Estructuras Abstractas I: Transformaciones Lineales, Matrices y Composición

Investigación: José Portillo

$$\Big\langle \langle K, +_g, \cdot \rangle,\ \langle E, +_E \rangle,\ \circ \Big\rangle$$

El sistema anterior se denomina *espacio vectorial* sobre el cuerpo K [2]. Los elementos de K se llaman *escalares*, y los elementos de E se llaman *vectores*. La operación $\circ : K \times E \to E$ representa la multiplicación escalar, y $+_E$ corresponde a la suma vectorial. En lo que sigue, los símbolos $\alpha_1, \dots, \alpha_k$ denotan siempre escalares de K, y los símbolos $x_1, \dots, x_k$ denotan vectores de E.

De manera intuitiva, un espacio vectorial es un conjunto de vectores que pueden sumarse y escalarse mediante elementos de un cuerpo.

Ejemplo (el espacio Γ^n, [3]).
Sea Γ un cuerpo. Considérese el conjunto

$$\Gamma^n = \Gamma \times \cdots \times \Gamma$$

de n-tuplas $(\xi^1, \dots, \xi^n)$, con $\xi^i \in \Gamma$. La suma se define como

$$(\xi^1, \dots, \xi^n) + (\eta^1, \dots, \eta^n) = (\xi^1 + \eta^1, \dots, \xi^n + \eta^n),$$

y la multiplicación escalar como

$$\lambda(\xi^1, \dots, \xi^n) = (\lambda\xi^1, \dots, \lambda\xi^n), \quad \lambda \in \Gamma.$$

Con estas operaciones, Γ^n constituye un espacio vectorial sobre Γ, denominado el n-*espacio sobre* Γ. En particular, Γ es un espacio vectorial sobre si mismo (Γ), donde la multiplicación escalar coincide con la multiplicación del cuerpo.

Dados vectores $x_1, \dots, x_k \in E$ y escalares $\alpha_1, \dots, \alpha_k \in K$, una expresión de la forma

$$\alpha_1 \circ x_1 +_E \alpha_2 \circ x_2 +_E \cdots +_\nu \alpha_k \circ x_k$$

se denomina *combinación lineal*.

Un subconjunto $S \subset E$ se llama un *sistema de generadores* (o un *conjunto generador*) de E si todo vector $x \in E$ puede escribirse como una combinación lineal de elementos de S [3]. Es decir, para cada $x \in E$ existen vectores $s_1, \dots, s_k \in S$ y escalares $\alpha_1, \dots, \alpha_k \in K$ tales que

$$x = \alpha_1 \circ s_1 +_E \cdots +_E \alpha_k \circ s_k.$$

Una familia finita de vectores $\{x_1, \dots, x_k\} \subset E$ se llama *linealmente independiente* si la relación

$$\alpha_1 \circ x_1 +_E \cdots +_E \alpha_k \circ x_k = 0$$

implica que

$$\alpha_1 = \cdots = \alpha_k = 0.$$

Si existe tal relación no trivial, la familia se llama *linealmente dependiente*.

Un conjunto finito de vectores $\{b_1, \ldots, b_n\} \subset E$ se llama *base* del espacio vectorial E si es tanto un sistema de generadores de E como linealmente independiente.

De manera equivalente, $\{b_1, \ldots, b_n\}$ es una base de E si cada vector $x \in E$ puede escribirse *de manera única* como combinación lineal

$$x = \alpha_1 \circ b_1 +_E \alpha_2 \circ b_2 +_E \cdots +_E \alpha_n \circ b_n, \quad \alpha_1, \ldots, \alpha_n \in K.$$

Una base, por lo tanto, proporciona un conjunto mínimo y no redundante de bloques de construcción para todo el espacio. Una vez que se fija una base, cada vector queda completamente determinado por sus coeficientes relativos a esa base, por lo que podemos decir que la base genera el espacio.

Si el espacio vectorial E admite una base finita, el número de vectores en cualquier base de E se llama la *dimensión* de E y se denota por $\dim E$.

Sean E y G espacios vectoriales sobre el mismo cuerpo K. Una función $\psi : E \to G$ se denomina *transformación lineal* si preserva combinaciones lineales [1]:

$$\psi(\alpha_1 \circ_E x_1 + \cdots + \alpha_k \circ_E x_k) = \alpha_1 \circ_G \psi(x_1) + \cdots + \alpha_k \circ_G \psi(x_k).$$

Las transformaciones lineales conservan la estructura algebraica de los espacios vectoriales, independientemente de cualquier elección de coordenadas.

Sean $B_E = \{e_1, \ldots, e_n\}$ y $B_G = \{g_1, \ldots, g_m\}$ bases de E y G. Cada imagen $\psi(e_i)$ puede escribirse de manera única como

$$\psi(e_i) = a_{1i} \circ g_1 +_G \cdots +_G a_{mi} \circ g_m.$$

Los escalares a_{ji} forman una matriz $m \times n$

$$A = (a_{ji}),$$

llamada la *matriz de* ψ *relativa a las bases elegidas*.

Así, las matrices no definen transformaciones lineales por sí mismas; son representaciones concretas de dichas transformaciones una vez que se han fijado bases (sistemas de coordenadas).

En muchas aplicaciones, un vector no representa simplemente un elemento abstracto de un espacio vectorial, sino el *estado* de un sistema relativo a un marco referencial elegido. Algebraicamente, un marco referencial no es más que la elección de una base.

Sea E un espacio vectorial de dimensión finita y sea

$$B = \{e_1, \ldots, e_n\}$$

una base de E. Todo vector $x \in E$ puede escribirse de manera única como

$$x = \alpha_1 \circ e_1 +_E \cdots +_E \alpha_n \circ e_n.$$

Los escalares $(\alpha_1, \ldots, \alpha_n)$ son las *coordenadas* del estado x respecto al marco referencial B. Representarlas como un vector columna

$$[x]_B = \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{bmatrix},$$

es una notación conveniente que codifica la acción de las combinaciones lineales relativas a la base elegida.

Si se elige un segundo marco referencial

$$B' = \{e'_1, \dots, e'_n\}$$

para el mismo espacio E, el vector x sigue siendo el mismo objeto abstracto, pero sus coordenadas respecto a B' forman un vector columna distinto $[x]_{B'}$.

El paso de $[x]_B$ a $[x]_{B'}$ no es arbitrario. Está determinado por una transformación lineal

$$T : K^n \longrightarrow K^n,$$

cuyo matriz depende únicamente de la relación entre las dos bases. Así,

$$[x]_{B'} = A\,[x]_B,$$

donde A es la matriz del cambio de marco referencial.

Este procedimiento constituye el origen operativo de la multiplicación matriz–vector: una matriz actúa sobre un vector columna porque representa la transformación lineal que opera sobre la representación en coordenadas del estado.

El requisito de que el número de columnas de A coincida con el número de componentes de $[x]_B$ no es una convención, sino que refleja la necesidad de que el codominio de una aplicación lineal coincida con el dominio de la siguiente.

En este sentido, los vectores columna representan estados, las matrices representan transformaciones entre marcos referenciales y la multiplicación de matrices corresponde a cambios sucesivos de coordenadas. Las reglas computacionales habituales derivan, por tanto, de la estructura abstracta introducida anteriormente, y no de trucos algebraicos independientes.

Esta formulación abstracta es independiente de la geometría, y sin embargo subyace en la mayoría de los modelos cuantitativos en ciencia e ingeniería. En el aprendizaje automático moderno (machine learning), los *embeddings* y las matrices de pesos son transformaciones lineales sobre espacios vectoriales de dimensión elevada, compuestas con activaciones no lineales. Incluso cuando la no linealidad domina, las transformaciones lineales proporcionan la columna vertebral estructural: la universalidad de una estructura lineal creada por el ser humano, aunque el universo mismo no sea lineal.

Nota: Desde un punto de vista operacional, un sistema de ecuaciones lineales de la forma

$$\begin{cases} a_{11}x_1 + \cdots + a_{1n}x_n = b_1 \\ \vdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n = b_m. \end{cases}$$

es visto como una colección de ecuaciones a resolver para las incógnitas $x_1, \dots, x_n$.

Desde el punto de vista abstracto, este sistema no es más que la expresión en coordenadas de una transformación lineal.

De hecho, sea

$$\psi : K^n \longrightarrow K^m$$

la transformación lineal cuya matriz relativa a las bases canónicas es

$$A = \begin{bmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{bmatrix}.$$

Escribiendo

$$x = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \quad \text{y} \quad b = \begin{bmatrix} b_1 \\ \vdots \\ b_m \end{bmatrix},$$

el sistema anterior se expresa de manera equivalente como la ecuación única

$$\psi(x) = b, \quad \text{o, en coordenadas,} \quad Ax = b.$$

Así, un sistema de ecuaciones lineales especifica un vector $b \in K^m$ *y pregunta si este se encuentra en la imagen de una transformación lineal* $\psi : K^n \to K^m$*. Preguntas sobre existencia, unicidad o multiplicidad de soluciones corresponden a propiedades geométricas de esta aplicación, como inyectividad y sobreyectividad, y son independientes de cualquier algoritmo particular usado para calcularlas.*

Algoritmos como la eliminación de Gauss son métodos para encontrar un vector $x \in K^n$ *que satisfaga* $Ax = b$ *relativo a un sistema de referencia elegido (base). Desde la perspectiva abstracta, son simplemente procedimientos para explorar la imagen y el núcleo de una transformación lineal en coordenadas. Es decir, la eliminación de Gauss explota la estructura del mapa lineal codificada por la matriz* A *para determinar si existe una solución, si es única, y cómo expresar todas las soluciones posibles, trabajando completamente en el sistema de coordenadas elegido.*

El conjunto de todas las transformaciones lineales de E al cuerpo base K se denota por

$$E^* = L(E, K)$$

y se llama el *espacio dual* de E. Un elemento $x^* \in E^*$ se denomina *funcional lineal*. Así,

$$x^* : E \to K$$

satisface

$$x^*(\alpha_1 \circ_E x_1 +_E \cdots +_E \alpha_k \circ_E x_k) = \alpha_1 x^*(x_1) + \cdots + \alpha_k x^*(x_k).$$

Los funcionales lineales asignan valores escalares a vectores conservando la estructura lineal.

Si E es de dimensión finita con base $B_E = \{e_1, \ldots, e_n\}$, cada $x^* \in E^*$ queda determinado de manera única por

$$x^*(e_1), \ldots, x^*(e_n).$$

Estos escalares forman un vector fila

$$[\, x^*(e_1) \;\ldots\; x^*(e_n) \,].$$

En contraste, los vectores de E se representan como vectores columna. Ambos objetos habitan espacios diferentes: E y E^*.

En muchas situaciones, un vector se utiliza para definir un funcional lineal. Esto requiere una estructura adicional. Si E está dotado de un producto interior $\langle \cdot, \cdot \rangle$, entonces cada vector $x \in E$ determina un funcional lineal mediante

$$x^*(y) = \langle x, y \rangle.$$

Así, el producto interior define una aplicación

$$E \longrightarrow E^*, \quad x \longmapsto x^*.$$

El espacio vectorial E introducido previamente se ve ahora enriquecido con una estructura geométrica adicional. Un *espacio con producto interior* es un espacio vectorial

$$\Big\langle \langle K, +_g, \cdot \rangle,\ \langle E, +_E \rangle,\ \circ,\ \langle \cdot, \cdot \rangle \Big\rangle$$

donde

$$\langle \cdot, \cdot \rangle : E \times E \to K$$

representa un producto interior.

Si se utiliza una base no ortonormal, la expresión en coordenadas del producto interior involucra coeficientes adicionales, típicamente representados mediante una matriz simétrica definida positiva.

Para comprender por qué el producto interior toma la forma de una suma de productos en coordenadas, es necesario hacer explícita la estructura geométrica impuesta.

Sea E un espacio vectorial de dimensión finita con producto interior $\langle \cdot, \cdot \rangle$, y sea

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

una base de E. La base B se denomina *ortonormal* si

$$\langle e_i, e_j \rangle = \delta_{ij},$$

donde δ_{ij} es la *delta de Kronecker*, definida por

$$\delta_{ij} = \begin{cases} 1, & \text{si } i = j, \\ 0, & \text{si } i \neq j. \end{cases}$$

Esta condición codifica dos hechos geométricos: cada vector de la base tiene longitud unitaria, y los vectores distintos son ortogonales.

Si los vectores $x, y \in E$ se escriben respecto a esta base como

$$x = \sum_{i=1}^{n} x_i e_i, \qquad y = \sum_{j=1}^{n} y_j e_j,$$

entonces, por la bilinealidad del producto interior,

$$\langle x, y \rangle = \sum_{i,j} x_i y_j \langle e_i, e_j \rangle.$$

La condición de ortonormalidad $\langle e_i, e_j \rangle = \delta_{ij}$ elimina todos los términos cruzados con $i \neq j$, dejando

$$\langle x, y \rangle = \sum_{i=1}^{n} x_i y_i.$$

Por tanto, la conocida fórmula de coordenadas para el producto escalar no es una definición, sino la expresión en coordenadas del producto interior respecto a una base ortonormal. Solo los componentes en la misma dirección geométrica contribuyen al valor del producto interior; los componentes en direcciones ortogonales no aportan nada.

Cuando los vectores se escriben en coordenadas como vectores columna, el funcional lineal correspondiente se representa como vector fila. La operación que realiza esta conversión en coordenadas es la *transpuesta*.

Ejemplo.
Sea

$$x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \in K^2.$$

Con el producto interno estándar en K^2 (por ejemplo, $\mathsf{K} = \mathbb{R}$),

$$\langle x, y \rangle = x_1 y_1 + x_2 y_2,$$

el funcional lineal asociado se representa mediante el vector fila

$$x^\mathsf{T} = [\, x_1 \ \ x_2 \,].$$

Su acción sobre un vector $y = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$ está dada por

$$x^\mathsf{T} y = \langle x, y \rangle.$$

La operación de transpuesta existe, por tanto, para convertir vectores en funcionales lineales cuando hay un producto interior. No es intrínseca a los espacios vectoriales, sino que surge de la estructura geométrica adicional impuesta sobre ellos.

La multiplicación de matrices no es una regla algebraica arbitraria; es la expresión en coordenadas de la *composición de aplicaciones lineales.*

Nota: Las transformaciones lineales son funciones lineales en un solo argumento. Más generalmente, pueden considerarse funciones lineales en varios argumentos, llamadas aplicaciones multilineales*. Una aplicación multilineal*

$$\mathsf{T} : \mathsf{E}_1 \times \cdots \times \mathsf{E}_k \longrightarrow \mathsf{G}$$

es lineal en cada argumento por separado. En particular, funciones de la forma

$$\mathsf{T} : \mathsf{E} \times \cdots \times \mathsf{E} \longrightarrow \mathsf{K}$$

generalizan los funcionales lineales y las formas bilineales. Un tensor *de tipo* (k, ℓ) *sobre* E *es una aplicación multilineal*

$$\mathsf{T} : \underbrace{\mathsf{E}^* \times \cdots \times \mathsf{E}^*}_{k} \times \underbrace{\mathsf{E} \times \cdots \times \mathsf{E}}_{\ell} \longrightarrow \mathsf{K}.$$

Casos especiales incluyen escalares $(0, 0)$*, vectores* $(1, 0)$*, funcionales lineales* $(0, 1)$*, transformaciones lineales* $(1, 1)$ *y formas bilineales* $(0, 2)$*. Una vez elegida una base de* E*, un tensor puede representarse mediante un arreglo multidimensional de escalares; estos arreglos dependen de la base elegida, pero el tensor en sí no. En computación numérica y aprendizaje automático, tales arreglos se denominan comúnmente* tensores*, aunque representan expresiones en coordenadas de objetos abstractos. Los tensores de NumPy o TensorFlow son arreglos multidimensionales que representan tensores tras la elección de una base; el tensor en sí es un objeto independiente de coordenadas. Esta aclaración se incluye solo para situar vectores, vectores duales y aplicaciones lineales dentro de un marco conceptual más amplio. No se requiere ni se utiliza cálculo tensorial o teoría multilineal en este documento.*

Un vector columna representa un elemento de K^n, un vector fila representa un elemento de $(\mathsf{K}^n)^*$, y una matriz $m \times n$ representa una aplicación lineal

$$\mathsf{K}^n \xrightarrow{\psi} \mathsf{K}^m.$$

Un producto como

$$[\, x^* \,]\, A$$

está definido porque corresponde a la composición

$$\mathsf{K}^n \xrightarrow{\psi} \mathsf{K}^m \xrightarrow{x^*} \mathsf{K}.$$

Las dimensiones coinciden exactamente porque el codominio de la primera aplicación es igual al dominio de la segunda.

En cambio, un producto de dos vectores fila intentaría componer

$$K^n \to K \quad \text{con} \quad K^n \to K,$$

lo cual carece de sentido: la salida de la primera aplicación no pertenece al dominio de la segunda. El producto, por tanto, no está definido por convención, sino por necesidad.

Las dimensiones de las matrices codifican la compatibilidad de dominios y codominios. Se pueden multiplicar matrices si y solo si las aplicaciones lineales correspondientes pueden componerse.

Ejemplo numérico concreto.
Sea

$$x^*(x, y) = 3x + 2y$$

un funcional lineal sobre K^2. Respecto a la base canónica, se representa mediante el vector fila

$$[3 \;\; 2].$$

Sea

$$A = \begin{bmatrix} 2 & 3 \\ 5 & 4 \end{bmatrix}$$

la matriz de una transformación lineal $\psi : K^2 \to K^2$.
El producto

$$[3 \;\; 2]A = [3 \;\; 2] \begin{bmatrix} 2 & 3 \\ 5 & 4 \end{bmatrix} = [19 \;\; 17]$$

está definido porque representa la composición

$$K^2 \xrightarrow{\psi} K^2 \xrightarrow{x^*} K.$$

El resultado es nuevamente un vector fila, correspondiente al funcional lineal

$$(x, y) \longmapsto 19x + 17y.$$

En cambio, el producto

$$[3 \;\; 2][a \;\; b]$$

no está definido porque intentaría componer dos aplicaciones de la forma

$$K^2 \to K,$$

lo cual es imposible: la salida de la primera aplicación es un escalar, no un vector en K^2.

Resumen conceptual. Las estructuras introducidas en este memorando—espacios vectoriales, transformaciones lineales, bases, sistemas de coordenadas, espacios duales y productos internos—son todas facetas de un mismo marco lineal subyacente. Las matrices codifican transformaciones lineales relativas a una base elegida, los vectores columna codifican estados relativos a un sistema de referencia, y los sistemas de ecuaciones lineales son simplemente expresiones en coordenadas de estas transformaciones preguntando si un vector pertenece a la imagen de un mapa. Los productos internos permiten que los vectores definan funcionales lineales, conectando geometría con álgebra. Algoritmos como la eliminación de Gauss, aunque esenciales para el cálculo, no crean estructuras nuevas: exploran las consecuencias a nivel de coordenadas de las relaciones lineales abstractas ya presentes. Juntos, estos conceptos muestran cómo las definiciones algebraicas abstractas dan lugar a las herramientas computacionales y geométricas familiares usadas en ciencia e ingeniería.

Nota. Este documento es la primera de tres partes sobre estructuras lineales y geométricas abstractas. Se utilizó un modelo de lenguaje de IA como herramienta de apoyo para discusión, clarificación de conceptos y refinamiento estilístico. Todas las decisiones matemáticas y la redacción final son responsabilidad del autor.

Referencias

[1] Iribarren Ignacio L., *Álgebra Lineal*, Editorial Equinoccio, 2009.

[2] Cohen, Leon and Ehrlich, Gertrude, *The Structure of the Real Number System*, D. Van Nostrand Company, 1963.

[3] Greub, Werner, *Linear Algebra*, Springer-Verlag, 1975.