorema 3.5.1) para representar transformaciones lineales entre espacios vectoriales finito dimensionales con matrices.

- Mostraremos que, fijando bases en los espacios de salida y de llegada, podremos establecer una biyección (que de hecho será un isomorfismo) que identificará una transformación lineal entre estos espacios con una matriz.
- Puesto que esta asociación preserva estructura, podremos realizar operaciones con matrices (que son objetos sencillos de almacenar y operar en una computadora) e interpretarlas como operaciones entre transformaciones lineales.

Podrá notar que, en lo que sigue, siempre que se defina un concepto para transformaciones lineales, se intentará traducir y relacionar con uno para matrices. Estaremos constantemente estudiando la relación entre espacios de transformaciones lineales con espacios matriciales.

Antes de explicar cómo almacenar la información de una transformación lineal en una matriz - situación que implicará trabajar con un F -espacio vectorial de matrices con coeficientes en F - estudiemos la estructura del espacio de transformaciones lineales.

4.1 El espacio $\mathcal{L}(V, W)$

A partir de ahora vamos a considerar dos F -espacios vectoriales V y W , con

$$\dim(V) = n, \quad \dim(W) = m,$$

y

$$\beta = \{v_1, \dots, v_n\} \subseteq V, \quad \gamma = \{w_1, \dots, w_m\} \subseteq W$$

bases de estos.

Proposición 4.1.1. *Definiendo la suma y multiplicación escalar puntualmente,*

$$\mathcal{L}(V, W) := \{T : V \longrightarrow W \mid T \text{ es lineal}\}$$

es un F -espacio vectorial.

Demostración. **Ejercicio.** Como sugerencia, considere a

$$X = \{T : V \longrightarrow W \mid T \text{ es función}\}.$$

Por herencia de W , X es un F -espacio vectorial. Si demuestra que el subconjunto $\mathcal{L}(V, W)$ de X es no vacío, cerrado bajo sumas y multiplicación escalar (c.f. ejercicio 17), entonces tendrá que es subespacio de X - luego, espacio vectorial en sí mismo. \square

En W no se tiene definida una multiplicación, por lo que no tendría sentido definir el producto de transformaciones lineales puntualmente, sin embargo, siempre podemos componer funciones: recuerda que, si $T : V_1 \longrightarrow V_2$, $U : V_2 \longrightarrow V_3$ son transformaciones lineales, su composición $U \circ T : V_1 \longrightarrow V_3$ se define como

$$(U \circ T)(v) = U(T(v)), \quad v \in V_1.$$

Proposición 4.1.2. *La composición de transformaciones lineales es una transformación lineal.*

Demostración. **Ejercicio.** \square

4.2 El isomorfismo $[\cdot]_\beta$

A partir de ahora, fijada una base β de un F -espacio vectorial V , nos interesarán no sólo sus elementos, sino también el orden en el que estos se enlistan. Por eso, ahora que digamos “base”, nos estaremos refiriendo a una “base ordenada”. Por ejemplo,

$$\beta_1 = \{(1, 0), (0, 1)\} \quad y \quad \beta_2 = \{(0, 1), (1, 0)\}$$

son, como conjuntos, iguales, pero como bases ordenadas de \mathbb{R}^2 son distintas.

Definición 4.2.1. Sea V un F -espacio vectorial n -dimensional, $\beta = \{v_1, \dots, v_n\}$ una base de este. Para $x \in V$, si $x = \sum_{i=1}^n a_i v_i$ es la única representación de x como combinación lineal de elementos de β , entonces, al vector de \mathbb{R}^n

$$[x]_\beta = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} \in \mathbb{R}^n \quad (4.1)$$

se le llamará el **vector coordenado de x respecto a β** .

Observe que la asignación

$$x \mapsto [x]_\beta$$

de V en F^n es una función, pues la representación de cada vector en V como combinación lineal de elementos de β es única (luego, a cada vector le estamos asignando una y sólo una columna ordenada). De hecho, $[\cdot]_\beta$ es la transformación lineal usada en la demostración de la Proposición 3.5.3 para demostrar que, si V tiene dimensión n , entonces es isomorfo a F^n . **Usando entonces a la base β de V , estamos identificando a cada vector de V con una n -tupla via este isomorfismo.** Esto es útil pues, aunque los elementos del espacio original V no sean elementos de \mathbb{R}^n , si V es de dimensión finita, ¡los podemos identificar como tal!

Proposición 4.2.2. Sean V un F -espacio vectorial de dimensión n , $\beta = \{v_1, v_2, \dots, v_n\}$ una base de V , y $[\cdot]_\beta : V \longrightarrow \mathbb{R}^n$ la transformación lineal dada por (4.1).

- $[\cdot]_\beta$ es un isomorfismo. En particular,

$$[x]_\beta = [y]_\beta \Rightarrow x = y.$$

- Para toda $1 \leq i \leq n$, $[v_i]_\beta = e_i$, el i -ésimo vector de la base canónica de \mathbb{R}^n .

Note que la identificación (4.1) de un vector del espacio con un elemento de \mathbb{R}^n depende de la base β fijada.

Ejercicio: Sean V un F -espacio vectorial finito dimensional, β, γ bases de este distintas entre sí. Demuestre que,

- existe $x \in V$ tal que $[x]_\beta \neq [x]_\gamma$, y que
- $[x]_\beta = [y]_\gamma$ no necesariamente implica $x = y$.

Nota 4.2.3. Fijada una base $\beta = \{v_1, \dots, v_n\}$ de V , es importante recordar la definición del isomorfismo $[\cdot]_\beta$;

$$\forall x \in V : [x]_\beta = (a_1, \dots, a_n) \text{ si y sólo si } x = a_1 v_1 + \dots + a_n v_n, \quad (4.2)$$

pues es con ella que recuperamos al vector $x \in V$ a partir de la n -tupla $(a_1, \dots, a_n) \in F^n$.

Ejercicio: resuelve los ejercicios 1 a 5 de la lista 4.10.

En (4.2), nota que la ecuación de la derecha es en V , y la ecuación de la izquierda es en F^n .

4.3 Representación matricial de una transformación lineal

Vamos a considerar ahora a dos F -espacios vectoriales finito dimensionales

$$V, \dim(V) = n, \beta = \{v_1, \dots, v_n\} \subseteq V \text{ base de } V,$$

$$W, \dim(W) = m, \gamma = \{w_1, \dots, w_m\} \subseteq W \text{ base de } W.$$

Recuerde que, si $T: V \rightarrow W$ es una transformación lineal, entonces

- Por el Teorema fundamental de las bases 3.5.1, la transformación T queda completamente determinada por los valores que toma en la base β de V , es decir, a partir de los valores

$$T(v_i), \quad 1 \leq i \leq n$$

podemos recuperar la definición de T en todo el espacio V .

- Por ser γ base de W , cada $T(v_i)$ puede expresarse de forma única como combinación lineal de los vectores w_1, \dots, w_m , es decir, para cada $1 \leq i \leq n$, existen únicos escalares a_{ji} tales que

$$T(v_i) = \sum_{j=1}^m a_{ji} w_j, \quad 1 \leq i \leq n. \quad (4.3)$$

En esta discusión, “ i ” es la variable con la que contamos a los elementos de la base β , y con “ j ” contamos a los elementos de la base γ .

Nota que en la ecuación (4.3) se fija el segundo índice de los escalares, es decir, el índice columna.

Así, una vez fijadas las bases β y γ , **la transformación lineal T queda completamente determinada por los $m \times n$ escalares a_{ji} , con $1 \leq i \leq n$ y $1 \leq j \leq m$.**

Almacenamos esta información en la siguiente matriz:

$$[T]_{\beta}^{\gamma} = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}. \quad (4.4)$$

Note que la i -ésima columna de $[T]_{\beta}^{\gamma}$ es el vector columna $[T(v_i)]_{\gamma}$. Entonces,

- La matriz $[T]_{\beta}^{\gamma}$ tiene n columnas, pues hay una por cada elemento de la base β de V , y
- tiene m filas, pues cada vector columna tiene m entradas, a saber, los coeficientes de las combinaciones lineales de las imágenes $T(v_i)$ respecto a γ .

$$\begin{array}{ccc} V & \xrightarrow{T} & W \\ \beta \subseteq V \text{ base} & & \gamma \subseteq W \text{ base} \\ \dim(V) = n & & \dim(W) = m \\ [T]_{\beta}^{\gamma} \in M_{m \times n}(\mathbb{R}) \end{array}$$

$$[T]_{\beta}^{\gamma} = \begin{pmatrix} T(v_1) & T(v_2) & \dots & T(v_n) \\ a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix} \begin{matrix} w_1 \\ w_2 \\ \vdots \\ w_m \end{matrix}$$

$T: V \rightarrow W$
 $\beta = \{v_j \mid 1 \leq j \leq n\}, \quad \gamma = \{w_k \mid 1 \leq k \leq m\}$

Figure 4.1: Este diagrama puede ayudarte a recordar cómo se construye la matriz $[T]_{\beta}^{\gamma}$

4.4 El isomorfismo Φ_β^γ

Agregar nota de iso de anillos: Golan.

Fijadas bases $\beta = \{v_j \mid 1 \leq j \leq n\} \subseteq V$ y $\gamma = \{w_k \mid 1 \leq k \leq m\} \subseteq W$ de dos F -espacios finito dimensionales, vamos a establecer una biyección entre

- el espacio $\mathcal{L}(V, W)$ de transformaciones lineales de V en W y
- el espacio de matrices $M_{m \times n}(F)$.

Denotemos por Φ_β^γ a la función que a cada transformación lineal de V en W le asigna su representación matricial respecto a β y γ , es decir,

$$\Phi_\beta^\gamma : \mathcal{L}(V, W) \longrightarrow M_{m \times n}(F)$$

$$\Phi_\beta^\gamma(T) = [T]_\beta^\gamma.$$

Del Teorema fundamental de las bases se sigue de inmediato que Φ_β^γ la función es una biyección.

Proposición 4.4.1. *La función Φ_β^γ es una biyección.*

Demostración.

- **Inyectividad de Φ_β^γ :** Sean $T, U : V \longrightarrow W$ lineales tales que

$$[T]_\beta^\gamma = \Phi_\beta^\gamma(T) = \Phi_\beta^\gamma(U) = [U]_\beta^\gamma.$$

El que las matrices $[T]_\beta^\gamma$ y $[U]_\beta^\gamma$ sean iguales significa que

$$\forall v \in \beta : T(v) = U(v),$$

pues la j -ésima columna de la matriz $[T]_\beta^\gamma$ (respectivamente, de $[U]_\beta^\gamma$), da los coeficientes de la combinación lineal en términos de γ igual a T (respectivamente, U) evaluada en el j -ésimo vector de la base β . Así, T y U coinciden en la base β de V , luego, según el Corolario 3.5.2, coinciden.

- **Suprayectividad de Φ_β^γ :** sea $A = (a_{ij}) \in M_{m \times n}(F)$. Por el Teorema fundamental de las bases, existe una única transformación lineal tal que

$$T(v_j) = \sum_{k=1}^m a_{kj} w_k, \quad 1 \leq j \leq n.$$

Dicha T es tal que $[T]_\beta^\gamma = A$.

□

Nota: se pidió a partir de ahora que las bases de los espacios vectoriales considerados fuesen ordenadas pues, para la construcción de las representaciones matriciales $[T]_\beta^\gamma$, la j -ésima columna corresponde al j -ésimo vector de la base β , luego, el orden en el que se enlistan los elementos de β importa (es cierto que un cambio de orden en los elementos de un conjunto no afecta su identidad, pero, el cambiar el orden de las entradas de una matriz definitivamente la altera). El orden de las columnas de una representación matricial A está directamente relacionado con el orden de los vectores en una base ordenada.

En la demostración de la inyectividad de Φ_β^γ , queremos inferir una igualdad de funciones a partir de una igualdad de matrices.

$$T: V \longrightarrow V, \quad \beta = \{v_1, v_2\}, \\ \beta' = \{v_2, v_1\} \text{ bases.}$$

$$T(v_1) = v_2, \quad T(v_2) = 2v_1 + 3v_2$$

$$[T]_{\beta}^{\beta} = \begin{bmatrix} 0 & 2 \\ 1 & 3 \end{bmatrix}, \quad [T]_{\beta'}^{\beta'} = \begin{bmatrix} 3 & 1 \\ 2 & 0 \end{bmatrix}$$

En la siguiente sección desarrollaremos la teoría suficiente para demostrar que, más que una biyección, Φ_{β}^{γ} es un isomorfismo. Esto implica que el espacio $\mathcal{L}(V, W)$ de transformaciones lineales entre dos espacios finito dimensionales es, salvo isomorfismo, un espacio de matrices. En lo que sigue, seguiremos estudiando y explotando la relación entre estos dos espacios, para interpretar las operaciones que realicemos en uno en términos de operaciones en el otro espacio (pues, para las aplicaciones, es muchísimo más sencillo trabajar con matrices a lidiar con transformaciones lineales entre espacios que, a pesar de ser finito dimensionales, pueden ser complicados).

4.5 Operaciones entre transformaciones lineales usando sus representaciones matriciales

Empecemos estableciendo la linealidad de la biyección Φ_{β}^{γ} estudiada en la Sección 4.4.

Proposición 4.5.1. Sean V y W dos F -espacios vectoriales con $\beta = \{v_1, \dots, v_n\} \subseteq V$ y $\gamma = \{w_1, \dots, w_m\} \subseteq W$ bases de estos. Para cualesquiera $T, U: V \longrightarrow W$ y $\lambda \in F$,

$$[T + U]_{\beta}^{\gamma} = [T]_{\beta}^{\gamma} + [U]_{\beta}^{\gamma}. \quad (4.5)$$

$$[\lambda T]_{\beta}^{\gamma} = \lambda [T]_{\beta}^{\gamma}. \quad (4.6)$$

Nota que el lado izquierdo de (4.5) involucra una suma de funciones, mientras que el lado derecho es una suma de matrices.

Demostración. Digamos que

$$[U]_{\gamma}^{\delta} = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}, \quad [T]_{\beta}^{\gamma} = \begin{pmatrix} b_{11} & b_{12} & \dots & b_{1n} \\ b_{21} & b_{22} & \dots & b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ b_{m1} & b_{m2} & \dots & b_{mn} \end{pmatrix}.$$

De estas representaciones matriciales inferimos que, para toda $1 \leq j \leq n$,

$$U(v_j) = \sum_{k=1}^m a_{kj} w_k, \quad T(v_j) = \sum_{k=1}^m b_{kj} w_k,$$

luego,

$$(T + U)(v_j) = \sum_{k=1}^m b_{kj} w_k + \sum_{k=1}^m a_{kj} w_k = \sum_{k=1}^m (b_{kj} + a_{kj}) w_k,$$

y

$$(\lambda T)(v_j) = \lambda \left(\sum_{k=1}^m b_{kj} w_k \right) = \sum_{k=1}^m (\lambda b_{kj}) w_k,$$

por lo tanto,

y

$$[\lambda T]_{\beta}^{\gamma} = \begin{pmatrix} \lambda b_{11} & \lambda b_{12} & \dots & \lambda b_{1n} \\ \lambda b_{21} & \lambda b_{22} & \dots & \lambda b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \lambda b_{m1} & \lambda b_{m2} & \dots & \lambda b_{mn} \end{pmatrix}$$

Si V y W son finito dimensionales, $\mathcal{L}(V, W)$ es un espacio de funciones finito dimensional.

$$\Phi_{\beta}^{\gamma} : \mathcal{L}(V, W) \longrightarrow M_{m \times n}(F)$$

$$\Phi_{\beta}^{\gamma}(T) = [T]_{\beta}^{\gamma}.$$

es un isomorfismo.

Corolario 4.5.3. *Si $\dim(V) = n, \dim(W) = m \in \mathbb{N}$, entonces $\dim(\mathcal{L}(V, W)) = mn$.*

Proposición 4.5.4. Sean V , W y Z tres F -espacios vectoriales, con $\beta = \{v_j : 1 \leq j \leq n\} \subseteq V$, $\gamma = \{w_k : 1 \leq k \leq m\} \subseteq W$ y $\delta = \{z_i : 1 \leq i \leq r\}$ bases de estos. Si $T : V \rightarrow W$ y $U : W \rightarrow Z$ son lineales, entonces

$$[U \circ T]_{\beta}^{\delta} = [U]_{\gamma}^{\delta} \cdot [T]_{\beta}^{\gamma}. \quad (4.7)$$

Nota que el lado izquierdo de (4.7) involucra una composición de funciones, mientras que el lado derecho es un producto de matrices. De hecho, el producto de matrices se definió de tal forma que la ecuación (4.7) fuese cierta.

Demostración. Tenemos que $[U]_\gamma^\delta \in M_{r \times m}(\mathbb{R})$ y que $[T]_\beta^\gamma \in M_{m \times n}(\mathbb{R})$, por lo tanto, el producto de matrices en (4.7) está bien definido.

$[U]_{\gamma}^{\delta}$ tiene $1 \leq i \leq r$ filas y $1 \leq k \leq m$ columnas,

$[T]_{\beta}^{\gamma}$ tiene $1 \leq k \leq m$ filas y $1 \leq j \leq n$ columnas;

entonces,

$[U \circ T]_{\beta}^{\delta}$ tiene $1 \leq i \leq r$ filas y $1 \leq j \leq n$ columnas.

$$\begin{array}{ccccc}
 & & U \circ T & & \\
 & \nearrow & & \searrow & \\
 V & \xrightarrow{T} & W & \xrightarrow{U} & Z \\
 \beta = \{v_j : 1 \leq j \leq n\} & & \gamma = \{w_k : 1 \leq k \leq m\} & & \delta = \{z_i : 1 \leq i \leq r\} \\
 \dim(V) = n & & \dim(W) = m & & \dim(Z) = r
 \end{array}$$

Digamos pues que

$$[U]_{\gamma}^{\delta} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{r1} & a_{r2} & \cdots & a_{rm} \end{pmatrix}, \quad [T]_{\beta}^{\gamma} = \begin{pmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ b_{k1} & b_{k2} & \cdots & b_{kn} \end{pmatrix};$$

estas dos ecuaciones matriciales equivalen a los siguientes dos grupos de ecuaciones:

$$\forall 1 \leq j \leq n : \quad T(v_j) = \sum_{k=1}^m b_{kj} w_k$$

y

$$\forall 1 \leq k \leq m : \quad U(w_k) = \sum_{i=1}^r a_{ik} z_i.$$

Ahora bien, la j -ésima columna de $[U \circ T]_{\beta}^{\delta}$ se consigue evaluando a $U \circ T$ en v_j :

$$\begin{aligned} (U \circ T)(v_j) &= U(T(v_j)) = U\left(\sum_{k=1}^m b_{kj} w_k\right) \\ &= \sum_{k=1}^m b_{kj} U(w_k) = \sum_{k=1}^m b_{kj} \left(\sum_{i=1}^r a_{ik} z_i\right) \\ &= \sum_{i=1}^r \left(\sum_{k=1}^m a_{ik} b_{kj}\right) z_i; \end{aligned}$$

reconociendo a $\sum_{k=1}^m a_{ik} b_{kj}$ como la ij -ésima entrada del producto $[U]_{\gamma}^{\delta} \cdot [T]_{\beta}^{\gamma}$, terminamos.

□

Lema 4.5.5. Sean $A = (a_{kj})_{\substack{1 \leq k \leq m \\ 1 \leq j \leq n}} \in M_{m \times n}(F)$, $b = (b_j)_{1 \leq j \leq n} \in M_{n \times 1}(F)$. Si por $A_{.,j}$ denotamos a la j -ésima columna de A , entonces

$$Ab = b_1 A_{.,1} + \dots + b_n A_{.,n}, \quad (4.8)$$

es decir, Ab es una combinación lineal de las columnas de A , donde los escalares de la combinación lineal son las entradas de b .

Demostración. En efecto, la k -ésima entrada tanto de Ab como de $\sum_{j=1}^n A_{.,j} b_j$ es $\sum_{j=1}^n a_{kj} b_j$ (k se mantiene fijo y se itera el índice de columnas j de A).

□

Teorema 4.5.6. Sean V y W dos F -espacios vectoriales finito dimensionales. Si $\beta \subseteq V$ y $\gamma \subseteq W$ son bases de estos espacios y $T : V \rightarrow W$ es una transformación lineal, entonces,

$$\forall x \in V : \quad [T(x)]_{\gamma} = [T]_{\beta}^{\gamma} [x]_{\beta}.$$

Demostración. En efecto, si $\beta = \{v_j : 1 \leq j \leq n\}$ y

$$[x]_{\beta} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \quad \text{i.e. si } x = \sum_{j=1}^n x_j v_j,$$

usando el Lema 4.5.5, se sigue de inmediato que

$$\begin{aligned} [T]_{\beta}^{\gamma} &= x_1[T(v_1)]_{\gamma} + x_2[T(v_2)]_{\gamma} + \cdots + x_n[T(v_n)]_{\gamma} \\ &= [x_1T(v_1) + \cdots + x_nT(v_n)]_{\gamma} \\ &= [T(x_1v_1 + \cdots + x_nv_n)] = [T(x)]_{\gamma}. \end{aligned}$$

□

Ejercicio: demuestra los siguientes resultados.

Proposición 4.5.7. Si V es finito dimensional y $\beta, \beta' \subseteq V$ son bases de V , entonces

- $[Id]_{\beta}^{\beta'} = I_n$ si y sólo si β y β' son iguales como bases ordenadas.
- La matriz $[Id]_{\beta}^{\beta'}$ es invertible, y

$$([Id]_{\beta}^{\beta'})^{-1} = [Id]_{\beta'}^{\beta}.$$

4.6 Transformaciones lineales de la forma L_A

Sea $A \in M_{m \times n}(F)$ una matriz de m filas y n columnas con coeficientes en F . Definimos a partir de ella una transformación lineal como sigue:

$$L_A : F^n \longrightarrow F^m \quad (4.9)$$

$$L_A(x) = Ax,$$

o sea, L_A es la función “multiplicar a A por la izquierda de vectores de F^n ”.

Usando esta nueva notación, podemos establecer el Teorema 4.5.6 en términos del siguiente diagrama de transformaciones lineales:

$$\begin{array}{ccc} V & \xrightarrow{T} & W \\ \downarrow [\cdot]_{\beta} & & \downarrow [\cdot]_{\gamma} \\ F^n & \xrightarrow{L_{[T]_{\beta}^{\gamma}}} & F^m \end{array}$$

Este diagrama representa lo explicado en el Teorema 4.5.6.

Es decir,

$$[\cdot]_{\gamma} \circ T = L_{[T]_{\beta}^{\gamma}} \circ [\cdot]_{\beta}.$$

Nótese que partimos de espacios vectoriales V y W *arbitrarios* (siendo la única condición impuesta el que ambos sean finito dimensionales), y que gracias a los isomorfismos $[\cdot]_{\beta}$ podemos usar transformaciones del tipo (4.9) en lugar de transformaciones $T : V \longrightarrow W$.

Algunas propiedades fáciles de probar sobre transformaciones de la forma (4.9) se enuncian y se dejan como ejercicio.

Proposición 4.6.1. Si L_A es como se definió en (4.9), entonces

- L_A es lineal
- Si β y γ son las bases canónicas de F^n y F^m respectivamente, entonces $[L_A]_\beta^\gamma = A$.
- $L_A = L_B$ si y sólo si $A = B$.
- $L_{A+B} = L_A + L_B$ y $L_{\lambda A} = \lambda L_A$ para toda $\lambda \in F$.
- Si $T : F^n \longrightarrow F^m$ es una transformación lineal, entonces existe una única matriz C de $m \times n$ tal que $T = L_C$
- Si E es una matriz de $n \times r$, entonces $L_{AE} = L_A \circ L_E$.

Veremos a continuación que la invertibilidad de una transformación lineal T (i.e. el que sea o no isomorfismo) está relacionada a la de cualquiera de sus representaciones matriciales $[T]_\beta^\gamma$.

Teorema 4.6.2. Sea $T : V \longrightarrow W$ una transformación lineal entre dos F -espacios vectoriales finito dimensionales V y W . Si $\beta \subseteq V$ y $\gamma \subseteq W$ son bases cualesquiera de estos, entonces T es un isomorfismo si y sólo si la matriz $[T]_\beta^\gamma$ es invertible.

Recuerda que una matriz A es invertible si y sólo si A es cuadrada y existe B otra matriz tal que $AB = I_n = BA$.

Demostración.

\Rightarrow) Supongamos que T es un isomorfismo, es decir, que es invertible. Entonces, según la Proposición 3.4.3, $\dim(V) = \dim(W)$. Sean $\beta = \{v_j \mid 1 \leq j \leq n\}$ y $\gamma = \{w_j \mid 1 \leq j \leq n\}$ bases de V y W , y sea $T^{-1} : W \longrightarrow V$ la inversa de T - que, recuerde, también es una transformación lineal. Entonces, según la proposición 4.5.4,

$$I_n = [I_V]_\beta^\beta = [T^{-1} \circ T]_\beta^\beta = [T^{-1}]_\gamma^\beta [T]_\beta^\gamma$$

y, análogamente,

$$I_n = [T]_\beta^\gamma [T^{-1}]_\gamma^\beta.$$

Así, $[T]_\beta^\gamma$ es invertible y, de hecho,

La ecuación (4.10) es importante en sí misma.

$$([T]_\beta^\gamma)^{-1} = [T^{-1}]_\gamma^\beta. \quad (4.10)$$

\Leftarrow) Sea $B \in M_{n \times n}(F)$ tal que

$$[T]_\beta^\gamma B = I_n = B [T]_\beta^\gamma.$$

Entonces, si $U : W \longrightarrow V$ es la única transformación lineal tal que $[U]_\gamma^\beta = B$,

$$[T \circ U]_\gamma^\gamma = [T]_\beta^\gamma [U]_\gamma^\beta = I_n.$$

Esto implica que $T \circ U$ es la identidad en W . De forma análoga se prueba que $U \circ T$ es la identidad en V . Así, T es invertible, luego, un isomorfismo.

□

Proposición 4.6.3. Fijadas $\beta \subseteq V$ y $\gamma \subseteq W$ bases de los F -espacios vectoriales finito dimensionales V y W , la función

$$\Phi_\beta^\gamma : \mathcal{L}(V, W) \longrightarrow M_{m \times n}(F)$$

$$\Phi_\beta^\gamma(T) = [T]_\beta^\gamma.$$

es un isomorfismo.

Demostración. Ya vimos en proposiciones anteriores la linealidad de Φ_β^γ .

- Inyectividad: si $T : V \longrightarrow W$ es tal que $[T]_\beta^\gamma$ es la matriz cero, entonces T evaluada en cualquier vector de la base β es cero, luego, T debe ser la transformación lineal cero. Así, $\text{Ker}(\Phi_\beta^\gamma) = \{0\}$.
- Suprayectividad: se estableció en la proposición ??.

□

4.7 Ejemplos

Sea

$$\mathcal{U}_2(\mathbb{R}) = \left\{ \begin{pmatrix} a & b \\ 0 & c \end{pmatrix} : a, b, c \in \mathbb{R} \right\} \leq M_{2 \times 2}$$

el subespacio de matrices triangulares superiores de 2×2 . No es difícil convencerse de que este es un \mathbb{R} -espacio vectorial de dimensión 3, de hecho,

$$\delta = \left\{ \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \right\}$$

es una base para $\mathcal{U}_2(\mathbb{R})$. Sean las transformaciones lineales $S : \mathcal{U}_2(\mathbb{R}) \longrightarrow \mathbb{R}^3$, $T : \mathcal{U}_2(\mathbb{R}) \longrightarrow \mathcal{U}_2(\mathbb{R})$ definidas como

$$T \begin{pmatrix} a & b \\ 0 & c \end{pmatrix} = \begin{pmatrix} a & b \\ 0 & a \end{pmatrix},$$

$$U \begin{pmatrix} a & b \\ 0 & c \end{pmatrix} = (a, b, c)$$

Terminar.

4.8 Cambio de sistema coordenado

Ya ha quedado claro el uso de una base en un F -espacio vectorial V ; estas actúan como un **sistema de coordenadas**, pues, como se estableció en el Teorema 2.10.3 (resultado usado una y otra vez en el desarrollo de la teoría anterior), si β es base de V , dado $x \in V$ cualquiera, existen

- **únicos** v_1, \dots, v_n de elementos de β , y
- **únicos** escalares a_1, \dots, a_n escalares tales que

$$x = \sum_{j=1}^n a_j x_j.$$

Recuerde que es por eso que, en el caso en el que V es finito dimensional (i.e. β es finita), usamos esto para construir el isomorfismo $[\cdot]_\beta$ entre V y F^n (justo asociándole a x su única colección de escalares a_1, \dots, a_n). Note que esta representación **depende de la base** β , es decir, si $\gamma \subseteq V$ es una base de V distinta a β , entonces los isomorfismos $[\cdot]_\beta$ y $[\cdot]_\gamma$ son distintos entre sí. Un caso particular de esta situación es el realizar un cambio de coordenadas en el plano (el espacio vectorial en esta situación es, por supuesto, \mathbb{R}^2).

En este caso, “fijar el plano/sistema coordenado” y “fijar la base” significan lo mismo. **Discusión del pizarrón. Fijar sistema de referencia, y usar nuevas bases en ese sistema.**

Ejemplo 4.8.1. Mostremos cómo en ocasiones es conveniente cambiar de sistema de coordenadas. En \mathbb{R}^2 , considere a la base canónica $\beta = \{e_1 = (1, 0), e_2 = (0, 1)\}$. Todo punto del espacio (x, y) se expresa de forma única como combinación lineal de elementos de β ;

$$(x, y) = x(1, 0) + y(0, 1).$$

Esta representación es muy sencilla - por eso resulta ventajoso en muchas ocasiones trabajar con la base canónica de \mathbb{R}^2 . Considere al lugar geométrico $\mathcal{E} \subseteq \mathbb{R}^2$ de puntos del plano que consta de los puntos que satisfacen la siguiente relación:

$$w = (x, y) \in \mathcal{E} \quad \text{si y sólo si} \quad 2x^2 - 4xy + 5y^2 = 1. \quad (4.11)$$

Puesto que en la ecuación para definir \mathcal{E} están involucrados cuadrados de las variables, parece que este lugar geométrico es una cónica, sin embargo, el término mixto $-4xy$ dificulta su identificación. Cambiemos de sistema coordenado; considérese a la base

$$\beta' = \left\{ u = \left(\frac{\sqrt{5}}{5}, -\frac{2\sqrt{5}}{5} \right), v = \left(\frac{2\sqrt{5}}{5}, \frac{\sqrt{5}}{5} \right) \right\}.$$

Si $w = (x, y) \in \mathbb{R}^2$, entonces

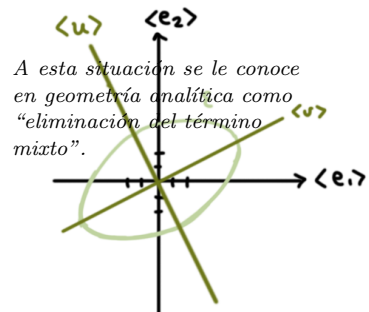
$$\begin{aligned} w = (x, y) &= xe_1 + ye_2 \\ &= x'u + y'v \\ &= x' \left(\frac{\sqrt{5}}{5}, -\frac{2\sqrt{5}}{5} \right) + y' \left(\frac{2\sqrt{5}}{5}, \frac{\sqrt{5}}{5} \right) \\ &= x' \left(\frac{\sqrt{5}}{5} e_1 - \frac{2\sqrt{5}}{5} e_2 \right) + y' \left(\frac{2\sqrt{5}}{5} e_1 + \frac{\sqrt{5}}{5} e_2 \right) \\ &= \left(\frac{\sqrt{5}}{5} x' + \frac{2\sqrt{5}}{5} y' \right) e_1 + \left(-\frac{2\sqrt{5}}{5} x' + \frac{\sqrt{5}}{5} y' \right) e_2. \end{aligned}$$

De la igualdad de los coeficientes al representar vectores respecto a una base se deduce que

$$x = \frac{\sqrt{5}}{5} x' + \frac{2\sqrt{5}}{5} y', \quad y = -\frac{2\sqrt{5}}{5} x' + \frac{\sqrt{5}}{5} y'. \quad (4.12)$$

Si usamos ahora el sistema coordenado β' , debemos sustituir los cambios (4.12) en la condición de definición (4.11). Haciendo esto, llegamos a que

$$w = x'u + y'v \in \mathcal{E} \quad \text{si y sólo si} \quad 6(x')^2 + (y')^2 = 1. \quad (4.13)$$



Ahora sí reconocemos al lugar geométrico como una elipse, pero **no** respecto al sistema coordenado β , sino respecto a β' .

Al dibujar el nuevo sistema coordenado en el plano (i.e. a las rectas $\{\{u\}\}$ y $\{\{v\}\}$, que son subespacios uno dimensionales de \mathbb{R}^2), notamos que este se obtiene rotando el sistema original β .

Observa que la relación (4.12) (que nos permitió pasar de una base a otra) puede representarse en forma matricial como

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \frac{\sqrt{5}}{5} & \frac{2\sqrt{5}}{5} \\ -\frac{2\sqrt{5}}{5} & \frac{\sqrt{5}}{5} \end{pmatrix} \begin{pmatrix} x' \\ y' \end{pmatrix};$$

nota que

$$[w]_{\beta} = \begin{pmatrix} x \\ y \end{pmatrix}, \quad [Id]_{\beta'}^{\beta} = \begin{pmatrix} \frac{\sqrt{5}}{5} & \frac{2\sqrt{5}}{5} \\ -\frac{2\sqrt{5}}{5} & \frac{\sqrt{5}}{5} \end{pmatrix}, \quad [w]_{\beta'} = \begin{pmatrix} x' \\ y' \end{pmatrix}.$$

Usando el que la transformación lineal identidad es el neutro de la composición y que multiplicación de matrices es la operación correspondiente a composición de funciones, se establece fácilmente cómo hacer cambios de base.

Teorema 4.8.2. Sean β, β' dos bases de un F -espacio vectorial finito dimensional V . Sea $Q = [Id_V]_{\beta'}^{\beta}$.

- La matriz Q es invertible.
- Para toda $v \in V$,

$$[v]_{\beta} = Q[v]_{\beta'}. \quad (4.14)$$

Demostración. Claro que Q es invertible, pues la transformación identidad Id_V es un isomorfismo (c.f. [cita](#)). Además, según el Teorema 4.5.6,

$$[v]_{\beta} = [Id_V(v)]_{\beta} = [Id]_{\beta'}^{\beta}[v]_{\beta'}.$$

□

La ecuación (4.14) nos explica cómo obtener las coordenadas de un vector v en términos de β cuando se tienen sus coordenadas en términos de β' .

Definición 4.8.3. Si β y β' son dos bases de un F -espacio vectorial finito dimensional V , a la matriz $[Id]_{\beta'}^{\beta}$ se le conoce como **matriz de cambio de base** de β' a β .

Poner ejemplo numérico.

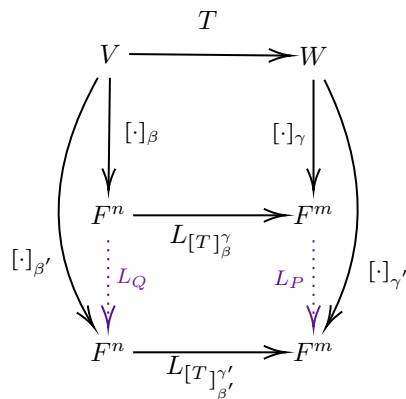
Sean V y W dos F -espacios vectoriales, y considere dos pares de bases

$$\beta \subseteq V, \gamma \subseteq W; \quad \beta' \subseteq V, \gamma' \subseteq W$$

de estos. Dada una transformación lineal $T : V \rightarrow W$, según el Teorema 4.5.6, tenemos los siguientes dos cuadrados conmutativos;

$$\begin{array}{ccc} V & \xrightarrow{T} & W \\ \downarrow [\cdot]_{\beta} & & \downarrow [\cdot]_{\gamma} \\ F^n & \xrightarrow{L[T]_{\beta}^{\gamma}} & F^m \end{array} \quad \begin{array}{ccc} V & \xrightarrow{T} & W \\ \downarrow [\cdot]_{\beta'} & & \downarrow [\cdot]_{\gamma'} \\ F^n & \xrightarrow{L[T]_{\beta'}^{\gamma'}} & F^m \end{array}$$

Ambos diagramas hacen referencia a T (su objetivo de hecho es decir cómo evaluar a T en vectores de V en términos de multiplicaciones de matrices). ¿Qué relación hay entre estos? En otras palabras, ¿Cómo conociendo la representación matricial $[T]_{\beta}^{\gamma}$ encontramos a $[T]_{\beta'}^{\gamma'}$?



Podemos usar matrices de cambio de base para pasar fácilmente de una representación matricial a la otra. En efecto, si

$$P = [Id_W]_{\gamma'}^{\gamma}, \quad Q = [Id_V]_{\beta}^{\beta'},$$

entonces, para toda $v \in V$,

$$P[T]_{\beta}^{\gamma}[v]_{\beta} = ([Id_W]_{\gamma'}^{\gamma}[T]_{\beta}^{\gamma})[v]_{\beta} = [Id_W \circ T]_{\beta}^{\gamma'}[v]_{\beta} = [T]_{\beta}^{\gamma'}[v]_{\beta} = [T(v)]_{\gamma},$$

y también

$$[T]_{\beta'}^{\gamma'}Q[v]_{\beta} = ([T]_{\beta'}^{\gamma'}[Id_V]_{\beta}^{\beta'})[v]_{\beta} = [T \circ Id_V]_{\beta}^{\gamma'}[v]_{\beta} = [T]_{\beta}^{\gamma'}[v]_{\beta} = [T(v)]_{\gamma}.$$

De esto se deduce que

$$P[T]_{\beta}^{\gamma} = [T]_{\beta'}^{\gamma'}Q. \quad (4.15)$$

Puesto que P es invertible, podemos despejar a $[T]_{\beta}^{\gamma}$ de (4.15). Hemos demostrado así el siguiente

Teorema 4.8.4. Sean V y W dos F -espacios vectoriales finito dimensionales, $T : V \rightarrow W$ una transformación lineal entre ellos. Si $\beta, \beta' \subseteq V$ y $\gamma, \gamma' \subseteq W$ son bases para estos espacios,

- $P = [Id_W]_{\gamma'}^{\gamma}$ es la matriz que transforma las coordenadas de γ en γ' , y
- $Q = [Id_V]_{\beta}^{\beta'}$ es la matriz que transforma las coordenadas de β en β' ,

entonces,

$$[T]_{\beta}^{\gamma} = P^{-1}[T]_{\beta'}^{\gamma'}Q. \quad (4.16)$$

Observación 4.8.5. Si V es un F -espacio vectorial finito dimensional y $\beta, \beta' \subseteq V$ son bases de V , entonces

$$([Id_V]_{\beta}^{\beta'})^{-1} = [Id_V]_{\beta'}^{\beta}. \quad (4.17)$$

Ejemplo numérico con observaciones.

Vale la pena escribir el Teorema 4.8.4 para el caso particular en el que el espacio de dominio coincide con el de codominio.

En efecto, toma a $[v]_{\beta}$ como un vector de la base canónica de F^n para probar que las correspondientes columnas de las matrices en (4.15) en efecto coinciden entre sí.

Corolario 4.8.6. Si V es un F -espacio vectorial finito dimensional y $\beta, \beta' \subseteq V$ son bases de V , entonces

$$[T]_{\beta}^{\beta} = Q^{-1} [T]_{\beta'}^{\beta'} Q, \quad (4.18)$$

donde $Q = [Id_V]_{\beta}^{\beta'}$.

4.8.1 Matrices similares

Sigue discusión y salta a p. 239 Friedb.

Definición 4.8.7. Sean $A, B \in M_{n \times n}(F)$. Decimos que las matrices A y B son **similares** si existe $Q \in M_{n \times n}(F)$ invertible tal que

$$A = Q^{-1} B Q.$$

En términos de esta definición, el Corolario 4.8.6 se resume como sigue:

Las representaciones matriciales de una transformación lineal $T : V \rightarrow V$ respecto a una misma base son todas similares entre sí.

No es difícil demostrar que la relación “ser similar con” es de equivalencia:

Ejercicio: Demuestra que la relación en el conjunto de las matrices cuadradas $n \times n$ dada por

$$\forall A, B \in M_{n \times n}(F) : A \sim B \Leftrightarrow \exists Q \in M_{n \times n}(F) \text{ invertible tal que } A = Q^{-1} B Q \quad (4.19)$$

es de equivalencia, es decir, que

- es reflexiva: $\forall A \in M_{n \times n}(F) : A \sim A$,
- simétrica: $\forall A, B \in M_{n \times n}(F) : A \sim B \Rightarrow B \sim A$, y
- transitiva: $\forall A, B, C \in M_{n \times n}(F) : A \sim B \wedge B \sim C \Rightarrow A \sim C$.

Como toda relación de equivalencia en el conjunto $M_{n \times n}(F)$, esta induce una partición en él.

Notación 4.8.8. Si $A \in M_{n \times n}(F)$, por \overline{A} denotaremos a la clase de equivalencia de A bajo la relación “ser similar a”;

$$\overline{A} = \{B \in M_{n \times n}(F) : A \sim B\}.$$

Sea V un F -espacio vectorial con $\dim(V) = n$ (por ejemplo, $V = F^n$). Si $T : V \rightarrow V$ es una transformación lineal y $\beta \subseteq V$ es una base para V , consideremos a la matriz $A = [T]_{\beta}^{\beta}$. Según el Corolario 4.8.6, si $\beta' \subseteq V$ es otra base cualquiera de V ,

$$[T]_{\beta'}^{\beta'} = Q^{-1} [T]_{\beta}^{\beta} Q = Q^{-1} A Q,$$

donde $Q = [Id_V]_{\beta}^{\beta'}$, es en efecto invertible (pues $Id_V : V \rightarrow V$ lo es). Esto significa que

$$\{[T]_{\beta'}^{\beta'} : \beta' \subseteq V \text{ es base}\} \subseteq \overline{A}.$$

¿Se tiene también la otra inclusión? Damos una respuesta afirmativa en el siguiente

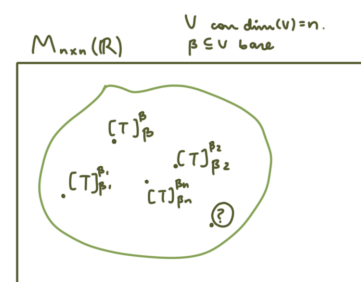


Figure 4.2

Teorema 4.8.9. Sea V un F -espacio vectorial n -dimensional, $T : V \rightarrow V$ lineal. Si

$$\beta = \{v_1, \dots, v_n\} \subseteq V$$

es base y $A = [T]_\beta^\beta$, entonces,

$$\forall B \in M_{n \times n}(F) : A \sim B \text{ si y sólo si existe } \beta' \subseteq V \text{ base tal que } B = [T]_{\beta'}^{\beta'}.$$

Es decir, la clase de equivalencia de $A = [T]_\beta^\beta$ bajo la relación de equivalencia (4.19) es

$$\{[T]_{\beta'}^{\beta'} \mid \beta' \subseteq V \text{ es base}\}.$$

Demostración. Sea $B \in M_{n \times n}(F)$ similar a A ; digamos pues que $Q \in M_{n \times n}(F)$ es una matriz invertible tal que

$$B = Q^{-1}AQ = Q^{-1}[T]_\beta^\beta Q. \quad (4.20)$$

Usando la matriz Q vamos a construir una base para V ; definamos a los vectores

$$v'_j := q_{1j}v_1 + q_{2j}v_2 + \dots + q_{nj}v_n, \quad 1 \leq j \leq n. \quad (4.21)$$

O sea, tomamos a las columnas de Q como listas de coeficientes. Sea $\beta' = \{v'_1, \dots, v'_n\}$. Si $a_1, \dots, a_n \in F$ tales que

$$\begin{aligned} 0 &= a_1v'_1 + \dots + a_nv'_n \\ &= a_1(q_{11}v_1 + \dots + p_{n1}v_n) + \dots + a_n(q_{1n}v_1 + \dots + p_{nn}v_n) \\ &= (a_1p_{11} + \dots + a_np_{1n})v_1 + \dots + (a_1p_{n1} + \dots + a_np_{nn})v_n, \end{aligned}$$

luego, como β es l.i., tenemos que los coeficientes de la combinación lineal anterior son todos cero, hecho que puede reescribirse como la siguiente ecuación matricial;

$$Qa = 0_{F^n}, \quad a := \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}.$$

Por ser Q invertible, de esta última igualdad se deduce que a es el vector cero de F^n . De esto se concluye que β' es l.i., luego, como tiene cardinalidad n (el que sea l.i. implica que los vectores v'_j son distintos entre sí), es base de V . La definición (4.21) de los vectores v'_i se reinterpreta ahora como

$$Q = [Id]_{\beta'}^\beta. \quad (4.22)$$

Sustituyendo a (4.22) en (4.20), concluimos que

$$B = [Id]_{\beta'}^{\beta'} [T]_\beta^\beta [Id]_{\beta'}^\beta = [T]_{\beta'}^{\beta'}.$$

□

4.9 Cálculo del rango de transformaciones usando representaciones matriciales

Definición 4.9.1. Sea $A \in M_{m \times n}(F)$ una matriz. Definimos su **rango** como la máxima cantidad de columnas linealmente independientes que tenga.

Recuerda que el rango de una transformación lineal $T : V \longrightarrow W$ se definió como la dimensión de su imagen.

Proposición 4.9.2. *El rango columna es igual al rango fila. Reformular y demostrar.*

Proposición 4.9.3. *Si $A \in M_{m \times n}(F)$, entonces su rango coincide con el rango de su transformación lineal asociada, o sea,*

$$\text{Rango}(A) = \dim(L_A(F^n)). \quad (4.23)$$

Demostración. Si e_i es el i -ésimo vector de la base canónica de F^n , note que $L_A(e_i) = Ae_i$ es la i -ésima columna de A . Además, como $\{e_1, \dots, e_n\}$ es base de F^n , $\{L_A(e_1), \dots, L_A(e_n)\}$ genera a $L_A(F^n)$ (c.f. Corolario 3.2.4). Es decir, los vectores columna generan al espacio imagen de L_A ; extraer de este generador una base es lo mismo que encontrar la mayor cantidad de columnas linealmente independientes de A .

□

Proposición 4.9.4. *Si V es finito dimensional y $U : V \longrightarrow W$ es un isomorfismo, entonces, para todo $X \leq V$, $\dim(X) = \dim(U(X))$.*

Demostración. Claro que la restricción $U|_X : X \longrightarrow U(X)$ es, por herencia de U , lineal e inyectiva, y por como hemos definido el codominio, también es suprayectiva, luego, es un isomorfismo, por lo tanto, según la proposición 3.4.3, $\dim(X) = \dim(U(X))$.

□

Proposición 4.9.5. *Sea V finito dimensional, $T : V \longrightarrow W$ lineal. Si $U : Z \longrightarrow V$ es un isomorfismo, entonces*

$$\dim(\text{Ker}(T \circ U)) = \dim(\text{Ker}(T)), \quad \dim(T \circ U(Z)) = \dim(T(V)),$$

es decir, componer por la derecha con un isomorfismo no altera la dimensión del Kernel o de la imagen de una transformación lineal.

Demostración. Como U es un isomorfismo, $\dim(Z) = \dim(V)$. Obsérvese que $x \in \text{Ker}(T \circ U)$ si y sólo si $T(U(x)) = 0$, o sea, si y sólo si $U(x) \in \text{Ker}(T)$, entonces,

$$\text{Ker}(T \circ U) = U^{-1}(\text{Ker}(T)).$$

Nota que, como U es invertible, aquí $U^{-1}(C)$ es la imagen bajo C de la transformación inversa U^{-1} .

Además, U^{-1} es un isomorfismo, luego, según la Proposición 4.9.4,

$$\dim(\text{Ker}(T)) = \dim(U^{-1}(\text{Ker}(T))) = \dim(\text{Ker}(T \circ U)).$$

Usando el teorema de la dimensión, también tenemos que

$$\dim((T \circ U)(Z)) = \dim(Z) - \dim(\text{Ker}(T \circ U)) = \dim(V) - \dim(\text{Ker}(T)) = \dim(T(V)).$$

□

Ejercicio: Ahora demuestra que componer por la izquierda con un isomorfismo no altera la dimensión del kernel o la imagen de la transformación lineal.

Proposición 4.9.6. *Si $T : V \longrightarrow W$ lineal, $\beta \subseteq V$, $\gamma \subseteq W$ bases, entonces,*

$$\text{Rango}(T) = \text{Rango}([T]_{\beta}^{\gamma}). \quad (4.24)$$

Demostración. Si $A = [T]_{\beta}^{\gamma}$, según la Proposición 4.9.3, basta ver que $\text{Rango}(T) = \text{Rango}(L_A)$. Puesto que $L_A \circ [\cdot]_{\beta} = [\cdot]_{\gamma} \circ T$, se tiene que

$$T = [\cdot]_{\gamma}^{-1} \circ L_A \circ [\cdot]_{\beta}.$$

Puesto que $[\cdot]_{\gamma}^{-1}$ y $[\cdot]_{\beta}$ son ambos isomorfismos, según la proposición anterior, terminamos. \square

Así, el calcular el rango de una transformación lineal es lo mismo que calcular el de *cualquiera* de sus representaciones matriciales. Por supuesto que algunas representaciones serán más útiles que otras para este propósito. **Nota que el cálculo del rango de una matriz se simplifica mucho si se crean ceros en sus entradas usando operaciones elementales.** Tales ceros pueden formarse al efectuar operaciones elementales. **Aquí recordamos lo que son operaciones elementales en filas y columnas, y las identificamos como hacer multiplicaciones (dependiendo de fila o col., por derecha o por izquierda) con matrices elementales.**

Definición 4.9.7. Fijado $n \in \mathbb{N}$, una **matriz elemental** es una matriz que se obtiene al realizar una operación elemental en filas o columnas de la matriz identidad I_n .

Si $A \in M_{m \times n}(F)$, entonces

- para realizar una operación elemental en las columnas de A , hay que multiplicar por la derecha de A por la matriz elemental obtenida al hacer esa operación elemental en la identidad I_n .
- Dualmente, para realizar una operación elemental en las filas de A , hay que multiplicar por la izquierda de A por la matriz elemental obtenida al hacer esa operación elemental en la identidad I_m .

$$\begin{pmatrix} a & b & c \\ d & e & f \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} = \begin{pmatrix} a & c & b \\ d & f & e \end{pmatrix}.$$

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \cdot \begin{pmatrix} a & b & c \\ d & e & f \end{pmatrix} = \begin{pmatrix} d & e & f \\ a & b & c \end{pmatrix}.$$

Observa que, realizar operaciones elementales en una matriz $A \in M_{m \times n}$ no afecta entonces su rango, pues, si E es la correspondiente matriz elemental, como $L_E : F^m \rightarrow F^m$ es invertible,

$$\text{Rango}(A) = \text{Rango}(L_A) = \text{Rango}(L_E \circ L_A) = \text{Rango}(L_{E \cdot A}) = \text{Rango}(E \cdot A).$$

Ejemplo 4.9.8. **Ejercicio:** Modificando a la matriz A con operaciones elementales, calcula su rango, con

$$A = \begin{pmatrix} 1 & 2 & 3 & 1 \\ 2 & 1 & 1 & 1 \\ 1 & -1 & 1 & 0 \end{pmatrix}.$$

En efecto, si β es la base canónica de F^m , $[L_E]_{\beta}^{\beta} = E$ es una matriz invertible, luego, la transformación lineal L_E es invertible.

4.10 Ejercicios IV

Cálculo del isomorfismo $[\cdot]_\beta : V \longrightarrow \mathbb{R}^n$

Si V es un F -espacio vectorial n -dimensional y $\beta = \{v_j : 1 \leq j \leq n\}$ es una base de V , entonces, se definió el isomorfismo

$$[\cdot]_\beta : V \longrightarrow F^n$$

como

$$T\left(\sum_{j=1}^n a_j v_j\right) = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} \quad \text{para toda } x = \sum_{j=1}^n a_j v_j \in V$$

Es decir, calcular $[\cdot]_\beta$ es lo mismo que expresar a un vector arbitrario x de V como combinación lineal de elementos de β

Si $T : V \longrightarrow V$ es lineal y en V se usa sólo una base β , a $[T]_\beta^\beta$ a veces se le denota como $[T]_\beta$.

Cálculo de representación matricial $[T]_\beta^\gamma : V \longrightarrow W$

Si V y W son F -espacios vectoriales finito dimensionales y

$$\beta = \{v_j : 1 \leq j \leq n\} \subseteq V,$$

$$\gamma = \{w_k : 1 \leq k \leq m\} \subseteq W$$

son bases de estos, entonces, $[T]_\beta^\gamma$ es una matriz de m filas con n columnas, donde su j -ésima columna es $[T(v_j)]_\gamma$.

Entonces, $[T]_\beta^\gamma$ tiene

- n columnas, pues hay una por cada elemento de la base β , y
- m filas, ya que cada $T(v_j)$ necesita m escalares para representarse como combinación lineal de elementos de γ .

Sean

$$\beta = \{(1, 0, 0), (0, 1, 0), (0, 0, 1)\}, \quad \beta' = \{(0, 1, 0), (1, 0, 0), (0, 0, 1)\}, \quad (4.25)$$

$$\gamma = \{(0, 3, 5), (1, 2, 0), (3, 4, 5)\}, \quad \gamma' = \{(3, 4, 5), (1, 2, 0), (0, 3, 5)\}, \quad (4.26)$$

$$\delta = \{(1, 2, 0), (0, 1, 1), (1, 0, 3)\}. \quad (4.27)$$

Nota que β y β' son iguales como conjuntos, pero no como bases ordenadas.

1. Demuestre que los subconjuntos de \mathbb{R}^3 dados en (4.25), (4.26) y (4.27) son bases de \mathbb{R}^3 .
2. Dado $(x, y, z) \in \mathbb{R}^3$ un elemento cualquiera de \mathbb{R}^3 , expréselo como combinación lineal de elementos de las bases del inciso anterior.
3. Dé las fórmulas de los isomorfismos $[\cdot]_\beta$, $[\cdot]_\gamma$, $[\cdot]_\delta$.
4. Encuentre $x \in \mathbb{R}^3$ tal que

- $[x]_\beta = (1, 2, 3)$,
- $[x]_\gamma = (1, 2, 3)$,
- $[x]_{\gamma'} = (1, 2, 3)$,
- $[x]_\delta = (1, 2, 3)$.

5. Sean $\beta = \{1, x, x^2\}$, $\gamma = \{-3, 2x - 1, x^2 + x + 1\} \subseteq P_2(\mathbb{R})$.

- Demuestra que γ es base de $P_2(\mathbb{R})$.
- Calcula explícitamente a los isomorfismos $[\cdot]_\beta, [\cdot]_\gamma : P_2(\mathbb{R}) \longrightarrow \mathbb{R}^3$.

6. Sean $T, U : \mathbb{R}^3 \longrightarrow \mathbb{R}^3$ las funciones

$$T(x, y, z) = (2x, 5y - z, x + y + z), \quad U(x, y, z) = (2x + 3y, -x, -z + y).$$

Calcule las matrices $[T]_\beta^\gamma, [T]_\gamma^\beta, [T]_\gamma^{\delta}, [T]_\delta^\gamma, [U]_\beta^\delta$ y $[U]_\delta^\beta$.

7. Compare a $[T]_\beta^\gamma$ con $[T]_{\beta'}^\gamma$ y a $[U]_\beta^\gamma$ con $[T]_\beta^{\gamma'}$. ¿Qué relación tienen una con otra? ¿Cómo el cambio de orden en una base afecta a las matrices?

8. Sea la matriz

$$A = \begin{pmatrix} 1 & 7 & 0 \\ 0 & 9 & 2 \\ 5 & 1 & 2 \end{pmatrix} \in M_{3 \times 3}(\mathbb{R}).$$

Explique por qué existe^(*) una única transformación lineal $L : \mathbb{R}^3 \longrightarrow \mathbb{R}^3$ tal que $[L]_\beta^\gamma = A$. Encuéntrala.

(*) Pista: usa el Teorema fundamental de las bases 3.5.1

De igual manera, encuentre a la única transformación lineal $S : \mathbb{R}^3 \longrightarrow \mathbb{R}^3$ lineal tal que $[S]_\gamma^\delta = A$. ¿Son S y L iguales?

Sea

$$\phi = \{(1, 2), (1, 1)\}. \quad (4.28)$$

- Demuestra que ϕ es base de \mathbb{R}^2 .
- Da explícitamente al isomorfismo $[\cdot]_\phi$.
- Sea $K : \mathbb{R}^3 \longrightarrow \mathbb{R}^2$ definida como

$$K(x, y, z) = (2x + z, -y - z).$$

- Da explícitamente la fórmula de las composiciones $K \circ U$ y $K \circ T$.
- Usa las fórmulas de estas composiciones para calcular las matrices $[K \circ U]_\delta^\phi$ y $[K \circ T]_\beta^\phi$.
- Calcula las matrices del inciso anterior usando representaciones matriciales apropiadas de T , U y K y multiplicándolas entre sí.

Además, revisa los ejercicios del capítulo 2.3 del Friedberg, pero en particular resuelve 3 y 4 de la página 93.

$$\begin{array}{ccccc} \mathbb{R}^3 & \xrightarrow{T} & \mathbb{R}^3 & \xrightarrow{K} & \mathbb{R}^2 \\ & \searrow & & \nearrow & \\ & & K \circ T & & \end{array}$$

4.11 Simulacro de examen, segundo parcial

Álgebra Lineal I, Otoño 2024 § Segundo examen parcial § Lic. Amélie Bernès

Es obligatorio que contestes los siguientes dos problemas; si tu respuesta es correcta, ganas los puntos correspondientes, pero si no los contestas o tu respuesta es incorrecta, pierdes los puntos correspondientes.

Problema 4.11.1. (+0.5/-0.25) *Enuncia la definición de transformación lineal.*

Problema 4.11.2. (+0.5/-0.25) *Si $T : V \rightarrow W$ es una transformación lineal y $X \subseteq V$, $Y \subseteq W$, escribe las definiciones de los conjuntos $T(X)$ y $T^{-1}(Y)$.*

Contesta los problemas que quieras de la siguiente lista. En los ejercicios numéricos, es recomendable que escribas limpiamente tu procedimiento, para evaluar tu procedimiento en caso de que la respuesta numérica final sea incorrecta.

Problema 4.11.3. (1 p) *Enuncia el teorema fundamental de las bases.*

Problema 4.11.4. (3 p.) *Enuncia y demuestra el teorema de la dimensión.*

Problema 4.11.5. (0.5 p) *Sean V y W dos F -espacios vectoriales finito dimensionales, $T : V \rightarrow W$ una transformación lineal entre ellos. Si $\dim(V) < \dim(W)$, explica por qué T no puede ser suprayectiva.*

Problema 4.11.6. (1 p) *Sea $T : V \rightarrow W$ lineal. Demuestra que T es inyectiva si y sólo si $\text{Ker}(T) = \{0_V\}$.*

Problema 4.11.7. (1.5 p) *Sea $T : V \rightarrow W$ lineal. Demuestra que, si $\beta = \{v_1, \dots, v_n\}$ es un subconjunto de V tal que $T(\beta) \subseteq W$ es l.i., entonces β también es l.i..*

Problema 4.11.8. (2 p) *Sean $\beta = \{(1, 0, 0), (0, 1, 0), (0, 0, 1)\} \subseteq \mathbb{R}^3$, $\gamma = \{(3, 0), (1, -1)\} \subseteq \mathbb{R}^2$ bases respectivas de \mathbb{R}^3 y \mathbb{R}^2 . Da la fórmula de la única transformación lineal $T : \mathbb{R}^3 \rightarrow \mathbb{R}^2$ tal que*

$$[T]_{\beta}^{\gamma} = \begin{pmatrix} 1 & 7 & 0 \\ 2 & 3 & 4 \end{pmatrix}.$$

Problema 4.11.9. *Sean*

$$\beta = \left\{ E_1 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, E_2 = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, E_3 = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, E_4 = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \right\} \subseteq M_{2 \times 2}(\mathbb{R}),$$

$$\beta' = \left\{ A = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}, B = \begin{pmatrix} 0 & 2 \\ 3 & 0 \end{pmatrix}, C = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}, D = \begin{pmatrix} 0 & 3 \\ 2 & 0 \end{pmatrix} \right\} \subseteq M_{2 \times 2}(\mathbb{R}),$$

$$\gamma = \{1\} \subseteq \mathbb{R}$$

bases de $M_{2 \times 2}(\mathbb{R})$ y \mathbb{R} , respectivamente. Defínase $T : M_{2 \times 2}(\mathbb{R}) \rightarrow \mathbb{R}$ como $T\left(\begin{pmatrix} a & b \\ c & d \end{pmatrix}\right) = a + d$.

- (1 p) *Demuestra que T es una transformación lineal.*
- (1 p) *Calcula $[T]_{\beta'}^{\gamma}$, a partir de la matriz $[T]_{\beta}^{\gamma}$ y una matriz de cambio de base.*

- (1 p) *Calcula el Kernel de T , su dimensión y la dimensión de la imagen de T .*

Problema 4.11.10. (1.5 p) *Considera a la transformación lineal $T : M_{2 \times 2}(\mathbb{R}) \longrightarrow \mathbb{R}^4$ dada por*

$$T\left(\begin{pmatrix} a & b \\ c & d \end{pmatrix}\right) = (4a + 3b + 10c + d, 4a + b - 2c - d, 8a - b + 8c + d, -3b - 8c - d).$$

¿Es T un isomorfismo? Argumenta tu respuesta.

Chapter 5

Diagonalización

Fijemos un F -espacio vectorial V de dimensión finita n y una transformación lineal $T: V \rightarrow V$ de V en sí mismo. En el capítulo 4 se estudió cómo, escogiendo una base β de V , podemos definir a la matriz $[T]_{\beta}^{\beta}$, que guarda la información de T en un arreglo de n^2 elementos del campo F . Nos interesa ahora no sólo el poder representar a T con una matriz, también queremos que tal representación matricial sea lo más sencilla posible - por ejemplo, podríamos desear que $[T]_{\beta}^{\beta}$ tenga “muchas” entradas nulas, pues eso reducirá significativamente las operaciones a realizar - y, aún más importante, reducirá los errores numéricos en el proceso de cálculo. Una situación ideal sería poder encontrar β tal que $[T]_{\beta}^{\beta}$ sea *diagonal*, digamos,

$$[T]_{\beta}^{\beta} = \text{Diag}(\lambda_1, \dots, \lambda_n) = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \lambda_n \end{pmatrix}, \quad (5.1)$$

- En este caso, el rango de T - o sea, la dimensión del espacio vectorial $T(V)$ - es igual a la cantidad de entradas no cero de $[T]_{\beta}^{\beta}$ (c.f. Proposición 4.9.6).
- Representaciones matriciales diagonales simplifican la evaluación de T a vectores del espacio; si $x \in V$ y $[x]_{\beta} = (x_1, \dots, x_n)$, entonces

$$[T(x)]_{\beta} = \text{Diag}(\lambda_1, \dots, \lambda_n) \cdot \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} x_1 \lambda_1 \\ x_2 \lambda_2 \\ \vdots \\ x_n \lambda_n \end{pmatrix}.$$

En general, si $m \geq 1$ y por T^m denotamos a la transformación lineal que resulta de componer a T consigo mismo m veces, es fácil demostrar (por ejemplo, por inducción matemática), que

$$[T^m(x)]_{\beta} = \begin{pmatrix} x_1 \lambda_1^m \\ x_2 \lambda_2^m \\ \vdots \\ x_n \lambda_n^m \end{pmatrix}.$$

- El determinante de la matriz (5.1) (que después usaremos para definir el determinante de la transformación T) es simplemente el producto de sus entradas diagonales;

$$\det(\text{Diag}(\lambda_1, \dots, \lambda_n)) = \prod_{j=1}^n \lambda_j.$$

Recuerda que, en la práctica, para hacer cálculos con una transformación $T: V \rightarrow V$, es mucho más cómodo usar representaciones matriciales de esta, es decir, pasar del espacio vectorial $\mathcal{L}(V, V)$ al espacio isomorfo $M_{n \times n}(F)$.

$\dim(V) < \infty$.
 • Espacio $\mathcal{L}(V, V)$.
 $T: V \rightarrow V$ lineal

• Espacio $M_{n \times n}(F)$

$\begin{pmatrix} [T]_{\beta_1}^{\beta_1} & \dots & [T]_{\beta_2}^{\beta_2} \end{pmatrix}$
 (clase de equivalencia \sim)

¿Cuál de todas estas representaciones tomar para estudiar a T ?

La pregunta ahora es qué transformaciones lineales $T : V \longrightarrow V$ admiten una representación matricial diagonal. En este último capítulo, daremos condiciones necesarias y suficientes para la existencia de tal representación, que será usada para **termina. Procesos de Markov, aplicación a Probabilidad.**

5.1 Operadores lineales diagonalizables

Fijemos a un F -espacio vectorial V con $\dim(V) < \infty$.

Definición 5.1.1. Toda transformación lineal $T : V \longrightarrow V$ de un F -espacio vectorial V en sí mismo es llamada un **operador lineal** sobre V .

Definición 5.1.2. Una matriz $D = (d_{ij}) \in M_{n \times n}(F)$ se dice **diagonal** si sólo sus entradas d_{ii} pueden no ser cero.

Recuerda que, dadas $T : V \longrightarrow V$ lineal y una base $\beta \subseteq V$, la clase de equivalencia de la matriz $[T]_{\beta}^{\beta}$ bajo la relación “ser similar a” (c.f. Definición 4.8.7) consta *exactamente* de otras representaciones matriciales $[T]_{\beta'}^{\beta'}$ de la misma transformación lineal T (c.f. Teorema 4.8.9). Nos gustaría saber cuándo T es tal que una de esas representaciones matriciales es diagonal.

Definición 5.1.3. Se dice que un operador lineal $T : V \longrightarrow V$ sobre un F -espacio vectorial finito dimensional V es **diagonalizable** si existe una base β de V tal que $[T]_{\beta}^{\beta}$ es una matriz diagonal. **Diagonalizar un operador** T se refiere a encontrar (cuando sea posible) una base β de V tal que $D := [T]_{\beta}^{\beta}$ es diagonal.

Tenemos una definición dual a la 5.1.3 para matrices:

Definición 5.1.4. Una matriz cuadrada A es **diagonalizable** si A es similar a una matriz diagonal. **Diagonalizar una matriz** A se refiere a encontrar (cuando sea posible) una matriz invertible Q tal que $D = Q^{-1}AQ$.

Cuidado: no confundas los conceptos de “matriz diagonal” y “matriz diagonalizable”.

$$\begin{array}{ll}
 \dim(V) < \infty & \\
 V \xrightarrow{T} V & \text{diagonalizable;} \\
 \exists \beta \subseteq V \text{ base : } D := [T]_{\beta}^{\beta} \text{ es diagonal} & A \in M_{n \times n}(F) \text{ diagonalizable} \\
 & \exists Q, D \in M_{n \times n}(F), Q \text{ invertible} \\
 & D \text{ diagonal } \Rightarrow D = Q^{-1}AQ
 \end{array}$$

Figure 5.1: Los dos conceptos de diagonalización en los espacios isomorfos $\mathcal{L}(V, V)$ y $M_{n \times n}(F)$.

Pensando siempre en establecer una conexión entre las definiciones hechas en el espacio $\mathcal{L}(V, V)$ y el espacio $M_{n \times n}(F)$, relacionamos las dos nociones de “ser diagonalizable” a continuación.

Proposición 5.1.5. Sean V un F -espacio vectorial finito dimensional, $T : V \longrightarrow V$ un operador lineal en V . Las siguientes son equivalentes:

- T es un operador diagonalizable
- Existe una base β de V tal que la matriz $[T]_{\beta}^{\beta}$ es diagonalizable
- Para toda base γ de V , la representación matricial $[T]_{\gamma}^{\gamma}$ es diagonalizable.



Demostración.

Recuerda que con \sim denotamos a la relación de equivalencia “ser similar a”.

$a) \Rightarrow b)$) Sea $\beta \subseteq V$ base tal que $A = [T]_\beta^\beta$ es una matriz diagonal. Como A es trivialmente similar a sí misma, β es una base que satisface el inciso $b)$.

$b) \Rightarrow c)$) Sea β una base de V tal que la matriz $[T]_\beta^\beta$ es diagonalizable; esto significa que existe D matriz diagonal tal que

$$[T]_\beta^\beta \sim D.$$

Si γ es cualquier otra base de V , entonces, por el Teorema 4.8.9, también se tiene

$$[T]_\gamma^\gamma \sim [T]_\beta^\beta.$$

De la transitividad de \sim se deduce que $[T]_\gamma^\gamma \sim D$, o sea, que $[T]_\gamma^\gamma$ es diagonalizable.

$c) \Rightarrow a)$) Sea γ una base cualquiera de V . Por hipótesis, $[T]_\gamma^\gamma$ es diagonalizable, es decir, existe D diagonal tal que $[T]_\gamma^\gamma \sim D$. Según el Teorema 4.8.9, esto implica que la matriz D también es una representación matricial de T respecto a una misma base del espacio. Así, T es diagonalizable.

□

Corolario 5.1.6. La matriz $A \in M_{n \times n}(F)$ es diagonalizable si y sólo si el operador $L_A: F^n \longrightarrow F^n$ es diagonalizable.

Demostración. Recuerda que si β es la base canónica de F^n entonces $[L_A]_\beta^\beta = A$, luego,

- si A es diagonalizable, es decir, si existe D diagonal con $D = Q^{-1}AQ$ para alguna matriz invertible Q , como D es similar a A , por el Teorema 4.8.9 existe otra base $\gamma \in F^n$ tal que $D = [L_A]_\gamma^\gamma$, luego, L_A es un operador diagonalizable;
- recíprocamente, si $L_A: F^n \longrightarrow F^n$ es un operador diagonalizable, entonces existe γ base de F^n tal que $D := [L_A]_\gamma^\gamma$ es diagonal. Como A también es una representación matricial de L_A , D y A son similares, de hecho,

$$D = Q^{-1}AQ, \quad \text{donde } Q = [Id]_\gamma^\beta,$$

luego, A es una matriz diagonalizable.

En este sentido, diagonalizar a la matriz A es lo mismo que diagonalizar al operador L_A .

□

Definidos ya los conceptos que deseamos estudiar, encaremos la pregunta con la que iniciamos el capítulo: ¿qué debe cumplir un operador $T : V \rightarrow V$ para que exista una base $\beta \subseteq V$ tal que $[T]_\beta^\beta$ sea diagonal? Observa que

- si tal $\beta = \{v_1, \dots, v_n\}$ existe, si

$$[T]_\beta^\beta = \begin{pmatrix} \lambda_1 & 0 & 0 & \dots & 0 \\ 0 & \lambda_2 & 0 & \dots & 0 \\ 0 & 0 & \lambda_3 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & \lambda_n \end{pmatrix}, \quad (5.2)$$

podemos usar esta representación para saber cómo actúa T en la base β , pues,

$$\forall 1 \leq j \leq n : T(v_j) = 0v_1 + \dots + \lambda_j v_j + \dots + 0v_n,$$

o sea,

$$\forall 1 \leq j \leq n : T(v_j) = \lambda_j v_j. \quad (5.3)$$

- Recíprocamente, si una base satisface la condición (5.3), entonces claro que su representación matricial respecto a esa base es diagonal, de hecho, es la matriz (5.2).

Queda demostrada así la siguiente

Proposición 5.1.7. Sean V un F -espacio vectorial finito dimensional, $T : V \rightarrow V$ un operador sobre este. T es diagonalizable si y sólo si existe una base $\beta = \{v_j \mid 1 \leq j \leq n\}$ de V para la que existan escalares λ_j , $1 \leq j \leq n$ tales que

$$\forall 1 \leq j \leq n : T(v_j) = \lambda_j v_j.$$

Esta caracterización de diagonalizabilidad motiva las siguientes definiciones.

Definición 5.1.8. Sea $T : V \rightarrow V$ un operador lineal. Un vector $v \in V - \{0\}$ es un **autovector** de T si existe $\lambda \in F$ tal que $T(v) = \lambda v$ (i.e. v y su imagen bajo T son paralelos). Al escalar λ se le llama el **autovalor** asociado a v .

Algunos autores usan los nombres “eigenvector” y “eigenvalue”.

Sea $A \in M_{n \times n}(F)$ una matriz cuadrada. Un vector $v \in F^n - \{0\}$ es un **autovector** de A si es un vector propio del operador L_A .

Puesto que

$$\forall \lambda \in F : T(0) = 0 = \lambda 0,$$

se excluye al vector cero de la definición de autovector.

Con esta nueva terminología, la Proposición 5.1.7 se reformula como “un operador $T : V \rightarrow V$ es diagonalizable si y sólo si existe una base de V compuesta de vectores propios de T ”.

Así, diagonalizar un operador consiste en encontrar una base del espacio que conste de vectores propios de este.

Definición 5.1.9. Sea $T : V \rightarrow V$ operador.

- Denotaremos al conjunto de valores propios de T por

$$\text{spec}(T) := \{\lambda \in F \mid \exists v \in V - \{0\} : T(v) = \lambda v\}. \quad (5.4)$$

A este se la llamará el **espectro** de T .

- Para cada $\lambda \in F$, denotaremos al conjunto de vectores de V que cumplen que $T(v) = \lambda v$, o sea,

$$E_\lambda := \{v \in V \mid T(v) = \lambda v\}. \quad (5.5)$$

Ejercicio: Demuestra que, para cada $\lambda \in F$, E_λ es un subespacio de V .

Ejercicio: Demuestra que un vector propio tiene un y sólo un valor propio. En otras palabras, demuestra que son equivalentes las siguientes proposiciones: para $\lambda, \lambda' \in F$,

- $\lambda \neq \lambda'$.
- $E_\lambda \cap E_{\lambda'} = \{0\}$.
- $E_\lambda \oplus E_{\lambda'}$.

Observación 5.1.10. Para todo $\lambda \in F$, $\lambda \in \text{spec}(T)$ si y sólo si $\{0\} \not\subseteq E_\lambda$.

Ejemplo 5.1.11. (De operadores en \mathbb{R}^2 , uno de ellos diagonalizable y el otro no).

- Considere a un ángulo $\theta \in [0, 2\pi[$, y sea R_θ la rotación respecto al origen de θ radianes *referencia*. ¿Es diagonalizable R_θ ? Determina $\text{Spec}(R_\theta)$.
- Considere a la matriz

$$A = \begin{pmatrix} 1 & 3 \\ 4 & 2 \end{pmatrix}.$$

Puede comprobar que

$$v_1 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}, \quad v_2 = \begin{pmatrix} 3 \\ 4 \end{pmatrix}$$

son ambos vectores propios de A , con respectivos valores propios -2 y 5 . Entonces, si $\beta = \{v_1, v_2\}$,

$$[L_A]_\beta^\beta = \begin{pmatrix} -2 & 0 \\ 0 & 5 \end{pmatrix}.$$

Ejercicio: Demuestra que $\text{Spec}(Id_V) = \{1\}$, y que, en este caso, $E_1 = V$.

Ejemplo 5.1.12. (De los autovalores de un operador en un espacio infinito dimensional).

Considere al espacio $\mathbb{R}^\mathbb{N}$ de sucesiones reales. En este espacio, sea $L : \mathbb{R}^\mathbb{N} \rightarrow \mathbb{R}^\mathbb{N}$ el operador left-shift:

$$L((x_0, x_1, x_2, \dots)) = (x_1, x_2, \dots).$$

Si a partir de $k \in \mathbb{R}$ definimos a la sucesión

$$\hat{k} = (1, k, k^2, k^3, \dots),$$

tenemos que

$$L(\hat{k}) = (k, k^2, k^3, \dots) = k\hat{k}.$$

Así, todo número real es autovalor de L .

Ejemplo 5.1.13. Observe que si $T : V \rightarrow V$ no es inyectivo, entonces tiene una infinidad de vectores propios con valor propio cero. En efecto, el que T no sea inyectivo significa que $\{0\} \neq \text{Ker}(T)$; sea pues $v \in \text{Ker}(T) - \{0\}$.

$$T(v) = 0_V = 0_F \cdot v.$$

De hecho,

$$\forall a \in F : T(av) = aT(v) = a \cdot 0_V = 0_V.$$

Así, todo elemento de $\text{span}(\{v\})$ es un vector propio con valor propio cero.

En el Ejemplo 5.1.11 se te dieron sugerencias de autovectores de la matriz A , ¿cómo encontramos (de existir) los autovalores y autovectores de un operador?

El siguiente resultado es extremadamente útil en la práctica, pues nos permite pasar de un proceso de diagonalización en un espacio vectorial finito dimensional cualquiera V (en el que puede ser difícil trabajar directamente) a un proceso de diagonalización en el espacio $M_{n \times n}(F)$ (espacio en el que nunca tenemos problemas para realizar los cálculos correspondientes).

Proposición 5.1.14. $T : V \longrightarrow V$ operador, $\beta \subseteq V$ base, $A = [T]_{\beta}^{\beta}$ su representación matricial respecto a β . T y A tienen los mismos valores propios; además, $v \in V$ es vector propio de T si y sólo si $x = [v]_{\beta}$ es vector propio de L_A .

Una aplicación importante de esta Proposición se muestra en el Ejemplo [referencia](#).

$$\begin{array}{ccc}
 v \in V & \xrightarrow{T} & V \\
 \downarrow [\cdot]_{\beta} & & \downarrow [\cdot]_{\beta} \\
 x = [v]_{\beta} \in F^n & \xrightarrow{L_A} & F^m
 \end{array}$$

Demostración. En efecto, dado $v \in V$, si $x = [v]_{\beta}$, entonces

$$\begin{aligned}
 T(v) = \lambda v & \text{ sii } [T(v)]_{\beta} = [\lambda v]_{\beta} \\
 & \text{ sii } [T]_{\beta}^{\beta} [v]_{\beta} = \lambda [v]_{\beta} \\
 & \text{ sii } Ax = \lambda x \\
 & \text{ sii } L_A(x) = \lambda x.
 \end{aligned}$$

La Proposición 5.1.14 nos permite cambiar la pregunta ¿es x un autovector de T ? por ¿es $[x]_{\beta}$ un autovector de L_{A3} ?

□

5.2 El determinante de un operador y su polinomio característico

Sea $T : V \longrightarrow V$ un operador. Observa que, dado $\lambda \in F$ cualquiera,

$$\begin{aligned}
 T(v) = \lambda v & \text{ sii } T(v) = \lambda Id(v) \\
 & \text{ sii } (T - \lambda Id)(v) = 0 \\
 & \text{ sii } v \in \text{Ker}(T - \lambda Id),
 \end{aligned}$$

luego, λ es un autovalor de T si y sólo si en el Kernel del operador en V

$$U_{\lambda} := T - \lambda Id$$

hay vectores no cero - es decir, si y sólo si U_{λ} no es inyectiva (c.f Proposición 3.3.1). Recordando las notaciones introducidas en la Definición 5.1.9, tenemos que

$$\forall \lambda \in F : E_{\lambda} = \text{Ker}(T - \lambda Id) = \text{Ker}(U_{\lambda}). \quad (5.6)$$

Por (5.6) identificamos a los espacios de autovalores E_{λ} para $\lambda \in \text{spec}(T)$ como subespacios de V .

Resumimos esto en la siguiente

Observación 5.2.1. Para todo $\lambda \in F$, $\lambda \in \text{spec}(T)$ si y sólo si $\{0\} \not\subseteq \text{Ker}(U_\lambda)$.

Este razonamiento nos permite identificar fácilmente los autovalores de T - y, a partir de ellos, a los autovectores de T .

Proposición 5.2.2. Sea $T : V \longrightarrow V$ un operador lineal en un espacio V finito dimensional. Sea $A = [T]_\beta^\beta$ una representación matricial de T . Las siguientes son equivalentes:

- $\lambda \in F$ es un autovalor de T .
- $\det(A - \lambda I_n) = 0$.

Demostración. Por lo discutido arriba, si $U_\lambda := T - \lambda Id : V \longrightarrow V$, entonces λ es un autovalor de T si y sólo si U_λ no es inyectiva; como V es finito dimensional, esto equivale a que U_λ no sea invertible (c.f. Proposición 3.4.4); esto, según el Teorema [cita](#) equivale a que su representación matricial

$$[U_\lambda]_\beta^\beta = [T - \lambda Id]_\beta^\beta = A - \lambda I_n$$

no sea invertible, lo que a su vez equivale a que el determinante de $A - \lambda I_n$ sea cero. \square

Lema 5.2.3. Sean V finito dimensional, $\beta, \beta' \subseteq V$ bases, $T : V \longrightarrow V$ un operador en V . Se tiene que

$$\det([T]_\beta^\beta) = \det([T]_{\beta'}^{\beta'}).$$

Demostración. En efecto, haciendo $A = [T]_\beta^\beta$ y $B = [T]_{\beta'}^{\beta'}$, sólo note que si Q es la matriz invertible $[Id_n]_{\beta'}^\beta$, entonces

$$\begin{aligned} \det(B) &= \det(Q^{-1} A Q) = \det(Q^{-1}) \det(A) \det(Q) \\ &= \det(Q)^{-1} \det(A) \det(Q) = \det(A). \end{aligned}$$

Recuerda que, si $A, B \in M_{n \times n}(F)$, entonces $\det(AB) = \det(A) \cdot \det(B)$. De esto se deduce además que, si A es invertible, entonces $\det(A^{-1}) = \det(A)^{-1}$.

\square

Lo demostrado en el Lema 5.2.3 hace legítima la siguiente

Definición 5.2.4. Dado $T : V \longrightarrow V$ en un espacio finito dimensional V , definimos su **determinante** como

$$\det(T) = \det([T]_\beta^\beta), \quad \beta \subseteq V \text{ base.}$$

Lema 5.2.5. Sea $T : V \longrightarrow V$ un operador. Si λ es un escalar, β es una base de V y $A = [T]_\beta^\beta$, entonces

$$\det(T - \lambda Id) = \det(A - \lambda I). \quad (5.7)$$

Demostración. Recordando que $[Id]_\beta^\beta = I_n$ y que la representación matricial de una suma es la suma de las representaciones, tenemos que

$$\det(T - \lambda Id) = \det([T - \lambda Id]_\beta^\beta) = \det([T]_\beta^\beta - \lambda [Id]_\beta^\beta) = \det(A - \lambda I_n).$$

\square

Podemos reformular a la Proposición 5.2.2 como sigue.

Corolario 5.2.6. Dado $T : V \rightarrow V$ operador en un F -espacio finito dimensional V , $\lambda \in F$ es un autovalor de T si y sólo si $\det(U_\lambda) = 0$, donde $U_\lambda = T - \lambda Id$.

Definición 5.2.7. Sea $T : V \rightarrow V$ un operador lineal en un espacio finito dimensional V . El polinomio

$$p(\lambda) := \det(T - \lambda Id)$$

es llamado el **polinomio característico de T** .

Proposición 5.2.8. Sean $T : V \rightarrow V$ operador, con $\dim(V) = n$, $p(\lambda)$ su polinomio característico.

- Las raíces de $p(\lambda)$ son exactamente los autovalores de T , o sea,

$$\forall \lambda \in F : \lambda \in \text{Spec}(T) \Leftrightarrow p(\lambda) = 0.$$

- El grado de $p(\lambda)$ es n .
- Si T es diagonalizable, entonces p se factoriza como un producto de n polinomios de grado 1 no necesariamente distintos entre sí.

Demostración. Si β es base de V tal que $[T]_\beta^\beta = \text{Diag}(\lambda_1, \dots, \lambda_n)$, entonces

$$p(\lambda) = \det(T - \lambda I) = \det \begin{pmatrix} \lambda_1 - \lambda & 0 & \dots & 0 \\ 0 & \lambda_1 - \lambda & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \lambda_n - \lambda \end{pmatrix} = \prod_{j=1}^n (\lambda_j - \lambda).$$

Según la Proposición 5.2.8, si el polinomio característico de un operador T no se descompone como producto de polinomios de grado 1, entonces T no es diagonalizable.

□

Mostremos con un ejemplo que la suficiencia en el tercer punto de la Proposición 5.2.8 no se tiene; aún cuando el polinomio característico de $T : V \rightarrow V$ se pueda factorizar como producto de polinomios de grado uno, puede ser que T no sea diagonalizable.

Ejemplo 5.2.9. Sea $T : P_2(\mathbb{R}) \rightarrow P_2(\mathbb{R})$ el operador derivación. Si consideramos a la base $\beta = \{1, x, x^2\}$ del espacio de polinomios reales de grado a lo más 2, se calcula que

$$[T]_\beta^\beta = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{pmatrix}.$$

A partir de esta representación matricial, calculamos al polinomio característico de p ;

$$p(\lambda) = \det \begin{pmatrix} -\lambda & 1 & 0 \\ 0 & -\lambda & 2 \\ 0 & 0 & -\lambda \end{pmatrix} = -\lambda^3.$$

Así, $\text{Spec}(T) = \{0\}$, o sea, T sólo tiene autovalores de autovector cero.

$$\begin{aligned} E_0 &= \{f \in P_2(\mathbb{R}) : f' = 0\} \\ &= \{f \in P_2(\mathbb{R}) : f \text{ es un polinomio constante}\}. \end{aligned}$$

Claro que $\dim(E_0) = 1$, luego, es imposible formar una base del espacio $P_2(\mathbb{R})$ - que tiene dimensión 3 - a partir de autovectores de T , es decir, T no es diagonalizable, pues no hay “suficientes” autovectores de T linealmente independientes. \diamond

Corolario 5.2.10. Sean $D_1, D_2 \in M_{n \times n}(F)$ matrices cuadradas diagonales, digamos

$$D_1 = \text{Diag}(\lambda_1, \dots, \lambda_n), \quad D_2 = \text{Diag}(\mu_1, \dots, \mu_n).$$

Si D_1 y D_2 son similares, entonces, salvo permutación, tienen las mismas entradas diagonales, es decir, existe $\rho: \{1, \dots, n\} \rightarrow \{1, \dots, n\}$ biyección tal que

$$\forall 1 \leq j \leq n: \quad \lambda_j = \mu_{\rho(j)}.$$

Demostración. Consideremos al operador $T = L_{D_1}$ en F^n . Si β es la base canónica de F^n ,

$$D_1 = [T]_{\beta}^{\beta}. \quad (5.8)$$

Por hipótesis, $D_1 \sim D_2$, luego, según el Teorema 4.8.9, existe γ base de V tal que

$$D_2 = [T]_{\gamma}^{\gamma}. \quad (5.9)$$

Podemos usar estas dos representaciones matriciales para calcular el polinomio característico de T ;

$$p(\lambda) = \det \begin{pmatrix} \lambda_1 - \lambda & 0 & 0 & \dots & 0 \\ 0 & \lambda_2 - \lambda & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & \lambda_n - \lambda \end{pmatrix} = \prod_{j=1}^n (\lambda_j - \lambda);$$

$$p(\lambda) = \det \begin{pmatrix} \mu_1 - \lambda & 0 & 0 & \dots & 0 \\ 0 & \mu_2 - \lambda & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & \mu_n - \lambda \end{pmatrix} = \prod_{j=1}^n (\mu_j - \lambda);$$

o sea,

$$\prod_{j=1}^n (\lambda_j - \lambda) = \prod_{j=1}^n (\mu_j - \lambda) \quad (\text{ecuación en el dominio entero } F[\lambda]). \quad (5.10)$$

Por ser F campo, los polinomios de grado uno son irreducibles (c.f. [Rot15]), luego, (5.10) implica que los polinomios de grado uno que aparecen en el producto de la derecha son, salvo un reordenamiento, exactamente los que aparecen en el producto de la izquierda; esto significa que las λ_j son, salvo un reordenamiento, las μ_j . \square

Ejercicio: Si $T: \mathbb{C}^n \rightarrow \mathbb{C}^n$ es un operador en el \mathbb{C} -espacio vectorial \mathbb{C}^n , demuestra que $\text{Spec}(T) \neq \emptyset$. Contrasta esto con el ejemplo 5.1.11.

Definición 5.2.11. Sean V finito dimensional, $T: V \rightarrow V$ un operador en V . Si $\lambda_0 \in F$ es un autovalor de T , definimos su **multiplicidad algebraica** como el mayor entero $m \geq 1$ tal que $(\lambda - \lambda_0)^m$ divide al polinomio característico $p(\lambda)$ de T .

Nota que, si T es diagonalizable - digamos que β es base de V compuesta de autovectores de T - y si $\lambda_1 \in \text{Spec}(T)$, entonces la multiplicidad algebraica de λ_1 corresponde a la cantidad de elementos de β que también están en el espacio E_{λ_1} . Observaciones de este tipo nos ayudarán a construir criterios y algoritmos para la diagonalización de operadores. Antes de seguir desarrollando la teoría, apliquémosla en algunos ejemplos.

Ejercicio: Fijado $\theta \in [0, 2\pi[$ ángulo - medido en radianes - calcula el polinomio característico de la rotación $R_{\theta}: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ por θ radianes, y úsalo para calcular el espectro de R_{θ} . Compara esto con el ejercicio [cita](#).

Nota que la ecuación análoga a (5.10) en el campo F sería

$$\prod_{j=1}^n \lambda_j = \prod_{j=1}^n \mu_j,$$

y de esta **no** se puede deducir de inmediato que las λ_j son las μ_j permutadas, pues no necesariamente son primos en F . Este problema se arregla si consideramos, como en nuestra demostración, a los irreducibles $(\lambda - \lambda_j)$ en $F[\lambda]$.

5.3 Ejemplos de diagonalización

Ilustremos con unos ejemplos cómo diagonalizar matrices u operadores lineales. Como se podrá apreciar,

- es mucho más conveniente buscar primero los autovalores de un operador (i.e. el conjunto $\text{spec}(T)$) para así saber qué espacios de autovectores $E_\lambda \leq V$ calcular, y
- cuando en el espacio V se presenten dificultades a la hora de calcular los espacios de autovectores E_λ , siempre podremos usar la Proposición 5.1.14 para en su lugar trabajar en un espacio de matrices, en el que nunca tendremos problemas para realizar los cálculos.

Ejemplo 5.3.1. Sea

$$A = \begin{pmatrix} 1 & 1 \\ 4 & 1 \end{pmatrix} \in M_{2 \times 2}(\mathbb{R}).$$

Vamos a diagonalizar a la matriz A ; para ello, consideremos al operador $L_A : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ dado por

$$L_A((x, y)) = \begin{pmatrix} 1 & 1 \\ 4 & 1 \end{pmatrix} \cdot \begin{pmatrix} x \\ y \end{pmatrix} = (x + y, 4x + y).$$

Usemos primero la Proposición 5.2.2 para determinar los autovalores de L_A .

$$\det(A - \lambda I_2) = \det \begin{pmatrix} 1 - \lambda & 1 \\ 4 & 1 - \lambda \end{pmatrix} = \lambda^2 - 2\lambda - 3 = (\lambda - 3)(\lambda + 1),$$

luego, $\det(A - \lambda I_2)$ es cero si y sólo si $\lambda = 3, -1$. De esto se concluye que

$$\text{spec}(A) = \{3, -1\}. \quad (5.11)$$

Calculemos ahora a los espacios de autovalores E_3 y E_{-1} . Por definición, $(x, y) \in E_3$ si y sólo si $L_A(x, y) = (3x, 3y)$. Usando la fórmula para L_A e igualando entradas, llegamos a sistema de ecuaciones

$$\begin{cases} x + y = 3x, \\ 4x + y = 3y, \end{cases}$$

que equivale a la única ecuación $-2x + y = 0$. Así,

$$E_3 = \{(x, 2x) \in \mathbb{R}^2 \mid x \in \mathbb{R}\} = \{(1, 2)\}.$$

Similarmente se calcula que

$$E_{-1} = \{(x, -2x) \in \mathbb{R}^2 \mid x \in \mathbb{R}\} = \{(1, -2)\}.$$

Tomando un vector de E_3 y otro de E_{-1} , formamos a

$$\beta = \{(1, 2), (1, -2)\},$$

que es una base de \mathbb{R}^2 ; puesto que

$$L_A((1, 2)) = 3(1, 2), \quad L_A((1, -2)) = -1(1, 2),$$

tenemos que

$$[L_A]_\beta^\beta = \begin{pmatrix} 3 & 0 \\ 0 & -1 \end{pmatrix}. \quad (5.12)$$

Como A y la matriz (5.12) representan al mismo operador L_A , tenemos que A y (5.12) son similares, luego, A es diagonalizable en el sentido de la definición 5.1.4.

Ejercicio: Calcula Q invertible tal que $D = Q^{-1}AQ$. \diamond

Observa que, en el ejemplo anterior, intentar adivinar los autovalores λ de A (si es que existen) y calcular los respectivos autovectores (i.e. los espacios E_λ) no es práctico. Por ejemplo, si quisiéramos ver si $\lambda = 2$ es un autovalor de A , tendríamos que calcular a E_2 y ver si tiene vectores distintos de cero. Por definición, $(x, y) \in E_2$ si y sólo si $L_A(x, y) = (2x, 2y)$, condición equivalente al sistema de ecuaciones

$$\begin{cases} x + y &= 2x, \\ 4x + y &= 2y, \end{cases}$$

que equivale a

$$\begin{cases} -x + y &= 0, \\ 4x - y &= 0, \end{cases}$$

cuya única solución es $x = 0 = y$, luego, $E_2 = \{(0, 0)\}$, y L_A no tiene autovectores de autovalor 2.

Ejemplo 5.3.2. Sea el operador $T : P_2(\mathbb{R}) \longrightarrow P_2(\mathbb{R})$ definido por

$$T(f(x)) = f(x) + xf'(x) + f'(x). \quad (5.13)$$

Queremos ver si T es diagonalizable o no. Fijemos a la base $\beta = \{1, x, x^2\}$. Puesto que

$$T(1) = 1, \quad T(x) = 1 + 2x, \quad T(x^2) = 2x + 3x^2,$$

se calcula que

$$A := [T]_\beta^\beta = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 2 & 2 \\ 0 & 0 & 3 \end{pmatrix}. \quad (5.14)$$

Para encontrar los autovalores de T , como dice la Proposición 5.2.2 vamos a calcular los ceros del siguiente polinomio en λ :

$$\det(A - \lambda I_3) = \det \begin{pmatrix} 1 - \lambda & 1 & 0 \\ 0 & 2 - \lambda & 2 \\ 0 & 0 & 3 - \lambda \end{pmatrix} = (1 - \lambda)(2 - \lambda)(3 - \lambda).$$

Claro entonces que

$$\text{spec}(T) = \{1, 2, 3\}.$$

Intentar caracterizar a los autovectores de autovalor 3 significa encontrar a los polinomios $f(x) = a + bx + cx^2$ tales que $T(f(x)) = 3f(x)$, o sea, los polinomios de grado a lo más dos que satisfacen la ecuación diferencial

$$-2f(x) + xf'(x) + f'(x) = 0, \quad f(x) \in P_2(\mathbb{R}). \quad (5.15)$$

Es mucho más sencillo usar la Proposición 5.1.14 para trasladar el problema de diagonalización del espacio $P_2(\mathbb{R})$ (en el que no sabemos cómo caracterizar a los autovectores) al espacio $M_{3 \times 3}(\mathbb{R})$ (en el que calcular espacios de autovectores se reduce a realizar cuentas sencillas). Sea el isomorfismo asociado a la base β $[\cdot]_\beta : P_2(\mathbb{R}) \longrightarrow \mathbb{R}^3$, dado por

$$\forall f(x) = a + bx + cx^2 : [f(x)]_\beta = (a, b, c).$$

Consideremos al operador $L_A : \mathbb{R}^3 \longrightarrow \mathbb{R}^3$ dado por

$$L_A((x, y, z)) = (x + y, 2y + 2z, 3z).$$

Según la Proposición 5.1.14, también se tiene

$$\text{spec}(L_A) = \{1, 2, 3\}.$$

Se tiene que $(x, y, z) \in \mathbb{R}^3$ es un autovector de autovalor 3 si y sólo si $L_A(x, y, z) = (3x, 3y, 3z)$, o sea, si y sólo si $(x + y, 2y + 2z, 3z) = (3x, 3y, 3z)$. Esto equivale al sistema

$$\begin{cases} x + y &= 3x \\ 2y + 2z &= 3y \\ 3z &= 3z \end{cases}$$

Una terna $(x, y, z) \in \mathbb{R}^3$ satisface este sistema si y sólo si $x, z = \frac{1}{2}y$. Así,

$$E_3 = \{(0.5y, y, 0.5y) \mid y \in \mathbb{R}\} = \{\{(1, 2, 1)\}\}.$$

Recuerda que lo que queremos es el espacio de autovectores **en el espacio** $P_2(\mathbb{R})$ **correspondiente a** T , el operador con el que estabamos trabajando originalmente. Según la Proposición 5.1.14, este es

$$[E_3]_{\beta}^{-1} = \{f(x) \in P_2(\mathbb{R}) \mid [f(x)]_{\beta} \in E_3\} = \{f(x) = a + (2a)x + ax^2 \mid a \in \mathbb{R}\}. \quad (5.16)$$

Así, estamos seguros de que el conjunto (5.16) consta de las soluciones de la ecuación diferencial (5.15). Por ejemplo, puedes comprobar que

$$T(1 + 2x + x^2) = 3 + 6x + 3x^2.$$

Ejercicio: usando esta misma técnica, encuentra los espacios de autovectores de T asociados a los autovalores 1 y 2. Usa a estos para encontrar una base de $P_2(\mathbb{R})$ que conste de autovalores de T , y calcula la representación matricial de T respecto a esta, que claro que será diagonal. \diamond

5.4 Criterios para determinar cuándo un operador es diagonalizable

En lo que sigue, siempre se va a suponer a V finito dimensional, $\dim(V) = n$. Diagonalizar al operador $T : V \rightarrow V$ es encontrar una base de V compuesta por autovectores de T . Demos un primer resultado sobre cómo construir subconjuntos linealmente independientes a partir de autovectores.

Proposición 5.4.1. Sean $T : V \rightarrow V$ lineal, $\lambda_1, \dots, \lambda_k \in \text{Spec}(T)$ autovalores de T distintos entre sí. Si x_1, \dots, x_k son autovectores correspondientes a estos valores, entonces $\{x_1, \dots, x_k\}$ es linealmente independiente.

Demostración. Por inducción sobre k .

- *Base de inducción:* si $k = 1$, entonces $x_1 \neq 0$ por ser autovector de T , luego, el singulete $\{x_1\}$ es l.i..
- *Paso inductivo:* supongamos el teorema cierto para $k-1 \geq 1$, y sean x_1, \dots, x_k autovectores de T con autovalores $\lambda_1, \dots, \lambda_k$, todos distintos entre sí. Sean $a_1, \dots, a_k \in F$ tales que

$$a_1x_1 + \dots + a_{k-1}x_{k-1} + a_kx_k = 0. \quad (5.17)$$

Aplicando T a ambos lados de la igualdad, tenemos que

$$a_1 \lambda_1 x_1 + \cdots + a_{k-1} \lambda_{k-1} x_{k-1} + a_k \lambda_k x_k = 0. \quad (5.18)$$

Multiplicando a (5.17) por λ_k y restando con (5.18), se llega a que

$$a_1 (\lambda_1 - \lambda_k) x_1 + \cdots + a_{k-1} (\lambda_{k-1} - \lambda_k) x_{k-1} = 0;$$

por hipótesis de inducción, $\{x_1, \dots, x_{k-1}\}$ es l.i., luego, los coeficientes de esta última combinación lineal son todos cero;

$$\forall 1 \leq i \leq k-1 : a_i (\lambda_i - \lambda_k) = 0. \quad (\text{ecuación en el campo } F).$$

Puesto que, para toda i , $\lambda_i - \lambda_k \neq 0$, concluimos que

$$\forall 1 \leq i \leq k-1 : a_i = 0.$$

Sustituyendo en (5.17), también se tiene que

$$a_k x_k = 0.$$

Como el vector x_k no es cero, concluimos que debe ocurrir también $a_k = 0$ (c.f. Proposición 2.4.2). Así, $\{x_1, \dots, x_j\}$ es l.i.

□

Desde aquí pon una cita a las equiv. de ser suma directa en general.

Corolario 5.4.2. Sea $T : V \longrightarrow V$ lineal. Si $\lambda_1, \dots, \lambda_k \in \text{Spec}(T)$ son autovalores de T distintos entre sí, entonces

$$E_{\lambda_1} \oplus \cdots \oplus E_{\lambda_k}.$$

Demostración. Para $1 \leq j \leq k$, sean $x_j \in E_{\lambda_j}$ tales que

$$x_1 + x_2 + \cdots + x_k = 0. \quad (5.19)$$

O sea, la suma de autoespacios de un operador siempre es directa.

Nota que, en (5.19), todos los sumandos deben ser cero; de lo contrario, (5.19) es una combinación lineal no trivial igual a cero de autovectores de T que corresponden a autovalores de T todos distintos entre sí, contradiciendo la independencia lineal establecida en la Proposición 5.4.1. □

La Proposición 5.4.1 nos da una idea de cómo formar subconjuntos l.i. de V a partir de autovectores de T ; como mostramos en el Ejemplo 5.2.9, a veces tales subconjuntos l.i. no podrán ser extendidos a una base del espacio respecto a la cual el operador tenga una representación diagonal.

Corolario 5.4.3. Si $\dim(V) = n$ y $T : V \longrightarrow V$ es un operador lineal sobre V , entonces

- $|\text{Spec}(T)| \leq n$
- Si $|\text{Spec}(T)| = n$, entonces T es diagonalizable.

La otra implicación del Corolario (5.4.3) no se cumple; el que $T : V \longrightarrow V$ sea diagonalizable no implica que T tenga n autovalores distintos - tómese, por ejemplo, $T = Id_V$.

Proposición 5.4.4. Sea $T : V \longrightarrow V$ operador, con $\dim(V) < \infty$. Si $\lambda \in \text{Spec}(T)$ y m es su multiplicidad algebraica, entonces

$$1 \leq \dim(E_\lambda) \leq m.$$

Demostración. En efecto, si $\lambda \in \text{Spec}(T)$, es porque existe $v \neq 0$ tal que $v \in E_\lambda$, luego, $\dim(E_\lambda) \geq 1$.

Sea ahora $\{x_1, \dots, x_p\}$ una base para E_λ ; extendamos este l.i. a una base $\beta = \{x_1, \dots, x_n\}$ del espacio. Claro que

$$[T]_\beta^\beta = \begin{pmatrix} \lambda I_p & B_2 \\ 0 & B_3 \end{pmatrix}$$

usando esta representación matricial de T para calcular su polinomio característico,

$$p(t) = \det \begin{pmatrix} (\lambda - t)I_p & B_2 \\ 0 & B_3 - tI_{n-p} \end{pmatrix} = (\lambda - t)^p \cdot \det(B_3 - tI_{n-p}),$$

por lo tanto, $p \leq m$.

□

Agrega en los ejercicios de antes equivalencias a ser suma directa.

Teorema 5.4.5. Sean V un F -espacio vectorial con $\dim(V) = n$, $T : V \longrightarrow V$ un operador lineal tal que su polinomio característico $p(\lambda)$ se descompone como producto de polinomios de grado 1. Sean

$$\text{Spec}(T) = \{\lambda_1, \dots, \lambda_k\},$$

m_j la multiplicidad algebraica del valor propio λ_j ($1 \leq j \leq k$), y

$$d_j = \dim(E_{\lambda_j}), \quad 1 \leq j \leq k.$$

Son equivalentes,

i) T es diagonalizable.

ii) $V = E_{\lambda_1} \oplus \dots \oplus E_{\lambda_k}$.

iii) $n = \sum_{j=1}^k d_j$.

iv) $\forall 1 \leq j \leq k: d_j = m_j$.

Demostración. Observa antes que, como el grado de $p(\lambda)$ es n ,

$$\sum_{j=1}^k m_j = n. \quad (5.20)$$

Además, según el Corolario 5.4.2, la suma de los autoespacios E_λ siempre es directa.

i) \Rightarrow ii) El que T sea diagonalizable significa que existe una base β de V tal que

$$\beta \subseteq \bigcup_{j=1}^k E_{\lambda_j},$$

luego, claro que $V = E_{\lambda_1} \oplus \dots \oplus E_{\lambda_k}$ (recuerda que el espacio suma es, por definición, el generado por la unión).

Es natural pedir en el Teorema 5.4.5 pedir que $p(\lambda)$ se pueda expresar como producto de polinomios de grado uno pues, según la Proposición 5.2.8, en caso contrario, T definitivamente no es diagonalizable.

ii) \Rightarrow iii) Se tiene que

$$\begin{aligned} n = \dim(V) &= \dim(E_{\lambda_1} \oplus \cdots \oplus E_{\lambda_k}) \\ &= \sum_{j=1}^k \dim(E_{\lambda_j}) = \sum_{j=1}^k d_j. \end{aligned}$$

iii) \Rightarrow iv) Supongamos que $\sum_{j=1}^k d_j = n$. Usando (5.20) y la Proposición 5.4.4, se tiene que

$$n = \sum_{j=1}^k d_j \leq \sum_{j=1}^k m_j = n,$$

luego,

$$\sum_{j=1}^k (m_j - d_j) = 0, \quad \text{donde, para toda } j, m_j - d_j \geq 0,$$

por lo tanto, para toda j se tiene $m_j = d_j$.

iv) \Rightarrow i) Supongamos que, para toda j , $d_j = m_j$. Considere al subespacio de V

$$W = E_{\lambda_1} \oplus \cdots \oplus E_{\lambda_k}.$$

Si β_j es base de E_{λ_j} , como la suma es directa, $\beta := \beta_1 \sqcup \cdots \sqcup \beta_k$ es base de W . Pero,

$$|\beta| = |\beta_1 \sqcup \cdots \sqcup \beta_k| = \sum_{j=1}^k |\beta_j| = \sum_{j=1}^k d_j = \sum_{j=1}^k m_j = n,$$

luego, el l.i. β es base de V . Así,

$$V = \text{span}(\beta) = W = E_{\lambda_1} \oplus \cdots \oplus E_{\lambda_k}.$$

Recuerda que con \sqcup denotamos la unión de una familia de conjuntos disjuntos dos a dos.

□

Usando el Teorema de la dimensión 3.2.3, podemos reescribir

$$\begin{aligned} m_j = \dim(E_{\lambda_j}) &= \dim(\text{Ker}(T - \lambda Id)) \\ &= n - \text{Rank}(T - \lambda Id). \end{aligned}$$

Termina demostrando que T es diagonalizable si y sólo si V puede descomponerse como suma directa de subespacios T -invariantes unidimensionales. Esto para conectar con los temas que van a ver en Álgebra Lineal II.

Bibliography

- [Axl] Sheldon Axler. *Linear Algebra Done Right*. fourth edition.
- [Bro] William C. Brown. *A second course in linear algebra*. John Wiley **and** Sons.
- [FSH03] Spence L. E. Friedberg S. H. Insel A. J. *Linear Algebra*. fourth edition. Prentice Hall, 2003.
- [God] Roger Godement. *Cours d'algèbre*. Hermann éditeurs.
- [Gol] Jonathan S. Golan. *Foundations of Linear Algebra*. Kluwer Academic Publishers.
- [Jac85] Nathan Jacobson. *Basic Algebra I*. second edition. W.H. Freeman **and** Company, 1985.
- [Mej11] Hugo Alberto Rincón Mejía. *Álgebra Lineal*. Las prensas de ciencias, 2011.
- [Rot15] Joseph J. Rotman. *Advanced Modern Algebra*. Graduate Texts in Mathematics. American Mathematical Society, 2015.