
ESPACIOS VECTORIALES Y APLICACIONES LINEALES

Espacios y subespacios vectoriales

Un *espacio vectorial* sobre un conjunto de números K es intuitivamente un conjunto en el que tenemos definida una suma y una multiplicación por números con las propiedades habituales. La definición rigurosa es más complicada requiriendo la estructura algebraica de grupo abeliano con la suma y cuatro propiedades que ligan la suma y la multiplicación. Los elementos de un espacio vectorial se llaman *vectores* y los elementos del cuerpo (los números) *escalares*.

En principio los vectores pueden diferir mucho de la idea habitual que tenemos acerca de ellos. Por ejemplo, los polinomios forman un espacio vectorial y también las matrices $\mathcal{M}_{m \times n}$. En este curso prácticamente sólo nos ocuparemos de \mathbb{R}^n y de sus subespacios.

Un *subespacio vectorial* es un espacio vectorial incluido en otro con las mismas operaciones. Las propiedades de espacio vectorial se cumplen inmediatamente en un subconjunto siempre que las operaciones estén bien definidas, por ello para demostrar que cierto subconjunto S de un espacio vectorial sobre K , digamos sobre \mathbb{R} para simplificar, es un subespacio basta comprobar que no nos salimos de él al sumar o multiplicar por números, esto es:

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

Ejemplo. Si definimos los siguientes subconjuntos de \mathbb{R}^3 :

$$\begin{aligned} V_1 &= \{(x, y, z) \in \mathbb{R}^3 : xyz = 0\}, & V_2 &= \{(x, y, z) \in \mathbb{R}^3 : x + y + z = 1\}, \\ V_3 &= \{(x, y, z) \in \mathbb{R}^3 : x + y + z = 0\}, & V_4 &= \{(x, y, z) \in \mathbb{R}^3 : x = 0, y + 2z = 0\}; \end{aligned}$$

el primero no es subespacio porque $(0, 1, 1)$ y $(1, 0, 0)$ están en V_1 pero no su suma; el segundo tampoco lo es porque $(1, 0, 0) \in V_2$ pero $2 \cdot (1, 0, 0) \notin V_2$; finalmente V_3 y V_4 sí son subespacios. La razón es simplemente que podemos separar sumas de cosas que dan cero con paréntesis o que podemos multiplicar algo igualado a cero por cualquier número.

Una conclusión del ejemplo anterior es que un sistema de ecuaciones lineales igualadas a cero en \mathbb{R}^n siempre definen un subespacio vectorial. De hecho, todos los subespacios de \mathbb{R}^n se pueden expresar de esta forma [HVZ12].

Dados vectores $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_n$, una *combinación lineal* de ellos es cualquier expresión del tipo $\lambda_1 \vec{v}_1 + \lambda_2 \vec{v}_2 + \dots + \lambda_n \vec{v}_n$ con $\lambda_1, \dots, \lambda_n \in K$. Los vectores que se obtienen como combinaciones lineales de elementos de un conjunto de vectores $C = \{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_k\}$ forman el *subespacio generado* por C que se denota con $\mathcal{L}(\{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_k\})$. Es fácil ver que realmente es un subespacio vectorial.

Se dice que los vectores $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_k$ son *linealmente independientes* si ninguno es combinación lineal de los otros. Otra forma de expresar esto es que $\lambda_1 = \lambda_2 = \dots = \lambda_k = 0$ es

la única solución de

$$\lambda_1 \vec{v}_1 + \lambda_2 \vec{v}_2 + \cdots + \lambda_k \vec{v}_k = \vec{0}.$$

Con ello decidir si ciertos vectores son linealmente independientes se reduce a estudiar si un sistema homogéneo (igualado a cero) tiene solución única. Si utilizamos reducción de Gauss, esto equivale a que al poner los vectores en columna haya tantos escalones como columnas en la matriz escalonada.

Ejemplo. Estudiemos si los vectores $(1, 2, 1)$, $(2, 1, 0)$, $(4, 5, 2)$ de \mathbb{R}^3 son linealmente independientes. Como acabamos de mencionar, los ponemos en columna y aplicamos reducción de Gauss:

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

El sistema tiene infinitas soluciones (porque el número de escalones, dos, no coincide con el de columnas, tres), por tanto son linealmente independientes.

Una *base* B de un espacio vectorial V es un subconjunto que verifica $\mathcal{L}(B) = V$ (*sistema de generadores*) y que es linealmente independiente.

Intuitivamente, es un conjunto de vectores que no tiene información redundante y que sirve para construir todos los vectores de un subespacio. Más formalmente, dada una base $B = \{\vec{b}_1, \vec{b}_2, \dots, \vec{b}_n\}$ cada vector \vec{v} se puede escribir de forma única como combinación lineal $\vec{v} = \lambda_1 \vec{b}_1 + \lambda_2 \vec{b}_2 + \cdots + \lambda_n \vec{b}_n$. Los números $\lambda_1, \lambda_2, \dots, \lambda_n$ se llaman *coordenadas* o *componentes* de \vec{v} en la base B .

Un espacio vectorial V puede tener muchas bases pero todas ellas tienen el mismo número de elementos, llamado *dimensión* que se indica con $\dim V$.

Ejemplo. El ejemplo más simple de una base en \mathbb{R}^n es la llamada *base canónica* $B = \{\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n\}$, donde \vec{e}_j es el vector con todas sus coordenadas cero excepto la j -ésima que vale 1. Las coordenadas de un vector con respecto a esta base son las coordenadas en el sentido habitual. Así en \mathbb{R}^3 ,

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

Si sabemos de antemano la dimensión de un subespacio, no es necesario comprobar la condición de sistema de generadores, dicho de otra forma, en un espacio vectorial de dimensión n siempre n vectores linealmente independientes forman una base.

Por otro lado, siempre que tengamos un subespacio definido por ecuaciones lineales igualadas a cero, al resolver el sistema habremos expresado las soluciones como combinación lineal de vectores multiplicados por parámetros arbitrarios. Si hemos utilizado reducción de Gauss (o cualquier método sin añadir información redundante), estos vectores siempre serán linealmente independientes y por tanto formarán una base.

Ejemplo. Para hallar una base del subespacio $\{(x, y, z) \in \mathbb{R}^3 : x + 2y + 3z = 0\}$, debemos resolver la ecuación $x + 2y + 3z = 0$. Obviamente podemos escoger las dos últimas variables como parámetros arbitrarios: $y = \lambda$, $z = \mu$ y consecuentemente $x = -2\lambda - 3\mu$. Entonces cada vector del subespacio es de la forma

$$(-2\lambda - 3\mu, \lambda, \mu) = \lambda(-2, 1, 0) + \mu(-3, 0, 1).$$

Los vectores $(-2, 1, 0)$ y $(-3, 0, 1)$ forman una base porque todos los vectores del subespacio son combinación lineal de ellos y porque son linealmente independientes (por la construcción o porque uno no es múltiplo de otro).

Por definición, $B = \{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_k\}$ es siempre un sistema de generadores del subespacio $\mathcal{L}(\{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_k\})$, pero podría no ser base porque algunos vectores fueran combinaciones lineales de otros (linealmente dependientes). A veces se presenta el problema de quitar algunos vectores para obtener una base. Si hemos comprobado que no son linealmente independientes usando reducción de Gauss, siempre los vectores de las columnas pivote dan lugar a una base.

El número de columnas pivote, esto es, el número de vectores linealmente independientes, se llama *rango*. Es bien conocido que la discusión de las soluciones de un sistema lineal se reduce a consideraciones sobre el rango. Aunque hay una definición con determinantes, ésta suele ser poco eficiente, especialmente para dimensión mayor que tres.

Ejemplo. Sabíamos por un ejemplo anterior que los vectores $\vec{v}_1 = (1, 2, 1)$, $\vec{v}_2 = (2, 1, 0)$, $\vec{v}_3 = (4, 5, 2)$ no son linealmente independientes. Por tanto no son base de $\mathcal{L}(\{\vec{v}_1, \vec{v}_2, \vec{v}_3\})$. Como al aplicar reducción de Gauss las columnas pivote eran la primera y la segunda, se tiene que $B = \{\vec{v}_1, \vec{v}_2\}$ es base de este subespacio.

Aplicaciones lineales

Una *aplicación lineal* (o *transformación lineal*) es una función entre espacios vectoriales que preserva las operaciones (suma de vectores y producto por escalares). Desde el punto de vista práctico podemos considerar que una aplicación lineal $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ es siempre una función de la forma $f(\vec{x}) = A\vec{x}$ con $A \in \mathcal{M}_{m \times n}$ cuando escribimos \vec{x} y $f(\vec{x})$ como vectores columna. Las filas de A están formadas por los coeficientes que aparecen en las coordenadas de $f(\vec{x})$.

Geoméricamente, una aplicación lineal es una manera de deformar los objetos [Gol86], en cierta forma como verlos en perspectiva. Dicho sea de paso, las aplicaciones lineales que dan la perspectiva son cruciales en el software 3D (por ejemplo, los videojuegos) y curiosamente se representan por matrices $\mathcal{M}_{4 \times 4}$.

Ejemplo. La aplicación $f : \mathbb{R}^3 \rightarrow \mathbb{R}^2$ dada por $f(x, y, z) = (x + y + 2z, x - y)$ es lineal y su matriz es

$$A = \begin{pmatrix} 1 & 1 & 2 \\ 1 & -1 & 0 \end{pmatrix} \quad \text{porque} \quad \begin{pmatrix} x + y + 2z \\ x - y \end{pmatrix} = \begin{pmatrix} 1 & 1 & 2 \\ 1 & -1 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix}.$$

En realidad la descripción anterior de las aplicaciones lineales es un poco simplista y con un poco más de rigor sólo se define la matriz de una aplicación lineal una vez que se ha fijado una base con respecto a la cual definir las coordenadas. Sin entrar en detalles (véase [HVZ12]), si pensamos en una aplicación lineal de \mathbb{R}^n en \mathbb{R}^n como una función que a cada $\vec{x} \in \mathbb{R}^n$ le asigna $\vec{y} = A\vec{x}$, otro “observador” que usase una base distinta vería $\vec{y}' = C\vec{y}$ y $\vec{x}' = C\vec{x}$ con lo cual $\vec{y}' = CAC^{-1}\vec{x}'$, entonces para él la matriz sería CAC^{-1} . Ésta es la fórmula de *cambio de base*. Aunque aquí no profundizaremos sobre ella, motiva las fórmulas que aparecen al diagonalizar.

Se llama *núcleo* de una aplicación lineal f , y se escribe $\text{Nuc}(f)$, a las soluciones de $f(\vec{x}) = \vec{0}$. Siempre forman un subespacio y, como se ha sugerido antes, todos los subespacios de \mathbb{R}^n se pueden conseguir así.

Ejemplo. El núcleo de $f(x, y, z) = (x + y + z, x + y, z)$ es la solución de $x + y + z = 0$, $x + y = 0$, $z = 0$. Es fácil ver que se tiene $x = \lambda$, $y = -\lambda$, $z = 0$, con λ arbitrario. Así $\text{Nuc}(f) = \{\lambda(1, -1, 0)\}$ que también es $\mathcal{L}(\{(1, -1, 0)\})$.

Las aplicaciones lineales más importantes son las que aplican \mathbb{R}^n en sí mismo y por tanto tienen una matriz cuadrada $A \in \mathcal{M}_{n \times n}$. Un problema natural que aparece en muchas aplicaciones es saber si utilizando una base adecuada (por así decirlo, cambiando el sistema de referencia) se podría simplificar A hasta hacer que sea muy sencilla. Esto es lo que motiva la *diagonalización de matrices*.

Dada A una matriz cuadrada, se llama *ecuación característica* a $|A - \lambda I| = 0$, y *polinomio característica* al primer miembro. Cada una de sus raíces se dice que es un *autovalor* (o *valor propio*) y, dado un autovalor λ , se llama *autovector* (o *vector propio*) a cualquier vector $\vec{v} \neq \vec{0}$ tal que $(A - \lambda I)\vec{v} = \vec{0}$.

Como estamos identificando aplicaciones lineales $\mathbb{R}^n \rightarrow \mathbb{R}^n$ y matrices cuadradas, cuando hablemos de los autovalores y autovectores de una aplicación lineal, nos estaremos refiriendo a los que corresponden a su matriz.

Ejemplo. Vamos a hallar los autovalores y autovectores de

$$f: \mathbb{R}^2 \rightarrow \mathbb{R}^2 \quad \text{con} \quad A = \begin{pmatrix} 2 & 2 \\ 1 & 3 \end{pmatrix} \\ \vec{x} \mapsto A\vec{x}$$

Los autovalores se obtienen resolviendo la ecuación característica:

$$\begin{vmatrix} 2 - \lambda & 2 \\ 1 & 3 - \lambda \end{vmatrix} = \lambda^2 - 5\lambda + 4 = 0 \quad \implies \quad \lambda_1 = 1, \quad \lambda_2 = 4.$$

Para $\lambda = 1$ los autovectores son los múltiplos (no nulos) de $(-2, 1)$ porque

$$\begin{pmatrix} 2 - 1 & 2 \\ 1 & 3 - 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \quad \implies \quad x = -2\mu, \quad y = \mu.$$

De la misma forma, para $\lambda = 4$ los autovectores son los múltiplos (no nulos) de $(1, 1)$ porque

$$\begin{pmatrix} 2-4 & 2 \\ 1 & 3-4 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \implies x = \mu, \quad y = \mu.$$

Una matriz $A \in \mathcal{M}_{n \times n}$, o la aplicación lineal a la que representa, se dice que es *diagonalizable* si podemos encontrar n autovectores linealmente independientes. Dicho de otra forma, si hay una base n -dimensional formada por autovectores. Si los autovalores respectivos son $\lambda_1, \lambda_2, \dots, \lambda_n$, entonces se cumple

$$(1) \quad A = P \begin{pmatrix} \lambda_1 & & \\ & \lambda_2 & \\ & & \ddots \\ & & & \lambda_n \end{pmatrix} P^{-1}$$

donde P es la matriz formada por los autovectores elegidos colocados ordenadamente en columna y la matriz central es diagonal (los elementos no indicados son nulos). La terminología al uso es decir que A *diagonaliza en la base* formada por las columnas de P .

Ejemplo. En el ejemplo anterior, si escogemos los autovectores $(-2, 1)$ y $(1, 1)$, respectivamente, para los autovalores 1 y 4, se tiene

$$A = \begin{pmatrix} 2 & 2 \\ 1 & 3 \end{pmatrix} = \begin{pmatrix} -2 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 4 \end{pmatrix} \begin{pmatrix} -2 & 1 \\ 1 & 1 \end{pmatrix}^{-1}.$$

Con otra elección de los autovectores, también se tendría una igualdad válida.

Un resultado asegura que los autovectores de autovalores distintos son siempre linealmente independientes. Por tanto, si la ecuación característica de una matriz $A \in \mathcal{M}_{n \times n}$ tiene n raíces distintas, es diagonalizable.

Ejemplo. La aplicación lineal $f(x, y) = (-y, x)$ tiene una matriz cuyo polinomio característico es $\lambda^2 + 1$. No tiene raíces reales pero sí dos raíces complejas distintas. Entonces es diagonalizable sobre \mathbb{C} (pero no sobre \mathbb{R}). Un cálculo como el de antes prueba que una posible elección de los autovectores correspondientes a los autovalores $\lambda = \pm i$ es $(\pm i, 1)$.

La única manera de que una matriz no sea diagonalizable sobre \mathbb{C} es que haya raíces múltiples y que para un autovalor no haya tantos autovectores independientes como la multiplicidad. Es decir, para ser diagonalizable debe haber dos autovectores independientes para una raíz doble de la ecuación característica, tres para una triple y así sucesivamente.

Ejemplo. Calculemos todos los autovalores y autovectores de $f : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ dada por $f(\vec{x}) = A\vec{x}$ donde

$$A = \begin{pmatrix} 4 & -1 & 6 \\ 2 & 1 & 6 \\ 2 & -1 & 8 \end{pmatrix}.$$

Unos cálculos, que se pueden simplificar con las propiedades de los determinantes, prueban que la ecuación característica es

$$|A - \lambda I| = -\lambda^3 + 13\lambda^2 - 40\lambda + 36 = (9 - \lambda)(\lambda - 2)^2 = 0.$$

Por tanto hay dos autovalores: $\lambda_1 = 9$ y $\lambda_2 = 2$. Resolviendo $(A - 9I)\vec{x} = \vec{0}$ se obtiene que $(1, 1, 1)$ es un autovector para $\lambda_1 = 9$. Para que sea diagonalizable tiene que haber dos autovectores independientes para $\lambda_2 = 2$ (pues es una raíz doble). Al resolver el sistema $(A - 2I)\vec{x} = \vec{0}$ obtenemos

$$\begin{pmatrix} 2 & -1 & 6 \\ 2 & -1 & 6 \\ 2 & -1 & 6 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \implies z = \lambda, \quad y = \mu, \quad x = \frac{1}{2}\mu - 3\lambda.$$

Es decir $(x, y, z) = \lambda(-3, 0, 1) + \mu(1/2, 1, 0)$ son autovectores y podemos elegir $(-3, 0, 1)$, $(1/2, 1, 0)$ y la aplicación lineal es diagonalizable. La expresión (1) sería en este caso

$$\begin{pmatrix} 4 & -1 & 6 \\ 2 & 1 & 6 \\ 2 & -1 & 8 \end{pmatrix} = \begin{pmatrix} 1 & -3 & 1/2 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix} \begin{pmatrix} 9 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix} \begin{pmatrix} 1 & -3 & 1/2 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}^{-1}.$$

A pesar de que la situación más común es que una matriz sea diagonalizable sobre \mathbb{C} , hay ejemplos sencillos en que no ocurre así porque faltan autovectores.

Ejemplo. Comprobemos que la aplicación lineal $f(x, y) = (x + y, y)$ no es diagonalizable. Se tiene

$$\begin{vmatrix} 1 - \lambda & 1 \\ 0 & 1 - \lambda \end{vmatrix} = (1 - \lambda)^2 = 0, \quad (A - I)\vec{x} = \vec{0} \implies \begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Esto implica que todos los autovectores son múltiplos de $(1, 0)$ y por tanto no hay suficientes para diagonalizar.

Las matrices simétricas de dimensión n siempre son diagonalizables y sus autovalores son siempre reales [Gol86]. Veremos de hecho más adelante, que para estas matrices no habría que hacer ningún esfuerzo para calcular P^{-1} si quisiéramos comprobar (1).

Una curiosidad bastante misteriosa que se explica en [Lax97] (y que parece que fue primero descubierta dentro de la física cuántica) es que cuando uno varía continuamente los elementos de una matriz simétrica las gráficas que representan los autovalores parecen evitar cortarse.

Referencias. Hay muchos libros de álgebra lineal y casi todos tienen contenidos parecidos. Uno con muchos ejemplos y buenas explicaciones es [HVZ12]. Una faceta del álgebra lineal, en la que desafortunadamente no incidimos en este curso, es la cantidad de aplicaciones que tiene. Éstas aplicaciones están en gran medida sustentadas por la posibilidad de

programar eficientemente muchos cálculos de álgebra lineal. Un libro que cubre las aplicaciones y los cálculos numéricos es [Str80]. Por otro lado, [Gol86] satisfará a los que tengan interés en la interpretación geométrica y física del álgebra lineal, aunque quizá no sea fácil de encontrar. Por último, para los estudiantes muy avanzados, [Lax97] es un libro escrito por un matemático de primera línea que constituye una excepción a la uniformidad de temas de los libros de álgebra lineal.

Referencias

- [Gol86] L. I. Golovina. *Algebra Lineal y Algunas de sus Aplicaciones*. Mir, 1986.
- [HVZ12] E. Hernández, M.J. Vázquez, and M.A. Zurro. *Álgebra lineal y Geometría*. Pearson/Addison Wesley, tercera edición, 2012.
- [Lax97] P. D. Lax. *Linear algebra*. Pure and Applied Mathematics (New York). John Wiley & Sons, Inc., New York, 1997. A Wiley-Interscience Publication.
- [Str80] G. Strang. *Linear algebra and its applications*. Academic Press, New York-London, second edition, 1980.