

Introducción

Este curso aborda lo esencial del **Álgebra Lineal**, la cual es una de las herramientas más ubicuas dentro de la ingeniería. El álgebra lineal es la teoría que establece qué son los *vectores*, cómo operan entre sí y qué propiedades cumplen. Define el concepto de *dimensión* de manera formal y estricta, permitiendo llevarlo a ámbitos más allá de lo espacial.

En el curso pasado estudiaron el conjunto \mathbb{R}^2 , como pares ordenados o vectores. En este curso vamos a generalizar los vectores a \mathbb{R}^n , pero también veremos otros objetos matemáticos, como las funciones, matrices, polinomios y sucesiones, cada uno en su ámbito será un caso particular de lo que llamaremos *espacio vectorial*.

También trata sobre las funciones que se pueden aplicar a los vectores, más precisamente de las *funciones lineales*, que son muy simples, permiten un análisis muy profundo y aplicaciones muy amplias. Están fuertemente ligados al concepto de *matriz*, de manera que repasaremos este objeto también. Aprenderemos conceptos como los *valores propios*, que juegan un rol central en el análisis de sistemas dinámicos, estructuras, y análisis de sistemas en general, pues rescatan parte esencial de las funciones.

Esperamos que les guste, avanzaremos con lentitud en cada concepto, deberán hacer consultas continuamente para ir entendiendo paso a paso, y respondiendo los tests formativos que les entregaremos cotidianamente.

Cuerpo

Definición 1 (Cuerpo) Considera un conjunto \mathbb{K} donde se definen dos operaciones:

$$\begin{array}{ll} + : \mathbb{K} \times \mathbb{K} \rightarrow \mathbb{K} & \cdot : \mathbb{K} \times \mathbb{K} \rightarrow \mathbb{K} \\ (x, y) \rightarrow x + y & (x, y) \rightarrow x \cdot y \end{array}$$

La tripleta $(\mathbb{K}, +, \cdot)$ se llamará Cuerpo Algebráico o Cuerpo si:

1. $(\forall x, y, z \in \mathbb{K}) x + (y + z) = (x + y) + z,$ *(asociatividad de +)*
2. $(\forall x, y \in \mathbb{K}) x + y = y + x,$ *(comutatividad de +)*
3. $(\exists \mathcal{O} \in \mathbb{K})(\forall x \in \mathbb{K}) \mathcal{O} + x = x,$ *(existencia de neutro para +)*
4. $(\forall x \in \mathbb{K})(\exists y \in \mathbb{K}) x + y = \mathcal{O}.$ *(existencia de inverso para +)*
5. $(\forall x, y, z \in \mathbb{K}) x \cdot (y \cdot z) = (x \cdot y) \cdot z,$ *(asociatividad de ·)*
6. $(\forall x, y \in \mathbb{K}) x \cdot y = y \cdot x,$ *(comutatividad de ·)*
7. $(\exists 1 \in \mathbb{K})(\forall x \in \mathbb{K}) 1 \cdot x = x,$ *(existencia de neutro para ·)*
8. $(\forall x \in \mathbb{K} \setminus \{\mathcal{O}\})(\exists y \in \mathbb{K}) x \cdot y = 1.$ *(existencia de inverso para ·)*
9. $(\forall x, y, z \in \mathbb{K}) x \cdot (y + z) = x \cdot y + x \cdot z.$ *(distributividad)*

Ejemplos de cuerpo son:

- $(\mathbb{R}, +, \cdot)$ con las operaciones usuales $+$ y \cdot en $\mathbb{R}.$
- $(\mathbb{C}, +, \cdot)$ donde $(a + bi) + (c + di) = (a + c) + (b + d)i$ y $(a + bi) \cdot (c + di) = (ac - bd) + (ad + bc)i$

Pero revisemos, ¿qué otros conjuntos conocen ustedes que cumplan estas propiedades?

- ¿los números naturales las cumplen?
- ¿los números racionales las cumplen?
- ¿el conjunto de los polinomios las cumplen?
- ¿los complejos las cumplen?

Mientras responden estas preguntas, les contamos que el interés de definir el concepto de *cuerpo* está en poder trabajar al mismo tiempo con todos los conjuntos que cumplen estas propiedades. La teoría que desarrollaremos será válida para todos los cuerpos que conocemos y que podamos llegar a encontrar alguna vez en la vida. Los teoremas que vamos a demostrar, serán válidos de momento que el conjunto con el que trabajemos cumpla las 9 propiedades enumeradas, y podremos hacer uso de los teoremas (sin tener que demostrarlos cada vez) con solo verificar las 9 propiedades, un gran ahorro, ¿no?

Analicemos el conjunto de los polinomios con la suma (+) y el producto (·) de polinomios, veamos, sean $p, q, r \in \mathcal{P}(\mathbb{K})$ la suma + es:

- *Commutativa ($p + q = q + p$). En efecto, dado $x \in \mathbb{K}$ cualquiera se cumple $(p + q)(x) = p(x) + q(x)$, entonces usando la commutatividad en \mathbb{K} , esto es igual a $q(x) + p(x) = (q + p)(x)$, y hemos probado que $\forall x \in \mathbb{K}$, $(p + q)(x) = (q + p)(x)$, es decir: $p + q = q + p$.*
- *Asociativa $((p + q) + r = p + (q + r))$ la misma idea.*
- *Existe el elemento neutro. En efecto, el polinomio nulo, constante igual a 0 es neutro para la suma.*
- *Existe el inverso aditivo. En efecto, dado p cualquiera, el polinomio $-p$, definido por $\forall x \in \mathbb{K}$, $(-p)(x) = -p(x)$ cumple $\forall x \in \mathbb{K}$, $(p + (-p))(x) = p(x) + (-p)(x) = p(x) - p(x) = 0$, es decir: $p + (-p)$ es el polinomio nulo.*

Por otro lado, el producto · es:

- *Asociativo $((p \cdot q) \cdot r = p \cdot (q \cdot r))$. En efecto, dado $x \in \mathbb{K}$ cualquiera se cumple $((p \cdot q) \cdot r)(x) = (p \cdot q)(x) \cdot r(x) = (p(x) \cdot q(x)) \cdot r(x)$, entonces usando la asociatividad de · en \mathbb{K} , esto es igual a $p(x) \cdot q(x) \cdot r(x) = p(x) \cdot (q(x) \cdot r(x)) = p(x) \cdot (q \cdot r)(x) = (p \cdot (q \cdot r))(x)$, y hemos probado que $(p \cdot q) \cdot r = p \cdot (q \cdot r)$.*
- *Commutativo ($p \cdot q = q \cdot p$). La misma idea...*
- *Existe elemento neutro. El polinomio constante igual a 1 es neutro para la multiplicación de polinomios.*
- *Solo los polinomios constantes no nulos tienen inverso multiplicativo: si p no es constante, entonces $\frac{1}{p}$ no es un polinomio. Aunque esta afirmación requiere una justificación menos dogmática: supongamos que $q = \frac{1}{p}$ es polinomio, o dicho de otra manera, $pq = 1$ entonces $gr(pq) = gr(p) + gr(q) = gr(1) = 0$, por lo tanto $gr(p) = gr(q) = 0$, es decir, ambos polinomios son constantes.*

Por lo tanto, el conjunto de los polinomios NO es cuerpo con las operaciones usuales. □

Para simplificar la notación, se considerará $\alpha\beta$ en lugar de $\alpha \cdot \beta$ cuando α y β pertenecen a un cuerpo.

La siguiente proposición es bastante básica, establece propiedades que ustedes también conocían para los números reales, pero que ahora veremos que se trata de propiedades que vienen dadas de forma “gratuita” partir de la definición de cuerpo.

Propiedad 1 *En cualquier cuerpo se cumplen las siguientes propiedades:*

- a) *El neutro aditivo y el neutro multiplicativo son únicos.*
- b) *Para cada $\alpha \in \mathbb{K}$, su inverso aditivo es único, y por o tanto lo denotaremos como $-\alpha$.*
- c) *El inverso aditivo de $-\alpha$ es α , es decir: $-(-\alpha) = \alpha$.*
- d) $-(\alpha\beta) = (-\alpha)\beta,$
- e) $0\alpha = 0,$
- f) $x + z = y + z \Rightarrow x = y.$
- g) $\alpha\beta = 0 \Rightarrow \alpha = 0 \vee \beta = 0,$
- h) *Si $\gamma \neq 0$ y $\alpha\gamma = \beta\gamma$, entonces $\alpha = \beta$.*

A modo de práctica, desarrollaremos la demostración de algunas partes de la propiedad 1.

Demostración. *Vamos a demostrar algunas de estas propiedades usando sólo la definición de cuerpo.*

- a) *Supongamos que existen dos números que cumplen con la definición de neutro aditivo: A y B , esto es, cumplen:*

$$\forall \alpha \in \mathbb{K}, A + \alpha = \alpha \wedge B + \alpha = \alpha.$$

Demostraremos que $A = B$ usando las propiedades de cuerpo: como la propiedad es válida para todo α , la aplicamos con $\alpha = B$, entonces,

$$A + B = B.$$

Pero si la aplicamos con $\alpha = A$ obtenemos,

$$B + A = A.$$

Juntando ambas y usando commutatividad, concluimos que $A = B$. □

- c) *La afirmación dice que “el inverso aditivo de $-\alpha$ es α ”, en efecto, por la definición de inverso aditivo y la commutatividad tenemos:*

$$0 = \alpha + (-\alpha) = -\alpha + \alpha, \text{ entonces } \alpha \text{ es el inverso de } -\alpha. \quad \square$$

e) Esta requiere trucos interesantes que aprender, partimos de una igualdad dada porque 0 es neutro, y aplicamos operaciones a ambos lados de la ecuación, y las propiedades de cuerpo:

$$\begin{array}{ll}
 0 + 0 = 0 & / \cdot \alpha \\
 (0 + 0)\alpha = 0\alpha & (\text{distributividad}) \\
 0\alpha + 0\alpha = 0\alpha & (\text{neutro aditivo}) \\
 0\alpha + 0\alpha = 0\alpha + 0 & /(-0\alpha)+ \\
 -0\alpha + 0\alpha + 0\alpha = -0\alpha + 0\alpha + 0 & (\text{comutatividad}) \\
 0\alpha + (-0\alpha) + 0\alpha = 0\alpha + (-0\alpha) + 0 & (\text{inverso aditivo}) \\
 0 + 0\alpha = 0 + 0 & (\text{neutro aditivo}) \\
 0\alpha = 0 &
 \end{array}$$

□

Tuplas

Recordamos \mathbb{R}^2 el plano, compuesto por pares ordenados de números reales, definimos $\mathbb{R}^2 = \mathbb{R} \times \mathbb{R}$, como el producto cartesiano del conjunto \mathbb{R} con sí mismo. Sus elementos son *pares ordenados*:

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

Donde (a, b) denota un par de números en que el orden importa, entonces $(a, b) \neq (b, a)$. La notación que distingue un par ordenado de un conjunto es el uso del paréntesis redondo “(”, que por convención matemática internacional se usará para situaciones en que el orden de aparición de los objetos importa; a diferencia del conjunto, que se denota siempre con llaves “{”, donde el orden no tiene relevancia: $\{a, b\} = \{b, a\}$.

Claramente esto es válido cualquiera sea el conjunto considerado: \mathbb{C}^2 será el conjunto de pares ordenados de números complejos, por ejemplo: $(1 + i, 2) \in \mathbb{C}^2$, o para un conjunto que nos demos: Si $B = \{0, 1\}$, entonces

$$B^2 = \{(0, 0), (0, 1), (1, 0), (1, 1)\}$$

El “2” de \mathbb{R}^2 viene de “par” ordenado, ¿cómo es entonces \mathbb{R}^3 ?

En efecto podemos definir \mathbb{R}^3 como el conjunto de las *tripletas ordenadas*: $\mathbb{R}^3 = \mathbb{R} \times \mathbb{R} \times \mathbb{R}$.

$$\mathbb{R}^3 = \{(a, b, c) : a, b, c \in \mathbb{R}\}$$

Por ejemplo $(2, 0, -1) \in \mathbb{Z}^3$. Podemos decir que en general, dado un conjunto cualquiera A , A^3 es el conjunto de todas las secuencias ordenadas de 3 elementos de A .

Las tripletas ordenadas se llaman también “*3-tuplas*”. En general podemos definir una n -tuple como una secuencia ordenada de n elementos. Definimos así:

$$\mathbb{R}^n = \mathbb{R} \times \cdots \text{(n veces)} \cdots \times \mathbb{R} = \{(a_1, \dots, a_n) : \text{para todo } i \in \{1, \dots, n\}, a_i \in \mathbb{R}\}$$

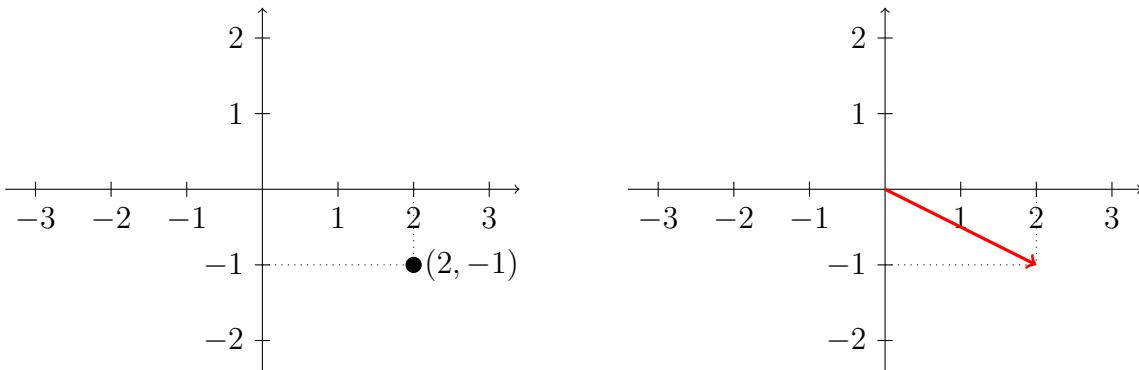
Algunos textos prefieren la notación vertical:

$$\mathbb{R}^n = \left\{ \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_{n-1} \\ a_n \end{pmatrix} : \text{para todo } i \in \{1, \dots, n\}, a_i \in \mathbb{R} \right\}$$

Ambas son válidas, solo es importante notar que cuando escribimos una tupla en forma horizontal sus componentes van siempre separadas por “;”, mientras que cuando la escribimos en forma vertical no.

Las tuplas tienen múltiples aplicaciones, por ejemplo cuando llevamos registro de la temperatura máxima en cada uno de los días de un año, lo que estamos haciendo de definir una 365-tupla, la cual nos servirá luego para hacer modelos matemáticos, comparaciones, promedios, etc...

Siempre será interesante poder imaginar las tuplas, para eso usamos la representación como flecha o como “vector” (aunque esta solo es posible para \mathbb{R}^2 o \mathbb{R}^3). Veamos un ejemplo, el caso \mathbb{R}^2 es más sencillo: representamos la primera coordenada en el eje horizontal, y la segunda en el eje vertical, el elemento $(2, -1)$ quedará representado en el punto cuya proyección vertical caiga en 2 sobre el eje horizontal y su proyección horizontal caiga en -1 en el eje vertical.

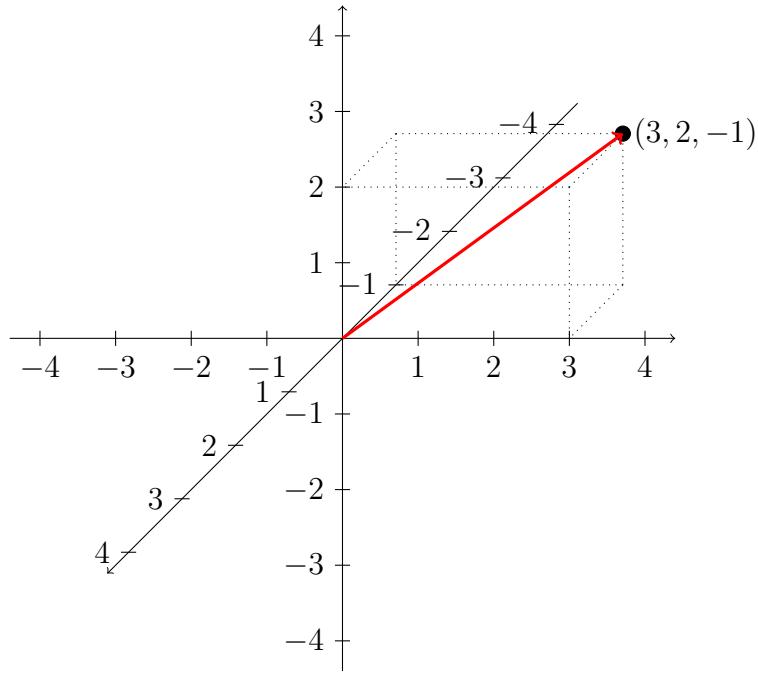


La representación de flecha o vectorial consiste en trazar una flecha partiendo desde el punto $(0,0)$ hasta el punto que representa a la tupla.

Para representar tuplas de \mathbb{R}^3 , necesitamos una astucia: la idea es representar un tercer eje, también horizontal pero que “salga del papel”, de manera perpendicular a los otros dos ejes. Por supuesto esto necesita hacer uso de la imaginación, los arquitectos¹ disponen de distintas maneras de hacer estos dibujos, pero acá solo usaremos lo que GeoGebra llama “*proyección oblicua*”, en la cual el tercer eje se dibuja en forma inclinada y a una escala más pequeña, emulando la perspectiva.

En la siguiente figura representamos la tupla $(3, 2, -1) \in \mathbb{R}^3$.

¹En el video sobre los emojis Ter habla también de estas representaciones.



Lo entretenido comienza cuando comenzamos a hacer operaciones entre tuplas.

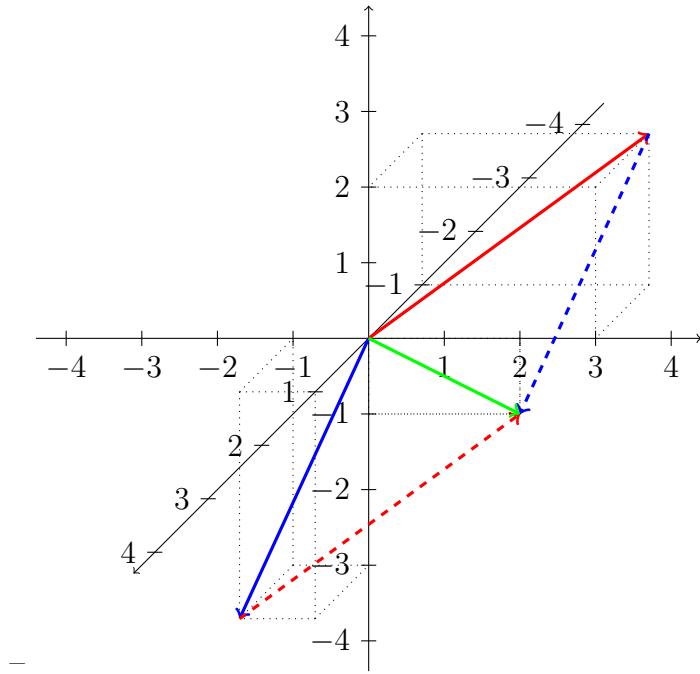
Se define la *suma* de tuplas como sigue. Dados dos vectores cualesquiera $\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}$ en \mathbb{K}^n , se define:

$$\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \oplus \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} x_1 + y_1 \\ x_2 + y_2 \\ \vdots \\ x_n + y_n \end{pmatrix},$$

que consiste en sumar coordenada a coordenada los números que componen la tupla. Vamos a usar el símbolo “ \oplus ” para denotar esta suma y así no confundirla con la suma del cuerpo “ $+$ ”.

Esta suma se torna muy interesante cuando la graficamos en la representación por flechas², pues la suma se obtiene de transladar paralelamente una de las flechas a la cabeza de la otra, y trazar el vector que va del origen hasta la cabeza de la segunda flecha. En la figura sumamos $(3, 2, -1) \oplus (-1, -3, 1) = (2, -1, 0)$

²Algo de esto ya lo habíamos visto cuando representamos a los números complejos como 2-tuplas.

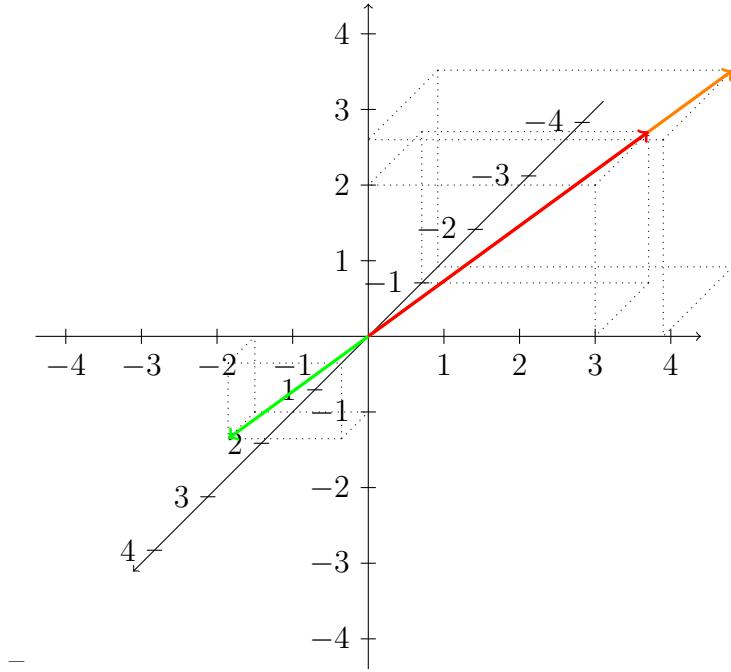


Otra operación que nos va a interesar es la *ponderación por escalar*. Esta consiste en tomar un vector y multiplicar cada una de sus componentes por un mismo número constante (el *escalar*). Usamos el símbolo “ \odot ” para denotar esta operación nueva.

$$\alpha \odot \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} \alpha x_1 \\ \alpha x_2 \\ \vdots \\ \alpha x_n \end{pmatrix}.$$

Gráficamente esta acción lo que hace es “amplificar” el vector (la flecha) por el escalar, si el escalar es positivo y mayor que uno, el vector crece, si el escalar es positivo y menor que 1 el vector se achica, si el escalar es negativo, el vector se invierte, y va en la dirección opuesta.

Veamos el efecto de ponderar la 3-tupla $v = (3, 2, -1)$ por $\alpha = -0,5$ (en verde) y por $\alpha = 1,3$ (en naranja).



Ya que hemos definido operaciones, puede ser interesante averiguar qué propiedades tienen estas operaciones.

Sean $x = (x_1, \dots, x_n)$, $y = (y_1, \dots, y_n)$, $z = (z_1, \dots, z_n) \in \mathbb{K}^n$, y sean $\alpha, \beta \in \mathbb{K}$, donde \mathbb{K} es un cuerpo cualquiera.

- Comutativa ($x \oplus y = y \oplus x$). En efecto,

$$\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \oplus \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} x_1 + y_1 \\ x_2 + y_2 \\ \vdots \\ x_n + y_n \end{pmatrix} = \begin{pmatrix} y_1 + x_1 \\ y_2 + x_2 \\ \vdots \\ y_n + x_n \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} \oplus \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

En cada componente es la suma del cuerpo la que se usa, por lo tanto podemos apoyarnos en la comutatividad de la suma en el cuerpo.

- Asociativa ($((x \oplus y) \oplus z) = x \oplus (y \oplus z)$). ... la misma idea.
- Existe el elemento neutro. En efecto, si sumamos 0 a cada componente no modificamos ningún vector, por lo tanto podemos definir el vector Θ como el nulo:

$$\Theta = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

- Existe el inverso aditivo. En efecto, el inverso aditivo de x es $-x = (-x_1, -x_2, \dots, -x_n)$, pues cumple que al sumarse a x da Θ .

$$\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \oplus \begin{pmatrix} -x_1 \\ -x_2 \\ \vdots \\ -x_n \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

Todas las buenas propiedades.

Para la ponderación no podemos hablar de conmutatividad pues los escalares y las tuplas son objetos diferentes, no se pueden intercambiar.

Pero tal vez hay otras propiedades interesantes:

- $\alpha \odot (\beta \odot x) = (\alpha\beta) \odot x$. En efecto,

$$\alpha \odot \left(\beta \odot \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \right) = \alpha \odot \begin{pmatrix} \beta x_1 \\ \beta x_2 \\ \vdots \\ \beta x_n \end{pmatrix} = \begin{pmatrix} \alpha\beta x_1 \\ \alpha\beta x_2 \\ \vdots \\ \alpha\beta x_n \end{pmatrix} = (\alpha\beta) \odot \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

- $\alpha \odot (x \oplus y) = \alpha \odot x \oplus \alpha \odot y$. Similar...
- $(\alpha + \beta) \odot x = \alpha \odot x \oplus \beta \odot x$. Parecido...
- $1 \odot x = x$. Donde 1 es el neutro multiplicativo del cuerpo \mathbb{K} .
- NO EXISTE inverso multiplicativo, ya que el resultado de la ponderación es siempre un vector, y por lo tanto nunca puede ser igual al neutro del cuerpo, y por la misma razón, el neutro para la ponderación no puede ser un vector.

Vamos a generalizar la noción de vector a algo más amplio, rescatando las operaciones de suma y ponderación y velando por que se preserven las propiedades más importantes que nos gustan. De esto nacerá el concepto de **Espacio Vectorial**, tema central del curso.

Espacios Vectoriales

Hemos visto que las tuplas se pueden *sumar* poniendo “uno delante del otro”, lo cual equivale a sumar coordenada a coordenada sus valores. También se pueden *ponderar* por un *escalar*, amplificando, reduciendo o incluso invirtiendo el sentido del vector.

Vamos a generalizar la noción de vector a algo más amplio, rescatando las operaciones de suma y ponderación y velando para que se preserven las propiedades más importantes que nos gustan.

Definición 1 (Espacio Vectorial) *Sea \mathbb{K} un cuerpo y V un conjunto sobre el cual se definen dos operaciones:*

$$\begin{array}{ll} \oplus : V \times V \rightarrow V & \odot : \mathbb{K} \times V \rightarrow V \\ \text{Adición Binaria Interna} & \text{Ponderación por escalar Binaria Externa} \end{array}$$

La estructura (V, \oplus, \odot) se llama espacio vectorial (e.v.) sobre \mathbb{K} , o simplemente \mathbb{K} -espacio vectorial, si se cumplen las siguientes propiedades (axiomas de espacio vectorial):

1. $\forall u, v, w \in V, u \oplus (v \oplus w) = (u \oplus v) \oplus w$
2. $\forall u, v \in V, u \oplus v = v \oplus u$
3. $\exists \Theta \in V, \forall u \in V, \Theta \oplus u = u$
4. $\forall u \in V, \exists v \in V, u \oplus v = \Theta,$
5. $\forall \alpha, \beta \in \mathbb{K}, \forall u, v \in V,$
 - a) $\alpha \odot (\beta \odot u) = (\alpha\beta) \odot u,$
 - b) $\alpha \odot (u \oplus v) = \alpha \odot u \oplus \alpha \odot v,$
 - c) $(\alpha + \beta) \odot u = \alpha \odot u \oplus \beta \odot u,$
6. $\forall u \in V, 1 \odot u = u.$

Donde 1 denota el neutro multiplicativo del cuerpo $(\mathbb{K}, +, \cdot)$.

Los elementos del espacio vectorial V se llaman *vectores*, a los elementos del cuerpo \mathbb{K} los llamamos *escalares*.

Observar que en la definición, $\alpha\beta$ corresponde a la multiplicación entre escalares (se están multiplicando dos escalares y por tanto el resultado es un escalar) y que $\alpha \odot u$ corresponde a la ponderación por escalar binaria externa entre un escalar y un vector (su resultado es un vector). Por otro lado,

$\alpha + \beta$ es la operación suma de escalares y $u \oplus v$ es la operación suma de vectores. Usaremos esta notación especial mientras nos acostumbramos a distinguir entre estas dos operaciones. Al neutro del espacio vectorial lo denotaremos por Θ para distinguirlo de $0 \in \mathbb{K}$.

Sea \mathbb{K} un cuerpo, ejemplos de espacios vectoriales son:

- El conjunto de las tuplas de un cuerpo, \mathbb{K}^n , con las operaciones usuales de suma entre tuplas y ponderación escalar:

- $x = (x_1, \dots, x_n)$ y $y = (y_1, \dots, y_n) \Rightarrow x \oplus y = (x_1 + y_1, \dots, x_n + y_n)$
- $\alpha \in \mathbb{K}$ y $x = (x_1, \dots, x_n) \Rightarrow \alpha \odot x = (\alpha x_1, \dots, \alpha x_n)$

La clase pasada vimos que se cumplían algunos de los 6 axiomas de espacio vectorial, los demás se verifican de manera similar (¡es un buen ejercicio!).

- El conjunto de los polinomios de grado menor o igual que n , denotado por $\mathcal{P}_n(\mathbb{K})$, es espacio vectorial.

- $p(x) = \sum_{i=0}^n a_i x^i$ y $q(x) = \sum_{i=0}^n b_i x^i \Rightarrow (p \oplus q)(x) = \sum_{i=0}^n (a_i + b_i) x^i$
- $\alpha \in \mathbb{K}$ y $p(x) = \sum_{i=0}^n a_i x^i \Rightarrow (\alpha \odot p)(x) = \sum_{i=0}^n (\alpha a_i) x^i$

- El conjunto de las funciones de un conjunto X sobre \mathbb{K} , denotado por

$$\mathcal{F}(X, \mathbb{K}) = \{f : X \rightarrow \mathbb{K} : f \text{ es función}\},$$

es espacio vectorial sobre \mathbb{K} con las operaciones usuales de suma de funciones y multiplicación por escalar.

- $(f \oplus g)(x) = f(x) + g(x)$, con $x \in X$.
- $\alpha \in \mathbb{K}$, $(\alpha \odot f)(x) = \alpha f(x)$, con $x \in X$.

Notamos que estas son las mismas operaciones que definimos para polinomios, ya que los polinomios son funciones. Muchos de los axiomas de espacio vectorial los demostramos ya en el capítulo de “Funciones Reales” en el curso de Álgebra I.

Tal como en el caso de los cuerpos, podemos demostrar propiedades que se desprenden de los axiomas de Espacio Vectorial. Notar que las propiedades de cuerpo no son las mismas que las propiedades de espacio vectorial, se requiere mucho más sigilo para trabajar con vectores, pues hay muchas cosas que ya no se pueden hacer, por ejemplo NO se puede “dividir” por un vector.

Propiedad 1 En cualquier espacio vectorial $(V, +, \odot)$ sobre un cuerpo $(\mathbb{K}, +, \cdot)$ se cumple que el neutro aditivo es único, al igual que el inverso aditivo. Si denotamos por 0 al neutro aditivo del cuerpo \mathbb{K} y por Θ al neutro aditivo de V ; y denotamos además por $-u$ al inverso aditivo de u , entonces para cualquier $u \in V$ y $\alpha \in \mathbb{K}$ se cumple lo siguiente.

- | | |
|--|--|
| a) $0 \odot u = \Theta$ | e) $(-1) \odot u = -u$ |
| b) $\alpha \odot \Theta = \Theta$ | f) $(-\alpha) \odot u = -(\alpha \odot u)$ |
| c) Si $\alpha \odot u = \Theta$, entonces $\alpha = 0$ o $u = \Theta$. | g) Si $\alpha \neq 0$ y $\alpha \odot u = \alpha \odot v$, entonces $u = v$. |
| d) $-(-u) = u$ | h) Si $u \neq \Theta$ y $\alpha \odot u = \beta \odot u$, entonces $\alpha = \beta$. |

Demostración. Sean $\alpha, \beta \in \mathbb{K}$ y $u \in V$ cualesquiera.

- c) Suponemos que es verdad el lado izquierdo de la implicancia, suponemos que es verdad que $\alpha \odot u = \Theta$. Tenemos que demostrar que ya sea $\alpha = 0$ o bien $u = \Theta$. Si α fuera 0, estaríamos listos, entonces supondremos que $\alpha \neq 0$ y demostrarímos que $u = \Theta$. En ese caso existe el inverso multiplicativo de α , podemos multiplicar la primera hipótesis por éste.

$$\begin{aligned} & \alpha \odot u = \Theta && / \alpha^{-1} \odot (\) \\ \Leftrightarrow & \alpha^{-1} \alpha \odot u = \alpha^{-1} \odot \Theta && (\text{por ítem b}) \\ & 1 \odot u = \Theta && (\text{por axioma 8 de e.v.}) \\ & u = \Theta && \end{aligned}$$

- h) Como la ponderación no tiene inverso, hacemos un truco.

$$\begin{aligned} \alpha \odot u &= \beta \odot u && / + (-\beta \odot u) \\ \alpha \odot u + (-\beta \odot u) &= \beta \odot u + (-\beta \odot u) && (\text{inverso aditivo e ítem f}) \\ \alpha \odot u + (-\beta) \odot u &= \Theta && (\text{distributividad}) \\ (\alpha + (-\beta)) \odot u &= \Theta && (\text{ítem c}) \text{ y } u \neq \Theta \\ \alpha + (-\beta) &= 0 && / + \beta \\ \alpha + (-\beta) + \beta &= 0 + \beta && (\text{neutral e inverso aditivo}) \\ \alpha &= \beta && \end{aligned}$$

Desarrollaremos una vasta teoría a través del curso, la cual será válida para cualquier conjunto que cumpla los axiomas de espacio vectorial. Para poder aplicar la teoría solo necesitaremos verificar que nuestro conjunto cumpla los axiomas de cuerpo, gracias a lo cual toda la teoría se cumplirá. Esta es la gran ventaja de trabajar en forma abstracta sobre un espacio vectorial V inespecífico, desarrollaremos la teoría una sola vez y será válida para todos los espacios vectoriales que se crucen en nuestro camino.

Subespacio Vectorial

Definición 2 (Subespacio Vectorial) *Dado un espacio vectorial (V, \oplus, \odot) sobre un cuerpo \mathbb{K} . Se dice que S es un subespacio vectorial (s.e.v) de V si S es subconjunto de V y, con las mismas operaciones binarias definidas en V , es también un espacio vectorial.*

En otras palabras, (S, \oplus, \odot) satisface:

$$\oplus : S \times S \rightarrow S$$

$$\odot : \mathbb{K} \times S \rightarrow S$$

- | | |
|--|--|
| 1. $(\forall u, v, w \in S) u \oplus (v \oplus w) = (u \oplus v) \oplus w$ | 5. $(\forall \alpha, \beta \in \mathbb{K})(\forall u, v \in S)$ |
| 2. $(\forall u, v \in S) u \oplus v = v \oplus u$ | a. $\alpha \odot (\beta \odot u) = (\alpha\beta) \odot u,$ |
| 3. $(\exists \Theta \in S)(\forall u \in S) \Theta \oplus u = u$ | b. $\alpha \odot (u \oplus v) = \alpha \odot u \oplus \alpha \odot v,$ |
| 4. $(\forall u \in S)(\exists v \in S) u \oplus v = \Theta,$ | c. $(\alpha + \beta) \odot u = \alpha \odot u \oplus \beta \odot u,$ |
| | 6. $(\forall u \in S) 1 \odot u = u.$ |

Ejemplo 1 Para probar que $S = \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3 : xy = 0 \right\} \subset \mathbb{R}^3$ es un subespacio vectorial de \mathbb{R}^3 , las operaciones $\oplus : S \times S \rightarrow S$ y $\odot : \mathbb{R} \times S \rightarrow S$ deben cumplir las propiedades.

En este caso, S no es subespacio vectorial puesto que el resultado de la suma \oplus de dos elementos de S no está en S (no se cumple: $\oplus : S \times S \rightarrow S$), por ejemplo:

$$\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \in S \text{ y } \begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix} \in S, \text{ su suma es } \begin{pmatrix} 1 \\ 1 \\ 3 \end{pmatrix} \notin S$$

Ejemplo 2 $S = \left\{ \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2 : x + y = 0 \right\} \subset \mathbb{R}^2$ es un subespacio vectorial de \mathbb{R}^2 sobre \mathbb{R} . De partida vemos que las operaciones de suma y ponderación donde participan los vectores de S arrojan resultados que también están en S , veamos.

Sea $(x, y), (a, b) \in S$, esto significa que $x + y = 0$ y $a + b = 0$. La suma: $(x + a, y + b)$ cumple $x + a + y + b = x + y + a + b = 0 + 0 = 0$, por lo tanto $(x, y) \oplus (a, b) \in S$.

Asimismo, si $\alpha \in \mathbb{R}$ y $(x, y) \in S$, entonces $\alpha \odot (x, y) = (\alpha x, \alpha y)$ verifica que $\alpha x + \alpha y = \alpha(x + y) = \alpha 0 = 0$, por lo tanto $\alpha \odot (x, y) \in S$.

Por otra parte, los axiomas 1,2,5 y 6 de espacio vectorial se cumplen para todos los elementos de \mathbb{R}^2 , en particular para los elementos de S también se cumplen, no necesitamos verificarlos nuevamente. Solo falta verificar los axiomas 3 y 4. Pero el neutro de \mathbb{R}^2 es $(0, 0) \in S$ ya que $0 + 0 = 0$. Por otra parte si $x + y = 0$, entonces $-x - y = 0$, así $(-x, -y) \in S$, se cumple el axioma 4.

De aquí en adelante, para simplificar notación se ocupará $+$ en lugar de \oplus y se entenderá que $\alpha + \beta$ es la suma en cuerpo cuando α y β son escalares y que $u + v$ es la suma de vectores cuando u y v son vectores. También, se considerará que $\alpha\beta$ es la multiplicación en cuerpo cuando α y β son escalares y que αu es la ponderación cuando u es un vector.

Una forma más fácil de comprobar si un conjunto es un subespacio vectorial es aplicando el siguiente lema.

Lema 1 *Dado V un \mathbb{K} -espacio vectorial y $S \subset V$, se cumple que: S es un s.e.v. de V sobre \mathbb{K} sí y sólo si se cumplen:*

1. $\Theta \in S$ (*el nulo de V está en S*).
2. $\forall v, u \in S, u + v \in S$ (*la suma es cerrada*).
3. $\forall \alpha \in \mathbb{K}, \forall u \in S, \alpha u \in S$ (*la multiplicación por escalar es cerrada*).

Demostración.

(\Rightarrow) Si S es un s.e.v., entonces es un \mathbb{K} -espacio vectorial, y por lo tanto es claro que cumple las tres propiedades pues son parte de estas, en efecto: 1. es el axioma 3 de espacio vectorial; 2. viene de la definición de la suma en S ; y 3. viene de la definición de la ponderación en S sobre \mathbb{K} .

(\Leftarrow) Si S cumple las 3 propiedades, entonces ya sabemos al menos que la suma y la ponderación están definidas dentro de S sobre \mathbb{K} . Por otra parte, como $S \subseteq V$ y V es un \mathbb{K} -espacio vectorial, es claro que se cumplen los axiomas 1,2,5 y 6 de espacio vectorial. La hipótesis incluye al axioma 3 de espacio vectorial, solo falta verificar el axioma 4: la existencia del inverso aditivo dentro de S .

Pero esta se deduce de lo siguiente: llamemos -1 al inverso aditivo de $1 \in \mathbb{K}$, si recordamos las propiedades de espacio vectorial que se desprenden de los axiomas obtenemos que

$$-u = -(1u) = (-1)u,$$

es decir, el inverso aditivo de un vector se obtiene ponderando al vector por -1 . Como la hipótesis nos dice que S es cerrado para la ponderación por cualquier escalar, entonces si $u \in S$, se tiene $-u \in S$. \square

Ejemplo 3 $S = \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3 : x + y = 1 \right\} \subset \mathbb{R}^3$ no es s.e.v porque el nulo de \mathbb{R}^3 no está en S .

Ejemplo 4 El conjunto de las funciones de \mathbb{R} en \mathbb{R} que toman solo valores mayores o iguales a 0 no es un s.e.v del conjunto de todas las funciones, pues si bien contiene a la función nula y es cerrado para la suma de funciones, NO es cerrado para la ponderación de funciones, ya que si pondero una función positiva por -1 , obtengo una función que toma valores negativos.

Ejemplo 5 ■ $\mathcal{P}_5(\mathbb{R})$ es s.e.v de $\mathcal{P}(\mathbb{R})$. *En efecto: $\mathcal{P}_5(\mathbb{R})$ es un espacio vectorial y $\mathcal{P}_5(\mathbb{R}) \subseteq \mathcal{P}(\mathbb{R})$*

- El conjunto de las funciones continuas de \mathbb{R} en \mathbb{R} es s.e.v. del conjunto de las funciones cualesquiera de \mathbb{R} en \mathbb{R} . *En efecto: la suma de funciones continuas es continua; la ponderación por escalar de una función continua es continua; además, la función constante igual a 0 es una función continua.*

Ejemplo 6

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

Cumple las tres propiedades:

1. $(0, 0, 0, 0) \in U$, para mostrarlo basta tomar $x = y = 0 \in \mathbb{C}$, obteniendo $(x, x, 0, y) = (0, 0, 0, 0)$.
2. Si tengo dos vectores arbitrarios en U , digamos $(x, x, 0, y)$ y $(a, a, 0, b)$, y los sumo, obtengo: $(x+a, x+a, 0, y+b)$, el cual también es un vector de U ya que tiene la forma de los elementos de U .
3. Si pondero un elemento $(x, x, 0, y)$ de U , con $x, y \in \mathbb{C}$ cualesquiera, por un escalar cualquiera α en \mathbb{C} , obtengo $(\alpha x, \alpha x, 0, \alpha y)$, que nuevamente se puede probar que está en U tomando $x' = \alpha x \in \mathbb{C}$ y $y' = \alpha y \in \mathbb{C}$, los cuales están en \mathbb{C} ya que $\alpha, x, y \in \mathbb{C}$ y la multiplicación de elementos de \mathbb{C} está en \mathbb{C} .

Ejemplo 7

$$W = \{(x, \bar{x}) \in \mathbb{C}^2 : x \in \mathbb{C}\}$$

Veamos si cumple o no las 3 propiedades:

1. $(0, 0) \in W$, para mostrarlo basta tomar $x = 0 \in \mathbb{C}$, obteniendo $(x, \bar{x}) = (0, \bar{0}) = (0, 0)$.
2. Si tengo dos vectores arbitrarios en W , digamos (x, \bar{x}) y (y, \bar{y}) , y los sumo, obtengo: $(x+y, \bar{x}+\bar{y}) = (x+y, \bar{x+y})$, el cual también es un vector de W ya que tiene la forma de los elementos de W .

3. Si pondro un elemento (x, \bar{x}) de W , con $x \in \mathbb{C}$ cualquiera, por un escalar cualquiera α en \mathbb{C} , obtengo $(\alpha x, \alpha \bar{x}) = (\alpha x, \overline{\alpha \bar{x}})$... no pareciera ser posible llevarlo a la forma (y, \bar{y}) para algún $y \in \mathbb{C}$...

Tomemos un contraejemplo entonces: si tomo $x = 1 + i$, obtengo $(1 + i, 1 - i) \in W$.

Si escojo ahora $\alpha = i$ obtengo

$$\begin{aligned}\alpha(1 + i, 1 - i) &= i(1 + i, 1 - i) \\ &= (i(1 + i), i(1 - i)) \\ &= (i - 1, i + 1) \notin W,\end{aligned}$$

ya que estoy obligada a tomar $y = i - 1$, pero entonces $\bar{y} = -i - 1 \neq i + 1$.

Trabajaremos mucho con los subespacios vectoriales, pues estos aparecen naturalmente en muchos modelos y aplicaciones de la ingeniería, y serán el **objeto central** de este curso.

En \mathbb{R}^2 y de \mathbb{R}^3 , los únicos subespacios vectoriales que encontraremos son:

- El espacio que solo contiene al nulo $\{\Theta\}$ (espacio trivial).
- Las rectas que pasan por el origen.
- Los planos que pasan por el origen.
- V es siempre subespacio de sí mismo.
- \mathbb{R}^2 NO es subespacio de \mathbb{R}^3 pues NO está contenido en este.

Operaciones entre Subespacios

Recuerdo

Dados dos subconjuntos W, U de V , su intersección se define como sigue.

$$W \cap U = \{s \in V \mid s \in W \text{ y } s \in U\}$$

Nos va a interesar particularmente intersectar subespacios y ver qué obtenemos. Por ejemplo la intersección entre dos planos diferentes de \mathbb{R}^3 que pasan por el origen nos da una recta que también pasa por el origen. La intersección entre dos rectas diferentes por el origen, nos da un punto: el origen.

Esta operación de conjuntos la conocemos, introduzcamos otra operación entonces.

Definición 3 Dados dos subconjuntos W, U de V , se define su suma como sigue.

$$W + U = \{w + u \mid w \in W \text{ y } u \in U\}$$

Ejemplo 8 Si tenemos por ejemplo el conjunto $A = \{1, 3, 4\}$ y el conjunto $B = \{1, 2\}$, entonces

$$\begin{aligned} A \cap B &= \{1\} \\ A + B &= \{1+1, 1+2, 3+1, 3+2, 4+1, 4+2\} \\ &= \{2, 3, 4, 5, 6\} \end{aligned}$$

Propiedad 2 Dados s. e. v. W, U de V , resulta que $W \cap U$ y $W + U$ son también s. e. v. de V .

Demuestra. Sean W, U s.e.v. de V sobre \mathbb{K} . Esto significa que ambos cumplen las tres propiedades que caracterizan a los s.e.v. (y todas las demás, por cierto):

1. $\Theta \in U$ y $\Theta \in W$.
2. U y W son cerrados para la suma.
3. U y W son cerrados para la ponderación.

Esta es nuestra hipótesis. Veamos ahora si el conjunto intersección y el conjunto suma cumplen también las 3 propiedades.

$W \cap U$ Analicemos cada una de las 3 propiedades. Recordamos la definición de la intersección:

$$W \cap U = \{s \in V \mid s \in W \text{ y } s \in U\}$$

1. Debemos demostrar que $\Theta \in W \cap U$, lo cual equivale a mostrar que $\Theta \in W$ y $\Theta \in U$; pero esto es cierto de la hipótesis 1.
2. Sean dos vectores $x, y \in W \cap U$, esto equivale a decir que $x, y \in W$ y $x, y \in U$. Debemos demostrar que $x + y \in W \cap U$, es decir, $x + y \in W$ y $x + y \in U$. Pero por la hipótesis 2, y dado que $x, y \in W$, tenemos que $x + y \in W$, por otra parte, y dado que $x, y \in U$, tenemos que $x + y \in U$, así $x + y \in W \cap U$.
3. Sea un vector $x \in W \cap U$, esto equivale a decir que $x \in W$ y $x \in U$, y sea un escalar $\alpha \in \mathbb{K}$. Debemos demostrar que $\alpha x \in W \cap U$, es decir, $\alpha x \in W$ y $\alpha x \in U$. Pero por la hipótesis 3, y dado que $x \in W$, tenemos que $\alpha x \in W$, por otra parte, y dado que $x \in U$, tenemos que $\alpha x \in U$, así $\alpha x \in W \cap U$.

$W + U$ Analicemos cada una de las 3 propiedades. Recordamos la definición de la suma:

$$W + U = \{w + u \mid w \in W \text{ y } u \in U\}$$

1. Debemos demostrar que $\Theta \in W + U$, lo cual equivale a mostrar que existe $w \in W$ y $u \in U$ tales que $\Theta = w + u$; pero si usamos la hipótesis 1, vemos que podemos tomar $w = \Theta$ y $u = \Theta$, se cumple $w + u = \Theta + \Theta = \Theta$, así $\Theta \in W + U$.
2. Sean dos vectores $x, y \in W + U$, esto equivale a decir que existen $w_1, w_2 \in W$ y $u_1, u_2 \in U$ tales que $x = w_1 + u_1$ e $y = w_2 + u_2$. Debemos demostrar que $x + y \in W + U$, es decir, que existen $w_3 \in W$ y $u_3 \in U$ tales que $x + y = w_3 + u_3$. Pero $x + y = w_1 + u_1 + w_2 + u_2 = w_1 + w_2 + u_1 + u_2$, entonces tomando $w_3 = w_1 + w_2$, el cual sabemos por la hipótesis 2 que está en W , y tomando $u_3 = u_1 + u_2$ que también sabemos que está en U , se tiene $x + y = w_3 + u_3 \in W + U$.
3. Sea un vector $x \in W + U$, esto equivale a decir que existen $w \in W$ y $u \in U$ tales que $x = w + u$, y sea un escalar $\alpha \in \mathbb{K}$. Debemos demostrar que $\alpha x \in W + U$, es decir, que existen $w' \in W$ y $u' \in U$ tales que $\alpha x = w' + u'$. Pero $\alpha x = \alpha(w + u) = \alpha w + \alpha u$. Por la hipótesis 3, y dado que $w \in W$, tenemos que $\alpha w \in W$, por otra parte, y dado que $u \in U$, tenemos que $\alpha u \in U$, así $\alpha w + \alpha u \in W + U$, en otras palabras, basta tomar $w' = \alpha w \in W$ y $u' = \alpha u \in U$ para demostrar que $\alpha x = w' + u'$ y por lo tanto $\alpha x \in W + U$. \square

Se observa que la unión de s. e. v. no es en general un s. e. v., es más, la única forma que la unión de dos s. e. v. dé como resultado un s.e.v. es que uno sea subconjunto del otro. Por eso es que la unión de s.e.v. no tiene uso en álgebra lineal.

Por otra parte, el subespacio más pequeño que contiene a $W \cup U$ es justamente $W + U$, tal como establece el siguiente resultado.

Propiedad 3 Si $W \cup U \subseteq S$ y S es subespacio vectorial de V , entonces $W + U \subseteq S$.

Demostración. Nuestras hipótesis son: $W \cup U \subseteq S$ y S es s.e.v. de V .

Debemos demostrar que $W + U \subseteq S$. Tomemos entonces un elemento arbitrario de $W + U$, para esto tomamos un elemento arbitrario $w \in W$ y otro $u \in U$, así $w + u$ es un elemento arbitrario de $W + U$; ahora debemos demostrar que $w + u \in S$.

Pero sabemos que $w, u \in W \cup U \subseteq S$, y como S es s.e.v., es cerrado para la suma, por lo tanto $w + u \in S$. \square

Combinaciones Lineales

Definición 1 (Combinación Lineal) *Dado un espacio vectorial V sobre un cuerpo \mathbb{K} , y dado $\{v_1, v_2, \dots, v_k\}$ un subconjunto de V .*

El vector $w \in V$ es combinación lineal de $\{v_1, v_2, \dots, v_k\}$ si existen $\alpha_1, \alpha_2, \dots, \alpha_k \in \mathbb{K}$ tales que:

$$w = \alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_k v_k$$

Ejemplo 1 Si consideramos el espacio vectorial \mathbb{R}^2 sobre \mathbb{R} y el conjunto $G = \{(2, 1), (-1, 1)\}$, una combinación lineal de este conjunto es el vector $(1, 5)$:

$$(1, 5) = 2(2, 1) + 3(-1, 1)$$

aquí $v_1 = (2, 1)$, $v_2 = (-1, 1)$, $\alpha_1 = 2$ y $\alpha_2 = 3$.

¿Qué otros vectores son c.l. de G ? ¿Es $(0, 0)$ c.l. de G ? ¿es $(2, 1)$ c.l. de G ?

Ejemplo 2 Sea el conjunto $H = \{(0, 1, 3), (1, 0, 1)\} \subseteq \mathbb{R}^3$, considerando a \mathbb{R}^3 como espacio vectorial sobre \mathbb{R} , una combinación lineal de H es $(-2, 1, 1)$:

$$(-2, 1, 1) = 1(0, 1, 3) - 2(1, 0, 1)$$

Sin embargo $(1, 1, 1)$ no es combinación lineal de H , para que lo fuera, necesitaríamos que existieran reales α, β tales que:

$$\begin{aligned} (1, 1, 1) &= \alpha(0, 1, 3) + \beta(1, 0, 1) \\ &= (0, \alpha, 3\alpha) + (\beta, 0, \beta) \\ &= (b, \alpha, 3\alpha + \beta) \end{aligned}$$

Dado que la igualdad de tuplas es una igualdad coordenada a coordenada, esta ecuación vectorial se traduce en 3 ecuaciones:

$$1 = \beta \tag{1}$$

$$1 = \alpha \tag{2}$$

$$1 = 3\alpha + \beta \tag{3}$$

$$1 = 3\alpha + \beta \tag{4}$$

De las dos primeras ecuaciones ya sabemos que $\beta = 1$ y $\alpha = 1$, pero si reemplazamos estos valores en la tercera ecuación obtenemos: $3\alpha + \beta = 4 \neq 1$, por lo tanto las tres ecuaciones no se pueden cumplir simultáneamente, α, β no existen.

¿Es $(1, -1, -2)$ combinación lineal de los elementos de H ?

Subespacio generado por un conjunto de vectores

Definición 2 (Conjunto Generador) *Dado un espacio vectorial V sobre un cuerpo \mathbb{K} , y $\{v_1, v_2, \dots, v_k\}$ un subconjunto de V . El conjunto de todas las combinaciones lineales de $\{v_1, v_2, \dots, v_k\}$ se denomina conjunto generado por $\{v_1, v_2, \dots, v_k\}$ y se denota por:*

$$\langle \{v_1, v_2, \dots, v_k\} \rangle,$$

luego, $u \in \langle \{v_1, v_2, \dots, v_k\} \rangle$ si y solo si existen $\alpha_1, \alpha_2, \dots, \alpha_k \in \mathbb{K}$ tales que:

$$u = \alpha_1 v_1 + \alpha_2 v_2 + \dots + \alpha_k v_k.$$

Dicho en otras palabras:

$$\langle \{v_1, v_2, \dots, v_k\} \rangle = \{\alpha_1 v_1 + \alpha_2 v_2 + \dots + \alpha_k v_k : \alpha_1, \alpha_2, \dots, \alpha_k \in \mathbb{K}\}.$$

Además, si $S = \langle \{v_1, v_2, \dots, v_k\} \rangle$ entonces se dice que S es generado por $\{v_1, v_2, \dots, v_k\}$, y que $\{v_1, v_2, \dots, v_k\}$ genera a S o que es generador de S .

Por último, se define $\langle \phi \rangle = \{\Theta\}$.

Ejemplo 3 No es difícil ver que $\mathbb{R}^2 = \langle \{(1, 0), (0, 1)\} \rangle$, en efecto:

$$(x, y) = x(1, 0) + y(0, 1)$$

Ejemplo 4 Si consideramos el espacio vectorial de los polinomios de grado menor o igual a 4: $\mathcal{P}_3(\mathbb{R})$, y consideramos los polinomios siguientes:

$$p(x) = 1, q(x) = x, r(x) = x^2, s(x) = x^3,$$

vemos fácilmente que $\mathcal{P}_3(\mathbb{R}) = \langle \{p, q, r, s\} \rangle$, en efecto, si tomamos un polinomio arbitrario de $\mathcal{P}_3(\mathbb{R})$: $t(x) = a_0 + a_1 x + a_2 x^2 + a_3 x^3$, tenemos que:

$$t = a_0 p + a_1 q + a_2 r + a_3 s$$

Ejemplo 5 Ya habíamos visto que

$$(1, 5) = 2(2, 1) + 3(-1, 1),$$

por tanto $(1, 5) \in \langle \{(2, 1), (-1, 1)\} \rangle$.

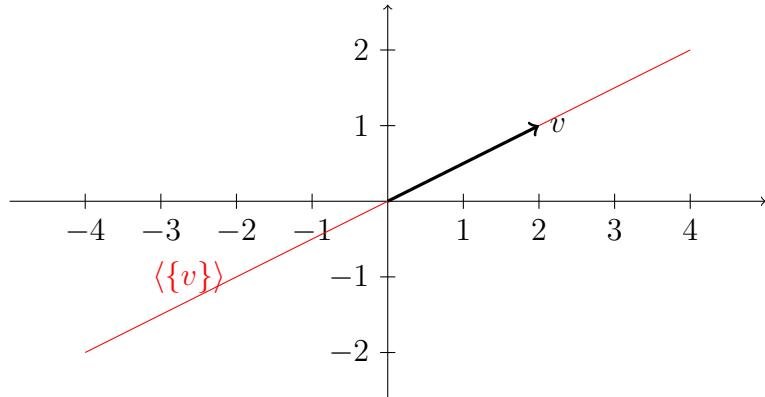
La notación es importante, $\{(2, 1), (-1, 1)\}$ es un conjunto de 2 elementos, en cambio $\langle \{(2, 1), (-1, 1)\} \rangle$ considera todas las combinaciones lineales de esos dos vectores.

- $(1, 5) \in \langle \{(2, 1), (-1, 1)\} \rangle$.
- $(1, 5) \notin \{(2, 1), (-1, 1)\}$.

También es importante no confundir confundir \in con \subseteq (ni ϕ con Θ).

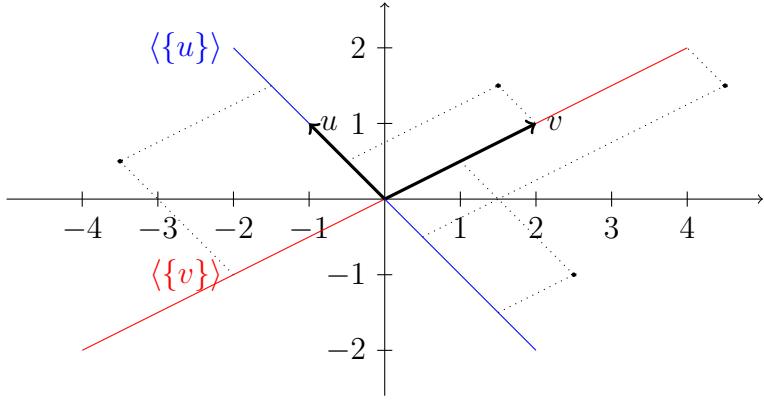
¿Qué otros vectores hay en $\langle \{(2, 1), (-1, 1)\} \rangle$? exploremos esta pregunta de manera gráfica primero para adquirir una intuición en este tema.

Por ejemplo, si $v = (2, 1) \in \mathbb{R}^2$, entonces $\langle \{v\} \rangle$ contendrá todas las ponderaciones de v por todos los escalares que podamos imaginar. El conjunto resultante se grafica así:



Es la recta por el origen que es paralela al vector v y es un subespacio vectorial.

Ahora si $u = (-1, 1)$ entonces $\langle \{u, v\} \rangle$ contendrá todas las ponderaciones de v , todas las ponderaciones de u y todos los vectores que resulten de esas ponderaciones:



Gráficamente vemos que $\langle\{u, v\}\rangle = \mathbb{R}^2$, sin embargo también podemos verlo algebraicamente:

$$\begin{aligned}
 (x, y) \in \langle\{(2, 1), (-1, 1)\}\rangle &\Leftrightarrow \exists \alpha, \beta \in \mathbb{R}, (x, y) = \alpha(2, 1) + \beta(-1, 1) \\
 &\Leftrightarrow \exists \alpha, \beta \in \mathbb{R}, x = 2\alpha - \beta \text{ e } y = \alpha + \beta \\
 &\Leftrightarrow \exists \alpha, \beta \in \mathbb{R}, x = 2\alpha - \beta \text{ e } x + y = 3\alpha \\
 &\Leftrightarrow \exists \alpha, \beta \in \mathbb{R}, x - 2\alpha = -\beta \text{ e } \frac{x+y}{3} = \alpha \\
 &\Leftrightarrow \exists \alpha, \beta \in \mathbb{R}, x - 2\frac{x+y}{3} = -\beta \text{ e } \frac{x+y}{3} = \alpha \\
 &\Leftrightarrow \exists \alpha, \beta \in \mathbb{R}, \frac{x-2y}{3} = -\beta \text{ e } \frac{x+y}{3} = \alpha \\
 &\Leftrightarrow \text{Verdadero}
 \end{aligned}$$

Como resolvimos el sistema, y encontramos que su solución existe para cualquier $x, y \in \mathbb{R}$, se concluye que todo par $(x, y) \in \langle\{(2, 1), (-1, 1)\}\rangle$, es decir $\in \langle\{(2, 1), (-1, 1)\}\rangle = \mathbb{R}^2$. Podemos decir entonces que $\{(2, 1), (-1, 1)\}$ genera a \mathbb{R}^2 .

Propiedad 1 $\langle\{v_1, \dots, v_k\}\rangle$ es un s. e. v.

Demostración. Primero vemos que el conjunto generado contiene al vector nulo (basta tomar todos los “alfas” iguales a 0):

$$\Theta = 0 \cdot v_1 + 0 \cdot v_2 + \cdots + 0 \cdot v_k.$$

Además la suma de dos combinaciones lineales es también una combinación lineal, en efecto, si $p, q \in \langle\{v_1, v_2, \dots, v_k\}\rangle$, entonces p se escribe como $\sum_{i=1}^k \alpha_i v_i$, y q como $\sum_{i=1}^k \beta_i v_i$, por lo tanto

$$p + q = \sum_{i=1}^k \alpha_i v_i + \sum_{i=1}^k \beta_i v_i = \sum_{i=1}^k (\alpha_i + \beta_i) v_i \in \langle\{v_1, \dots, v_k\}\rangle.$$

Por otra parte, la ponderación por escalar de una combinación lineal también es una combinación lineal, en efecto, si $\lambda \in \mathbb{K}$, entonces

$$\lambda p = \lambda \sum_{i=1}^k \alpha_i v_i = \sum_{i=1}^k \lambda \alpha_i v_i \in \langle \{v_1, \dots, v_k\} \rangle.$$

□

Pero podemos decir todavía más, les dejamos esta propiedad adicional para que la saboreen.

Propiedad 2 $\langle \{v_1, \dots, v_k\} \rangle$ coincide con el s. e. v. más pequeño que contiene a $\{v_1, \dots, v_k\}$.

Demostración. Ya vimos que $\langle \{v_1, \dots, v_k\} \rangle$ es un s.e.v.

Además, contiene a $\{v_1, \dots, v_k\}$, para verlo basta tomar igual a 1 el “alfa” que acompaña el vector que queremos, y nulos los demás, por ejemplo: $v_2 = 0 \cdot v_1 + 1 \cdot v_2 + 0 \cdot v_3 + \dots + 0 \cdot v_k$.

Para ver que es el más pequeño de todos los s.e.v que contienen a $\{v_1, \dots, v_k\}$, consideremos otro s.e.v S que también contiene a $\{v_1, \dots, v_k\}$.

Como S es cerrado para la ponderación, entonces contiene a $\alpha_i v_i$, para cualquier i y cualquier $\alpha_i \in \mathbb{K}$.

Como S es cerrado para la suma, entonces contiene $\sum_{i=1}^k \alpha_i v_i$.

Por lo tanto contiene a $\langle \{v_1, \dots, v_k\} \rangle$. □

Independencia Lineal

Definición 3 (Dependencia e Independencia Lineal) Sea V un espacio vectorial sobre un cuerpo \mathbb{K} . El subconjunto $\{v_1, v_2, \dots, v_k\} \subset V$ se dice que es linealmente independiente (l.i.) si y sólo si la única manera de obtener

$$\alpha_1 v_1 + \alpha_2 v_2 + \dots + \alpha_k v_k = \Theta$$

es mediante

$$\alpha_1 = \alpha_2 = \dots = \alpha_k = 0$$

Por el contrario, el conjunto $\{v_1, v_2, \dots, v_k\}$ se dice que es linealmente dependiente (l.d.) cuando NO es linealmente independiente, es decir, si y sólo si existen $\alpha_1, \alpha_2, \dots, \alpha_k \in \mathbb{K}$, no todos iguales a cero, tales que $\alpha_1 v_1 + \alpha_2 v_2 + \dots + \alpha_k v_k = \Theta$.

Ejemplo 6 ■ El conjunto $\{(1, 0, 0), (0, 1, 0), (0, 0, 1)\}$ es linealmente independiente en \mathbb{R}^3 sobre \mathbb{R} , pues si planteo la ecuación

$$\alpha(1, 0, 0) + \beta(0, 1, 0) + \gamma(0, 0, 1) = (0, 0, 0)$$

Obtengo directamente que $\alpha = 0$, $\beta = 0$, y $\gamma = 0$.

- Si nos permitimos denotar un polinomio a través de su fórmula, entonces el siguiente es un conjunto de polinomios y es linealmente independiente $\{1, x, x^2, x^3\}$. En efecto, si planteamos la ecuación

$$\forall x \in \mathbb{R}, \alpha(1) + \beta(x) + \gamma(x^2) + \delta(x^3) = 0$$

vemos que la única manera que se cumpla en todo x es que todos los coeficientes del polinomio resultante sean nulos, es decir que $\alpha = \beta = \gamma = \delta = 0$.

- El conjunto $\{(2, 1), (-1, 1)\}$ también es l.i. en \mathbb{R}^2 sobre \mathbb{R} . Para probarlo planteamos la ecuación de independencia lineal:

$$\alpha(2, 1) + \beta(-1, 1) = (0, 0)$$

Esta equivale a: $2\alpha - \beta = 0$ y $\alpha + \beta = 0$. De la primera obtengo $\beta = 2\alpha$, si reemplazo en la segunda, obtengo $\alpha + 2\alpha = 3\alpha = 0$, de donde $\alpha = 0$ y por consiguiente $\beta = 0$.

El interés principal de la independencia lineal es que nos dice que si un vector es generado por un conjunto l.i., entonces hay una sola manera de obtenerlo como combinación lineal del conjunto.

Propiedad 3 Si $\{v_1, \dots, v_n\}$ es l. i. y $\sum_{i=1}^n \alpha_i v_i = \sum_{i=1}^n \beta_i v_i$, entonces $\alpha_1 = \beta_1, \alpha_2 = \beta_2, \dots, \alpha_n = \beta_n$.

Demostración. Analicemos la hipótesis.

$$\alpha_1 v_1 + \alpha_2 v_2 + \dots + \alpha_n v_n = \beta_1 v_1 + \beta_2 v_2 + \dots + \beta_n v_n$$

Restemos lado derecho a la ecuación y usemos la distributividad.

$$(\alpha_1 - \beta_1)v_1 + (\alpha_2 - \beta_2)v_2 + \dots + (\alpha_n - \beta_n)v_n = \Theta$$

Ahora usamos la hipótesis de independencia lineal del conjunto $\{v_1, \dots, v_n\}$, que nos dice que si tenemos una ecuación como la anterior, entonces los escalares que ponderan a cada vector del conjunto deben ser nulos:

$$\forall i \in \{1, 2, \dots, n\}, \alpha_i - \beta_i = 0$$

Es decir, $\forall i \in \{1, 2, \dots, n\}, \alpha_i = \beta_i$, que era lo que se quería demostrar. \square

Cuando el conjunto es linealmente dependiente la representación de los vectores que genera no es única.

Ejemplo 7 ■ El conjunto $\{(1, 5), (2, 1), (-1, 1)\}$ es linealmente dependiente en \mathbb{R}^2 sobre \mathbb{R} , en efecto:

$$-(1, 5) + 2(2, 1) + 3(-1, 1) = (0, 0),$$

es decir, $(0, 0)$ se puede obtener de dos maneras diferentes, la que acabamos de dar, y la trivial:

$$0(1, 5) + 0(2, 1) + 0(-1, 1) = (0, 0).$$

■ El conjunto $\{(0, 0, 0)\}$ es linealmente dependiente, por la misma razón anterior:

$$\begin{aligned} 3(0, 0, 0) &= (0, 0, 0) \\ 0(0, 0, 0) &= (0, 0, 0) \end{aligned}$$

En general cualquier conjunto de vectores que contenga al nulo será linealmente dependiente ya que permitirá obtener el nulo ponderando por un escalar no nulo.

■ El conjunto $\{(1, -1, 0), (0, 1, -1), (-1, 0, 1)\}$ es linealmente dependiente, de hecho:

$$(-1, 0, 1) = -(0, 1, -1) - (1, -1, 0)$$

Lo cual me permite llegar a que:

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

lo cual es una combinación no trivial que me da el vector nulo.

■ Si $\{u, v\} \subseteq V$ es l.d., entonces existen $\alpha, \beta \in \mathbb{K}$ no ambos nulos tales que $\alpha u + \beta v = \Theta$. Esto implica que ya sea $u = \frac{\beta}{\alpha}v$ o bien $v = \frac{\alpha}{\beta}u$, en cualquier caso uno de los dos vectores es ponderación del otro.

Recíprocamente, si $u = \gamma v$, entonces $\gamma v - u = \Theta$, por lo tanto $\{u, v\}$ es l.d.

En los ejemplos anteriores vimos que en los conjuntos dependientes uno de los vectores era combinación lineal de los demás. Este es siempre el caso, tal como establece el siguiente lema.

Propiedad 4 (Lema de Dependencia Lineal) *Sea $\{v_1, v_2, \dots, v_k\}$ un conjunto linealmente dependiente en V sobre \mathbb{K} . Entonces existe $j \in \{1, 2, \dots, k\}$ que cumple las dos afirmaciones siguientes.*

$$1. v_j \in \langle \{v_1, v_2, \dots, v_{j-1}\} \rangle$$

$$2. \langle \{v_1, v_2, \dots, v_k\} \setminus \{v_j\} \rangle = \langle \{v_1, v_2, \dots, v_k\} \rangle$$

Demostración. Dado que $\{v_1, v_2, \dots, v_k\}$ es linealmente dependiente, han de existir escalares $\alpha_1, \alpha_2, \dots, \alpha_k \in \mathbb{K}$, no todos nulos, tales que

$$\alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_k v_k = \Theta$$

Sea j el índice más grande tal que $\alpha_j \neq 0$, entonces tenemos:

$$\alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_j v_j = \Theta.$$

Despejemos v_j de la ecuación dividiendo por α_j .

$$\frac{\alpha_1}{\alpha_j} v_1 + \frac{\alpha_2}{\alpha_j} v_2 + \cdots + \frac{\alpha_j}{\alpha_j} v_j = \Theta \quad (5)$$

$$v_j = -\frac{\alpha_1}{\alpha_j} v_1 - \frac{\alpha_2}{\alpha_j} v_2 - \cdots - \frac{\alpha_{j-1}}{\alpha_j} v_{j-1} \quad (6)$$

Esto prueba que $v_j \in \langle \{v_1, v_2, \dots, v_{j-1}\} \rangle$.

Por otra parte es claro que $\langle \{v_1, v_2, \dots, v_k\} \setminus \{v_j\} \rangle \subseteq \langle \{v_1, v_2, \dots, v_k\} \rangle$ ya que en el conjunto de la derecha siempre puedo ponderar al vector v_j por 0, dando así lugar a las combinaciones lineales del conjunto de la izquierda.

Para probar la inclusión en el otro sentido tomemos un elemento arbitrario $u \in \langle \{v_1, v_2, \dots, v_k\} \rangle$. Esto significa que existen escalares $\beta_1, \beta_2, \dots, \beta_k \in \mathbb{K}$ tales que

$$u = \beta_1 v_1 + \beta_2 v_2 + \cdots + \beta_k v_k.$$

Usamos (6) para reemplazar a v_j en esta ecuación, por una c.l. de los demás vectores, obtenemos

$$u = \left(\beta_1 - \frac{\beta_j \alpha_1}{\alpha_j} \right) v_1 + \left(\beta_2 - \frac{\beta_j \alpha_2}{\alpha_j} \right) v_2 + \cdots + \left(\beta_{j-1} - \frac{\beta_j \alpha_{j-1}}{\alpha_j} \right) v_{j-1} + \beta_{j+1} v_{j+1} + \cdots + \beta_k v_k.$$

así probamos que u está también en $\langle \{v_1, v_2, \dots, v_k\} \setminus \{v_j\} \rangle$. \square

La primera consecuencia de este lema es que cada conjunto contiene un subconjunto linealmente independiente que genera el mismo espacio.

Ejemplo 8 El conjunto $\{(1, -1, 0), (0, 1, -1), (-1, 0, 1)\}$ es linealmente dependiente, y de

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

podemos despejar $(-1, 0, 1) = -(0, 1, -1) - (1, -1, 0)$; según el lema anterior tenemos que

$$\langle \{(1, -1, 0), (0, 1, -1), (-1, 0, 1)\} \rangle = \langle \{(1, -1, 0), (0, 1, -1)\} \rangle.$$

Notamos que en la primera ecuación podemos despejar otro vector como c.l. de los demás, entonces obtenemos:

$$\langle \{(1, -1, 0), (0, 1, -1), (-1, 0, 1)\} \rangle = \langle \{(1, -1, 0), (0, 1, -1)\} \rangle = \langle \{(1, -1, 0), (-1, 0, 1)\} \rangle.$$

En otras palabras, los 3 vectores que participan de la combinación lineal $(1, -1, 0) + (0, 1, -1) - (-1, 0, 1) = (0, 0, 0)$ ponderados por un escalar no nulo son intercambiables entre sí.

El siguiente teorema será clave para poder definir el concepto de *dimensión* la próxima clase.

Teorema 1 Sea S un s.e.v de V y $B = \{u_1, \dots, u_k\}$ un generador de S . Si $D = \{v_1, \dots, v_m\} \subset S$ con $m > k$ entonces D es l.d.

Demostración. La presente demostración es bastante delicada, la idea consiste en ir reemplazando uno a uno los vectores de B por los de D , para lo cual usaremos el Lema de Dependencia Lineal repetidas veces. La clave para lograr esto está en la hipótesis de que B es generador de S , por lo tanto todos los elementos de S son combinación lineal de los elementos de B , en particular los elementos de D son c.l. de los de B .

Para ayudarnos, razonaremos por reducción al absurdo, vamos a suponer que D es linealmente independiente, y llegaremos a una contradicción. Esta hipótesis implica en particular que D no contiene al vector nulo.

- Paso 1. Por lo dicho anteriormente, v_1 es c.l. de B , por lo tanto $B \cup \{v_1\} = \{v_1, u_1, \dots, u_k\}$ es l.d. Usando el Lema de Independencia Lineal, y el supuesto de que $v_1 \neq \Theta$, sabemos que existe $j \in \{1, \dots, k\}$ tal que $u_j \in \langle \{v_1, u_1, \dots, u_k\} \rangle$, el cual puede ser eliminado de la lista, con lo cual tenemos un nuevo generador de S conformado por v_1 y $k - 1$ vectores de B .
- Paso 2. Podemos repetir el argumento, tomamos ahora v_2 , dado que está en S , es c.l. del conjunto generador definido en el Paso 1, podemos agregarlo a este. Por el Lema de Independencia Lineal sabemos que existe un vector que se puede sacar de este nuevo conjunto por ser c.l. de los anteriores. Tal vector NO es v_1 , ya que estamos suponiendo que v_1 no es c.l. de v_2 . Entonces el vector que podemos sacar es uno de los vectores de B , disminuyendo así su participación en este nuevo generador.
- Paso n . Podemos repetir el argumento previo n veces: tomamos v_n , con $n \leq k$, y lo agregamos al último generador obtenido con el procedimiento antes descrito. Como $v_n \in S$, entonces al agregarlo al generador de S obtenemos un conjunto l.d. Nuevamente el Lema de Independencia Lineal nos sirve para escoger uno de los vectores de este generador que sea c.l. de los primeros vectores de la lista. Dado que estamos suponiendo que D es l.i., el vector que escojamos lo podemos escoger fuera de D . Lo sacamos, y obtenemos un nuevo generador con más elementos de D que el anterior.

Como el procedimiento descrito lo podemos repetir siempre, hasta $n = k$, entonces siempre podemos reemplazar todos los vectores de B por los primeros k vectores de D . Como hemos supuesto que D tiene estrictamente más de k vectores, y todos están en S , queda claro que los $m - k$ vectores de D que no fueron agregados al generador son c.l. de los primeros k , lo cual contradice el supuesto de independencia lineal de D , demostrando que este supuesto es falso, y por lo tanto el teorema es verdadero. \square

Base y coordenadas

Definición 1 (Base) Una base de un sub espacio U es una familia ordenada de vectores $\mathcal{B} = \{v_1, v_2, \dots, v_k\}$ que cumple que

- \mathcal{B} es conjunto generador de U ($U = \langle \mathcal{B} \rangle$) y
- \mathcal{B} es linealmente independiente.

Cuando \mathcal{B} genera a U , tenemos que todo vector $u \in U$ es combinación lineal de \mathcal{B} , y además, si \mathcal{B} es l.i., los escalares que lo producen son únicos, es decir: para todo $u \in U$, existen únicos $\alpha_1, \alpha_2, \dots, \alpha_k \in \mathbb{K}$ tales que:

$$u = \alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_k v_k.$$

La unicidad la conocemos gracias a la propiedad 3 del apunte de la Semana 3. Gracias a esto podemos definir la siguiente noción.

Definición 2 (Vector de coordenadas) Si $\mathcal{B} = \{v_1, v_2, \dots, v_k\}$ es una base de U , y $u \in U$, entonces definimos el vector de coordenadas de u respecto a \mathcal{B} como

$$[u]_{\mathcal{B}} = \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_k \end{pmatrix} \Leftrightarrow u = \alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_k v_k.$$

Ejemplo 1 ■ Una base para el espacio \mathbb{K}^2 sobre \mathbb{K} es $\mathcal{C}_2 = \{(1, 0), (0, 1)\}$ y todo elemento de \mathbb{K}^2 , a saber, $(x, y) \in \mathbb{K}^2$ se escribe como sigue.

$$(x, y) = x(1, 0) + y(0, 1), \text{ entonces: } [(x, y)]_{\mathcal{C}_2} = \begin{pmatrix} x \\ y \end{pmatrix}$$

$$(5, -2) = 5(1, 0) + (-2)(0, 1), \text{ entonces } [(5, -2)]_{\mathcal{C}_2} = \begin{pmatrix} 5 \\ -2 \end{pmatrix}$$

$$(i, 1 - 2i) = i(1, 0) + (1 - 2i)(0, 1), \text{ entonces } [(i, 1 - 2i)]_{\mathcal{C}_2} = \begin{pmatrix} i \\ 1 - 2i \end{pmatrix}$$

Notamos que las coordenadas se toman del cuerpo, por eso es importante siempre especificar sobre qué cuerpo se está, si cambia el cuerpo, seguramente cambiará la base.

- $\tilde{\mathcal{B}}_2 = \{(1, 2), (1, 0)\}$ también es base de \mathbb{K}^2 sobre \mathbb{K} , en este caso, todo vector $(x, y) \in \mathbb{R}^2$ se escribe como:

$$(x, y) = \frac{y}{2}(1, 2) + \left(x - \frac{y}{2}\right)(1, 0), \text{ entonces: } [(x, y)]_{\mathcal{B}_2} = \begin{pmatrix} \frac{y}{2} \\ x - \frac{y}{2} \end{pmatrix}$$

- Los conjuntos $\mathcal{C}_3 = \{(1, 0, 0), (0, 1, 0), (0, 0, 1)\}$ y $\mathcal{C}_4 = \{(1, 0, 0, 0), (0, 1, 0, 0), (0, 0, 1, 0), (0, 0, 0, 1)\}$ son base de \mathbb{K}^3 y \mathbb{K}^4 sobre \mathbb{K} respectivamente. Las bases \mathcal{C}_2 , \mathcal{C}_3 y \mathcal{C}_4 se denominan bases canónicas de \mathbb{K}^2 , \mathbb{K}^3 y \mathbb{K}^4 sobre \mathbb{K} respectivamente, así para cualquier $n \in \mathbb{N}$, la base canónica de \mathbb{K}^n sobre \mathbb{K} está dada por:

$$\mathcal{C}_n = \{(1, 0, 0, \dots, 0), (0, 1, 0, \dots, 0), \dots, (0, 0, 0, \dots, 1)\}.$$

El vector de coordenadas de un vector de \mathbb{K}^n sobre \mathbb{K} respecto a la base canónica de este espacio es siempre el mismo vector.

Ejemplo 2 Sea $S = \{(z_1, z_2, z_3) \in \mathbb{C}^3 : iz_1 - z_2 = 0\}$ un s.e.v de \mathbb{C}^3 sobre \mathbb{C} . Encuentre una base para S .

Para obtener una base de S se debe calcular un conjunto generador y éste debe ser un conjunto linealmente independiente.

Conjunto Generador: Sea $u \in S$, $u = (z_1, z_2, z_3)$ tal que $iz_1 - z_2 = 0$ y por tanto $z_2 = iz_1$, así todo elemento de S es de la forma:

$$u = (z_1, z_2, z_3) = (z_1, iz_1, z_3) = z_1(1, i, 0) + z_3(0, 0, 1)$$

por tanto $S = \langle \{(1, i, 0), (0, 0, 1)\} \rangle$. Sea $\mathcal{B} = \{(1, i, 0), (0, 0, 1)\}$, \mathcal{B} es conjunto generador de S .

Independencia lineal: Sea $\alpha_1, \alpha_2 \in \mathbb{C}$ tales que:

$$\alpha_1(1, i, 0) + \alpha_2(0, 0, 1) = (0, 0, 0).$$

De esta ecuación se obtiene 3: $\alpha_1 = 0$, $i\alpha_1 = 0$ y $\alpha_2 = 0$, es decir, $\alpha_1 = \alpha_2 = 0 \in \mathbb{C}$, es decir, \mathcal{B} es l.i.

Como \mathcal{B} es conjunto generador y l.i entonces $\mathcal{B} = \{(1, i, 0), (0, 0, 1)\}$ es base de S .

Si ahora tomamos, por ejemplo, vector $v = (1 + i, i - 1, 2 + i)$ que pertenece a S (verificar que cumple la ecuación), no es difícil ver que para obtener

$$\alpha_1(1, i, 0) + \alpha_2(0, 0, 1) = (1 + i, i - 1, 2 + i),$$

basta con tomar $\alpha_1 = 1 + i$ y $\alpha_2 = 2 + i$, así:

$$\left[\begin{pmatrix} 1+i \\ i-1 \\ 2+i \end{pmatrix} \right]_{\mathcal{B}} = \begin{pmatrix} 1+i \\ 2+i \end{pmatrix}.$$

En $\mathcal{P}_n(\mathbb{R})$ sobre \mathbb{R} también podemos definir una base canónica, aunque a veces no hay acuerdo entre los libros respecto a cuál es esta. Acá consideraremos que la base canónica de $\mathcal{P}_n(\mathbb{R})$ sobre \mathbb{R} es la siguiente.

$$\mathcal{C}_{\mathcal{P}_n} = \{1, x, x^2, \dots, x^n\}$$

De donde tenemos que el vector de coordenadas del polinomio $p(x) = ax^2 + bx + c$ respecto a $\mathcal{C}_{\mathcal{P}_2}$ es:

$$[ax^2 + bx + c]_{\mathcal{C}_{\mathcal{P}_2}} = \begin{pmatrix} c \\ b \\ a \end{pmatrix}.$$

Esto nos evidencia la importancia de los coeficientes del polinomio, con estos basta para conocerlo completamente, sus coeficientes son sus coordenadas respecto a la base canónica de los polinomios.

Notamos que el orden es importante en una base, pues al cambiar el orden de los vectores de la base, cambia el orden de las coordenadas del vector de coordenadas.

Dimensión

Teorema 1 *Sea S un s.e.v de V . Si $\{v_1, \dots, v_k\}$ y $\{u_1, \dots, u_m\}$ son bases de S entonces $k = m$.*

Demostración. Si $m \neq k$ entonces se ha de tener que $m < k$ o bien $m > k$. Supongamos sin pérdida de generalidad que $m > k$. Entonces podemos usar el Teorema 1 del apunte de la Semana 3 para concluir que $\{u_1, \dots, u_m\}$ es l.d., pero esto contradice la hipótesis, por lo tanto ha de tenerse que $m = k$. \square

Nos dice que *todas* las bases de un espacio vectorial tienen exactamente la misma cantidad de elementos. El tamaño de sus bases caracteriza a un s. e. v. sobre un determinado cuerpo \mathbb{K} , por lo tanto podemos darle nombre a este atributo.

Definición 3 (Dimensión) *Sea S un s.e.v de V con base $\{v_1, \dots, v_k\}$. La dimensión de S , la cual se denota por $\dim(S)$, es la cardinalidad de su base, es decir $\dim(S) = \text{card}(\{v_1, \dots, v_k\}) = k$.*

Ejemplo 3 *Calcule la dimensión de $S = \{(z_1, z_2, z_3) \in \mathbb{C}^3 : iz_1 - z_2 = 0\}$.*

Más arriba se mostró que $\mathcal{B} = \{(1, i, 0), (0, 0, 1)\}$ es una base de S , luego $\dim(S) = \text{card}(\{(1, i, 0), (0, 0, 1)\}) = 2$.

Acá la notación también es importante, escribir $\dim(\{(1, i, 0), (0, 0, 1)\})$ o $\dim(\mathcal{B})$ está equivocado puesto que el concepto de dimensión se aplica a un subespacio vectorial y no a un conjunto cualquiera, $\mathcal{B} = \{(1, i, 0), (0, 0, 1)\}$ no es un sub espacio vectorial.

De igual forma, escribir $\text{card}(S)$ o $\text{card}(\langle \mathcal{B} \rangle)$ no está correcto ya que $S = \langle \mathcal{B} \rangle$ es un s.e.v., por lo tanto su cardinalidad es mucho mayor a su dimensión (en general es infinita).

Observación 1 Así tenemos que $\dim(\mathbb{K}^n) = n$, ya que su base canónica, cuando el cuerpo es \mathbb{K} , tiene n elementos.

Notar que $\{\Theta\}$ no es un conjunto l. i., ya que $1 \cdot \Theta = \Theta$ es una c.l. de $\{\Theta\}$ no trivial que produce Θ . Pero el conjunto vacío \emptyset sí lo es, y ya que $\langle \emptyset \rangle = \{\Theta\}$, tenemos: $\dim(\{\Theta\}) = \text{card}(\emptyset) = 0$: ¡la dimensión del espacio nulo es 0!

También vimos una base canónica para $\mathcal{P}_n(\mathbb{K})$ sobre \mathbb{K} :

$$\mathcal{C}_{\mathcal{P}_n} = \{1, x, x^2, \dots, x^n\},$$

la cual posee $n + 1$ vectores, por lo tanto $\dim(\mathcal{P}_n(\mathbb{K})) = n + 1$.

Pero atención: la dimensión, igual que la base dependen del cuerpo que se está considerando, y aunque la notación “dim” no menciona el cuerpo, este es importante y debe tenerse en cuenta a la hora de concluir.

¿Cómo obtenemos una base para un subespacio vectorial dado? primero supongamos que ya tenemos un sistema generador. Vimos que cuando el sistema no es l. i., entonces hay vectores que son prescindibles, los podemos eliminar del sistema sin pérdida.

Propiedad 1 Si un conjunto finito \mathcal{B} genera a un s.e.v. U , entonces \mathcal{B} contiene a una base de U

Demostración. En efecto, gracias al Lema de Independencia Lineal, sabemos que cuando un conjunto no es l.i., podemos escoger uno de sus vectores para quitárselo sin que pierda su poder generativo; si aún así no obtenemos un conjunto l.i., podemos escoger otro vector y sacarlo; y así sucesivamente, sin que nunca el conjunto deje de dejar de generar a U . En algún momento obtendremos un conjunto l.i. y por lo tanto una base de U . \square

A partir de esta propiedad deducimos que cada vector que resulte ser c.l. de los demás puede ser eliminado, uno a uno, del sistema generador, sin que el espacio generado se vea disminuido. Así, para obtener una base, basta con eliminar estos vectores, uno a uno.

¿Cómo sabemos qué vectores eliminar? Veamos un ejemplo para ilustrarlo.

Se vió que $B = \left\{ \begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix}, \begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} -4 \\ -3 \\ 1 \end{pmatrix} \right\}$ es linealmente dependiente. Resolviendo el sistema de independencia lineal se obtiene que $\alpha_2 = 3\alpha_3$ y $\alpha_1 = -2\alpha_3$ con $\alpha_3 \in \mathbb{R}$.

Podemos escoger $\alpha = 1$, y entonces obtenemos:

$$\begin{pmatrix} -4 \\ -3 \\ 1 \end{pmatrix} = 2 \begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix} - 3 \begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix}$$

Por lo tanto podemos eliminar $\begin{pmatrix} -4 \\ -3 \\ 1 \end{pmatrix}$, ya que los otros dos vectores lo pueden reemplazar sin problemas.

Ejemplo 4 Calcule la dimensión de $H = \langle \{x + 1, x(x + 1), (x - 1)(x + 1)\} \rangle$ subespacio de $\mathcal{P}_2(\mathbb{R})$ sobre \mathbb{R} .

El generador de H que nos están dando es $\mathcal{D} = \{x + 1, x(x + 1), (x - 1)(x + 1)\}$, veamos primero si es l.i. o no. Planteamos la ecuación de independencia lineal:

$$\forall x \in \mathbb{R}, a(x + 1) + bx(x + 1) + c(x - 1)(x + 1) = 0.$$

desarrollamos los productos y la suma para obtener los coeficientes del polinomio resultante.

$$\begin{aligned} a(x + 1) + bx(x + 1) + c(x - 1)(x + 1) &= ax + a + bx^2 + bx + cx^2 - c \\ &= (b + c)x^2 + (a + b)x + a - c \end{aligned}$$

Igualando cada coeficiente a 0 obtenemos: $a - c = 0$, $a + b = 0$ y $b + c = 0$. De la primera ecuación obtenemos $a = c$, reemplazamos esto en las demás ecuaciones para eliminar a , y obtenemos: $c + b = 0$ $b + c = 0$, vemos que las dos ecuaciones quedan iguales, por lo tanto concluimos que $b = -c$ y no podemos seguir trabajando.

Por lo tanto una solución al sistema es $a = 1, c = 1, b = -1$, lo que demuestra que \mathcal{D} no es l.i.

Esta solución me dice que

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

es decir que $x(x + 1) = (x + 1) + (x - 1)(x + 1)$, podemos entonces eliminar este vector del conjunto generador, para obtener $\mathcal{D}' = \{x + 1, (x - 1)(x + 1)\}$, que gracias al Lema de Independencia Lineal sigue generando a H .

Además, vemos que \mathcal{D}' es l.i. ya que ninguno de estos polinomio es ponderación del otro por escalar, entonces \mathcal{D}' es base de H , y por lo tanto $\dim(H) = 2$.

Teorema 2 (Completación) Si $\{u_1, \dots, u_k\} \subset V$ es l. i. y $u \notin \langle \{u_1, \dots, u_k\} \rangle$, entonces $\{u, u_1, \dots, u_k\}$ es también l. i.

Demostración. Suponemos que no es así, y tomamos una combinación lineal de $\langle \{u, u_1, \dots, u_k\} \rangle$ que da Θ , con los escalares a, a_1, \dots, a_k .

$$au + \sum_{i=1}^k a_i u_i = \Theta, \quad \text{despejando } au : \quad au = - \sum_{i=1}^k a_i u_i$$

Caso 1 ($a = 0$): entonces $\sum_{i=1}^k a_i u_i = \Theta$, pero como los vectores participantes de esta suma son l.i., entonces $a_1 = \dots = a_k = 0$.

Caso 2 ($a \neq 0$): dividimos por a en ambos lados de la ecuación y obtenemos:

$$u = \sum_{i=1}^k \left(\frac{-a_i}{a} \right) u_i,$$

de donde se deduce que $u \in \langle\{u_1, \dots, u_k\}\rangle$, lo cual es una contradicción. \square

¿Cómo obtenemos una base de un s.e.v. U si solo tenemos un conjunto l.i. dentro de éste, pero que no es generador?

Si el conjunto $\{u_1, \dots, u_k\}$ es l.i., y está en U pero no lo genera, significa que existe un vector u en U que no es c.l. de $\{u_1, \dots, u_k\}$. Podemos agregarlo al conjunto, y por el Teorema 2, el nuevo conjunto sigue siendo l.i. y estando en U .

Si aún con haber agregado este vector, no logramos generar todo U , es porque aún quedan vectores que agregar: buscamos uno, lo agregamos y obtenemos un generador más grande; y así sucesivamente continuamos hasta obtener un generador de U .

El proceso antes descrito termina cuando hayamos obtenido tantos vectores l.i. dentro de U como sea su dimensión... ¿pero es esta finita?¹

Si U es s.e.v. de V y V tiene dimensión finita, entonces U también ha de tenerla, de hecho, por el Teorema 1 del apunte de la Semana 3, sabemos que V no puede contener ningún conjunto l.i. que tenga más de $\dim(V)$ elementos, por lo tanto nuestro proceso termina antes de $\dim(V)$ pasos.

Este argumento muestra además que $\dim(U) \leq \dim(V)$

Ejemplo 5 Sea $\mathcal{D} = \{(2, 3, 0, 1), (0, 0, 1, 0)\} \subseteq W = \{(a, b, c, d) \in \mathbb{R}^4 : a + d = b\}$. En efecto, vemos que $2 + 1 = 3$ y $0 + 0 = 0$.

Es fácil ver que $\langle \mathcal{D} \rangle \subseteq W$, ya que W es un s.e.v. y por lo tanto es cerrado para combinaciones lineales. Por otra parte, \mathcal{D} es l.i., ya que ninguno de sus dos vectores es ponderación del otro (cuando hay solo dos vectores, es muy fácil chequear independencia lineal).

Veamos si es que $\langle \mathcal{D} \rangle = W$ o no. Para ello calculamos estudiemos $\langle \mathcal{D} \rangle$.

$$\begin{aligned}\langle \mathcal{D} \rangle &= \{((2a, 3a, b, a) : a, b \in \mathbb{R}\} \\ &= \{(x, y, z, t) \in \mathbb{R}^4 : x = 2t \text{ e } y = 3t, z \in \mathbb{R} \text{ libre}\}\end{aligned}$$

No es difícil encontrar un vector en $W \setminus \langle \mathcal{D} \rangle$, por ejemplo $(1, 1, 0, 0)$ cumple la ecuación de W , pero no las ecuaciones de $\langle \mathcal{D} \rangle$.

Entonces $\mathcal{D} \cup \{(1, 1, 0, 0)\}$ es l.i., gracias al Teorema 2. $\langle \mathcal{D} \cup \{(1, 1, 0, 0)\} \rangle$ es un subespacio más grande que $\langle \mathcal{D} \rangle$, y sigue contenido en W , ¿será igual a W ?

Propiedad 2 Si U y W son s.e.v. de V sobre \mathbb{K} y $U \subseteq W$, entonces

$$\dim(U) \leq \dim(W). \tag{1}$$

Si además $\dim(U) = \dim(W)$, entonces

$$U = W. \tag{2}$$

¹Hay espacios vectoriales de dimensión “infinita”, pero de eso no hablaremos en este curso...

Demostración. Ya demostramos (1) más arriba.

Demostremos ahora (2): supongamos por el contrario, que $U \not\subseteq W$. Esto significa que existe $w \in W \setminus U$. Si U tiene una base \mathcal{B} de tamaño $k = \dim(U) = \dim(W)$, entonces $w \notin \langle \mathcal{B} \rangle = U$. El teorema 2 de la clase de hoy nos dice que entonces $\{w\} \cup \mathcal{B}$ es l.i. Por lo tanto $\{w\} \cup \mathcal{B}$ es base de $\langle \{w\} \cup \mathcal{B} \rangle$, y como $\{w\} \cup \mathcal{B} \subset W$, $\langle \{w\} \cup \mathcal{B} \rangle$ sería un s.e.v. de W de dimensión $k+1 > k = \dim(W)$, lo cual contradice (1) recién demostrado. \square

Ejemplo 6 Retomando el ejemplo 5. Tenemos que $\mathcal{D} \cup \{(1, 1, 0, 0)\} \subseteq W$, y entonces $\langle \mathcal{D} \cup \{(1, 1, 0, 0)\} \rangle \subseteq W$. Como $\mathcal{D} \cup \{(1, 1, 0, 0)\}$ es l.i., entonces $\dim(\langle \mathcal{D} \cup \{(1, 1, 0, 0)\} \rangle) = |\mathcal{D} \cup \{(1, 1, 0, 0)\}| = 3$, por lo tanto $3 \leq \dim(W)$.

Por otra parte, es claro que $W \neq \mathbb{R}^4$, ya que por ejemplo $(1, 0, 0, 0) \notin W$. Por lo tanto $\dim(W) < 4$.

En conclusión $\dim(W) = 3$. De donde concluimos que $W = \langle \mathcal{D} \cup \{(1, 1, 0, 0)\} \rangle$, gracias a la propiedad 2 recién demostrada.

Con todo lo trabajado se pueden deducir las siguientes propiedades (pensar por qué).

Propiedad 3 *Dado un e.v. V de dimensión finita, se cumple lo siguiente.*

- a) *El conjunto l. i. más grande de V , tiene $\dim(V)$ elementos.*
- b) *Todo conjunto l. i. de V con $\dim(V)$ elementos, es base de V .*
- c) *El conjunto generador de V más pequeño tiene $\dim(V)$ elementos.*
- d) *Todo conjunto generador de V con $\dim(V)$ elementos es base de V .*

Aplicemos lo aprendido.

Ejemplo 7 Sea $\mathcal{B} = \{x^2 + 2x\} \subseteq \mathcal{P}_2(\mathbb{R})$ sobre \mathbb{R} , complete \mathcal{B} hasta obtener una base de $\mathcal{P}_2(\mathbb{R})$.

Solución: La dimensión de $\mathcal{P}_2(\mathbb{R})$ sobre \mathbb{R} es 3, luego todo conjunto linealmente independiente que tenga 3 vectores será base de $\mathcal{P}_2(\mathbb{R})$ sobre \mathbb{R} , por lo tanto a \mathcal{B} hay que agregarle dos vectores: $p(x)$ y $q(x)$, de manera que el conjunto resultante $\mathcal{B} \cup \{p(x), q(x)\}$ sea l.i.

Ya se sabe cuántos vectores agregar, ahora la pregunta es: ¿cuáles son esos vectores?.

El teorema 2 dice que si a un conjunto linealmente independiente se le agrega un vector que no es combinación lineal de los elementos de éste, entonces el nuevo conjunto es l.i. Por tanto, para este caso, $p(x)$ puede ser $p(x) = x^2$ ya que x^2 no es ponderación de $x^2 + 2x$, luego el nuevo conjunto $\{x^2 + 2x, x^2\}$ es l.i.

El segundo vector a agregar puede ser $q(x) = 1$ puesto que $1 \notin \langle \{x^2 + 2x, x^2\} \rangle$, esto es evidente del hecho que toda c.l. de estos vectores tendrá el término libre nulo. Luego $\{x^2 + 2x, x^2, 1\}$ es l.i y tiene 3 elementos, en consecuencia, es base de $\mathcal{P}_2(\mathbb{R})$ sobre \mathbb{R} .

Teorema 3 (Grassmann, XIX) *Dados dos s.e.v. U, W de dimensión finita de un espacio vectorial V sobre \mathbb{K} , se cumple:*

$$\dim(U + W) = \dim(U) + \dim(W) - \dim(U \cap W).$$

Demuestra. Espero que la disfruten la demostración como espectadores, es muy elaborada, y hace uso de los teoremas que hemos visto. Sin embargo no tiene utilidad en sí para los ejercicios que vemos en este curso (lo que sí tiene utilidad es el teorema).

Antes que nada, llamemos $k = \dim(U \cap W)$, $m = \dim(U)$ y $n = \dim(W)$.

Tomemos una base de $U \cap W$, digamos $\mathcal{B} = \{u_1, u_2, \dots, u_k\}$. Usando el teorema de completación, y dado que $(U \cap W) \subseteq U$, completemos esta base hasta obtener una base de U , digamos que hemos agregado un conjunto de vectores $\mathcal{C} = \{u_{k+1}, \dots, u_m\}$ a \mathcal{B} , es decir, que la base de U es: $\mathcal{B} \sqcup \mathcal{C}$.

Ahora, dado que $(U \cap W) \subseteq W$, completemos \mathcal{B} para obtener una base de W , digamos que agregamos el conjunto de vectores $\mathcal{D} = \{w_1, w_2, \dots, w_{n-k}\}$, es decir que la base de W es: $\mathcal{B} \sqcup \mathcal{D}$.

Entonces, por la propiedad anterior, tenemos un generador para $U + W$:

$$U + W = \langle (\mathcal{B} \sqcup \mathcal{C}) \cup (\mathcal{B} \sqcup \mathcal{D}) \rangle = \langle \mathcal{B} \sqcup \mathcal{C} \sqcup \mathcal{D} \rangle = \langle \{u_1, u_2, \dots, u_m, w_1, \dots, w_{n-k}\} \rangle.$$

este generador tiene justamente $m + n - k = \dim(U) + \dim(W) - \dim(U \cap W)$ vectores, si fuera l.i., habríamos entonces demostrado el Teorema de Grassmann. Estudiemos entonces su independencia lineal. Como hipótesis tenemos que \mathcal{B} , \mathcal{C} y \mathcal{D} son l.i. Sean entonces escalares $\alpha_1, \dots, \alpha_m, \beta_1, \dots, \beta_{n-k}$ tales que:

$$\sum_{i=1}^m \alpha_i u_i + \sum_{j=1}^{n-k} \beta_j w_j = \Theta.$$

Si restamos la segunda sumatoria nos queda:

$$\sum_{i=1}^m \alpha_i u_i = - \sum_{j=1}^{n-k} \beta_j w_j \in W.$$

Pero esto indicaría que $\sum_{j=1}^{n-k} \beta_j w_j$ pertenece a $U \cap W$, entonces debería ser combinación lineal de los vectores de $\mathcal{B} = \{u_1, u_2, \dots, u_k\}$, pero estos no aparecen en la suma, y por su parte los vectores de $\mathcal{D} = \{w_1, \dots, w_{n-k}\}$ los escogimos justamente fuera de $\langle \mathcal{B} \rangle$, entonces la única manera de que esto sea posible es que sea el vector nulo:

$$\sum_{j=1}^{n-k} \beta_j w_j = \Theta,$$

y dado que \mathcal{D} es l.i., esto implica que $\beta_1 = \beta_2 = \dots = \beta_{n-k} = 0$. Pero esto nos deja con la siguiente igualdad:

$$\sum_{i=1}^m \alpha_i u_i = \Theta,$$

y esta es una c.l. de $\mathcal{B} \sqcup \mathcal{C}$ que es base de U , entonces $\alpha_1 = \alpha_2 = \dots = \alpha_m = 0$. Hemos demostrado que $\mathcal{B} \sqcup \mathcal{C} \sqcup \mathcal{D}$ es l.i. \square

Definición 4 Dos s. e. v. W , U de V se dicen en suma directa si $W \cap U = \{\Theta\}$, en tal caso denotamos $W \oplus U$ en lugar de $W + U$.

La suma directa es de particular importancia debido a la siguiente propiedad, la cual dice que cuando la suma de dos espacios es del tipo “directa”, entonces los vectores del espacio suma, se descomponen de manera única como suma de vectores de los espacios.

Propiedad 4 Si W y U son dos s. e. v. de V que están en suma directa, entonces se cumple:

$$(\forall v \in W \oplus U)(\exists! w \in W)(\exists! u \in U) v = w + u.$$

Demostración. Nuestra hipótesis principal es que $W \cap U = \{\Theta\}$.

Sea $v \in W \oplus U$, supongamos que hay dos maneras diferentes de escribir v como suma de un elemento de W más uno de U , es decir, supongamos que existen $w_1, w_2 \in W$ y $u_1, u_2 \in U$ tales que:

$$v = w_1 + u_1 \quad \text{y} \quad v = w_2 + u_2.$$

Entonces podemos afirmar que:

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

Sumamos $-u_1 - w_2$ a ambos lados:

$$w_1 - w_2 = u_2 - u_1 \in U.$$

Pero esto significa que $w_1 - w_2 \in U \cap W = \{\Theta\}$, por lo tanto $w_1 - w_2 = \Theta$, es decir $w_1 = w_2$. Además, por la igualdad, también tenemos que $u_2 - u_1 = \Theta$, por lo tanto $u_1 = u_2$.

En otras palabras, las dos descomposiciones no eran dos si no una sola, lo que prueba la propiedad. \square

Última observación: si U y W están en suma directa, entonces

$$\dim(U \oplus W) = \dim(U) + \dim(W).$$

Transformaciones Lineales

El estudio de funciones juega un rol fundamental en la ciencia y en la ingeniería. Una función es una operación que transforma un objeto en otro. Estudiaremos funciones entre espacios vectoriales, es decir funciones que a un vector le asocian otro vector (eventualmente de un espacio vectorial distinto).

En este capítulo nos centraremos en una clase particular de funciones, llamadas *lineales*. Veremos que son simples y que hay una vasta teoría que nos permite comprenderlas profundamente. No es este el caso de las funciones en general, por esto mismo, las funciones lineales juegan un rol fundamental en la ingeniería, pues son prácticamente las únicas que podemos representar totalmente en un computador, y las únicas que podemos “resolver” de manera directa. Verán ustedes dentro de unos años más que, este es uno de los temas que más usarán en sus carreras profesionales.

Definición 1. *Dados dos espacios vectoriales V y W sobre un mismo cuerpo \mathbb{K} , decimos que la función $T : V \rightarrow W$ es una Transformación Lineal si cumple las siguientes propiedades.*

1. $(\forall u, v \in V) T(u + v) = T(u) + T(v)$.
2. $(\forall v \in V)(\forall \alpha \in \mathbb{K}) T(\alpha v) = \alpha T(v)$.

En otras palabras, una transformación lineal respeta la suma y respeta la ponderación, y por lo tanto preserva los subespacios vectoriales, preserva las líneas rectas.

Ejemplo 1. Las funciones lineales más básicas y que merecen un nombre propio son las siguientes.

- **La función constante igual a Θ .** Si llamamos Θ_W al vector nulo de W , podemos definir: $\mathcal{O} : V \rightarrow W$ como $\mathcal{O}(v) = \Theta_W$, para todo $v \in V$.

Es lineal ya que si $u, v \in V$, entonces $\mathcal{O}(u + v) = \Theta_W = \Theta_W + \Theta_W = \mathcal{O}(u) + \mathcal{O}(v)$.

Por su parte si $\alpha \in \mathbb{K}$, entonces $\mathcal{O}(\alpha v) = \Theta_W = \alpha \Theta_W = \alpha \mathcal{O}(v)$. Respeta la suma y la ponderación.

- **La función identidad.** $id : V \rightarrow V$ definida por $id(v) = v$ es lineal. En efecto.

$$id(u + v) = u + v = id(u) + id(v)$$

$$id(\alpha v) = \alpha v = \alpha id(v)$$

Ejemplo 2. La funciones que estudiamos, transforman un vector en otro vector. Consideremos por ejemplo la transformación $L : \mathbb{R}^2 \rightarrow \mathbb{R}$ dada por $L(x, y) = 3x - 2y$. Lo primero que debemos hacer para entenderla es calcularla en alguna parte, veamos algunos ejemplos.

$$L(1, -1) = 3 \cdot 1 - 2 \cdot (-1) = 3 + 2 = 5$$

$$L(0, 10) = 3 \cdot 0 - 2 \cdot 10 = 0 - 20 = -20$$

$$L(3a, a + b) = 3 \cdot 3a - 2(a + b) = 9a - 2a - 2b = 7a - 2b$$

Resulta que L es lineal, ya que cumple la definición, si $u = (x, y)$ y $v = (a, b)$ entonces

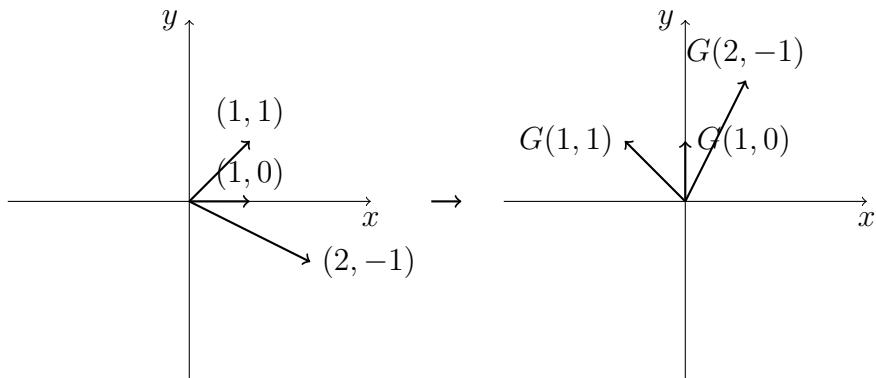
$$L(u+v) = L(x+a, y+b) = 3(x+a) - 2(y+b) = 3x + 3a - 2y - 2b = 3x - 2y + 3a - 2b = L(u) + L(v)$$

y si $\alpha \in \mathbb{R}$

$$L(\alpha u) = L(\alpha x, \alpha y) = 3\alpha x - 2\alpha y = \alpha(3x - 2y) = \alpha L(u).$$

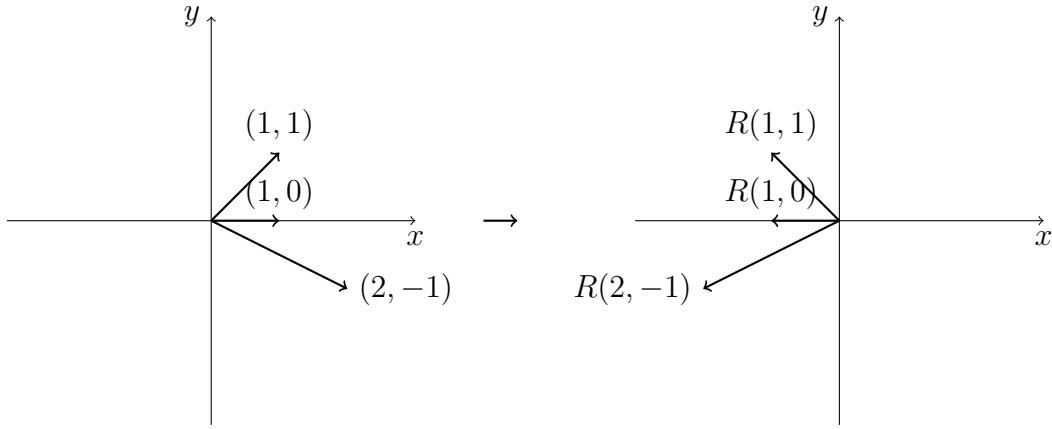
Ejemplos célebres de transformaciones lineales de \mathbb{R}^2 en \mathbb{R}^2 , son las reflexiones, y rotaciones.

Ejemplo 3. Rotación en 90 grados La función $G : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ definida por $G(x, y) = (-y, x)$ rota cada vector de \mathbb{R}^2 en 90 grados, para verlo, basta calcular y dibujar.



Vemos que esta transformación respeta todo, la magnitud, los ángulos relativos entre los vectores, también la suma y la ponderación.

Reflexión respecto al eje Y La función $R : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ definida por $R(x, y) = (-x, y)$ refleja cada vector de \mathbb{R}^2 , como si en el eje Y hubiera un espejo, para verlo, basta calcular y dibujar.



¿Cómo haría usted para definir la reflexión respecto a otras rectas, por ejemplo respecto al eje X? ¿Cómo haría usted para definir una rotación en otro ángulo, por ejemplo 180 grados, o en 45 grados?

Ejemplo 4. Otros ejemplos clásicos de funciones lineales son los siguientes.

- $D : \mathcal{P}(\mathbb{R}) \rightarrow \mathcal{P}(\mathbb{R})$, definida por $D(p) = p'$, donde p' es la *derivada* de p .
Es lineal, en efecto:

$$D(p+q) = (p+q)' = p' + q' = D(p) + D(q), \text{ y}$$

$$D(\lambda p) = (\lambda p)' = \lambda p' = \lambda D(p).$$

- $I : \mathcal{P}(\mathbb{R}) \rightarrow \mathbb{R}$, definida por $I(p) = \int_0^1 p(x) dx \in \mathbb{R}$.

Es lineal, en efecto:

$$I(p+q) = \int_0^1 (p+q)(x) dx = \int_0^1 p(x) + q(x) dx = \int_0^1 p(x) dx + \int_0^1 q(x) dx = I(p) + I(q), \text{ y}$$

$$I(\lambda p) = \int_0^1 (\lambda p)(x) dx = \int_0^1 \lambda p(x) dx = \lambda \int_0^1 p(x) dx = \lambda I(p).$$

- $A : \mathbb{R}^2 \rightarrow \mathbb{R}^3$, $A(x, y) = (2x - y, x + 2y, y)$.

Es lineal, y la demostración se parece mucho a la del Ejemplo 2, así que mejor veamos un pequeño teorema que sirva para todas las funciones de este tipo de una sola vez.

Propiedad 1. Considerando a \mathbb{K}^m y \mathbb{K}^n como espacios vectoriales sobre \mathbb{K} . Si $T : \mathbb{K}^m \rightarrow \mathbb{K}^n$ es una función definida como sigue:

$$T(x_1, x_2, \dots, x_m) = (a_{11}x_1 + a_{12}x_2 + \dots + a_{1m}x_m, a_{21}x_1 + \dots + a_{2m}x_m, \dots, a_{n1}x_1 + \dots + a_{nm}x_m),$$

donde $a_{ij} \in \mathbb{K}$, con $i \in \{1, \dots, n\}$ y $j \in \{1, \dots, m\}$, son nm constantes fijas, entonces T es lineal.

Demostración. Tomamos dos vectores arbitrarios de \mathbb{K}^m : $u = (x_1, \dots, x_m)$ y $v = (y_1, \dots, y_m)$, tomamos también un escalar cualquiera $\alpha \in \mathbb{K}$, calculamos:

$$\begin{aligned}
T(u+v) &= T(x_1 + y_1, x_2 + y_2, \dots, x_m + y_m) \\
&= (a_{11}(x_1 + y_1) + a_{12}(x_2 + y_2) + \dots + a_{1m}(x_m + y_m), \dots, a_{n1}(x_1 + y_1) + \dots + a_{nm}(x_m + y_m)) \\
&= (a_{11}x_1 + a_{11}y_1 + a_{12}x_2 + a_{12}y_2 + \dots + a_{1m}x_m + a_{1m}y_m, \dots, a_{n1}x_1 + a_{n1}y_1 + \dots + a_{nm}x_m + a_{nm}y_m) \\
&= (a_{11}x_1 + a_{12}x_2 + \dots + a_{1m}x_m, a_{21}x_1 + \dots + a_{2m}x_m, \dots, a_{n1}x_1 + \dots + a_{nm}x_m) \\
&\quad + (a_{11}y_1 + a_{12}y_2 + \dots + a_{1m}y_m, a_{21}y_1 + \dots + a_{2m}y_m, \dots, a_{n1}y_1 + \dots + a_{nm}y_m) \\
&= T(u) + T(v)
\end{aligned}$$

Respeta la suma, veamos ahora qué pasa con la ponderación.

$$\begin{aligned}
T(\alpha u) &= (a_{11}\alpha x_1 + a_{12}\alpha x_2 + \dots + a_{1m}\alpha x_m, a_{21}\alpha x_1 + \dots + a_{2m}\alpha x_m, \dots, a_{n1}\alpha x_1 + \dots + a_{nm}\alpha x_m) \\
&= \alpha(a_{11}x_1 + a_{12}x_2 + \dots + a_{1m}x_m, a_{21}x_1 + \dots + a_{2m}x_m, \dots, a_{n1}x_1 + \dots + a_{nm}x_m) \\
&= \alpha T(u)
\end{aligned}$$

□

¿Qué funciones NO son lineales? Bueno, si $V = W = \mathbb{R}$, un ejemplo clásico de función que no es lineal es sen , ya que $\text{sen}(x+y) \neq \text{sen}(x) + \text{sen}(y)$. ¿Qué otros ejemplos conocen?

Ejemplo 5. ■ $F : \mathbb{R}^2 \rightarrow \mathbb{R}^2$, definida por $F(x, y) = (x+1, y-1)$.

No es lineal, veamos:

$$F((x, y) + (a, b)) = F(x+a, y+b) = (x+a+1, y+b-1), \text{ y}$$

$$F(x, y) + F(a, b) = (x+1, y-1) + (a+1, b-1) = (x+a+2, y+b-2).$$

No son iguales, por lo tanto F no cumple la definición de función lineal, no es lineal.

■ $E : \mathbb{R}^3 \rightarrow \mathbb{R}^2$, definida por $E(x, y, z) = (xy, xz, zx)$.

No es lineal, veamos:

$$E(\alpha(x, y, z)) = E(\alpha x, \alpha y, \alpha z) = (\alpha x \alpha y, \alpha y \alpha z, \alpha z \alpha x) = \alpha^2(xy, xz, zx)$$

$\alpha E(x, y, z) = \alpha(xy, xz, zx)$ No son iguales para $\alpha = 2$ por ejemplo, por lo tanto E no es lineal.

Resulta que las funciones lineales heredan las siguientes propiedades.

Propiedad 2. Si $T : V \rightarrow W$ es lineal, entonces:

■ $T(\Theta_V) = \Theta_W$, donde Θ_V es el vector neutro de V y Θ_W el de W .

■ $\forall v \in V, T(-v) = -T(v)$.

■ $\forall \alpha_1, \dots, \alpha_k \in \mathbb{K}, \forall u_1, \dots, u_k \in V, T\left(\sum_{i=1}^k \alpha_i u_i\right) = \sum_{i=1}^k \alpha_i T(u_i)$.

Demostración. Usamos que T es lineal en cada paso.

- $T(\Theta_V) = T(0 \cdot \Theta_V) = 0 \cdot T(\Theta_V) = \Theta_W$.
- Sea $v \in V$, $T(-v) = T(-1 \cdot v) = -1 \cdot T(v) = -T(v)$.
- Sean $\alpha_1, \dots, \alpha_k \in \mathbb{K}$, y sean $u_1, \dots, u_k \in V$, razonamos de manera inductiva:

$$T\left(\sum_{i=1}^k \alpha_i u_i\right) = T\left(\sum_{i=1}^{k-1} \alpha_i u_i\right) + T(\alpha_k u_k) = \dots = \sum_{i=1}^k T(\alpha_i u_i) = \sum_{i=1}^k \alpha_i T(u_i). \quad \square$$

Ejemplo 6. Decida cuáles de las siguientes funciones son lineales y cuáles no.

1. $G : \mathbb{C}^2 \rightarrow \mathbb{C}^3$, $G(x, y) = (ix + y, 0, iy - x)$
2. $H : \mathbb{C}^3 \rightarrow \mathbb{C}^2$, $H(x, y, z) = (\bar{x}, y + z)$
3. $Q : \mathcal{P}(\mathbb{R}) \rightarrow \mathcal{P}(\mathbb{R})$, $Q(p)(x) = x^2 p(x)$
4. $M : \mathcal{P}(\mathbb{R}) \rightarrow \mathcal{P}(\mathbb{R})$, $M(p)(x) = p(x) + x^2$
5. $N : \mathcal{P}_2(\mathbb{R}) \rightarrow \mathbb{R}^2$, $N(ax^2 + bx + c) = (2a + b, a + b + c)$

1. Si. 2. depende si el cuerpo es \mathbb{R} o \mathbb{C} . 3. Si. 4. No. 5. Si.

Subespacios núcleo e imagen de una transformación lineal

Definición 1. Dada $T : V \rightarrow W$ lineal, se definen los siguientes conjuntos.

$$Ker(T) = \{v \in V : T(v) = \Theta_W\}$$

$$\begin{aligned} Im(T) &= \{T(v) \in W : v \in V\} \\ &= \{w \in W : (\exists v \in V) T(v) = w\} \end{aligned}$$

$Ker(T)$ se llama nucleo de T y $Im(T)$ se llama imagen de T .

Estos conceptos no son nuevos, en realidad $Im(T) = rec(T)$, el recorrido de la función T , y $Ker(T)$ es el conjunto de ceros o raíces de T , solo que en el caso de las funciones lineales, el conjunto de raíces suele ser infinito, como se verá.

Ejemplo 1. Considerando $R : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ dada por $R(x, y) = (-x, y)$, entonces,

$$\begin{aligned} Ker(R) &= \{(x, y) \in \mathbb{R}^2 : R(x, y) = (0, 0)\} \\ &= \{(x, y) \in \mathbb{R}^2 : (-x, y) = (0, 0)\} \\ &= \{(x, y) \in \mathbb{R}^2 : x = 0 \wedge y = 0\} \\ &= \{(0, 0)\}, \end{aligned}$$

$$\begin{aligned} Im(R) &= \{R(x, y) : (x, y) \in \mathbb{R}^2\} \\ &= \{(-x, y) : x, y \in \mathbb{R}\} \\ &= \mathbb{R}^2. \end{aligned}$$

Ejemplo 2. Considerando $D : \mathcal{P}(\mathbb{R}) \rightarrow \mathcal{P}(\mathbb{R})$, definida por $D(p) = p'$, entonces,

$$\begin{aligned} Ker(D) &= \{p \in \mathcal{P}(\mathbb{R}) : \forall x \in \mathbb{R}, p'(x) = 0\} \\ &= \{p \in \mathcal{P}(\mathbb{R}) : p \text{ es constante}\} \\ &= \langle\{a\}\rangle, \end{aligned}$$

$$\begin{aligned} Im(D) &= \{p \in \mathcal{P}(\mathbb{R}) : \exists q \in \mathcal{P}(\mathbb{R}), p = q'\} \\ &= \{p \in \mathcal{P}(\mathbb{R}) : p \text{ tiene primitiva en } \mathcal{P}(\mathbb{R})\} \\ &= \mathcal{P}(\mathbb{R}). \end{aligned}$$

Ya que todo polinomio tiene primitiva y ésta es también un polinomio.

Propiedad 1. Dada $T : V \rightarrow W$ lineal, se cumple que

- $\text{Ker}(T)$ es un subespacio vectorial de V , e
- $\text{Im}(T)$ es un subespacio vectorial de W .

Demostración: Sea $T : V \rightarrow W$. Probaremos las 3 propiedades necesarias y suficientes para cada subespacio.

Primero $\text{Ker}(T) = \{v \in V : T(v) = \Theta_W\}$.

1. $\Theta_V \in \text{Ker}(T)$ pues $T(\Theta_V) = \Theta_W$ por propiedad de transformación lineal.
2. Sean $u, v \in \text{Ker}(T)$, esto es $T(u) = T(v) = \Theta_W$, luego $T(u+v) = T(u)+T(v) = \Theta_W + \Theta_W = \Theta_W$ y por tanto $u+v \in \text{Ker}(T)$.
3. Sea $\alpha \in \mathbb{K}$ y sea $u \in \text{Ker}(T)$, luego $T(u) = \Theta_W$ y por tanto $T(\alpha u) = \alpha T(u) = \alpha \Theta_W = \Theta_W$ y en consecuencia $\alpha u \in \text{Ker}(T)$.

Por otro lado, $\text{Im}(T) = \{w \in W : (\exists v \in V) T(v) = w\}$.

1. $\Theta_W \in \text{Im}(T)$ ya que existe una preimagen: $T(\Theta_V) = \Theta_W$.
2. Sean $r, s \in \text{Im}(T)$, esto es, existen $u, v \in V$ tales que $T(u) = r$ y $T(v) = s$. Luego, $T(u+v) = T(u) + T(v) = r + s$, es decir, para $r + s$ existe una preimagen $u + v \in V$, por consiguiente $r + s \in \text{Im}(T)$.
3. Sea $\alpha \in \mathbb{K}$ y sea $r \in \text{Im}(T)$, esto es, existe $u \in V$ tal que $T(u) = r$. De aquí $T(\alpha u) = \alpha T(u) = \alpha r$, es decir, αr tiene preimagen por T y es $\alpha u \in V$, en consecuencia $\alpha r \in \text{Im}(T)$. \square

Dado que son subespacios, es posible calcular su dimensión, y ésta dirá mucho sobre la función:

Definición 2. Dada $T : V \rightarrow W$ lineal, se define:

- $\eta(T) = \dim(\text{Ker}(T))$, y se llama nulidad de T , y
- $r(T) = \dim(\text{Im}(T))$, y se llama rango de T .

Ejemplo 3. En los ejemplos anteriores vemos que:

- $\eta(R) = 0$ y $r(R) = 2$

- $\eta(D) = 1$ y $r(D)$ no está definida pues $\mathcal{P}(R)$ NO tiene dimensión finita.

Ejemplo 4. Si $Q : \mathbb{R}^2 \rightarrow \mathbb{R}^3$ está dada por $Q(x, y) = (x - y, 0, y - x)$, entonces,

$$\begin{aligned} Ker(Q) &= \{(x, y) \in \mathbb{R}^2 : Q(x, y) = (0, 0, 0)\} \\ &= \{(x, y) \in \mathbb{R}^2 : (x - y, 0, y - x) = (0, 0, 0)\} \\ &= \{(x, y) \in \mathbb{R}^2 : x - y = 0 \wedge y - x = 0\} \\ &= \{(x, y) \in \mathbb{R}^2 : x = y\} \\ &= \{(x, x) \in \mathbb{R}^2 : x \in \mathbb{R}\} \\ &= \langle\{(1, 1)\}\rangle \end{aligned}$$

$$\begin{aligned} Im(Q) &= \{Q(x, y) \in \mathbb{R}^3 : (x, y) \in \mathbb{R}^2\} \\ &= \{(x - y, 0, y - x) \in \mathbb{R}^3 : (x, y) \in \mathbb{R}^2\} \\ &= \{(x - y, 0, y - x) = x(1, 0, -1) + y(-1, 0, 1) \in \mathbb{R}^3 : x, y \in \mathbb{R}\} \\ &= \langle\{(1, 0, -1), (-1, 0, 1)\}\rangle \\ &= \langle\{(1, 0, -1)\}\rangle \end{aligned}$$

Así la nulidad de Q es 1 y su rango también es 1.

La siguiente propiedad es muy útil y es la manera que realmente se usa para calcular la imagen de una transformación lineal.

Propiedad 2. Dada $T : V \rightarrow W$ lineal, y $B = (v_1, \dots, v_m)$ una base de V , se tiene que:

$$Im(T) = \langle\{T(v_1), \dots, T(v_m)\}\rangle.$$

Demostración: Si $\{v_1, \dots, v_m\}$ es una base de V , entonces se cumple

$$v \in V \Leftrightarrow \exists \alpha_1, \dots, \alpha_m \in \mathbb{K}, v = \sum_{i=1}^m \alpha_i v_i,$$

de aquí:

$$\begin{aligned} Im(T) &= \{T(v) : v \in V\} \\ &= \left\{ T(v) : \exists \alpha_1, \dots, \alpha_m \in \mathbb{K}, v = \sum_{i=1}^m \alpha_i v_i \right\} \\ &= \left\{ T\left(\sum_{i=1}^m \alpha_i v_i\right) : \alpha_1, \dots, \alpha_m \in \mathbb{K} \right\} \\ &= \left\{ \sum_{i=1}^m \alpha_i T(v_i) : \alpha_1, \dots, \alpha_m \in \mathbb{K} \right\} \\ &= \langle\{T(v_1), \dots, T(v_m)\}\rangle. \end{aligned}$$

□

Ejemplo 5. En uno de los ejemplos anteriores, $R : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ con $R(x, y) = (-x, y)$, si tomamos la base canónica $\mathcal{C}_2 = \{(1, 0), (0, 1)\}$ de \mathbb{R}^2 sobre \mathbb{R} , entonces:

$$\begin{aligned} Im(R) &= \langle \{R(1, 0), R(0, 1)\} \rangle \\ &= \langle \{(-1, 0), (0, 1)\} \rangle \\ &= \mathbb{R}^2. \end{aligned}$$

Ejemplo 6. Consideremos la función $P : \mathcal{P}_2(\mathbb{R}) \rightarrow \mathbb{R}^3$ definida por:

$$P(p) = (p(0), p'(0), p''(0)).$$

Calculemos su núcleo e imagen nulidad y rango.

$$\begin{aligned} Ker(P) &= \{p \in \mathcal{P}_2 : (p(0), p'(0), p''(0)) = (0, 0, 0)\} \\ &= \{ax^2 + bx + c \in \mathcal{P}_2 : (c, b, 2a) = (0, 0, 0)\} \\ &= \{0\} \end{aligned}$$

Vemos entonces que $\eta(P) = 0$.

$$\begin{aligned} Im(P) &= \{(p(0), p'(0), p''(0)) \in \mathbb{R}^3 : p \in \mathcal{P}_2(\mathbb{R})\} \\ &= \{(c, b, 2a) \in \mathbb{R}^3 : a, b, c \in \mathbb{R}\} \\ &= \langle \{(1, 0, 0), (0, 1, 0), (0, 0, 2)\} \rangle \\ &= \mathbb{R}^3 \end{aligned}$$

Vemos entonces que $r(P) = 3$.

Teorema Núcleo-Imagen

En los ejemplos anteriores comprobamos que la nulidad es igual al rango menos la dimensión del espacio de partida m . El siguiente es uno de los teoremas más importantes de esta área, y el más importante de la unidad.

Teorema 1 (Núcleo-Imagen). *Dada $T : V \rightarrow W$ lineal, se cumple:*

$$\eta(T) + r(T) = \dim(V)$$

Demostración: Primero que nada tomemos una base de $\text{Ker}(T)$, digamos: $\{u_1, \dots, u_k\}$, donde $k = \eta(T)$.

- Si $k = \dim(V)$, significa que $\text{Ker}(T) = V$, lo cual solo ocurre si se trata de la función nula, en ese caso $\text{Im}(T) = \{\Theta_W\}$, por lo tanto $r(T) = 0$, así $\eta(T) + r(T) = \dim(V) + 0 = \dim(V)$, y se cumple el teorema trivialmente.

Por ejemplo la función nula de \mathbb{K}^5 en \mathbb{K}^3 tiene fórmula: $T(x, y, z, t, s) = (0, 0, 0)$.

- Si $k < \dim(V)$, agregamos $\dim(V) - k$ vectores a esta base, hasta obtener una base de $\dim(V)$, esto siempre es posible hacerlo, tal como vimos en el capítulo anterior. Obtenemos así una base de $V : \{u_1, \dots, u_k, u_{k+1}, \dots, u_m\}$, donde $m = \dim(V)$

A partir de la última propiedad de la clase anterior se obtiene la imagen de T :

$$\begin{aligned} \text{Im}(T) &= \langle \{T(u_1), \dots, T(u_k), T(u_{k+1}), \dots, T(u_m)\} \rangle \\ &= \langle \{\Theta_W, \dots, \Theta_W, T(u_{k+1}), \dots, T(u_m)\} \rangle \\ &= \langle \{T(u_{k+1}), \dots, T(u_m)\} \rangle \end{aligned}$$

Se demostrará que este conjunto de $m - k$ vectores generadores de $\text{Im}(T)$ es l. i. y así se concluirá el teorema.

Sean $\lambda_{k+1}, \dots, \lambda_m \in \mathbb{K}$ escalares tales que $\sum_{i=k+1}^m \lambda_i T(u_i) = \Theta_W$.

$$\begin{aligned}
\sum_{i=k+1}^m \lambda_i T(u_i) = \Theta_W &\Leftrightarrow T\left(\sum_{i=k+1}^m \lambda_i u_i\right) = \Theta_n \\
&\Leftrightarrow \sum_{i=k+1}^m \lambda_i u_i \in \text{Ker}(T) \\
&\Leftrightarrow \sum_{i=k+1}^m \lambda_i u_i \in \langle\{u_1, \dots, u_k\}\rangle \\
&\Leftrightarrow \exists \alpha_1, \dots, \alpha_k \sum_{i=k+1}^m \lambda_i u_i = \sum_{j=1}^k \alpha_j u_j \\
&\Leftrightarrow \exists \alpha_1, \dots, \alpha_k \sum_{j=1}^k \alpha_j u_j - \sum_{i=k+1}^m \lambda_i u_i = \Theta_m \quad \text{ya que } \{u_1, \dots, u_m\} \text{ es base} \\
&\Leftrightarrow \forall i \in \{k+1, \dots, m\}, \forall j \in \{1, \dots, k\}, \lambda_i = 0 \wedge \alpha_j = 0
\end{aligned}$$

Así se demuestra que $\eta(T) + r(T) = k + m - k = m = \dim(V)$. \square

Ejemplo 7. En el ejemplo 10 tenemos que $\eta(P) = 0$ y $r(P) = 3$, se verifica en efecto que $0 + 3 = \dim(\mathcal{P}_2(\mathbb{R}))$.

Si hubiésemos conocido el teorema Núcleo-Imagen previamente habríamos podido adivinar $\text{Im}(P)$ sin necesidad de calcularlo, ya que al saber que $\eta(P) = 0$, el teorema Núcleo-Imagen nos adelanta que $r(P) = \dim(\mathcal{P}_2(\mathbb{R})) - \eta(P) = 3 - 0 = 3$, así $\dim(\text{Im}(P)) = 3 = \dim(\mathbb{R}^3)$, y como $\text{Im}(P)$ es subespacio de \mathbb{R}^3 , se concluye que $\text{Im}(P) = \mathbb{R}^3$.

Ejemplo 8. Invete una aplicación lineal G tal que $\eta(G) = 3$ y $r(G) = 2$.

Primero que nada el espacio de partida debe tener dimensión 5, por ejemplo \mathbb{R}^5 .

Segundo, el espacio de llegada debe tener dimensión mayor o igual a 2, por ejemplo \mathbb{R}^4 .

Para no complicarnos la vida podemos definir G simplemente como sigue.

$$G : \mathbb{R}^5 \rightarrow \mathbb{R}^4$$

$$G(a, b, c, d, e) = (a, b, 0, 0)$$

Por las propiedades vistas, es evidente que $\text{Im}(G) = \langle\{(1, 0, 0, 0), (0, 1, 0, 0)\}\rangle$, por lo tanto $r(G) = \dim(\text{Im}(G)) = 2$.

Aplicando el teorema núcleo-imagen concluimos que $\eta(G) = \dim(\mathbb{R}^5) - 2 = 3$.

Inyectividad y sobreyectividad

Recordemos ahora conceptos importantes de la teoría de funciones.

- $T : A \rightarrow B$ es *sobreyectiva* si y solo si $rec(T) = B$.
- $T : A \rightarrow B$ es *inyectiva* si y solo sí para todo $u \neq v$, $T(u) \neq T(v)$.

En otras palabras, una función es sobreyectiva si todo elemento del conjunto de llegada tiene al menos una preimagen, y una función es inyectiva si y solo si todo elemento del recorrido tiene una única preimagen.

Observación 1. Dada $T : V \rightarrow W$ lineal, se cumple la siguiente equivalencia.

$$T \text{ es sobreyectiva} \Leftrightarrow Im(T) = W \Leftrightarrow r(T) = \dim(W)$$

Demostración. Dado que $Im(T)$ es subespacio de W , si resultara que $r(T) = \dim(W)$, tendría que tenerse que $Im(T) = W$. La recíproca es obvia. \square

La inyectividad por su parte está asociada a la nulidad de la transformación, gracias al siguiente teorema.

Teorema 2. Dada $T : V \rightarrow W$ lineal, se cumple la siguiente equivalencia.

$$T \text{ es inyectiva} \Leftrightarrow Ker(T) = \{\Theta_V\} \Leftrightarrow \eta(T) = 0$$

Demostración.

\Rightarrow Si T es inyectiva, Θ_W tiene una única preimagen, que es Θ_V gracias a la Propiedad ??, lo cual implica que $Ker(T) = \{\Theta_V\}$, así $n(T) = 0$.

\Leftarrow Supongamos que $\eta(T) = 0$, esto significa que $Ker(T) = \{\Theta_V\}$, ya que el espacio nulo es el único espacio de dimensión nula. Estudiemos ahora la inyectividad de T . Para esto tomemos dos vectores arbitrarios $u, v \in V$, tales que $T(u) = T(v)$. Trabajando esta ecuación obtenemos que: $T(u) - T(v) = \Theta_W$, entonces $T(u - v) = \Theta_W$, por lo tanto $u - v \in Ker(T) = \{\Theta_V\}$. Pero esto significa que $u - v = \Theta_V$, ya que este es el único elemento dentro de $Ker(T)$. Así, $u = v$, lo cual prueba que T es inyectiva. \square

Ejemplo 9. La función reflexión R es inyectiva claramente pues su nulidad es 0. Si lo pensamos ahora desde la intuición, al tratarse se una reflexion, es claro que dos vectores distintos no pueden hacerse iguales cuando se reflejan, reflejarse es una acción que no pierde información, por lo tanto es inyectiva.

Ejemplo 10. Consideremos nuevamente el ejemplo 10.

Teníamos que $\eta(P) = 0$ por lo tanto P es inyectiva. Dicho de otro modo, si conocemos la posición, primera derivada y segunda derivada de una parábola en 0, entonces la conocemos completamente.

Por otra parte calculamos que $r(P) = 3$, por lo tanto P es sobreyectiva. Dicho de otro modo, la posición, primera derivada (velocidad) y segunda derivada (aceleración) de una parábola en 0 pueden tomar cualquier valor en \mathbb{R} .

Si combinamos esto con el teorema Núcleo-Imagen, vemos que en algunos casos algunas propiedades son imposibles.

Corolario 1. *Dada una transformación lineal $T : V \rightarrow W$, podemos afirmar lo siguiente.*

- Si $\dim(V) < \dim(W)$, entonces T no puede ser sobreyectiva.
- Si $\dim(V) > \dim(W)$, entonces T no puede ser inyectiva.
- Si $\dim(V) = \dim(W)$, entonces, T es inyectiva si y solo si T es sobreyectiva.

Demostración.

- Si $\dim(V) < \dim(W)$, entonces $r(T) = \dim(V) - \eta(T) \leq \dim(V) < \dim(W)$, y para ser sobreyectiva su rango debe ser igual a $\dim(W)$.
- Si $\dim(V) > \dim(W)$, entonces $\eta(T) = \dim(V) - r(T) \geq \dim(V) - \dim(W) > 0$, y para ser inyectiva su nulidad debe ser 0.
- Supongamos que $\dim(V) = \dim(W)$.
 - \Rightarrow Si T es inyectiva, entonces $\eta(T) = 0$, entonces $r(T) = \dim(V) = \dim(W)$ por lo tanto T es sobreyectiva.
 - \Leftarrow Si T es sobreyectiva, entonces $r(T) = \dim(W)$, entonces $\eta(T) = \dim(V) - r(T) = \dim(V) - \dim(W) = 0$ por lo tanto T es inyectiva.

□

Ejemplo 11. Consideremos la función $E : \mathbb{K}^5 \rightarrow \mathbb{K}^6$ definida por

$$E(x, y, z, s, t) = (x - z, 2x + y - s - 2t, y - s, x - t, x - 2z + t, z - t),$$

Sin necesidad de calcular nada, ya sabemos que esta función no puede ser sobreyectiva.

Corolario 2. *Si $T : V \rightarrow W$ es lineal e inyectiva, y $B = \{v_1, \dots, v_k\}$ in conjunto l.i., entonces se tiene que el conjunto $\{T(v_1), \dots, T(v_k)\}$ es también l.i.*

Demostración. Primero que nada restrinjamos el dominio de T al subespacio $U = \langle B \rangle$ de V , así tenemos que $T|_U : U \rightarrow W$ también es lineal e inyectiva.

Observamos que $\dim(U) = k$ dado que B sería su base y esta tiene k elementos.

Como $T|_U$ es inyectiva, se tiene que $\dim(\text{Ker}(T|_U)) = 0$, entonces el Teorema Núcleo-Imagen nos dice que

$$\dim(U) = 0 + r(T|_U) = r(T|_U) = \dim(\text{Im}(T|_U))$$

Dado que $\text{Im}(T|_U) = \langle \{T(v_1), \dots, T(v_k)\} \rangle$ está generado por $k = \dim(U)$ vectores y tiene dimensión k , no queda otra opción mas que su generador, $\{T(v_1), \dots, T(v_k)\}$, sea l.i. \square

Ejemplo 12. Considere la siguiente función $L : \mathcal{P}_2(\mathbb{R}) \rightarrow \mathbb{R}^3$ definida por:

$$L(p) = (p(-1), p(3), p(10)).$$

Calculemos su núcleo.

$$\begin{aligned} \text{Ker}(L) &= \{p \in \mathcal{P}_2(\mathbb{R}) : (p(-1), p(3), p(10)) = (0, 0, 0)\} \\ &= \{p \in \mathcal{P}_2(\mathbb{R}) : p(-1) = 0, p(3) = 0, p(10) = 0\} \end{aligned}$$

Es el conjunto de todos los polinomios de grado menor o igual a 2 que se anulan en $x = -1, 3$ y 10 , es decir, que tienen 3 raíces y estas son $-1, 3$ y 10 . Pero esto es imposible, los polinomios de grado menor o igual a 2 tienen a lo más 2 raíces, solo el polinomio nulo puede tener más de dos raíces... por lo tanto.

$$\text{Ker}(L) = \{0\} \text{ y } \eta(L) = 0$$

De aquí vemos que L es inyectiva, y además

$$r(L) = \dim(\mathcal{P}_2(\mathbb{R})) - 0 = 3,$$

de donde se deduce también que $\text{Im}(L) = \mathbb{R}^3$, y que L es sobreyectiva.

Este resultado es interesante pues nos dice que para conocer completamente un parábola, basta conocerla en 3 puntos diferentes ($-1, 3$ y 10 son circunstanciales), y que dados 3 puntos cualesquiera (con abscisas distintas) siempre existirá una parábola que los visite.

Matrices

Recordemos la siguiente propiedad vista y demostrada en la semana 5.

Propiedad 1. Considerando a \mathbb{K}^m y \mathbb{K}^n como espacios vectoriales sobre \mathbb{K} . Si $T : \mathbb{K}^m \rightarrow \mathbb{K}^n$ es una función definida como sigue:

$$T(x_1, x_2, \dots, x_m) = (a_{11}x_1 + a_{12}x_2 + \dots + a_{1m}x_m, a_{21}x_1 + \dots + a_{2m}x_m, \dots, a_{n1}x_1 + \dots + a_{nm}x_m),$$

donde $a_{ij} \in \mathbb{K}$, con $i \in \{1, \dots, n\}$ y $j \in \{1, \dots, m\}$, son nm constantes fijas, entonces T es lineal.

Ejemplo 1. $G : \mathbb{R}^4 \rightarrow \mathbb{R}^3$, $G(x_1, x_2, x_3, x_4) = (2x_1 - x_2 + x_4, x_1 + 2x_2, x_2 - x_4)$. vemos que $m = 4$ y $n = 3$, además G tiene la forma que pide la Propiedad 1, basta tomar

$$a_{11} = 2, a_{12} = -1, a_{13} = 0, a_{14} = 1,$$

$$a_{21} = 1, a_{22} = 2, a_{23} = 0, a_{24} = 0,$$

$$a_{31} = 0, a_{32} = 1, a_{33} = 0, \text{ y } a_{34} = -1.$$

Aplicando la Propiedad 1, obtenemos que G es lineal.

Ahora tratemos de entender mejor esta notación, ¿por qué hay dos subíndices?, ¿por qué están definidos de esa forma?, etc.. Tomemos un ejemplo general con $m = 4$ y $n = 3$ y consideremos la base canónica de \mathbb{R}^4 :

$$e_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, e_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, e_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, e_4 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}.$$

Tomemos también un vector arbitrario $v = (x_1, x_2, x_3, x_4)$, lo podemos escribir como columna (ya dijimos que los vectores son indistintamente columnas o tuplas), y desglosarlo como combinación lineal de la base canónica:

$$v = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = x_1 \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} + x_2 \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} + x_3 \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} + x_4 \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}.$$

Por lo tanto, si aplicamos el T y usamos su linealidad obtenemos lo siguiente:

$$\begin{aligned}
T(v) &= T \left(x_1 \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} + x_2 \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} + x_3 \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} + x_4 \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} \right) \\
&= T \left(x_1 \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} \right) + T \left(x_2 \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} \right) + T \left(x_3 \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} \right) + T \left(x_4 \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} \right) \\
&= x_1 T \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} + x_2 T \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} + x_3 T \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} + x_4 T \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}.
\end{aligned}$$

Si ahora aplicamos la fórmula que tiene T en la Propiedad 1, con $m = 4$ y $n = 3$:

$$T(x_1, x_2, x_3, x_4) = (a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + a_{14}x_4, a_{21}x_1 + a_{22}x_2 + a_{23}x_3 + a_{24}x_4, a_{31}x_1 + a_{32}x_2 + a_{33}x_3 + a_{34}x_4),$$

y calculamos cuánto vale T en la base canónica, vemos que:

$$T(e_1) = \begin{pmatrix} a_{11} \\ a_{21} \\ a_{31} \end{pmatrix}, T(e_2) = \begin{pmatrix} a_{12} \\ a_{22} \\ a_{32} \end{pmatrix}, T(e_3) = \begin{pmatrix} a_{13} \\ a_{23} \\ a_{33} \end{pmatrix}, T(e_4) = \begin{pmatrix} a_{14} \\ a_{24} \\ a_{34} \end{pmatrix}.$$

Por lo tanto

$$\begin{aligned}
T(v) &= x_1 T \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} + x_2 T \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} + x_3 T \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} + x_4 T \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} \\
&= x_1 \begin{pmatrix} a_{11} \\ a_{21} \\ a_{31} \end{pmatrix} + x_2 \begin{pmatrix} a_{12} \\ a_{22} \\ a_{32} \end{pmatrix} + x_3 \begin{pmatrix} a_{13} \\ a_{23} \\ a_{33} \end{pmatrix} + x_4 \begin{pmatrix} a_{14} \\ a_{24} \\ a_{34} \end{pmatrix} \\
&= \begin{pmatrix} a_{11}x_1 \\ a_{21}x_1 \\ a_{31}x_1 \end{pmatrix} + \begin{pmatrix} a_{12}x_2 \\ a_{22}x_2 \\ a_{32}x_2 \end{pmatrix} + \begin{pmatrix} a_{13}x_3 \\ a_{23}x_3 \\ a_{33}x_3 \end{pmatrix} + \begin{pmatrix} a_{14}x_4 \\ a_{24}x_4 \\ a_{34}x_4 \end{pmatrix} \\
&= \begin{pmatrix} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + a_{14}x_4 \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 + a_{24}x_4 \\ a_{31}x_1 + a_{32}x_2 + a_{33}x_3 + a_{34}x_4 \end{pmatrix}
\end{aligned}$$

Ahora se ve más clara la estructura de los subíndices de a_{ij} . Ellos describen lo que conocemos como “matriz”.

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \end{pmatrix}$$

Entonces la operación de “*evaluar la transformación en el vector*” se define como “*multiplicación de matriz por vector*”:

$$\begin{aligned}
 Av &= \begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} \\
 &= x_1 \begin{pmatrix} a_{11} \\ a_{21} \\ a_{31} \end{pmatrix} + x_2 \begin{pmatrix} a_{12} \\ a_{22} \\ a_{32} \end{pmatrix} + x_3 \begin{pmatrix} a_{13} \\ a_{23} \\ a_{33} \end{pmatrix} + x_4 \begin{pmatrix} a_{14} \\ a_{24} \\ a_{34} \end{pmatrix} \\
 &= \begin{pmatrix} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + a_{14}x_4 \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 + a_{24}x_4 \\ a_{31}x_1 + a_{32}x_2 + a_{33}x_3 + a_{34}x_4 \end{pmatrix}
 \end{aligned}$$

A es la llamada *matriz asociada a la transformación*.

En el ejemplo del comienzo, $G(x_1, x_2, x_3, x_4) = (2x_1 - x_2 + x_4, x_1 + 2x_2, x_2 - x_4)$, tenemos:
 $G(1, 0, 0, 0) = (2, 1, 0)$, $G(0, 1, 0, 0) = (-1, 2, 1)$, $G(0, 0, 1, 0) = (0, 0, 0)$, $G(0, 0, 0, 1) = (1, 0, -1)$.
Así:

$$G \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 2 & -1 & 0 & 1 \\ 1 & 2 & 0 & 0 \\ 0 & 1 & 0 & -1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 2x_1 - x_2 + x_4 \\ x_1 + 2x_2 \\ x_2 - x_4 \end{pmatrix}.$$

La matriz asociada a G es:

$$\begin{pmatrix} 2 & -1 & 0 & 1 \\ 1 & 2 & 0 & 0 \\ 0 & 1 & 0 & -1 \end{pmatrix}.$$

Ejemplo 2. Considere una transformación lineal $T : \mathbb{K}^3 \rightarrow \mathbb{K}^3$ tal que $T(1, 0, 0) = (1, 0, 1)$, $T(0, 1, 0) = (2, 2, 1)$ y $T(0, 0, 1) = (0, 1, 1)$. Usando la propiedad 1 (en particular, la tercera) es posible obtener la transformación para cualquier vector de \mathbb{K}^3 . En efecto, los vectores $(1, 0, 0)$, $(0, 1, 0)$ y $(0, 0, 1)$ forman la base canónica de \mathbb{K}^3 , luego todo vector (x, y, z) puede ser escrito como una combinación lineal de ellos:

$$(x, y, z) = x(1, 0, 0) + y(0, 1, 0) + z(0, 0, 1),$$

donde $x, y, z \in \mathbb{K}$, y por tanto:

$$\begin{aligned}
 T(x, y, z) &= T(x(1, 0, 0) + y(0, 1, 0) + z(0, 0, 1)) \\
 &= T(x(1, 0, 0)) + T(y(0, 1, 0)) + T(z(0, 0, 1)) \\
 &= xT(1, 0, 0) + yT(0, 1, 0) + zT(0, 0, 1) \\
 &= x(1, 0, 1) + y(2, 2, 1) + z(0, 1, 1) \\
 &= (x + 2y, 2y + z, x + y + z),
 \end{aligned}$$

es decir, $T(x, y, z) = (x + 2y, 2y + z, x + y + z)$.

O dicho de otra forma:

$$T \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 1 & 2 & 0 \\ 0 & 2 & 1 \\ 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x + 2y \\ 2y + z \\ x + y + z \end{pmatrix}$$

La matriz asociada a T es

$$\begin{pmatrix} 1 & 2 & 0 \\ 0 & 2 & 1 \\ 1 & 1 & 1 \end{pmatrix}.$$

Observamos que las columnas de la matriz son las imágenes de los vectores de la base canónica. Esto implica además que basta conocer la función en la base canónica para poder conocerla en todo el espacio. Más adelante veremos un teorema que dice que basta conocer la función en cualquier base del espacio para conocerla en todas partes, y ese teorema se conoce como el "*Teorema fundamental del álgebra lineal*".

Definición 1. Una matriz es un arreglo rectangular de números. Una matriz de orden $m \times n$ a coeficientes en un conjunto \mathbb{K} es un arreglo de elementos de \mathbb{K} de m filas y n columnas.

$$A = \underbrace{\begin{pmatrix} A_{11} & A_{12} & \cdots & A_{1n} \\ A_{21} & a_{22} & \cdots & A_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ A_{m1} & A_{m2} & \cdots & A_{mn} \end{pmatrix}}_{n \text{ columnas}} \Bigg\} m \text{ filas}$$

El coeficiente A_{ij} está en la fila i y en la columna j .

El conjunto de todas las matrices de orden $m \times n$ a coeficientes en \mathbb{K} se denota $\mathcal{M}_{m \times n}(\mathbb{K})$.

Definición 2 (Suma). Sólo se pueden sumar matrices de igual orden, y la suma es coeficiente a coeficiente. Si $A, B \in \mathcal{M}_{m \times n}(\mathbb{K})$, entonces

$$(A + B)_{ij} = A_{ij} + B_{ij}, \text{ para cada } i \in \{1, \dots, m\}, j \in \{1, \dots, n\}.$$

Si \mathbb{K} es cuerpo, resulta (y es fácil verlo) que la suma es **conmutativa**, **asociativa**, tiene **neutro aditivo** (el cual es la matriz cuyos coeficientes son todos iguales a 0), y cada matriz tiene un **inverso aditivo**, que se obtiene reemplazando cada coeficiente por su inverso aditivo. En otras palabras, si $A, B, C \in \mathcal{M}_{m \times n}(\mathbb{K})$, y $\Theta_{m \times n}$ denota la matriz de orden $m \times n$ cuyos coeficientes son todos nulos, entonces:

$$\blacksquare A + B = B + A \quad \blacksquare (A + B) + C = A + (B + C) \quad \blacksquare A + \Theta_{m \times n} = A$$

Definición 3 (Ponderación). Sólo se puede ponderar una matriz por un escalar del cuerpo, coeficiente a coeficiente. Si $A \in \mathcal{M}_{m \times n}(\mathbb{K})$, y $\lambda \in \mathbb{K}$, entonces

$$(\lambda A)_{ij} = \lambda A_{ij}, \text{ para cada } i \in \{1, \dots, m\}, j \in \{1, \dots, n\}.$$

Si $A, B \in \mathcal{M}_{m \times n}(\mathbb{K})$, y $\alpha, \beta \in \mathbb{K}$, entonces:

- $\alpha(\beta(A)) = (\alpha\beta)A$
- $(\alpha + \beta)A = \alpha A + \beta B$
- $A + (-1)A = \Theta_{m \times n}$
- $\alpha(A + B) = \alpha A + \alpha B$
- $1A = A$

En otras palabras, $(\mathcal{M}_{m \times n}(\mathbb{K}), +, \cdot)$ es un espacio vectorial sobre \mathbb{K} .

Como espacio vectorial, $\mathcal{M}_{m \times n}(\mathbb{K})$ sobre \mathbb{K} también tiene su base canónica, y esta consiste en tomar las matrices que tienen un 1 en una posición y 0 en las demás, e ir las poniendo en orden de izquierda a derecha y de arriba a abajo. Por ejemplo, la base canónica del espacio de las matrices de orden 3×2 es la siguiente, en el ordenamiento que sigue:

$$\mathcal{C}_{3 \times 2}(\mathbb{K}) = \left\{ \begin{pmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 1 \end{pmatrix} \right\}.$$

Esto nos demuestra que $\dim(\mathcal{M}_{m \times n}(\mathbb{K})) = mn$. Por ejemplo $\dim(\mathcal{M}_{3 \times 2}(\mathbb{K})) = 6$.

Multiplicación de matrices

Al inicio de esta semana definimos la multiplicación de matriz por vector de manera tal que se tuviera que “evaluar la transformación en el vector fuera igual a multiplicar la matriz por el vector”, e decir que se tuviera la ecuación:

$$T(v) = Av.$$

Las matrices se definen para representar bien a las transformaciones lineales. Recordamos que si $T : \mathbb{K}^m \rightarrow \mathbb{K}^n$, entonces su matriz asociada $A \in \mathcal{M}_{n \times m}(\mathbb{K})$.

Ahora vamos a definir la multiplicación entre matrices. Para hacerlo con sentido, vamos a asociar esta operación a la composición de transformaciones lineales.

Sean $T : \mathbb{K}^m \rightarrow \mathbb{K}^n$ y $F : \mathbb{K}^l \rightarrow \mathbb{K}^m$, dado que el espacio de llegada de F es igual al espacio de partida de T podemos definir la composición $T \circ F : \mathbb{K}^l \rightarrow \mathbb{K}^n$, por $(T \circ F)(v) = T(F(v))$.

Supongamos ahora que A es la matriz asociada a T , es decir cumple: $T(v) = Av$. Se debe tener que $A \in \mathcal{M}_{n \times m}(\mathbb{K})$.

Supongamos que B es la matriz asociada a F , es decir que cumple: $F(u) = Bu$. Se debe tener que $B \in \mathcal{M}_{m \times l}(\mathbb{K})$.

Tenemos entonces que $(T \circ F)(v) = T(F(v)) = T(Bu) = A(Bu) = (AB)u$. En otras palabras, la matriz asociada a la composición de T con F es igual a la multiplicación de las matrices A y B . Como $T \circ F : \mathbb{K}^l \rightarrow \mathbb{K}^n$, se debe tener $AB \in \mathcal{M}_{n \times l}(\mathbb{K})$.

Notamos que:

$$\underbrace{A}_{n \times m} \underbrace{B}_{m \times l} = \underbrace{AB}_{n \times l}$$

¿Pero cómo es AB ? Recordemos que las columnas de la matriz asociada a una transformación corresponden a las imágenes de los vectores canónicos.

La i -ésima columna de AB es entonces $T(F(e_i))$. Pero $F(e_i)$ es la i -ésima columna de B , por lo tanto $T(F(e_i)) = T(Col_i(B)) = ACol_i(B)$, es la multiplicación de A por la i -ésima columna de B , y esa operación la sabemos hacer pues es multiplicación de matriz por vector.

Ejemplo 3. Si tomamos $l = 2$, $m = 2$ y $n = 3$, $A \in \mathcal{M}_{3 \times 2}(\mathbb{K})$, $B \in \mathcal{M}_{2 \times 2}(\mathbb{K})$, $AB \in \mathcal{M}_{3 \times 2}(\mathbb{K})$.

$$A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \\ A_{31} & A_{32} \end{pmatrix}, \quad B = \begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix}$$

$$Col_1(B) = \begin{pmatrix} B_{11} \\ B_{21} \end{pmatrix}, \quad Col_2(B) = \begin{pmatrix} B_{12} \\ B_{22} \end{pmatrix}$$

$$Col_1(AB) = A \begin{pmatrix} B_{11} \\ B_{21} \end{pmatrix} = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \\ A_{31} & A_{32} \end{pmatrix} \begin{pmatrix} B_{11} \\ B_{21} \end{pmatrix} = \begin{pmatrix} A_{11}B_{11} + A_{12}B_{21} \\ A_{21}B_{11} + A_{22}B_{21} \\ A_{31}B_{11} + A_{32}B_{21} \end{pmatrix}$$

$$Col_2(AB) = A \begin{pmatrix} B_{12} \\ B_{22} \end{pmatrix} = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \\ A_{31} & A_{32} \end{pmatrix} \begin{pmatrix} B_{12} \\ B_{22} \end{pmatrix} = \begin{pmatrix} A_{11}B_{12} + A_{12}B_{22} \\ A_{21}B_{12} + A_{22}B_{22} \\ A_{31}B_{12} + A_{32}B_{22} \end{pmatrix}$$

Juntando todo obtenemos AB :

$$AB = \begin{pmatrix} A_{11}B_{11} + A_{12}B_{21} & A_{11}B_{12} + A_{12}B_{22} \\ A_{21}B_{11} + A_{22}B_{21} & A_{21}B_{12} + A_{22}B_{22} \\ A_{31}B_{11} + A_{32}B_{21} & A_{31}B_{12} + A_{32}B_{22} \end{pmatrix}.$$

En otras palabras, tenemos que el coeficiente ij de AB tiene la siguiente fórmula

$$(AB)_{ij} = A_{i1}B_{1j} + A_{i2}B_{2j}$$

Definición 4 (Multiplicación). *Sólo se puede calcular AB , si A tiene tantas columnas como filas tiene B , es decir, si $A \in \mathcal{M}_{m \times n}(S)$ y $B \in \mathcal{M}_{k \times l}(S)$, sólo se pueden multiplicar si $n = k$. En ese caso $AB \in \mathcal{M}_{m \times l}$ y:*

$$(AB)_{ij} = \sum_{p=1}^n A_{ip}B_{pj}, \text{ para cada } i \in \{1, \dots, m\}, j \in \{1, \dots, l\}.$$

En otras palabras, para calcular el coeficiente ij de AB , debemos multiplicar cada coeficiente de la i -ésima fila de A , con cada coeficiente respectivo de la j -ésima columna de B , y luego sumar los n productos obtenidos.

Dado que la composición de funciones es asociativa, la multiplicación también resulta asociativa. Hay un neutro multiplicativo, el cual corresponde a la matriz asociada a la función identidad. Además, la multiplicación distribuye respecto a la adición (no es difícil verlo). Sin embargo NO es conmutativa en general, y NO toda matriz tiene inverso multiplicativo; peor aún, existen matrices no nulas que al multiplicarlas da como resultado la matriz nula $\Theta_{m \times n}$.

Veamos primero que la multiplicación NO es conmutativa. Podríamos tomar dos matrices al azar $A \in \mathcal{M}_{2 \times 3}(\mathbb{R})$, and $B \in \mathcal{M}_{3 \times 4}(\mathbb{R})$, entonces $AB \in \mathcal{M}_{2 \times 4}(\mathbb{R})$. Sin embargo BA no se puede definir, obviamente no puede haber conmutatividad entre estas matrices.

Pero tomemos mejor dos matrices cuadradas, $C, D \in \mathcal{M}_{2 \times 2}(\mathbb{R})$, en este caso CD y DC están definidas y ambas son de orden 2×2 . La pregunta es entonces: ¿serán iguales?. Veámoslo con números.

$$C = \begin{pmatrix} 1 & 2 \\ -1 & 1 \end{pmatrix} \quad D = \begin{pmatrix} 2 & -2 \\ 1 & 2 \end{pmatrix}$$

Las multiplicamos y obtenemos.

$$\begin{aligned} CD &= \begin{pmatrix} 1 & 2 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} 2 & -2 \\ 1 & 2 \end{pmatrix} = \begin{pmatrix} 4 & 2 \\ -1 & 4 \end{pmatrix} \\ DC &= \begin{pmatrix} 2 & -2 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ -1 & 1 \end{pmatrix} = \begin{pmatrix} 4 & 2 \\ -1 & 4 \end{pmatrix} \end{aligned}$$

¡¡Jaja!! ¡¡conmutan!!...

¿Pero será cierto con cualquier par de matrices, o será que tuvimos “suerte”?

Tomemos otra matriz:

$$E = \begin{pmatrix} -1 & 2 \\ 3 & 1 \end{pmatrix}.$$

$$\begin{aligned} CE &= \begin{pmatrix} 1 & 2 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} -1 & 2 \\ 3 & 1 \end{pmatrix} = \begin{pmatrix} 5 & 4 \\ 4 & -1 \end{pmatrix} \\ EC &= \begin{pmatrix} -1 & 2 \\ 3 & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ -1 & 1 \end{pmatrix} = \begin{pmatrix} -3 & 0 \\ 2 & 7 \end{pmatrix} \end{aligned}$$

... era casualidad. Veámoslo algebraicamente ahora.

$$(AB)_{ij} = \sum_{k=1}^m A_{ik}B_{kj} \quad (BA)_{ij} = \sum_{k=1}^n B_{ik}A_{kj} = \sum_{k=1}^n A_{kj}B_{ik}$$

Son expresiones distintas, era de esperarse que no dieran el mismo resultado. La multiplicación de matrices no es conmutativa.

La otra cosa que sorprende es que dos matrices distintas de la nula pueden producir una matriz nula al multiplicarse, veamos un ejemplo donde esto ocurre.

$$\begin{pmatrix} 2 & -3 \\ -4 & 6 \end{pmatrix} \begin{pmatrix} 3 & 9 \\ 2 & 6 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

Exploraremos ahora el neutro multiplicativo, primero observamos que debe ser una matriz cuadrada, para que no modifique las dimensiones de la que multiplica. Resulta que la matriz que funciona como neutro por la derecha para una matriz de orden $m \times n$ es una matriz de orden $n \times n$ que llamamos $I_{(n)}$ y definimos como sigue.

$$I_{(n)} = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix}, \text{ es decir } I_{(n)ij} = \begin{cases} 1 & \text{si } i = j, \\ 0 & \text{si } i \neq j. \end{cases}$$

Verificamos que la primera columna de $I_{(n)}$ no hace más que replicar la primera columna de la matriz que multiplica, cuando lo hace por la derecha (ojo que por la izquierda ni siquiera se va a poder multiplicar si $m \neq n$), la segunda replica la segunda columna, y etc.

Por la izquierda se deberá usar $I_{(m)}$ (Ejercicio: verificarlo).

Estudiemos ahora la asociatividad, como pretendemos que es una propiedad verdadera, debemos demostrarla algebraicamente, ya no son válidos los contra-ejemplos. Tomemos A de orden $m \times n$, B de orden $n \times l$, y C de orden $l \times k$.

$$\begin{aligned}
(A(BC))_{ij} &= \sum_{r=1}^n A_{ir}(BC)_{rj} \\
&= \sum_{r=1}^n A_{ir} \left(\sum_{t=1}^l B_{rt} C_{tj} \right) \\
&= \sum_{r=1}^n \sum_{t=1}^l (A_{ir} B_{rt} C_{tj})
\end{aligned}$$

$$\begin{aligned}
((AB)C)_{ij} &= \sum_{t=1}^l (AB)_{it} C_{tj} \\
&= \sum_{t=1}^l \left(\sum_{r=1}^n A_{ir} B_{rt} \right) C_{tj} \\
&= \sum_{t=1}^l \sum_{r=1}^n (A_{ir} B_{rt} C_{tj}) \\
&= \sum_{r=1}^n \sum_{t=1}^l (A_{ir} B_{rt} C_{tj})
\end{aligned}$$

La distributividad usa las mismas técnicas. Supongamos que D es de orden $n \times l$.

$$\begin{aligned}
(A(B+D))_{ij} &= \sum_{r=1}^n A_{ir}(B+D)_{rj} \\
&= \sum_{r=1}^n A_{ir}(B_{rj} + D_{rj}) \\
&= \sum_{r=1}^n (A_{ir}B_{rj} + A_{ir}D_{rj}) \\
&= \sum_{r=1}^n A_{ir}B_{rj} + \sum_{r=1}^n A_{ir}D_{rj} \\
&= (AB)_{ij} + (AD)_{ij}
\end{aligned}$$

En resumen, si A es de orden $m \times n$, B, D es de orden $n \times l$, C es de orden $l \times k$, entonces

- $(AB)C = A(BC)$
- $AI_{(n)} = A$
- $A(B+D) = AB + AD$
- $I_{(m)}A = A$

Como tenemos neutro multiplicativo podemos definir la inversa multiplicativa de una matriz (aunque no todas las matrices son invertibles, como veremos).

Antes de definirlo, volvamos a las transformaciones lineales, ya que la inversa multiplicativa de una matriz A asociada a una transformación $T : \mathbb{K}^m \rightarrow \mathbb{K}^n$, debería estar asociada a la transformación inversa de T .

Por lo tanto solo las matrices asociadas a transformaciones invertibles pueden ser invertibles. Si recordamos de Álgebra I que una función es invertible si y solo si es inyectiva y sobreyectiva, y si recordamos de la semana pasada que una transformación lineal $T : \mathbb{K}^m \rightarrow \mathbb{K}^n$ es sobreyectiva e inyectiva si y solo si $n(T) = 0$ y $r(T) = \dim(\mathbb{K}^n) = n$, y que $n(T) + r(T) = \dim(\mathbb{K}^m) = m$, concluimos que solo se puede ser invertible si $r(T) = m = n$ y $n(T) = 0$.

Así que si vamos a hablar de la inversa multiplicativa de una matriz A , necesitamos que A sea una matriz cuadrada.

Definición 5. Una matriz B de orden $m \times m$ es la inversa de una matriz A de orden $m \times m$ si y solo si se cumple:

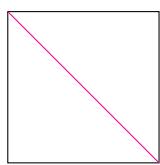
$$AB = I_{(m)} \text{ y } BA = I_{(m)}.$$

En general obtener la inversa de una matriz es trabajoso, incluso decidir si un matriz es invertible o no, requiere trabajo, aprenderemos a hacerlo durante el semestre.

Taxonomía de matrices cuadradas

En el mundo de las matrices cuadradas, es decir, las matrices de orden $n \times n$ (igual ancho que alto), hay tipos particulares de matrices que facilitan algunos cálculos, estos tipos se refieren a la “forma” que tiene la matriz.

En una matriz llamamos *diagonal principal* a la diagonal que parte de la esquina superior izquierda y termina en la esquina inferior derecha. Corresponde a los coeficientes de coordenadas iguales.



$$A = \begin{pmatrix} A_{11} & A_{12} & \cdots & A_{1n} \\ A_{21} & A_{22} & \cdots & A_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ A_{n1} & A_{n2} & \cdots & A_{nn} \end{pmatrix}$$

Matriz Diagonal

Decimos que una matriz es *diagonal* si es cuadrada y fuera de la diagonal principal todos sus coeficientes son 0. Es decir, la matriz A es diagonal si y solo si es cuadrada y cumple

$$\forall i \neq j, \quad A_{ij} = 0.$$

Decida cuáles de las siguientes matrices son diagonales:

$$\begin{pmatrix} 0 & -1 \\ 2 & 0 \end{pmatrix} \quad \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 \\ 0 & 0 & -10,5 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad \begin{pmatrix} 1 & 0 & 0 \\ 0 & 3 & 0 \end{pmatrix} \quad \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

La matriz identidad también es una matriz diagonal

Matriz Triangular Superior

Decimos que una matriz es *triangular superior* si bajo la diagonal principal todos sus coeficientes son 0. Es decir, la matriz A es triangular superior si y solo si cumple

$$\forall i > j, A_{ij} = 0.$$

Decida cuáles de las siguientes matrices son triangulares superior:

$$\begin{pmatrix} 0 & -1 \\ 0 & 3 \end{pmatrix} \quad \begin{pmatrix} 0 & -100 & 0 & 2 \\ 0 & 3 & 1 & 0 \\ 0 & 0 & -10,5 & 0 \\ 0 & 0 & 1 & 1 \end{pmatrix} \quad \begin{pmatrix} 1 & 0 & 3 \\ 0 & 0 & 1 \end{pmatrix} \quad \begin{pmatrix} 1 & 0 & -2 \\ 0 & 10 & -5 \\ 0 & 0 & 0 \end{pmatrix}$$

Matriz Triangular Inferior

De manera análoga se dice que una matriz es *triangular inferior* si *bajo* sobre la diagonal principal todos sus coeficientes son 0. Es decir, la matriz A es triangular inferior si y solo si cumple

$$\forall i < j, A_{ij} = 0.$$

Observamos que solo las matrices diagonales pueden ser triangular inferior y superior al mismo tiempo.

Matriz Simétrica

Decimos que una matriz es *simétrica* si es simétrica respecto a la diagonal principal. Es decir, la matriz A es simétrica si y solo si cumple

$$\forall i, j, A_{ij} = A_{ji}.$$

Decida cuáles de las siguientes matrices son simétricas:

$$\begin{pmatrix} 0 & -1 \\ -1 & 3 \end{pmatrix} \quad \begin{pmatrix} 0 & 3 & -4 & 2 \\ 3 & 1 & 1 & 0 \\ -4 & 1 & 0,2 & 1 \\ 2 & 0 & 1 & 1 \end{pmatrix} \quad \begin{pmatrix} 1 & 0 & 3 \\ 3 & 0 & 1 \end{pmatrix} \quad \begin{pmatrix} 1 & 0 & -2 \\ -2 & 10 & -5 \\ -5 & 0 & 0 \end{pmatrix}$$

Observamos además que toda matriz diagonal es simétrica.

Matriz Antisimétrica

Decimos que una matriz es *antisimétrica* si cada coeficiente es el opuesto aditivo de su coeficiente en posición simétrica respecto a la diagonal principal. Es decir, la matriz A es antisimétrica si y solo si cumple

$$\forall i, j, A_{ij} = -A_{ji}.$$

Notamos que entonces $A_{ii} = -A_{ii}$, lo cual implica que $A_{ii} = 0$ para todo i .

Decida cuáles de las siguientes matrices son antisimétricas:

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad \begin{pmatrix} 0 & 3 & -4 & 2 \\ -3 & 0 & 1 & 0 \\ 4 & -1 & 0 & 1 \\ -2 & 0 & -1 & 0 \end{pmatrix} \quad \begin{pmatrix} 0 & 0 & -2 \\ 0 & 0 & -5 \\ 2 & 5 & 0 \end{pmatrix} \quad \begin{pmatrix} 0 & 3 & 0 \\ -2 & 0 & 2 \\ 0 & -3 & 0 \end{pmatrix}$$

Observamos que sólo la matriz nula puede ser simétrica y antisimétrica.

Para ayudarnos a entender mejor estos últimos tipos de matrices definimos un operador en el conjunto de las matrices, que lo que hace es “girar” la matriz respecto a la diagonal principal, visualizaremos así más fácilmente las simetrías.

Dada una matriz cualquiera $A \in \mathcal{M}_{m \times n}(\mathbb{R})$, se define su **traspuesta** por $A^\top \in \mathcal{M}_{n \times m}(\mathbb{R})$ como sigue:

$$\forall i, \forall j, (A^\top)_{ij} = A_{ji}$$

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \\ \pi & 0 & \sqrt{3} \end{pmatrix} \quad \rightarrow \quad A^\top = \begin{pmatrix} 1 & 4 & 7 & \pi \\ 2 & 5 & 8 & 0 \\ 3 & 6 & 9 & \sqrt{3} \end{pmatrix}$$

Propiedad 2. ■ A es simétrica si y solo si $A^\top = A$

- A es antisimétrica si y solo si $A^\top = -A$
- La traspuesta de una matriz triangular inferior es triangular superior.
- Toda matriz diagonal es simétrica.
- La traspuesta de una matriz columna, es una matriz fila.

Propiedad 3. ■ *La suma de matrices simétricas es simétrica.*

- *La suma de matrices antisimétricas es antisimétrica*
- *La suma (y la multiplicación) de matrices triangulares inferiores es triangular inferior.*
- *La inversa de una matriz diagonal e invertible A es la matriz diagonal en que cada coeficiente de la diagonal es el inverso multiplicativo del respectivo coeficiente de A .*
- *Si B se multiplica por una matriz diagonal A por la derecha, entonces cada una de sus columnas resulta multiplicada por el coeficiente de la misma columna de A .*
- *Si B se multiplica por una matriz diagonal A por la izquierda, entonces cada una de sus filas resulta multiplicada por el coeficiente de la misma fila de A .*

Sistemas de ecuaciones lineales

En esta cápsula veremos lo que es un sistema lineal, y cómo este se representa de manera matricial. Entenderemos lo que significa que dos sistemas sean equivalentes, y pasando de un sistema equivalente a otro, aprenderemos a resolver estos sistemas mediante el método de Gauss.

Definición 1. *Un sistema lineal de m ecuaciones y n incógnitas es un conjunto de m ecuaciones lineales en las cuales participan a lo más n incógnitas. Se suele ordenar de manera que los términos con variables queden del lado izquierdo, y los términos independientes del lado derecho:*

$$\left| \begin{array}{cccccc} a_{11}x_1 & + & a_{12}x_2 & + & \cdots & + & a_{1n}x_n = b_1 \\ a_{21}x_1 & + & a_{22}x_2 & + & \cdots & + & a_{2n}x_n = b_2 \\ \vdots & & \vdots & & \ddots & & \vdots \\ a_{m1}x_1 & + & a_{m2}x_2 & + & \cdots & + & a_{mn}x_n = b_m \end{array} \right|$$

donde los coeficientes a_{ij} y b_i son constantes dadas en un cuerpo \mathbb{K} y x_1, \dots, x_n denotan las incógnitas.

El conjunto solución del sistema es el conjunto de vectores $(x_1, \dots, x_n) \in \mathbb{K}^n$ que hacen que las m ecuaciones se cumplan simultáneamente.

Este conjunto puede ser vacío si no hay solución, tener un solo elemento, si hay unicidad o tener muchos elementos.

Considerando la multiplicación de matrices, es posible representar todo sistema lineal de forma matricial, en que colocamos las incógnitas en un vector columna, como sigue.

$$\left(\begin{array}{cccc} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{array} \right) \left(\begin{array}{c} x_1 \\ x_2 \\ \vdots \\ x_n \end{array} \right) = \left(\begin{array}{c} b_1 \\ b_2 \\ \vdots \\ b_m \end{array} \right)$$

Lo cual lo resumimos como $Ax = b$, si es que A es la matriz de coeficientes, x el vector con las incógnitas, y b es el vector con los términos independientes (el lado derecho).

Dado un sistema lineal $Ax = b$, decimos que:

- Es *homogéneo* si $b = \Theta$.
- Es *incompatible* si su conjunto solución es el vacío.
- Es *compatible* si tiene alguna solución.
 - Es *compatible determinado* si su solución es única.
 - Es *compatible indeterminado* si tiene más de una solución.

Recurramos un poco a la materia de Transformaciones Lineales para entender con qué se relaciona esta clasificación. Recordemos que una matriz de orden $m \times n$ a coeficientes en un cuerpo \mathbb{K} está asociada a una transformación $T : \mathbb{K}^n \rightarrow \mathbb{K}^m$ definida como sigue: $T(v) = Av$. Observamos entonces que:

$$Ax = b \Leftrightarrow T(x) = b$$

El conjunto de soluciones del sistema $Ax = b$ es entonces el conjunto de preimágenes del vector b a través de la transformación T . Obtenemos entonces:

$$Ax = b \text{ es incompatible } \Leftrightarrow b \notin Im(T)$$

$$Ax = b \text{ es compatible } \Leftrightarrow b \in Im(T)$$

Nuevamente entonces, determinar para qué vectores b el sistema $Ax = b$ tiene solución equivale a calcular el subespacio imagen de T .

Por su parte, la determinación del sistema está relacionada con el núcleo de T . Esto es evidente cuando $b = \Theta$, pues entonces

$$Ax = \Theta \Leftrightarrow T(x) = \Theta \Leftrightarrow x \in \text{Ker}(T).$$

Es decir, el conjunto de soluciones del sistema homogéneo coincide con el núcleo de T .

Ahora bien, cuando no es homogéneo, pero es compatible tenemos algo que nos sirve de todas formas. Supongamos que tenemos v una solución del sistema, es decir, tenemos que $Av = b$. Entonces el conjunto de soluciones del sistema será:

$$\begin{aligned}
\{x \in \mathbb{K}^n : Ax = b\} &= \{x : Ax = Av\} \\
&= \{x : Ax - Av = \Theta\} \quad \text{multiplicación de matriz por vector es distributiva} \\
&= \{x : A(x - v) = \Theta\} \\
&= \{x : x - v \in \text{Ker}(T)\} \\
&= \{y + v : y = x - v \in \text{Ker}(T)\} \\
&= \{y + v : y \in \text{Ker}(T)\} \\
&= \{v\} + \text{Ker}(T)
\end{aligned}$$

En otras palabras, toda solución de $Ax = b$ es resultado de una solución de $Ax = \Theta$ más la solución v con la que partimos. Esto nos lleva a concluir que la solución de $Ax = b$ es única si y solo si la solución de $Ax = \Theta$ es única. Por lo tanto:

$Ax = b$ es compatible determinado $\Leftrightarrow T$ es inyectiva y $b \in \text{Im}(T) \Leftrightarrow \text{Ker}(T) = \{\Theta\}$ y $b \in \text{Im}(T)$

$Ax = b$ es compatible indeterminado $\Leftrightarrow T$ no es inyectiva y $b \in \text{Im}(T) \Leftrightarrow \text{Ker}(T) \neq \{\Theta\}$ y $b \in \text{Im}(T)$

Esto nos lleva al primer teorema interesante.

Teorema 1. *Si un sistema lineal tiene más incógnitas que ecuaciones, entonces en caso de ser compatible será compatible indeterminado.*

En particular, todo sistema lineal homogéneo con más incógnitas que ecuaciones es indeterminado.

Demostración. Supongamos que hay m ecuaciones y n incógnitas. La matriz asociada al sistema será una matriz de orden $n \times m$, y estará asociada a una transformación lineal $T : \mathbb{K}^n \rightarrow \mathbb{K}^m$. Si hay más incógnitas que ecuaciones quiere decir que $m < n$. Es decir, el conjunto de llegada de T tiene dimensión inferior al conjunto de partida de T . Por el teorema núcleo imagen $\dim(\text{Ker}(T)) \geq n - m > 0$, T no puede ser inyectiva. Por lo tanto el sistema no puede ser determinado. \square

Para no referirnos siempre a las propiedades de la transformación asociada a la matriz asociada al sistema, generalicemos los conceptos de nulidad y rango a las matrices.

Definición 2. *Dada una matriz $A \in \mathcal{M}_{n \times m}(\mathbb{K})$, se define:*

- la nulidad de A como la dimensión del núcleo de la transformación asociada a A , y
- el rango de A como la dimensión de la imagen de la transformación asociada a A .

Método de eliminación de Gauss

Vamos a estar particularmente interesados en conocer el conjunto solución del sistema. Para ello vamos a realizar modificaciones en el sistema que lo van a simplificar sucesivamente sin cambiar su conjunto solución, lo transformarán en un sistema *equivalente* cada vez más simple.

Del liceo sabemos que si reemplazamos una ecuación por la suma de esta y otra de las ecuaciones, o bien si ponderamos una ecuación completa por una constante, no afectamos la solución del sistema. Tampoco lo hacemos al cambiar el orden de aparición de las ecuaciones en el sistema. Para ver esto se puede razonar de manera conjuntista.

Si tenemos una solución del sistema, ésta hace que todas las ecuaciones se cumplan, entonces cuando hacemos sumas entre las ecuaciones o ponderamos ambos lados de una de estas, las igualdades siguen satisfaciéndose, por lo tanto la solución sigue siendo solución del nuevo sistema.

La otra dirección es más difícil verla, interesa que las operaciones que hagamos no *agreguen* soluciones al sistema, pues esto modificaría el conjunto solución y no es lo que deseamos. Evitamos esto haciendo siempre operaciones que sean *reversibles*, es decir, que puedan *deshacerse*.

Por ejemplo, si multiplicamos por 0 perdemos la ecuación original, multiplicar por 0 no es reversible. Pero multiplicar por cualquier otra constante sí lo es pues para recuperar la ecuación original basta multiplicar por el inverso multiplicativo de la constante.

Si reemplazo dos ecuaciones por la suma de estas, pierdo información, ya no puedo recuperar las dos ecuaciones originales, pero si reemplazo solo una de las ecuaciones por la suma de ella con otra, sí puedo, pues al disponer de la otra, basta restarla a la ecuación obtenida para recuperar la original.

Si solo hacemos operaciones reversibles, el sistema original también se podrá obtener a partir del nuevo sistema a través de operaciones reversibles, así las soluciones del nuevo sistema también serán soluciones del sistema original. En resumen, ambos sistemas tendrán igual conjunto solución.

Matriz ampliada del sistema

En un sistema $Ax = b$, lo relevante es la matriz A y el lado derecho b . Creamos entonces una nueva matriz, llamada *matriz ampliada* del sistema, que consiste en colocar A al lado de b separados por una barra vertical, como sigue: $(A|b)$,

$$\left(\begin{array}{cccc|c} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} & b_m \end{array} \right) \in \mathcal{M}_{m \times n+1}$$

Cada fila de esta matriz representa una ecuación del sistema.

Operaciones elementales de fila

Las operaciones reversibles entre ecuaciones de las que hablamos se traducen en operaciones entre las filas de la matriz ampliada, y ellas son:

- intercambiar dos filas,
- ponderar una fila completa por un escalar no nulo,
- sumar a una fila, otra fila ponderada por un escalar.

Con éstas modificamos el sistema, pero no su conjunto solución.

Usando estas operaciones intentaremos hacer que la matriz ampliada quede *escalonada*, esto es, a *groso modo*, que cada fila tenga más 0s que la anterior, para así finalmente poder encontrar la solución mediante *substitución regresiva*.

Escalonamiento

En una fila dada, llamamos *pivote* al primer coeficiente no nulo que encontramos en la fila partiendo de izquierda a derecha. Si es una fila completamente nula, entonces no tiene pivote.

Consideramos el conjunto de los pivotes de todas las filas.

Una matriz se dice *escalonada* si cada columna tiene a lo más un pivote.

$$\left(\begin{array}{ccc} 0 & (2) & 3 \\ (1) & 0 & 4 \\ 0 & 0 & 0 \\ (-1) & 0 & 0 \end{array} \right) \quad \text{Matriz no escalonada}$$
$$\left(\begin{array}{ccc} 0 & (2) & 3 \\ (1) & 0 & 4 \\ 0 & 0 & 0 \\ 0 & 0 & (4) \end{array} \right) \quad \text{Matriz escalonada}$$

Si la matriz está escalonada, entonces, al volver la forma de sistema lineal, podremos despejar la variable correspondiente a cada pivote en cada ecuación. Partiendo por las ecuaciones cuyo pivote está más a la derecha podremos ir reemplazando en las demás ecuaciones hasta haber despejado todas las variables correspondientes a un pivote, y las demás quedarán como parámetro.

Para lograr escalar la matriz, se parte con una fila que tenga un pivote lo más a la izquierda posible, y se usa este pivote para *anular* a los demás pivotes que estén en la misma columna.

Luego se repite el proceso con el siguiente pivote hacia la derecha, y así sucesivamente hasta que cada comuna termine con un solo pivote.

Será mejor ver todo esto con un ejemplo.

Ejemplo 1. Sistema lineal.

$$\left| \begin{array}{l} 2\alpha + 4\beta = 4 \\ -\alpha - 3\beta = 1 \\ 4\alpha + \beta = 29 \\ 5\beta = -15 \end{array} \right.$$

Forma matricial.

$$\left(\begin{array}{cc} 2 & 4 \\ -1 & -3 \\ 4 & 1 \\ 0 & 5 \end{array} \right) \left(\begin{array}{c} \alpha \\ \beta \end{array} \right) = \left(\begin{array}{c} 4 \\ 1 \\ 29 \\ -15 \end{array} \right).$$

Método de Gauss.

$$\left(\begin{array}{cc|c} \textcircled{2} & 4 & 4 \\ \textcircled{-1} & -3 & 1 \\ \textcircled{4} & 1 & 29 \\ 0 & \textcircled{5} & -15 \end{array} \right) \xrightarrow{\begin{array}{l} f_2 + \frac{1}{2}f_1 \\ f_3 - 2f_1 \end{array}} \sim \left(\begin{array}{cc|c} \textcircled{2} & 4 & 4 \\ 0 & \textcircled{-1} & 3 \\ 0 & \textcircled{-7} & 21 \\ 0 & \textcircled{5} & -15 \end{array} \right) \xrightarrow{\begin{array}{l} f_3 - 7f_2 \\ f_4 + 5f_2 \end{array}} \sim \left(\begin{array}{cc|c} \textcircled{2} & 4 & 4 \\ 0 & \textcircled{-1} & 3 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{array} \right).$$

Sistema lineal asociado a la matriz ampliada escalonada.

$$\left. \begin{array}{l} 2\alpha + 4\beta = 4 \\ -\beta = 3 \\ 0 = 0 \\ 0 = 0 \end{array} \right| \rightarrow \begin{array}{l} \beta = -3 \\ \alpha = \frac{4-4\beta}{2} = \frac{4+12}{2} = 8 \end{array} \quad \begin{array}{l} \text{Cjto. Solución:} \\ \{(8, -3)\} \\ \text{Sistema compatible determinado} \end{array}$$

Ejemplo 2. Sistema lineal.

$$\left| \begin{array}{l} -x + z = 0 \\ 2x + y - z = 0 \\ y + z = 0 \end{array} \right.$$

Forma matricial.

$$\left(\begin{array}{ccc} -1 & 0 & 1 \\ 2 & 1 & -1 \\ 0 & 1 & 1 \end{array} \right) \left(\begin{array}{c} x \\ y \\ z \end{array} \right) = \left(\begin{array}{c} 0 \\ 0 \\ 0 \end{array} \right)$$

Método de Gauss.

$$\left(\begin{array}{ccc|c} \textcircled{-1} & 0 & 1 & \\ \textcircled{2} & 1 & -1 & \\ 0 & \textcircled{1} & 1 & \end{array} \right) \xrightarrow{f_2 + 2f_1} \left(\begin{array}{ccc|c} \textcircled{-1} & 0 & 1 & \\ 0 & \textcircled{1} & 1 & \\ 0 & \textcircled{1} & 1 & \end{array} \right) \xrightarrow{f_3 - f_2} \left(\begin{array}{ccc|c} \textcircled{-1} & 0 & 1 & \\ 0 & \textcircled{1} & 1 & \\ 0 & 0 & 0 & \end{array} \right)$$

Sistema lineal asociado a la matriz ampliada escalonada.

$$\begin{array}{lcl} \begin{array}{l|l} -x + z = 0 \\ y + z = 0 \\ 0 = 0 \end{array} & \rightarrow z \text{ libre} & \text{Cjto. Solución:} \\ & \rightarrow y = -z & \{(z, -z, z) : z \in \mathbb{R}\} \\ & \rightarrow x = z & \text{Sistema compatible indeterminado} \end{array}$$

Ejemplo 3. Sistema lineal.

$$\begin{array}{lcl} x + 2y + z & = & 1 + i \\ 2x - y + z & = & i \\ 3x + y + 2z & = & 1 \end{array}$$

Forma matricial.

$$\left(\begin{array}{ccc} 1 & 2 & 1 \\ 2 & -1 & 1 \\ 3 & 1 & 2 \end{array} \right) \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 1+i \\ i \\ 1 \end{pmatrix}$$

Método de Gauss.

$$\begin{array}{lcl} \left(\begin{array}{ccc|c} \textcircled{1} & 2 & 1 & 1+i \\ \textcircled{2} & -1 & 1 & i \\ \textcircled{3} & 1 & 2 & 1 \end{array} \right) & \xrightarrow{f_2 - 2f_1} & \left(\begin{array}{ccc|c} \textcircled{1} & 2 & 1 & 1+i \\ 0 & \textcircled{-5} & -1 & -2-i \\ 0 & \textcircled{-5} & -1 & -2-3i \end{array} \right) \\ & \xrightarrow{f_3 - f_2} & \left(\begin{array}{ccc|c} \textcircled{1} & 2 & 1 & 1+i \\ 0 & \textcircled{-5} & -1 & -2-i \\ 0 & 0 & 0 & -2i \end{array} \right) \end{array}$$

Sistema lineal asociado a la matriz ampliada escalonada.

$$\begin{array}{lcl} x + 2y + z & = & 1+i \\ -5y - z & = & -2-i \\ 0 & = & -2i \end{array}$$

Pero

$$0 \neq -2i$$

Cjto. Solución: \emptyset

Sistema incompatible

Podemos distinguir el conjunto solución volviendo al sistema lineal desde la matriz escalonada, pues allí podemos evidenciar si hay contradicción (conjunto vacío), se puede resolver explícitamente o si quedan parámetros libres.

Pero si miramos con atención nos damos cuenta que no es necesario volver al sistema para darnos cuenta que tipo de conjunto solución obtendremos.

- El caso más evidente es cuando hay contradicción. Esto ocurre cuando en la matriz hay **filas nulas enfrentadas a valores no nulos** en el lado derecho. En ese caso el sistema no tiene solución, es **incompatible**.
- Otro caso claro es cuando hay una sola solución. Esto ocurre cuando queda una escalera perfecta, cuando **hay tantas incógnitas como filas no nulas** después de escalar. En ese caso el sistema es **compatible determinado**.
- Queda el caso en que quedan “escalones anchos”, lo cual se evidencia en que aparecen variables que no se podrán despejar, serán *libres*. Esto ocurre cuando **hay más incógnitas que filas no nulas** después de escalar, en ese caso el sistema es **compatible indeterminado**.

Una vez que hemos aclarado de qué tipo de conjunto, de todas formas tendremos que volver al sistema para encontrar la solución, si es que es un sistema compatible y si es que la solución nos interesa.

Interesante es también destacar que el lado derecho solo sirve para saber si el sistema es compatible o no, pero no influye en si este es determinado o no, pues esto último solo depende del número de filas no nulas de la matriz y el número de incógnitas.

Rango de una matriz

En lo que sigue vamos a exportar a las matrices las definiciones que vimos para transformaciones lineales. Para ello adoptemos una notación.

Definición 1. *Dada una matriz $A \in \mathcal{M}_{m \times n}(\mathbb{K})$, definimos la transformación lineal a la cual está directamente asociada por:*

$$T_A : \mathbb{K}^n \rightarrow \mathbb{K}^m$$

$$T_A(v) = Av$$

Exportamos entonces los conceptos de rango y nulidad al mundo de las matrices de la manera más natural.

Definición 2. *Dada una matriz $A \in \mathcal{M}_{m \times n}(\mathbb{K})$, definimos su nulidad $\eta(A)$ y su rango $r(A)$ como sigue:*

$$\eta(A) = \eta(T_A) \quad r(A) = r(T_A).$$

Resulta que la nulidad de $\eta(A)$ es la dimensión del espacio solución del sistema lineal homogéneo que define la matriz A , en otras palabras es lo que llamamos “grados de libertad” del sistema de A .

Por su parte, el rango de $r(A)$ tiene también una interesante interpretación.

Sea $\{e_1, \dots, e_n\}$ la base canónica de \mathbb{K}^n . Del Propiedad 2 de la Semana 6 obtenemos lo siguiente.

$$\begin{aligned} \text{Im}(T_A) &= \langle \{T_A(e_1), \dots, T_A(e_n)\} \rangle \quad \text{por definición de } T_A \\ &= \langle \{Ae_1, \dots, Ae_n\} \rangle \quad \text{por la forma de } e_i \text{ y por cómo se multiplica} \\ &= \langle \{\text{Col}_1(A), \dots, \text{Col}_n(A)\} \rangle \end{aligned}$$

Resulta que $\text{Im}(T_A)$ es entonces igual al espacio generado por las columnas de A . Tal espacio es llamado *Espacio Columna de A* .

Vemos entonces que el rango de A es la dimensión del espacio columna de A .

Interesante es ver que pasa con el rango cuando multiplicamos dos matrices. Lo cierto es que en general puede cambiar mucho, pero hay un caso en que no cambia y es cuando multiplicamos por una matriz invertible. Antes de demostrar esto, enunciemos algo más básico y directo, que ya habíamos usado cuando introdujimos el concepto de invertible en la Semana 7.

Propiedad 1. Una matriz $B \in \mathcal{M}_{m \times m}(\mathbb{K})$ es invertible si y solo si T_B es invertible.

Recordemos también, que definimos la multiplicación de matrices de manera que se tuviera la siguiente propiedad.

Propiedad 2. Dadas dos matrices $A \in \mathcal{M}_{m \times n}(\mathbb{K})$ y $B \in \mathcal{M}_{l \times m}(\mathbb{K})$ se cumple que:

$$T_{BA} = T_B \circ T_A.$$

Con esto en mano podemos estudiar qué pasa con el rango cuando multiplicamos por una matriz invertible.

Teorema 1. Si $B \in \mathcal{M}_{m \times m}(\mathbb{K})$ es una matriz invertible y $A \in \mathcal{M}_{m \times n}(\mathbb{K})$ una matriz cualquiera, entonces se cumple:

$$r(BA) = r(A).$$

Demostración Partamos avanzando en cómo es el espacio imagen de T_{BA} .

$$\begin{aligned} \text{Im}(T_{BA}) &= T_{BA}(\mathbb{K}^n) \\ &= T_B(T_A(\mathbb{K}^n)) \\ &= T_B(\text{Im}(T_A)) \end{aligned}$$

Entonces consideremos una base de $\text{Im}(T_A)$ (la cual podemos obtener por ejemplo a partir de las columnas de A , eliminando columnas que sean c.l. de las demás). Digamos que $k = r(A)$ y que la base está dada por los vectores $w_1, \dots, w_k \in \mathbb{K}^m$. Con esto en mano podemos proseguir con nuestro cálculo:

$$T_B(\text{Im}(T_A)) = \langle \{T_B(w_1), \dots, T_B(w_k)\} \rangle.$$

Como vimos en el Corolario 2 de la Semana 6, cuando una transformación es inyectiva, respecta la independencia lineal de los conjuntos, así el conjunto $\{T_B(w_1), \dots, T_B(w_k)\}$ es linealmente independiente y $r(BA) = \dim(T_B(\text{Im}(T_A))) = k = r(A)$. \square

Análogamente se preserva el rango al multiplicar por la derecha. La demostración en ese caso es más directa y se deja como ejercicio a el o la lectora.

Teorema 2. Si $C \in \mathcal{M}_{n \times n}(\mathbb{K})$ es una matriz invertible y $A \in \mathcal{M}_{m \times n}(\mathbb{K})$ una matriz cualquiera, entonces se cumple:

$$r(AC) = r(A).$$

Con todo esto podemos obtener una caracterización de la invertibilidad de matrices a través del rango, basándonos en la Observación 1 de la Semana 6. Teniendo en cuenta que solo las matrices cuadradas pueden ser invertibles, se cumple lo siguiente.

Corolario 1. Dada $A \in \mathcal{M}_{m \times m}(\mathbb{K})$, se cumple que:

$$A \text{ es invertible si y solo si } r(A) = m \text{ si y solo si } \eta(A) = 0.$$

El concepto de subespacio imagen y rango también es interesante en el análisis de sistemas de ecuaciones lineales. Para verlo observamos primero lo siguiente, si $v = (x_1, x_2, \dots, x_n)$, entonces:

$$Av = x_1 Col_1(A) + x_2 Col_2(A) + \cdots + x_n Col_n(A) \in \text{Im}(T_A)$$

Así tenemos la siguiente cadena de resultados

$$\begin{aligned} Av = b \text{ es compatible} &\quad \text{si y solo si } b \in \text{Im}(T_A) \\ &\quad \text{si y solo si } \langle \{Col_1(A), \dots, Col_n(A), b\} \rangle = \langle \{Col_1(A), \dots, Col_n(A)\} \rangle \\ &\quad \text{si y solo si } \dim(\langle \{Col_1(A), \dots, Col_n(A), b\} \rangle) = \dim(\langle \{Col_1(A), \dots, Col_n(A)\} \rangle) \\ &\quad \text{si y solo si } r(A) = r(A|b) \end{aligned}$$

Por otra parte, en caso que $Av = b$ fuera compatible, que el número de grados de libertad del conjunto de solución puede ser definido como la dimensión del núcleo de A , y este es igual a $\eta(A) = n - r(A)$, donde n es el número de columnas de A , es decir, es igual al número de incógnitas del sistema:

El número de grados de libertad del conjunto solución es: $\eta(A) = \#\text{incógnitas} - r(A)$

Por lo tanto,

$$Av = b \text{ es determinado si y solo si } \eta(A) = 0.$$

Más aún, si el sistema $Av = b$ es compatible determinado entonces tenemos que la única solución que existe está dada por la fórmula:

$$v = A^{-1}b.$$

Matrices elementales

Aprendimos que para resolver sistemas lineales conviene hacer “operaciones elementales de filas”. Resulta que estas operaciones pueden hacerse mediante multiplicaciones por matrices por la izquierda. Veamos cómo.

Sea $A \in \mathcal{M}_{m \times n}(\mathbb{K})$ una matriz a la cual deseamos escalar.

- 1. Ponderación de una fila por un escalar no nulo.** Si concebimos una matriz de $m \times m$ igual a la matriz identidad de orden m , $I_{(m)}$, pero con el i -ésimo valor de la diagonal igual a un escalar dado λ , obtenemos una matriz como la que sigue, que llamaremos $E_i(\lambda)$.

$$E_i(\lambda) = \begin{pmatrix} 1 & 0 & \cdots & & 0 \\ 0 & \ddots & \ddots & & \vdots \\ \vdots & \ddots & 1 & & \\ & & & \lambda & \\ & & & & 1 & \ddots \\ & & & & & \ddots & \ddots & 0 \\ 0 & \cdots & & & 0 & 1 \end{pmatrix}; \quad E_i(\lambda)_{jj} = \begin{cases} 1 & \text{si } j \neq i \\ \lambda & \text{si } j = i \end{cases}$$

No es difícil ver que el producto $E_i(\lambda)A$ es igual a la matriz A pero con la i -ésima fila ponderada por λ .

2. **Intercambio de filas.** Definamos ahora una matriz parecida a la identidad pero con dos filas intercambiadas

$$P_{ij} = \begin{pmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & & 0 & \dots & \dots & \dots & 1 \\ & & & \vdots & 1 & & & \vdots \\ & & & & & 1 & & \\ & & & & & & 1 & \\ & & & & & & & 1 \\ & & & & & & & & \ddots \\ & & & & & & & & & 1 \end{pmatrix}; \quad (P_{ij})_{ii} = (P_{ij})_{jj} = 0; \quad (P_{ij})_{ij} = (P_{ij})_{ji} = 1$$

$i \uparrow \qquad \qquad j \uparrow$

Resulta que el producto $P_{ij}A$ es igual a la matriz A pero con la i -ésima y la j -ésima filas intercambiadas.

3. **Suma de una fila ponderada por un escalar a otra fila.** Finalmente dados $j > i$, definimos una matriz $E_{ij}(\lambda)$, similar a la matriz identidad, pero con un solo coeficiente diferente: el coeficiente ji el cual es igual a λ .

$$E_{ij}(\lambda) = \begin{pmatrix} 1 & & & & & & & \\ & \ddots & & & & & & \\ & & 1 & & & & & \\ & & & \vdots & & & & \\ 0 & \dots & \lambda & \dots & & 1 & & 1 \\ & & & & & & 1 & \\ 0 & \dots & 0 & \dots & \dots & \dots & \dots & 0 & 1 \\ & & & i \uparrow & & j \uparrow & & & \vdots \end{pmatrix}$$

Resulta que el producto $E_{ij}(\lambda)A$ es igual a la matriz A pero donde a la j -ésima fila se le ha sumado la fila i ponderada por λ .

Estas matrices nos ayudan a demostrar por qué el método de Gauss funciona.

Propiedad 3. Si $\lambda \in \mathbb{K}$ no es nulo, entonces las matrices elementales, $E_i(\lambda)$, P_{ij} y $E_{ij}(\lambda) \in \mathcal{M}_{m \times m}(\mathbb{K})$ son todas invertibles.

Demostración. Basta proveer su inversa:

$$1. (E_i(\lambda))^{-1} = E_i(\lambda^{-1}).$$

$$2. (P_{ij})^{-1} = P_{ij}.$$

$$3. (E_{ij}(\lambda))^{-1} = E_{ij}(-\lambda).$$

□

Notamos que aplicar las operaciones de fila a la matriz ampliada es equivalente a multiplicar por la correspondiente matriz elemental a cada lado de la ecuación:

$$\begin{aligned} Av = b &\quad /E_{ij}(\lambda) \cdot () \\ E_{ij}(\lambda)Av = E_{ij}(\lambda)b \end{aligned}$$

Como se trata de operaciones reversibles, no se pierden ni se ganan soluciones.

El Método de Gauss se preocupa además de avanzar de manera ordenada, llenando de 0s debajo de cada pivote, partiendo de los pivotes que están más a la izquierda y avanzando luego uno a uno hacia la derecha.

La ventaja de hacer esto es que los 0s que se coloquen en los pivotes de más a la izquierda no van a ser perturbados por las operaciones de fila posteriores, ya que los pivotes que siguen corresponden, por definición, a filas que tienen solo 0s a la izquierda del pivote, y al sumarlas a las demás filas, no modificarán los primeros coeficientes de estas.

En todo momento, la matriz que obtendremos será igual a A multiplicada por un producto de matrices elementales, es decir, multiplicada por una matriz invertible.

Por lo visto anteriormente, tal matriz tendrá el mismo rango que A . Es decir, ¡¡las operaciones elementales de filas preservan el rango de la matriz!!

Propiedad 4. *Las operaciones elementales de filas no modifican el rango de la matriz sobre la cual se aplican.*

En una matriz escalonada, es evidente que su rango es igual al número de filas no nulas que tenga. Pero este es igual al rango de la matriz original, así, una manera de calcular el rango de una matriz es escalaronarla, para luego contar su número de filas no nulas.

Propiedad 5. *El rango de una matriz es igual al número de filas no nulas que tiene la matriz que resulta al escalaronarla completamente.*

Por lo tanto:

Observación 1. *$Av = b$ es compatible si y solo si el número de filas no nulas de la matriz ampliada $(A|b)$ escalaronada es igual al número de filas no nulas de A escalaronada. Lo cual equivale a que no hayan ecuaciones de la forma $0 = f$, con f constante no nula.*

Atención que para que esto sea cierto, la matriz debe estar *realmente* escalaronada.

Ahora bien, el Método de Gauss también provee un método para invertir matrices invertibles, y las matrices elementales juegan un rol muy interesante en esto.

Método de Gauss-Jordan

En la Semana 7 definimos la *inversa de una matriz* y demostramos que si una matriz tiene inversa es necesario que sea cuadrada. Recordamos aquí la definición dada entonces.

Definición 3. *Una matriz B de orden $m \times m$ es la inversa de una matriz A de orden $m \times m$ si y solo si se cumple:*

$$AB = I_{(m)} \text{ y } BA = I_{(m)}.$$

Para invertir una matriz lo que haremos será aplicarle operaciones de fila hasta obtener una matriz identidad. Como hacer operaciones de fila equivale a multiplicar por una matriz invertible, cuando hayamos llegado a la identidad será debido a que la multiplicación de las operaciones de fila que

hacemos resulta en la matriz Inversa de A, dicho de otra manera, si las operaciones que aplicamos las denotamos por las matrices $E^{(t)}, E^{(t-1)}, \dots, E^{(1)}$ y lo que obtuvimos fue la identidad, entonces tenemos:

$$E^{(t)} E^{(t-1)} \dots E^{(1)} A = I_{(n)}$$

Por lo tanto: $A^{-1} = E^{(t)} E^{(t-1)} \dots E^{(1)}$.

Conviene ir anotando las matrices elementales por las que multiplicaos entonces... o al menos ir anotando su producto. Pero hemos observado que multiplicar por matrices elementales es lo mismo que hacer operaciones de fila. Solo que si ahora solo nos interesa el producto de las matrices elementales por “nada”, esto equivale a hacer el producto de las matrices elementales por la matriz identidad, es decir, lo que podemos hacer es aplicar operaciones de fila a la matriz identidad.

Ejemplo 1. Veamos un ejemplo simple de esto para la matriz $A = \begin{pmatrix} 1 & 2 \\ -1 & 1 \end{pmatrix}$.

$$\left(\begin{array}{cc|cc} 1 & 2 & 1 & 0 \\ -1 & 1 & 0 & 1 \end{array} \right) \xrightarrow{f_2 + f_1} \left(\begin{array}{cc|cc} 1 & 2 & 1 & 0 \\ 0 & 3 & 1 & 1 \end{array} \right)$$

$$\left(\begin{array}{cc|cc} 1 & 2 & 1 & 0 \\ 0 & 3 & 1 & 1 \end{array} \right) \xrightarrow{f_2/3} \left(\begin{array}{cc|cc} 1 & 2 & 1 & 0 \\ 0 & 1 & \frac{1}{3} & \frac{1}{3} \end{array} \right) \xrightarrow{f_1 - 2f_2} \left(\begin{array}{cc|cc} 1 & 0 & \frac{1}{3} & -\frac{2}{3} \\ 0 & 1 & \frac{1}{3} & \frac{1}{3} \end{array} \right)$$

Según nuestros análisis la matriz $B = \begin{pmatrix} \frac{1}{3} & -\frac{2}{3} \\ \frac{1}{3} & \frac{1}{3} \end{pmatrix}$ debiera ser la inversa de A, veamos:

$$AB = \begin{pmatrix} 1 & 2 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} \frac{1}{3} & -\frac{2}{3} \\ \frac{1}{3} & \frac{1}{3} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Funciona.

Esto es lo que se conoce como el Método de Gauss-Jordan, cuando se agregan operaciones de fila que van poniendo 0s arriba de la diagonal, para obtener una matriz identidad.

Ejemplo 2. Veamos otro ejemplo, con $A = \begin{pmatrix} 2 & -3 \\ -4 & 6 \end{pmatrix}$.

$$\left(\begin{array}{cc|cc} 2 & -3 & 1 & 0 \\ -4 & 6 & 0 & 1 \end{array} \right) \xrightarrow{f_2 + 2f_1} \left(\begin{array}{cc|cc} 2 & -3 & 1 & 0 \\ 0 & 0 & 2 & 1 \end{array} \right)$$

¿Qué pasó en este caso? Lo que ocurre es que A no es invertible, por eso el proceso falla.

Traza

Para las matrices cuadradas se define una función muy fácil de calcular y que tiene muy bonitas propiedades, que descubriremos poco a poco en lo que queda del semestre: la *traza*.

Definición 4. Dada una matriz A cuadrada de orden $m \times m$ se define la traza de A , que se denota $\text{tr}(A) \in \mathbb{K}$ como la suma de los elementos de la diagonal de A :

$$\text{tr}(A) = \sum_{i=1}^m A_{ii}$$

Ejemplo 3. Calculemos la traza de las siguientes matrices:

$$\text{tr} \begin{pmatrix} 0 & -1 \\ -1 & 3 \end{pmatrix} = 3 \quad \text{tr} \begin{pmatrix} 0 & 3 & -4 & 2 \\ 3 & 1 & 1 & 0 \\ -4 & 1 & 0,2 & 1 \\ 2 & 0 & 1 & 1 \end{pmatrix} = 2,2 \quad \text{tr} \begin{pmatrix} 1 & 0 & -2 \\ -2 & 10 & -5 \\ -5 & 0 & 0 \end{pmatrix} = 11$$

Determinante

La siguiente es una herramienta extremadamente útil en el análisis de matrices : el *determinante*. Se trata de un escalar que es posible asociar a una matriz cuadrada, y que ofrece información relevante respecto a ésta. Su cálculo es medianamente complejo, pero es un concepto muy profundo que vale la pena aprender. En términos abstractos el determinante se define a través de las siguientes propiedades.

Definición 5. Una función $\det : \mathcal{M}_{n \times n}(\mathbb{K}) \rightarrow \mathbb{K}$ se llama determinante si para cualquier serie de filas $v_1, \dots, v_n, u \in \mathbb{K}^n$ y cualquier $\alpha, \beta \in \mathbb{K}$ cumple:

$$1. \det \begin{pmatrix} - & v_1 & - \\ - & \dots & - \\ - & \alpha v_i + \beta u & - \\ - & \dots & - \\ - & v_n & - \end{pmatrix} = \alpha \det \begin{pmatrix} - & v_1 & - \\ - & \dots & - \\ - & v_i & - \\ - & \dots & - \\ - & v_n & - \end{pmatrix} + \beta \det \begin{pmatrix} - & v_1 & - \\ - & \dots & - \\ - & u & - \\ - & \dots & - \\ - & v_n & - \end{pmatrix}$$

$$2. \det \begin{pmatrix} - & v_1 & - \\ - & \dots & - \\ - & v_i & - \\ - & \dots & - \\ - & v_j & - \\ - & \dots & - \\ - & v_n & - \end{pmatrix} = - \det \begin{pmatrix} - & v_1 & - \\ - & \dots & - \\ - & v_j & - \\ - & \dots & - \\ - & v_i & - \\ - & \dots & - \\ - & v_n & - \end{pmatrix}$$

$$3. \det(I_n) = 1$$

La matemática demuestra, mediante un muy largo proceso, que existe una sola función que cumple estas 3 propiedades, y ésta se puede calcular recursivamente como sigue:

Definición 6 (Cálculo del determinante). *El determinante de una matriz cuadrada se obtiene así:*

- Si $A = (a_{11}) \in \mathcal{M}_{1 \times 1}(\mathbb{K})$, entonces: $\det(A) = a_{11}$.
- Si $A \in \mathcal{M}_{n \times n}(\mathbb{K})$, con $n \geq 2$, entonces:

$$\det(A) = \sum_{i=1}^n (-1)^{i+1} a_{i1} \det(A^{i1})$$

Donde A^{kl} representa la (sub)matriz que se obtiene de A al eliminar la fila k y la columna l .

Ejemplo: Si a la matriz $\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$ le borramos la fila 1 y la columna 2, nos queda la matriz de 1×1 : (a_{21}) . Así

$$\det \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = (-1)^{1+1} a_{11} \det(a_{22}) + (-1)^{2+1} a_{21} \det(a_{12}) = a_{11}a_{22} - a_{12}a_{21}$$

$$\text{Por ejemplo, } \det \begin{pmatrix} 2 & 4 \\ -3 & 5 \end{pmatrix} = 2 \cdot 5 - 4 \cdot (-3) = 10 + 12 = 22.$$

Ejemplo: Si $A \in \mathcal{M}_{3 \times 3}(\mathbb{K})$, entonces:

$$\begin{aligned} A &= \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \longrightarrow A^{11} = \begin{pmatrix} & a_{22} & a_{23} \\ & a_{32} & a_{33} \end{pmatrix} = \begin{pmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{pmatrix} \\ &\longrightarrow A^{21} = \begin{pmatrix} a_{12} & a_{13} \\ a_{32} & a_{33} \end{pmatrix} = \begin{pmatrix} a_{12} & a_{13} \\ a_{32} & a_{33} \end{pmatrix} \\ &\longrightarrow A^{31} = \begin{pmatrix} a_{12} & a_{13} \\ a_{22} & a_{23} \end{pmatrix} = \begin{pmatrix} a_{12} & a_{13} \\ a_{22} & a_{23} \end{pmatrix} \end{aligned}$$

Ejemplo. Calculemos el determinante de la siguiente matriz:

$$\begin{aligned} \det(A) &= \det \begin{pmatrix} 1 & 1 & -2 \\ 3 & 0 & -6 \\ -5 & 1 & 8 \end{pmatrix} \\ &= (-1)^{1+1} 1 \det \begin{pmatrix} 0 & -6 \\ 1 & 8 \end{pmatrix} + (-1)^{2+1} 3 \det \begin{pmatrix} 1 & -2 \\ 1 & 8 \end{pmatrix} + (-1)^{3+1} (-5) \det \begin{pmatrix} 1 & -2 \\ 0 & -6 \end{pmatrix} \\ &= (0 + 6) - 3(8 + 2) - 5(-6 + 0) \\ &= 6 - 30 + 30 \\ &= 6 \end{aligned}$$

El determinante tiene muchos usos, pero en esta parte del curso el que nos servirá más tiene que ver con la invertibilidad de una matriz.

Teorema 3. *Si $A \in \mathcal{M}_{n \times n}(\mathbb{K})$, entonces:*

$$A \text{ es invertible si y solo si } \det(A) \neq 0.$$

No estamos en condiciones de demostrar esta propiedad, pero podemos adelantar que el determinante se puede usar para calcular la inversa misma de una matriz mediante el Método de Cramer, aunque la manera más eficiente por ahora de calcular la inversa seguirá siendo aplicar el método de Gauss-Jordan.

No es difícil ver que el determinante cumple propiedades de la Definición 5, de hecho esas propiedades se tienen que ver con operaciones elementales de filas, y el determinante se porta muy bien con las operaciones de fila, puesto que estas son multiplicaciones matriciales y el determinante respeta la multiplicación de matrices

Propiedad 6. *Dadas $A, B \in \mathcal{M}_{n \times n}(\mathbb{K})$, $\lambda \in \mathbb{K}$, se cumple:*

1. *Si A es triangular, entonces $\det(A) = a_{11}a_{22} \dots a_{nn}$.*
2. $\det(\lambda A) = \lambda^n \det(A)$
3. *Si A tiene una fila nula, entonces $\det(A) = 0$.*
4. *Si A tiene dos filas iguales, entonces $\det(A) = 0$.*
5. *Si a una fila de A le sumamos otra fila ponderada, el determinante no cambia.*
6. *Si al aplicar operaciones de fila a A , una fila se anula, entonces $\det(A) = 0$.*
7. $\det(A^\top) = \det(A)$
8. $\det(AB) = \det(A) \det(B)$
9. *Si A es invertible, entonces $\det(A^{-1}) = (\det(A))^{-1}$.*

Teorema Fundamental del Álgebra Lineal

La clase de las transformaciones lineales es muy reducida en el mundo de las transformaciones. No cualquiera es una transformación lineal, y por ésto no podemos pedirle cualquier cosa, es más, si especificamos la imagen de una transformación lineal tan solo en unos cuantos puntos ya podemos deducir su valor en el espacio completo.

Ejemplo 1. Sea una transformación lineal $T : \mathbb{K}^3 \rightarrow \mathbb{K}^3$ tal que $T(1, 0, 0) = (1, 0, 1)$, $T(0, 1, 0) = (2, 2, 1)$ y $T(0, 0, 1) = (0, 1, 1)$. Encuentre $T(x, y, z)$.

Observar que los vectores $(1, 0, 0)$, $(0, 1, 0)$ y $(0, 0, 1)$ forman la base canónica de \mathbb{K}^3 , luego todo vector (x, y, z) puede ser escrito como una combinación lineal ellos:

$$(x, y, z) = x(1, 0, 0) + y(0, 1, 0) + z(0, 0, 1)$$

donde $x, y, z \in \mathbb{K}$, y por tanto:

$$\begin{aligned} T(x, y, z) &= T\left(x(1, 0, 0) + y(0, 1, 0) + z(0, 0, 1)\right) \\ &= T(x(1, 0, 0)) + T(y(0, 1, 0)) + T(z(0, 0, 1)) \\ &= xT(1, 0, 0) + yT(0, 1, 0) + zT(0, 0, 1), \end{aligned}$$

y puesto que $T(1, 0, 0) = (1, 0, 1)$, $T(0, 1, 0) = (2, 2, 1)$ y $T(0, 0, 1) = (0, 1, 1)$, entonces:

$$\begin{aligned} &= x(1, 0, 1) + y(2, 2, 1) + z(0, 1, 1) \\ &= (x + 2y, 2y + z, x + y + z), \end{aligned}$$

es decir, $T(x, y, z) = (x + 2y, 2y + z, x + y + z)$, o bien:

$$T \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 1 & 2 & 0 \\ 0 & 2 & 1 \\ 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

Cuando los vectores donde conocemos la función conforman una base, el sistema que impongamos tendrá solución única, y por lo tanto podremos conocer la función completamente. El ejemplo anterior nos sugiere el siguiente resultado, cuya demostración sigue los mismos pasos que hemos hecho en el ejemplo, veamos.

Teorema 1 (Fundamental del Álgebra Lineal). *Para especificar completamente una transformación lineal de V en W basta conocer su imagen en una base de V .*

Demostración Supongamos que tenemos una base $B_V = \{v_1, \dots, v_m\}$ de V y que $\dim(V) = m$. Supongamos que tenemos una transformación lineal T que conocemos bien en B_V , es decir, que sabemos que $T(v_j) = u_j$, con $u_1, \dots, u_m \in W$ vectores conocidos dados.

Queremos calcular T en un vector arbitrario de V , sea v un vector cualquiera, y supongamos que

$$[v]_{B_V} = \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_m \end{pmatrix} \in \mathbb{K}^m, \text{ es decir que:}$$

$$v = \sum_{j=1}^m \alpha_j v_j.$$

Entonces:

$$T(v) = T \left(\sum_{j=1}^m \alpha_j v_j \right) = \sum_{j=1}^m \alpha_j T(v_j) = \sum_{j=1}^m \alpha_j u_j.$$

Hemos demostrado que conocemos $T(v)$, ya que este queda expresado a partir de los datos dados y de las coordenadas de v respecto a la base B_V . \square

Ahora bien, el ejemplo anterior tiene además la gracia que hemos podido representar la función como una multiplicación de matriz por vector. Habíamos dicho que esto siempre es posible cuando la función va de \mathbb{K}^m en \mathbb{K}^n , con cuerpo \mathbb{K} , pero en realidad es posible en todos los espacios, mientras tengan dimensión finita, todo está bien, veamos.

Definición 1 (Matriz Representante). *Dada $T : V \rightarrow W$ lineal, $B_V = \{v_1, \dots, v_n\}$ una base de V y $B_W = \{w_1, \dots, w_m\}$ una base de W , se define la matriz representante de T respecto a B_V y B_W como sigue:*

$$[T]_{B_V}^{B_W} = \left[\begin{array}{cccc} | & | & | & | \\ [T(v_1)]_{B_W} & [T(v_2)]_{B_W} & \cdots & [T(v_n)]_{B_W} \\ | & | & & | \end{array} \right] \in \mathcal{M}_{m \times n}(\mathbb{K})$$

Ejemplo 2. Si tomamos la transformación $D : \mathcal{P}_3(\mathbb{R}) \rightarrow \mathcal{P}_2(\mathbb{R})$, definida por $D(p) = p'$, donde p' es la derivada de p , y consideramos la base canónica de $\mathcal{P}_3(\mathbb{R})$ y $\mathcal{P}_2(\mathbb{R})$, $\mathcal{C}_{\mathcal{P}_3} = \{1, x, x^2, x^3\}$ y $\mathcal{C}_{\mathcal{P}_2} = \{1, x, x^2\}$, respectivamente.

Entonces calculamos la matriz representante de D respecto a $\mathcal{C}_{\mathcal{P}_3}$ como sigue.

p	$D(p)$	descomposición c/r $\mathcal{C}_{\mathcal{P}_2}$	$[D(p)]_{\mathcal{C}_{\mathcal{P}_2}}$
1	0	$= 0 \cdot 1 + 0x + 0x^2$	(0, 0, 0)
x	1	$= 1 \cdot 1 + 0x + 0x^2$	(1, 0, 0)
x^2	$2x$	$= 0 \cdot 1 + 2x + 0x^2$	(0, 2, 0)
x^3	$3x^2$	$= 0 \cdot 1 + 0x + 3x^2$	(0, 0, 3)

Así, la matriz representante de D respecto a la base $\mathcal{C}_{\mathcal{P}_3}$ es:

$$[D]_{\mathcal{C}_{\mathcal{P}_3}}^{\mathcal{C}_{\mathcal{P}_2}} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \end{pmatrix}.$$

Ejemplo 3. Si tomamos la transformación $T : \mathcal{M}_{2 \times 3}(\mathbb{R}) \rightarrow \mathcal{M}_{3 \times 2}(\mathbb{R})$, definida por $T(A) = A^T$, donde A^T es la traspuesta de A , y consideramos las bases canónicas de $\mathcal{M}_{2 \times 3}(\mathbb{R})$ y $\mathcal{M}_{3 \times 2}(\mathbb{R})$, $\mathcal{C}_{\mathcal{M}_{2 \times 3}}$ y $\mathcal{C}_{\mathcal{M}_{3 \times 2}}$, respectivamente.

Entonces calculamos la matriz representante de T respecto a $\mathcal{C}_{\mathcal{M}_{2 \times 3}}$ y $\mathcal{C}_{\mathcal{M}_{3 \times 2}}$ como sigue.

A	$T(A)$	descomposición c/r $\mathcal{C}_{\mathcal{M}_{3 \times 2}}$
$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$	$\begin{pmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 0 \end{pmatrix} = 1 \begin{pmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 0 \end{pmatrix} + 0 \begin{pmatrix} 0 & 1 \\ 0 & 0 \\ 0 & 0 \end{pmatrix} + 0 \begin{pmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 0 \end{pmatrix} + 0 \begin{pmatrix} 0 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix} + 0 \begin{pmatrix} 0 & 0 \\ 0 & 0 \\ 1 & 0 \end{pmatrix} + 0 \begin{pmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 1 \end{pmatrix}$	
$\begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$	$\begin{pmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 0 \end{pmatrix} = 0 \begin{pmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 0 \end{pmatrix} + 0 \begin{pmatrix} 0 & 1 \\ 0 & 0 \\ 0 & 0 \end{pmatrix} + 1 \begin{pmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 0 \end{pmatrix} + 0 \begin{pmatrix} 0 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix} + 0 \begin{pmatrix} 0 & 0 \\ 0 & 0 \\ 1 & 0 \end{pmatrix} + 0 \begin{pmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 1 \end{pmatrix}$	
$\begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$	$\begin{pmatrix} 0 & 0 \\ 0 & 0 \\ 1 & 0 \end{pmatrix} = 0 \begin{pmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 0 \end{pmatrix} + 0 \begin{pmatrix} 0 & 1 \\ 0 & 0 \\ 0 & 0 \end{pmatrix} + 0 \begin{pmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 0 \end{pmatrix} + 0 \begin{pmatrix} 0 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix} + 1 \begin{pmatrix} 0 & 0 \\ 0 & 0 \\ 1 & 0 \end{pmatrix} + 0 \begin{pmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 1 \end{pmatrix}$	
$\begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$	$\begin{pmatrix} 0 & 1 \\ 0 & 0 \\ 0 & 0 \end{pmatrix} = 0 \begin{pmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 0 \end{pmatrix} + 1 \begin{pmatrix} 0 & 1 \\ 0 & 0 \\ 0 & 0 \end{pmatrix} + 0 \begin{pmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 0 \end{pmatrix} + 0 \begin{pmatrix} 0 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix} + 0 \begin{pmatrix} 0 & 0 \\ 0 & 0 \\ 1 & 0 \end{pmatrix} + 0 \begin{pmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 1 \end{pmatrix}$	
$\begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$	$\begin{pmatrix} 0 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix} = 0 \begin{pmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 0 \end{pmatrix} + 0 \begin{pmatrix} 0 & 1 \\ 0 & 0 \\ 0 & 0 \end{pmatrix} + 0 \begin{pmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 0 \end{pmatrix} + 1 \begin{pmatrix} 0 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix} + 0 \begin{pmatrix} 0 & 0 \\ 0 & 0 \\ 1 & 0 \end{pmatrix} + 0 \begin{pmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 1 \end{pmatrix}$	
$\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$	$\begin{pmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 1 \end{pmatrix} = 0 \begin{pmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 0 \end{pmatrix} + 0 \begin{pmatrix} 0 & 1 \\ 0 & 0 \\ 0 & 0 \end{pmatrix} + 0 \begin{pmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 0 \end{pmatrix} + 0 \begin{pmatrix} 0 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix} + 0 \begin{pmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 1 \end{pmatrix} + 1 \begin{pmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 1 \end{pmatrix}$	

Así, la matriz representante de T respecto a las bases $\mathcal{C}_{\mathcal{M}_{2 \times 3}}$ y $\mathcal{C}_{\mathcal{M}_{3 \times 2}}$ es:

$$[T]_{\mathcal{C}_{\mathcal{M}_{2 \times 3}}}^{\mathcal{C}_{\mathcal{M}_{3 \times 2}}} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}.$$

Ahora, pero ¿para qué sirve esta matriz? ¿cómo se usa?, pues igual como lo hicimos en el Ejemplo 1, pero teniendo el cuidado de usar el vector coordenada del vector donde queremos aplicar la transformación.

Propiedad 1. *La matriz representante de T es la matriz que cumple, para todo $v \in V$:*

$$[T(v)]_{B_W} = [T]_{B_V}^{B_W} [v]_{B_V}.$$

Demuestra. Supongamos primero que $B_V = \{v_1, \dots, v_m\}$ y que $B_W = \{w_1, \dots, w_n\}$, y que $m = \dim(V)$ y $n = \dim(W)$. Supongamos además que $[T]_{B_V}^{B_W} = R \in \mathcal{M}_{n \times m}(\mathbb{K})$. Esto nos dice que

$$[T(v_j)] = j - \text{ésima columna de } R,$$

y por lo tanto,

$$T(v_j) = \sum_{i=1}^n R_{ij} w_i.$$

Finalmente, supongamos que $[v]_{B_V} = \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_m \end{pmatrix} \in \mathbb{K}^m$, es decir, se tiene que:

$$v = \sum_{j=1}^m \alpha_j v_j.$$

Entonces:

$$\begin{aligned} T(v) &= T\left(\sum_{j=1}^m \alpha_j v_j\right) \\ &= \sum_{j=1}^m \alpha_j T(v_j) \\ &= \sum_{j=1}^m \alpha_j \sum_{i=1}^n R_{ij} w_i \\ &= \sum_{i=1}^n \left(\sum_{j=1}^m \alpha_j R_{ij} \right) w_i. \end{aligned}$$

Por lo tanto,

$$[T(v)]_{B_W} = \begin{pmatrix} \sum_{j=1}^m \alpha_j R_{1j} \\ \sum_{j=1}^m \alpha_j R_{2j} \\ \vdots \\ \sum_{j=1}^m \alpha_j R_{nj} \end{pmatrix} = \begin{pmatrix} R_{11} & R_{12} & \cdots & R_{1m} \\ R_{21} & R_{22} & \cdots & R_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ R_{n1} & R_{n2} & \cdots & R_{nm} \end{pmatrix} \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_m \end{pmatrix} = [T]_{B_V}^{B_W} [v]_{B_V}$$

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

Para ver que no hay otra matriz que cumpla esto, basta evaluar T en v_j , como $[v_j]_{B_V} =$

entonces $[T(v_j)]_{B_W} = \begin{pmatrix} R_{1j} \\ \vdots \\ R_{nj} \end{pmatrix}$, y como las coordenadas de un vector respecto a una base dada son únicas, cualquier otra matriz que cumpla la igualdad de este teorema, deberá tener sus columnas iguales a las de R . \square

Es decir, la matriz representante de una transformación lineal en efecto sirve para calcular la transformación mediante una multiplicación de matriz por vector, pero actúa sobre el vector de coordenadas del vector donde se quiere aplicar, por ello es que depende de que estén especificadas las bases respecto a las cuales se usa.

Ejemplo 4. En el Ejemplo 1, la matriz que encontramos funciona con los vectores escritos respecto a las bases canónicas:

$$[T]_{C_3}^{C_3} = \begin{pmatrix} 1 & 2 & 0 \\ 0 & 2 & 1 \\ 1 & 1 & 1 \end{pmatrix},$$

si $v = (x, y, z)$, se cumple:

$$T(v) = [T(v)]_{C_3} = \begin{pmatrix} 1 & 2 & 0 \\ 0 & 2 & 1 \\ 1 & 1 & 1 \end{pmatrix} [v]_{C_3} = \begin{pmatrix} 1 & 2 & 0 \\ 0 & 2 & 1 \\ 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

Ejemplo 5. En el ejemplo de la derivada verificamos lo siguiente.

$$[ax^3 + bx^2 + cx + d]_{C_{P_3}} = \begin{pmatrix} d \\ c \\ b \\ a \end{pmatrix}$$

$$[D(ax^3 + bx^2 + cx + d)]_{C_{P_3}} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \end{pmatrix} \begin{pmatrix} d \\ c \\ b \\ a \end{pmatrix} = \begin{pmatrix} c \\ 2b \\ 3a \end{pmatrix}$$

Por lo tanto $D(ax^3 + bx^2 + cx + d) = c + 2bx + 3ax^2$

Ejemplo 6. En el ejemplo de la traspuesta verificamos que.

$$\left[\begin{pmatrix} a & b & c \\ d & e & f \end{pmatrix} \right]_{C_{2 \times 3}} = \begin{pmatrix} a \\ b \\ c \\ d \\ e \\ f \end{pmatrix}$$

$$\left[T \begin{pmatrix} a & b & c \\ d & e & f \end{pmatrix} \right]_{C_{2 \times 3}} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \\ d \\ e \\ f \end{pmatrix} = \begin{pmatrix} a \\ d \\ b \\ e \\ c \\ f \end{pmatrix}$$

Como $C_{3 \times 2} = \left\{ \begin{pmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 1 \end{pmatrix} \right\}$, verificamos que

$$T \begin{pmatrix} a & b & c \\ d & e & f \end{pmatrix} = \begin{pmatrix} a & d \\ b & e \\ c & f \end{pmatrix}.$$

Vamos a terminar esta sección estableciendo resultados importantes que relacionan la composición de funciones con la multiplicación de matrices.

Propiedad 2. 1. Dada B_V una base de V , se cumple:

$$[id]_{B_V}^{B_V} = I_n.$$

2. Dada $T : V \rightarrow W$ lineal e invertible, dadas B_V base de V y B_W base de W , resulta que $T^{-1} : W \rightarrow V$ es también lineal y

$$[T^{-1}]_{B_W}^{B_V} = ([T]_{B_V}^{B_W})^{-1}.$$

3. Dadas $L : U \rightarrow V$ y $T : V \rightarrow W$ lineales, y dadas B_U base de U , B_V base de V y B_W base de W , resulta que $T \circ L : U \rightarrow W$ es también lineal y

$$[T \circ L]_{B_U}^{B_W} = [T]_{B_V}^{B_W} [L]_{B_U}^{B_V}.$$

Demostración.

1. Es claro que la identidad es la única matriz que cumple:

$$[id(v)]_{B_V} = [v]_{B_V} = I_n [v]_{B_V}.$$

2. Supongamos ahora que T es invertible, esto significa que T es inyectiva y sobreyectiva. Usando el teorema núcleo-imagen sabemos que si es inyectiva, entonces $\dim(V) \leq \dim(W)$, y si es sobreyectiva, entonces $\dim(V) \geq \dim(W)$, por lo tanto $\dim(V) = \dim(W)$. Así la matriz representante de T es cuadrada, es además de rango completo, por lo tanto es una matriz invertible.

Sea ahora $w \in W$ y $v \in V$ tales que $T(v) = w$, es decir: $v = T^{-1}(w)$. Esto lo podemos expresar en la matriz representante a través de los vectores coordenada de v y w :

$$[w]_{B_W} = [T(v)]_{B_W} = [T]_{B_V}^{B_W} [v]_{B_V}.$$

Pero despejando v en términos de w , obtenemos:

$$([T]_{B_V}^{B_W})^{-1} [w]_{B_W} = [v]_{B_V} = [T^{-1}(w)]_{B_V} = [T^{-1}]_{B_W}^{B_V} [w]_{B_W},$$

como esto se cumple para todo w , demuestra que $[T^{-1}]_{B_W}^{B_V} = ([T]_{B_V}^{B_W})^{-1}$.

3. Sea $u \in U$, se cumple:

$$[(T \circ L)]_{B_U}^{B_W}[u]_{B_U} = [(T \circ L)(u)]_{B_W} = [T(L(u))]_{B_W} = [T]_{B_V}^{B_W}[L(u)]_{B_V} = [T]_{B_V}^{B_W}[L]_{B_U}^{B_V}[u]_{B_U},$$

lo que prueba que $[T \circ L]_{B_U}^{B_W} = [T]_{B_V}^{B_W}[L]_{B_U}^{B_V}$. \square

Ejemplo 7. Sean $H : \mathbb{R}^2 \rightarrow \mathbb{R}^4$ definida por $H(x, y) = (x + 2y, x, -y, x + y)$, y $T : \mathbb{R}^4 \rightarrow \mathbb{R}^3$ definida por $T(a, b, c, d) = (a + b, 2a + b, c + d)$.

La composición de H con T es la función $T \circ H : \mathbb{R}^2 \rightarrow \mathbb{R}^3$ definida por:

$$\begin{aligned} (T \circ H)(x, y) &= T(H(x, y)) \\ &= T(\textcolor{red}{x+2y}, \textcolor{blue}{x}, \textcolor{orange}{-y}, \textcolor{magenta}{x+y}) \\ &= (\textcolor{red}{x+2y+x}, 2(\textcolor{red}{x+2y})+\textcolor{blue}{x}, \textcolor{orange}{-y+x+y}) \\ &= (2x + 2y, 3x + 4y, x) \end{aligned}$$

Matricialmente vemos que:

$$M_H = \begin{pmatrix} 1 & 2 \\ 1 & 0 \\ 0 & -1 \\ 1 & 1 \end{pmatrix} \quad M_T = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 2 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{pmatrix},$$

por lo tanto:

$$M_{T \circ H} = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 2 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 1 & 0 \\ 0 & -1 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 2 & 2 \\ 3 & 4 \\ 1 & 0 \end{pmatrix},$$

que coincide con lo que habríamos obtenido directamente de la fórmula de $T \circ H$.

Ejemplo 8. Consideremos ahora la función $G : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ definida por $G(a, b) = (a + 2b, 2a + 4b)$. Su matriz asociada es claramente:

$$M_G = \begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix}.$$

La cual es una matriz no invertible, de donde obtenemos que G tampoco lo es.

Matriz de cambio de base

Observación 1. Si consideramos dos bases distintas de V , B_1 y B_2 , entonces la matriz representante de la función identidad de V en V ya no será la matriz identidad, será una matriz interesante que llamamos matriz de paso de la base B_1 a la base B_2 , y la denotamos por:

$$P_{B_1}^{B_2} = [id]_{B_1}^{B_2}$$

Cumple la interesante propiedad que para todo $v \in V$:

$$P_{B_1}^{B_2}[v]_{B_1} = [v]_{B_2},$$

además

$$P_{B_1}^{B_2} = (P_{B_2}^{B_1})^{-1}.$$

Ejemplo 9. Considere el espacio de los polinomios de grado menor o igual que 3, $\mathcal{P}_3(\mathbb{R})$, y su base canónica, $\mathcal{C}_{\mathcal{P}_3} = \{1, x, x^2, x^3\}$. Considere también la base $\mathcal{B} = \{x^3 - 1, x^2 + x - 1, x + 1, 1\}$.

La matriz de paso de \mathcal{B} a $\mathcal{C}_{\mathcal{P}_3}$ corresponde a la matriz representante de la función identidad respecto a estas bases, veamos.

p	$id(p)$	descomposición c/r $\mathcal{C}_{\mathcal{P}_2}$	$[P]_{\mathcal{C}_{\mathcal{P}_2}}$
$x^3 - 1$	$x^3 - 1$	$= -1 \cdot 1 + 0 \cdot x + 0 \cdot x^2 + 1 \cdot x^3$	$(-1, 0, 0, 1)$
$x^2 + x - 1$	$x^2 + x - 1$	$= -1 \cdot 1 + 1 \cdot x + 1 \cdot x^2 + 0 \cdot x^3$	$(-1, 1, 1, 0)$
$x + 1$	$x + 1$	$= 1 \cdot 1 + 1 \cdot x + 0 \cdot x^2 + 0 \cdot x^3$	$(1, 1, 0, 0)$
1	1	$= 1 \cdot 1 + 0 \cdot x + 0 \cdot x^2 + 0 \cdot x^3$	$(1, 0, 0, 0)$

Por lo tanto, la matriz es:

$$P_{\mathcal{B}}^{\mathcal{C}_{\mathcal{P}_3}} = \begin{pmatrix} -1 & -1 & 1 & 1 \\ 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}.$$

Ahora bien, la matriz de paso de $\mathcal{C}_{\mathcal{P}_3}$ a \mathcal{B} es la inversa de esta matriz.

$$P_{\mathcal{C}_{\mathcal{P}_3}}^{\mathcal{B}} = (P_{\mathcal{B}}^{\mathcal{C}_{\mathcal{P}_3}})^{-1}$$

Calculémosla.

$$\begin{aligned} & \left(\begin{array}{cccc|ccccc} -1 & -1 & 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{array} \right) \sim \left(\begin{array}{cccc|ccccc} -1 & -1 & 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & -1 & 1 & 1 & 1 & 0 & 0 & 1 \end{array} \right) \sim \left(\begin{array}{cccc|ccccc} -1 & -1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 2 & 1 & 1 & 0 & 0 & 1 & 1 \end{array} \right) \\ & \sim \left(\begin{array}{ccc|ccccc} -1 & -1 & 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 2 & 1 & 1 & 0 & 0 & 1 \end{array} \right) \sim \left(\begin{array}{ccc|ccccc} -1 & -1 & 1 & 0 & 0 & 2 & -3 & -1 \\ 0 & 1 & 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 & -1 & 1 & 0 \\ 0 & 0 & 2 & 1 & 1 & -2 & 3 & 1 \end{array} \right) \\ & \sim \left(\begin{array}{ccc|ccccc} -1 & -1 & 0 & 0 & 0 & 1 & -2 & -1 \\ 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & -1 & 0 & 0 & -1 & 1 & 0 \\ 0 & 0 & 2 & 1 & 1 & -2 & 3 & 1 \end{array} \right) \sim \left(\begin{array}{ccc|ccccc} -1 & 0 & 0 & 0 & 0 & 1 & -1 & -1 \\ 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & -1 & 0 & 0 & -1 & 1 & 0 \\ 0 & 0 & 2 & 1 & 1 & -2 & 3 & 1 \end{array} \right) \end{aligned}$$

$$\sim \left(\begin{array}{cccc|cccc} 1 & 0 & 0 & 0 & 0 & -1 & 1 & 1 \\ 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 1 & 1 & -2 & 3 & 1 \end{array} \right)$$

$$P_{\mathcal{C}_{\mathcal{P}_3}}^{\mathcal{B}} = \left(\begin{array}{cccc} 0 & -1 & 1 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & -1 & 0 \\ 1 & -2 & 3 & 1 \end{array} \right)$$

Entonces ahora podemos usar esta matriz para encontrar el vector de coordenadas de un polinomio respecto a la base \mathcal{B} .

$$\begin{aligned} [ax^3 + bx^2 + cx + d]_{\mathcal{B}} &= P_{\mathcal{C}_{\mathcal{P}_3}}^{\mathcal{B}} [ax^3 + bx^2 + cx + d]_{\mathcal{C}_{\mathcal{P}_3}} \\ &= \begin{pmatrix} 0 & -1 & 1 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & -1 & 0 \\ 1 & -2 & 3 & 1 \end{pmatrix} \begin{pmatrix} d \\ c \\ b \\ a \end{pmatrix} \\ &= \begin{pmatrix} a + b - c \\ c \\ c - b \\ a + 3b - 3c + d \end{pmatrix} \end{aligned}$$

Por ejemplo para $x^3 - 2$, su vector de coordenadas serían:

$$[x^3 - 2]_{\mathcal{B}} = \begin{pmatrix} 1 \\ 0 \\ 0 \\ -1 \end{pmatrix},$$

lo cual implica que:

$$x^3 - 2 = 1 \cdot (x^3 - 1) + 0 \cdot (x^2 + x - 1) + 0 \cdot (x + 1) + (-1) \cdot 1 = x^3 - 1 - 1$$

Esta matriz sirve para obtener fácilmente el vector de coordenadas. Pero también sirve para ayudar en la búsqueda y cálculo de la matriz representante de una transformación lineal gracias a la siguiente observación.

Observación 2. Si tenemos dos bases de W , B_{1W} y B_{2W} , y dos bases de V , B_{1V} y B_{2V} , entonces:

$$[T]_{B_{1V}}^{B_{1W}} = P_{B_{2W}}^{B_{1W}} [T]_{B_{2V}}^{B_{2W}} P_{B_{1V}}^{B_{2V}}.$$

Es decir, la matriz de paso sirve para encontrar una matriz relativa a unas bases a partir de la matriz relativa a otras bases, y esto también es útil para encontrar la fórmula de la transformación a partir de su valor en vectores dados.

Ejemplo 10. Supongamos que tenemos una transformación lineal $F : \mathbb{R}^2 \rightarrow \mathbb{R}^3$ tal que $F(1, 1) = (1, 1, 1)$, $F(1, 2) = (1, 2, 1)$, entonces, para poder conocer a F necesitamos que $\mathcal{D} = \{(1, 1), (1, 2)\}$ sea base de \mathbb{R}^2 , afortunadamente lo es. Luego, la matriz representante de F respecto a la base \mathcal{D} de \mathbb{R}^2 y la base canónica \mathcal{C}_3 de \mathbb{R}^3 es

$$[F]_{\mathcal{D}}^{\mathcal{C}_3} = \begin{pmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 1 \end{pmatrix}.$$

Si buscamos su matriz representante respecto a las bases canónicas de ambos espacios respectivamente, necesitamos la matriz de paso de la base canónica \mathcal{C}_2 de \mathbb{R}^2 a la base \mathcal{D} .

Como antes, es más sencillo calcular la matriz de paso de \mathcal{D} a \mathcal{C}_2 , pues basta escribir los vectores de \mathcal{D} en términos de los vectores canónicos, y eso son los mismos vectores tal cual, luego ponerlos como columna, entonces:

$$P_{\mathcal{D}}^{\mathcal{C}_2} = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix},$$

luego ocupar la fórmula de la inversa para obtener la matriz de paso inverso:

$$P_{\mathcal{C}_2}^{\mathcal{D}} = (P_{\mathcal{D}}^{\mathcal{C}_2})^{-1} = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}^{-1} = \begin{pmatrix} 2 & -1 \\ -1 & 1 \end{pmatrix}.$$

Entonces, aplicando la fórmula de la observación anterior tenemos:

$$[F]_{\mathcal{C}_2}^{\mathcal{C}_3} = [F]_{\mathcal{D}}^{\mathcal{C}_3} P_{\mathcal{C}_2}^{\mathcal{D}} == \begin{pmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 2 & -1 \\ -1 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

Con esto podemos obtener la fórmula de F fácilmente:

$$F \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x \\ y \\ x \end{pmatrix}$$

Que en efecto cumple:

$$F \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \quad F \begin{pmatrix} 1 \\ 2 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}$$

tal como se pedía al comienzo.

Definición 2. Dos matrices A, B cuadradas en $\mathcal{M}_{n \times n}(\mathbb{K})$ se dicen semejantes si existe una matriz $M \in \mathcal{M}_{n \times n}(\mathbb{K})$ tal que $A = M^{-1}BM$.

Propiedad 3. Si A y B son semejantes, entonces:

- $tr(A) = tr(B)$,
- $det(A) = det(B)$,
- $R(A) = R(B)$,
- $\eta(A) = \eta(B)$.

Vectores y Valores Propios

Una función lineal $f : \mathbb{R} \rightarrow \mathbb{R}$ se puede escribir $f(x) = mx$, con m su pendiente. Esta pendiente nos dice si la función es “rápida” o “lenta”, cuánto “estira” o “estrecha” a sus preimágenes, etc. El concepto de valor propio generaliza la idea de pendiente a una transformación lineal $T : V \rightarrow V$ o a una matriz cuadrada, a través de una ecuación similar.

En este capítulo trabajamos con transformaciones lineales de un espacio en sí mismo.

Definición 1. Dado V un espacio vectorial sobre un cuerpo \mathbb{K} y una aplicación lineal $T : V \rightarrow V$, decimos que $\lambda \in \mathbb{K}$ es un valor propio de T si existe $v \neq \Theta$ tal que

$$T(v) = \lambda v,$$

en tal caso, decimos que v es un vector propio de T asociado a λ .

El conjunto de valores propios de T se llama espectro de T y se denota por $\sigma(T)$.

Análogamente, dada una matriz $A \in \mathcal{M}_{n \times n}(\mathbb{K})$, decimos que $\lambda \in \mathbb{K}$ es un valor propio de A si existe $v \in \mathbb{K}^n$ tal que

$$Av = \lambda v \quad \wedge \quad v \neq \Theta$$

en tal caso, decimos que v es un vector propio de A asociado a λ . El conjunto de valores propios de A se llama espectro de A y se denota por $\sigma(A)$.

Ejemplo 1. Si $A = \begin{pmatrix} 1 & 2 \\ 3 & 6 \end{pmatrix}$, entonces: $\lambda_1 = 7$ es un valor propio de A con vector propio asociado $u_1 = (1, 3)$:

$$\begin{pmatrix} 1 & 2 \\ 3 & 6 \end{pmatrix} \begin{pmatrix} 1 \\ 3 \end{pmatrix} = \begin{pmatrix} 7 \\ 21 \end{pmatrix} = 7 \begin{pmatrix} 1 \\ 3 \end{pmatrix}$$

y $\lambda_2 = 0$ es un valor propio de A con el vector propio asociado $u_2 = (2, -1)$:

$$\begin{pmatrix} 1 & 2 \\ 3 & 6 \end{pmatrix} \begin{pmatrix} 2 \\ -1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} = 0 \begin{pmatrix} 2 \\ -1 \end{pmatrix}$$

así $0 \in \sigma(A)$ y $7 \in \sigma(A)$.

En otras palabras, un vector propio es un vector al cual la transformación trata con una simple ponderación por escalar. Por supuesto no todos los vectores serán tratados así en general, por eso cuando ocurre se destaca particularmente. Como veremos la clase siguiente, estos vectores donde la función es *simple* ayudarán a entenderla mejor y a simplificar algunos cálculos.

Propiedad 1. *Dada $T : V \rightarrow V$ lineal, V espacio vectorial sobre el cuerpo \mathbb{K} y \mathcal{B} una base de V , las siguientes afirmaciones son equivalentes para todo $\lambda \in \mathbb{K}$:*

1. λ es valor propio de T
2. $\text{Ker}(T - \lambda id) \neq \{\Theta_V\}$
3. $n(T - \lambda id) \neq 0$
4. $n([T]_B^B - \lambda Id) \neq 0$
5. $\det([T]_B^B - \lambda Id) = 0$

Demostración Sea $T : V \rightarrow V$ una transformación lineal, y $\mathcal{B} = \{v_1, \dots, v_n\}$ una base de V .

$1 \Rightarrow 2$ Si λ es valor propio de T , entonces existe $v \in V$ no nulo tal que $T(v) = \lambda v$. Esto implica que $T(v) - \lambda id(v) = (T - \lambda id)(v) = \Theta_V$, por lo tanto $v \in \text{Ker}(T - \lambda id)$, y como $v \neq \Theta_V$, entonces $\text{Ker}(T - \lambda id) \neq \{\Theta_V\}$.

$2 \Rightarrow 3$ Si $\text{Ker}(T - \lambda id) \neq \{\Theta_V\}$, significa que la dimensión de $\text{Ker}(T - \lambda id)$ es 1 o más, por lo tanto $n(T - \lambda id) \neq 0$.

$3 \Rightarrow 4$ Si $n(T - \lambda id) \neq 0$, entonces hay algún vector no nulo $v \in \text{Ker}(T - \lambda id)$, esto implica que $T(v) - \lambda id(v) = \Theta_V$. Pasando esta igualdad a la matriz representante de estos operadores lineales obtenemos: $[T]_{\mathcal{B}}^{\mathcal{B}}[v]_{\mathcal{B}} - \lambda Id[v]_{\mathcal{B}} = \Theta_{\mathbb{K}^n}$. Por lo tanto $[v]_{\mathcal{B}} \in \text{Ker}([T]_B^B - \lambda Id)$, y como $[v]_{\mathcal{B}} \neq \Theta_{\mathbb{K}^n}$, esto significa que $\text{Ker}([T]_B^B - \lambda Id)$ es no trivial, y por lo tanto su dimensión es mayor que 0, es decir, $n([T]_B^B - \lambda Id) \neq 0$.

$4 \Rightarrow 5$ Si $n([T]_B^B - \lambda Id) \neq 0$, tenemos que la matriz $[T]_B^B - \lambda Id$ es singular, no es invertible, por lo tanto $\det([T]_B^B - \lambda Id) = 0$.

$5 \Rightarrow 1$ Si $\det([T]_B^B - \lambda Id) = 0$, entonces la matriz $[T]_B^B - \lambda Id$ es singular, por lo tanto existe $x = (x_1, \dots, x_n) \in \mathbb{K}^n \setminus \{\Theta_{\mathbb{K}^n}\}$ tal que $([T]_B^B - \lambda Id)x = \Theta_{\mathbb{K}^n}$. Tomando $v = \sum_{i=1}^n x_i v_i \in V$, tenemos que $[v]_{\mathcal{B}} = x$, así:

$$[T(v) - \lambda v]_{\mathcal{B}} = [T(v)]_{\mathcal{B}} - \lambda[id(v)]_{\mathcal{B}} = [T]_B^B[v]_{\mathcal{B}} - \lambda[v]_{\mathcal{B}} = ([T]_B^B - \lambda Id)[v]_{\mathcal{B}} = \Theta_{\mathbb{K}^n}.$$

Por lo tanto, $T(v) - \lambda v = \Theta_V$, entonces v es vector propio de T asociado a λ y λ es valor propio de T .

□

Definición 2. Dada una matriz $A \in \mathcal{M}_{n \times n}(\mathbb{K})$, se denomina **polinomio característico** de A a:

$$p_A(\lambda) = \det(A - \lambda Id),$$

el cual es un polinomio de grado n . Además a $p_A(\lambda) = 0$ se le llama **ecuación característica de A** .

Análogamente, dada una transformación lineal $T : V \rightarrow V$, su polinomio característico es:

$$p_T(\lambda) = \det([T]_{\mathcal{B}}^{\mathcal{B}} - \lambda Id),$$

para cualquier base \mathcal{B} de V , ya que en realidad no depende de ésta, siempre que se tome la misma a la partida y a la llegada.

Propiedad 2. ■ Dada una matriz $A \in \mathcal{M}_{n \times n}(\mathbb{K})$, $\lambda \in \mathbb{K}$ es valor propio de A si y sólo si $p_A(\lambda) = 0$.

■ Dada una transformación lineal $T : V \rightarrow V$, $\lambda \in \mathbb{K}$ es valor propio de T si y sólo si $p_T(\lambda) = 0$.

Ejemplo 2. Sea $A = \begin{pmatrix} 3 & 0 & 2 \\ 0 & 2 & 1 \\ 1 & 0 & 2 \end{pmatrix}$.

- a) Calcule el polinomio característico de A .
- b) Encuentre el espectro de A .

Solución.

a) El polinomio característico de A se define como $p_A(\lambda) = \det(A - \lambda Id_3)$, es decir:

$$\begin{aligned} p_A(\lambda) = \det(A - \lambda Id_3) &= \det \left(\begin{pmatrix} 3 & 0 & 2 \\ 0 & 2 & 1 \\ 1 & 0 & 2 \end{pmatrix} - \lambda \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \right) \\ &= \det \left(\begin{pmatrix} 3 - \lambda & 0 & 2 \\ 0 & 2 - \lambda & 1 \\ 1 & 0 & 2 - \lambda \end{pmatrix} \right) \\ &= (3 - \lambda)(2 - \lambda)(2 - \lambda) + (-2(2 - \lambda)) \\ &= (2 - \lambda)((3 - \lambda)(2 - \lambda) - 2) \\ &= (2 - \lambda)(4 - 5\lambda + \lambda^2) \end{aligned}$$

$$p_A(\lambda) = (2 - \lambda)(\lambda - 1)(\lambda - 4)$$

b) El espectro de A es el conjunto de todos los valores propios de A y se denota por $\sigma(A)$. Por su parte, los valores propios de A son las raíces del polinomio característico, esto es, son los $\lambda \in \mathbb{R}$ tales que $p_A(\lambda) = 0$, por tanto los valores propios de A son: $\lambda_1 = 1$, $\lambda_2 = 2$ y $\lambda_3 = 4$, así:

$$\sigma(A) = \{1, 2, 4\}$$

Definición 3. Si λ es un valor propio de T , se define el espacio propio asociado a λ por

- $S_\lambda(T) = \{v \in V \mid T(v) = \lambda v\} = \text{Ker}(T - \lambda id)$.

Lo análogo se define para matrices cuadradas.

Propiedad 3. Si λ es un valor propio de un operador lineal $T : V \rightarrow V$, entonces su espacio propio es un s. e. v de V .

Ejemplo 3. Para $A = \begin{pmatrix} 3 & 0 & 2 \\ 0 & 2 & 1 \\ 1 & 0 & 2 \end{pmatrix}$ vista en el ejemplo anterior, calcule el subespacio propio asociado al valor propio $\lambda = 1$.

Solución.

$$\begin{aligned} S_{\lambda=1}(A) &= \text{Ker}(A - \lambda Id_3) \\ &= \left\{ v \in \mathbb{R}^3 : (A - \lambda Id_3)v = \theta_{\mathbb{R}^3} \right\} \\ &= \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3 : \left(\begin{pmatrix} 3 & 0 & 2 \\ 0 & 2 & 1 \\ 1 & 0 & 2 \end{pmatrix} - 1 \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \right) \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \right\} \\ &= \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3 : \left(\begin{pmatrix} 3 & 0 & 2 \\ 0 & 2 & 1 \\ 1 & 0 & 2 \end{pmatrix} - 1 \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \right) \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \right\} \\ &= \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3 : \begin{pmatrix} 3-1 & 0 & 2 \\ 0 & 2-1 & 1 \\ 1 & 0 & 2-1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \right\} \\ &= \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3 : \begin{pmatrix} 2 & 0 & 2 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \right\} \end{aligned}$$

Escalonando la matriz y resolviendo el sistema, se llega a:

$$\begin{aligned}
S_{\lambda=1}(A) &= \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3 : y = -z \wedge x = -z \right\} \\
&= \left\{ \begin{pmatrix} -z \\ -z \\ z \end{pmatrix} \in \mathbb{R}^3 : z \in \mathbb{R} \right\} \\
&= \left\{ z \begin{pmatrix} -1 \\ -1 \\ 1 \end{pmatrix} \in \mathbb{R}^3 : z \in \mathbb{R} \right\} \\
&= \left\langle \left\{ \begin{pmatrix} -1 \\ -1 \\ 1 \end{pmatrix} \right\} \right\rangle.
\end{aligned}$$

Definición 4. Si el polinomio característico de T se factoriza como sigue:

$$p_T(x) = (x - \lambda_1)^{r_1}(x - \lambda_2)^{r_2} \dots (x - \lambda_k)^{r_k} q(x),$$

donde $q(x)$ es un polinomio sin raíces en \mathbb{K} , entonces r_i se llama multiplicidad algebraica del valor propio λ_i de T .

Por otra parte, la multiplicidad geométrica de λ_i se define como $g_i = \dim(S_{\lambda_i})$.

Conceptos análogos se definen en el caso de matrices cuadradas.

Ejemplo 4. Para la matriz A , vista en el ejemplo anterior, el polinomio característico era $p_A(\lambda) = (2 - \lambda)(\lambda - 1)(\lambda - 4)$, por lo tanto los 3 valores propios tienen multiplicidad algebraica igual a 1. Además, dado que $S_1(A)$ tiene dimensión 1, podemos afirmar que el valor propio 1 tiene multiplicidad geométrica igual a 1.

Ejemplo 5. Considere la transformación lineal $F : \mathcal{P}_2(\mathbb{R}) \rightarrow \mathcal{P}_2(\mathbb{R})$, definida por $F(ax^2 + bx + c) = (b + 2a)x^2 - ax - c$. Determine espectro, polinomio característico y multiplicidades.

Para calcular el polinomio característico, primero que nada debemos tomar una base de $\mathcal{P}_2(\mathbb{R})$ y calcular la matriz representante de F respecto a esta. Tomamos la base canónica para no complicarnos: $\mathcal{C}_{\mathcal{P}_2} = \{1, x, x^2\}$.

p	$F(p)$	descomposición c/r $\mathcal{C}_{\mathcal{P}_2}$	$[F(p)]_{\mathcal{C}_{\mathcal{P}_2}}$
1	-1	= $-1 \cdot 1 + 0 \cdot x + 0 \cdot x^2$	(-1, 0, 0)
x	x^2	= $0 \cdot 1 + 0 \cdot x + 1 \cdot x^2$	(0, 0, 1)
x^2	$2x^2 - x$	= $0 \cdot 1 + (-1) \cdot x + 2 \cdot x^2$	(0, -1, 2)

Así,

$$[F]_{\mathcal{C}_{\mathcal{P}_2}}^{\mathcal{C}_{\mathcal{P}_2}} = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 2 \end{pmatrix},$$

por lo tanto su polinomio característico es:

$$p_F(\lambda) = \det \left(\begin{pmatrix} -1 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 2 \end{pmatrix} - \lambda \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \right) = -(\lambda + 1)(\lambda - 1)^2$$

luego sus valores propios son $\lambda_1 = -1$ y $\lambda_2 = 1$, con multiplicidad algebraica 1 y 2 respectivamente.

Para calcular las multiplicidades geométricas, calculamos los subespacios propios.

$$\begin{aligned} S_{-1}(F) &= \left\{ ax^2 + bx + c \in \mathcal{P}_2(\mathbb{R}) : \left(\begin{pmatrix} -1 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 2 \end{pmatrix} - (-1) \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \right) \begin{pmatrix} c \\ b \\ a \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \right\} \\ &= \left\{ ax^2 + bx + c \in \mathcal{P}_2(\mathbb{R}) : \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & -1 \\ 0 & 1 & 3 \end{pmatrix} \begin{pmatrix} c \\ b \\ a \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \right\} \\ &= \{ax^2 + bx + c \in \mathcal{P}_2(\mathbb{R}) : a = b = 0\} \\ &= \langle \{1\} \rangle \end{aligned}$$

$$\begin{aligned} S_1(F) &= \left\{ ax^2 + bx + c \in \mathcal{P}_2(\mathbb{R}) : \left(\begin{pmatrix} -1 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 2 \end{pmatrix} - (1) \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \right) \begin{pmatrix} c \\ b \\ a \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \right\} \\ &= \left\{ ax^2 + bx + c \in \mathcal{P}_2(\mathbb{R}) : \begin{pmatrix} -2 & 0 & 0 \\ 0 & -1 & -1 \\ 0 & 1 & 1 \end{pmatrix} \begin{pmatrix} c \\ b \\ a \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \right\} \\ &= \{ax^2 + bx + c \in \mathcal{P}_2(\mathbb{R}) : c = 0 \wedge a = -b\} \\ &= \langle \{x - x^2\} \rangle \end{aligned}$$

por tanto la multiplicidad geométrica de ambos valores propios es 1.

En este caso vemos que la multiplicidad geométrica de cada valor propio es menor o igual a su multiplicidad algebraica. Esto es cierto para cualquier matriz, y lo vamos a estudiar en la clase siguiente.

Ejemplo 6. Considere el operador derivada D en el subespacio de funciones generadas por el seno y el coseno:

$$S = \langle \{\sin, \cos\} \rangle, \text{ s.e.v. del espacio de las funciones de } \mathbb{R} \text{ en } \mathbb{R} \text{ sobre el cuerpo } \mathbb{R}.$$

Es claro que $B = \{\sin, \cos\}$ es l.i y por lo tanto es base de S ya que \sin no es ponderación de \cos por escalar.

Calculamos la matriz representante del operador derivada en este espacio.

f	$D(f)$	descomposición c/r B	$[D(f)]_B$
sen	$\begin{matrix} \cos \\ -\operatorname{sen} \end{matrix}$	$= \begin{matrix} 0 \cdot \operatorname{sen} + 1 \cdot \cos \\ -1 \cdot \operatorname{sen} + 0 \cdot \cos \end{matrix}$	$(0, 1)$
cos			$(-1, 0)$

Así,

$$[D]_B^B = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix},$$

por lo tanto su polinomio característico es:

$$p_D(\lambda) = \det \left(\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} - \lambda \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \right) = \lambda^2 + 1$$

luego sus valores propios son $\lambda_1 = i$ y $\lambda_2 = -i$, cada uno con multiplicidad algebraica 1.

Sin embargo, como el cuerpo del operador es \mathbb{R} , estos valores solo son valores propios de la matriz representante de D , pero no son valores propios de D , y éste no tiene ningún valor propio, ni vector propio: $\sigma(D) = \emptyset$ y $\sigma([D]_B^B) = \{i, -i\}$.

D tampoco tendrá vectores propios, pero la matriz sí, veamos.

$$\begin{aligned} S_i([D]_B^B) &= \left\{ \begin{pmatrix} a \\ b \end{pmatrix} \in \mathbb{R}^2 : \left(\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} - i \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \right) \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \right\} \\ &= \left\{ \begin{pmatrix} a \\ b \end{pmatrix} \in \mathbb{R}^2 : \begin{pmatrix} -i & -1 \\ 1 & -i \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \right\} \\ &= \left\{ \begin{pmatrix} a \\ b \end{pmatrix} \in \mathbb{R}^2 : -ai - b = 0 \wedge a - ib = 0 \right\} \\ &= \left\{ \begin{pmatrix} a \\ b \end{pmatrix} \in \mathbb{R}^2 : b = -ai \wedge a - i(-ai) = 0 \right\} \\ &= \left\{ \begin{pmatrix} a \\ b \end{pmatrix} \in \mathbb{R}^2 : b = -ai \wedge 0 = 0 \right\} \\ &= \left\{ \begin{pmatrix} a \\ -ai \end{pmatrix} \in \mathbb{R}^2 : a \text{ libre en } \mathbb{R} \right\} \\ &= \left\langle \left\{ \begin{pmatrix} 1 \\ -i \end{pmatrix} \right\} \right\rangle \end{aligned}$$

Es un subespacio de \mathbb{C}^2 .

$$\begin{aligned}
S_{-i}([D]_B^B) &= \left\{ \begin{pmatrix} a \\ b \end{pmatrix} \in \mathbb{R}^2 : \left(\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} + i \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \right) \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \right\} \\
&= \left\{ \begin{pmatrix} a \\ b \end{pmatrix} \in \mathbb{R}^2 : \begin{pmatrix} i & -1 \\ 1 & i \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \right\} \\
&= \left\{ \begin{pmatrix} a \\ b \end{pmatrix} \in \mathbb{R}^2 : ai - b = 0 \wedge a + ib = 0 \right\} \\
&= \left\{ \begin{pmatrix} a \\ b \end{pmatrix} \in \mathbb{R}^2 : b = ai \wedge a + i(ai) = 0 \right\} \\
&= \left\{ \begin{pmatrix} a \\ b \end{pmatrix} \in \mathbb{R}^2 : b = ai \wedge 0 = 0 \right\} \\
&= \left\{ \begin{pmatrix} a \\ ai \end{pmatrix} \in \mathbb{R}^2 : a \text{ libre en } \mathbb{R} \right\} \\
&= \left\langle \left\{ \begin{pmatrix} 1 \\ i \end{pmatrix} \right\} \right\rangle
\end{aligned}$$

Dijimos que los valores propios servían para conocer propiedades claves de las transformaciones lineales. Un primer ejemplo de ello es la siguiente observación:

Observación 1. *Dada $T : V \rightarrow V$ lineal, se cumple que*

$$T \text{ es invertible} \Leftrightarrow 0 \notin \sigma(T)$$

En efecto, dado que T va del espacio en sí mismo, resulta que es inyectiva si y solo si es sobreyectiva, lo cual implica que resulta que T es invertible si y solo si es inyectiva. Pero habíamos visto que la inyectividad equivale a que $\text{Ker}(T) = \{\Theta\}$, como $T = T - 0\text{id}$, esto equivale a que 0 NO sea valor propio de T .

En los ejemplos que vimos la clase pasada, vimos que la multiplicidad geométrica resultó menor o igual a la multiplicidad algebraica. Esto no es una casualidad, de hecho hay un teorema que lo establece para cualquier transformación lineal.

Teorema 1. *Si λ es un valor propio de T , y r y g son su multiplicidad algebraica y geométrica, respectivamente, entonces:*

$$1 \leq g \leq r.$$

La demostración lamentablemente excede la dificultad de este curso, pero el resultado es de suma importancia y utilidad en los ejercicios.

Ejemplo 7. Sea $A = \begin{pmatrix} 3 & 0 & 2 \\ 0 & 2 & 1 \\ 1 & 0 & 2 \end{pmatrix}$. Calcule las multiplicidades algebraicas y geométricas de cada uno de los valores propios de A .

Solución. En el apunte anterior calculamos el polinomio característico de A :

$$p_A(\lambda) = (2 - \lambda)(\lambda - 1)(\lambda - 4),$$

de donde obtuvimos que el espectro de A es: $\sigma(A) = \{1, 2, 4\}$.

Dado que el exponente de cada factor lineal en la factorización de p_A es 1, concluimos que la multiplicidad algebraica de cada uno de sus valores propios es 1.

Gracias al Teorema 1, podemos concluir que entonces la multiplicidad geométrica de cada uno de los valores propios es también 1.

No siempre se tiene tanta suerte, si una multiplicidad algebraica resultara igual a 2 o más, entonces la única manera de conocer la multiplicidad geométrica es calcular el espacio propio asociado para obtener su dimensión, tal como hicimos en el último ejemplo de la última clase.

Otra propiedad que es útil para verificar si andamos bien o mal con nuestros cálculos es la siguiente.

Propiedad 4. Sea $T : V \rightarrow V$, donde V es un espacio vectorial sobre \mathbb{K} de dimensión n . Si $\lambda_1, \lambda_2, \dots, \lambda_k$ son valores propios distintos de T , entonces

$$\sum_{i=1}^k r_i \leq n,$$

con igualdad en el caso $\mathbb{K} = \mathbb{C}$.

Demostración: Suponiendo, en el caso general que

$$p_T(x) = (x - \lambda_1)^{r_1}(x - \lambda_2)^{r_2} \dots (x - \lambda_k)^{r_k} q(x),$$

tenemos que

$$n = gr(p_T(x)) = \sum_{i=1}^k r_i + gr(q(x)).$$

De donde es claro el resultado ya que $gr(q(x)) \geq 0$. □

Verificamos que en los ejemplos vistos se cumple con igualdad, aunque no siempre es así.

Ejemplo 8. Calcule el polinomio característico, valores propios y multiplicidades geométricas y algebraicas del siguiente operador.

$$L : \mathcal{M}_{2 \times 2}(\mathbb{R}) \rightarrow \mathcal{M}_{2 \times 2}(\mathbb{R})$$

$$L \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} a & d \\ b & c \end{pmatrix}$$

Solución. Primero calculamos su matriz representante respecto a la base canónica.

M	$L(M)$	$[L(M)]_{\mathcal{C}_{M_{2 \times 2}}}$
$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$	$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$	$(1, 0, 0, 0)$
$\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$	$\begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$	$(0, 0, 1, 0)$
$\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$	$\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$	$(0, 0, 0, 1)$
$\begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$	$\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$	$(0, 1, 0, 0)$
$\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$	$\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$	

$$[L]_{\mathcal{C}_{M_{2 \times 2}}}^{\mathcal{C}_{M_{2 \times 2}}} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}$$

Ahora podemos calcular el polinomio característico.

$$\begin{aligned} p_L(x) &= \det \begin{pmatrix} 1-x & 0 & 0 & 0 \\ 0 & -x & 0 & 1 \\ 0 & 1 & -x & 0 \\ 0 & 0 & 1 & -x \end{pmatrix} \\ &= (-1)^{1+1}(1-x)\det \begin{pmatrix} -x & 0 & 1 \\ 1 & -x & 0 \\ 0 & 1 & -x \end{pmatrix} \\ &= (1-x) \left[(-1)^{1+1}(-x)\det \begin{pmatrix} -x & 0 \\ 1 & -x \end{pmatrix} + (-1)^{2+1}1\det \begin{pmatrix} 0 & 1 \\ 1 & -x \end{pmatrix} \right] \\ &= (1-x)[-x(x^2 - 0) - (0 - 1)] \\ &= (1-x)(1-x^3) \end{aligned}$$

el polinomio aun no está totalmente factorizado ya que sabemos que 1 es raíz de $1-x^3$, puesto que al evaluar en $x = 1$ da 0, pero ¿cuáles serán los otros factores? Para encontrarlos realizamos división de polinomios.

$$\begin{array}{r} -x^3 + \\ \hline -(-x^3 + x^2) \\ \hline -x^2 + \\ \hline -(-x^2 + x) \\ \hline -x + 1 \\ \hline 0 \end{array} \quad \begin{array}{l} 1 : -x + 1 = x^2 + x + 1 \\ 1 \\ 1 \\ 1 \end{array}$$

El polinomio x^2+x+1 no tiene raíces reales, de hecho su discriminante es: $b^2-4ac = 1-4 = -3 < 0$. Por lo tanto la factorización total de p_L es:

$$p_L(x) = (x-1)^2(x^2+x+1).$$

Así, el espectro de L es $\sigma(L) = \{1\}$, con su único valor propio $\lambda = 1$ de multiplicidad 2. Entonces, la suma de las multiplicidades algebraicas de sus valores propios es $2 \leq 4$.

La siguiente propiedad servirá de cimiento para el tema de la clase de hoy. Dijimos que la teoría desarrollada permitía entender mejor la transformación ya que la transformación se comporta de

manera simple sobre sus subespacios propios. Conviene entonces mirar la transformación en el subespacio generado por todos sus vectores propios. Esto no siempre se puede hacer, pero al menos algunas cosas podemos hacer. Lo primero es observar que la unión de las bases de los subespacios propios es l.i., tal como dice el siguiente teorema.

Propiedad 5. *Sea $T : V \rightarrow V$. Si v_1, \dots, v_k son vectores propios asociados a valores propios distintos $\lambda_1, \lambda_2, \dots, \lambda_k$, entonces $\{v_1, \dots, v_k\}$ es l. i.*

Demostración: Haremos una demostración por contradicción. Es una demostración complicada del punto de vista lógico, usted podrá seguirla sólo si toma una actitud muy crítica hacia ella. Es una demostración interesante, aunque no es esencial para la comprensión de esta unidad. Espero que la disfrute.

Primero suponemos que existe un operador $T : V \rightarrow V$ que no cumple la proposición, es decir, que posee un conjunto de valores propios $\lambda_1, \lambda_2, \dots, \lambda_k$, y un conjunto de vectores propios asociados a éstos: $\{v_1, \dots, v_k\}$, que no es l. i. Suponemos sin pérdida de generalidad que λ_1 es no nulo, pues si $k \geq 2$ ha de haber un valor no nulo, y si $k = 1$, el resultado es cierto pues todo conjunto de un solo vector no nulo es l. i., y los vectores propios son siempre no nulos.

Vamos a suponer también que este conjunto es el más pequeño posible que cumple tal propiedad, es decir, vamos a suponer que cualquier subconjunto de éste sí es l. i., siempre podemos suponer esto pues si tuviera un subconjunto l.d., podríamos tomar aquél en vez de éste.¹

Como $\{v_1, \dots, v_k\}$ no es l. i., existen escalares $\alpha_1, \dots, \alpha_k$ tales que

$$\sum_{i=1}^k \alpha_i v_i = \Theta. \quad (1)$$

Ninguno de estos escalares es nulo, ya que si hubiera uno nulo, el conjunto de los vectores restantes seguiría siendo l. d. y sería un conjunto más pequeño, lo cual contradice nuestro supuesto.

Aplicamos ahora T a ambos lados en la ecuación 1, y obtenemos:

$$\begin{aligned} T \left(\sum_{i=1}^k \alpha_i v_i \right) &= \Theta \\ \sum_{i=1}^k \alpha_i T(v_i) &= \Theta \\ \sum_{i=1}^k \alpha_i \lambda_i v_i &= \Theta / \frac{1}{\lambda_1} \\ \alpha_1 v_1 + \sum_{i=2}^k \alpha_i \frac{\lambda_i}{\lambda_1} v_i &= \Theta. \end{aligned} \quad (2)$$

¹Este tipo de razonamiento es una forma escondida de inducción matemática.

Ahora miramos (1)–(2):

$$\sum_{i=2}^k \alpha_i \left(1 - \frac{\lambda_i}{\lambda_1}\right) v_i = \Theta$$

Como $\lambda_i \neq \lambda_1$ para todo $i \geq 2$, tenemos $\alpha_i \left(1 - \frac{\lambda_i}{\lambda_1}\right) \neq 0$, para todo $i \geq 2$, y por lo tanto el conjunto $\{v_2, \dots, v_k\}$ es l. d., lo cual contradice nuestro supuesto, y evidencia una contradicción. \square

vspace0.3 cm

Ejemplo 9. Se vió que la matriz

$$C = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 2 \end{pmatrix}$$

tiene espectro $\sigma(C) = \{-1, 1\}$ y subespacios:

$$S_{-1} = \left\langle \left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \right\} \right\rangle \quad y \quad S_1 = \left\langle \left\{ \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix} \right\} \right\rangle$$

La observación 1 dice que C es invertible y la propiedad 5 nos asegura que el conjunto $\{(1, 0, 0), (0, 1, -1)\}$ es l.i.

En general la unión de las bases de los distintos subespacios de un operador será siempre l.i., propiedad que vamos a explotar fuertemente en lo que sigue.

Diagonalización

El caso más interesante ocurre cuando la unión de las bases de los subespacios propios es una base del espacio completo, pues entonces podremos reescribir la transformación de manera mucho más simple, gracias al concepto de “diagonalización”.

Definición 5. Una transformación lineal $T : V \rightarrow V$ se dice diagonalizable si existe una base B de V tal que $[T]_B^B$ es diagonal.

Si $[T]_B^B$ es diagonal, será mucho más fácil operar con ella, elevarla a potencia e invertible cuando se pueda, será muy simple, bastará con elevar a potencia los valores de la diagonal o invertir los valores de la diagonal. De aquí el interés en encontrar la base que diagonaliza a una transformación lineal. Pero no siempre existe, y el siguiente teorema esclarece este asunto.

Teorema 2. T es diagonalizable si y solo si existe una base de V formada solo por vectores propios de T .

Demostración:

- (\Leftarrow) Sea $\mathcal{B} = \{v_1, \dots, v_n\}$ una base de vectores propios de T asociados a valores propios $\lambda_1, \dots, \lambda_n$, respectivamente (no necesariamente distintos). Esto es, $T(v_i) = \lambda_i v_i$ para todo i . Demostremos que entonces $[T]_{\mathcal{B}}^{\mathcal{B}}$ es diagonal.

En efecto, $[T(v_i)]_{\mathcal{B}} = (0, \dots, 0, \lambda_i, 0, \dots, 0)$, con el λ_i en la i -ésima coordenada, ya que solo el vector v_i se usa para describir v_i en términos de la base \mathcal{B} . Pero esto significa que su i -ésima columna tendrá solo la i -ésima coordenada no nula, justo en la diagonal. Por lo tanto la matriz $[T]_{\mathcal{B}}^{\mathcal{B}}$ será diagonal.

- (\Rightarrow) En el sentido inverso, supongamos que existe una base $\mathcal{C} = \{u_1, \dots, u_n\}$ tal que $[T]_{\mathcal{C}}^{\mathcal{C}} = D$, con D una matriz diagonal. Llámese $\{d_{ii}\}_{i=1}^n$ los coeficientes de la matriz diagonal.

Mostremos que $\{u_1, \dots, u_n\}$ son vectores propios de T . Para ello observamos primero que $[u_i]_{\mathcal{C}} = e_i$, el i -ésimo vector de la base canónica de \mathbb{K}^n , ya que u_i es el i -ésimo vector de la base \mathcal{C} , y $u_i = 1u_i$.

Ahora usaremos el TFAL:

$$[T(u_i)]_{\mathcal{C}} = [T]_{\mathcal{C}}^{\mathcal{C}}[u_i] = De_i = \text{columna } i\text{-ésima de } D.$$

Como D es diagonal, solo el i -ésimo valor de la columna i -ésima de D es no nulo, y como dijimos vale d_{ii} , por lo tanto $T(u_i) = \sum_{j=1}^n d_{ij}u_j = d_{ii}u_i$. Lo cual prueba que u_i es un vector propio asociado a d_{ii} , por lo tanto $\mathcal{C} = \{u_1, \dots, u_n\}$ es una base de V formada solo por vectores propios de T y demuestra el teorema. \square

Ejemplo 10. La matriz $A = \begin{pmatrix} 1 & 2 \\ 3 & 6 \end{pmatrix}$ tiene espectro $\sigma(A) = \{0, 7\}$, con subespacios propios:

$$S_0 = \langle \{(2, -1)\} \rangle \text{ y } S_7 = \langle \{(1, 3)\} \rangle$$

El conjunto $\{(1, 3), (2, -1)\}$ es l. i. y, al tener cardinalidad 2, resulta ser una base de \mathbb{R}^2 .

Luego el teorema 2, dice que A es diagonalizable con $P = \begin{pmatrix} 2 & 1 \\ -1 & 3 \end{pmatrix}$ y $D = \begin{pmatrix} 0 & 0 \\ 0 & 7 \end{pmatrix}$, en efecto:

$$A = PDP^{-1}$$

$$A = \begin{pmatrix} 2 & 1 \\ -1 & 3 \end{pmatrix} \begin{pmatrix} 0 & 0 \\ 0 & 7 \end{pmatrix} \begin{pmatrix} 3/7 & -1/7 \\ 1/7 & 2/7 \end{pmatrix}$$

$$A = \begin{pmatrix} 0 & 7 \\ 0 & 21 \end{pmatrix} \begin{pmatrix} 3/7 & -1/7 \\ 1/7 & 2/7 \end{pmatrix}$$

$$A = \begin{pmatrix} 1 & 2 \\ 3 & 6 \end{pmatrix}$$

Entonces calcular las potencias de A es simple:

$$\begin{aligned} A^n &= \begin{pmatrix} 2 & 1 \\ -1 & 3 \end{pmatrix} \begin{pmatrix} 0 & 0 \\ 0 & 7^n \end{pmatrix} \begin{pmatrix} 3/7 & -1/7 \\ 1/7 & 2/7 \end{pmatrix} \\ A^n &= \begin{pmatrix} 0 & 7^n \\ 0 & 3 \cdot 7^n \end{pmatrix} \begin{pmatrix} 3/7 & -1/7 \\ 1/7 & 2/7 \end{pmatrix} \\ A^n &= \begin{pmatrix} 7^{n-1} & 2 \cdot 7^{n-1} \\ 3 \cdot 7^{n-1} & 6 \cdot 7^{n-1} \end{pmatrix} = 7^{n-1} A \end{aligned}$$

Fórmula que de otra manera sería imposible obtener.

La demostración del teorema anterior es útil cuando se quiere diagonalizar una matriz, es decir cuando queremos conocer explícitamente la matriz P que la diagonaliza. Ahora bien, si solo se desea saber si es o no diagonalizable, basta con calcular las multiplicidades geométricas de cada valor propio, como muestra el siguiente resultado.

Teorema 3. Si g_1, \dots, g_k son las multiplicidades geométricas de los valores propios de T , entonces T es diagonalizable si y solo si $\sum_{i=1}^k g_i = \dim(V)$.

Demostración: De la propiedad 5 sabemos que la unión de las bases de los subespacios propios es l.i. La base del subespacio propio asociado al valor propio λ_i tiene g_i vectores, pues la dimensión de este espacio es g_i , por lo tanto la unión de las bases de todos los subespacios propios tiene $\sum_{i=1}^k g_i$, por lo tanto esta es base de V si y solo si su dimensión es igual a la dimensión de V . \square

Teorema 4. T es diagonalizable si y solo si $p_T(x)$ se factoriza en factores lineales y para cada valor propio λ de T se cumple $g_\lambda = r_\lambda$.

Demostración: De la demostración de la propiedad 4, vemos que si $p_T(x)$ se factoriza en factores lineales, entonces el polinomio $q(x)$ de la demostración no existe (es 1), y por lo tanto se cumple: $\sum_{i=1}^k r_i = n$. Así, si para todo valor propio λ_i se cumple $g_i = r_i$, entonces $\sum_{i=1}^k g_i = n$, de donde el teorema 3 nos permite concluir que T es diagonalizable.

Recíprocamente, si T es diagonalizable, entonces $\sum_{i=1}^k g_i = n$, y como $g_i \leq r_i$, la única manera que esto ocurra es que $\sum_{i=1}^k r_i = n$. \square

Ejemplo 11. Determine si la siguiente matriz es diagonalizable o no:

$$M = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 2 & -1 & 2 & 0 & 0 \\ -1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 3 & 2 \\ 0 & 0 & 0 & -4 & -3 \end{pmatrix}.$$

Solución: Su polinomio característico es:

$$\begin{aligned} p_M(\lambda) &= \det(M - \lambda Id) = (1 - \lambda)(-1 - \lambda)(1 - \lambda)[(3 - \lambda)(-3 - \lambda) + 8] \\ &= -(1 - \lambda)^2(1 + \lambda)(\lambda^2 - 9 + 8) = -(\lambda - 1)^3(\lambda + 1)^2 \end{aligned}$$

Los valores propios serán $\lambda_1 = -1$ y $\lambda_2 = 1$, con multiplicidades algebraicas iguales a 2 y 3 respectivamente. Habrá dos subespacios propios:

$$S_{\lambda_1=-1} = \text{Ker}(M - (-1)Id)$$

$$\begin{aligned} &= \left\{ \begin{pmatrix} x \\ y \\ z \\ s \\ t \end{pmatrix} \in \mathbb{R}^5 : \left(\begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 2 & -1 & 2 & 0 & 0 \\ -1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 3 & 2 \\ 0 & 0 & 0 & -4 & -3 \end{pmatrix} - (-1) \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} \right) \begin{pmatrix} x \\ y \\ z \\ s \\ t \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \right\} \\ &= \left\{ \begin{pmatrix} x \\ y \\ z \\ s \\ t \end{pmatrix} \in \mathbb{R}^5 : \begin{pmatrix} 2 & 0 & 0 & 0 & 0 \\ 2 & 0 & 2 & 0 & 0 \\ -1 & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 4 & 2 \\ 0 & 0 & 0 & -4 & -2 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \\ s \\ t \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \right\} \\ &= \left\{ \begin{pmatrix} x \\ y \\ z \\ s \\ t \end{pmatrix} \in \mathbb{R}^5 : 2x = 0 \wedge 2x + 2z = 0 \wedge -x + 2z = 0 \wedge y \in \mathbb{R} \wedge 4s + 2t = 0 \right\} \\ &= \left\{ \begin{pmatrix} x \\ y \\ z \\ s \\ t \end{pmatrix} \in \mathbb{R}^5 : x = z = 0, y \in \mathbb{R}, t = -2s \right\} \\ &= \left\langle \left\{ \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \\ -2 \end{pmatrix} \right\} \right\rangle. \end{aligned}$$

el cual es de dimensión 2, por tanto $g_1 = 2$. Por otro lado:

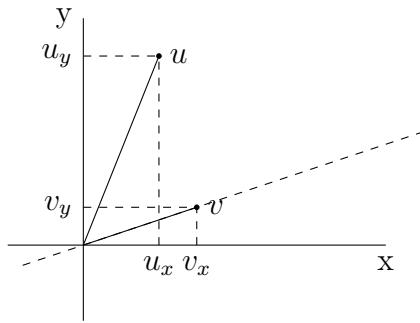
$$\begin{aligned}
S_{\lambda_2=1} &= \text{Ker} (M - (1)Id) \\
&= \left\{ \begin{pmatrix} x \\ y \\ z \\ s \\ t \end{pmatrix} \in \mathbb{R}^5 : \left(\begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 2 & -1 & 2 & 0 & 0 \\ -1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 3 & 2 \\ 0 & 0 & 0 & -4 & -3 \end{pmatrix} - (1) \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} \right) \begin{pmatrix} x \\ y \\ z \\ s \\ t \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \right\} \\
&= \left\{ \begin{pmatrix} x \\ y \\ z \\ s \\ t \end{pmatrix} \in \mathbb{R}^5 : \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 2 & -2 & 2 & 0 & 0 \\ -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2 & 2 \\ 0 & 0 & 0 & -4 & -4 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \\ s \\ t \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \right\} \\
&= \left\{ \begin{pmatrix} x \\ y \\ z \\ s \\ t \end{pmatrix} \in \mathbb{R}^5 : 2x - 2y + 2z = 0 \wedge -x = 0 \wedge 2s + 2t = 0 \right\} \\
&= \left\{ \begin{pmatrix} x \\ y \\ z \\ s \\ t \end{pmatrix} \in \mathbb{R}^5 : x = 0 \wedge y = z \wedge t = -s \right\} \\
&= \left\langle \left\{ \begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \\ -1 \end{pmatrix} \right\} \right\rangle.
\end{aligned}$$

y por tanto $g_2 = 2$. Luego por teorema 3, la matriz no es diagonalizable pues $g_1 + g_2 = 4$ y $n = 5$. También el teorema 4 nos lleva a la misma conclusión ya que $g_2 = 2$ mientras $r_2 = 3$.

Espacios Vectoriales con Producto Interior

En el presente capítulo vamos a generalizar una aspecto de la *geometría euclíadiana* a espacios vectoriales abstractos.

La geometría que ustedes conocen hasta ahora trabaja con rectas, distancias y ángulos en el plano. Acá vamos a exportar esos conceptos a dimensiones superiores. Pero primero recordemos algo de trigonometría.



Del teorema de Pitágoras, se tiene que la magnitud de un vector $u = (u_x, u_y)$ está dada por $\|u\| = \sqrt{u_x^2 + u_y^2}$. Entonces, podemos valernos del Teorema del Coseno para calcular el ángulo \angle_v^u :

$$\|u - v\|^2 = \|u\|^2 + \|v\|^2 - 2\|u\|\|v\|\cos(\angle_v^u)$$

De éste obtenemos que:

$$\begin{aligned} \cos(\angle_v^u) &= \frac{u_x^2 + u_y^2 + v_x^2 + v_y^2 - (u_x - v_x)^2 - (u_y - v_y)^2}{2\|u\|\|v\|} \\ &= \frac{2u_x v_x + 2u_y v_y}{2\|u\|\|v\|} \\ &= \frac{u_x v_x + u_y v_y}{\|u\|\|v\|} \end{aligned}$$

Entonces el ángulo se puede calcular directamente a partir de las coordenadas de los vectores u y v . Además, observamos que en el numerador aparece una expresión interesante que puede escribirse como multiplicación de matrices:

$$u_x v_x + u_y v_y = \begin{pmatrix} u_x & u_y \end{pmatrix} \begin{pmatrix} v_x \\ v_y \end{pmatrix} = \begin{pmatrix} u_x \\ u_y \end{pmatrix}^\top \begin{pmatrix} v_x \\ v_y \end{pmatrix}$$

Este producto es el tema del presente capítulo. Buscamos definir el producto interior en espacios vectoriales abstractos, y de dimensiones arbitrarias. Como buscamos total generalidad y total libertad, y como lo que nos gusta del producto interno son sus “cualidades”, no su forma precisa, entonces lo definiremos a través de las propiedades que nos interesa que cumpla.

Las propiedades que escogemos para definirlo son el resultado de décadas de investigación matemática, que logra sintetizar lo deseado de manera óptima. Al igual que en el caso de los espacios vectoriales, la definición que presentaremos asegura que todo lo demás se pueda obtener “gratis”. Este enfoque abstracto tiene la ventaja de otorgarnos la libertad de definir varios productos internos para un mismo espacio según nuestra conveniencia, lo cual puede ser pertinente en algunas aplicaciones *no* geométricas.

Este capítulo se restringe siempre a los cuerpos \mathbb{R} y \mathbb{C} , dado que el concepto de producto interno pierde sentido en otros cuerpos.

Definición 1 (Producto Interior): *Dado un espacio vectorial sobre un cuerpo \mathbb{K} , con $\mathbb{K} = \mathbb{R}$ o \mathbb{C} , decimos que la operación $\langle \cdot ; \cdot \rangle : V \times V \rightarrow \mathbb{K}$ es un producto interno (p. i.) si satisface las siguientes propiedades.*

1. $(\forall u, v, w \in V)(\forall \alpha \in \mathbb{K})$
 - a) $\langle \alpha u; v \rangle = \alpha \langle u; v \rangle$
 - b) $\langle u + v; w \rangle = \langle u; w \rangle + \langle v; w \rangle$
2. $(\forall u, v \in V) \langle u; v \rangle = \overline{\langle v; u \rangle}$
3. $(\forall u \in V) \langle u; u \rangle \geq 0$
4. $(\forall u \in V) [\langle u; u \rangle = 0 \Leftrightarrow u = \Theta]$

Observamos que la propiedad 2 asegura que $\langle u; u \rangle \in \mathbb{R}$, lo que da sentido a la propiedad 3. También observamos que si $\mathbb{K} = \mathbb{R}$, entonces el *conjugado* queda sin efecto en la propiedad 2, y en ese caso la operación es simétrica.

Ejemplo 1 Consideré \mathbb{R}^3 , la aplicación $\langle \cdot; \cdot \rangle : \mathbb{R}^3 \times \mathbb{R}^3 \rightarrow \mathbb{R}$ definida como:

$$\langle x; y \rangle = x_1 y_1 + x_2 y_2 + x_3 y_3$$

donde $x = (x_1, x_2, x_3)$ e $y = (y_1, y_2, y_3)$, es producto interior sobre \mathbb{R}^3 .

Demostración:

Sea $\alpha \in \mathbb{R}$ y sean $v = (v_1, v_2, v_3)$, $u = (u_1, u_2, u_3)$ y $w = (w_1, w_2, w_3) \in \mathbb{R}^3$.

$$1.a) \langle \alpha u; w \rangle = \langle (\alpha u_1, \alpha u_2, \alpha u_3); (w_1, w_2, w_3) \rangle = (\alpha u_1)w_1 + (\alpha u_2)w_2 + (\alpha u_3)w_3$$

$$= \alpha(u_1 w_1 + u_2 w_2 + u_3 w_3)$$

$$= \alpha \langle u; w \rangle$$

$$1.b) \langle u + v; w \rangle = \langle (u_1 + v_1, u_2 + v_2, u_3 + v_3); (w_1, w_2, w_3) \rangle$$

$$= (u_1 + v_1)w_1 + (u_2 + v_2)w_2 + (u_3 + v_3)w_3$$

$$= (u_1 w_1 + u_2 w_2 + u_3 w_3) + (v_1 w_1 + v_2 w_2 + v_3 w_3)$$

$$= \langle u; w \rangle + \langle v; w \rangle$$

$$2) \langle u; v \rangle = \langle (u_1, u_2, u_3); (v_1, v_2, v_3) \rangle$$

$$= u_1 v_1 + u_2 v_2 + u_3 v_3$$

$$= v_1 u_1 + v_2 u_2 + v_3 u_3$$

$$= \langle v; u \rangle$$

$$3) \langle u; u \rangle = \langle (u_1, u_2, u_3); (u_1, u_2, u_3) \rangle$$

$$= u_1 u_1 + u_2 u_2 + u_3 u_3$$

$$= u_1^2 + u_2^2 + u_3^2 \geq 0$$

$$4) \langle u; u \rangle = 0 \Rightarrow u_1^2 + u_2^2 + u_3^2 = 0 \Rightarrow u_1 = u_2 = u_3 = 0 \Rightarrow u = (0, 0, 0)$$

Por otro lado, $u = (0, 0, 0) \Rightarrow \langle u; u \rangle = u_1^2 + u_2^2 + u_3^2 = 0$.

Por lo tanto, la aplicación definida es un producto interior en \mathbb{R}^3 .

Observación 1 Recordemos algunas propiedades de los números complejos que necesitaremos en este capítulo.

Para todo $x, y \in \mathbb{C}$, y $a, b \in \mathbb{R}$ se cumple lo siguiente.

$$1. \overline{a+bi} = a - bi$$

$$4. \bar{\bar{x}} = x$$

$$7. x = \bar{x} \Leftrightarrow x \in \mathbb{R}$$

$$2. \bar{x}\bar{y} = \bar{x}\bar{y}$$

$$5. |x| = \sqrt{\bar{x}x}$$

$$8. x + \bar{x} = 2\operatorname{Re}(x)$$

$$3. \overline{x+y} = \bar{x} + \bar{y}$$

$$6. x^{-1} = \frac{\bar{x}}{|x|^2}$$

$$9. x - \bar{x} = 2i \operatorname{Im}(x)$$

Ejemplo 2 Considere \mathbb{R}^3 , la aplicación $\langle \cdot; \cdot \rangle : \mathbb{C}^3 \times \mathbb{C}^3 \rightarrow \mathbb{C}$ definida como:

$$\langle x; y \rangle = x_1y_1 + x_2y_2 + x_3y_3$$

donde $x = (x_1, x_2, x_3)$ e $y = (y_1, y_2, y_3) \in \mathbb{C}^3$.

Lo único que aquí cambia es que ahora estamos en el cuerpo de los números complejos.

Por lo tanto la propiedad 1.a) y 1.b) se cumplen igualmente, ya que en su demostración no usamos ninguna propiedad particular de los reales.

La propiedad 2, NO se cumple, ya que $\bar{z} \neq z$ si es que $z \notin \mathbb{R}$, y es claro que en algunos casos se puede tener que $\langle x; y \rangle \notin \mathbb{R}$, por ejemplo si $x = (i, i, i)$, a $y = (0, 1-i, 1+i)$, entonces:

$$\langle x; y \rangle = 0 + i + 1 + i - 1 = 2i = \langle y; x \rangle \neq \overline{\langle y; x \rangle} = -2i.$$

Pero más grave aún, la propiedad 3 no se cumple:

$$\langle x; x \rangle = -1 - 1 - 1 = -3 \not\geq 0$$

Tampoco la propiedad 4:

$$\langle y; y \rangle = (1-i)^2 + (1+i)^2 = 1 - i - i - 1 + 1 + i + i - 1 = 0$$

Por lo tanto en \mathbb{C}^3 tendremos que considerar otra definición.

Definición 2 ■ En \mathbb{R}^n sobre \mathbb{R} , el producto interno usual es $\langle \cdot; \cdot \rangle : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$ definido como:

$$\langle x; y \rangle = \sum_{j=1}^n x_j y_j$$

donde $x = (x_1, x_2, \dots, x_n)$ e $y = (y_1, y_2, \dots, y_n)$.

■ En \mathbb{C}^n sobre \mathbb{C} , el producto interno usual es $\langle \cdot; \cdot \rangle : \mathbb{C}^n \times \mathbb{C}^n \rightarrow \mathbb{R}$ definido como:

$$\langle x; y \rangle = \sum_{j=1}^n x_j \bar{y}_j$$

donde $x = (x_1, x_2, \dots, x_n)$ e $y = (y_1, y_2, \dots, y_n)$.

Propiedad 1 Si $\langle \cdot; \cdot \rangle : V \times V \rightarrow \mathbb{K}$ es un producto interno, entonces se cumple lo siguiente:

$$(\forall u, v, w \in V)(\forall \alpha \in \mathbb{K})$$

1. $\langle u; \alpha v \rangle = \bar{\alpha} \langle u; v \rangle$
2. $\langle w; u + v \rangle = \langle w; u \rangle + \langle w; v \rangle$

Demostración:

$$\langle u; \alpha v \rangle = \overline{\langle \alpha v; u \rangle} = \overline{\alpha \langle v; u \rangle} = \bar{\alpha} \overline{\langle v; u \rangle} = \bar{\alpha} \langle u; v \rangle$$

Por otro lado:

$$\langle u; v + w \rangle = \overline{\langle v + w; u \rangle} = \overline{\langle v; u \rangle + \langle w; u \rangle} = \overline{\langle v; u \rangle} + \overline{\langle w; u \rangle} = \langle u; v \rangle + \langle u; w \rangle$$

□

Propiedad 2 (Desigualdad de Cauchy-Schwarz) Si $\langle \cdot, \cdot \rangle : V \times V \rightarrow \mathbb{K}$ es un producto interno, entonces se cumple la siguiente desigualdad.

$$(\forall u, v \in V) |\langle u; v \rangle| \leq \sqrt{\langle u; u \rangle} \sqrt{\langle v; v \rangle}$$

Demostración:

Partiremos haciendo un tremendo truco: tomaremos $\alpha \in \mathbb{C}$ arbitrario. De la propiedad 3 de p.i. tenemos la siguiente desigualdad, que podemos desarrollar.

$$\begin{aligned} 0 &\leq \langle u - \alpha v; u - \alpha v \rangle \\ &= \langle u; u \rangle - \bar{\alpha} \langle u; v \rangle - \alpha \langle v; u \rangle + |\alpha|^2 \langle v; v \rangle \\ &= \langle u; u \rangle - \bar{\alpha} \langle u; v \rangle - \overline{\alpha \langle v; u \rangle} + |\alpha|^2 \langle v; v \rangle \\ &= \langle u; u \rangle - \bar{\alpha} \langle u; v \rangle - \overline{\alpha} \langle u; v \rangle + |\alpha|^2 \langle v; v \rangle \\ &= \langle u; u \rangle - 2\operatorname{Re}(\bar{\alpha} \langle u; v \rangle) + |\alpha|^2 \langle v; v \rangle \end{aligned}$$

Acá procede el truco maestro que consistente en tomar $\alpha = \frac{\langle u; v \rangle}{\langle v; v \rangle}$, claro que esto solo es posible si $v \neq \Theta$, de manera que el caso $v = \Theta$ se debe analizar por separado, por el momento sigamos.

$$\begin{aligned} 0 &\leq \langle u; u \rangle - 2\operatorname{Re} \left(\frac{\overline{\langle u; v \rangle}}{\langle v; v \rangle} \langle u; v \rangle \right) + \frac{|\langle u; v \rangle|^2}{\langle v; v \rangle^2} \langle v; v \rangle \\ &= \langle u; u \rangle - 2 \frac{|\langle u; v \rangle|^2}{\langle v; v \rangle} + \frac{|\langle u; v \rangle|^2}{\langle v; v \rangle} \\ &= \langle u; u \rangle - \frac{|\langle u; v \rangle|^2}{\langle v; v \rangle}, \end{aligned}$$

lo cual nos lleva a concluir que,

$$|\langle u; v \rangle|^2 \leq \langle v; v \rangle \langle u; u \rangle.$$

Como todos son números positivos, podemos concluir la desigualdad buscada tomando raíz cuadrada, ya que al ser esta una función creciente, respeta la desigualdad. \square

Se dijo que el producto interno servía para calcular ángulos, vemos que la desigualdad de Cauchy-Schwarz permite definir correctamente un coseno.

Definición 3 Dado un e.v. con p.i. V se define el coseno entre dos vectores $u, v \in V$ como sigue.

$$\cos(\angle_v^u) = \frac{\langle u; v \rangle}{\sqrt{\langle u; u \rangle} \sqrt{\langle v; v \rangle}} \in [-1, 1]$$

Cumple la buena característica que si $u = \beta v$, entonces $\cos(\angle_v^u) = \text{sign}(\beta) = \pm 1$.

Por supuesto, en espacios abstractos, este ángulo no tiene un sentido geométrico. Tampoco tendrá un sentido geométrico si no tomamos el producto *usual* o si estamos en \mathbb{C}^n . Pero como vimos al comienzo, en \mathbb{R}^n sí corresponderá al ángulo definido propiamente tal.

Norma y Norma Inducida

Otra noción importante en geometría es el “tamaño”, que en matemática llamamos “norma”. En la sección anterior vimos una fuerte relación entre la magnitud de un vector y el producto interior de este consigo mismo.

Sin embargo, es posible definir tamaños con otros criterios, por ejemplo, en una ciudad la distancia entre dos puntos estará dada por el número de cuadras necesarias para llegar de un punto al otro, lo cual difiere sustancialmente de la longitud del trazo recto que une esos dos puntos. Por lo tanto, necesitaremos abrirnos a criterios más flexibles para definir el concepto de norma.

Tal como hemos hecho hasta ahora, optamos por definirlo en términos de sus propiedades (no de su fórmula).

Definición 4 Dado un espacio vectorial sobre \mathbb{K} , con $\mathbb{K} = \mathbb{R}$ o \mathbb{C} , se dice que la función $\| \cdot \| : V \rightarrow \mathbb{R}_+ \cup \{0\}$ es una norma si satisface las siguientes propiedades.

1. $(\forall u \in V) \|u\| \geq 0$
2. $(\forall u \in V) [\|u\| = 0 \Leftrightarrow u = \Theta]$
3. $(\forall \alpha \in \mathbb{K})(\forall u \in V) \|\alpha u\| = |\alpha| \|u\|$
4. $(\forall u, v \in V) \|u + v\| \leq \|u\| + \|v\|$

Parece razonable pedir que una norma sea no negativa, y que solo el vector nulo tenga norma 0. También es razonable que si ponderamos por un escalar, su norma crezca proporcionalmente.

Finalmente la cuarta propiedad se conoce como *desigualdad triangular*, ya que expresa que siempre en un triángulo, la longitud de cualquier lado es menor o igual que la suma de las longitudes de los otros dos lados.

Ejemplo 3 En $V = \mathbb{R}^n$, la función $\|u\| = \sum_{i=1}^n |u_i|$ es una norma. Justamente esta norma corresponde a la distancia medida en una ciudad cuadriculada como la nuestra, es conocida también como la *distancia de Manhattan*. Verificamos sus propiedades.

1. $\|u\| = \sum_{i=1}^n |u_i| \geq 0$
2. Si $\|u\| = \sum_{i=1}^n |u_i| = 0$, entonces para cada i se debe tener $|u_i| = 0$, por lo tanto $u = \Theta$.
3. $\|\lambda u\| = \sum_{i=1}^n |\lambda u_i| = |\lambda| \sum_{i=1}^n |u_i| = |\lambda| \|u\|$.
4. Observamos primero que si $a, b \in \mathbb{R}$, entonces $|a + b| \leq |a| + |b|$, ya que si a y b tienen el mismo signo se tendrá la igualdad, y en caso contrario $|a + b|$ será estrictamente más pequeño que $|a| + |b|$. Así:

$$\|u + v\| = \sum_{i=1}^n |u_i + v_i| \leq \sum_{i=1}^n |u_i| + |v_i| = \|u\| + \|v\|.$$

Propiedad 3 Dado un espacio vectorial V con producto interior $\langle \cdot, \cdot \rangle$, se cumple que la función definida por

$$(\forall u \in V) \quad \|u\| = \sqrt{\langle u; u \rangle},$$

es una norma. Tal norma se llamará norma inducida por $\langle \cdot, \cdot \rangle$.

Demostración:

1. $\|u\| = \sqrt{\langle u; u \rangle} \geq 0$
2. Si $\|u\| = 0$, entonces $\sqrt{\langle u; u \rangle} = 0$, y por propiedad del producto interior concluimos que $u = \Theta$.
3. $\|\lambda u\| = \sqrt{\langle \lambda u; \lambda u \rangle} = \sqrt{\lambda \bar{\lambda} \langle u; u \rangle} = \sqrt{|\lambda|^2 \langle u; u \rangle} = |\lambda| \|u\|$
4. La desigualdad triangular no es tan simple, nuevamente requiere un truco, y además usa la desigualdad de Cauchy-Schwarz. Lo primero es observar que es equivalente probar que la desigualdad se cumple al cuadrado, demostraríremos entonces que:

$$\|u + v\|^2 \leq (\|u\| + \|v\|)^2$$

En efecto,

$$\begin{aligned}
\|u + v\|^2 &= \langle u + v; u + v \rangle \\
&= \langle u; u \rangle + \langle u; v \rangle + \langle v; u \rangle + \langle v; v \rangle \\
&= \|u\|^2 + \langle u; v \rangle + \overline{\langle u; v \rangle} + \|v\|^2 \\
&= \|u\|^2 + 2\operatorname{Re}(\langle u; v \rangle) + \|v\|^2, \quad z + \bar{z} = 2\operatorname{Re}(z) \\
&\leq \|u\|^2 + 2|\langle u; v \rangle| + \|v\|^2, \quad \operatorname{Re}(z) \leq |z| \\
&\leq \|u\|^2 + 2\|u\|\|v\| + \|v\|^2 \\
&= (\|u\| + \|v\|)^2
\end{aligned}$$

La primera desigualdad se justifica ya que la parte real de un complejo es siempre menor que su módulo.

La segunda desigualdad se vale de la desigualdad de Cauchy-Schwarz. \square

Propiedad 4 Dado un espacio vectorial V con producto interior $\langle \cdot, \cdot \rangle$ y norma inducida $\| \cdot \|$, se cumplen los siguientes:

1. (Pitágoras) $(\forall u, v \in V) \|u + v\|^2 = \|u\|^2 + \|v\|^2$ si y solo si $\operatorname{Re}\langle u, v \rangle = 0$.
2. (Ley del Paralelógramo) $(\forall u, v \in V) \|u + v\|^2 + \|u - v\|^2 = 2\|u\|^2 + 2\|v\|^2$.
3. $\|u\| = \|v\| \Leftrightarrow \langle u + v; u - v \rangle = 0$.

Ejemplo 4 Considere \mathbb{R}^2 y los productos interiores:

$$\langle x; y \rangle_2 = x_1 y_1 + x_2 y_2 \quad \wedge \quad \langle x; y \rangle_* = x_1 y_1 + 3x_2 y_2$$

y considere las siguientes normas:

$$\|x\|_\infty = \max\{|x_i| : i = 1, 2\}, \quad \|x\|_2 = \sqrt{x_1^2 + x_2^2}, \quad \|x\|_* = \sqrt{x_1^2 + 3x_2^2}$$

- a) Calcule las normas anteriores para $(2, -1)$.
- b) Pruebe que la norma $\|x\|_\infty$ no cumple la propiedad del paralelogramo pero que $\|x\|_2$ y $\|x\|_*$ sí lo hacen.
- c) Grafique las bolas de centro en $(0, 0)$ y radio 1 para las 3 normas.
- d) Calcule el producto interior entre $(1, 2)$ y $(2, -1)$ con ambos productos.

Solución:

a)

- $\|(2, -1)\|_\infty = \max\{|2|, |-1|\} = 2$
- $\|(2, -1)\|_2 = \sqrt{(2)^2 + (-1)^2} = \sqrt{5}$
- $\|(2, -1)\|_* = \sqrt{(2)^2 + 3(-1)^2} = \sqrt{7}$

b) La propiedad del paralelogramo es ($\forall u, v \in V$) $\|u + v\|^2 + \|u - v\|^2 = 2\|u\|^2 + 2\|v\|^2$.

Considere $u = (1, 0)$ y $v = (0, 1)$, entonces:

- $\|u + v\|_\infty = \|(1, 1)\|_\infty = \max\{|1|, |1|\} = 1$
- $\|u - v\|_\infty = \|(1, -1)\|_\infty = \max\{|1|, |-1|\} = 1$
- $\|u\|_\infty = \|(1, 0)\|_\infty = \max\{|1|, |0|\} = 1$
- $\|v\|_\infty = \|(0, 1)\|_\infty = \max\{|0|, |1|\} = 1$

luego $\|u + v\|_\infty^2 + \|u - v\|_\infty^2 \neq 2\|u\|_\infty^2 + 2\|v\|_\infty^2$ y por tanto $\|\cdot\|_\infty$ no es una norma inducida (es norma pero no es una norma inducida por algún producto interior).

Por otra parte, sea $u = (u_1, u_2)$ y $v = (v_1, v_2)$:

- $\|u + v\|_*^2 = (u_1 + v_1)^2 + 3(u_2 + v_2)^2$
- $\|u - v\|_*^2 = (u_1 - v_1)^2 + 3(u_2 - v_2)^2$
- $\|u\|_*^2 = u_1^2 + 3u_2^2$
- $\|v\|_*^2 = v_1^2 + 3v_2^2$

luego:

$$\begin{aligned} \|u + v\|_*^2 + \|u - v\|_*^2 &= (u_1 + v_1)^2 + 3(u_2 + v_2)^2 + (u_1 - v_1)^2 + 3(u_2 - v_2)^2 \\ &= u_1^2 + 2u_1v_1 + v_1^2 + 3u_2^2 + 6u_2v_2 + 3v_2^2 + u_1^2 - 2u_1v_1 + v_1^2 + 3u_2^2 - 6u_2v_2 + 3v_2^2 \end{aligned}$$

$$\begin{aligned}
&= 2u_1^2 + 2v_1^2 + 6u_2^2 + 6v_2^2 \\
&= 2(u_1^2 + 3u_2^2) + 2(v_1^2 + 3v_2^2) \\
&= 2\|u\|_*^2 + 2\|v\|_*^2
\end{aligned}$$

c) Ver video adjunto.

d)

- $\langle(1, 2); (2, -1)\rangle_2 = (1)(2) + (2)(-1) = 0$
- $\langle(1, 2); (2, -1)\rangle_* = (1)(2) + 3(2)(-1) = -4$

Ejemplo 5 En el espacio vectorial $\mathcal{P}_n(\mathbb{R})$ sobre \mathbb{R} hay dos productos internos clásicos, o más bien dicho, dos familias de productos internos clásicos.

Dados $a < b \in \mathbb{R}$, fijos, se define el producto:

$$\langle p; q \rangle_{[a,b]} = \int_a^b p(x)q(x)dx.$$

Resulta que es un producto interno, ya que claramente es lineal respecto a $p(x)$, es simétrico, además, dado que $p(x)^2 \geq 0$, resulta que la integral de $p(x)^2$ también es mayor o igual a 0.

La última propiedad de producto interno es más delicada, para verla, supongamos que $\int_a^b p(x)^2 dx = 0$, como $p(x)^2 \geq 0$, y dado que p es una función continua, un teorema de cálculo nos dice que la única manera que esta integral sea nula es que $\forall x \in [a, b]$, $p(x) = 0$ (no podemos decir nada de lo que pasa con p fuera de $[a, b]$). Pero como p es un polinomio, solo puede anularse en una cantidad finita de puntos, salvo que sea el polinomio nulo, n a lo más. El intervalo $[a, b]$ tiene infinitos puntos, por lo tanto deducimos que p debe ser el polinomio nulo en este caso.

Dados $n + 1$ puntos diferentes en \mathbb{R} : x_0, x_1, \dots, x_n , se define el producto:

$$\langle p; q \rangle_{\{x_0, x_1, \dots, x_n\}} = \sum_{i=0}^n p(x_i)q(x_i).$$

Resulta que es un producto interno en $\mathcal{P}_n(\mathbb{R})$, por argumentos muy similares a los anteriores, notar que es importante tomar $n + 1$, y no menos, para que así el último argumento funcione.

Un caso particular de este producto en $\mathcal{P}_2(\mathbb{R})$ sería por ejemplo:

$$\langle p; q \rangle_{\{-1,0,1\}} = p(-1)q(-1) + p(0)q(0) + p(1)q(1).$$

Ejemplo 6 Por su parte, en el subespacio $\mathcal{M}_{m \times n}(\mathbb{C})$ consideramos un producto análogo al producto usual de \mathbb{C}^n .

$$\langle A; B \rangle = \sum_{i=1}^m \sum_{j=1}^n A_{ij} \overline{B_{ij}}$$

Es un producto interno por las mismas razones que lo es el producto usual de \mathbb{C}^n , ya que funciona como si miráramos a la matriz como una gran columna.

Resulta que este producto también se puede escribir de una manera más sucinta, usando la función *traza*.

$$tr : \mathcal{M}_{m \times m}(\mathbb{C}) \rightarrow \mathbb{C}, tr(M) = \sum_{i=1}^n M_{ii}$$

La traza es una función lineal de $\mathcal{M}_{m \times m}(\mathbb{R})$ en \mathbb{R} , que cumple con las siguientes propiedades.

1. Si $\alpha \in \mathbb{C}$, $tr(\alpha M) = \alpha tr(M)$.
2. Si $M, H \in \mathcal{M}_{m \times m}(\mathbb{C})$, $tr(M + H) = tr(M) + tr(H)$.
3. Si $M \in \mathcal{M}_{m \times m}(\mathbb{C})$, $tr(\overline{M}) = \overline{tr(M)}$.
4. Si $M \in \mathcal{M}_{m \times m}(\mathbb{C})$, $tr(M^\top) = tr(M)$.
5. Si $L \in \mathcal{M}_{k \times l}(\mathbb{C})$, $P \in \mathcal{M}_{l \times k}(\mathbb{C})$ $tr(LP) = tr(PL) = \sum_{i=1}^k \sum_{j=1}^l L_{ij} P_{ij}$.

Gracias a estas propiedades tenemos la siguiente igualdad, que puede ser útil para hacer resultados teóricos respecto al producto interno usual de las matrices.

$$\langle A; B \rangle = tr(A^\top \overline{B}).$$

Por ejemplo, en $\mathcal{M}_{2 \times 3}(\mathbb{C})$, tenemos.

$$\begin{aligned} \left\langle \begin{pmatrix} 1 & 0 & 1-i \\ 1 & -1 & 1+i \end{pmatrix}; \begin{pmatrix} 0 & 1+2i & -3 \\ 0 & i & -i \end{pmatrix} \right\rangle &= 0 + 0 + -3(1-i) + 0 - 1(-i) + (1+i)i \\ &= -3 + 3i + i + i - 1 \\ &= -4 + 5i \\ tr \left(\begin{pmatrix} 1 & 1 \\ 0 & -1 \\ 1-i & 1+i \end{pmatrix} \left(\begin{pmatrix} \bar{0} & \bar{1+2i} & \bar{-3} \\ \bar{0} & \bar{i} & \bar{-i} \end{pmatrix} \right) \right) &= tr \left(\begin{pmatrix} 1 & 1 \\ 0 & -1 \\ 1-i & 1+i \end{pmatrix} \begin{pmatrix} 0 & 1-2i & -3 \\ 0 & -i & i \end{pmatrix} \right) \\ &= tr \begin{pmatrix} 0 & ? & ? \\ ? & i & ? \\ ? & ? & -3(1-i) + (1+i)i \end{pmatrix} \\ &= 0 + i - 3 + 3i + i - 1 \\ &= -4 + 5i \end{aligned}$$

Ortogonalidad, y distancia

Como dijimos antes, gracias a la desigualdad de Cauchy-Schwarz podemos considerar el siguiente número y llamarle “coseno”, y deducir un ángulo entre vectores a partir de éste:

$$\cos(\angle_v^u) = \frac{\langle u; v \rangle}{\|u\|\|v\|}.$$

Cumple con las buenas propiedades deseado y en particular permite concebir los ángulos de $\frac{\pi}{2}$ y $-\frac{\pi}{2}$ como aquellos cuyo coseno es 0, definiendo así la importante noción de *perpendicularidad* u *ortogonalidad*.

Definición 1 Dado un espacio vectorial V con producto interior $\langle \cdot, \cdot \rangle$ y norma inducida $\| \cdot \|$, decimos que:

1. u y v son ortogonales si $\langle u; v \rangle = 0$ y entonces denotamos $u \perp v$,
2. el conjunto $\{u_1, \dots, u_k\}$ es ortogonal si $(\forall i \neq j)$ $u_i \perp u_j$ y $(\forall i \in \{1, \dots, k\})$ $u_i \neq \Theta$,
3. el conjunto $\{u_1, \dots, u_k\}$ es ortonormal si es ortogonal y $(\forall i \in \{1, \dots, k\})$ $\|u_i\| = 1$.

Ejemplo 1 Considerando los productos que vimos la clase pasada:

$$\langle x; y \rangle_2 = x_1y_1 + x_2y_2 \quad \wedge \quad \langle x; y \rangle_* = x_1y_1 + 3x_2y_2$$

y tomamos dos vectores cuales quiera; $(1, 2)$ y $(6, -1)$ con ambos productos.

- $\langle (1, 2); (6, -1) \rangle_2 = (1)(6) + (2)(-1) = 4$
- $\langle (1, 2); (6, -1) \rangle_* = (1)(6) + 3(2)(-1) = 0$

Para saber si dos vectores son ortogonales basta calcular su producto interno, si da 0 son ortogonales, si no, no.

Vemos que la propiedad de ortogonalidad depende del producto que una toma como base.

$(1, 2)$ y $(6, -1)$ son ortogonales respecto a $\langle x; y \rangle_*$, pero no respecto a $\langle x; y \rangle_2$.

La noción de conjunto ortogonal exige dos cosas: que todos los productos entre vectores distintos sean nulos y que Θ no esté en el conjunto.

Con el producto usual de \mathbb{C}^3 tenemos que el conjunto $\{(1, i, 1), (i, 1, 0), (1, i, -2)\}$ es ortogonal:

- $\langle(1, i, 1); (i, 1, 0)\rangle = (1)(\bar{i}) + (i)(\bar{1}) + (1)(\bar{0}) = -i + i + 0 = 0$
- $\langle(1, i, 1); (1, i, -2)\rangle = (1)(\bar{1}) + (i)(\bar{i}) + (1)(\bar{-2}) = 1 + 1 - 2 = 0$
- $\langle(1, i, -2); (i, 1, 0)\rangle = (1)(\bar{i}) + (i)(\bar{1}) + (-2)(\bar{0}) = -i + i + 0 = 0$
- $\langle(i, 1, 0); (1, i, 1)\rangle = \overline{\langle(1, i, 1); (i, 1, 0)\rangle} = \bar{0} = 0$, etc.

Pero no es ortonormal ya que $\langle(1, i, 1); (1, i, 1)\rangle = 1\bar{1} + i\bar{i} + 1\bar{1} = 1 + 1 + 1 = 3$, entonces $\|(1, i, 1)\| = \sqrt{3} \neq 1$.

Sin embargo $\{\frac{1}{\sqrt{3}}(1, i, 1), \frac{1}{\sqrt{2}}(i, 1, 0), \frac{1}{\sqrt{6}}(1, i, -2)\}$ sí es ortonormal. Un conjunto ortonormal se obtiene de un conjunto ortogonal dividiendo cada vector por su norma, acción que se denomina “normalizar”.

Propiedad 1 Si $\{u_1, \dots, u_k\}$ es ortogonal, entonces es l.i.

Demostración. Supongamos que no es así, y que existen escalares $\lambda_1, \lambda_2, \dots, \lambda_k \in \mathbb{K}$ no todos nulos tales que

$$\sum_{i=1}^k \lambda_i u_i = \Theta$$

Si aplicamos producto interno a ambos lados por uno de los vectores u_j tal que $\lambda_j \neq 0$, obtenemos.

$$\begin{aligned} \left\langle \sum_{i=1}^k \lambda_i u_i; u_j \right\rangle &= \langle \Theta; u_j \rangle \\ \sum_{i=1}^k \lambda_i \langle u_i; u_j \rangle &= 0 \\ \lambda_j \langle u_j; u_j \rangle &= 0 \end{aligned}$$

Como u_j es ortogonal a todo u_i con $i \neq j$, todos los productos de la suma valen 0, salvo $\langle u_j; u_j \rangle$. Sin embargo, como la suma completa vale 0, podemos concluir que $\lambda_j \|u_j\| = 0$, y como $u_j \neq \Theta$, concluimos que $\lambda_j = 0$, lo cual es una contradicción. \square

Ejemplo 2 Por ejemplo, el siguiente conjunto es ortogonal respecto al producto interno usual de matrices: $\left\{ \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}, \begin{pmatrix} 1 & -1 \\ 1 & -1 \end{pmatrix} \right\}$. Por lo tanto de inmediato sabemos que es l.i.

Por otra parte, observamos que esta base no es ortonormal, ya que la norma inducida por el producto interno usual de matrices, que consiste en tomar la raíz de la suma de los cuadrados de cada coeficiente, es raíz de $4 = 2$.

Propiedad 2 Si $\{u_1, \dots, u_k\}$ es ortogonal, entonces es base de $U = \langle\{u_1, \dots, u_k\}\rangle$ y para todo $v, w \in U$ se cumple que la i -ésima coordenada de v en la base $\{u_1, \dots, u_k\}$ es $\frac{\langle v; u_i \rangle}{\|u_i\|^2}$ y además:

$$\langle v; w \rangle = \sum_{i=1}^k \frac{\langle v; u_i \rangle \overline{\langle w; u_i \rangle}}{\|u_i\|^2}, \quad y \quad \|v\|^2 = \sum_{i=1}^k \frac{|\langle v; u_i \rangle|^2}{\|u_i\|^2}.$$

Demostración. Supongamos que las coordenadas de v respecto a $\{u_1, \dots, u_k\}$ son $(\alpha_1, \alpha_2, \dots, \alpha_k)$, es decir: $\sum_{i=1}^k \alpha_i u_i = v$, si, como antes, hacemos producto con un vector u_j cualquiera de la base y obtenemos: $\langle v; u_j \rangle = \left\langle \sum_{i=1}^k \alpha_i u_i; u_j \right\rangle = \sum_{i=1}^k \alpha_i \langle u_i; u_j \rangle = \alpha_j \langle u_j; u_j \rangle$. De donde es claro que

$$\alpha_j = \frac{\langle v; u_j \rangle}{\langle u_j; u_j \rangle} = \frac{\langle v; u_j \rangle}{\|u_j\|^2}.$$

Usamos esto para el siguiente cálculo:

$$\langle v; w \rangle = \left\langle \sum_{i=1}^k \frac{\langle v; u_i \rangle}{\|u_i\|^2} u_i; \sum_{j=1}^k \frac{\langle w; u_j \rangle}{\|u_j\|^2} u_j \right\rangle = \sum_{i=1}^k \sum_{j=1}^k \frac{\langle v; u_i \rangle}{\|u_i\|^2} \frac{\overline{\langle w; u_j \rangle}}{\|u_j\|^2} \langle u_i; u_j \rangle = \sum_{i=1}^k \frac{\langle v; u_i \rangle \overline{\langle w; u_i \rangle}}{\|u_i\|^4} \|u_i\|^2.$$

El último resultado se obtiene directamente de aplicar éste a $w = v$ y que $\|v\|^2 = \langle v; v \rangle$. \square

Este resultado nos dice que solo con la ortogonalidad obtenemos la independencia lineal de manera gratuita, por lo tanto obtenemos una base des espacio que el conjunto genera.

Además, nos da una manera directa de obtener las coordenadas, sin necesidad de resolver ningún sistema lineal, sólo calcular dos productos y ya está. Ese es la mayor ventaja de tener una *base ortogonal*.

Por otra parte, también nos dice que el producto entre dos vectores cualesquiera, respecto a un producto cualquiera, tiene la forma del producto interno usual entre las coordenadas de los vectores respecto a la base ortogonal. De alguna manera, todos los productos son como el producto usual si uno hace el cambio de coordenadas apropiado.

Ejemplo 3 Consideremos $\mathcal{B} = \left\{ \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}, \begin{pmatrix} 1 & -1 \\ 1 & -1 \end{pmatrix} \right\}$ y $S = \langle \mathcal{B} \rangle$. Ya sabemos que \mathcal{B} es base de S . Si tomamos $A = \begin{pmatrix} 1 & 2 \\ 3 & 0 \end{pmatrix} \in S$, ¿cómo obtenemos sus coordenadas respecto a \mathcal{B} ? Tiene 3 coordenadas $(\alpha_1, \alpha_2, \alpha_3)$, gracias a que \mathcal{B} es ortogonal, las obtenemos como sigue:
 $\alpha_1 = \frac{1}{4} \left\langle \begin{pmatrix} 1 & 2 \\ 3 & 0 \end{pmatrix}; \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} \right\rangle = \frac{1}{4}(1 - 2 - 3) = -1$
 $\alpha_2 = \frac{1}{4} \left\langle \begin{pmatrix} 1 & 2 \\ 3 & 0 \end{pmatrix}; \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \right\rangle = \frac{1}{4}(1 + 2 + 3) = \frac{3}{2}$

$$\alpha_3 = \frac{1}{4} \left\langle \begin{pmatrix} 1 & 2 \\ 3 & 0 \end{pmatrix}; \begin{pmatrix} 1 & -1 \\ 1 & -1 \end{pmatrix} \right\rangle = \frac{1}{4}(1 - 2 + 3) = \frac{1}{2}$$

Así:

$$A = \begin{pmatrix} 1 & 2 \\ 3 & 0 \end{pmatrix} = -1 \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} + \frac{3}{2} \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} + \frac{1}{2} \begin{pmatrix} 1 & -1 \\ 1 & -1 \end{pmatrix}.$$

Proyección

Definición 2 Dado un s.e.v. U , se define la proyección ortogonal de v en U como el vector $u \in U$ que está más cerca de v , es decir el vector $u \in U$ que cumple:

$$\forall w \in U, \|v - u\| \leq \|v - w\|,$$

tal vector lo denotamos por $u = \text{Proy}_U(v)$.

En general no es fácil encontrar u , es un proceso de búsqueda, además, a priori, no tendría por qué haber un único u , la definición podría no tener sentido entonces... sin embargo sí lo tiene gracias al siguiente teorema, al menos en aquellos subespacios que posean una base ortogonal.

Teorema 1 Si $\|\cdot\|$ es una norma inducida por un p.i. $\langle \cdot; \cdot \rangle$, U es un s.e.v. generado por el conjunto ortogonal $\{u_1, \dots, u_k\}$, y $v \in V$, entonces se cumplen las siguientes propiedades.

1. $\forall y \in U, y \perp (v - \text{Proy}_U(v))$ respecto a $\langle \cdot; \cdot \rangle$

$$2. \text{Proy}_U(v) = \sum_{i=1}^k \frac{\langle v, u_i \rangle}{\|u_i\|^2} u_i$$

Demostración Primero definamos $x = \sum_{i=1}^k \frac{\langle v, u_i \rangle}{\|u_i\|^2} u_i$ y observemos que está en U , es claro ya que se trata de una c.l. de elementos de U .

Esta definición nos dice también que las coordenadas de x respecto a la base de U son justamente los $\frac{\langle v, u_i \rangle}{\|u_i\|^2}$ pero estas también son iguales a $\frac{\langle x, u_i \rangle}{\|u_i\|^2}$ por la propiedad anterior. Entonces $\langle v, u_i \rangle = \langle x, u_i \rangle$, para todo $i \in \{1, \dots, k\}$.

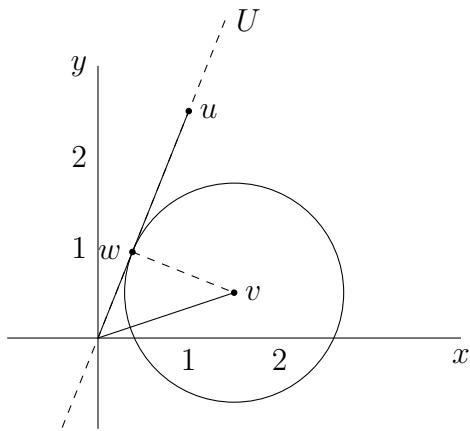
Esta igualdad se hereda a todos los vectores de U , en efecto si $y = \sum_{i=1}^k \alpha_i u_i$, entonces $\langle v, y \rangle = \langle v, \sum_{i=1}^k \alpha_i u_i \rangle = \sum_{i=1}^k \overline{\alpha_i} \langle v, u_i \rangle = \sum_{i=1}^k \overline{\alpha_i} \langle x, u_i \rangle = \langle x, \sum_{i=1}^k \alpha_i u_i \rangle = \langle x, y \rangle$.

Esto implica directamente que $\langle v - x, y \rangle = 0$, es decir que $(v - x) \perp y$, para cualquier $y \in U$ (Parte 1).

Tomemos ahora un vector $w \in U$, vamos a demostrar que x cumple lo necesario para ser la proyección de v en U , es decir, que x está más cerca de v que w . Primero observemos que $v - w = (v - x) + (x - w)$, pero $x - w \in U$, entonces por lo demostrado antes resulta que $(x - w) \perp (v - x)$, así podemos aplicar pitágoras a esta suma de vectores:

$$\|v - w\|^2 = \|v - x\|^2 + \|x - w\|^2 \geq \|v - x\|^2, \text{ por lo tanto } \|v - w\| \geq \|v - x\|, \text{ para todo } w \in U. \quad \square$$

En otras palabras, el teorema anterior nos da dos cosas: 1) un método para calcular la proyección, y 2) nos da una propiedad geométrica muy bonita e intuitiva: la proyección se logra allí donde se forma un ángulo recto con U .



$$\begin{aligned}
 U &= \langle\{(1, 1, 1), (1, 1, 0)\}\rangle \text{ con } u = \left(1; \frac{5}{2}\right) \text{ y } v = \left(\frac{3}{2}; \frac{1}{2}\right) \\
 w &= Proy_U(v) = \frac{\langle v; u \rangle}{\langle u; u \rangle} u = \frac{\frac{3}{2} \cdot 1 + \frac{1}{2} \cdot \frac{5}{2}}{1^2 + (\frac{5}{2})^2} u \\
 &= \frac{11}{29} u = \left(\frac{11}{29}, \frac{55}{58}\right)
 \end{aligned}$$

Ejemplo. $U = \langle\{(1, 1, 1), (1, 1, 0)\}\rangle$, tomemos el p.i. usual.

¿Son ortogonales? $\langle(1, 1, 1); (1, 1, 0)\rangle = 1 + 1 + 0 = 1$ NO, no podemos aplicar la teoría.

Tomo $(0, 0, 1) \in U$ ya que $(0, 0, 1) = (1, 1, 1) - (1, 1, 0)$ entonces

$$U = \langle\{(1, 1, 1), (1, 1, 0), (0, 0, 1)\}\rangle = \langle\{(1, 1, 0), (0, 0, 1)\}\rangle$$

¿Son ortogonales? $\langle(1, 1, 0); (0, 0, 1)\rangle = 0 + 0 + 0 = 0$ SI!!

Calculemos pa proyección de $v = (1, 2, 4)$ sobre U .

$$\begin{aligned}
 Proy_U(1, 2, 4) &= \frac{\langle(1, 2, 4); (1, 1, 0)\rangle}{\|(1, 1, 0)\|^2} (1, 1, 0) + \frac{\langle(1, 2, 4); (0, 0, 1)\rangle}{\|(0, 0, 1)\|^2} (0, 0, 1) \\
 &= \frac{3}{2}(1, 1, 0) + \frac{4}{1}(0, 0, 1) = \left(\frac{3}{2}, \frac{3}{2}, 4\right)
 \end{aligned}$$

$$v - Proy_U(v) = \left(-\frac{1}{2}, \frac{1}{2}, 0\right) \perp \{(1, 1, 1), (1, 1, 0)\}$$

$$\langle(-\frac{1}{2}, \frac{1}{2}, 0), (1, 1, 1)\rangle = 0; \quad \langle(-\frac{1}{2}, \frac{1}{2}, 0), (1, 1, 0)\rangle = 0$$

Observaciones finales:

- Veremos que cualquier espacio de dimensión finita tiene una base ortogonal respecto a su producto interior.
- En cualquier espacio U hay a lo más $\dim(U)$ vectores ortogonales entre sí.
- Tener una base ortogonal de U sirve para calcular más fácilmente las coordenadas de sus vectores respecto a esa base.
- Tener una base ortogonal de U sirve para obtener la proyección de cualquier vector del espacio en el subespacio U .
- La proyección de un v es la mejor aproximación que hay en U de v respecto a la norma inducida por el producto interno.

- Esto solo es verdad cuando tenemos una norma inducida por un producto interno, para las otras normas, la proyección puede no ser única, y es más complicado calcularla.