Resumen del capítulo 3

Espacios y subespacios vectoriales. Un *espacio vectorial* sobre un cuerpo es intuitivamente un conjunto en el que tenemos definida una suma y una multiplicación por números (los elementos de un cuerpo) 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*.

Los ejemplos de espacios vectoriales de espacios vectoriales sobre un cuerpo K más importantes en este curso son:

1. K^n que es el producto cartesiano $K \times \overset{n \text{ veces}}{\cdots\cdots} \times K$ con la suma y el producto por elementos de K definidos de la manera obvia. Por razones que serán claras más adelante, escribiremos en columna la colección de n elementos de K que representan cada vector de K^n. El ejemplo más común en el curso será $\mathbb{R}^n$.
2. $\mathbb{P}[x]$, el conjunto de polinomios en la variable x con coeficientes en K.
3. $\mathcal{M}_{m\times n}(K)$.

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 es un subespacio basta comprobar:

$$1)\ \vec{u},\vec{v}\in S \Rightarrow \vec{u}+\vec{v}\in S \qquad \text{y} \qquad 2)\ \vec{u}\in S,\ \lambda\in K \Rightarrow \lambda\vec{u}\in S$$

Por ejemplo, $\mathbb{R}_n[x]$, el conjunto de todos los polinomios con coeficientes reales de grado menor o igual que n (incluyendo al polinomio cero) son un subespacio de $\mathbb{P}[x]$ y en particular un espacio vectorial. Sin embargo los polinomios de grado exactamente 3 no lo son porque $1-x^3$ y $x+x^3$ están en este conjunto pero su suma no.

Un ejemplo de subespacio vectorial muy frecuente este curso es el subconjunto de K^n dado por las soluciones $\vec{x}\in K^n$ del sistema $A\vec{x}=\vec{0}$ donde $A\in\mathcal{M}_{m\times n}(K)$ (el convenio de los vectores columna hace que el producto de matrices tenga sentido).

Dados vectores $\vec{v}_1,\vec{v}_2,\dots\vec{v}_n$, una *combinación lineal* de ellos es cualquier expresión del tipo $\lambda_1\vec{v}_1+\lambda_2\vec{v}_2+\cdots+\lambda_n\vec{v}_n$ con $\lambda_1,\dots\lambda_n\in K$.

Los vectores que se obtienen como combinaciones lineales de elementos de un conjunto de vectores C forman el *subespacio generado por C* que se denota con $\langle C\rangle$ ó $\mathcal{L}(C)$. Es fácil ver que realmente es un subespacio vectorial.

Dado un subconjunto finito de C de K^n se puede escribir $\langle C\rangle$ "con ecuaciones", es decir, como $\{\vec{x}\in K^n\ :\ A\vec{x}=\vec{0}\}$ eliminando parámetros por reducción de Gauss en una combinación lineal genérica de los elementos de C.

Por ejemplo

$$C=\left\{\begin{pmatrix}1\\0\\-1\end{pmatrix},\begin{pmatrix}1\\-2\\3\end{pmatrix}\right\}\quad\Rightarrow\quad\langle C\rangle=\left\{\begin{pmatrix}x\\y\\x\end{pmatrix}\in\mathbb{R}^3\ :\ \begin{pmatrix}x\\y\\x\end{pmatrix}=\lambda\begin{pmatrix}1\\0\\-1\end{pmatrix}+\mu\begin{pmatrix}1\\-2\\3\end{pmatrix}\right\}.$$

Empleando la reducción de Gauss

$$\left(\begin{array}{cc|c}1&1&x\\0&\text{-2}&y\\\text{-1}&3&z\end{array}\right)\longrightarrow\left(\begin{array}{cc|c}1&1&x\\0&\text{-2}&y\\0&4&x+z\end{array}\right)\longrightarrow\left(\begin{array}{cc|c}1&1&x\\0&\text{-2}&y\\0&0&x+2y+z\end{array}\right).$$

Entonces la ecuación que define $\langle C\rangle$ es $x+2y+z=0$ (la matriz A sería en este caso (1 2 1), una matriz fila).

Recíprocamente, dado un subespacio de la forma $S=\{\vec{x}\in K^n\ :\ A\vec{x}=\vec{0}\}$, resolviendo esta ecuación se llegará a una solución en términos de parámetros (o simplemente la solución $\vec{0}$ que da lugar al subespacio trivial) que puede escribirse como una combinación lineal de ciertos vectores de K^n. El conjunto formado por dichos vectores satisface $\langle C\rangle=S$.

Aplicaciones lineales. Una *aplicación lineal* es una función entre espacios vectoriales que preserva las operaciones. Es decir, es una función $f:V\longrightarrow W$ tal que

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

El ejemplo que más utilizaremos en el curso (y al cual reduciremos prácticamente todos los ejemplos) es $f:K^n\longrightarrow K^m$ dada por $f(\vec{x})=A\vec{x}$ con $A\in\mathcal{M}_{m\times n}(K)$.

Hay dos subespacios asociados a una aplicación lineal $f:V\longrightarrow W$, el *núcleo* y la *imagen*, definidos respectivamente como:

$$\operatorname{Nuc}(f)=\{\vec{v}\in V\ :\ f(\vec{v})=\vec{0}\},\qquad\operatorname{Im}(f)=\{\vec{w}\in W\ :\ \exists\vec{v}\text{ con }f(\vec{v})=\vec{w}\}.$$

La aplicación lineal $f:V\longrightarrow W$ es inyectiva si y sólo si $\operatorname{Nuc}(f)=\{\vec{0}\}$ y es sobreyectiva si y sólo si $\operatorname{Im}(f)=W$. Las funciones que son inyectivas y sobreyectivas se dice que son biyectivas (y el caso de aplicaciones lineales se dice que establecen un *isomorfismo* entre V y W). Las funciones biyectivas tiene una *función inversa*, que en nuestro caso será la aplicación lineal $f^{-1}:W\longrightarrow V$ tal que $f^{-1}\circ f$ es la identidad en V y $f\circ f^{-1}$ es la identidad en W (esto significa que dejan todos los elementos invariantes en estos espacios).

Por ejemplo, si $f:\mathbb{R}^n\longrightarrow\mathbb{R}^n$ viene dada por $f(\vec{x})=A\vec{x}$ con $A\in\mathcal{M}_{n\times n}(\mathbb{R})$ invertible, entonces la función inversa es $f^{-1}(\vec{x})=A^{-1}\vec{x}$.

En el caso $f:K^n\longrightarrow K^m$ con $A\in\mathcal{M}_{m\times n}(K)$, utilizando la definición obtenemos una descripción de $\operatorname{Nuc}(f)$ con ecuaciones, $A\vec{x}=\vec{0}$, que una vez resueltas conduce a un conjunto que genera este subespacio. Por otro lado, las columnas de A siempre generan $\text{Im}(f)$ porque la estructura de la multiplicación de matrices hace que $A\vec{x} = \sum x_i \vec{c}_i$ con $\vec{c}_i$ la columna i-ésima de A. Con el procedimiento introducido antes, a partir del conjunto de columnas se pueden obtener las ecuaciones que determinan $\text{Im}(f)$.

Por ejemplo, dada la aplicación lineal $f : \mathbb{R}^2 \longrightarrow \mathbb{R}^3$ definida por

$$f \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 0 & -2 \\ -1 & 3 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix},$$

hallemos su núcleo y su imagen.

Al igualar a $\vec{0}$ y resolver el sistema se obtiene únicamente la solución trivial $x = y = 0$, por tanto $\text{Nuc}(f) = \{\vec{0}\}$. Las columnas forman un conjunto C que genera un subespacio de $\mathbb{R}^3$ que según hemos visto antes responde a la ecuación $x + 2y + z = 0$. De aquí concluimos que f es una aplicación lineal inyectiva pero no sobreyectiva, en particular no tiene inversa.