Álgebra lineal básica

Ingeniería Biomédica Curso: Matemáticas I

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

1. Espacios vectoriales

El concepto que da nombre a este apartado es, como la mayoría de las estructuras algebraicas, bastante abstracto. Ante su definición completa, la reacción típica de los que no tienen una fe matemática inquebrantable es de extrañeza al contemplar una lista de propiedades abstractas sin aparente motivación ni utilidad.

Por ello, aquí no vamos a hacer hincapié en la definición rigurosa. La idea básica, la definición no rigurosa, es que un *espacio vectorial* V es cualquier conjunto de cosas, que denominaremos *vectores*, tales que se pueden multiplicar por números y sumar entre sí. En breve, tal que dados $\vec{u}, \vec{v} \in V$ y números λ y μ, se tenga que $\lambda\vec{u} + \mu\vec{v}$ pertenece a V, que también sea un vector.

Cuando los números escogidos forman un conjunto K, se dice que el espacio vectorial es sobre K. En este curso siempre será $K = \mathbb{R}$ o $K = \mathbb{C}$, todos los espacios serán sobre $\mathbb{R}$ o sobre $\mathbb{C}$.

La definición anterior no es rigurosa porque es demasiado imprecisa: no especifica las propiedades básicas que deben tener una suma de vectores y una multiplicación por números que en contextos más o menos exóticos no tienen nada que ver con las que nos son familiares. Entre otras cosas, hay que delimitar las características de K para que consideremos que son números válidos. Curiosamente, hay espacios de interés en ingeniería sobre conjuntos K de números raros que no has visto en ningún curso anterior.

A pesar de no tener una definición a prueba de matemáticos, veamos tres familias de ejemplos fundamentales con nuestra definición no rigurosa.

El espacio de vectores de toda la vida $\mathbb{R}^n$ es un espacio vectorial sobre $\mathbb{R}$ con la suma coordenada a coordenada y la multiplicación simultánea de todas las coordenadas por un número real. De la misma forma, $\mathbb{C}^n$ es un espacio vectorial sobre $\mathbb{C}$. De hecho, si decretamos que solo está permitida la multiplicación de números reales, todavía $\mathbb{C}^n$ sería un espacio vectorial, esta vez sobre $\mathbb{R}$, que a veces se llama $\mathbb{C}^n_{\mathbb{R}}$ para distinguirlo del habitual, ya que $\vec{u}, \vec{v} \in \mathbb{C}^n$ y $\lambda, \mu \in \mathbb{R}$ implican $\lambda\vec{v} + \mu\vec{v} \in \mathbb{C}^n$.

También son espacios vectoriales $\mathcal{M}_{m\times n}(\mathbb{R})$ y $\mathcal{M}_{m\times n}(\mathbb{C})$ pues a efectos de hacer sumas o multiplicaciones por números siempre podemos recolocar los elementos en una columna para tener vectores de $\mathbb{R}^{mn}$ o $\mathbb{C}^{mn}$.

Un ejemplo general que contiene a muchos de los que se te puedan ocurrir son las funciones $f : X \longrightarrow \mathbb{R}$ o $f : X \longrightarrow \mathbb{C}$ con X un conjunto fijado. Si X es el conjunto de seres humanos, un vector de este espacio sería asignarles a cada uno su número de pasaporte o su altura o a unos π y a otros e. El caso es que aunque no sepamos sumar personas para que den otra persona o multiplicarlas por números, sí podemos hacerlo con los números que les asignamos y eso es suficiente. En ingeniería, el caso $X = \mathbb{R}$ se identifica con las señales continuas y $X = \mathbb{Z}$ con el de las señales discretas, el espacio vectorial formado por tiras infinitas a izquierda y derecha $(x_n)_{n=-\infty}^{\infty}$, que, siendo pedante, se podrían llamar sucesiones bilaterales.

Aunque nuestra definición de espacio vectorial no sea rigurosa, existe un resultado que afirma que *si sabemos que V es un espacio vectorial sobre K y $W \subset V$ es un subconjunto (no vacío) que cumple*

$$\lambda\vec{v} + \mu\vec{w} \in W \quad \textit{para todo} \quad \lambda, \mu \in K \text{ y } \vec{v}, \vec{w} \in W,$$

entonces W es un espacio vectorial sobre K.

Por razones obvias, en la situación anterior se dice que W es un *subespacio vectorial* de V. En pocas palabras, la definición no rigurosa se vuelve rigurosa si sabemos de antemano que hay un espacio vectorial por encima. Así pues, los ejemplos de espacios vectorial que en parte nos hemos creído, nos sirven para construir otros muchos ejemplos.

Uno de los más comunes este curso, en cierta forma el más fundamental, será el conjunto de soluciones de un sistema homogéneo:

$$V = \{\vec{x} \in K^n \,:\, A\vec{x} = \vec{0}\} \qquad \text{con } A \in \mathcal{M}_{m\times n}(K) \text{ fijada y } K = \mathbb{R} \text{ o } \mathbb{C}.$$

Claramente $V \subset K^n$ y, según lo dicho, para comprobar que es espacio vectorial basta ver que es subespacio vectorial de K^n. Si $\vec{u}, \vec{v} \in V$ se cumple $A\vec{u} = A\vec{v} = \vec{0}$, por tanto $A(\lambda\vec{u} + \mu\vec{v}) = \lambda A\vec{u} + \mu A\vec{v} = \vec{0}$, así pues $\lambda\vec{u} + \mu\vec{v} \in V$.

Dados $A \in \mathcal{M}_{m\times n}(\mathbb{R})$ y $\vec{b} \neq \vec{0}$, el conjunto $V = \{\vec{x} \in \mathbb{R}^n \,:\, A\vec{x} = \vec{b}\}$ no es un espacio vectorial. Si $\vec{u} \in V$ entonces $\vec{u} + \vec{u} = 2\vec{u} \notin V$ ya que $A(2\vec{u}) = 2A\vec{u} = 2\vec{b} \neq \vec{b}$. Esto significa que multiplicar por dos no está bien definido en V. No tenemos una operación $\cdot : \mathbb{R} \times V \longrightarrow V$. Una demostración más rápida, pero más rara, es notar simplemente que $\vec{0} \notin V$.

Los polinomios con coeficientes reales $\mathbb{R}[x]$ forman un espacio vectorial sobre $\mathbb{R}$. Son un subconjunto de todas las funciones $\mathbb{R} \longrightarrow \mathbb{R}$, que ya sabíamos que son espacio vectorial sobre $\mathbb{R}$, así que, de nuevo, basta comprobar las propiedades de subespacio. Si $P, Q \in \mathbb{R}[x]$ y $\lambda, \mu \in \mathbb{R}$ entonces es obvio que $\lambda P + \mu Q \in \mathbb{R}[x]$. Los polinomios con coeficientes complejos $\mathbb{C}[x]$ forman un espacio vectorial sobre $\mathbb{C}$.

De la misma forma los polinomios de $\mathbb{R}[x]$ con grado menor o igual que 3, denotados a veces con $\mathbb{R}_3[x]$, también forman un espacio vectorial. Sin embargo los de grado exactamente 3 no, ya que por ejemplo $P = 2x^3 + x$ y $Q = 1 - 3x^3$ tienen grado 3, pero $3P + 2Q$ no lo tiene. En general, los polinomios reales de grado menor o igual que n, denotados $\mathbb{R}_n[x]$, forman un espacio vectorial sobre $\mathbb{R}$. Lo mismo se aplica, sobre $\mathbb{C}$, para $\mathbb{C}_n[x]$ definido de la misma manera con coeficientes complejos.

Las matrices cuadradas simétricas reales $V = \{A \in \mathcal{M}_n(\mathbb{R}) : A = A^t\}$ forman un espacio vectorial sobre $\mathbb{R}$ porque $V \subset \mathcal{M}_n(\mathbb{R})$ y si $A, B \in V$ entonces $(\lambda A + \mu B)^t = \lambda A^t + \mu B^t = \lambda A + \mu B$.

La definición rigurosa [Opcional]

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 $+ : V \times V \longrightarrow V$, $\cdot : K \times V \longrightarrow V$. 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 propiedades:

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

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 $\mathbb{R}$ y $\mathbb{C}$, que usaremos este curso, y de $\mathbb{Q}$, es difícil que se te ocurran más posibilidades para K, aunque existen otras importantes. 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 inverso[1]. 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.

Ahora, si queremos probar de verdad que fijado X el conjunto de funciones $V = \{f : X \longrightarrow \mathbb{R}\}$ con la suma y producto habituales es un espacio vectorial sobre $\mathbb{R}$, tenemos que comprobar en principio un montón de cosas. El truco para evitarlo es apelar a lo que ya conocemos o supuestamente conocemos. Primero, todos sabemos sumar, restar, multiplicar y dividir números reales y operar paréntesis, por tanto $\mathbb{R}$ es un cuerpo. ¿Es $(V, +)$ un grupo abeliano? Sí y resulta que probarlo no entraña nada nuevo porque sumar funciones es sumar sus valores por tanto se sigue de que $(\mathbb{R}, +)$ es grupo abeliano por ser K cuerpo. De la misma forma, las cuatro propiedades

[1] Si 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 $\vec{u} + (-\vec{u}) = \vec{0}$.

al final de la definición de espacio vectorial se seguirían de que $\mathbb{R}$ es un cuerpo si $\vec{u}, \vec{v} \in \mathbb{R}$ y eso es todo lo que necesitamos porque, como acabamos de decir, las funciones se operan operando sus valores por separado.

Si $X = \{1, 2, \dots, n\}$ podemos identificar $f : X \longrightarrow \mathbb{R}$ con un vector de toda la vida de $\mathbb{R}^n$, donde la i-ésima coordenada dada por el valor de $f(i)$. De lo anterior se deduce que $\mathbb{R}^n$ es un espacio vectorial sobre $\mathbb{R}$. De la misma forma se tiene que $\mathcal{M}_{m\times n}(\mathbb{R})$ es un espacio vectorial sobre $\mathbb{R}$ tomando como X todas las posibles posiciones de los elementos, $X = \{(1,1), (1,2), \dots, (m,n)\}$ e identificando $f(i,j)$ con a_{ij}. Por supuesto lo mismo se aplica a $\mathbb{C}^n$ y a $\mathcal{M}_{m\times n}(\mathbb{C})$.

2. Independencia lineal

Si alguien nos preguntara por qué $\mathbb{R}^2$ tiene dimensión 2, identificándolo con los puntos del plano, 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. Pospondremos el concepto de dimensión al próximo apartado y nos centraremos ahora en este problema. La siguiente definición captura en términos matemáticos la idea de que un conjunto de vectores no sea redundante.

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

$$\lambda_1\vec{v}_1 + \lambda_2\vec{v}_2 + \cdots + \lambda_n\vec{v}_n = \vec{0} \tag{1}$$

con $\lambda_1, \dots, \lambda_n \in K$ es $\lambda_1 = \cdots = \lambda_n = 0$ (aunque es irrelevante en este curso, para una conjunto infinito de vectores, la independencia lineal se define pidiendo que ocurra para cualquiera de sus subconjuntos finitos).

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.

Por cierto, una expresión del tipo $\lambda_1\vec{v}_1 + \lambda_2\vec{v}_2 + \cdots + \lambda_n\vec{v}_n$ 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,

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

Es muy fácil ver que para cualquier subconjunto C de un espacio vectorial se cumple que $\mathcal{L}(C)$ es subespacio vectorial y, por tanto, espacio vectorial. Se dice que $\mathcal{L}(C)$ es el *espacio generado* por C, Seguro que no te asombra que el estudio de la dependencia o independencia lineal se reduzca habitualmente a discutir un sistema homogéneo.

Por ejemplo, estudiemos si los vectores de $\mathbb{R}^3$

$$\vec{v}_1 = (2, 1, -1)^t, \qquad \vec{v}_2 = (3, 0, 1)^t, \qquad \text{y} \qquad \vec{v}_3 = (-1, 4, -7)^t$$

son linealmente independientes. La ecuación (1) se puede escribir como

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

En general en $\mathbb{R}^n$ o $\mathbb{C}^n$ el sistema será $A\vec{\lambda} = \vec{0}$ con A la matriz formada por los vectores en columna. Es fácil ver que $\operatorname{rg}(A) = 2$, concretamente aplicando eliminación de Gauss intercambiando las dos primeras filas:

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

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

Si cambiamos $\vec{v}_3$ por $\vec{w}_3 = (-1, 4, -6)^t$, entonces $\vec{v}_1$, $\vec{v}_2$ 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.

La nueva definición permite reformular el concepto de rango de una matriz. Recordemos que lo habíamos introducido como el número de escalones al aplicar reducción de Gauss. Por otro lado, 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 nos dice el número máximo de columnas de una matriz que son linealmente independientes.

La dependencia lineal de dos vectores se reduce a comprobar que uno es múltiplo del otro. En $\mathbb{R}^n$ esto normalmente se decide a ojo mientras que en $\mathbb{C}^n$ quizá nuestro ojo tenga que trabajar más y pedir ayuda a nuestra mano. Por ejemplo, los vectores $\vec{v}_1 = (1+i, 2i, i-1)^t$ y $\vec{v}_2 = (1+2i, 3i-1, i-2)^t$ de $\mathbb{C}^3$ son linealmente dependientes si y solo si $\vec{v}_2 = \frac{1+2i}{1+i}\vec{v}_1$, para que cuadren las primeras coordenadas, y un cálculo muestra que esta relación se cumple. Si esto suena truculento y uno quiere hacer todos los problemas de la misma forma, podríamos seguir el mismo esquema que antes aplicando reducción de Gauss para ver que el sistema $\lambda_1\vec{v}_1 + \lambda_2\vec{v}_2 = \vec{0}$ tiene infinitas soluciones (λ_1, λ_2):

$$\begin{pmatrix} 1+i & 1+2i \\ 2i & 3i-1 \\ i-1 & i-2 \end{pmatrix} \underset{f_1 \mapsto \frac{f_1}{1+i}}{\longrightarrow} \begin{pmatrix} 1 & \frac{3+i}{2} \\ 2i & 3i-1 \\ i-1 & i-2 \end{pmatrix} \underset{\substack{f_2 \mapsto f_2 - 2if_1 \\ f_3 \mapsto f_3 - (i-1)f_1}}{\longrightarrow} \begin{pmatrix} 1 & \frac{3+i}{2} \\ 0 & 0 \\ 0 & 0 \end{pmatrix}.$$

El rango es, por tanto, 1 y al ser menor que el número de incógnitas, hay infinitas soluciones (el sistema es compatible indeterminado).

De la misma forma, los polinomios $P_1 = 2 + x + x^2$ y $P_2 = 3 - x^2$ son linealmente independientes en $\mathbb{R}[x]$ porque no son proporcionales, pero $\{P_1, P_2, P_3\}$ con $P_3 = 5 + x$ ya no lo son porque $P_3 = P_1 + P_2$, o equivalentemente $P_1 + P_2 - P_3 = 0$.

La abstracción no nos debe confundir. Decidir la independencia lineal de unos cuantos polinomios o matrices lleva a considerar sistemas homogéneos. Es algo tan simple como en $\mathbb{R}^n$. Por ejemplo, consideremos los polinomios $P_1 = 1 + x + x^2$, $P_2 = 3 + 5x + x^2$, $P_3 = 7 - x^2 + x^3$ del espacio vectorial $\mathbb{R}[x]$. La ecuación relevante es

$$\lambda_1\big(1 + x + x^2\big) + \lambda_2\big(3 + 5x + x^2\big) + \lambda_3\big(7 - x^2 + x^3\big) = 0 \qquad \text{con} \quad \lambda_1, \lambda_2, \lambda_3 \in \mathbb{R}.$$

Agrupando los términos del mismo grado,

$$(\lambda_1 + 3\lambda_2 + 7\lambda_3) + (\lambda_1 + 5\lambda_2)x + (\lambda_1 + \lambda_2 - \lambda_3)x^2 + \lambda_3 x^3 = 0.$$

El polinomio nulo tiene todos sus coeficientes nulos, por tanto llegamos al sistema homogéneo

$$A\vec{\lambda} = \vec{0} \qquad \text{con} \quad A = \begin{pmatrix} 1 & 3 & 7 \\ 1 & 5 & 0 \\ 1 & 1 & -1 \\ 0 & 0 & 1 \end{pmatrix}, \quad \vec{\lambda} = \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \lambda_3 \end{pmatrix}.$$

Está claro que $\mathrm{rg}(A) = 3$ simplemente considerando las tres últimas filas (el rango no puede ser mayor que tres) y aplicando la transformación elemental $f_3 \mapsto f_3 - f_2$. Entonces el sistema es compatible determinado, solo tiene la solución trivial, y los polinomios P_1, P_2, P_3 resultan ser linealmente independientes.

Por tener un ejemplo más fuera de $\mathbb{R}^n$ y $\mathbb{C}^n$, estudiemos la dependencia o independencia lineal de

$$\begin{pmatrix} 1 & 2 \\ 1 & -1 \end{pmatrix}, \qquad \begin{pmatrix} -1 & 0 \\ 1 & 1 \end{pmatrix} \qquad \text{y} \qquad \begin{pmatrix} -1 & 4 \\ 5 & 1 \end{pmatrix}$$

en el espacio vectorial $\mathcal{M}_{2\times 2}(\mathbb{R})$. Según la definición, hay que considerar la ecuación

$$\lambda_1 \begin{pmatrix} 1 & 2 \\ 1 & -1 \end{pmatrix} + \lambda_2 \begin{pmatrix} -1 & 0 \\ 1 & 1 \end{pmatrix} + \lambda_3 \begin{pmatrix} -1 & 4 \\ 5 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

que da lugar al sistema lineal

$$A\vec{\lambda} = \vec{0} \qquad \text{con} \quad A = \begin{pmatrix} 1 & -1 & -1 \\ 2 & 0 & 4 \\ 1 & 1 & 5 \\ -1 & 1 & 1 \end{pmatrix}.$$

No es extraño que las columnas de A estén formadas por los elementos de las matrices de partida escritos en vertical. Ahora aplicamos reducción de Gauss para calcular el rango:

$$A \underset{\substack{f_2\mapsto f_2-2f_1\\ f_3\mapsto f_3-f_1\\ f_4\mapsto f_4+f_1}}{\longrightarrow} \begin{pmatrix} 1 & -1 & -1\\ 0 & 2 & 6\\ 0 & 2 & 6\\ 0 & 0 & 0 \end{pmatrix} \underset{f_3\mapsto f_3-f_2}{\longrightarrow} \begin{pmatrix} 1 & -1 & -1\\ 0 & 2 & 6\\ 0 & 0 & 0\\ 0 & 0 & 0 \end{pmatrix}.$$

El rango es entonces 2, menor que el número de incógnitas. Por tanto, hay infinitas soluciones y las matrices son linealmente dependientes en $\mathcal{M}_{2\times 2}(\mathbb{R})$.

3. Bases y dimensión

Después de que hemos entendido la independencia lineal, ya estamos 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, 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, un espacio vectorial cumple $\dim V = n$ cuando hay una base con n elementos.

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.

Antes de seguir, tres observaciones en miniatura para los más exigentes con el rigor: 1) La definición habitual de base no excluye que pueda contener infinitos vectores, esto es, que $\dim V = \infty$. Sin embargo, en este curso nos limitaremos, sin recordarlo cada vez, al caso de dimensión finita. 2) 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. 3) Se puede demostrar que todas las bases de un espacio vectorial tienen el mismo número de elementos, por tanto la dimensión está unívocamente definida, no depende de cómo la calculemos.

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 $\mathbb{R}^n$ y $\mathbb{C}^n$ 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 $\dim\mathbb{R}^n = \dim\mathbb{C}^n = n$.

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 $\vec{u} = \lambda_1\vec{v}_1 + \cdots + \lambda_n\vec{v}_n$ porque si hubiera dos combinaciones lineales que dieran $\vec{u}$, al restarlas llegaríamos a algo del tipo (1) que contradiría la independencia lineal. Estos λ_j se dice que son las *coordenadas* de $\vec{u}$ en la base $\mathcal{B}$ e incluso a veces se escribe, abusando de la notación, $\vec{u} = (\lambda_1, \dots, \lambda_n)^t$ especificando la base en la que trabajamos. Aquí evitaremos este abuso.

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 $\vec{u} = (1,2)^t \in \mathbb{R}^2$ tiene obviamente coordenadas 1 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 1 y 1, pues $\vec{u} = \vec{v}_1 + \vec{v}_2$.

La pregunta fundamental a este nivel es cómo hallar una base de un espacio vectorial. La respuesta depende de cómo esté presentado. O bien pasando a coordenadas o bien directamente, muchos de los ejemplos del curso vienen dados por ecuaciones homogéneas. Es decir, son del tipo:

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

Pues bien, si leemos 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, su dimensión es** $n - \operatorname{rg}(A)$. Recordemos que estos $\vec{v}_j$ se hallaban mediante reducción de Gauss. Entonces es inexcusable aprender ese algoritmo en este curso.

Un ejemplo extremadamente simple es $V = \{\vec{x} \in \mathbb{R}^3 \,:\, x_1 + 2x_2 + 3x_3 = 0\}$. La matriz del sistema es la matriz fila $(1,2,3)$ y por tanto $\operatorname{rg}(A) = 1$, lo que implica $\dim V = 3 - 1 = 2$. A nadie sorprenderá que un plano tenga dimensión dos. Asignando $x_2 = \lambda_1$, $x_3 = \lambda_2$, que corresponden a las columnas no pivote, se tiene que la solución general es

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

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

Continuando el ejemplo, hallemos las coordenadas del vector $\vec{u} = (-7,2,1)^t$ en la base que acabamos de calcular. 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 $\mathcal{B}$. Si consideramos $\mathcal{B}$ como nuestro "sistema de medida", podemos pensar en $\vec{u}$ como en el vector $(2,1)^t$ de $\mathbb{R}^2$.

Como ejemplo con números complejos, calculemos la dimensión y una base del subespacio de $\mathbb{C}^3$

$$W = \{\vec{x} \in \mathbb{C}^3 \,:\, A\vec{x} = \vec{0}\} \qquad \text{con} \quad A = \begin{pmatrix} 2 & 1 & 5+i \\ 3 & 2i & 4+2i \\ 4 & -1+4i & 3+3i \end{pmatrix}.$$

Con el fin de evitar fracciones cuando anulamos a_{21} por eliminación de Gauss, combinamos la primera y la segunda transformaciones elementales en $f_2 \mapsto 2f_2 - 3f_1$.

$$\xrightarrow[\substack{f_2\mapsto 2f_2-3f_1\\ f_3\mapsto f_3-2f_1}]{} \begin{pmatrix} 1 & 1 & 5+i \\ 0 & -3+4i & -7+i \\ 0 & -3+4i & -7+i \end{pmatrix} \xrightarrow[f_3\mapsto f_3-f_2]{} \begin{pmatrix} 2 & 1 & 5+i \\ 0 & -3+4i & -7+i \\ 0 & 0 & 0 \end{pmatrix}.$$

De aquí $\text{rg}(A) = 2$ y, por tanto, $\dim W = 3 - 2 = 1$. Una base estará formada por cualquier vector no nulo de W. Tomando $x_3 = 1$ se deduce de la forma escalonada $x_2 = -1-i$ y $x_1 = -2$ que produce $\mathcal{B} = \{(-2, -1-i, 1)^t\}$. Esto equivale al procedimiento habitual de tomar $x_3 = \lambda$ y después sacar factor común λ. Aquí están las cosas preparadas para que $-7+i$ sea "divisible" por $-3+4i$. Si no fuera así y quisiéramos evitar las fracciones y minimizar las cuentas con números complejos, una elección natural sería $x_3 = -3+4i$.

Hallar bases en espacios que vienen dados por ecuaciones y no son subespacios de $\mathbb{R}^n$ o $\mathbb{C}^n$, dan lugar a problemas similares en este curso una vez que escribimos las ecuaciones como un sistema lineal. Un ejemplo con polinomios es

$$V = \big\{P \in \mathbb{R}_2[x] \,:\, 3P'' + (x+1)P' - 2P = 0\big\}.$$

Todos los polinomios $P \in \mathbb{R}_2[x]$ son de la forma $P = a + bx + cx^2$ y se tiene $P' = b + 2cx$ y $P'' = 2c$. Así pues, la ecuación que define V se reescribe como

$$6c + (x+1)(b+2cx) - 2a - 2bx - 2cx^2 = 0$$

que operando un poco resulta

$$-2a + b - 6c + (-b+2c)x = 0.$$

Si un polinomio es idénticamente cero, sus coeficientes son nulos y, por tanto, llegamos al sistema lineal homogéneo en a, b y c

$$\begin{aligned} -2a + b + 6c &= 0 \\ -b + 2c &= 0 \end{aligned} \qquad \text{que tiene por matriz} \quad A = \begin{pmatrix} -2 & 1 & 6 \\ 0 & -1 & 2 \end{pmatrix}.$$

En definitiva, el espacio viene determinado por una ecuación $A\vec{x} = \vec{0}$ con $\vec{x} = (a,b,c)^t$ y podemos proceder como en ejemplos anteriores. En realidad, como A ya está en forma escalonada, es muy fácil obtener que la solución es $(a,b,c)^t = \lambda(4,2,1)^t$. Este sería el vector que da la base si estuviéramos en $\mathbb{R}^3$, pero como estamos en un subespacio de $\mathbb{R}_2[x]$ hay que escribir el resultado como un polinomio. En definitiva, $\mathcal{B} = \{4 + 2x + x^2\}$ es base y $\dim V = 1$.

Esencialmente, los ejercicios que se suelen poner acerca de bases consisten o bien en hallar una base (como en los ejemplos anteriores) o en comprobar que algo es base. Para esto último, hay que usar la definición. La parte de comprobar que es sistema de generadores puede ser engorrosa o redundante 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 ya sabes que el número de elementos de la posible base coincide con la dimensión solo hace falta que compruebes que sus vectores son linealmente independientes para confirmar que es base.

Como ejemplo, estudiemos si

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

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 $\dim V = n - \operatorname{rg}(A) = 4 - 2 = 2$ 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 enunciado de este ejemplo 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.

Para terminar, explotemos esta idea para comprobar que en $\mathcal{M}_2(\mathbb{R})$ el conjunto

$$\mathcal{B} = \{I, M\} \quad \text{con} \quad M = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \quad \text{es base de} \quad V = \left\{N \ : \ M^2 N = N M^2\right\}.$$

Si queremos seguir un procedimiento como el anterior, tenemos que obtener un sistema lineal que caracterice las matrices $N \in V$, para ello calculamos

$$M^2 = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix} \qquad \text{y escribimos} \qquad N = \begin{pmatrix} x_1 & x_2 \\ x_3 & x_4 \end{pmatrix}.$$

Al imponer $M^2 N = N M^2$, las cuatro igualdades que se obtienen son $2x_3 = 0$, $-2x_1 + 2x_4 = 0$, $0 = 0$ y $x_3 = 0$. La tercera ecuación es trivial y la primera equivale a la cuarta. Así pues, nos quedamos con la segunda (simplificada por 2) y la cuarta, para deducir

$$V = \left\{ \begin{pmatrix} x_1 & x_2 \\ x_3 & x_4 \end{pmatrix} \ : \ \begin{array}{r} -x_1 + x_4 = 0 \\ x_3 = 0 \end{array} \right\} = \left\{ \begin{pmatrix} x_1 & x_2 \\ x_3 & x_4 \end{pmatrix} \ : \ A\vec{x} = \vec{0} \right\} \quad \text{con} \quad A = \begin{pmatrix} -1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}$$

y, por supuesto, $\vec{x} = (x_1, x_2, x_3, x_4)^t$. La matriz A ya está en forma escalonada y tiene dos escalones, por tanto, $\dim V = 4 - 2 = 2$ que coincide con el número de vectores (matrices) de $\mathcal{B}$. Está claro que $\mathcal{B} \subset V$ porque $M^2 I = IM^2$ y $M^2 M = MM^2$, entonces solo queda comprobar que los elementos de $\mathcal{B}$ son linealmente independientes. Escribiendo $\lambda_1 I + \lambda_2 M = O$ los elementos de la primera fila dan $\lambda_1 + \lambda_2 = \lambda_2 = 0$. Así pues, $\lambda_1 = \lambda_2 = 0$ es la única solución y realmente son linealmente independientes.

4. Aplicaciones lineales

En la mayor parte de los modelos de los fenómenos físicos si una cantidad depende de varios parámetros, pequeñas variaciones de estos se suman con distintas constantes de proporcionalidad para dar lugar a una aproximación de la variación de dicha cantidad. Por ejemplo, la variación de la intensidad gravitatoria debida a la Tierra tiene una fórmula más o menos complicada, pero cerca de la superficie, para variaciones pequeñas de la altura Δh es aproximadamente $9{,}8\Delta h$. La siguiente definición introduce un concepto matemático motivado por esta situación.

Dados espacios vectoriales V y W sobre un mismo cuerpo K, se dice que una función $f : V \longrightarrow W$ es una *aplicación lineal* si preserva las combinaciones lineales, es decir, si para $\vec{u}, \vec{v} \in V$, $\lambda, \mu \in K$ cualesquiera $f(\lambda\vec{u} + \mu\vec{v}) = \lambda f(\vec{u}) + \mu f(\vec{v})$. Se emplea el nombre *endomorfismo* para indicar $V = W$. Es decir, un endomorfismo es una aplicación lineal de un espacio vectorial en sí mismo.

Si recordamos la definición intuitiva de espacio vectorial, nos daremos cuenta de que, hablando sin rigor, las aplicaciones lineales son las funciones que pasan espacios vectoriales a espacios vectoriales.

Las buenas noticias es que, independientemente de que la definición sea más o menos enrevesada, para $V = K^n$ y $W = K^m$, el caso más común, todo se reduce a multiplicar por una matriz. Concretamente, cualquier aplicación lineal $f : K^n \longrightarrow K^m$ es de la forma $f(\vec{x}) = A\vec{x}$ con $A \in \mathcal{M}_{m\times n}(K)$. Además, las columnas de A son los vectores $f(\vec{e}_j)$ con $\{\vec{e}_1, \dots, \vec{e}_n\}$ la base canónica de K^n.

Ahora estamos en condiciones de entender plenamente por qué las matrices se multiplican de manera tan rara. La única manera de usar una aplicación lineal f sobre el resultado de otra aplicación g, lo que se llama *composición* de funciones $f \circ g$, es que g acabe donde comienza f. Digamos $K^{n_1} \xrightarrow{g} K^{n_2} \xrightarrow{f} K^{n_3}$. Si $A \in \mathcal{M}_{n_3\times n_2}(K)$ y $B \in \mathcal{M}_{n_2\times n_1}(K)$ son las matrices de f y g, respectivamente, $(f \circ g)(\vec{x}) = f\big(g(\vec{x})\big) = A(B\vec{x})$. La función $f \circ g$ es también una aplicación lineal, por tanto $(f \circ g)(\vec{x}) = C\vec{x}$. ¿Qué relación hay entre A, B y C? La i-ésima coordenada de $(f \circ g)(\vec{x})$ es $\sum_k a_{ik} m_k$ con m_k la k-ésima coordenada de $B\vec{x}$, que es $\sum_j b_{kj} x_j$. Por consiguiente $c_{ij} = \sum_k a_{ik} b_{kj}$. Esta es la manera de operar matrices para que indique la composición. Decimos que es una multiplicación, aunque no comparta con la habitual todas sus propiedades.

Según lo anterior, todos los ejemplos de aplicaciones lineales $K^n \longrightarrow K^m$ son sencillos. Así, la siguiente función $f : \mathbb{R}^3 \longrightarrow \mathbb{R}^2$

$$f(\vec{x}) = \begin{pmatrix} x_1 - 2x_2 + x_3 \\ 2024x_2 \end{pmatrix} \quad \text{es aplicación lineal porque} \quad f(\vec{x}) = \begin{pmatrix} 1 & -2 & 1 \\ 0 & 2024 & 0 \end{pmatrix} \vec{x}.$$

Por otro lado, cambiando $2024x_2$ por $2024x_2^2$ o por $\operatorname{sen}(2024x_2)$ o por x_1x_2 no sería aplicación lineal porque ya no vendría dada por la multiplicación por una matriz.

La "mala noticia" es que las aplicaciones lineales no son exclusivas de K^n. Por ejemplo, la derivada $D : \mathbb{R}_2[x] \longrightarrow \mathbb{R}_1[x]$ (o $D : \mathbb{R}_2[x] \longrightarrow \mathbb{R}_2[x]$ si queremos que sea endomorfismo), la integral en un intervalo $I_{[a,b]} : \mathbb{R}_2[x] \longrightarrow \mathbb{R}$ o el llamado *conmutador* $\mathcal{C}_T : \mathcal{M}_2(\mathbb{C}) \longrightarrow \mathcal{M}_2(\mathbb{C})$ dado por $\mathcal{C}_T(A) = AT - TA$ con T una matriz fijada, definen aplicaciones lineales (si no te parece claro, compruébalo con la definición).

En breve veremos que hay una manera de interpretar todas las aplicaciones lineales con matrices transformando los ejemplos difíciles en fáciles. Antes de ello, vamos a introducir dos subespacios destacados asociados a cada aplicación lineal.

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

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

Este "Ker" es la abreviatura de *kernel*, núcleo en inglés. A veces en textos en español se escribe $\operatorname{Nuc}(f)$ en vez de $\operatorname{Ker}(f)$.

No es difícil probar que

$$\begin{aligned} \operatorname{Ker}(f) = \{\vec{0}\} \quad &\Leftrightarrow \quad f \text{ es } \textit{inyectiva}, \\ \operatorname{Im}(f) = W \quad &\Leftrightarrow \quad f \text{ es } \textit{sobreyectiva}. \end{aligned}$$

En realidad, lo segundo es la propia definición de función sobreyectiva, aquella que alcanza a todos los elementos del conjunto final. Respecto a las funciones inyectivas, son las que asignan imágenes diferentes a elementos diferentes. Si $\operatorname{Ker}(f) \neq \{\vec{0}\}$, digamos $\vec{0} \neq \vec{v} \in \operatorname{Ker}(f)$, entonces $f(\vec{v}) = f(\vec{0}) = \vec{0}$ y no sería inyectiva.

Las funciones que son simultáneamente inyectivas y sobreyectivas se dice que son *biyectivas*. Estas admiten una *función inversa*, una función que lleva para atrás las imágenes. En nuestro contexto, todo lo que necesitas observar es que si $f : K^n \longrightarrow K^n$ es un endomorfismo dado por $f(\vec{x}) = A\vec{x}$ entonces existe una función inversa $f^{-1} : K^n \longrightarrow K^n$ si y solo si A es una matriz invertible, en cuyo caso $f^{-1}(\vec{x}) = A^{-1}\vec{x}$.

A menudo la inyectividad y la sobreyectividad se comprueba examinando dimensiones. Concretamente, una aplicación lineal $f : V \longrightarrow W$ es inyectiva si y solo si $\dim \operatorname{Ker}(f) = 0$ y es sobreyectiva si y solo si $\dim \operatorname{Im}(f) = \dim W$. Respecto a lo primero, nótese que acogiéndonos a la literalidad de la definición, el subespacio trivial $\{\vec{0}\}$ tiene como base el conjunto vacío con cero elementos.

Seguro que te estás preguntando qué gracia tiene decidir estas cuestiones utilizando dimensiones. La respuesta es que hay unas fórmulas que permiten calcularlas en el caso fácil de aplicaciones lineales $f : K^n \longrightarrow K^m$. En ese caso se cumple

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

Además los vectores correspondientes a columnas pivote de A son una base de $\operatorname{Im}(f)$.

La primera fórmula, en realidad, ya la conocías por el cálculo de la dimensión de un subespacio de K^n definido por $A\vec{x} = \vec{0}$.

Por ejemplo, supongamos que queremos saber si la aplicación lineal

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

es inyectiva o sobreyectiva. 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 $\operatorname{rg}(A) = 2$, dos escalones. Según las fórmulas, $\dim \operatorname{Im}(f) = 2$ y $\dim \operatorname{Ker}(f) = 4 - 2 = 2$. Por tanto, f es sobreyectiva y no es inyectiva. Si tienes el superpoder de la intuición matemática, esto último está claro 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.

Dentro de este mismo ejemplo, vamos a apurar un poco más y a calcular bases del núcleo y la imagen. Según lo dicho, para lo segundo hay que considerar 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 $\operatorname{Im}(f)$ 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 eliminación de Gauss. Este proceso solo sirve para calcular dónde están los pivotes.

Una base de $\operatorname{Ker}(f)$ 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:

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

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

Practiquemos ahora con un endomorfismo proveniente de un examen pasado. Queremos hallar bases de $\text{Ker}(f)$ e $\text{Im}(f)$ para

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

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 $\{(-1, 1, 1, 1)\}$ es una base de $\text{Ker}(f)$. Las columnas pivote de A son las tres primeras y conforman una base de $\text{Im}(f)$.

El cabo suelto que nos queda es cómo tratar las aplicaciones lineales $f : V \longrightarrow W$ cuando V y W son espacios vectoriales más raros que K^n, por ejemplo de polinomios o de matrices.

El truco es que, aunque estemos en un espacio vectorial raro, una vez fijada una base, cada vector se puede identificar por sus coordenadas en dicha base que componen un vector de números de toda la vida (en K^n) y todo funciona como antes. Escrito de una manera formal: **Si $f : V \longrightarrow W$ es una aplicación lineal, fijadas bases $\mathcal{B}_V$ y $\mathcal{B}_W$ de V y W, las coordenadas de $\vec{x}$ en la base $\mathcal{B}_V$ y las de $f(\vec{x})$ en $\mathcal{B}_W$ están relacionadas mediante la multiplicación por cierta matriz cuyas columnas son las coordenadas en $\mathcal{B}_W$ de las imágenes de los elementos de $\mathcal{B}_V$**. Esta matriz se dice que es la *matriz de la aplicación lineal*, aunque en la letra pequeña de nuestra mente deberíamos añadir “una vez fijadas bases”.

El truco no es tan novedoso, ya usamos algo parecido para dar ecuaciones definidas por sistemas lineales a subespacios de polinomios o matrices.

Por ejemplo, consideremos la derivada $D : \mathbb{R}_2[x] \longrightarrow \mathbb{R}_1[x]$ y las bases $\mathcal{B} = \{1, x, x^2\}$ y $\mathcal{B}' = \{1, x\}$ de estos espacios, respectivamente. Se tiene $f'(1) = 0$, $f'(x) = 1$, $f'(x^2) = 2x$. Estos tres vectores en la base $\mathcal{B}'$ tienen coordenadas $(0, 0)$, $(1, 0)$ y $(0, 2)$, por tanto la matriz de D con estas bases es

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

y podríamos derivar multiplicando por ella (aunque esto sería ridículo).

Una variante del ejemplo anterior es la función $f : \mathbb{R}_2[x] \longrightarrow \mathbb{R}_2[x]$ dada por $f(P) = (1-x)P' + P$, con P' la derivada. Es una aplicación lineal (concretamente, endomorfismo) porque para $P, Q \in \mathbb{R}_2[x]$, $\lambda, \mu \in \mathbb{R}$ se tiene que $f(\lambda P + \mu Q)$ es

$$(1-x)(\lambda P' + \mu Q') + \lambda P + \mu Q = \lambda\big((1-x)P' + P\big) + \mu\big((1-x)Q' + Q\big) = \lambda f(P) + \mu f(Q).$$

Calculemos la matriz A de f cuando se fija la base $\{1, x, x^2\}$. Según lo que sabemos, tenemos que poner en columnas las coordenadas de $f(1)$, $f(x)$ y $f(x^2)$:

$$\begin{array}{lll} f(1) & = 1 & = 1 + 0x + 0x^2 \\ f(x) & = 1 - x + x & = 1 + 0x + 0x^2 \\ f(x^2) & = (1-x)2x + x^2 & = 0 + 2x - x^2 \end{array} \quad \Longrightarrow \quad A = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & -1 \end{pmatrix}.$$

Una vez que tenemos la matriz, las dimensiones del núcleo y la imagen se calculan de la misma forma. A riesgo de ser un pesado, aquí va el mismo enunciado adaptado al caso general:

Si A es la matriz de una aplicación lineal $f : V \longrightarrow W$ para ciertas bases de V y W, se cumple

$$\dim \operatorname{Ker}(f) = \dim V - \operatorname{rg}(A) \qquad \text{y} \qquad \dim \operatorname{Im}(f) = \operatorname{rg}(A).$$

Además, los vectores cuyas coordenadas en la base elegida para W corresponden a columnas pivote de A son una base de $\operatorname{Im}(f)$.

Así, como en el último ejemplo hemos obtenido una matriz A de rango 2 (¿lo ves claro?, requiere solo un paso de eliminación de Gauss), se deduce que $\dim \operatorname{Im}(f) = 2$ y $\dim \operatorname{Ker}(f) = 3 - 2 = 1$, lo cual responde a una pregunta de un examen pasado. En particular, f no es ni inyectiva (porque $\dim \operatorname{Ker}(f) \neq 0$) ni sobreyectiva (porque $\dim \operatorname{Im}(f) = 2 \neq \dim W = 3$).

Usando coordenadas, el cálculo de bases de $\operatorname{Ker}(f)$ e $\operatorname{Im}(f)$ siempre puede pasarse al caso fácil. Lo único con lo que hay que tener cuidado es que al presentar el resultado final, debemos hacerlo como verdaderos vectores de los espacios, no como sus coordenadas.

Continuando el último ejemplo, el mencionado paso de eliminación de Gauss da

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

Las columnas pivote son la primera y la tercera, por ello podemos conformar una base de la imagen empleando $(1,0,0)^t$ y $(0,2,-1)^t$, que son la primera y la tercera columnas de A, pero eso da las coordenadas en la base escogida del espacio, $\{1,x,x^2\}$. La solución de verdad requiere pasar estas coordenadas a polinomios, resultando la base de la imagen

$$\mathcal{B}_I = \{1+0x+0x^2, 0+2x-x^2\} = \{1, 2x-x^2\}.$$

Para el núcleo haríamos algo similar. Resolviendo el sistema $A\vec{x}=\vec{0}$ se tiene $x_3=0$, $x_2=\lambda$, $x_1=-\lambda$. Es decir, $\vec{x}=\lambda(-1,1,0)^t$. No debemos decir que $(-1,1,0)^t$ es una base del núcleo porque estamos en un espacio de polinomios, debemos usar la base del espacio para traducir estas coordenadas a un polinomio. Con ello se tiene $\mathcal{B}_K=\{-1+x+0x^2\}=\{-1+x\}$. A modo de comprobación, nota que realmente $f(-1+x)=0$, es decir, es un elemento del núcleo. Nota también que el número de elementos de $\mathcal{B}_I$ y de $\mathcal{B}_K$ es coherente con las dimensiones halladas de la imagen y el núcleo.

Para finalizar, veamos un ejemplo que involucra subespacios de matrices y números complejos, aunque no de forma sustancial. Consideramos la aplicación lineal

$$f: V=\big\{A\in\mathcal{M}_2(\mathbb{C})\ :\ a_{11}=a_{22}\big\}\longrightarrow W=\big\{A\in\mathcal{M}_2(\mathbb{C})\ :\ a_{12}=a_{21}\big\}$$

dada por la fórmula

$$f(A)=A\begin{pmatrix}1-i & 0\\ 0 & i-1\end{pmatrix}+\begin{pmatrix}1-i & 0\\ 0 & i-1\end{pmatrix}A^t.$$

Las bases más sencillas de V y W son

$$\mathcal{B}_V=\left\{\begin{pmatrix}1&0\\0&1\end{pmatrix},\begin{pmatrix}0&1\\0&0\end{pmatrix},\begin{pmatrix}0&0\\1&0\end{pmatrix}\right\}=\{M_1,M_2,M_3\}$$

y

$$\mathcal{B}_W=\left\{\begin{pmatrix}1&0\\0&0\end{pmatrix},\begin{pmatrix}0&1\\1&0\end{pmatrix},\begin{pmatrix}0&0\\0&1\end{pmatrix}\right\}=\{N_1,N_2,N_3\}.$$

Digamos que nos dan estas bases (¿sabrías comprobar que lo son?) y nos piden calcular la matriz A de f. Para ello, debemos calcular las imágenes de los elementos de $\mathcal{B}_V$ y poner sus coordenadas en la base $\mathcal{B}_W$ en columna:

$$\begin{array}{l} f(M_1)=\left(\begin{smallmatrix}2-2i & 0\\ 0 & 2i-2\end{smallmatrix}\right)=(2-2i)N_1+0N_2+(2i-2)N_3\\ f(M_2)=\left(\begin{smallmatrix}0 & -1+i\\ -1+i & 0\end{smallmatrix}\right)=0N_1+(-1+i)N_2+0N_3\\ f(M_3)=\left(\begin{smallmatrix}0 & 1-i\\ 1-i & 0\end{smallmatrix}\right)=0N_1+(1-i)N_2+0N_3\end{array}\quad\Rightarrow\ A=\begin{pmatrix}2-2i & 0 & 0\\ 0 & -1+i & 1-i\\ 2i-2 & 0 & 0\end{pmatrix}.$$

Con esto ya tenemos la matriz buscada. Si ahora nos piden la dimensión del núcleo y la imagen, aplicamos eliminación de Gauss para hallar el rango:

$$A = \begin{pmatrix} 2-2i & 0 & 0 \\ 0 & -1+i & 1-i \\ 2i-2 & 0 & 0 \end{pmatrix} \underset{f_3 \mapsto f_3+f_1}{\longrightarrow} \begin{pmatrix} 2-2i & 0 & 0 \\ 0 & -1+i & 1-i \\ 0 & 0 & 0 \end{pmatrix},$$

de donde $\operatorname{rg}(A) = 2$ y se sigue $\dim \operatorname{Im}(f) = 2$ y $\dim \operatorname{Ker}(f) = 3 - 2 = 1$.

Por último, si quisiéramos bases de estos subespacios, como las columnas pivote son las dos primeras, debemos escoger las coordenadas que dictan esas columnas de A para obtener una base de la imagen: $\mathcal{B}_I = \{(2-2i)N_1 + (2i-2)N_3, (-1+i)N_2\}$. Dicho sea de paso, para los que les guste simplificar, una base no cambia su carácter cuando se multiplican sus vectores por una constante y, por tanto, también es válida la base más sencilla $\{N_1 - N_3, N_2\}$. Para hallar una base del núcleo, debemos resolver $A\vec{x} = \vec{0}$. tomamos $x_2 = \lambda$, por ser la segunda columna no pivote, y despejamos para obtener $\vec{x} = \lambda(0,1,1)^t$. El vector $(0,1,1)^t \in \mathbb{C}^3$ estrictamente no es una base del núcleo, debemos pasar de este vector de coordenadas a una matriz de V utilizando $\mathcal{B}_V$. Con ello, obtenemos la base del núcleo $\mathcal{B}_K = \{M_2 + M_3\}$ porque el vector de coordenadas $(0,1,1)^t$ significa ponme nada de M_1, uno de M_2 y otro de M_3.

5. Ecuaciones de subespacios de K^n

A veces se presenta el problema de que un subespacio $V \subset K^n$ definido a través de una base o de unos cuantos vectores que lo generan, queremos escribirlo como

$$V = \left\{\vec{x} \in K^n \, : \, A\vec{x} = \vec{0}\right\} \qquad \text{con} \quad A \in \mathcal{M}_{m\times n}(K).$$

Es decir, deseamos calcular unas ecuaciones $A\vec{x} = \vec{0}$ que caractericen el espacio.

El procedimiento consiste en aplicar eliminación de Gauss a las combinaciones lineales igualadas a un vector genérico y estudiar qué condiciones surgen de que haya solución para los coeficientes. En vez de intentar desentrañar esta frase, es mejor ir directamente a un ejemplo.

Consideramos el subespacio $V \subset \mathbb{R}^4$ con base $\mathcal{B} = \{\vec{v}_1, \vec{v}_2\}$ con $\vec{v}_1 = (-1,2,2,-3)^t$ y $\vec{v}_2 = (5,-8,-6,23)^t$ y queremos escribirlo como $V = \{\vec{x} \in \mathbb{R}^4 \, : \, A\vec{x} = \vec{0}\}$. Para ello aplicamos eliminación de Gauss a $\lambda_1\vec{v}_1 + \lambda_2\vec{v}_2 = \vec{x}$ entendiéndolo como un sistema de ecuaciones lineales en las incógnitas λ_1 y λ_2.

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

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 $\vec{x} \in V$, es que

$$\begin{cases} -2x_1 - 2x_2 + x_3 = 0, \\ -11x_1 - 4x_2 + x_4, \end{cases} \qquad \text{esto es,} \quad A = \begin{pmatrix} -2 & -2 & 1 & 0 \\ -11 & -4 & 0 & 1 \end{pmatrix}.$$

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

Seguramente el "a veces se presenta el problema" con el que comienza este apartado te resulte muy dudoso. No hay que irse a situaciones muy raras para encontrarlo. Veamos un ejemplo que es una simple cuestión sobre combinaciones lineales.

Nos preguntamos cuáles son los vectores de $\mathbb{R}^3$ que son simultáneamente combinación lineal de

$$C_1 = \big\{(3,3,4)^t, (2,1,5)^t\big\} \qquad \text{y de} \qquad C_2 = \big\{(1,1,2)^t, (1,3,-4)^t\big\}.$$

Para ello, hallamos las ecuaciones del subespacio generado por el primer conjunto:

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

y la ecuación que define $\mathcal{L}(C_1)$ es $-11x_1 + 7x_2 + 3x_3 = 0$. Con el segundo conjunto procedemos de la misma forma:

$$\left(\begin{array}{cc|c} 1 & 1 & x_1 \\ 1 & 3 & x_2 \\ 2 & -4 & x_3 \end{array}\right) \underset{\substack{f_2 \mapsto f_2 - f_1 \\ f_3 \mapsto f_3 - 2f_1}}{\longrightarrow} \left(\begin{array}{cc|c} 1 & 1 & x_1 \\ 0 & 2 & x_2 - x_1 \\ 0 & -6 & x_3 - 2x_1 \end{array}\right) \underset{f_3 \mapsto f_3 + 3f_2}{\longrightarrow} \left(\begin{array}{cc|c} 1 & 1 & x_1 \\ 0 & 2 & x_2 - x_1 \\ 0 & 0 & -5x_1 + 3x_2 + x_3 \end{array}\right)$$

para obtener la ecuación del subespacio $-5x_1 + 3x_2 + x_3 = 0$. Los vectores que son combinación lineal de ambos conjuntos son los que satisfacen ambas ecuaciones. Esto nos lleva a un sistema de ecuaciones lineales que debemos resolver:

$$\begin{pmatrix} -11 & 7 & 3 \\ -5 & 3 & 1 \end{pmatrix} \underset{f_2 \mapsto 11f_2 - 5f_1}{\longrightarrow} \begin{pmatrix} -11 & 7 & 3 \\ 0 & -2 & -4 \end{pmatrix} \quad \text{y} \quad x_3 = \lambda \;\Rightarrow\; x_2 = -2\lambda,\; x_1 = -\lambda.$$

Entonces los vectores buscados son los de la forma $\lambda(-1,-2,1)^t$.