

Espacios vectoriales

Ingeniería informática Curso: Álgebra

Fernando Chamizo <https://matematicas.uam.es/~fernando.chamizo/>

Comentarios

Estudiaremos subconjuntos de vectores que son cerrados con respecto a la suma y la multiplicación por números. Tales subconjuntos están siempre determinandos por sistemas de ecuaciones lineales homogéneos y admiten “sistemas de referencia” en los que expresar económicamente cada vector. También recuperaremos la idea de las matrices como funciones que actúan sobre vectores.

1. Subespacios del espacio euclídeo

En diferentes aplicaciones dentro y fuera de las matemáticas aparecen conjuntos en los que es posible sumar elementos y multiplicarlos por números, en este sentido se comportan como los vectores de toda la vida. Ejemplos físicos son las fuerzas o las ondas, para las que la suma corresponde a una superposición. Con esta motivación, a finales del siglo XIX se creó una estructura algebraica que hoy llamamos *espacio vectorial*. Lo que se pide para aplicar esta denominación a un conjunto es una lista de propiedades básicas que deben satisfacer la suma y la multiplicación por números que se hayan definido, incluso especifica qué cosas se pueden considerar números (en muchas aplicaciones informáticas de los espacios vectoriales es importante que los elementos de \mathbb{Z}_p con p primo sean números admisibles).

La definición matemática completa de espacio vectorial es bastante abstracta y complicada (si tienes interés, se resume en un comentario más adelante). En este curso solo vamos a trabajar con espacios vectoriales que están dentro de \mathbb{R}^n , el espacio euclídeo habitual, lo cual simplifica mucho las cosas porque \mathbb{R}^n con la suma de vectores y el producto por números reales, llamados *escalares* en este contexto, cumple las propiedades y los subconjuntos con las mismas operaciones las heredan siempre que dichas operaciones estén bien definidas. Concretamente, un subconjunto no vacío $V \subset \mathbb{R}^n$ es un espacio vectorial con la suma de vectores y el producto por escalares habituales si y solo si se cumple

$$1) \quad \vec{u}, \vec{v} \in V \Rightarrow \vec{u} + \vec{v} \in V \quad \text{y} \quad 2) \quad \lambda \in \mathbb{R} \wedge \vec{v} \in V \Rightarrow \lambda\vec{v} \in V.$$

Si uno lo prefiere, estas dos propiedades se pueden reducir a una sola:

$$3) \quad \lambda, \mu \in \mathbb{R} \wedge \vec{u}, \vec{v} \in V \Rightarrow \lambda\vec{u} + \mu\vec{v} \in V.$$

Se tiene que 1), 2) \Leftrightarrow 3), por tanto, es cuestión de gustos utilizar las dos primeras propiedades o la tercera. Las dos primeras son mucho más comunes en los textos porque transmiten la idea de que todo lo que se pide es que las operaciones deben estar bien definidas en el subconjunto. La tercera, tiene la pequeña ventaja de que hay que comprobar una cosa en lugar de dos.

Al ser V con las propiedades anteriores un espacio vectorial con operaciones heredadas de las de \mathbb{R}^n , se dice que es un *subespacio vectorial* de \mathbb{R}^n .

Ejemplo. Decidir si los siguientes conjuntos son subespacios vectoriales de \mathbb{R}^2 :

$$V_1 = \{\vec{x} \in \mathbb{R}^2 : x_1 + x_2 = 0\}, \quad V_2 = \{\vec{x} \in \mathbb{R}^2 : x_1 + x_2 = 1\}, \quad V_3 = \{\vec{x} \in \mathbb{R}^2 : x_1^2 - x_2^2 = 0\},$$

donde se usa la notación habitual $\vec{x} = (x_1, x_2)^t$.

Si $\vec{u} = (u_1, u_2)^t$, $\vec{v} = (v_1, v_2)^t$ y $\lambda, \mu \in \mathbb{R}$, entonces $\lambda\vec{u} + \mu\vec{v} = (\lambda u_1 + \mu v_1, \lambda u_2 + \mu v_2)^t$. Debemos comprobar si un vector de este tipo está siempre en cada V_j cuando \vec{u} y \vec{v} lo están.

Se tiene que $\vec{u}, \vec{v} \in V_1$ equivale a $u_1 + u_2 = v_1 + v_2 = 0$, de aquí, $(\lambda u_1 + \mu v_1) + (\lambda u_2 + \mu v_2) = \lambda(u_1 + u_2) + \mu(v_1 + v_2) = 0 + 0 = 0$, por consiguiente, $\lambda\vec{u} + \mu\vec{v}$ está en V_1 y se concluye que V_1 es subespacio vectorial.

No es cierto que V_2 sea espacio vectorial. Posiblemente el contraejemplo más sencillo de la tercera propiedad (o de la segunda) consista en tomar $\lambda = \mu = 0$ porque $\vec{0} \notin V_2$.

Tampoco V_3 es subespacio vectorial. Por ejemplo, $\vec{u} = (1, 1)^t$ y $\vec{v} = (2, -2)^t$ están en V_3 , pero $\vec{u} + \vec{v} = (3, -1)^t \notin V_3$, lo que contradice la primera propiedad o la tercera con $\lambda = \mu = 1$.

Ejemplo. Comprobar que

$$V = \left\{ \vec{x} \in \mathbb{R}^4 : \begin{array}{l} 2x_1 + x_2 - x_3 + 11x_4 = b_1 \\ 7x_1 + 3x_2 + x_3 + 2x_4 = b_2 \end{array} \right\}$$

es subespacio vectorial de \mathbb{R}^4 si y solo si $b_1 = b_2 = 0$.

Recordando la notación de los sistemas lineales, las ecuaciones se escriben como

$$A\vec{x} = \vec{b} \quad \text{con} \quad A = \begin{pmatrix} 2 & 1 & -1 & 11 \\ 7 & 3 & 1 & 2 \end{pmatrix}.$$

Si $\vec{b} \neq \vec{0}$, como antes $\vec{u} \in V$ y $\lambda = 0$ dan $\lambda\vec{u} = \vec{0}$ y se contradice la segunda propiedad (o la tercera, poniendo también $\mu = 0$). Por tanto, la única posibilidad para que sea espacio vectorial es que $b_1 = b_2 = 0$. En ese caso, $\vec{u}, \vec{v} \in V$ significa $A\vec{u} = A\vec{v} = \vec{0}$ y para $\lambda, \mu \in \mathbb{R}$ cualesquiera se tiene $A(\lambda\vec{u} + \mu\vec{v}) = \lambda A\vec{u} + \mu A\vec{v} = \vec{0} + \vec{0} = \vec{0}$, así que $\lambda\vec{u} + \mu\vec{v} \in V$ y es subespacio vectorial porque hemos verificado la tercera propiedad.

En el ejemplo anterior no hemos usado para nada la forma especial de la matriz A , por tanto, se generaliza inmediatamente para concluir que las soluciones de un sistema lineal de ecuaciones homogéneo forman un subespacio vectorial. Con fórmulas,

Para $A \in \mathcal{M}_{m \times n}$ fijada, $\{\vec{x} \in \mathbb{R}^n : A\vec{x} = \vec{0}\}$ es subespacio vectorial de \mathbb{R}^n .

Más adelante veremos que todos los subespacios de \mathbb{R}^n son de esta forma. Además, el conjunto de soluciones de un sistema no homogéneo no forma un subespacio vectorial.

Hay otra manera natural de construir subespacios vectoriales, esta vez a partir de unos cuantos vectores de \mathbb{R}^n . Una expresión del tipo $\lambda_1 \vec{v}_1 + \lambda_2 \vec{v}_2 + \dots + \lambda_n \vec{v}_n$ con $\vec{v}_j \in \mathbb{R}^n$ y $\lambda_j \in \mathbb{R}$ se dice que es una *combinación lineal* de los vectores $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_n$. Corresponde a “caminar” λ_1 en la dirección de \vec{v}_1 , λ_2 en la de \vec{v}_2 , etc. El conjunto de todas las combinaciones lineales de un conjunto de vectores lo indicaremos con una \mathcal{L} , es decir,

$$\mathcal{L}(\{\vec{v}_1, \dots, \vec{v}_n\}) = \{\lambda_1 \vec{v}_1 + \lambda_2 \vec{v}_2 + \dots + \lambda_n \vec{v}_n : \lambda_1, \dots, \lambda_n \in \mathbb{R}\}.$$

Es muy fácil ver que para cualquier subconjunto finito $C \subset \mathbb{R}^n$ se cumple que $\mathcal{L}(C)$ es subespacio vectorial. Se dice que $\mathcal{L}(C)$ es el *espacio generado* por C .

Por el resultado teórico del tema anterior, las soluciones de un sistema homogéneo son de la forma $\mathcal{L}(C)$, simplemente resolviendo el sistema. Sin embargo, no está claro cómo proceder en sentido contrario. Es decir, dado C hallar un sistema homogéneo $A\vec{x} = \vec{0}$ cuyo conjunto de soluciones sea $\mathcal{L}(C)$. En otras palabras, lo que buscamos son las *ecuaciones* de $\mathcal{L}(C)$. Si $C = \{\vec{v}_1, \dots, \vec{v}_n\}$ el procedimiento que se sigue es aplicar eliminación de Gauss a la matriz $(\vec{v}_1 \dots \vec{v}_n | \vec{x})$, que corresponde al sistema $\lambda_1 \vec{v}_1 + \dots + \lambda_n \vec{v}_n = \vec{x}$, con λ_j las incógnitas, y se estudia qué condiciones surgen en \vec{x} para que tenga solución. Dichas condiciones vendrán de las últimas filas nulas que aparecen al final de la matriz escalonada correspondiente a las n primeras columnas. Posiblemente sea más efectivo repasar el siguiente ejemplo, que intentar desentrañar esta explicación en frío.

Ejemplo. Sea $C = \{\vec{v}_1, \vec{v}_2\}$ con $\vec{v}_1 = (-1, 2, 2, -3)^t$ y $\vec{v}_2 = (5, -8, -6, 23)^t$. Hallar las ecuaciones del subespacio $\mathcal{L}(C)$.

Según el esquema anterior, aplicamos eliminación de Gauss a $(\vec{v}_1 \dots \vec{v}_n | \vec{x})$.

$$\left(\begin{array}{cc|c} -1 & 5 & x_1 \\ 2 & -8 & x_2 \\ 2 & -6 & x_3 \\ -3 & 23 & x_4 \end{array} \right) \rightarrow \left(\begin{array}{cc|c} -1 & 5 & x_1 \\ 0 & 2 & 2x_1 + x_2 \\ 0 & 4 & 2x_1 + x_3 \\ 0 & 8 & -3x_1 + x_4 \end{array} \right) \rightarrow \left(\begin{array}{cc|c} -1 & 5 & x_1 \\ 0 & 2 & 2x_1 + x_2 \\ 0 & 0 & -2x_1 - 2x_2 + x_3 \\ 0 & 0 & -11x_1 - 4x_2 + x_4 \end{array} \right).$$

Se sigue que la condición necesaria y suficiente para que existan λ_1 y λ_2 con $\lambda_1 \vec{v}_1 + \lambda_2 \vec{v}_2 = \vec{x}$, o, equivalentemente, $\vec{x} \in \mathcal{L}(C)$, es que

$$\begin{cases} -2x_1 - 2x_2 + x_3 = 0, \\ -11x_1 - 4x_2 + x_4, \end{cases} \quad \text{esto es, } \mathcal{L}(C) = \{\vec{x} \in \mathbb{R}^4 : A\vec{x} = \vec{0}\} \text{ con } A = \begin{pmatrix} -2 & -2 & 1 & 0 \\ -11 & -4 & 0 & 1 \end{pmatrix}.$$

Para comprobar el resultado es conveniente verificar que \vec{v}_1 y \vec{v}_2 realmente cumplen las ecuaciones.

Ejemplo. Hallar las ecuaciones que definen $\mathcal{L}(C)$ con $C = \{(3, 3, 4)^t, (2, 1, 5)^t\}$.

Procedemos como antes,

$$\left(\begin{array}{cc|c} 3 & 2 & x_1 \\ 3 & 1 & x_2 \\ 4 & 5 & x_3 \end{array} \right) \xrightarrow[\substack{f_2 \mapsto f_2 - f_1 \\ f_3 \mapsto 3f_3 - 4f_1}]{} \left(\begin{array}{cc|c} 3 & 2 & x_1 \\ 0 & -1 & x_2 - x_1 \\ 0 & 7 & 3x_3 - 4x_1 \end{array} \right) \xrightarrow{f_3 \mapsto f_3 + 7f_2} \left(\begin{array}{cc|c} 3 & 2 & x_1 \\ 0 & -1 & x_2 - x_1 \\ 0 & 0 & -11x_1 + 7x_2 + 3x_3 \end{array} \right)$$

y la ecuación que define $\mathcal{L}(C)$ es $-11x_1 + 7x_2 + 3x_3 = 0$.

Si considerásemos un ejemplo como $C = \{(1, 0)^t, (0, 1)^t\}$, no nos saldría ninguna ecuación. Eso es natural, ya que está claro que $\mathcal{L}(C) = \mathbb{R}^2$ y hemos obtenido la trivialidad de que \mathbb{R}^2 sin ninguna restricción es \mathbb{R}^2 . Si uno lo prefiere, puede decir que \mathbb{R}^2 son las soluciones de $A\vec{x} = \vec{0}$ con A la matriz nula.

Comentario. El paso de $\mathcal{L}(C)$ a sus ecuaciones y viceversa está estrechamente relacionado con dos operaciones entre subespacios vectoriales. Si V y W son subespacios de \mathbb{R}^n , su *suma* se define como el subespacio formado por las sumas de los vectores de V y de W , mientras que su *intersección* se define como el subespacio dado por la intersección conjuntista. Es decir,

$$V + W = \{\vec{v} + \vec{w} \in \mathbb{R}^n : \vec{v} \in V \wedge \vec{w} \in W\} \quad \text{y} \quad V \cap W = \{\vec{x} \in \mathbb{R}^n : \vec{x} \in V \wedge \vec{x} \in W\}.$$

No es difícil comprobar que si $V = \mathcal{L}(C_1)$ y $W = \mathcal{L}(C_2)$ entonces $V + W = \mathcal{L}(C_1 \cup C_2)$, mientras que si $V = \{\vec{x} \in \mathbb{R}^n : A_1\vec{x} = \vec{0}\}$ y $W = \{\vec{x} \in \mathbb{R}^n : A_2\vec{x} = \vec{0}\}$ entonces $V \cap W = \{\vec{x} \in \mathbb{R}^n : A\vec{x} = \vec{0}\}$ con A la matriz obtenida añadiendo a las filas de A_1 las de A_2 . El cálculo de sumas e intersección motiva, entonces, las dos presentaciones que hemos visto de los subespacios vectoriales de \mathbb{R}^n y el paso de una a otra.

Comentario. [Avanzado] Es raro que conozcas todos los términos que aparecen en la temida definición rigurosa de espacio vectorial. La plantearemos primero sin paliativos e iremos desentrañando poco a poco el significado de cada cosa.

Sean V un conjunto, K un cuerpo y dos operaciones $+: V \times V \rightarrow V$ y $\cdot: K \times V \rightarrow V$. Se dice que V es un *espacio vectorial* sobre K con estas operaciones si $(V, +)$ es un grupo abeliano y además para $\vec{u}, \vec{v} \in V$ y $a, b \in K$ cualesquiera se cumplen las cuatro propiedades:

$$\begin{aligned} 1 \cdot \vec{u} &= \vec{u}, & (a + b) \cdot \vec{u} &= a \cdot \vec{u} + b \cdot \vec{u}, \\ a \cdot (b \cdot \vec{u}) &= (ab) \cdot \vec{u}, & a \cdot (\vec{u} + \vec{v}) &= a \cdot \vec{u} + a \cdot \vec{v}. \end{aligned}$$

A los elementos de K se les suele llamar *escalares* para diferenciarlos de los de V que son *vectores*. Aparte de esto, un par de observaciones con respecto a la notación: incluso si el símbolo “ \cdot ” no representara el producto de toda la vida, se suele omitir y en la primera propiedad 1 es el elemento unidad de K , el que cumple $1a = a$ para $a \in K$.

Seguramente tu primera duda es qué es un *cuerpo*. Es un conjunto donde hay definidas una suma, una resta, una multiplicación y una división (salvo por cero) con las propiedades habituales. Si esto te

resulta vago, en rigor para que K sea un cuerpo, $(K, +)$ y $(K - \{0\}, \times)$ deben ser grupos abelianos (sigue leyendo) y además se deben cumplir las *propiedades distributivas*, las reglas de operar paréntesis. Aparte de \mathbb{R} , de \mathbb{C} y de \mathbb{Q} , es difícil que se te ocurran más posibilidades para K , aunque existen otras importantes (una antes mencionada que ha aparecido en el curso es \mathbb{Z}_p con p primo). No son válidos $K = \mathbb{N}$ ni $K = \mathbb{Z}$ porque, en general, en el primero no se puede restar sin salirse del conjunto y en el segundo no se puede dividir.

El otro término que te puede resultar extraño en la definición es lo de *grupo abeliano*. Un conjunto con una operación, en nuestro caso V con $+$, se dice que es un grupo abeliano si satisface la *propiedad conmutativa*, la *propiedad asociativa*, tiene elemento neutro y tiene elemento inverso¹. Estos dos últimos se suelen indicar en espacios vectoriales con $\vec{0}$ y $-\vec{u}$.

Las cuatro propiedades al final de la definición podrán parecer arbitrarias, pero se dejan leer, no involucran términos nuevos.

Hay que dejar claro que todo esto no es una manera retorcida de hablar de \mathbb{R}^n o \mathbb{C}^n para matemáticos masoquistas. Por ejemplo, $V = \{(x_1, \dots, x_n) : \sum_{j=1}^n x_j = \bar{0} \text{ con } x_j \in \mathbb{Z}_2\}$ es un espacio vectorial sobre $K = \mathbb{Z}_2$ que representa cadenas de n bits con un número par de unos. La estructura algebraica de espacio vectorial aparece y es relevante en muchos ámbitos de las aplicaciones de las matemáticas tan dispares como la corrección de errores o la física cuántica.

2. Bases y coordenadas

Si alguien nos preguntara por qué el conjunto de puntos del plano representado por \mathbb{R}^2 tiene dimensión 2, responderíamos que necesitamos especificar una x y una y en cada punto, un número de pasos al este y un número de pasos al norte. Alguien con educación matemática elemental nos podría decir que hay cuatro puntos cardinales y nosotros convencerle de que usando números negativos hay redundancia en ellos: 3 pasos al sur son -3 pasos al norte. Todavía hay más redundancia si introducimos las direcciones intermedias de la rosa de los vientos, como el noreste. Un primer problema que debemos abordar antes de hablar de dimensiones es saber detectar y eliminar redundancias. La siguiente definición captura en términos matemáticos la idea de que un conjunto de vectores no sea redundante.

En un espacio vectorial V se dice que los vectores $\vec{v}_1, \dots, \vec{v}_n$ son *linealmente independientes* si la única solución de

$$\lambda_1 \vec{v}_1 + \lambda_2 \vec{v}_2 + \dots + \lambda_n \vec{v}_n = \vec{0}$$

con $\lambda_1, \dots, \lambda_n \in \mathbb{R}$ es $\lambda_1 = \dots = \lambda_n = 0$.

En un alarde de originalidad, cuando hay más soluciones se dice que los vectores son *linealmente dependientes*. En este caso, existe algún $\lambda_i \neq 0$ y se puede despejar \vec{v}_i en términos del resto de los vectores por medio de $\vec{v}_i = -\sum_{j \neq i} \lambda_i^{-1} \lambda_j \vec{v}_j$. Es ahí donde está la redundancia antes mencionada.

¹Si no conoces estas propiedades de memoria, en fórmulas son $\vec{u} + \vec{v} = \vec{v} + \vec{u}$, $\vec{u} + (\vec{v} + \vec{w}) = (\vec{u} + \vec{v}) + \vec{w}$, $\vec{u} + \vec{0} = \vec{u}$ y $\vec{u} + (-\vec{u}) = \vec{0}$.

La definición anterior reduce el estudio de la dependencia o independencia lineal a discutir un sistema homogéneo. En nuestro contexto, la matriz de coeficientes del sistema es $(\vec{v}_1 \vec{v}_2 \dots \vec{v}_n)$, la matriz cuyas columnas son los vectores, y entonces $\vec{v}_1, \dots, \vec{v}_n$ son *linealmente independientes* si y solo si el rango de dicha matriz es n .

Ejemplo. Estudiar si los vectores $\vec{v}_1 = (2, 1, -1)^t$, $\vec{v}_2 = (3, 0, 1)^t$ y $\vec{v}_3 = (-1, 4, -7)^t$ son linealmente independientes.

La ecuación que define la independencia lineal se puede escribir como

$$A\vec{\lambda} = \vec{0} \quad \text{con} \quad A = \begin{pmatrix} 2 & 3 & -1 \\ 1 & 0 & 4 \\ -1 & 1 & -7 \end{pmatrix}, \quad \vec{\lambda} = \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \lambda_3 \end{pmatrix}.$$

Es fácil ver que $\text{rg}(A) = 2$, concretamente aplicando eliminación de Gauss intercambiando las dos primeras filas:

$$\begin{pmatrix} 1 & 0 & 4 \\ 2 & 3 & -1 \\ -1 & 1 & -7 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & 0 & 4 \\ 0 & 3 & -9 \\ 0 & 1 & -3 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & 0 & 4 \\ 0 & 3 & -9 \\ 0 & 0 & 0 \end{pmatrix}.$$

Por tanto, el sistema tiene infinitas soluciones y los vectores son linealmente dependientes.

Ejemplo. Decidir qué ocurre en el ejemplo anterior al cambiar \vec{v}_3 por $\vec{w}_3 = (-1, 4, -6)^t$.

Se tiene que \vec{v}_1 , \vec{v}_2 y \vec{w}_3 son linealmente independientes porque las cuentas en la eliminación de Gauss son las mismas salvo que en cada paso el valor de a_{33} está incrementado en 1. Así el rango es 3 y la solución es única.

Los ejemplos anteriores muestran que para decidir la independencia lineal lo único que hacer es discutir el $A\vec{\lambda} = \vec{0}$ con A la matriz formada por los vectores en columna.

Recordemos que el rango era el número de escalones al aplicar reducción de Gauss. Necesitamos que este coincida con el número de columnas para que el sistema homogéneo asociado tenga solo una solución, la trivial. Por tanto,

El rango de una matriz indica el número máximo de columnas que son linealmente independientes.

De hecho, *los que corresponden a columnas pivote (columnas donde hay pivotes) siempre son linealmente independientes.* En el ejemplo anterior con \vec{v}_1 , \vec{v}_2 , \vec{v}_3 , las columnas pivote son la primera y la segunda, por tanto, \vec{v}_1 y \vec{v}_2 son linealmente independientes.

Comentario. Esta reinterpretación del rango sirve para probar $\text{rg}(A) = \text{rg}(A^t)$ sin apelar a los determinantes. El esquema es el siguiente: según lo dicho, $\text{rg}(A^t)$ es el número máximo de filas linealmente independiente de A , con un poco de trabajo se deduce que tal número no varía con las transformaciones T1, T2 y T3, por consiguiente, $\text{rg}(A^t)$ es también el número de filas linealmente independiente en la

forma escalonada reducida de A y es fácil comprobar, usando que sobre los pivotes hay ceros, que estas son todas las filas no nulas, esto es, el número de escalones $\text{rg}(A)$.

La dependencia lineal de dos vectores se reduce a comprobar que uno es múltiplo del otro, lo cual es fácil de hacer a ojo, sin lápiz y papel.

Ejemplo. Explicar por qué $(2, 3)^t$, $(2025, 2026)^t$, $(-5, 7)^t$ son linealmente independientes dos a dos y son linealmente dependientes conjuntamente.

Claramente ninguno es múltiplo de otro y se sigue la independencia lineal dos a dos. Por otro lado, el sistema $\lambda_1 \vec{v}_1 + \lambda_2 \vec{v}_2 + \lambda_3 \vec{v}_3 = \vec{0}$ con \vec{v}_j los vectores del enunciado tiene matriz con rango menor o igual que dos (hay dos filas) y hay tres incógnitas, lo que implica que hay infinitas soluciones sin hacer ningún cálculo.

Una vez que hemos entendido la independencia lineal, ya estamos preparados para dar una definición matemática que concrete la idea de dimensión a través de un sistema de medida, una elección de puntos cardinales, en el que no hay redundancia y que permite representar cualquier elemento, cualquier observación, en nuestro espacio.

Dado un espacio vectorial V (en este curso, siempre un subespacio de \mathbb{R}^n) se dice que un conjunto $\mathcal{B} \subset V$ es una *base* de V si los vectores de \mathcal{B} son linealmente independientes y $\mathcal{L}(\mathcal{B}) = V$. Si \mathcal{B} tiene n elementos se dice que V tiene *dimensión* n y se indica con $\dim V = n$. Esto es, en un espacio vectorial $\dim V$ es el cardinal de la base.

Se puede demostrar que las bases siempre existen (conviniendo $\mathcal{B} = \emptyset$ para $V = \{\vec{0}\}$, el único espacio vectorial sin vectores linealmente independientes) y que todas tienen el mismo cardinal, por tanto la dimensión está unívocamente definida, no depende de cómo la calculemos.

Para referirse a la segunda propiedad de las bases, $\mathcal{L}(\mathcal{B}) = V$, se suele decir que \mathcal{B} es un *sistema de generadores* de V . La razón es obvia: pedimos que todo el espacio generado por \mathcal{B} sea V .

Comentario. Una cosa por la que suelen pasar de puntillas los textos de matemáticas es que es conveniente que los vectores de las bases tengan una ordenación, por tanto, con todo rigor, más que conjuntos son listas finitas. Este es uno de los pocos casos en que los matemáticos suelen ser imprecisos de forma generalizada en una definición.

En \mathbb{R}^2 la base más sencilla es $\mathcal{B} = \{\vec{e}_1, \vec{e}_2\}$ con $\vec{e}_1 = (1, 0)^t$ y $\vec{e}_2 = (0, 1)^t$, llamada la *base canónica*. Se extiende a \mathbb{R}^n de manera natural tomando \vec{e}_j con un 1 en el lugar j y ceros en el resto. Comprobar que \mathcal{B} es base es sencillo. Los vectores \vec{e}_1 y \vec{e}_2 son linealmente independientes y está claro que cualquier vector $(x, y)^t \in \mathbb{R}^2$ se escribe como $x\vec{e}_1 + y\vec{e}_2$. De ello se deduce que $\dim \mathbb{R}^2 = 2$. En general, se tiene $\dim \mathbb{R}^n = n$.

La parte de la definición que pide que sea sistema de generadores puede ser engorrosa o redundante de comprobar y el siguiente resultado habitualmente nos evita este trabajo:

Si V es un espacio vectorial con $\dim V = n$, entonces $\mathcal{B} \subset V$ es base de V si y solo si \mathcal{B} está formada por n vectores linealmente independientes.

En otras palabras, si sabemos que el número de elementos de la posible base coincide con la dimensión solo hace falta comprobar que sus vectores son linealmente independientes para confirmar que es base.

Ejemplo. Estudiar si $\mathcal{B} = \{(11, 13)^t, (5, -7)^t\}$ es una base de \mathbb{R}^2 .

Lo es porque $\dim \mathbb{R}^2 = 2$ y \mathcal{B} cuenta con dos vectores linealmente independientes (uno no es múltiplo de otro).

En general, siempre que tengamos n vectores linealmente independientes en \mathbb{R}^n , tendremos una base.

La existencia de bases implica que todos los subespacios vectoriales de \mathbb{R}^n son de la forma $\mathcal{L}(C)$ y, según lo visto en el apartado anterior, también son de la forma $\{\vec{x} \in \mathbb{R}^n : A\vec{x} = \vec{0}\}$. Es decir, en nuestro contexto todos los espacios vectoriales son conjuntos de soluciones de sistemas homogéneos.

Fijada una base $\mathcal{B} = \{\vec{v}_1, \dots, \vec{v}_n\}$ de un espacio vectorial V cada vector $\vec{u} \in V$ se escribe de forma única como $\vec{u} = \lambda_1 \vec{v}_1 + \dots + \lambda_n \vec{v}_n$ porque si hubiera dos combinaciones lineales que dieran \vec{u} , al restarlas llegaríamos a una combinación lineal nula no trivial que contradiría la independencia lineal. Estos λ_j se dice que son las *coordenadas* de \vec{u} en la base \mathcal{B} . Incluso a veces se escribe, con un abuso de la notación que evitaremos aquí, $\vec{u} = (\lambda_1, \dots, \lambda_n)^t$ cuando se especifica la base en la que se trabaja.

Por ejemplo, si consideramos la base canónica $\mathcal{B} = \{\vec{e}_1, \vec{e}_2\}$ en \mathbb{R}^2 , que corresponde a las direcciones este y norte, el vector $\vec{u} = (5, 2)^t \in \mathbb{R}^2$ tiene obviamente coordenadas 5 y 2 en esta base, pues $\vec{u} = \vec{e}_1 + 2\vec{e}_2$. Sin embargo, si usamos la base $\mathcal{B}' = \{\vec{v}_1, \vec{v}_2\}$ con $\vec{v}_1 = (1, 1)^t$ y $\vec{v}_2 = (0, 1)^t$, que corresponde a las direcciones noreste y el norte, las coordenadas pasan a ser 5 y -3 , pues $\vec{u} = 5\vec{v}_1 - 3\vec{v}_2$.

La pregunta fundamental a este nivel es cómo hallar una base de un espacio vectorial. La respuesta depende de cómo esté presentado. Si

$$V = \{\vec{x} \in \mathbb{R}^n : A\vec{x} = \vec{0}\} \quad \text{con} \quad A \in \mathcal{M}_{m \times n}(\mathbb{R}),$$

leyendo con ojos matemáticos el resultado de estructura de las soluciones de ecuaciones lineales, entenderemos que, con la notación empleada allí, $\mathcal{B} = \{\vec{v}_1, \dots, \vec{v}_{n-r}\}$ es una base de V . En particular,

$$\dim \{\vec{x} \in \mathbb{R}^n : A\vec{x} = \vec{0}\} = n - \text{rg}(A).$$

Recordemos que estos \vec{v}_j se hallaban mediante reducción de Gauss.

Por otro lado, si $V = \mathcal{L}(C)$ con $C = \{\vec{v}_1, \dots, \vec{v}_n\}$, una base vendrá dada por un subconjunto de C de vectores linealmente independientes con cardinal máximo. Según lo visto antes,

$$\dim V = \text{rg}(A) \quad \text{con} \quad A = (\vec{v}_1 \cdots \vec{v}_n)$$

y una base se consigue eligiendo los \vec{v}_j que corresponden a las columnas pivote.

Sea cual sea la presentación, el cálculo de bases y dimensiones requiere el paso a la forma escalonada de una matriz, entonces es inexcusable aprender la eliminación de Gauss en este curso.

Comentario. Al hilo del cálculo de dimensiones, hay una fórmula sencilla que relaciona las dimensiones de los espacios vectoriales obtenidos con las operaciones de suma e intersección introducidas en un comentario anterior. Es la *fórmula de Grassmann* $\dim(V_1 + V_2) + \dim(V_1 \cap V_2) = \dim V_1 + \dim V_2$.

Ejemplo. Hallar una base de $V = \{\vec{x} \in \mathbb{R}^3 : x_1 + 2x_2 + 3x_3 = 0\}$.

La matriz del sistema es la matriz fila $(1, 2, 3)$ y por tanto $\text{rg}(A) = 1$, lo que implica $\dim V = 3 - 1 = 2$. A nadie sorprenderá que un plano tenga dimensión dos. Asignando parámetros $x_2 = \lambda_1$, $x_3 = \lambda_2$, que corresponden a las columnas no pivote, se tiene que la solución general es

$$\vec{x} = \begin{pmatrix} -2\lambda_1 - 3\lambda_2 \\ \lambda_1 \\ \lambda_2 \end{pmatrix} = \lambda_1 \vec{v}_1 + \lambda_2 \vec{v}_2 \quad \text{con} \quad \vec{v}_1 = \begin{pmatrix} -2 \\ 1 \\ 0 \end{pmatrix} \quad \text{y} \quad \vec{v}_2 = \begin{pmatrix} -3 \\ 0 \\ 1 \end{pmatrix}.$$

Según lo anterior, una base de V es $\mathcal{B} = \{\vec{v}_1, \vec{v}_2\}$.

Ejemplo. Hallar las coordenadas del vector $\vec{u} = (-7, 2, 1)^t$ en la base obtenida en el ejemplo anterior.

El sistema $\vec{u} = \lambda_1 \vec{v}_1 + \lambda_2 \vec{v}_2$ trivialmente conduce a $\lambda_1 = 2$, $\lambda_2 = 1$. Estas son las coordenadas de \vec{u} en la base \mathcal{B} .

Un problema algo diferente es comprobar si un conjunto de vectores dado es base.

Ejemplo. Estudiar si

$$\mathcal{B} = \left\{ \vec{v}_1 = \begin{pmatrix} 3 \\ 1 \\ -2 \\ -2 \end{pmatrix}, \vec{v}_2 = \begin{pmatrix} 0 \\ 1 \\ -1 \\ 0 \end{pmatrix} \right\} \quad \text{es base de} \quad V = \left\{ \vec{x} \in \mathbb{R}^4 : \begin{array}{l} x_1 + x_2 + x_3 + x_4 = 0 \\ x_1 - x_2 - x_3 + 2x_4 = 0 \end{array} \right\}.$$

Se cumple $\mathcal{B} \subset V$ porque \vec{v}_1 y \vec{v}_2 satisfacen las ecuaciones de V . El espacio es de la forma $V = \{\vec{x} \in \mathbb{R}^4 : A\vec{x} = \vec{0}\}$ con A de rango dos, porque un paso de reducción de Gauss da

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

Entonces la dimensión de V es $\dim V = n - \text{rg}(A) = 4 - 2 = 2$ que coincide con el número de vectores de \mathcal{B} , por tanto, para decidir si es base, solo tenemos que estudiar si \vec{v}_1 y \vec{v}_2 son linealmente independientes. Esto es, si $\lambda_1 \vec{v}_1 + \lambda_2 \vec{v}_2 = \vec{0}$ implica $\lambda_1 = \lambda_2 = 0$. Esto se puede comprobar a ojo, viendo que \vec{v}_1 y \vec{v}_2 no son uno múltiplo del otro o resolviendo el sistema.

Una variante tonta del ejemplo anterior sería decidir si $\{\vec{v}_1, \vec{v}_2, (6, -1, -1, -4)^t\}$ es base de V . La respuesta es automáticamente que no, porque $\dim V = 2$ implica que todas las bases deben contener exactamente dos vectores.

Una variante todavía más tonta sería decidir si $\{\vec{v}_1, (0, 1, 1, 0)^t\}$ es base de V . No puede serlo porque el segundo vector ni siquiera está en V , no cumple las ecuaciones.

3. Aplicaciones lineales

Una función entre espacios vectoriales $f : V \rightarrow W$ se dice que es una *aplicación lineal* si preserva las operaciones. Es decir, si

$$1) \quad \vec{u}, \vec{v} \in V \Rightarrow f(\vec{u} + \vec{v}) = f(\vec{u}) + f(\vec{v}) \quad \text{y} \quad 2) \quad \lambda \in \mathbb{R} \wedge \vec{v} \in V \Rightarrow f(\lambda\vec{v}) = \lambda f(\vec{v}).$$

Al igual que en el caso de la definición de subespacio vectorial, estas dos propiedades se pueden integrar en una sola:

$$3) \quad \lambda, \mu \in \mathbb{R} \wedge \vec{u}, \vec{v} \in V \Rightarrow f(\lambda\vec{u} + \mu\vec{v}) = \lambda f(\vec{u}) + \mu f(\vec{v}).$$

En este sentido, las aplicaciones lineales son las funciones que preservan las combinaciones lineales.

Se emplea el nombre *endomorfismo* para indicar $V = W$. Es decir, un endomorfismo es una aplicación lineal de un espacio vectorial en sí mismo.

En este curso nos ocupamos solo de espacios vectoriales V y W que son subespacios de \mathbb{R}^n y ahora nos especializaremos más considerando únicamente aplicaciones lineales $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$.

Comentario. No es una gran especialización, pues lo único que tendríamos que hacer para pasar a subespacios es restringir el dominio y el conjunto imagen. Por otro lado, cualquier aplicación lineal $f : V \rightarrow W$ con $V \subset \mathbb{R}^n$ y $W \subset \mathbb{R}^m$ se puede extender a $\tilde{f} : \mathbb{R}^n \rightarrow \mathbb{R}^m$.

Lo que tienen de particular las aplicaciones lineales $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ es que son exactamente las funciones de la forma $f(\vec{x}) = A\vec{x}$ con $A \in \mathcal{M}_{m \times n}$, son tan sencillas como multiplicar por una matriz. Se dice que A es la *matriz de la aplicación lineal* f . Así, la siguiente función $f : \mathbb{R}^3 \rightarrow \mathbb{R}^2$

$$f(\vec{x}) = \begin{pmatrix} x_1 - 2x_2 + x_3 \\ 2025x_2 \end{pmatrix} \quad \text{es aplicación lineal porque} \quad f(\vec{x}) = \begin{pmatrix} 1 & -2 & 1 \\ 0 & 2025 & 0 \end{pmatrix} \vec{x}.$$

Por otro lado, cambiando $2025x_2$ por $2025x_2^2$ o por $\sin(2025x_2)$ o por x_1x_2 no sería aplicación lineal porque ya no vendría dada por la multiplicación por una matriz. De este modo, en nuestro contexto especializado, la definición original de aplicación lineal pierde fuerza porque decidir si una f lo es se vuelve tan trivial como mirar si es o no de la forma $f(\vec{x}) = A\vec{x}$.

Ejemplo (Teórico). Demostrar que si $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ es aplicación lineal entonces es de la forma $f(\vec{x}) = A\vec{x}$ donde la columna j de A es $f(\vec{e}_j)$ con $\{\vec{e}_1, \dots, \vec{e}_n\}$ la base canónica de \mathbb{R}^n .

Escribiendo $\vec{x} = x_1\vec{e}_1 + \dots + x_n\vec{e}_n$ se tiene $f(\vec{x}) = x_1f(\vec{e}_1) + \dots + x_nf(\vec{e}_n)$ por ser f aplicación lineal (empleando sucesivas veces la tercera propiedad). El segundo miembro se puede escribir en forma matricial como $A\vec{x}$ con $A = (f(\vec{e}_1) \dots f(\vec{e}_n))$.

Hay dos subespacios destacados asociados a cada aplicación lineal. Escribiremos la definición general, aunque, según lo dicho, todos nuestros ejemplos serán con $V = \mathbb{R}^n$ y $W = \mathbb{R}^m$.

Dada una aplicación lineal $f : V \rightarrow W$ se definen el *núcleo* de f y la *imagen* de f como los subespacios de V y W dados respectivamente por

$$\text{Ker}(f) = \{\vec{x} \in V : f(\vec{x}) = \vec{0}\}, \quad \text{Im}(f) = \{\vec{y} \in W : \vec{y} = f(\vec{x}) \text{ con } \vec{x} \in V\}.$$

Este “Ker” es la abreviatura de *kernel*, núcleo en inglés. A veces en textos en español se escribe $\text{Nuc}(f)$ en vez de $\text{Ker}(f)$. La imagen es la imagen conjuntista que ya conocemos. Se cumple

$$\text{Im}(f) = W \Leftrightarrow f \text{ es sobreyectiva} \quad \text{y} \quad \text{Ker}(f) = \{\vec{0}\} \Leftrightarrow f \text{ es inyectiva}.$$

En realidad, lo primero es la propia definición de función sobreyectiva. Respecto a lo segundo, si $\text{Ker}(f) \neq \{\vec{0}\}$, digamos $\vec{0} \neq \vec{v} \in \text{Ker}(f)$, entonces $f(\vec{v}) = f(\vec{0}) = \vec{0}$ y no sería inyectiva. Por otro lado, si f no es inyectiva, $f(\vec{x}) = f(\vec{y})$ con $\vec{x} \neq \vec{y}$ y se deduce que $\vec{x} - \vec{y}$ es un elemento no nulo del núcleo.

Si $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$ es un endomorfismo dado por $f(\vec{x}) = A\vec{x}$ entonces es biyectivo (inyectivo y sobreyectivo) si y solo si A es una matriz invertible, en cuyo caso la función inversa $f^{-1} : \mathbb{R}^n \rightarrow \mathbb{R}^n$ es $f^{-1}(\vec{x}) = A^{-1}\vec{x}$.

A menudo la inyectividad y la sobreyectividad se comprueban examinando dimensiones. Concretamente, una aplicación lineal $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ es inyectiva si y solo si $\dim \text{Ker}(f) = 0$ y es sobreyectiva si y solo si $\dim \text{Im}(f) = m$. La razón para utilizar dimensiones es que si A es la matriz de f

$$\dim \text{Ker}(f) = n - \text{rg}(A) \quad \text{y} \quad \dim \text{Im}(f) = \text{rg}(A).$$

Además, los vectores correspondientes a columnas pivote de A son una base de $\text{Im}(f)$. Por otro lado, una base de $\text{Ker}(f)$ se obtiene resolviendo $A\vec{x} = \vec{0}$ por eliminación de Gauss para obtener los \vec{v}_j del resultado de estructura del conjunto de soluciones.

Ambas fórmulas para la dimensión son consecuencia de resultados anteriores: la primera viene del cálculo de la dimensión de un subespacio de \mathbb{R}^n definido por $A\vec{x} = \vec{0}$ y la segunda del análisis que hicimos sobre el rango y la independencia de los vectores columna.

Ejemplo. Decidir si es biyectivo el endomorfismo

$$f : \mathbb{R}^2 \rightarrow \mathbb{R}^2 \quad \text{dado por} \quad f(\vec{x}) = \begin{pmatrix} 2 & 3 \\ 3 & 5 \end{pmatrix} \vec{x}.$$

Es biyectivo porque un paso de eliminación de Gauss muestra que la matriz tiene rango 2 y, por consiguiente, $\dim \text{Ker}(f) = 2 - 2 = 0$ (es inyectiva) y $\dim \text{Im}(f) = 2$ (es sobreyectiva). Por lo dicho anteriormente, este resultado está ligado a que la matriz sea invertible.

Ejemplo. Estudiar si la siguiente aplicación lineal es inyectiva o sobreyectiva:

$$f : \mathbb{R}^4 \longrightarrow \mathbb{R}^2 \quad \text{dada por} \quad f(\vec{x}) = \begin{pmatrix} 1 & 3 & 1 & 0 \\ 1 & 2 & 3 & 0 \end{pmatrix} \vec{x}.$$

La eliminación de Gauss solo requiere un paso:

$$A = \begin{pmatrix} 1 & 3 & 1 & 0 \\ 1 & 2 & 3 & 0 \end{pmatrix} \xrightarrow{f_2 \mapsto f_2 - f_1} \begin{pmatrix} 1 & 3 & 1 & 0 \\ 0 & -1 & 2 & 0 \end{pmatrix}$$

obteniéndose $\text{rg}(A) = 2$, dos escalones. Según las fórmulas, $\dim \text{Im}(f) = 2$ y $\dim \text{Ker}(f) = 4 - 2 = 2$. Por tanto, f es sobreyectiva y no es inyectiva.

Comentario. Si tienes el superpoder de la intuición matemática, la falta de inyectividad en el ejemplo anterior está clara sin cálculos: si “aplastamos” \mathbb{R}^4 en \mathbb{R}^2 , vamos a tener que pegar imágenes de vectores porque cuatro dimensiones no caben en dos.

Ejemplo. Hallar una base del núcleo y de la imagen en el ejemplo anterior.

Según lo visto, se puede obtener una base de la imagen considerando las columnas pivote. Estas son las que dan lugar a los inicios de un escalón, esto es, la primera y la segunda. Con ello, una base válida de $\text{Im}(f)$ es $\{(1, 1)^t, (3, 2)^t\}$. Es importante notar que las columnas a elegir son las de A , no las obtenidas tras aplicar el proceso de eliminación de Gauss. Este proceso solo sirve para calcular dónde están los pivotes.

Una base de $\text{Ker}(f)$ es por definición una base del subespacio determinado por $A\vec{x} = \vec{0}$ y ya sabemos cómo se halla. Asignamos a las incógnitas correspondientes a columnas no pivote parámetros arbitrarios y despejamos las otras por sustitución regresiva:

$$\begin{pmatrix} 1 & 3 & 1 & 0 \\ 0 & -1 & 2 & 0 \end{pmatrix} \Rightarrow \begin{matrix} x_3 = \lambda, & x_4 = \mu, \\ x_2 = 2\lambda, & x_1 = -7\lambda \end{matrix} \Rightarrow \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \lambda \begin{pmatrix} -7 \\ 2 \\ 1 \\ 0 \end{pmatrix} + \mu \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}.$$

En definitiva, $\{(-7, 2, 1)^t, (0, 0, 0, 1)^t\}$ es base de $\text{Ker}(f)$.

Ejemplo. Hallar bases de $\text{Ker}(f)$ e $\text{Im}(f)$ para la aplicación lineal

$$f : \mathbb{R}^4 \longrightarrow \mathbb{R}^4 \quad \text{dada por} \quad f(\vec{x}) = A\vec{x} \quad \text{con} \quad A = \begin{pmatrix} 4 & 2 & -1 & 3 \\ 2 & 1 & 0 & 1 \\ 1 & 0 & 3 & -2 \\ 9 & 2 & -1 & 8 \end{pmatrix}$$

y decidir si es inyectiva o sobreyectiva.

Intercambiando la primera y la tercera fila la eliminación de Gauss en la primera columna se simplifica:

$$A \xrightarrow{f_1 \leftrightarrow f_3} \begin{pmatrix} 1 & 0 & 3 & -2 \\ 2 & 1 & 0 & 1 \\ 4 & 2 & -1 & 3 \\ 9 & 2 & -1 & 8 \end{pmatrix} \xrightarrow{\substack{f_2 \mapsto f_2 - 2f_1 \\ f_3 \mapsto f_3 - 4f_1 \\ f_4 \mapsto f_4 - 9f_1}} \begin{pmatrix} 1 & 0 & 3 & -2 \\ 0 & 1 & -6 & 5 \\ 0 & 2 & -13 & 11 \\ 0 & 2 & -28 & 26 \end{pmatrix}$$

y los otros pasos son también sencillos

$$\xrightarrow{\substack{f_3 \mapsto f_3 - 2f_2 \\ f_4 \mapsto f_4 - 2f_2}} \begin{pmatrix} 1 & 0 & 3 & -2 \\ 0 & 1 & -6 & 5 \\ 0 & 0 & -1 & 1 \\ 0 & 0 & -16 & 16 \end{pmatrix} \xrightarrow{f_4 \mapsto f_4 - 16f_3} \begin{pmatrix} 1 & 0 & 3 & -2 \\ 0 & 1 & -6 & 5 \\ 0 & 0 & -1 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Por consiguiente $\text{rg}(A) = 3$ y las fórmulas para las dimensiones dan $\dim \text{Ker}(f) = 1$ y $\dim \text{Im}(f) = 3$. De ello se deduce que f no es inyectiva ni sobreyectiva, esto último requeriría $\dim \text{Im}(f) = \dim \mathbb{R}^4 = 4$. Eligiendo $x_4 = 1$ en el sistema $A\vec{x} = \vec{0}$ se obtiene $x_2 = x_3 = 1$, $x_1 = -1$, por tanto $\{(-1, 1, 1, 1)^t\}$ es una base de $\text{Ker}(f)$. Las columnas pivote de A son las tres primeras y conforman una base de $\text{Im}(f)$.