Apuntes [un esbozo] de Ampliación de Matemáticas

Pedro Fortuny Ayuso

2009-2010

Correo electrónico: `fortunypedro@uniovi.es`

Índice

Introducción general a la asignatura

De momento no introduzco nada.

CAPíTULO 1

Sitemas de Ecuaciones. Gauss

1. Introducción

Comenzamos antes que nada por definir lo que es un sistema de ecuaciones lineales y el resto de términos técnicos que necesitamos para trabajar.

DEFINICIÓN 1.1. Una *ecuación* es una expresión de igualdad entre dos cosas. Estas dos cosas se llaman *miembros* y se distinguen como *el miembro de la izquierda* (o primer miembro) y el miembro de la derecha (o segundo miembro).

Así, por ejemplo,

$$
\begin{aligned}
1+1 &= 2 \\
7 &= 2\times 3+1 \\
x &= 33 \\
3y &= y+z \\
\cos\pi &= -1 \\
\sin\pi &= 18
\end{aligned}
$$

son ecuaciones. Unas son verdad, otras (como la última) son falsas y otras son simplemente enunciados que dependen del valor de algo que no se conoce (la x de la tercera, las y, z de la cuarta). En realidad, el término ecuación se *suele* reservar para las que contienen este tipo de términos, pero lo vamos a usar indistintamente.

Repito que una ecuación puede ser *falsa* (lo cual no indica más que lo que se ha escrito es incorreto). Esto aparecerá muchas veces en esta asignatura y nos servirá para determinar cuándo un problema seguro que no tiene solución.

DEFINICIÓN 1.2. Una *incógnita* en una ecuación es un elemento cuyo valor se desconoce.

Por ejemplo, las x, y, z que aparecen arriba son incógnitas. Las incógnitas a veces se llaman *variables*, pero esta es una nomenclatura algo confusa.

DEFINICIÓN 1.3. Un *coeficiente* es un número que multiplica a una variable.

En la siguiente ecuación, si suponemos que el número a se conoce y la x no, entonces a es un coeficiente y x una variable.

$$ax + x^2 = 32.$$

DEFINICIÓN 1.4. Un *sumando* que no contiene variables se denomina *término independiente*.

En la ecuación anterior, 32 es un término independiente. Si en una ecuación solo hay un término independiente (y esto se puede conseguir siempre sumándolos todos), entonces de denominará *el* término independiente.

DEFINICIÓN 1.5. Una ecuación se llama *lineal* si las variables que aparecen en ella aparecen sin exponentes, sin multiplicar (ni dividir) por otras variables, ni parámetros de funciones.

Por ejemplo, si x e y son incógnitas:

$$\begin{aligned} x + \operatorname{sen} y &= 23 - x \\ 2x + xy &= -\pi \\ x &= y \end{aligned}$$

Las dos primeras ecuaciones *no* son lineales, mientras que la última sí. En la primera, la y aparece como parámetro de la función seno, mientras que en la segunda la x y la y se multiplican.

DEFINICIÓN 1.6. Un *sistema de ecuaciones* es una colección finita de ecuaciones (es decir, una lista finita de ecuaciones). Si todas ellas son lineales, entonces el sistema se llama *sistema de ecuaciones lineales*.

Siempre que uno tiene una ecuación ó un sistema, hace falta que a uno le indiquen cuáles son las variables. Por ejemplo:

$$\begin{aligned} ax + by &= 33 \\ 2x + 3by &= 0 \end{aligned} \tag{1}$$

Puede entenderse como un sistema (no lineal) con cuatro incógnitas (que serían a, b, x, y) o bien como un sistema lineal en el que a, b son números (indeterminados pero conocidos, por ejemplo, pesos ó cargas ó fuerzas) y las incógnitas son x, y, o de muchas otras maneras.

NOMENCLATURA 1. *En nuestros problemas las incógnitas tendrán por nombre x, y, z, t, u, v, w (si con estas es suficiente) o bien letras x, y, z con* subíndices*: x_1, x_2, y_1, y_3, z_4.*
Utilizaremos las letras a, b, c (o ellas mismas con subíndices) para indicar coeficientes *y/o* términos independientes*.*

Así pues, el sistema (1) *lo entenderemos* como un sistema lineal de dos ecuaciones con dos incógnitas (o variables): x, y. Tanto a como b se entenderá que son datos del problema.

Finalmente,

Definición 1.7. El *conjunto de soluciones de un sistema* es el conjunto de familias de números tales que al sustituir cada incógnita por uno de ellos, se satisfacen todas las ecuaciones.

Ejemplo 1. Se tienen dos pesos x, y desconocidos, pero experimentalmente se sabe que están en equilibrio en las condiciones de la figura *. Donde 2

Falta la figura, claro

Figura 1. Dos modos de equilibrio

indica un peso de $2N$ y las cotas son distancias en horizontal. ¿Podemos calcular los pesos x, y? Para ello, expresamos las condiciones de equilibrio como ecuaciones. Puesto que es un balancín (como una polea), las ecuaciones de equilibrio son ecuaciones de momentos:

$$10 \times x + 2 \times 5 = 15 \times y$$
$$7{,}5 \times 2 + 5 \times y = 15 \times x$$

La respuesta es: podemos. ¿Cómo? La manera más sencilla en este caso es el *método de sustitución* (del que hablaremos más adelante). La segunda ecuación sirve para concer x "en función de" y:

$$x = 1 + \frac{1}{3}y \tag{2}$$

(lo cual quiere decir que, si conociéramos y, conoceríamos x). Ahora bien, sabiendo esto, podemos *sustituir* este valor de x en la primera ecuación:

$$10(1 + \frac{1}{3}y) + 10 = 15y$$

de donde, fácilmente:

$$\frac{35}{3}y = 20$$

es decir, $y = \frac{12}{7}$ y, volviendo a *sustituir* ahora en la ecuación (2), obtenemos:

$$x = 1 + \frac{1}{3}\frac{12}{7} = \frac{11}{7}.$$

Así pues, el sistema *solo tiene una solución*, que expresamos

$$x = \frac{11}{7},\ y = \frac{12}{7}.$$

Pero el objetivo de este capítulo es mostrar una vía de resolución que es *generalizable fácilmente* y, sobre todo, *fácilmente automatizable*. El conocido *método de Gauss* o de reducción a fórma de escalón (ó triangular superior).

2. Aritmética de los sistemas

Para poder expresar el método de Gauss, hace falta enunciar unos principios previos que son más o menos evidentes. Vamos a hacer uso del mismo sistema de ecuaciones que antes

$$\begin{aligned} 10x + 10 &= 15y \\ 15 + 5y &= 15x \end{aligned}$$

El primer paso es *dar un orden a las variables* y *escribir el sistema de manera ordenada*, con los términos variables a la izquierda y los términos independientes a la derecha. Así, si el orden es x, y, queda:

$$\begin{aligned} 10x - 15y &= -10 \\ -15x + 5y &= -15 \end{aligned}$$

Está claro que "podemos simplificar" (todos los números son múltiplos de 5), pero no vamos a hacerlo.

La primera regla que vamos a utilizar es

Regla 1. *Si en un sistema una ecuación se sutituye por la misma multiplicada por un número distinto de* 0*, las soluciones del sistema son las mismas.*

Es decir, da igual usar una ecuación que una proporcional a ella. Eso es lo que llamamos "simplificar". Dividiendo la primera ecuación por 10, queda

$$\begin{aligned} x - 1{,}5y &= -1 \\ -15x + 5y &= -15 \end{aligned}$$

Dividenddo la segunda por 5:

$$\begin{aligned} x - 1{,}5y &= -1 \\ -3x + y &= -3 \end{aligned}$$

La siguiente regla es:

Regla 2. *Si en un sistema una ecuación se sustituye por la suma de ella misma con otra multiplicada por un número, las soluciones del sistema son las mismas.*

(No voy a explicar por qué es cierto esto). En el sistema en cuestión, podemos sumarle a la segunda ecuación la primera multiplicada por 3 y queda entonces:

$$\begin{aligned} x - 1{,}5y &= -1 \\ - 3{,}5y &= -6 \end{aligned}$$

Llegados a este punto, el sistema es sencillísimo de resolver. La y se despeja en la ecuación de abajo y queda $y = 12/7$, y la x se calcula sustituyendo el valor de la y, y queda $x = 11/7$ (claro).

Un principio que no hemos utilizado pero que es evidente, es que

Regla 3. *En un sistema de ecuaciones* no importa el orden de las ecuaciones.

Es decir, podemos colocar las ecuaciones más arriba o más abajo, según nos convenga, y el conjunto de soluciones (que *es lo único que nos interesa*) no varía.

Juntas, las tres reglas que hemos enunciado son: *

problemas aquí con indent

Regla. *Reglas para el manejo de sistemas de ecuaciones:*

1. *En un sistema de ecuaciones, el orden de las ecuaciones no importa.*
2. *Si en un sistema una ecuación se sutituye por la misma multiplicada por un número distinto de* 0*, las soluciones del sistema son las mismas.*
3. *Si en un sistema una ecuación se sustituye por la suma de ella misma con otra multiplicada por un número, las soluciones del sistema son las mismas.*

Con estas tres reglas veremos cómo somos capaces de resolver cualquier sistema de ecuaciones lineales. Vamos a hacer antes otro ejemplo.

Partimos del sistema

Ejemplo 2. Tomemos una reacción química: la producción de TNT y agua ($C_7H_5O_6N_3$ y H_2O) a partir de tolueno y ácido nítrico (C_7H_8 y HNO_3). La reacción se escribe:

$$C_7H_8 + HNO_3 \to C_7H_5O_6N_3 + H_2O.$$

En la realidad uno habitualmente quiere producir una cantidad z de TNT y necesita saber cuánto tolueno y cuánto ácido necesita. Esto se escribiría:

$$xC_7H_8 + yHNO_3 \to zC_7H_5O_6N_3 + wH_2O.$$

El problema consiste en saber cuánto valen x, y, z, w.

Para plantear bien el sistema, hemos de igualar la cantidad de cada elemento en ambos lados de la ecuación. Es decir:

$$\begin{aligned} C &\equiv & 7x &= 7z \\ H &\equiv & 8x + y &= 5z + 2w \\ O &\equiv & 3y &= 6z + w \\ N &\equiv & y &= 3z \end{aligned}$$

Reordenando todas las ecuaciones, obtenemos el suiguiente sistema:

$$\begin{array}{rrrrl} 7x & & -\,7z & & = 0 \\ 8x\,+ & y & -\,5z & -\,2w & = 0 \\ & 3y & +\,6z & +\ \ w & = 0 \\ & y & +\,3z & & = 0 \end{array}$$

que pasamos a intentar resolver, de la manera que más tarde explicaremos en detalle (el método de Gauss).

La siguiente explicación es *muy exhaustiva* y no se repetirá.

Para empezar, simplificamos la primera ecuación dividiendo por 7:

$$\begin{aligned} x \quad - \; z \quad &= 0 \\ 8x + \; y - 5z - 2w &= 0 \\ 3y - 6z - \; w &= 0 \\ y - 3z \quad &= 0 \end{aligned}$$

Sustituyamos la segunda ecuación por ella misma *menos 8 veces la primera* (de manera que "la x desaparece"):

$$\begin{aligned} x \quad - \; z \quad &= 0 \\ y + 3z - 2w &= 0 \\ 3y - 6z - \; w &= 0 \\ y - 3z \quad &= 0 \end{aligned}$$

Ahora podemos restarle a la tercera ecuación tres veces la segunda:

$$\begin{aligned} x \quad - \quad z \quad &= 0 \\ y + \; 3z - 2w &= 0 \\ -15z + 5w &= 0 \\ y - \; 3z \quad &= 0 \end{aligned}$$

Y restarle a la cuarta ecuación, la segunda:

$$\begin{aligned} x \quad - \quad z \quad &= 0 \\ y + \; 3z - 2w &= 0 \\ -15z + 5w &= 0 \\ - \; 6z + 2w &= 0 \end{aligned}$$

Finalmente, si a la cuarta ecuación se le resta la tercera multiplicada por 6/15, queda el sistema:

$$\begin{aligned} x \quad - \quad z \quad &= 0 \\ y + \; 3z - 2w &= 0 \\ -15z + 5w &= 0 \\ 0 &= 0 \end{aligned} \tag{3}$$

Llegados a este punto, podríamos seguir simplificando pero no vamos a hacerlo porque el método de Gauss ya ha terminado. Como se ve, las ecuaciones han quedado "como un triángulo" (por debajo del triángulo los coeficientes son 0).

En el sistema en cuestión aparece una ecuación "tonta", $0 = 0$ (la última). Esto significa siempre que *hay alguna ecuación redundante* en el sistema original: a partir de las ecuaciones iniciales, que son cuatro, hemos obtenido tres ecuaciones y una "tautología" (algo que no aporta información). Es decir, el sistema inicial de cuatro ecuaciones es "equivalente" a uno que solo tiene tres.

Las soluciones del sistema son fáciles de calcular, *utilizando una variable como si fuera un término independiente*. A partir de ahora, la w la vamos a entender como "un número cualquiera" (es decir, como un *parámetro*) y vamos a escribir todas las demás ecuaciones en función de ella:

Para empezar, de la tercera ecuación de (3), obtenemos que

$$z = \frac{1}{3}w$$

Como en la segunda ecuación de (3) tenemos y, z y w, podemos despejar la y en función de z y w, y por tanto, en función solo de w:

$$y = -3z + 2w = -3\frac{1}{3}w + 2w = w$$

Y, por fin, la x también la podemos escribir en función solo de w, utilizando la primera ecuación de (3):

$$x = z = \frac{1}{3}w$$

Por tanto, las soluciones del sistema del trinitrotolueno son:

$$x = \frac{1}{3}w,\ y = w,\ z = \frac{1}{3}w$$

y w es un *parámetro* (puede ser cualquier valor). Esto tiene su lógica, pues hay muchas maneras de producir la reacción, todas ellas proporcionales. Si conocemos, por ejemplo (como en la expresión de arriba) cuánta agua se va a producir en el proceso (la incógnita w), sabemos de inmediato cuánto tolueno y cuánto acido hacen falta y también cuánto TNT se va a producir en la reacción.

3. El método de Gauss

Introducimos ahora el método de Gauss en general. La notación será pesada, pero ha de tenerse en cuenta que *la notación no es más que una manera de escribir en general lo que acabamos de hacer*. No hay ninguna dificultad añadida, solo que se trata el problema en todos los casos.

3.1. El algoritmo de Gauss. Para enunciar con propiedad el algoritmo, necesitamos partir de unos datos precisos. En concreto, suponemos que tenemos una colección finita de ecuaciones lineales:

$$\begin{array}{rcl} E_1 \equiv & a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n & = b_1 \\ E_2 \equiv & a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n & = b_2 \\ \vdots & \ddots & \vdots \\ E_m \equiv & a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n & = b_m \end{array} \tag{4}$$

donde las incógnitas son $x_1, \ldots, x_n$ y los términos independientes son $b_1, \ldots, b_m$. Los números a_{ij} son *coeficientes*, que pueden ser cualesquiera (incluso 0).

Partimos de una suposición (es decir, una *hipótesis*): para cualquier j, $j = 1, \ldots, n$, alguno de los coeficientes a_{ij} es distinto de cero. Es decir, la variable x_j aparece de hecho (esto es una suposición que puede parecer tonta, pero cuando se implementa el algoritmo en un ordenador, hay que estar atento). Si no fuera así, nos habrían "colado" un sistema con una variable "que no está". Esa variable sobra (no aporta información).

El sistema de arriba tiene m ecuaciones y n incógnitas. No hacemos ninguna suposición sobre si n es mayor, menor o igual que m: es indiferente.

El algoritmo se describe así. Hace falta una variable f, que indica en qué fila estamos en cada paso y una p para saber en qué columna. Ahora lo veremos.

Inicio: Póngase $f = 1$.

Intercamio$_f$: Se busca entre todas las ecuaciones a partir de la f la que tenga un coeficiente a_{ij} distinto de cero para j mínimo. Hágase p igual a este j mínimo. Intercámbiense la ecuación f con la encontrada. Nota: En el primer caso, se tendrá ahora que $p = 1$, por la hipótesis de que la primera incógnita *aparece*.

Bucle de reducciones: Para cada $k > f$, se sustituye la ecuación E_k por la nueva ecuación:

$$E_k - \frac{a_{kp}}{a_{fp}} \cdot E_f$$

(división que se puede hacer porque debido al paso anterior, $a_{fp} \neq 0$). Como se ve, este bucle hace que los coeficientes a_{kp} “desaparezcan” en las ecuaciones por debajo de E_f.

Test de fin: ¿Es $f = m$? Si no, increméntese f en 1 y sáltese hasta el paso **Intercambio$_f$**. En otro caso, se ha terminado.

Está claro que el algoritmo termina en un número finito de pasos, siempre y cuando comencemos con un número finito de ecuaciones con un número finito de incógnitas.

Por ejemplo, si suponemos que estamos empezando con el sistema (4) y que ya es $a_{11} \neq 0$. Entonces, el bucle de reducciones transforma dicho sistema en

$$\begin{array}{rllll}
E_1 \equiv & a_{11}x_1 + & a_{12}x_2 + \cdots + a_{1n}x_n & = b_1 \\
E_2 \equiv & & a_{22}x_2 + \cdots + a_{2n}x_n & = b_2 \\
\vdots & & \ddots & \vdots \\
E_m \equiv & & a_{m2}x_2 + \cdots + a_{mn}x_n & = b_m
\end{array}$$

en cuya primera columna solo aparece un coeficiente diferente de 0, que es el primero. Por supuesto, *todos* los a_{ij} pueden ser distintos de los inciales. En este punto, si $m > n$, hemos de incrementar f y ponerlo a 2 y volver al punto **Intercambio$_f$**.

La idea para el segundo paso sería *repetir lo mismo con la segunda columna*. Pero aquí aparece el primer problema serio: puede ocurrir que *en la segunda columna* todos los coeficientes sean 0 salvo el de la primera fila. Por eso se describe el algoritmo como se describe. Hay que encontrar el el menor índice j tal que la variable x_j aparezca a partir de la segunda ecuación. Este índice es lo que llamamos p. Intercambiamos la ecuación correspondiente con la segunda, y repetimos el proceso de reducción, pero ahora desde la segunda fila hacia abajo.

Si *nos fijamos solo en los coeficientes*, lo que hemos conseguido hacer con todos ellos es obtener una expresión como sigue:

$$
\begin{array}{|c|c|c|c|c|c|c|c|c|c|c|c|}
\hline
a_{11} & a_{12} & \dots & a_{1p_2} & \dots & a_{1p_3} & \dots & a_{1p_4} & \dots & \dots & \dots & = b_1 \\
\hline
 & & & a_{2p_2} & \dots & a_{2p_3} & \dots & a_{2p_4} & \dots & \dots & \dots & = b_2 \\
\hline
 & & & & & a_{3p_3} & \dots & a_{3p_4} & \dots & \dots & \dots & = b_3 \\
\hline
 & & & & & & & a_{4p_4} & \dots & \dots & \dots & = b_4 \\
\hline
 & & & & & & & & \ddots & \vdots & \vdots & \vdots \\
\hline
 & & & & & & & & & a_{mp_m} & \dots & = b_m \\
\hline
\end{array}
\tag{5}
$$

que, por razones obvias, recibe el nombre de *forma escalonada* del sistema (4). En realidad, faltan las incógnitas, pero como todo el mundo se ha dado cuenta a estas alturas, *escribir las incógintas es más un engorro que otra cosa.* Al fin y al cabo, ya sabemos cuáles son y no hace falta repetirlas en cada paso. Esto es el comienzo del cálculo matricial.

DEFINICIÓN 3.1. Los índices $p_2, p_3, \dots p_m$ se llaman *posiciones de los pivotes* y los coeficientes (o si se quiere las incógnitas) correspondientes, se denominan *pivotes*. Cada pivote está en una *altura* (que no es más que la fila en la que está).

3.2. Algunos casos singulares. Aunque esto es mejor verlo haciendo ejercicios, hay casos singulares que conviene indicar expresamente, para la práctica posterior.

3.2.1. Ecuaciones redundantes. En toda la discusión de arriba, nada impide que en un momento ocurra que una de las ecuaciones se convierta en una que no tenga más que ceros a la izquierda y un término independiente nulo. De esta forma, al terminar todo el proceso, quedaría al final

$$
\begin{array}{rlrl}
E_1 \equiv & a_{11}x_1 + & a_{12}x_2 + \cdots + a_{1n}x_n & = b_1 \\
E_2 \equiv & & a_{22}x_2 + \cdots + a_{2n}x_n & = b_2 \\
\vdots & & \ddots & \vdots \\
E_{m-1} \equiv & & a_{m-1,2}x_2 + \cdots + a_{m-1,n}x_n & = b_m \\
E_m \equiv & & 0 & = 0
\end{array}
$$

La última igualdad ($0 = 0$) significa que entre todas las ecuaciones que teníamos al principio, sobraba alguna. Nótese que esta igualdad puede aparecer *más de una vez*. Aparecerá tantas veces como "ecuaciones sobrantes" tenía el sistema original.

3.2.2. Sistemas incompatibles. Por otro lado, también puede ocurrir que en el primer miembro de una ecuación aparezca un cero pero el término independiente correspondiente no sea nulo:

$$\begin{array}{rlll}
E_1 \equiv & a_{11}x_1 + & a_{12}x_2 + \cdots + a_{1n}x_n & = b_1 \\
E_2 \equiv & & a_{22}x_2 + \cdots + a_{2n}x_n & = b_2 \\
\vdots & & \ddots & \vdots \\
E_{m-1} \equiv & & a_{m-1,2}x_2 + \cdots + a_{m-1,n}x_n & = b_m \\
E_m \equiv & & 0 & = b_m
\end{array}$$

con $b_m \neq 0$. Esto quiere decir que el sistema que se ha planteado *no tiene sentido*. En otras palabras, el sistema no tiene solución, pues si tuviera alguna, se tendría también que $0 = b_m$, con b_m no nulo.

De este modo, el método de Gauss sirve para llegar a una contradicción si el sistema no tiene soluciones (y antes de haber realmente empezado a "resolverlo").

3.2.3. El caso general. Por lo general, los pivotes tienen como índice la altura, es decir, están puestos en diagonal ($p_2 = 2, p_3 = 3, \dots, p_m = m$), siempre y cuando $m \leq n$. Si $m > n$ entonces a partir de la fila $m+1$, seguro que el primer miembro de la ecuación es un 0.

Si hay más incógnitas que ecuaciones ($n > m$), quedará algo así:

$$\begin{array}{rl}
E_1 \equiv & a_{11}x_1 + a_{12}x_2 + \cdots + a_{1m}x_m + \cdots + a_{1n}x_n = b_1 \\
E_2 \equiv & \phantom{a_{11}x_1 +{}} a_{22}x_2 + \cdots + a_{2m}x_m + \cdots + a_{2n}x_n = b_2 \\
\vdots & \qquad\qquad \ddots \qquad\qquad\qquad\qquad \vdots \\
E_m \equiv & \qquad\qquad\qquad a_{mm}x_m + \cdots + a_{mn}x_n = b_m
\end{array}$$

Mientras que si hay más ecuaciones que incógnitas, la forma final será:

$$\begin{array}{rl}
E_1 \equiv & a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\
E_2 \equiv & \phantom{a_{11}x_1 +{}} a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \\
\vdots & \qquad\qquad \ddots \qquad\quad \vdots \\
E_m \equiv & \qquad\qquad\qquad\quad a_{mn}x_n = b_m
\end{array}$$

En este último caso, si hay igualdad, se tendrá exactamente la forma que se pone arriba, mientras que *si hay más ecuaciones que incógnitas*, siempre aparecerán ecuaciones incompatibles o ecuaciones redundantes.

4. El conjunto de soluciones

Una vez terminado el algoritmo de Gauss y transformado el sistema de ecuaciones en uno con forma de escalón, se puede ya describir el conjunto de soluciones del sistema de forma sencilla.

Para empezar, un poco de notación:

DEFINICIÓN 4.1. Un *vector* es una familia ordenada de n elementos $v_1, \dots, v_n$. Se escribirá como una lista de números entre paréntesis: $(v_1, \dots, v_n)$. $\blacksquare$

DEFINICIÓN 4.2. Dado un sistema de ecuaciones lineales E, de n incógnitas $x_1, \dots, x_n$, el *conjunto de soluciones* es el conjunto de vectores $(v_1, \dots, v_n)$ $\blacksquare$ tales que al sustituir x_i por v_i se satisfacen todas las ecuaciones de E.

Como se vio al principio, da igual resolver un sistema que otro obtenido del primero haciendo las transformaciones de Gauss. Así que el conjunto de soluciones de un sistema no cambia al hacer transformaciones *elementales por filas* (que es el nombre técnico de las trasnformaciones que se han descrito antes).

La ventaja de la forma en escalón es que permite describir el conjunto de soluciones de manera muy sencilla.

Hay que darse cuenta, antes que nada, de que a partir de la forma en escalón, unas variables se pueden despejar en función de las otras. En concreto, las variables *iniciales* de cada fila se pueden despejar en función de las demás que aparecen en esa fila. El proceso de "despejado" se realiza *de abajo arriba*, como es obvio dada la estructura del sistema *en escalón*.

4.1. Sistemas incompatibles. El caso más sencillo es aquel en que al final del proceso de Gauss aparecen una o más igualdades del tipo $0 = b_k$ con $b_k \neq 0$; como esta ecuación es absurda, pero ha sido obtenida del sistema original, resulta que el conjunto de soluciones es vacío (es decir, *no hay ninguna* solución).

DEFINICIÓN 4.3. Un sistema de ecuaciones lineales que no tiene solución se denomina *incompatible*.

4.2. Sistemas compatibles determinados. El caso siguiente en "tamaño" del conjunto de soluciones es aquel que solo tiene una. Esto solo puede ocurrir (como se verá cuando estudiemos el caso siguiente) si al final del algoritmo de Gauss, se obtiene una lista de ecuaciones en que los *pivotes* están en las posiciones $1, 2, \dots, n$ (donde n es el número de variables). Por tanto, en este caso ha de ser $m \geq n$ (ha de haber al menos n ecuaciones) y, si $m > n$ entonces al final del proceso de reducción se obtienen igualdades $0 = 0$ a partir de la posición $n + 1$.

En fin, en este caso lo que resulta al terminar el proceso de Gauss es un sistema como:

$$
\begin{array}{rlr}
E_1 \equiv & a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = & b_1 \\
E_2 \equiv & \qquad\quad a_{22}x_2 + \cdots + a_{2n}x_n = & b_2 \\
\vdots & \qquad\qquad\qquad \ddots & \vdots \\
E_n \equiv & \qquad\qquad\qquad\qquad a_{nn}x_n = & b_n \\
E_{n+1} \equiv & \qquad\qquad\qquad\qquad\quad 0 = & 0 \\
\vdots & \qquad\qquad\qquad\qquad\quad \vdots = & 0 \\
E_m \equiv & \qquad\qquad\qquad\qquad\quad 0 = & 0
\end{array}
\tag{6}
$$

con —y esto es importante en todo el proceso de Gauss— $a_{ii} \neq 0$. Fácilmente se comprueba que x_{nn} se puede despejar:

$$x_{nn} = \frac{-b_n}{a_{nn}}$$

y, a partir de este valor, se obtiene

$$x_{n-1,n-1} = \frac{b_{n-1} - a_{n-1,n}x_n}{a_{n-1,n-1}}$$

y, de manera sucesiva se obtiene *un valor único* para cada x_i. De modo que el sistema tiene como conjunto de soluciones uno formado por un solo vector.

DEFINICIÓN 4.4. Un sistema de ecuaciones lineales que tiene una solución se denomina *compatible determinado.*

4.3. Sistemas compatibles indeterminados. Por último, uno puede llegar a un sistema en forma de escalón (un sistema reducido) en el que haya algún pivote que no esté en la diagonal. Es decir, en alguno de los pasos de la reducción, la incógnita que aparece con coeficiente distinto de cero en la siguiente fila *no es la siguiente incógnita.* Esto hace que haya incógnitas *olvidadas*, en cierto modo (¿qué pasa con ellas?). Suponemos que *no aparece ninguna igualdad del tipo* $0 = b_k$ con $b_k \neq 0$, pues este caso es el incompatible.

Si uno analiza el sistema, se da cuenta de que lo que ocurre es que *esas incógnitas pueden tomar cualquier valor.* Por ejemplo, si en un sitema llegamos a una situación como

$$\begin{aligned} 7x_5 - 2x_6 + 8x_7 - 9x_8 &= -3 \\ 3x_7 - 6x_8 &= 1 \end{aligned}$$

uno observa que x_7 se puede despejar *sea cual sea el valor de* x_8. Y, una vez escrito así

$$x_7 = \frac{1 + 6x_8}{3},$$

la incógnita x_5 se puede despejar *sea cual sea el valor de* x_6:

$$x_5 = \frac{-3 + 9x_8 - 8x_7 + 2x_6}{7} = \frac{-17}{21} + \frac{2}{7}x_6 - x_8$$

(la última igualdad porque sabemos cuánto vale x_7 en función de x_8).

Esto que hemos mostrado con un ejemplo es general (pero no vamos a hacer la descripción detallada).

DEFINICIÓN 4.5. Las incógnitas que aparecen en un sistema reducido y que están entre dos pivotes se llaman *variables libres* del sistema. Las variables libres *se tratan como si fueran términos independientes* a la hora de resolver el sistema.

Es decir, una vez alcanzada la forma en escalón, las variables libres *se pasan a la derecha* y se manejan como si fueran constantes. Las otras incógnitas ahora se pueden despejar, como en los sistemas compatibles determinados y quedan todas ellas en función de las variables libres. Como se ve, *hay infinitas soluciones* (una para cada valor que se dé a las variables libres).

DEFINICIÓN 4.6. Un sistema de ecuaciones lineales con más de una solución se llama *compatible indeterminado*. El conjunto de soluciones es la familia de vectores en que las variables libres toman cualquier valor y las otras se calculan a partir de la forma en escalón del sistema.

Como es obvio, cuanto "menos diagonal" es un sistema reducido, más variables libres aparecen en él y por tanto *mayor* (en un sentido impreciso) es el conjunto de soluciones (tiene más *parámetros*).

4.4. Ejemplos, ejemplos, ejemplos. La única manera de aprender es enfrentarse a la vida real...

CAPíTULO 2

De ecuaciones a matrices

1. No necesitamos las variables

En todo el capítulo anterior hemos estado tratando con ecuaciones y sistemas de ecuaciones, escribiéndolas una y otra vez y, al final, cualquiera se ha dado cuenta de que en todas las transformaciones que hemos hecho de las ecuaciones y de los sistemas, una vez que estaban puestas las incógnitas en orden, escribirlas solo aportaba incomodidad (pues lo que interesaba realmente son los coeficientes: *ya sabemos qué variable hay que poner en cada lugar*). Esto es más claro todavía mirando la ecuación ó tabla (5): uno sabe dónde va cada incógnita, pero es irrelevante escribirlas o no.

Por eso, como es más fácil no escribirlas, aparece una manera más sencilla de tratar con los sistemas de ecuaciones: las matrices.

DEFINICIÓN 1.1. Un *vector columna* es una lista ordenada de m números. Se escribirá en columna y entre paréntesis. Una matriz es una lista ordenada de n vectores. Se escribirá entre paréntesis ordenando los vectores de izquierda a derecha.

Ejemplo 3. Los siguientes son vectores:

$$\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}, \begin{pmatrix} 2 \end{pmatrix}, \begin{pmatrix} x \\ 1 \\ 0{,}32 \\ e^{\pi} \\ 21 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \\ 3 \\ \dots \\ n \end{pmatrix}, \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix}$$

Ejemplo 4. Lo siguiente son matrices:

$$\begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 0 \\ 23 & 1 & x \\ e & \pi & \frac{3}{4} \end{pmatrix}, \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1m} \\ a_{21} & a_{22} & \dots & a_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nm} \end{pmatrix}$$

Nótese que, por definición, un vector es una matrix (claro).

Una matriz que solo tiene una fila se llama ordinariamente *vector fila*. De hecho, la manera habitual de escribir los vetores es, como todo el mundo sabe, como vectores fila. La forma en columna se utiliza sobre todo en el contexto de los sistemas de ecuaciones y en otros lugares singulares.

2. Escritura abreviada de un sistema

Vista la definición de matriz, podemos pasar a escribir los sistemas de ecuaciones de manera *abreviada*, sin repetir la escritura de las incógnitas por doquier.

Partimos de un sistema de m ecuaciones con n incógnitas:

$$\begin{array}{rl} E_1 \equiv & a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n = b_1 \\ E_2 \equiv & a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n = b_2 \\ \vdots & \qquad\qquad \ddots \qquad\qquad \vdots \\ E_m \equiv & a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n = b_m \end{array} \tag{7}$$

Este sistema se escribe *abreviadamente* o *en forma matricial* como:

$$\left(\begin{array}{cccc|c} a_{11} & a_{12} & \dots & a_{1n} & b_1 \\ a_{21} & a_{22} & \dots & a_{2n} & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{mn} & a_{m2} & \dots & a_{mn} & b_m \end{array}\right) \tag{8}$$

lo cual es inteligible a primera vista.

Definición 2.1. La matriz formada solo por los coeficientes (la que queda a la izquierda) se denomina *matriz de coeficientes*. El vector de las b_i se llama *vector de términos independientes*. La matriz entera se denomina también *matriz ampliada*.

Por ejemplo:

$$\begin{array}{rrrrl} 4x + & 2y + & 3z + & w = & 7 \\ - & y & - & 2w = & 0 \\ 2x & + & z - & w = & 22 \\ 2x - & y + & 7z - & 4w = & 1 \end{array}$$

se reescribe como sigue:

$$\left(\begin{array}{cccc|c} 1 & 2 & 3 & 1 & 7 \\ 0 & -1 & 0 & -2 & 0 \\ 2 & 0 & 1 & -1 & 22 \\ 2 & -1 & 7 & -4 & 1 \end{array}\right) \tag{9}$$

La ventaja de usar matrices para hacer las transformaciones de Gauss es evidente (hay que escribir del orden de $n \times m$ menos caracteres cada vez, por lo menos).

Ejemplo 5. Apliquemos el agoritmo de Gauss al sistema (9):

$$\left(\begin{array}{rrrr|r} 1 & 2 & 3 & 1 & 7 \\ 0 & -1 & 0 & -2 & 0 \\ 2 & 0 & 1 & -1 & 22 \\ 2 & -1 & 7 & -4 & 1 \end{array}\right) \overset{\rho_3-2\rho_1}{\longrightarrow} \left(\begin{array}{rrrr|r} 1 & 2 & 3 & 1 & 7 \\ 0 & -1 & 0 & -2 & 0 \\ 0 & -4 & -5 & 3 & 8 \\ 2 & -1 & 7 & -4 & 1 \end{array}\right) \overset{\rho_4-2\rho_1}{\longrightarrow}$$

$$\left(\begin{array}{rrrr|r} 1 & 2 & 3 & 1 & 7 \\ 0 & -1 & 0 & -2 & 0 \\ 0 & -4 & -5 & 3 & 8 \\ 0 & -5 & 1 & -6 & -13 \end{array}\right) \overset{\rho_3-4\rho_2}{\longrightarrow} \left(\begin{array}{rrrr|r} 1 & 2 & 3 & 1 & 7 \\ 0 & -1 & 0 & -2 & 0 \\ 0 & 0 & -5 & 11 & 8 \\ 0 & -5 & 1 & -6 & -13 \end{array}\right) \overset{\rho_4-5\rho_2}{\longrightarrow}$$

$$\left(\begin{array}{rrrr|r} 1 & 2 & 3 & 1 & 7 \\ 0 & -1 & 0 & -2 & 0 \\ 0 & 0 & -5 & 11 & 8 \\ 0 & 0 & 1 & 4 & -13 \end{array}\right) \overset{\rho_4+1/5\rho_3}{\longrightarrow} \left(\begin{array}{rrrr|r} 1 & 2 & 3 & 1 & 7 \\ 0 & -1 & 0 & -2 & 0 \\ 0 & 0 & -5 & 11 & 8 \\ 0 & 0 & 0 & \frac{21}{5} & \frac{-27}{5} \end{array}\right)$$

El sistema es compatible determinado.

2.1. Nociones relativas a vectores.

Definición 2.2. Una ecuación lineal $a_1x_1+\cdots+a_nx_n = d$ con incógnitas $x_1, \ldots, x_n$ *es satisfecha* por el vector columna

$$\begin{pmatrix} s_1 \\ \vdots \\ s_n \end{pmatrix}$$

si $a_1s_1 + \cdots + a_ns_n = d$.

Para describir las soluciones es conveniente (y natural, como se verá al hablar de espacios vectoriales) definir una operación suma y una operación “producto por escalar” para vectores.

Definición 2.3. Dados dos vectores columna u y v, la suma $u+v$ es la suma componente a componente:

$$u + v = \begin{pmatrix} u_1 \\ \vdots \\ u_n \end{pmatrix} + \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} = \begin{pmatrix} u_1 + v_1 \\ \vdots \\ u_n + v_n \end{pmatrix}$$

y, en general, dos matrices con el mismo número de filas y columnas *se suman componente a componente.*

Finalmente:

Definición 2.4. Dado un vector u y un número real λ (real o complejo, pero en este curso solo utilizamos números reales), el producto $\lambda \cdot u$ es el vector producto componente a componente:

$$\lambda \cdot \begin{pmatrix} u_1 \\ \vdots \\ u_n \end{pmatrix} = \begin{pmatrix} \lambda u_1 \\ \vdots \\ \lambda u_n \end{pmatrix}$$

y, en general, una matriz se multiplica por un escalar *componente a componente*.

3. Solución = particular + homogéneo

Dado un sistema de ecuaciones como (7), tras realizar la reducción de Gauss y obtener un sistema en forma de escalón, pueden aparecer variables libres (cuando los pivotes no están en la diagonal, por ejemplo). Cuando no hay (y el sistema es compatible) ya dijimos que la solución es única. Cuando hay variables libres (y, de nuevo, el sistema es compatible), hay una familia infinita de soluciones.

Ahora bien, cada solución (como se ve en los ejemplos) se expresa siempre de manera análoga a lo siguiente:

$$s = \begin{pmatrix} p_1 \\ \vdots \\ p_n \end{pmatrix} + w_1 \begin{pmatrix} e_{11} \\ \vdots \\ e_{1n} \end{pmatrix} + \cdots + w_r \begin{pmatrix} e_{r1} \\ \vdots \\ e_{rn} \end{pmatrix} \tag{10}$$

es decir, como *un vecto fijo* más una suma "arbitraria" (los w_i son los parámetros) de otros vectores, tantos como variables libres.

De hecho, hay una relación más profunda entre las soluciones de un sistema lineal de ecuaciones:

Proposición 3.1. *Dadas dos soluciones s_1 y s_2 de un sistema de ecuaciones, su diferencia es una solución del sistema homogéneo asociado al original.*

A partir del sistema original, escrito en notación matricial como

$$\left(\begin{array}{cccc|c} a_{11} & a_{12} & \dots & a_{1n} & b_1 \\ a_{21} & a_{22} & \dots & a_{2n} & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{mn} & a_{m2} & \dots & a_{mn} & b_m \end{array}\right) \tag{11}$$

se construyel el denominado *sistema homogéneo asociado*: es el mismo sistema sustituyendo las b_i por ceros:

$$\left(\begin{array}{cccc|c} a_{11} & a_{12} & \dots & a_{1n} & 0 \\ a_{21} & a_{22} & \dots & a_{2n} & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{mn} & a_{m2} & \dots & a_{mn} & 0 \end{array}\right) \tag{12}$$

Comprobar la veracidad de la proposición anterior es sencillo. Supongamos que s_1 y s_2 son dos soluciones de (11). Esto significa que, para todo $i = 1, \dots, m$ (para cada ecuación):

$$a_{i1}s_{11} + a_{i2}s_{12} + \cdots + a_{in}s_{1n} = b_i$$
$$a_{i1}s_{21} + a_{i2}s_{22} + \cdots + a_{in}s_{2n} = b_i$$

(el mismo segundo miembro en ambas). Si restamos:

$$a_{i1}(s_{11} - s_{21}) + \cdots + a_{in}(s_{1n} - s_{2n}) = 0$$

y esto ocurre para todos los i (para todas las ecuaciones). Por tanto, el vector $s_1 - s_2$ es *solución del sistema que tiene los mismos coeficientes que* (11) pero los términos independientes iguales a 0, es decir, el sistema homogéneo asociado (12).

Por otro lado, es claramente cierto (y no lo vamos a probar) que si el vector s es una solución cualquiera (lo que llamaremos en adelante una *solución particular*) del sistema (11) y el vector r es una solución del sistema homogéneo asociado (12), entonces el vector $s+r$ es otra solución del sistema inicial. Podemos por tanto enunciar el siguiente

TEOREMA 3.2. *Dado un sistema de ecuaciones lineales. Si s es una solución particular, cualquier otra solución r es la suma de s con una solución del sistema homogéneo asociado.*

Pero hay otra propiedad aún más interesante: los sistemas homogéneos tienen conjuntos de soluciones muy simples: son sumas de productos de vectores:

DEFINICIÓN 3.3. Dada una familia de vectores $u_1, \dots, u_n$, una *combinación lineal* de ellos es cualquier suma del tipo $\alpha_1 u_1 + \cdots + \alpha_n u_n$.

Hablar de combinaciones lineales es mucho más sencillo que hablar de “sumas de múltiplos”. El adejtivo lineal viene de que los vectores *no aparecen ni multiplicados ni con exponentes, etc...* (De hecho no sabemos cómo se pueden “multiplicar” vectores).

El resultado para sistemas homogéneos es el siguiente:

TEOREMA 3.4. *Dado un sistema homogénos como (12), su conjunto de soluciones es de la forma:*

$$\{\alpha_1 s_1 + \cdots + \alpha_r s_r | \alpha_i \in \mathbb{R}\}$$

donde $s_1, \dots, s_r$ son vectores.

DEMOSTRACIÓN. La prueba de este resultado es fácil. Primero de todo, cualquier sistema homogéneo tiene como solución el vector *nulo* (el vector cuyas componentes son todas 0). Esto debería estar claro. Si el sistema homogéneo es compatible determinado (solo tiene una solución), entonces tomando $r = 1$ y s_1 el vector nulo, se ha terminado.

En cualquier otro caso, basta reducir el sistema homogéneo a forma de escalón y despejar las incógnitas de los pivotes en función de las variables libres. Lo que queda como expresión de las soluciones es algo así:

$$\begin{array}{rcccccc} x_1 & = & c_{11}w_1 & + & \dots & + & c_{1r}w_r \\ x_2 & = & c_{21}w_1 & + & \dots & + & c_{2r}w_r \\ \vdots & & \vdots & & \ddots & \vdots & \\ x_k & = & c_{k1}w_1 & + & \dots & + & c_{kr}w_r \\ w_1 & = & w_1 & & & & \\ \vdots & & \vdots & & & & \\ w_r & = & w_r & & & & \end{array}$$

donde $w_1, \dots, w_r$ son las *variables libres* "cambiadas de nombre" y de orden por comodidad (está claro que, como partimos siempre de n incógnitas, ha de ser $k + r = n$).

El hecho de que *no haya términos independientes* es porque el sistema es homogéneo, así que en el proceso de reducción de Gauss siempre quedan ceros en los segundos miembros.

Lo que la igualdad anterior dice no es más que que cualquier solución s es combinación lineal de los vectores

$$s = \begin{pmatrix} s_1 \\ s_2 \\ \vdots \\ s_k \\ s_{k+1} \\ s_{k+2} \\ \vdots \\ s_n \end{pmatrix} = \begin{pmatrix} c_{11} \\ c_{21} \\ \vdots \\ c_{k1} \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \begin{pmatrix} c_{12} \\ c_{22} \\ \vdots \\ c_{k2} \\ 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \dots \begin{pmatrix} c_{1r} \\ c_{2r} \\ \vdots \\ c_{kr} \\ 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix},$$

Téngase en cuenta, repito, que para simplificar la prueba *he reordenado las ecuaciones*. Los parámetros $w_1, \dots, w_r$ aparecen por lo general "en medio del sistema", no al final. □

Nota 3.5. Atención: esto *no quiere decir* que la forma de resolver los sistemas de ecuaciones deba consistir en "encontrar una solución del sistema y luego resolver el homogéneo". Aunque esto pueda hacerse, los dos resultados anteriores nos sirven para dar una *descripción general* de las soluciones de los sistemas de ecuaciones lineales, que puede ser útil en otros contextos.

El caso más normal de sistema de ecuaciones es aquél en que hay tantas ecuaciones como incógnitas, es decir, $m = n$. Hay además un caso especial, aquel en que al hacer la reducción de Gauss, los pivotes quedan todos en la diagonal. En el caso homogéneo quedaría:

$$\left(\begin{array}{cccc|c} a_{11} & a_{12} & \dots & a_{1n} & 0 \\ 0 & a_{22} & \dots & a_{1n} & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & a_{nn} & 0 \end{array}\right)$$

en este caso, el sistema tiene una única solución (como se ve, pues x_n es único y yendo de abajo arriba se resuelve de manera única, no hay variables libres).

DEFINICIÓN 3.6. Una matriz cuadrada se llama *no singular* si el sistema homogéneo asociado a ella tiene solución única. Se llama *singular* si el sistema homogéneo asociado a ella tiene infinitas soluciones.

CAPíTULO 3

Sistemas II: Gauss-Jordan

1. Reducir todavía más

La forma de escalón que se obtiene de un sistema de ecuaciones o, como es equivalente, de la matriz asociada es útil pero no es la forma más sencilla de *terminar de resolver el sistema*, pues una vez alcanzada hay que ir despejando todas las incógnitas, tras pasar las variables libres a los segundos miembros. Existe una manera de obtener de forma más explícita las soluciones, haciendo como en el método de Gauss, operaciones solo con las ecuacione (es decir, operaciones solo con las filas de la matriz). Este es el método de Gauss-Jordan, que da lugar a la forma *reducida* de una matriz (i.e. de un sistema de ecuaciones).

En este capítulo damos por supuesto que los sistemas se escriben en forma matricial y, por tanto, hablaremos solo de matrices, vectores, filas, columnas, etc. en lugar de hablar de ecuaciones y variables.

Partimos de un sistema cualquiera

$$\left(\begin{array}{cccc|c} a_{11} & a_{12} & \dots & a_{1n} & b_1 \\ a_{21} & a_{22} & \dots & a_{2n} & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{mn} & a_{m2} & \dots & a_{mn} & b_m \end{array}\right) \tag{13}$$

que pretendemos resolver. Primero de todo lo reducimos a forma escalonada, con el método de Gauss:

$$\left(\begin{array}{cccc|c} \bar{a}_{11} & \bar{a}_{12} & \dots & \bar{a}_{1n} & \bar{b}_1 \\ 0 & \bar{a}_{22} & \dots & \bar{a}_{2n} & \bar{b}_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & \bar{a}_{mn} & \bar{b}_m \end{array}\right) \tag{14}$$

ahora, sabemos que *cada pivote* es diferente de cero. Por tanto, podemos dividir cada fila por su pivote (seguimos con las barras para no usar cientos de símbolos):

$$\left(\begin{array}{cccc|c} 1 & \bar{a}_{12} & \dots & \bar{a}_{1n} & \bar{b}_1 \\ 0 & 1 & \dots & \bar{a}_{2n} & \bar{b}_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & \bar{a}_{mn} & \bar{b}_m \end{array}\right) \tag{15}$$

(atención: estamos suponiendo que el pivote de la segunda ecuación es 2, pero podría ser cualquier otro: no perderse en el ejemplo).

Ahora se trata de operar *hacia arriba*: comenzando desde la última ecuación, se trata de *poner ceros* encima de todos los pivotes (es decir, si estamos en el pivote de la fila m, que resulta ser k, a la fila $i < m$ se le resta la fila m multiplicada por $\bar{a}_{ik}$, de manera que el nuevo $\bar{a}_{ik}$ es 0).

Un buen ejercicio (difícil) es describir el proceso de arriba con propiedad, como hicimos con el algoritmo de reducción de Gauss.

Un ejemplo es, claro, mucho más explicativo. El siguiente está copiado de [**1**] y es autoexplicativo:

$$\left(\begin{array}{cccc|c} 2 & 6 & 1 & 2 & 5 \\ 0 & 3 & 1 & 4 & 1 \\ 0 & 3 & 1 & 2 & 5 \end{array}\right) \xrightarrow{\rho_3-\rho_2} \left(\begin{array}{cccc|c} 2 & 6 & 1 & 2 & 5 \\ 0 & 3 & 1 & 4 & 1 \\ 0 & 0 & 0 & -2 & 4 \end{array}\right) \text{(ya en escalón)} \xrightarrow[\substack{\frac{1}{3}\rho_2 \\ -\frac{1}{2}\rho_3}]{\frac{1}{2}\rho_1} \dots$$

$$\dots \xrightarrow[-\rho_3+\rho_1]{\frac{4}{3}\rho_3+\rho_2} \dots \xrightarrow{-3\rho_2+\rho_1} \left(\begin{array}{cccc|c} 1 & 0 & -1/2 & 0 & -9/2 \\ 0 & 1 & 1/3 & 0 & 3 \\ 0 & 0 & 0 & 1 & -2 \end{array}\right)$$

En el ejemplo se observa que la tercera incógnita (la tercera columna) corresponde a una variable libre (será un parámetro de las soluciones). Más aun, vista la forma del sistema, ahora es trivial calcular las soluciones: basta "pasar los parámetros hacia la derecha". Donde pone λ, podría poner perfectamente x_3:

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} -\frac{9}{2} \\ 3 \\ 0 \\ -2 \end{pmatrix} + \lambda \cdot \begin{pmatrix} \frac{1}{2} \\ \frac{1}{3} \\ 1 \\ 0 \end{pmatrix}$$

Una última observación: la incógnita x_4 *no depende del parámetro*. El hecho de que aparezcan parámetros no quiere decir que todas las incógnitas vayan a depender de ellos.

Definición 1.1. Una *operación elemental* en una matriz o en un sistema de ecuaciones lineales es cualquier operación de las descritas en la página 11.

Lema 1.2. *Cualquier operación elemental es reversible.*

Es decir, si en una matriz M se hace una operación elemental y se obtiene M', existe otra operación elemental que lleva M' a M. Esto es claro: cambiar de orden o multiplicar por un número no nulo son obviamente reversibles y sustituir una fila por la suma de ella con el producto de otra por un número, también.

Lo que significa el lema previo es que si "puedo llegar de una matriz a otra por operaciones elementales, entonces puedo también llegar de la segunda a la primera por operaciones elementales".

DEFINICIÓN 1.3. Se dice que dos matrices son *equivalentes por filas* si se puede llegar de una a otra mediante operaciones elementales.

NOTA 1.4. La relación que se ha definido es una *relación de equivalencia.* No vamos a entrar en detalles ahora.

DEFINICIÓN 1.5. Se dice que una matriz M está en forma de Gauss-Jordan (o en forma escalonada reducida) si por encima de los pivotes solo hay ceros.

El resultado principal es el siguiente:

TEOREMA 1.6. *Una matriz es equivalente por filas a una y solo una matriz en forma de Gauss-Jordan.*

Es decir, si mediante operaciones elementales por filas de una matriz M se alcanza *una* forma de Gauss-Jordan, entonces *esa es la única forma de Gauss-Jordan* que se puede alcanzar a partir de M.

No vamos a probar este resultado (es mucho más sencillo de lo que parece, pero requiere un trabajo de notación exhaustivo que no debemos hacer en este curso). Lo importante es que, de alguna manera, las formas de Gauss-Jordan "describen todos los sistemas lineales que existen". Y, por tanto, basta con saber resolver estos sistemas (lo que es trivial, como se ha visto arriba) para poder resolver (haciendo operaciones elementales) todos los sistemas de ecuaciones lineales.

CAPíTULO 4

Determinantes I: el volumen

Es posible que el estudiante posea una noción de determinante aprendida en algún otro curso. Usualmente (aunque no tiene por qué ser así), la noción es —a los ojos del estudiante— bastante arbitraria y aparece como un complicado método de asignar un valor a una matriz cuadrada. Luego se aplica algo (aunque de manera ineficiente) a la discusión y posible resolución de los sistemas de ecuaciones lineales.

En estos apuntos nos acercamos de manera diversa. El determinante se presenta como la noción "general" de volumen, trasladada a dimensión arbitraria "teniendo en cuenta la orientación".

1. Área de un paralelogramo

Considérese la figura 1, en la que se explica el cálculo del área de un paralelogramo sin recurrir a análisis de las ecuaciones de rectas, rectas perpendiculares, etc.

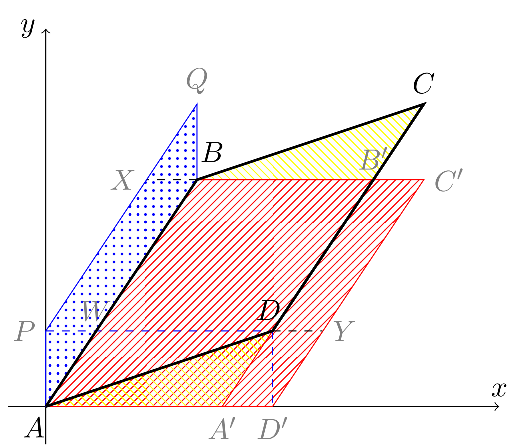

Figura 1. Área de un paralelogramo

Se parte de dos vectores, v_1 y v_2 que indican, respectivamente, los puntos D y B y se pretende calcular el área delimitada por ambos (el paralelogramo $ABCD$). Como los lados no son perpendiculares y además no están sobre los ejes coordenados, lo mejor es "proyectar" uno de ellos sobre ambos ejes. Es lo que se hace con el punto D (el vector v_1), y se obtienen los puntos D' y P respectivamente.

Si ahora se considera el vector v_2 (el lado AB) con el vector AD', se obtiene el paralelogramo $ABC'D'$ (en rojo). Si se utiliza el lado AP (por la izquierda) aparece el paralelogramo $ABQP$ (en azul).

Por un lado, el triángulo $BB'C$ es igual que $AA'D$, por semejanza y tener un lado igual (el punto A' es el corte con el eje horizontal del lado DC prolongado). Con esto se tiene que el área del paralelogramo original $ABCD$ es el área de $ABC'D'$ menos "la parte sobrante" $A'B'C'D'$.

Ahora bien, otro razonamiento de semejanza y longitud nos dice que $DD'Y$ y PWA son iguales. Lo mismo ocurre con $AD'D$ y XBQ.

Así pues, el área "sobrante" (en rojo a la derecha) es igual que el área del paralelogramo $ABQP$.

Por tanto:

$$\acute{a}rea(ABCD) = \acute{a}rea(ABC'D') - \acute{a}rea(ABQP)$$

que pasamos a escribir en coordenadas. Si $D = v_1 = (v_{11}, v_{12})$ y $B = v_2 = (v_{21}, v_{22})$, entonces

$$\acute{a}rea(ABCD) = v_{11}v_{22} - v_{21}v_{12}$$

porque, facilmente, el área de $ABC'D'$ es "base" por "altura", que es $v_{11}v_{22}$ y análogamente, el área de $ABQP$ es "base" por "altura", $v_{21}v_{12}$.

Entrando antes de lo debido en materia, resulta que

$$\acute{a}rea\begin{pmatrix} v_{11} & v_{21} \\ v_{12} & v_{22} \end{pmatrix} = \begin{vmatrix} v_{11} & v_{21} \\ v_{12} & v_{22} \end{vmatrix}$$

que es el determinante. Hemos escrito el área de "dos vectores" para indicar el área del "paralelogramo que contienen".

Otra idea importante que se obtiene: el área es una función de grado dos de las coordenadas (esto es normal, pues el área "mide" objetos "cuadráticos", no "lineales").

1.1. Propiedades del área. Cualquier definición de área que tenga relación con vectores (es decir, no el área como "valor absoluto de la región delimitada por cuatro lados", sino el área como "lo contenido entre dos *vectores*") debería comportarse de manera "natural" con las operaciones con vectores. De hecho, sabemos (esto no es más que geometría básica) que si multiplicamos *cualquiera* de los lados por un número real, el área se multiplica por ese número.

De ahora en adelante fijaremos dos vectores v_1 y v_2 planos (es decir, con dos componentes solo):

$$v_1 = \begin{pmatrix} v_{11} \\ v_{12} \end{pmatrix},\ v_2 = \begin{pmatrix} v_{21} \\ v_{22} \end{pmatrix}.$$

Lo que hemos dicho sobre la multiplicación por un escalar se escribe:

$$\acute{a}rea(\lambda v_1, v_2) = \lambda\, \acute{a}rea(v_1, v_2),\ \ \acute{a}rea(v_1, \mu v_2) = \mu\, \acute{a}rea(v_1, v_2).$$

Por otro lado, parece natural exigir que al sumar un vector w a cualquiera de las componentes, el área sea igual a la suma de las áreas correspondientes:

$$\begin{aligned}
\acute{a}rea(v_1+w,v_2) &= \acute{a}rea(v_1,v_2)+\acute{a}rea(w,v_2)\\
\acute{a}rea(v_1,v_2+w) &= \acute{a}rea(v_1,v_2)+\acute{a}rea(v_1,w),
\end{aligned}$$

esto último es razonable (la prueba, utilizando un diagrama similar pero más complejo que la figura 1, es sencilla).

También es natural pedir que el área de "dos veces el mismo vector" sea 0, pues un solo vector "no define ningún área":

$$\acute{a}rea(v_1,v_1)=0.$$

Nótese que esto, junto con la condición sobre la suma, implica necesariamente que *hay áreas negativas*, pues:

$$\acute{a}rea(v_1+v_2,v_1+v_2)=\acute{a}rea(v_1,v_1)+\acute{a}rea(v_1,v_2)+\acute{a}rea(v_2,v_1)+\acute{a}rea(v_2,v_2)$$

y, como el área de un vector consigo mismo es 0, queda:

$$0=\acute{a}rea(v_1,v_2)+\acute{a}rea(v_2,v_1),$$

de donde *el área de un par de vectores cambia de signo si se cambian de posición.*

De modo que "lo que estamos buscando" (una definición de área que sea consistente con la geometría y a la vez con la linealidad) lleva incluida la *orientación de los vectores.*

Por último, para fijar unidades, le pediremos al área que

$$\acute{a}rea\begin{pmatrix}1&0\\0&1\end{pmatrix}=1$$

(el área del cuadrado unidad es 1, por convenio). Hemos escrito el par de vectores unidad como una matriz por comodidad. De aquí que

$$\acute{a}rea\begin{pmatrix}0&1\\1&0\end{pmatrix}=-1$$

(el área de los vectores unidad en orientación levógira es -1).

Pues bien, resulta que la ecuación a que llegamos arriba

$$\acute{a}rea\begin{pmatrix}v_{11}&v_{21}\\v_{12}&v_{22}\end{pmatrix}=v_{11}v_{22}-v_{12}v_{21}$$

cumple todas las condiciones requeridas para definir un área. Además, *es la única.*

Finalmente, $\acute{a}rea(0,v)=\acute{a}rea(v,0)=0$, sea quien sea el vector v, donde 0 indica aquí el *vector* $(0,0)$.

2. Volumen de un paralelepípedo

Del mismo modo que hemos podido definir un área en el plano para un par de vectores como una función que, asigna a dos vectores un número real que dice "cuánto" área delimitan (el valor absoluto) y cuál es su orientación (el área es positiva si están orientados en sentido antihorario, negativa si en sentido horario), podemos en el espacio definir una función para 3 vectores, que vale el "volumen" que ocupa el paralelepípedo que delimitan, y que será positiva o negativa según su orientación dextrógira o levógira.

Igual que con el área, para definir de manera sensata (coherente con la realidad experimental, que *eso* es lo sensato) un volumen, hemos de pedir que, dados tres vectores v_1, v_2, v_3, un escalar λ y otro vector w, se cumpla:

(16)
$$
\begin{aligned}
volumen(v_1+w, v_2, v_3) &= volumen(v_1, v_2, v_3) + volumen(w, v_2, v_3),\\
volumen(v_1, v_2+w, v_3) &= volumen(v_1, v_2, v_3) + volumen(v_1, w, v_3),\\
volumen(v_1, v_2, v_3+w) &= volumen(v_1, v_2, v_3) + volumen(v_1, v_2, w),\\
volumen(\lambda v_1, v_2, v_3) &= volumen(v_1, \lambda v_2, v_3) = volumen(v_1, v_2, \lambda v_3) =\\
&= \lambda\, volumen(v_1, v_2, v_3).
\end{aligned}
$$

Todas estas propiedades son fáciles (aunque un poco pesadas) de probar "geométricamente", pero no vamos a hacerlo. Lo que sí vamos a hacer es, a partir de ellas, encontrar una manera sencilla de calcular el volumen: utilizar el método de Gauss de reducción de un sistema de ecuaciones. Esto es lo que tienen las buenas ideas: son sencillas y utilizables en múltiples contextos.

Antes de seguir, emplazamos al lector a probar que si dos de los vectores son iguales, entonces el volumen es cero. Y, una vez probado esto, que si dos vectores se intercambian (cada uno se pone en la posición del otro), el volumen cambia de signo.

Supongamos ahora que se nos dan tres vectores, $v_1 = (v_{11}, v_{12}, v_{13})$, $v_2 = (v_{21}, v_{22}, v_{23})$ y $v_3 = (v_{31}, v_{32}, v_{33})$ y queremos calcular el volumen delimitado por ellos, que escribiremos:

$$
volumen \begin{pmatrix} v_{11} & v_{21} & v_{31} \\ v_{12} & v_{22} & v_{32} \\ v_{13} & v_{23} & v_{33} \end{pmatrix}
$$

Por las propiedades exigidas en (16) y las inmediatamente deducibles, sabemos que

$$
volumen(v_1, v_2, v_3) = volumen(v_1, v_2 - \frac{v_{21}}{v_{11}} v_1, v_3),
$$

(suponiendo que $v_{11} \neq 0$): esto es por la linealidad y porque $volumen(v_1, v_1, v_3) = 0$. De aquí que, si ponemos $\tilde{v}_2 = v_2 - (v_{21}/v_{11})v_1)$, es:

$$
volumen \begin{pmatrix} v_{11} & 0 & v_{31} \\ v_{12} & \tilde{v}_{22} & v_{32} \\ v_{13} & \tilde{v}_{23} & v_{33} \end{pmatrix},
$$

(es decir, podemos hacer un cero "a la derecha de v_{11}". Por la misma regla de tres, podemos hacer un cero "en v_{31}" y más tarde (suponiendo $v_{22} =\neq 0$), podremos hacer un cero en la segunda fila de la tercera columna. En fin, que *haciendo lo mismo que para reducir un sistema a forma de Gauss*, llegamos a:

$$volumen \begin{pmatrix} v_{11} & v_{21} & v_{31} \\ v_{12} & v_{22} & v_{32} \\ v_{13} & v_{23} & v_{33} \end{pmatrix} = volumen \begin{pmatrix} v_{11} & 0 & 0 \\ v_{12} & \tilde{v}_{22} & 0 \\ v_{13} & \tilde{v}_{23} & \tilde{v}_{33} \end{pmatrix}$$

para ciertos $\tilde{v}_{22}, \tilde{v}_{23}$ y $\tilde{v}_{33}$.

Antes de seguir, observemos que los tres vectores que nos han quedado están dispuestos como en la figura 2.

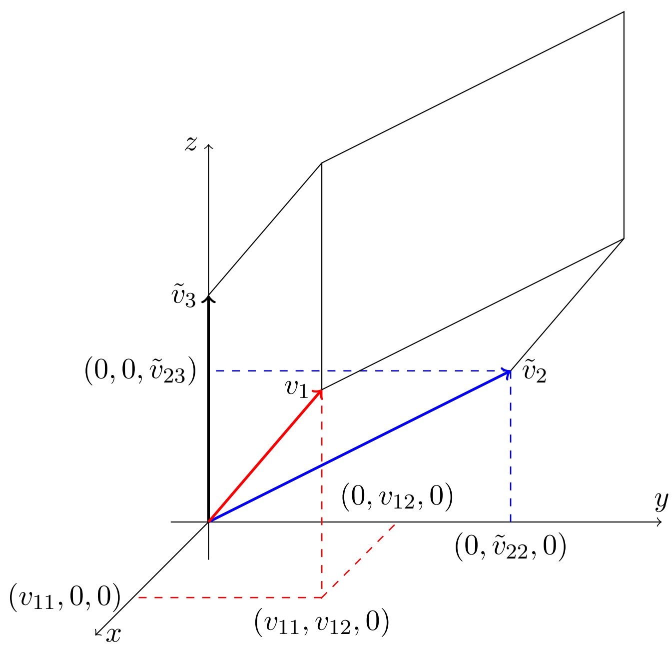

FIGURA 2. Volumen delimitado por tres vectores casi ortogonales.

El volumen del paraleleípedo es *área de la base por altura*. Tal como está dispuesto, se puede tomar la base como el paralelogramo delimitado por $\tilde{v}_2$ y $\tilde{v}_3$, cuya área es, claramente $\tilde{v}_{33}\tilde{v}_{22}$. El vector v_1, en rojo, es el único "que sale hacia afuera". La altura del paralelepípedo respecto de la base que hemos seleccionado es nada más que "la componente de v_1 perpendicular al plano yz", que es, claramente, la componente de la x, es decir, v_{11}. Así que:

$$volumen(v_1, v_2, v_3) = volumen(v_1, \tilde{v}_2, \tilde{v}_3) = v_{11}\tilde{v}_{22}\tilde{v}_{33}.$$

En fin: para calcular el volumen basta con "hacer ceros por encima de la diagonal" y calcular el producto de los elementos que quedan en la diagonal.

Esto es universal, como veremos y es realmente la forma más fácil de calcular lo que llamaremos *determinante*.

Algunas propiedades del volumen, aparte de las especificadas arriba de linealidad, etc:

1. Si un vector se puede escribir como suma de los otros dos $v_3 = \alpha v_1 + \beta v_2$, entonce $volumen(v_1, v_2, v_3) = 0$.
2. El volumen de los vectores unidad $(1, 0, 0)$, $(0, 1, 0)$ y $(0, 0, 1)$ es 1.
3. Si dos vectores de cambian de orden, el volumen cambia de signo: $volumen(v_1, v_2, v_3) = -\,volumen(v_2, v_1, v_3)$ y así.
4. Si uno de los vectores es $(0, 0, 0)$, el volumen es 0.

3. El determinante no es más que una "forma de volumen"

Podríamos seguir intentando definir volúmenes en dimensión cuatro, etc... pero pienso que han quedado claras las propiedades fundamentales que se han de especificar para definir un "volumen en dimensión n".

Para empezar, *hacen falta tantos vectores como su tamaño*. Esto no lo repetiré lo suficiente: para definir un "volumen" en un espacio de "tamaño" n (lo que llamaremos cuando lo veamos, *dimensión*), *hacen falta n vectores*. Para que en el plano haya un área, hacen falta dos vectores; para que en el espacio se delimite un volumen, hacen falta 3 vectores, ...

Lo segundo: el "volumen" consiste en *asingar un número (un escalar) a una familia de vectores*. Si trabajamos con vectores de longitud n, el "volumen" delimitado por n vectores es *un número* (que indica el "tamaño" del "espacio rodeado" por esos n vectores). Así, este "volumen" es una *aplicación* que asigna a n vectores un número. Esto se denomina *forma*.

Tercero: si a cualquiera de esos n vectores (digamos al k) se le suma otro, el volumen resultante es la suma del volumen de los n primeros con el volumen generado por los mismos *sustituyendo el vector k por el que se ha sumado*:

$$\begin{aligned} volumen(v_1, \dots, v_k + w, \dots, v_n) = {}& volumen(v_1, \dots, v_k, \dots, v_n) + \\ & + volumen(v_1, \dots, w, \dots, v_n). \end{aligned}$$

Cuarto: si cualquiera de los vectores se multiplica por un escalar λ, el volumen queda multiplicado por ese escalar:

$$volumen(v_1, \dots, \lambda v_k, \dots, v_n) = \lambda\, volumen(v_1, \dots, v_k, \dots, v_n).$$

Además, si dos vectores son iguales, el volumen es 0 (porque "no llegan a generar un espacio del tamaño necesario").

Y finalmente, se especifica (para fijar unidades) que el volumen de los vectores "unitarios" es 1.

De estas cuatro condiciones se deduce que *solo hay una manera coherente de definir este "volumen"*. A esta única aplicación se la denomina *determinante*:

Definición 3.1. El *determinante* de n vectores de longitud n es la única función multilineal alternada que asigna a los n vectores unitarios el valor 1.

Tal como se ha escrito, no significa nada porque no hemos definido lo que es una función multilineal alternada. Así que lo reescribimos.

Si V es el conjunto de vectores de longitud n (con coordenadas números reales),

DEFINICIÓN 3.2. El *determinante* es la única función:

$$|\ | : V^n \to \mathbb{R}$$

tal que

1. Dados $v_1, \dots, v_n, w \in V$, y para cualquier $k \in \{1, \dots, n\}$:
$$|v_1, \dots, v_k + w, \dots, v_n| = |v_1 \dots, v_k, \dots, v_n| + |v_1, \dots, w, \dots, v_n|$$
2. Dados $v_1, \dots, v_n \in V$, $\lambda \in \mathbb{R}$ y $k \in \{1, \dots, n\}$:
$$|v_1, \dots, \lambda v_k, \dots, v_n| = \lambda |v_1, \dots, v_k, \dots, v_n|$$
3. Dados $v_1, \dots, v_n \in V$, $\lambda \in \mathbb{R}$ y dos índices $i, j \in \{1, \dots, n\}$:
$$|v_1, \dots, v_i, \dots, v_j, \dots, v_n| = -|v_1, \dots, v_j, \dots, v_i, \dots, v_n|.$$
(Se ha cambiado el orden de los vectores i y j).
4. Si $v_1 = (1, 0, \dots, 0)$, $v_2 = (0, 1, 0, \dots, 0)$, $\dots$, $v_n = (0, 0, \dots, 1)$, entonces:
$$|v_1, \dots, v_n| = 1.$$

Aunque no lo probemos, la cuarta condición es la que hace que el determinante *esté bien definido*. Las dos primeras condiciones (que en realidad son dos condiciones *para cada componente*) se resumen diciendo que *el determinante es una función multilineal*. La tercera condición se expresa diciendo que *el determinante es una función alternada*. Así, el determinante es una función multilineal alternada, como dijimos en la definición previa.

Finalmente, puesto que al fin y al cabo una familia de n vectores de longitud n determina una matriz $n \times n$, se puede definir el determinante de una matriz $n \times n$ sin “pensar” en vectores. Dada una matriz

$$A = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \dots & a_{nn} \end{pmatrix}$$

se define el *determinante de A* como el determinante de los vectores columna:

$$|A| := |v_1, \dots, v_n|,$$

donde

$$v_1 = \begin{pmatrix} a_{11} \\ \vdots \\ a_{n1} \end{pmatrix}, \dots, v_n = \begin{pmatrix} a_{1n} \\ \vdots \\ a_{nn} \end{pmatrix}.$$

Se utilizará con asiduidad la siguiente proposición que no demostramos:

Proposición 3.3. *Si* ^{t}A *es la matriz* traspuesta *de* A*, entonces*

$$|A| = |{}^{t}A|.$$

3.1. Cálculo del determinante I: Gauss. Uno de los problemas computacionales más complejos es precisamente el cálculo del determinante de n vectores. De hecho, *siempre que sea posible un determinante ha de calcularse con un ordenador*. Cualquier otro esfuerzo se ha de dejar a los calculistas profesionales (pues pretender hacerlo bien, como se verá, es asumir demasiados riesgos).

De la multilinealidad del determinante y de su propiedad de alternancia se deduce el siguiente resultado que explica *cómo calcularemos habitualmente los determinantes*:

Teorema 3.4. *Sea* $v_1, \dots, v_n$ *una familia de* n *vectores de longitud* n *y sea* $A = (v_1, \dots, v_n)$ *la matriz formada por ellos "en columna". Sea* $\tilde{A}$ *una matriz triangular obtenida a partir de* A *haciendo ceros en filas (o en columnas) como en el método de Gauss para sistemas de ecuaciones lineales y sea* r *el número de "trasposiciones" de filas (o columnas) que se ha necesitado. Entonces, si* $\tilde{a}_1, \dots, \tilde{a}_n$ *son los elementos de la diagonal de* $\tilde{A}$*, se tiene:*

$$|A| = (-1)^r \tilde{a}_1 \cdot \dots \cdot \tilde{a}_n.$$

Por *trasposición* se entiende "intercambio", que es necesario cuando el elemento pivote no está en la fila correspondiente, sino un poco más abajo.

Ejemplo 6. Lo mejor es hacer ejemplos. Uno de orden (i.e. tamaño) 3:

$$\begin{vmatrix} 2 & 4 & 6 \\ 4 & 1 & 0 \\ 2 & 3 & 7 \end{vmatrix}$$

Vamos a "trabajar por filas", como en el método de Gauss. Para hacer ceros debajo del 2 hay que restar a la segunda fila dos veces por la primera y a la tercera, la primera:

$$\begin{vmatrix} 2 & 4 & 6 \\ 4 & 1 & 0 \\ 2 & 3 & 7 \end{vmatrix} \underset{\rho_3-\rho_1}{\overset{\rho_2-2\rho_1}{=}} \begin{vmatrix} 2 & 4 & 6 \\ 0 & -7 & -12 \\ 0 & -1 & 1 \end{vmatrix} \overset{\rho_3-\frac{1}{7}\rho_2}{=} \begin{vmatrix} 2 & 4 & 6 \\ 0 & -7 & -12 \\ 0 & 0 & 19/7 \end{vmatrix} = 2\times(-7)\times\frac{19}{7} = -38.$$

Ejemplo 7. Vamos a hacer uno en que aparezcan *trasposiciones*. La indicamos con una doble flecha:

$$\begin{vmatrix} 0 & 1 & 2 \\ 3 & -4 & 5 \\ 6 & 7 & 8 \end{vmatrix} \overset{\rho_1\leftrightarrow\rho_2}{=} - \begin{vmatrix} 3 & -4 & 5 \\ 0 & 1 & 2 \\ 6 & 7 & 8 \end{vmatrix} \overset{\rho_3-2\rho_1}{=} - \begin{vmatrix} 3 & -4 & 5 \\ 0 & 1 & 2 \\ 0 & 15 & -2 \end{vmatrix} \overset{\rho_3-15\rho_2}{=} - \begin{vmatrix} 3 & -4 & 5 \\ 0 & 1 & 2 \\ 0 & 0 & -32 \end{vmatrix} = 96$$

Finalmente, vamos a calcular un determinante 4×4:

$$\begin{vmatrix} 0 & 1 & 2 & -1 \\ 1 & 4 & 3 & 2 \\ 1 & 4 & 1 & 3 \\ 0 & 2 & 1 & 1 \end{vmatrix} \overset{\rho_1 \leftrightarrow \rho_2}{=} - \begin{vmatrix} 1 & 4 & 3 & 2 \\ 0 & 1 & 2 & -1 \\ 1 & 4 & 1 & 3 \\ 0 & 2 & 1 & 1 \end{vmatrix} \overset{\rho_3 - \rho_1}{=} - \begin{vmatrix} 1 & 4 & 3 & 2 \\ 0 & 1 & 2 & -1 \\ 0 & 0 & -2 & 1 \\ 0 & 2 & 1 & 1 \end{vmatrix} \overset{\rho_4 - 2\rho_2}{=} - \begin{vmatrix} 1 & 4 & 3 & 2 \\ 0 & 1 & 2 & -1 \\ 0 & 0 & -2 & 1 \\ 0 & 0 & -3 & 3 \end{vmatrix} =$$

$$\overset{\rho_4 - 3/2\rho_3}{=} - \begin{vmatrix} 1 & 4 & 3 & 2 \\ 0 & 1 & 2 & -1 \\ 0 & 0 & -2 & 1 \\ 0 & 0 & 0 & 3/2 \end{vmatrix} = 3$$

Esta es realmente la única manera razonable de calcular el determinante. Una salvedad son los determinantes de orden 2 y 3, cuya fórmula puede aprenderse fácilmente.

3.2. El determinante de orden 2. Este es fácil. Mostramos la fórmula (positivo al caer, negativo al subir, yendo de izquierda a derecha siempre):

$$\begin{vmatrix} v_{11} & v_{21} \\ v_{12} & v_{22} \end{vmatrix} = v_{11}v_{22} - v_{12}v_{21}.$$

Otra manera de decirlo es $ad - bc$:

$$\begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc.$$

3.3. El determinante de orden 3: Sarrus etc. Para calcular el determinante de orden 3 se suelen utilizar dos fórmulas. Una es la *regla de Sarrus*, que es justamente el desarrollo completo del determinante una vez se "calcula" su fórmula general. De todos modos

$$\text{suman: } \begin{array}{ccc} u_1 & v_1 & w_1 \\ u_2 & v_2 & w_2 \\ u_3 & v_3 & w_3 \end{array} \qquad \text{restan: } \begin{array}{ccc} u_1 & v_1 & w_1 \\ u_2 & v_2 & w_2 \\ u_3 & v_3 & w_3 \end{array}$$

Es decir: $u_1v_2w_3 + v_1w_2u_3 + w_1u_2v_3 - u_3v_2w_1 - u_2v_1w_3 - u_1v_3w_2$.

FIGURA 3. Regla de Sarrus

Esa es la manera primera que se aprende, por lo general. Hay otra, que está basada en una propiedad general de los determinantes, que consiste en que *se pueden desarrollar por filas*. Lo habitual es desarrollarlo por la primera fila:

$$\begin{vmatrix} u_1 & v_1 & w_1 \\ u_2 & v_2 & w_2 \\ u_3 & v_3 & w_3 \end{vmatrix} = u_1 \begin{vmatrix} v_2 & w_2 \\ v_3 & w_3 \end{vmatrix} - v_1 \begin{vmatrix} u_2 & w_2 \\ u_3 & w_3 \end{vmatrix} + w_1 \begin{vmatrix} u_1 & v_1 \\ u_2 & v_2 \end{vmatrix}.$$

Teniendo *mucho cuidado* con el signo menos del segundo término. Los determinantes de orden dos ya se saben calcular, así que no hace falta explicitarlos.

3.4. El caso general: desarrollo por una fila. Está claro que un vector se puede escribir como suma de n vectores, cada una conteniendo una sola de sus componentes:

$$\begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix} = \begin{pmatrix} v_1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ v_2 \\ \vdots \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ \vdots \\ v_n \end{pmatrix}.$$

Lo que hace que, utilizando la multilinealidad del determinante, se tenga la siguiente (larga pero casi evidente) igualdad:

$$\begin{vmatrix} v_{11} & v_{21} & \dots & v_{n1} \\ v_{12} & v_{22} & \dots & v_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ v_{1n} & v_{2n} & \dots & v_{nn} \end{vmatrix} = \begin{vmatrix} v_{11} & v_{21} & \dots & v_{n1} \\ 0 & v_{22} & \dots & v_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & v_{2n} & \dots & v_{nn} \end{vmatrix} + \begin{vmatrix} 0 & v_{21} & \dots & v_{n1} \\ v_{12} & v_{22} & \dots & v_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & v_{2n} & \dots & v_{nn} \end{vmatrix} + \begin{vmatrix} 0 & v_{21} & \dots & v_{n1} \\ 0 & v_{22} & \dots & v_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ v_{1n} & v_{2n} & \dots & v_{nn} \end{vmatrix}$$

que, utilizando un razonamiento por inducción *que no vamos a explicitar*, se simplifica a:

$$v_{11}\begin{vmatrix} v_{22} & \dots & v_{n2} \\ \vdots & \ddots & \vdots \\ v_{2n} & \dots & v_{nn} \end{vmatrix} - v_{12}\begin{vmatrix} v_{21} & \dots & v_{n1} \\ \vdots & \ddots & \vdots \\ v_{2n} & \dots & v_{nn} \end{vmatrix} + (-1)^{1+n}v_{1n}\begin{vmatrix} v_{21} & \dots & v_{n1} \\ \vdots & \ddots & \vdots \\ v_{2(n-1)} & \dots & v_{(n-1)(n-1)} \end{vmatrix},$$

donde hay que *tener mucho cuidado con los signos*. Lo que se ha hecho es, en cada uno de los determinantes de la ecuación anterior, tachar la columna primera y la fila con el término no nulo y multiplicar el determinante pequeño (de orden $n-1$) por el coeficiente correspondiente y *por* (-1) *elevado a la suma de las coordenadas del coeficiente*.

Eso que se ha hecho para la primera columna puede hacerse, claro está, para cualquiera (y para cualquier fila), siempre y cuando *se lleve buena cuenta de los signos*.

Definición 3.5. Dada una matriz M de tamaño $n \times n$ y dos índices i, j, se define el *determinante adjunto* del elemento (i, j) (ó simplemente *determinante adjunto* (i, j)) como el determinante de la matriz que resulta de M eliminando la fila i y la columna j. Si la matriz M se escribe $M = (a_{ij})$, el adjunto se suele escribir A_{ij}, pero esto es irrelevante.

Por ejemplo, el adjunto del elemento $(1, 2)$ en una matriz de orden tres $M = (a_{ij})$ es

$$A_{12} = \begin{vmatrix} a_{12} & a_{32} \\ a_{13} & a_{33} \end{vmatrix}$$

(el determinante de los elementos que resultan de quitar la primera fila y la segunda columna).

Teorema 3.6. *Sea $M = (a_{ij})$ una matriz $n \times n$. Fijemos $k \in \{1, \dots, n\}$. Entonces el determinante se puede calcular desarrollando por la* columna k:

$$|M| = (-1)^{k+1}a_{k1}A_{k1} + (-1)^{k+2}a_{k2}A_{k2} + \dots + (-1)^{k+n}a_{kn}A_{kn},$$

donde A_{ij} es el determinante adjunto (i,j). Lógicamente, la igualdad análoga se cumple desarrollando por la fila k:

$$|M| = (-1)^{k+1}a_{1k}A_{1k} + (-1)^{k+2}a_{2k}A_{2k} + \cdots + (-1)^{k+n}a_{nk}A_{nk}.$$

De este resultado se deduce el método más común para calcular determinantes (que, *es lo mismo que Gauss*, pero sin llevar las cosas a la diagonal):

PROPOSICIÓN 3.7 (Hacer ceros en una fila para calcular el determinante). *Sea $M = (a_{ij})$ una matriz $n\times n$. Supongamos que un elemento a_{kl} es distinto de cero (el elemento de la columna k y fila l). Entonces el determinante se calcula recursivamente de la siguiente manera* haciendo ceros la columna k:

1. *Utilizando las propiedadas básicas de los determinantes y que $a_{kl} \neq 0$, se puede conseguir (sumando* a cada fila *distinta de la l un múltiplo de esta) que en la columna k el único elemento no nulo sea a_{kl}.*
2. *Ahora $|M| = (-1)^{k+l}a_{kl}A_{kl}$, pues todos los demás elementos de dicha fila son 0 y en el teorema 3.6 la suma del miembro derecho solo tiene este término.*
3. *Para calcular el determinante A_{kl} se procede igual (salvo que sea de orden 2 ó 3, pues estos ya se conocen).*

Lo mismo que se dice para la columna k se puede hacer con la fila l, trabajando por columnas, claro.

Ejemplo 8. Como excepción, vamos a calcular ¡un determinante de orden 5! con todo detalle.

$$\begin{vmatrix} 1 & 2 & 1 & 3 & 0 \\ 2 & 1 & 1 & 4 & 1 \\ 0 & 1 & 0 & 2 & 1 \\ 3 & 1 & 1 & 2 & 0 \\ 1 & 2 & 3 & 0 & 1 \end{vmatrix}$$

Antes de operar conviene pensar. Se aprecia que tanto la fila 3 como la columna 5 tienen dos ceros. *Esto es lo que se hace antes de nada: buscar el camino más fácil.* Pero la fila 3 tiene un 2, mientras que la columna 5 tiene solo 0 y 1, así que no va a haber que multiplicar por nada para hacer ceros. Elegimos operar en la columna 5 y, por comodidad, elegimos el elemento $(5,2)$ para "centrar" la operación. Indicamos, como siempre, con ρ las filas, etc…

$$D = \begin{vmatrix} 1 & 2 & 1 & 3 & 0 \\ 2 & 1 & 1 & 4 & 1 \\ 0 & 1 & 0 & 2 & 1 \\ 3 & 1 & 1 & 2 & 0 \\ 1 & 2 & 3 & 0 & 1 \end{vmatrix} \underset{\rho_5-\rho_2}{\overset{\rho_3-\rho_2}{=}} \begin{vmatrix} 1 & 2 & 1 & 3 & 0 \\ 2 & 1 & 1 & 4 & 1 \\ -2 & 0 & -1 & -2 & 0 \\ 3 & 1 & 1 & 2 & 0 \\ -1 & 1 & 2 & -4 & 0 \end{vmatrix} = (-1)^{5+2=7}\times 1\times \begin{vmatrix} 1 & 2 & 1 & 3 \\ -2 & 0 & -1 & -2 \\ 3 & 1 & 1 & 2 \\ -1 & 1 & 2 & -4 \end{vmatrix}$$

$(-1)^7 = -1$, así que el determinante original nos ha quedado igual a

$$D = -\begin{vmatrix} 1 & 2 & 1 & 3 \\ -2 & 0 & -1 & -2 \\ 3 & 1 & 1 & 2 \\ -1 & 1 & 2 & -4 \end{vmatrix}$$

para cuyo cálculo, siguiendo el sentido común, vamos a elegir como pivote el elemento $(2,3)$ (el 1 de la segunda columna, tercera fila). Igual que antes:

$$D = -\begin{vmatrix} 1 & 2 & 1 & 3 \\ -2 & 0 & -1 & -2 \\ 3 & 1 & 1 & 2 \\ -1 & 1 & 2 & -4 \end{vmatrix} = - \underset{\rho_1 - 2\rho_3}{\overset{\rho_4 - \rho_3}{=}} \begin{vmatrix} -5 & 0 & -1 & -1 \\ -2 & 0 & -1 & -2 \\ 3 & 1 & 1 & 2 \\ -4 & 0 & 1 & -6 \end{vmatrix} = -(-1)^5 \times 1 \times \begin{vmatrix} -5 & -1 & -1 \\ -2 & -1 & -2 \\ -4 & 1 & -6 \end{vmatrix}$$

(nótese el cuidado que hay que tener con los signos). Este determinante es de orden 3, así que sabemos calcularlo:

$$D = \begin{vmatrix} -5 & -1 & -1 \\ -2 & -1 & -2 \\ -4 & 1 & -6 \end{vmatrix} = -30 + (-8) + 2 - (-4) - (-12) - 10 = -30.$$

Esto no se repetirá.

CAPíTULO 5

Espacios vectoriales

1. Introducción

En este capítulo introducimos el lenguaje abstracto para manejar los objetos de que venimos hablando durante todo el curso de una manera más cómoda. Hasta ahora hemos estado utilizando palabras poco técnicas, o palabras técnicas de modo impreciso, como ecuación, columna, vector, matriz... En este capítulo comenzamos a hablar ya de manera general, sin particularismos. Trabajar con listas de vectores de longitud 4 ó 7 es indiferente, igual que trabajar con matrices de dimensiones 2×3 ó 40×23.

Aun así, es crucial darse cuenta y tener siempre presente que el lenguaje se introduce como una mera herramienta *que posibilita hablar en general*. Pero de lo que se habla es sistematicamente de lo mismo siempre: "vectores", como "flechas en el espacio". En realidad, como veremos, la introducción del lenguaje abstracto hace posible hablar de objetos más generales (como se verá con el ejemplo de las funciones de $\mathbb{R}$ en $\mathbb{R}$). Pero lo que se busca primariamente es "generalizar la idea de flecha de fuerza aplicada en un punto", ni más ni menos. Y cuando a lo largo del curso se hable de vectores y transformaciones de vectores, etc..., lo propio será situarse en ese contexto, aunque no sea el único ni necesariamente el más utilizado.

2. Espacio Vectorial: definición y ejemplos

Todo este curso, salvo excepción, utilizaremos como conjunto de escalares, los números reales. Un *escalar* es, como su nombre indica, un valor que se puede utilizar para "multiplicar" (cambiar de *escala*): cambiar algo de tamaño para hacerlo más grande o pequeño en proporción (escala). Damos por supuesto que el alumno conoce lo suficiente de los números reales (como un conjunto, denotado $\mathbb{R}$, que posee dos operaciones, suma y producto, que cumplen las condiciones habituales).

DEFINICIÓN 2.1. Un *espacio vectorial* sobre $\mathbb{R}$ es un conjunto E dotado de una operación interna $+$ y una externa $\cdot$ que cumplen las siguientes propiedades:

Interna: Existe un elemento $\overline{0}$ tal que para todo $u \in E$, es $u + \overline{0} = \overline{0} + u = u$. Para cada $u \in E$ existe un elemento $v \in E$ tal que $u+v = v+u = \overline{0}$. Además, $(u+v)+w = u+(v+w)$ y $u+v = v+u$, para cualesquiera $u, v, w \in E$.

Externa e interna: Dados $\lambda, \mu \in \mathbb{R}$ y $u, v \in E$, se tiene que $(\lambda+\mu)\cdot u = \lambda \cdot u + \mu \cdot v$; además, $(\lambda\mu)\cdot u = (\lambda)\cdot(\mu \cdot u)$. También se ha de cumplir que $1 \cdot u = u$.

La definición parece larga y la lista de propiedades quizás demasiado extensa, pero en la práctica "en seguida" se ve si un conjunto es o no un espacio vectorial.

2.1. Ejemplos. Comenzamos con el ejemplo más complicado para (intentar) hacer ver al alumno que un lenguaje suficientemente preciso permite hablar de conceptos dispares sin demasiado esfuerzo sintáctico.

Ejemplo 9. Consideremos el conjunto F de las funciones reales de variable real. Es decir, los elementos de F son las funciones que asignan números reales a números reales. El alumno puede imaginarse este conjunto como el conjunto de todas las "gráficas de funciones", pues se supone que en cursos previos ha realizado "representaciones gráficas".

Técnicamente, se podría escribir

$$F = \{f | f : \mathbb{R} \to \mathbb{R}\}$$

(F es el conjunto de funciones reales de variable real).

Para poder tener un espacio vectorial hace falta una "operación interna" (que se suele denominar *suma*) y una "externa" (que se suele denominar *producto por un escalar*).

La operación natural para sumar las funciones de este tipo es la de sumar "punto a punto". Pero es importante notar la diferencia. Vamos a tomar dos funciones, f y g del conjunto, y vamos a definir *una tercera función* $f+g$. Como se ve, no hemos mencionado ni variables ni valores, porque *las funciones son los objetos que manejamos ahora mismo.*

Así pues, tomemos $f, g \in F$ y definamos una función "suma", que denotaremos $f+g$.

¿Qué es una función? Una *manera de asingar un número real a cada número real* (un valor y a cada $x \in \mathbb{R}$). Lo que tendremos que hacer para definir $f+g$ es explicitar *qué valor se le asigna mediante $f+g$ a cada número x*. Esto se enuncia diciendo *cuánto vale $f+g$ en x*. Por definición (y porque es lo natural)

$$\begin{aligned} (f+g) : \mathbb{R} &\to \mathbb{R} \\ x &\mapsto f(x)+g(x) \end{aligned}$$

que se lee: *la función $f+g$ consiste en asignar a cada número x el valor $f(x)+g(x)$, que es una suma de números reales, que "se sabe hacer".*. Es crucial no confundir la función (que se escribe sin paréntesis) con el valor en cada punto (que es la expresión con paréntesis). Dicho de otro modo, la *función* es la gráfica entera, mientras que $f(x)$ no es más que *el valor de la función en el punto x*, por así decir, "el punto de la gráfica $(x, f(x))$. No insistieremos suficiente en este asunto, pero sería conveniente que el alumno hiciera un breve esfuerzo mental por intentar comprender la diferencia.

También hemos de definir, para poder hablar de espacio vectorial, un "producto por un escalar". Para esto necesitamos un escalar $\lambda \in \mathbb{R}$ y un elemento de F, $f \in F$. Dados estos dos, hemos de definir una función $\lambda \cdot f$... Igual que antes, lo hacemos describiendo cómo se comporta para cada número real:

$$\begin{array}{rcl} \lambda \cdot f : \mathbb{R} & \to & \mathbb{R} \\ x & \mapsto & \lambda \cdot f(x) \end{array}$$

y, de la misma manera que antes, $\lambda \cdot f(x)$ es una multiplicación corriente de números reales.

Antes de nada, nótese que por un lado la suma $+$ que hemos definido para funciones es una operación interna (obviamente, dadas dos funciones hemos producido una función) y el "producto" $\cdot$ es una operación que asigna a un escalar y una función, otra función (así la hemos definido). En algunos casos esto puede no ser obvio.

Con estas dos operaciones, hemos de comprobar que se cumplen las condiciones de espacio vectorial. Vamos a hacerlo solo para algunas, obviamente.

Existencia de neutro para la suma: Hemos de comprobar si existe una función $\overline{0}$ que cumpla que, dada cualquiera otra función f, sea $f + \overline{0} = \overline{0} + f = f$.

¿Cuál es la función que puede cumplir estas condiciones? Ha de ser una que cumpla que en cualquier punto, su valor sumado al de otra función sea el de esta. Es decir, ha de ser en todos los puntos, $f(x) + \overline{0}(x) = \overline{0}(x) + f(x) = f(x)$. Estas dos igualdades son de números reales, y solo se cumplen si $\overline{0}(x) = 0$. Así que tendrá que ser:

$$\begin{array}{rcl} \overline{0} : \mathbb{R} & \to & \mathbb{R} \\ x & \mapsto & 0 \end{array}$$

Con esta definición ahora:

$$(\overline{0} + f)(x) \stackrel{\text{def.}}{=} \overline{0}(x) + f(x) = f(x),$$

para cualquier x. Como *dos funciones son iguales si y solo si valen lo mismo en todos los puntos*, resulta que $\overline{0} + f = f$. Del mismo modo se comprueba que $f + \overline{0} = f$.

Existencia de opuesto para la suma: Para comprobar esta propiedad se ha de encontrar, para cada $f \in F$ un elemento $g \in F$ tal que $f + g = g + f = \overline{0}$.

Como antes y puesto que dos funciones son iguales si y solo si valen lo mismo en todos los puntos, ha de ser

$$f(x) + g(x) = \overline{0}(x) = 0,$$

para todos los puntos. Si tomamos g como la función siguiente:

$$\begin{array}{rcl} \lambda \cdot g : \mathbb{R} & \to & \mathbb{R} \\ x & \mapsto & \lambda \cdot -f(x) \end{array}$$

resulta ahora que $(f + g)(x) = f(x) + (-f(x)) = 0$ y lo mismo para $(g + f)$. De este modo, resulta que $f + g = g + f = \overline{0}$, que es lo que se le exige al elemento opuesto. Obviamente, esta g se denotará de ahora en adelante como $-f$.

Propiedad distributiva del producto por escalares: Se trata de comprobar que, dados $\lambda \in \mathbb{R}$ y $f, g \in F$, se cumple que $\lambda \cdot (f+g) = \lambda \cdot f + \lambda \cdot g$. Quizás convenga detenerse un momento a comentar esta ecuación.

Como ya se ha dicho, f y g son funciones. También $f+g$ es una función, al igual que $\lambda \cdot f$ y $\lambda \cdot g$, y por fin, $\lambda \cdot (f+g)$. La igualdad que tratamos de comprobar es una *igualdad de funciones*, no de números. Estamos "multiplicando gráficas por un escalar" y "sumando gráficas": no es evidente que estas operaciones se comporten como uno desea, y es lo que estamos verificando. Igual que antes, para comprobar que dos funciones son iguales, hay que ir "punto a punto". Por un lado:

$$(\lambda \cdot (f+g))\,(x) \overset{\text{def.}}{=} \lambda\left((f+g)(x)\right) \overset{\text{def.}}{=} \lambda(f(x)+g(x)).$$

La primera igualdad es por definición del "producto por un escalar". La segunda por definición de "suma de funciones". Nótese, finalmente, que siempre que aparece una multiplicación de números reales, se pone *sin* el puntito $\cdot$.

Por otro lado:

$$(\lambda \cdot f + \lambda \cdot g)\,(x) \overset{\text{def.}}{=} (\lambda \cdot f)(x) + (\lambda \cdot g)(x) \overset{\text{def.}}{=} \lambda f(x) + \lambda g(x)$$

La primera igualdad por definición de suma de funciones y la segunda por definición de producto por escalar.

Ambos miembros valen lo mismo *para cada valor de x*, así que, por definición de función, ambas funciones (el miembro de la izquierda $\lambda \cdot (f+g)$ y el miembro de la derecha $\lambda \cdot f + \lambda \cdot g$) son iguales.

En teoría habría que comprobar muchas más condiciones, pero no vamos a hacerlo porque el modo de funcionar es claro ya.

Ejemplo 10. El espacio vectorial $\mathbb{R}^3$. Se considera el conjunto de los "vectores de longitud 3" sobre $\mathbb{R}$. Es decir, el conjunto

$$E = \left\{ \begin{pmatrix} a \\ b \\ c \end{pmatrix} \,\middle|\, a, b, c \in \mathbb{R} \right\}$$

dotado de las operaciones normales:

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} + \begin{pmatrix} u \\ v \\ w \end{pmatrix} = \begin{pmatrix} x+u \\ y+v \\ z+w \end{pmatrix}, \quad \lambda \cdot \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} \lambda x \\ \lambda y \\ \lambda z \end{pmatrix}$$

para vectores y $\lambda \in \mathbb{R}$. ¿Es este conjunto un espacio vectorial?

Claramente (y no vamos a deternos), la suma que se ha definido es una *operación interna* (es decir, *la suma de dos vectores es un vector*) y el producto por un escalar es una "operación externa" (al hacer el producto de un escalar por un vector se obtiene un vector).

Es sencillo comprobar que $\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$ es el elemento neutro para la suma. No lo vamos a detallar.

Comprobemos que existe elemento *opuesto* para la suma. Dado un vector $\begin{pmatrix} x \\ y \\ z \end{pmatrix}$, debemos encontrar otro $\begin{pmatrix} u \\ v \\ w \end{pmatrix}$ tal que

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} + \begin{pmatrix} u \\ v \\ w \end{pmatrix} = \overline{0} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

Puesto que

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} + \begin{pmatrix} u \\ v \\ w \end{pmatrix} \overset{\text{def.}}{=} \begin{pmatrix} x+u \\ y+v \\ z+w \end{pmatrix},$$

entonces tiene que ser $u = -x, v = -y, w = -z$. Y de ese modo:

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} + \begin{pmatrix} -x \\ -y \\ -z \end{pmatrix} = \begin{pmatrix} -x \\ -y \\ -z \end{pmatrix} + \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \overline{0}.$$

Comprobemos en este caso la propiedad asociativa del producto por un escalar. Tomemos $\lambda, \mu \in \mathbb{R}$ y $\begin{pmatrix} x \\ y \\ z \end{pmatrix} \in E$. Por un lado,

$$(\lambda\mu) \cdot \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} \lambda\mu x \\ \lambda\mu y \\ \lambda\mu z \end{pmatrix},$$

(en el segundo miembro las multiplicaciones son de números reales y da igual el orden). Por otro lado,

$$\lambda \cdot \left(\mu \cdot \begin{pmatrix} x \\ y \\ z \end{pmatrix} \right) \overset{\text{def.}}{=} \lambda \cdot \begin{pmatrix} \mu x \\ \mu y \\ \mu z \end{pmatrix} = \begin{pmatrix} \lambda\mu x \\ \lambda\mu y \\ \lambda\mu z \end{pmatrix},$$

que es lo mismo, *independientemente de* x, que hemos computado antes, así que se cumple la propiedad "asociativa" del producto por un escalar.

Ejemplo 11. Considérese ahora el conjunto

$$E = \{(x, y, z) | x - y + z = 0\}$$

de vectores del espacio $\mathbb{R}^3$ (escritos en fila, pero esto es irrelevante) tales que sus coordenadas cumplen la ecuación escrita. En E se consideran, claro, las *mismas* operaciones que en $\mathbb{R}^3$ que se han definido antes (sumas por componentes y producto por escalar por componentes). ¿Es este conjunto E un espacio vectorial?

La respuesta debería ser obvia, si el lector recuerda algo del conjunto de soluciones de un sistema de ecuaciones lineales.

Para empezar, hemos de comprobar que tanto la suma de vectores como el producto por escalares "se quedan dentro de E". Es decir, que dados $u, v \in E$ y $\lambda \in \mathbb{R}$, se tiene que $u + v \in E$ y $\lambda \cdot u \in E$.

Considérese la ecuación siguiente:

$$x - y + z = 0.$$

Que es una ecuación lineal (es la que define el conjunto E). Nótese que es una ecuación lineal *homogénea*. En el capítulo 2 se probó que *la suma de dos soluciones de un sistema de ecuaciones lineales homogéneo* es también solución del mismo sistema. Por ello, la suma de dos elementos de E es un elemento de E. Por otra parte, si $v = (x, y, z)$ es solución de una ecuación lineal homogénea, también lo es (claramente) $\lambda \cdot v = (\lambda x, \lambda y, \lambda z)$ (¿por qué es obvio?), así que también pertenece a E el producto λv. Por tanto, ambas operaciones (suma y producto por escalares) "se quedan en E". Esta condición, que en los ejemplos anteriores no hemos estudiado *porque las operaciones las definimos arriba para que dieran resultado en el conjunto adecuado* es la primera que hay que comprobar, como se verá en los ejemplos que siguen.

Por tanto, $+$ y $\cdot$ son operaciones que "dan valores" en E.

Ahora se podría pensar que hay que comprobar el resto de propiedades. Pero esto no es así. ¿Por qué?

En el ejemplo anterior hemos visto que dados cualesquiera $u, v \in \mathbb{R}^3$ y $\lambda, \mu \in \mathbb{R}$, se cumplen todas las propiedades necesarias para la definición de espacio vectorial (la propiedad asociativa, distributivas, etc…). Si eso ocurre para todos los elementos de $\mathbb{R}^3$, pues también ocurre cuando esos elementos se toman en $E \subset \mathbb{R}^3$.

Se podría decir que la propiedad del elemento opuesto para la suma no cumple esta condición, pero realmente no es así, pues siempre ocurre, como vamos a probar a continuación que

$$-u = (-1) \cdot u,$$

para cualquier vector de un espacio vectorial (por $-u$ se entiende justamente el elemento opuesto a u para la suma). Así que, como $(-1) \cdot u$ es un elemento del espacio E que estamos considerando si $u \in E$ (porque es producto de un elemento de E por un número), resulta que $-u \in E$ si $u \in E$, y no hace falta comprobar que el opuesto de un elemento de E está en E.

Definición 2.2. Dado $u \in E$, se denota $-u$ al elemento opuesto a u para la suma.

Proposición 2.3. *Si E es un espacio vectorial y $u \in E$ es un vector, entonces*

$$-u = (-1) \cdot u.$$

Demostración. Es un clásico del razonamiento formal. Se tiene la siguiente cadena de iguldades, bien por las propiedades de los espacios vectoriales, bien por las de los números reales. Por un lado:

$$0 \cdot u = (0+0) \cdot u = 0 \cdot u + 0 \cdot u,$$

de donde ha de ser $0 \cdot u = \overline{0}$. De aquí sale que:

$$\overline{0} = 0 \cdot u = (1 + (-1)) \cdot u = 1 \cdot u + (-1) \cdot u = u + (-1) \cdot u$$

de donde, por definición de $-u$, resulta que $-u = (-1) \cdot u$. $\square$

La demostración anterior se pone como ejemplo, no se espera ni que se entienda ni que se memorice.

Ejemplo 12. El espacio vectorial *trivial*. Dado un espacio vectorial cualquiera E, se considera el conjunto

$$Z = \left\{\overline{0}\right\}$$

(es decir, el conjunto que solo consiste del vector nulo). La suma es $\overline{0}+\overline{0} = \overline{0}$ y el producto por escalares $\lambda \cdot \overline{0} = \overline{0}$. Es *muy sencillo* comprobar que Z es un espacio vectorial. Como todos los espacios vectoriales que tienen un solo elemento son "iguales", se les denomina a todos *el espacio vectorial trivial.* (Trivial en el sentido de que es lo más simple que hay).

Ejemplo 13. Considérese el conjunto

$$E = \left\{(x, y, z) \in \mathbb{R}^3 | 2x + 3y - 7z = 7\right\}$$

(las soluciones de la ecuación lineal dada). ¿Es E un espacio vectorial?

Es sencillo comprobar que *no*: como sabemos por el capítulo 2, la diferencia entre dos soluciones de una ecuación lineal (ó de un sistema) *no homogéneo* es una solución del sistema homogéneo asociado. Así que, si $u, v \in E$ son dos vectores de E, el vector $u - v = u + (-1) \cdot v$ *no está* en E (pues no cumple la ecuación, que no es homogénea). Por tanto, o bien $(-1) \cdot v$ no está en E o bien la suma de dos elementos de E (u y $-v$) no está en E.

Ejemplo 14. El espacio $\mathbb{R}^n$ de vectores de longitud n, para cualquier número natural n, es —igual que $\mathbb{R}^3$— un espacio vectorial. La prueba es igual (pero con vectores de longitud n) que para el ejemplo de $\mathbb{R}^3$.

Ejemplo 15. El conjunto de polinomios de grado menor o igual que un número. Sea k un número natural. Sea

$$\mathcal{P}_k = \left\{a_0 + a_1 x + \cdots + a_k x^k | a_i \in \mathbb{R}\right\}$$

el conjunto de polinomios de grado menor o igual que k. Téngase en cuenta que los a_i pueden ser 0, así que son los polinomios de grado *como mucho* k, no los de grado *exactamente* k. Como se sabe, este conjunto tiene definida una suma:

$$(a_0 + a_1x + \cdots + a_kx^k) + (b_0 + b_1x + \cdots + b_kx^k) = \\ (a_0 + b_0) + (a_1 + b_1)x + \cdots + (a_k + b_k)x^k$$

y un producto por escalares:

$$\lambda \cdot (a_0 + a_1x + \cdots + a_kx^k) = \lambda a_0 + \lambda a_1x + \cdots + \lambda a_kx^k.$$

Ambas operaciones dan valores dentro de $\mathcal{P}_k$ —la suma de polinomios no aumenta el grado y multiplicar por un escalar tampoco—. Así pues, podemos intentar verificar que $\mathcal{P}_k$ con estas dos operaciones es un espacio vectorial.

El opuesto de un elemento $a_0 + \cdots + a_kx^k$ es, claramente, $-a_0 - \cdots - a_kx^k$. El resto de propiedades son consecuencia de las equivalentes para los números reales (y del uso de la x como "variable").

Este ejemplo es notable porque —y esto se deja como un ejercicio interesante al lector—, no hay diferencia alguna, *como espacio vectorial*, entre $\mathcal{P}_k$ y el conjunto de vectores de longitud $k + 1$:

$$\mathbb{R}^{k+1} = \{(a_0, \ldots, a_k) | a_i \in \mathbb{R}\}$$

(que es el del Ejemplo 14, con $k+1$ en lugar de n), pues el término independiente "funciona como" la coordenada de lugar "uno", el término de grado uno como la "segunda" coordenada, etc... Ambos espacios vectoriales, como veremos más adelante, son *isomorfos*.

3. Subespacios, conjuntos generadores, dependencia lineal

En el Ejemplo 11 se mostró un espacio vectorial de un tipo especial: un subconjunto (en el ejemplo, un plano) de un espacio vectorial más grande (en el ejemplo, $\mathbb{R}^3$). Es un caso particular de un objeto de gran interés que requiere una definición:

DEFINICIÓN 3.1. Un *subespacio vectorial* de un espacio vectorial E es un subconjunto $F \subset E$ que es espacio vectorial (utilizando las operaciones *heredadas* de E).

Con la expresión *operaciones heredadas* se quiere decir que, si v_1 y v_2 son elementos de F, entonces $v_1 + v_2$ *vale lo mismo* que si se calcula su suma como elementos de E. Esto es clave, pues por alguna razón podría ocurrir que F tuviera una cierta "operación" con el mismo nombre que E, pero que valiera distinto.

PROPOSICIÓN 3.2. *Si F es un subconjunto de un espacio vectorial E, entonces F es un subespacio vectorial si y solo si se cumple que:*

- *Para todos $u, v \in F$, su suma $u + v$ está en F.*
- *Para cualquier $\lambda \in \mathbb{R}$ y $u \in F$, se tiene que $\lambda u \in F$.*

Demostración. La prueba es la explicación que se dio en el Ejemplo 11: basta que se cumplan esas dos propiedades y el resto son consecuencia de que se cumplan en todo el espacio. □

Ejemplo 16. El eje de las x es un subespacio vectorial de $\mathbb{R}^2$:

$$E = \{(x,0)|x \in \mathbb{R}\}.$$

Esto es claro, pues coincide con:

$$E = \{(x,y)|y = 0\},$$

que es el conjunto de soluciones de una ecuación lineal homogénea, y como en el ejemplo 11, es un espacio vectorial con las operaciones heredadas.

Ejemplo 17. La generalización del ejemplo anterior es clara: *cualquier subconjunto de un espacio $\mathbb{R}^n$ definido por ecuaciones homogéneas es un subespacio vectorial de $\mathbb{R}^n$.*

Este ejemplo es, de hecho, el *más general*, pero esto lo veremos más adelante.

Ejemplo 18. Si $\mathcal{P}_3$ es el conjunto de polinomios de grado menor o igual a 3, el siguiente subconjunto ¿es un subespacio vectorial?

$$E = \{P \in \mathcal{P}_3 | P(1) = 0\}$$

Para comprobarlo habría que verificar las dos condiciones de la Proposicion 3.2:

- Tomemos dos polinomios P, Q tales que $P(1) = 0$ y $Q(1) = 0$. Claramente, $(P+Q)(1) = P(1) + Q(1) = 0 + 0 = 0$, así que la suma está en F.
- Por definición de producto por escalar, $\lambda \cdot P$ cumple que $(\lambda \cdot P)(1) = \lambda P(1) = \lambda 0 = 0$, así que también está incluido.

Por tanto, E es un subespacio vectorial de $\mathcal{P}_3$.

De hecho, la proposición 3.2 puede redactarse de otra manera:

Proposición 3.3. *Sea $S \subset E$ un subconjunto de un espacio vectorial. Se consideran en S las mismas operaciones que en E. Entonces las siguientes afirmaciones son equivalentes:*

1. *S es un subespacio vectorial de E.*
2. *Para cualesquiera $v_1, v_2 \in S$ y $c_1, c_2 \in \mathbb{R}$, se tiene que $c_1v_1 + c_2v_2 \in S$.*
3. *Para cualesquiera $v_1, \dots, v_n \in S$ y $c_1, \dots, c_n \in \mathbb{R}$, se tiene que $c_1v_1 + \dots + c_nv_n \in S$.*

Nótese que la multiplicación por escalares la hemos escrito "sin el punto" por comodidad.

Antes de continuar hemos de definir un concepto que ya usamos en los sistemas de ecuaciones para vectores columna. Se supone que E es un espacio vectorial:

Definición 3.4. Dados n vectores $v_1, \dots, v_n \in E$ y escalares $c_1, \dots, c_n \in E$, se llama *combinación lineal* de $v_1, \dots, v_n$ con los coeficientes $c_1, \dots, c_n$ al vector:

$$v = c_1 v_1 + \cdots + c_n v_n.$$

Y *una combinación lineal* de $v_1, \dots, v_n$ es un vector que se obtiene de esa manera.

Dicho de otra manera (imprecisa), una combinación lineal es cualquier suma de los vectores $v_1, \dots, v_n$ multiplicados por algún coeficiente.

Nota 3.5. En la combinación lineal, no es solo importante el vector final, sino también los coeficientes. De hecho, se dice que una combinación lineal es *no trival* si alguno de los coeficientes es no nulo.

Quizás el concepto clave de toda la teoría de espacios vectoriales es el siguiente:

Definición 3.6. Se dice que un vector $v \in E$ *depende linealmente* de un conjunto de vectores $S \subset E$ si existen vectores $v_1, \dots, v_n \in S$ tales que v es combinación lineal de estos.

Nótese que el conjunto S puede tener más de n vectores: lo que se pide es que v "se escriba" como combinación de algunos de S.

Ejemplo 19. El primer y más importante ejemplo es el trivial. El vector $\overline{0}$ *es siempre linealmente dependiente* de un conjunto de vectores no vacío:

$$\overline{0} = 0 \cdot v_1 + \cdots + 0 \cdot v_n,$$

cualesquiera que sean los v_i.

Ejemplo 20. De los polinomios $\mathcal{P}_4$ de grado menor o igual que 4, se consideran los vectores $v_1 = x - x^4$ y $v_2 = x + x^3$.

- El vector $v = 1$ *no es combinación lineal de* v_1 *y* v_2. Supongamo que sí. Entonces existen $a, b \in \mathbb{R}$ tales que

  $$1 = a(x - x^4) + b(x + x^3) = (a + b)x + bx^3 - ax^4,$$

  pero el coeficiente de grado tres de 1 es cero (así que b ha de ser 0) y el coeficiente de grado cuatro igual (así que a ha de ser cero), de donde quedaría

  $$1 = 0(x - x^4) + 0(x + x^3) = 0,$$

  que es imposible. Así que 1 no depende linealmente de v_1 y v_2.
- El vector $x^3 + x^4$ *sí* depende linealmente de v_1 y v_2:

  $$x^3 + x^4 = -1(x - x^4) + 1(x + x^3).$$

Las ideas de combinación lineal y de dependencia lineal llevan de la mano a la siguiente definición:

DEFINICIÓN 3.7. Sea E un espacio vectorial y $S \subset E$ un subconjunto. Se denomina *subespacio generado por* S, y lo denotaremos $[S]$ al conjunto de todas las combinaciones lineales de familias finitas de vectores de S:

$$[S] = \{c_1 \cdot v_1 + \cdots + c_k \cdot v_k | v_i \in S, c_i \in \mathbb{R}, k \in \mathbb{N}\}.$$

Es decir, todos los vectores que se pueden escribir como combinación lineal de vectores de S (teniendo en cuenta que las combinaciones lineales son *siempre* de una colección finita de vectores).

LEMA 3.8. *Dado un subconjunto $S \subset E$ de un espacio vectorial, el espacio que genera es el conjunto de vectores que dependen linealmente de S.*

DEMOSTRACIÓN. Es obvio, tal como se han definido dependencia lineal y espacio generado. $\square$

El nombre de *subespacio generado* tiene sentido:

LEMA 3.9. *El subespacio generado es,* de hecho*, un subespacio vectorial.*

DEMOSTRACIÓN. Sea $[S]$ el subespacio generado por S. Tomemos dos vectores $v_1, v_2 \in [S]$ y dos escalares cualesquiera c_1, c_2. Por la definición de $[S]$, existen $v_1^1, \dots, v_k^1$, $v_1^2, \dots, v_l^2$ y los escalares correspondientes tales que

$$v_1 = \sum_{i=1}^{k} c_i^1 v_i^1, \quad v_2 = \sum_{i=1}^{l} c_i^2 v_i^2.$$

Hemos de comprobar que $c_1 v_1 + c_2 v_2 \in [S]$. Pero obviamente:

$$c_1 v_1 + c_2 v_2 = c_1 \sum_{i=1}^{k} c_i^1 v_i^1 + c_2 \sum_{i=1}^{l} c_i^2 v_i^2 = \sum c_1 c_i^1 v_i^1 + \sum c_2 c_i^2 v_i^2,$$

que es, claramente, una combinación lineal de vectores de S y por tanto pertenece a $[S]$. $\square$

Ejemplo 21. Como se mostró en el teorema 3.4, las soluciones de un sistema de ecuaciones lineales homogéneo se escriben de manera paramétrica como *combinaciones lineales de vectores columna*. Es decir, las soluciones de un sistema homogéneo son *el espacio generado* por los vectores que sirven de parámetros.

NOTA 3.10. Aunque parezca obvio, no lo es tanto y se utiliza con mucha frecuencia el siguiente hecho: el subespacio generado por un conjunto que ya es un subespacio es el mismo subespacio: supongamos que $F \subset E$ es un subespacio vectorial de E. Entonces $[F] = F$. Por tanto, $[[S]] = [S]$, para cualquier subconjuto de E. Este hecho significa exactamente que $[S]$ es "lo más grande que se puede conseguir con vectores de S".

El siguiente resultado sobre dependencia lineal es también sencillo pero importante:

LEMA 3.11. *Sea $S \subset E$ un subconjunto de un espacio vectorial E y $v \in E$ un vector. Entonces*

$$[S] = [S \cup \{v\}] \text{ si y solo si } v \in [S].$$

Es decir, lo generado por un conjunto y un vector es lo mismo que lo generado solo por el conjunto si y solo si el vector ya podía ser generado por el conjunto.

DEMOSTRACIÓN. Es casi obvio. Para demostrar una equivalencia, hay que probar dos implicaciones: hacia la derecha y hacia la izquierda.

- **$\Rightarrow$:** Supongamos que $[S \cup \{v\}] = [S]$. Esto quiere decir, por ejemplo, que cualquier vector de los generados por el primer miembro está entre los del segundo. Pero el vector v es uno de ellos, así que ha de ser $v \in [S]$.
- **$\Leftarrow$:** En este caso sabemos que v está generado por S. Obviamente, $[S] \subset [S \cup \{v\}]$, así que solo hay que probar que $[S] \supset [S \cup \{v\}]$. Tomemos un vector del segundo miembro, w. Por definición, existen $v_1, \dots, v_k \in S$ y $c_1, dots, c_k, d$ tales que

$$w = c_1 v_1 + c_k v_k + dv$$

Pero sabemos, por hipótesis, que $v = c_{k+1} v_{k+1} + \cdots + c_{k+r} v_{k+r}$, para ciertos $v_{k+i} \in S$ y $c_{k+i} \in \mathbb{R}$ (pues $v \in [S]$). De donde:

$$w = c_1 v_1 + c_k v_k + d(c_{k+1} v_{k+1} + \cdots + c_{k+r} v_{k+r}) =$$
$$c_1 v_1 + \cdots + c_k v_k + dc_{k+1} v_{k+1} + \cdots + dc_{k+r} v_{k+r},$$

y por tanto $w \in [S]$ (depende linealmente de S).

$\square$

DEFINICIÓN 3.12. Un subconjunto S de un espacio vectorial se denomina *linealmente dependiente* ó *ligado* si existe algún elemento de S que es combinación lineal de los demás. Si no es linealmente dependiente, se dice *linealmente independiente* ó *libre* (sistema libre).

Ejemplo 22. Si un subconjunto contiene el vector $\overline{0}$, entonces es ligado, pues

$$\lambda \cdot \overline{0} = \overline{0},$$

para cualquier $\lambda \neq 0$. Por otro lado, si un subconjunto está formado por un solo elemento *no nulo*, entonces *es libre*, pues

$$\lambda \cdot v \neq \overline{0}$$

si $\lambda \neq 0$ (de otro modo $v = 1/\lambda\lambda \cdot v = \overline{0}$, lo cual no pasa).

Conceptualmente, un conjunto ligado es uno en el que los vectores están relacionados linealmente (claro). Otra manera (que es la usual) de comprobar si un conjunto es ligado viene dada por la siguiente proposición:

Proposición 3.13. *Sea S un subconjunto del espacio vectorial E. Es ligado si y solo si existe una combinación lineal no trivial (ver la nota 3.5) que vale cero:*

$$c_1v_1 + \cdots + c_nv_n = 0, \quad \textit{con } v_i \in S \textit{ y algún } c_i \neq 0.$$

Esta es la manera corriente de comprobar la dependencia lineal. No vamos a probar este enunciado. Un ejemplo muestra por qué:

Ejemplo 23. ¿Son los vectores siguientes linealmente dependientes?

$$\begin{pmatrix}1\\0\\2\\1\end{pmatrix}, \begin{pmatrix}0\\1\\4\\2\end{pmatrix}, \begin{pmatrix}1\\1\\3\\5\end{pmatrix}.$$

La manera de comprobarlo es ver si hay alguna combinación lineal entre ellos igual a cero con coeficientes no nulos:

$$c_1\begin{pmatrix}1\\0\\2\\1\end{pmatrix} + c_2\begin{pmatrix}0\\1\\4\\2\end{pmatrix} + c_3\begin{pmatrix}1\\1\\3\\5\end{pmatrix} = \begin{pmatrix}0\\0\\0\\0\end{pmatrix},$$

que reescrito da

$$\begin{aligned} c_1 + c_3 &= 0\\ c_2 + c_3 &= 0\\ 2c_1 + 4c_2 + 3c_3 &= 0\\ c_1 + 2c_2 + 5c_3 &= 0 \end{aligned}$$

que es un sistema de ecuaciones homogéneo. El conjunto de vectores es libre si y solo si (por definición) el sistema de ecuaciones tiene alguna solución distinta de la trivial (pues la trivial es precisamente la que *no* sirve para discriminar los sistemas ligados). Este es un primer ejemplo importante de la fuerte relación entre las nociones relativas a espacios vectoriales y los sistemas de ecuaciones lineales homogéneos.

Antes de seguir, demos una definición clave

Definición 3.14. Se dice que un conjunto de vectores S es un *sistema generador* de E, ó que *genera* E ó, simplemente, que *es un sistema generador* (si E se entiende por el contexto) si cualquier vector de E es combinación lineal de vectores de S.

La noción de dependencia lineal es fructífera porque lleva a enunciar resultados útiles:

Proposición 3.15. *Sea S un subconjunto* finito *de un espacio vectorial E. Existe un subconjunto S'* libre *de S que genera el mismo subespacio que S.*

Es decir, todo lo que se puede generar con un conjunto finito se puede generar también con una parte *más pequeña* pero “libre”. Este es el primer paso hacia la noción de dimensión, que veremos más adelante.

Demostración. La prueba requiere un poco de atención. Para empezar, si S es solo el vector $\overline{0}$, entonces genera el espacio trivial, que se puede generar con el conjunto vacío (esto no lo hemos dicho pero es un convenio). Suponemos entonces que en S hay algún vector distinto del $\overline{0}$.

Consideremos los subconjuntos de S que son sistemas libres. Por lo menos hay uno, pues como se vio en el ejemplo 22, cualquier vector aislado es libre. De entre todos ellos tomemos uno que tenga cardinal máximo (número de elementos máximo, que eso significa cardinal) y llamémoslo S'. Este S' existe porque como mucho tiene tantos elementos como S. Supongamos que $[S] \neq [S']$. Como es obvio, esto quiere decir que hay un vector $w \in S$ que no está en $[S']$. Sea $S^{\cdot} = S' \cup \{w\}$. El cardinal de $S^{\cdot}$ es uno más que el de S': por tanto $S^{\cdot}$ es ligado. Esto quiere decir que hay una combinación lineal

$$c_1v_1 + \cdots + c_kv_k + dw = 0,$$

con algún coeficiente no nulo (con $v_i \in S'$). Como S' es libre, tiene que ser $d \neq 0$. De donde:

$$w = -\frac{c_1}{d}v_1 - \cdots - \frac{c_k}{d}v_k,$$

que significa que $w \in [S']$, lo cual contradice nuestro razonamiento. Así pues, tiene que ser $[S] = [S']$. □

Como consecuencia obvia:

Corolario 1. Cualquier subconjunto de un sistema libre es libre y cualquier conjunto que contenga un sistema ligado es ligado.

Un razonamiento similar al anterior prueba que

Proposición 3.16. *Sea S un sistema libre de E y $v \in E$ un vector* que no está en S. *Se tiene que*

$$S \cup \{v\} \text{ es ligado si y solo si } v \in [S].$$

Demostración. Como antes, hay que probar la implicación en las dos direcciones.

$\Rightarrow$) Si el sistema $S \cup \{v\}$ es ligado, hay una combinación lineal

$$c_1v_1 + \cdots + c_kv_k + dv$$

con $d \neq 0$ *porque S es libre*. De aquí se deduce que $v \in [S]$, despejando y dividiendo por d.

$\Leftarrow$) Es obvia: si $v \in [S]$, entonces el conjunto $S \cup \{v\}$ es ligado por definición (hay un vector, v, que se escribe como combinación lineal de otros de S). $\square$

Una consecuencia importante es la siguiente:

Corolario 2. Sea $\mathcal{S} = \{v_1, \dots, v_k\}$ una *sucesión* de vectores (i.e. el orden es relevante, no solo su existencia). Entonces S es un sistema ligado si y solo si existe un i tal que $v_i \in [\{v_1, \dots, v_{i-1}\}]$. En otras palabras, si hay un vector de la sucesión que se escribe como combinación lineal de los anteriores.

Todo lo que hemos hecho hasta ahora no es más que el preludio para estudiar la noción más importante relativa a los espacios vectoriales: la de base, y su concepto colateral: *dimensión*.

4. Bases y dimensión

Los sistemas libres más importantes de un espacio vectorial son, claro, los que *generan* todo el espacio. Supongamos que E es un espacio vectorial

Definición 4.1. Una *base* de E es una sucesión de vectores $P = (v_1, \dots, v_k)$ que es un sistema libre y genera todo el espacio.

Son importantes no solo porque generen todo el espacio, sino porque lo hacen "sin que haya redundancias" (como en el corolario 2). Nótese que *una base no es un conjunto*: para tener una base los vectores tienen que darse *en un orden*:

Ejemplo 24. Los vectores $v = (1,0)$ y $w = (0,1)$ generan todo el espacio $\mathbb{R}^2$. Pero las bases $\mathcal{B} = (v, w)$ y $\mathcal{B}' = (v, w)$ son *distintas*. Porque el orden en que están los vectores es distinto..

Ejemplo 25. Los vectores siguientes forman una base de $\mathbb{R}^2$, distinta de la anterior:

$$\begin{pmatrix} 1 \\ 4 \end{pmatrix}, \begin{pmatrix} 7 \\ 2 \end{pmatrix}.$$

Es sencillo comprobar que generan todo el espacio y que son independientes.

La diferencia entre los anteriores es de "sencillez". La base más sencilla para $\mathbb{R}^n$ se llama *estándar*:

Definición 4.2. La base de $\mathbb{R}^n$ formada por los vectores

$$e_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, e_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \dots, e_n = \begin{pmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}$$

en ese orden se denomina *base estándar* de $\mathbb{R}^n$. Los vectores se denotarán $e_1, \dots, e_n$ (siempre que se escriba e_i se entenderá que es el correspondiente de la base estándar del $\mathbb{R}^n$ que esté en discusión).

Nótese que es trivial comprobar que esos vectores forman una base.

En $\mathbb{R}^3$, a veces se utiliza la notación $e_1 = \vec{i}$, $e_2 = \vec{j}$, $e_3 = \vec{k}$, sobre todo en Cálculo y en Electromagnetismo.

Nótese también que el contexto indica de qué se está hablando, pues en $\mathbb{R}^3$, $e_2 = (0, 1, 0)$, mientras que en $\mathbb{R}^2$ se tiene $e_2 = (0, 1)$. No son el mismo vector, pero los distinguiremos por el contexto.

El concepto de base es la abstracción de "los ejes de coordenadas". Con las coordenadas de los ejes, cada punto del espacio se escribe de manera única (esa es precisamente la utilidad de los ejes coordenados). Las bases realizan esta función en los espacios vectoriales generales:

TEOREMA 4.3. *Una sucesión de vectores* $(v_1, \dots, v_k)$ *de un espacio vectorial* E *es una base si y solo si cualquier vector de* E *se puede escribir* de manera única *como combinación lineal de* $(v_1, \dots, v_k)$.

DEMOSTRACIÓN. Como siempre, hay que probar las dos implicaciones:

$\Rightarrow$) Por definición de base, cualquier vector de E es combinación lineal de vectores del sistema. Si la combinación no fuera única, se tendrían dos igualdades como sigue, para algún vector $v \in E$:

$$v = c_1 v_1 + \cdots + c_k v_k$$
$$v = d_1 v_1 + \cdots + d_k v_k$$

con algún $c_i \neq d_i$. Restándolas, da

$$0 = (c_1 - d_1)v_1 + \cdots + (c_k - d_k)v_k,$$

con algún coeficiente no nulo. Eso implica que el sistema no era libre.

$\Leftarrow$) El hecho de que cualquier vector se escriba como combinación lineal de los del sistema quiere decir que genera todo el espacio. El hecho de que la combinación lineal sea única, por el razonamiento anterior, implica que el sistema es libre. Así que es una base. $\square$

Por este resultado se puede dar la definición siguiente:

DEFINICIÓN 4.4. Dada una base $\mathcal{B} = (v_1, \dots, v_k)$ de un espacio vectorial y un vector $v \in E$, la *representación de* v *en la base* $\mathcal{B}$ ó *las coordenadas de* v *en la base* $\mathcal{B}$ es la sucesión de escalares $(c_1, \dots, c_k)$ tal que

$$v = c_1 v_1 + \cdots + c_k v_k.$$

Lo denotaremos como:

$$\mathrm{Rep}_{\mathcal{B}}(v) = \begin{pmatrix} c_1 \\ \vdots \\ c_k \end{pmatrix}.$$

NOTA 4.5. No se puede insistir suficiente en la necesidad de *especificar la base* cuando se habla de las coordenadas de un vector. Una expresión como "las coordenadas de v" *es siempre incorrecta* si no va acompañada de una explicación "en la base…".

Ejemplo 26. Consideremos la base $\mathcal{B} = ((1,7),(2,3))$ de $\mathbb{R}^2$ y el vector del plano $(3,10)$. ¿Qué coordenadas tiene en la base $\mathcal{B}$?. Pues habrá que calcular los coeficientes c_1, c_2 tales que

$$\begin{pmatrix} 3 \\ 10 \end{pmatrix} = c_1 \begin{pmatrix} 1 \\ 7 \end{pmatrix} + c_2 \begin{pmatrix} 2 \\ 3 \end{pmatrix}$$

Es decir, resolver el sistema

$$\begin{aligned} c_1 + 2c_2 &= 3 \\ 7c_1 + 3c_2 &= 10 \end{aligned}$$

que claramente tiene como solución $c_1 = 1$ y $c_2 = 1$. Es decir,

$$\mathrm{Rep}_{\mathcal{B}}(v) = \mathrm{Rep}_{\mathcal{B}} \begin{pmatrix} 3 \\ 10 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

Ejemplo 27. Se deja como ejercicio previo comprobar que el sistema $(x+1, x^2-1, x^2+x+1)$ es una base de $\mathcal{P}_2$, que llamaremos $\mathcal{B}$. ¿qué coordenadas tiene el polinomio $2x-3$ en esa base? Se ha de calcular la sucesión (c_1, c_2, c_3) tal que

$$c_1(x+1) + c_2(x^2-1) + c_3(x^2+x+1) = 2x-3.$$

Igualando grado a grado queda el sistema

$$\begin{aligned} c_2 + c_3 &= 0 \\ c_1 \qquad + c_3 &= 2 \\ c_1 - c_2 + c_3 &= -3 \end{aligned}$$

que tiene como solución: $c_1 = 7, c_2 = 5, c_3 = -5$. Así pues

$$\mathrm{Rep}_{\mathcal{B}}(2x-3) = \begin{pmatrix} 7 \\ 5 \\ -5 \end{pmatrix}$$

que, como se observa, no tiene nada que ver con los coeficientes originales del polinomio.

Otra manera de clarificar la base (que es la que usaremos cuando haya confusión) es indicarla como subíndice después del segundo paréntesis. Así, en el ejemplo anterior se podría escribir

$$\mathrm{Rep}_{\mathcal{B}}(2x-3) = (7,5,-5)_{\mathcal{B}}.$$

Esto será clave en el futuro para evitar malentendidos.

4.1. Dimensión. El siguiente resultado (que el estudiante por lo menos ha empezado a intuir a estas alturas) es justamente la base del concepto de *dimensión*:

TEOREMA 4.6. *Si un espacio vectorial E tiene una base con un número finito de elementos k, entonces cualquier otra base de E tiene el mismo número de elementos.*

Para probarlo necesitamos un resultado obvio pero tedioso que denominaremos *lema de sustitución*:

LEMA 4.7 (Lema de sustitución). *Sea $\mathcal{B} = (v_1, \dots, v_k)$ una base de un espacio vectorial E y sea $v \in E$ un vector. Escribamos*

$$v = c_1 v_1 + \cdots + c_k v_k$$

y supongamos que $c_i \neq 0$, para cierto i. Entonces el sistema

$$\mathcal{B}' = (v_1, \dots, v_{i-1}, v, v_{i+1}, \dots, v_k)$$

es también una base de E.

Es decir, al sustituir un elemento de la base por un vector cuya coordenada correspondiente a ese vector es no nula, se obtiene una base.

DEMOSTRACIÓN. Como decimos, es tedioso más que difícil. Se trata de comprobar que $\mathcal{B}'$ es una base, así que hemos de ver que es un sistema de generadores libre. Antes de seguir, nótese que

$$v_i = \frac{1}{c_i} v - \frac{c_1}{c_i} v_1 - \cdots - \frac{c_k}{c_i} v_k \tag{17}$$

simplemente despejando.

Sistema de generadores: por la igualdad anterior sabemos que $v_i \in [\mathcal{B}']$. Como todos los demás elementos de $\mathcal{B}$ están en $\mathcal{B}'$ resulta que $[\mathcal{B}'] \supset [\mathcal{B}]$, así que es un sistema de generadores.

Libre: tomemos una combinación lineal igual a cero:

$$d_1 v_1 + \cdots + d_{i-1} v_{i-1} + d v + d_{i+1} v_{i+1} + \cdots + d_k v_k = 0.$$

Sustituyendo (17) en esta igualdad, queda una expresión de la forma

$$\left(d_1 - d\frac{c_1}{c_i}\right) v_1 + \cdots + \frac{d}{c_i} v_i + \left(d_k - d\frac{c_k}{c_i}\right) v_k = 0$$

ya sin v, solo elementos de $\mathcal{B}$. Puesto que $\mathcal{B}$ es una base, ha de ser, específicamente, $d = 0$, y por tanto $d_i = 0$ para todo i. □

Ahora tenemos las herramientas para probar el teorema de cardinalidad de las bases 4.6.

DEMOSTRACIÓN. Tomemos dos bases de E, $\mathcal{B} = (v_1, \dots, v_n)$ y $\mathcal{B}' = (w_1, \dots, w_n, w_{n+1}, \dots, w_m)$, suponiendo que $m > n$. Como solo nos interesa el número de elementos, el orden de las bases es irrelevante y las iremos reordenando como nos sea necesario.

Como $\mathcal{B}$ es base, el vector w_1 se escribe como combinación lineal de los de $\mathcal{B}$:

$$w_1 = c_{11} v_1 + \cdots + c_{1n} v_n.$$

Por comodidad, podemos suponer (reordenando $\mathcal{B}$) que $c_{11} \neq 0$. Utilizando el lema de sustitución, sabemos que el sistema siguiente

$$\mathcal{B}_1 = (w_1, v_2, \dots, v_n)$$

(sustituir v_1 por w_1 en $\mathcal{B}$) es una base. De aquí que se puede escribir

$$w_2 = \lambda_1 w_1 + c_{22} v_2 + \cdots + c_{2n} v_n.$$

Utilizando el corolario 2, ha de ser alguno de los $c_{2i} \neq 0$, pues $w_2 \notin [\{w_1\}]$. Podemos suponer que es v_2, cambiando de orden. Otra vez usando el lema de sustitución obtenemos un sistema

$$\mathcal{B}_2 = (w_1, w_2, v_3, \ldots, v_n),$$

que es una base.

Siguiendo así, llegamos hasta sustituir v_n por w_n de manera que

$$\mathcal{B}_n = (w_1, \ldots, w_n)$$

es una base. Pero esto quiere decir que w_{n+1} depende linealmente de $(w_1, \ldots, w_n)$, lo cual no puede ser por el mismo corolario 2. Así que es imposible que haya más vectores en $\mathcal{B}'$ que en $\mathcal{B}$. □

El teorema anterior permite dar la siguiente

Definición 4.8. Sea E un espacio vectorial. Si admite una base con un número finito n de elementos, se dice que E tiene *dimensión n*. Si no, se dice que E tiene dimensión infinita. Se denotará $\dim(E)$ ó $\dim E$.

Ejemplo 28. Cualquier base de $\mathbb{R}^n$ tiene n vectores.

Ejemplo 29. Una base del espacio trivial solo puede ser vacía, así que la dimensión des espacio trivial es 0.

Ejemplo 30. El espacio $\mathcal{P}_k$ de polinomios de grado menor o igual que k es de dimensión $k+1$. Por ejemplo, el siguiente sistema es una base:

$$v_0 = 1, v_1 = x, \ldots, v_k = x^{k+1}.$$

Nótese que hemos "empezado a contar" en cero. Esto es irrelevante, siempre y cuando se vaya de uno en uno y esté claro el principio y el final de los índices.

La dimensión es, precisamente, una noción relativa al tamaño:

Lema 4.9. *Sea E un espacio vectorial de dimensión n. Cualquier sistema libre S tiene como mucho n elementos.*

Demostración. O bien el sistema libre S genera todo el espacio E, en cuyo caso es ya una base y por tanto tiene exactamente n elementos por el Teorema 4.6, o bien no genera todo el espacio E. Sea $\mathcal{B}$ una base de E. Si S tuviera más de n elementos, podríamos ir sustituyendo (como en la demostración del Teorema 4.6, utilizando el Lema de sustitución 4.7) los elementos de $\mathcal{B}$ por elementos de S. Al terminar con los n elementos de $\mathcal{B}$ tendríamos una base formada a partir de elementos de S, de donde S debería generar todo el espacio, en contra de lo que acabamos de suponer. □

LEMA 4.10. *Sea E un espacio vectorial y F un subespacio. La dimensión de F es menor o igual que la de E.*

DEMOSTRACIÓN. Sea $\mathcal{B}_F$ una base de F. Se tiene que, o bien $\mathcal{B}_F$ genera todo E, con lo cual ya es una base de E, o bien no lo hace, con lo cual, por el lema 4.9, tiene cardinal menor que la dimensión de E (y, por tanto, la dimensión de F es menor que la de E). $\square$

TEOREMA 4.11. *Si F es un subespacio vectorial de dimensión finita E y $\dim F = \dim E$, entonces $F = E$.*

DEMOSTRACIÓN. Si $F \neq E$ quiere decir que existe un vector v de E tal que $v \notin F$. Sea $\mathcal{B}_F$ una base de F. Por la Proposición 3.16, como $v \notin F$, resulta que al añadir v a $\mathcal{B}_F$ se obtiene un sistema libre, con lo que $\dim F > \dim E$, lo cual es imposible por el Lema 4.10. $\square$

Ejemplo 31. El conjunto F de las funciones de $\mathbb{R}$ en $\mathbb{R}$ es un espacio vectorial de dimensión infinita. ¿Cómo se puede comprobar esto? Pues por ejemplo observando que F incluye a todos los espacios $\mathcal{P}_k$. Como cada $\mathcal{P}_k$ es de dimensión $k+1$ (ejemplo 30), no puede ser que E tenga dimensión finita.

Todos los resultados que siguen se utilizan continuamente de manera "natural", es decir, *casi sin pensar*. Nos ha costado llegar hasta aquí, pero el esfuerzo merece la pena porque ahora pensar en bases, dimensiones y sistemas libres y generadores es mucho más sencillo.

Corolario 3. Cualquier sistema libre se puede expandir a una base.

DEMOSTRACIÓN. Llamemos S a dicho sistema libre y sea $F = [S]$ el subespacio que genera. Sea k el número de elementos de S. Tómese una base $\mathcal{B}$ de E, que por el Lema 4.10, tiene cardinal mayor o igual que k. Utilizando el Lema de sustitución 4.7, se pueden reemplazar los k primeros elementos de $\mathcal{B}$, *reordenándolos si es preciso*, por los elementos de S, obteniéndose así una base compuesta de los vectores de S y eventualmente otros. $\square$

Corolario 4. De cualquier sistema de generadores se puede extraer una base.

DEMOSTRACIÓN. Se van tomando vectores linealmente independientes del conjunto hasta alcanzar la dimensión. Esta se alcanza precisamente porque el sistema es generador. $\square$

Corolario 5. En un espacio vectorial de dimensión n, un subconjunto de cardinal n es libre si y solo si es un sistema de generadores.

DEMOSTRACIÓN. $\Rightarrow$) Como el sistema es libre, podría expandirse a una base, pero como ya tiene n elementos, no puede ser que una base tenga más de n, así que *ya es una base.*

$\Leftarrow$) Si no fuera libre, se podría quitar de él un vector (uno que dependiera de otros), manteniendo la propiedad de generación. Haciendo esto hasta que se obtuviera un sistema libre, se alcanzaría una base de tamaño menor que n, lo cual es imposible. $\square$

En fin, todo se resume en que *dos bases de un espacio vectorial tienen el mismo número de elementos*. Este resultado es un Teorema profundo sobre la estructura "dimensional" de los espacios vectoriales, y es la base de todos los isomorfismos que veremos en el futuro, que se resumen en "los espacios vectoriales de dimensión finita son todos iguales que $\mathbb{R}^n$".

5. Espacios vectoriales y sistemas de ecuaciones lineales

Volvemos a tratar las ecuaciones lineales. Porque, al fin y al cabo, toda la asignatura trata de ellas. Vamos a aplicar toda la teoría que hemos estudiado a los sistemas de ecuaciones y a las matrices asociadas; obtendremos resultados sorprendentes (quizás el más sorprendente es que el rango de una matriz es el mismo que el de su traspuesta). En realidad, vamos a hablar de matrices continuamente, más que de sistemas, pues sabemos ir de unas a otras ya.

Antes de comenzar con las filas y las columnas, enunciamos un resultado importante:

LEMA 5.1. *Si la matriz de un sistema homogéneo de n incógnitas tiene k parámetros en su forma de escalón, entonces el espacio de soluciones tiene dimensión k.*

DEMOSTRACIÓN. Los parámetros son precisamente las coordenadas que "pueden tomar cualquier valor". Si hay k parámetros, digamos las variables $x_{i_1}, \dots, x_{i_k}$, entonces los vectores con coordenadas:

$$\begin{pmatrix} x_{i_1} \\ x_{i_2} \\ \vdots \\ x_{i_k} \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \dots, \begin{pmatrix} x_{i_1} \\ x_{i_2} \\ \vdots \\ x_{i_k} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}$$

(y el resto de coordenadas *como corresponda*) son obviamente independientes y obviamente generan el espacio de soluciones (que consiste en *asignar a cada parámetro un valor* y calcular el resto). Así que son una base del espacio de soluciones. $\square$

DEFINICIÓN 5.2. Sea M una matriz $n \times m$ (n filas y m columnas). El *espacio por filas* es el subespacio de $\mathbb{R}^m$ generado por los vectores fila. El *rango por filas* es la diemnsión del espacio generado por las filas, que es el mismo que el número máximo de filas independientes.

Ejemplo 32. Tómese

$$\begin{pmatrix} 1 & 2 & -1 \\ 3 & 6 & -3 \end{pmatrix},$$

El espacio generado por las filas es el generado por los vectores $(1, 2, -1)$ y $(3, 6, -3)$, que es, claramente, un espacio de dimensión 1, con base (por ejemplo) $\mathcal{B} = ((1, 2, 1))$. *. El rango por filas es, por tanto, 1.

Poner las bases con ⟨ y ⟩?

Lema 5.3. *Si las matrices A y B están relacionadas por una operación por filas del tipo*

$$A \xrightarrow{\rho_i \leftrightarrow \rho_j} B,\ A \xrightarrow{k\rho_i} B,\ A \xrightarrow{\rho_i + \lambda\rho_j} B$$

entonces el rango por filas de A es igual que el rango por filas de B.

Demostración. Es evidente, pues por construcción los espacios generados por las filas de A y los generados por las filas de B son el mismo (y el rango por filas es la dimensión de este espacio). □

Lema 5.4. *Las filas* no nulas *de una matriz en forma de escalón forman un sistema libre (linealmente independiente).*

Demostración. Si no fuera así, habría una manera de "hacer ceros" toda una fila de esas. Pero una sencilla mirada a la estructura escalonada muestra que esto no es posible *. □

Esto debería estar enunciado de alguna manera en el algoritmo de Gauss?

Por supuesto, cuando se estudian matrices, las filas tienen "los mismos derechos" que las columnas:

Definición 5.5. Sea M una matriz $n \times m$ (n filas y m columnas). El *espacio por columnas* es el subespacio de $\mathbb{R}^n$ generado por los vectores columna. El *rango por columnas* es la diemnsión del espacio generado por las columnas, que es el mismo que el número máximo de columnas independientes.

Pero las matrices vistas en columnas tienen un interés especial. Considérese un sistema cualquiera

$$\begin{aligned} c_1 \qquad\quad + 2c_3 - 7c_4 &= d_1 \\ 2c_1 - 3c_2 \qquad\quad + \ c_4 &= d_2 \\ - 2c_2 \qquad\quad - \ c_4 &= d_3 \end{aligned}$$

donde d_1, d_2, d_3 son números reales, que tiene como matriz asociada

$$\begin{pmatrix} 1 & 0 & 2 & -7 & d_1 \\ 2 & -3 & 0 & 1 & d_2 \\ 0 & -2 & 0 & -1 & d_3 \end{pmatrix}$$

El sistema de ecuaciones *significa* ni más ni menos que el vector columna

$$\begin{pmatrix} d_1 \\ d_2 \\ d_3 \end{pmatrix}$$

es combinación lineal de los vectores

$$\begin{pmatrix}1\\2\\0\end{pmatrix}, \begin{pmatrix}0\\-3\\-2\end{pmatrix}, \begin{pmatrix}2\\0\\0\end{pmatrix}, \begin{pmatrix}-7\\1\\-1\end{pmatrix},$$

es decir, de las columnas de A. Esto, *exactamente esto*, es la manera adecuada de *comenzar a entender* un sistema de ecuaciones lineales. Es decir,

LEMA 5.6. *Un sistema de ecuaciones lineales tiene solución si y solo si el espacio generado por la matriz de coeficientes es el mismo que el generado por la matriz ampliada (la de coeficientes y términos independientes). Por tanto, tiene solución si y solo si el rango por columnas de la matriz de coeficientes es el mismo que el de la matriz ampliada.*

Lo que vamos a hacer ahora es probar que el rango por filas y el rango por columnas son iguales. Así, para calcular el rango por columnas basta reducir la matriz a forma de escalón y calcular el rango por filas.

Como al final todo se resume en trabajar en la forma de escalón, vamos a *irnos allí*. Para ello necesitamos probar:

LEMA 5.7. *Las operaciones por filas no modifican el rango por columnas.*

DEMOSTRACIÓN. El rango por columnas es lo mismo que el número máximo de columnas linealmente independientes. Supongamos que son k y que son las k primeras (da igual cuáles sean). Esto significa que el sistema homogéneo *formado por las k primeras columnas* (y nada más, poniendo como términos independientes 0) tiene solución única (esta es la definición de independencia lineal). Nótese que este sistema tiene *solo k* incógnitas, no es exactamente el mismo que antes.

Ahora bien, una operación por filas en el sistema original "se traslada" como la misma operación por filas en el segundo sistema (el de las k primeras columnas). Es decir, las primeras k columnas son independientes en el sistema original si y solo si lo son en el "pequeño". Pero como en el pequeño lo eran, lo siguen siendo en el original. Es decir, el rango por columnas tras una operación por filas (de las admitidas para resolver sistemas) no puede cambiar. □

De aquí se deduce de manera trivial la consecuencia clave:

Corolario 6. El rango por filas es el mismo que el rango por columnas.

DEMOSTRACIÓN. Una vez alcanzada la forma de escalón, el rango por filas es el número de filas distintas de la $(0, 0, \dots, 0)$ (evidentemente). Es fácil de comprobar (vista la estructura en escalón) que el rango por columnas es el número de pivotes que aparecen en la forma de escalón. Pero estos dos números son obviamente los mismos, pues hay pivote si y solo si hay filas por debajo que no son nulas. (Esto es más sencillo de lo que parece y no hay trampa ni cartón). □

Por esto se puede dar la siguiente:

Definición 5.8. El *rango* de una matriz es su rango por filas o su rango por columnas.

También es obvio de aquí que el espacio generado por las filas y el generado por las columnas *tienen la misma dimensión*.

Como consecuencia de todo este estudio, la discusión de los sistemas lineales se reduce a:

Teorema 5.9. *Dado un sistema de ecuaciones lineales con m incógnitas, con matriz de coeficientes A. Los siguientes enunciados son equivalentes:*

- *El rango de A es r.*
- *El espacio de soluciones del sistema homogéneo asociado es de dimensión $n - r$.*

Demostración. Si el rango de A es r, hay r pivotes. Para alcanzar el número de incógnitas hacen falta $n - r$ parámetros, pero ya vimos (lema 5.1) que el número de parámetros es exactamente la dimensión del sistema homogéneo asociado. Es decir, el número de pivotes más el número de parámetros es n. O sea: el rango más la dimensión del espacio de soluciones del sistema homogéneo asociado es n. □

Teorema 5.10. *En el caso particular en que A (la matriz de coeficientes) es $n \times n$ (hay n ecuaciones y n incógnitas), son equivalentes:*

1. *El rango de A es n.*
2. *A es no singular (no hay parámetros en la forma en escalón).*
3. *Las filas de A son un sistema libre.*
4. *Las columnas de A son un sistema libre.*
5. *El sistema tiene solución única.*

Finalmente,

Teorema 5.11 (de Rouché-Frobenius). *Un sistema de ecuaciones tiene solución si y solo si el rango de la matriz de coeficientes es igual que el rango de la matriz ampliada.*

CAPíTULO 6

Aplicaciones lineales

1. Introducción: un "ejemplo"

Las aplicaciones lineales son el modo "normal" de relacionar dos espacios vectoriales. Lo interesante de las aplicaciones lineales es precisamente que son aquellas que "trasladan" la estructura lienal (es decir, estructura con suma y con producto por escalares) de un "lugar" (espacio origen) a "otro" (espacio final o de llegada). Esta propiedad hace que sean especialmente sencillas de manejar y de entender.

Son posiblemente las aplicaciones más importantes de todas las matemáticas, así que hay que hacer un esfuerzo por entenderlas bien. En general, todo lo que no son aplicaciones lineales termina intentando aproximarse por varias de estas.

1.1. Un ejemplo. Vamos a estudiar la proyección de los vectores del espacio en un plano. Tomamos el conjunto original como $E = \mathbb{R}^3$, el espacio vectorial de vectores con tres coordenadas. Vamos a "proyectar" todos los vectores sobre un plano (igual que se hace, por ejemplo, al dibujar en *perspectiva axonométrica*). De todos modos, vamos a considerar que el espacio de llegada es también $\mathbb{R}^3$, pero como lo estamos considerando como algo "diferente", le asignamos otra letra, $F = \mathbb{R}^3$.

La aplicación de proyección necesita, como todo el mundo sabe, un *plano* Π sobre el que proyectar (la "pizarra" o el "papel") y una *dirección* $\bar{d}$ de proyección (que es *transversal* al plano). Antes de seguir necesitamos utilizar coordenadas, así que fijamos una base $\mathcal{B}_1 = \{e_1, e_2, e_3\}$ en E y otra $\mathcal{B}_2\,\{e_1, e_2, e_3\}$, en F. Para nuestro ejemplo utilizamos en abmos casos la base estándar, pero esto *no es ni mucho menos necesario.* Ahora podemos fijar el plano de proyección Π y la dirección $\bar{d}$.

Supondremos que queremos proyectar el espacio en el plano Π dado por la ecuación $\Pi \equiv x + y - z = 0$. La dirección de proyección será la dada por el vector $\bar{d} = (1, 1, -1)$[1].

Antes de seguir, vamos a hacer unos cuantos comentarios

1. Está claro que *cada vector* tiene una imagen (se trata de "trazar" la recta que pasa por el extremo del vector y cortar con el plano). Esta proyección es, ciertamente una *aplicación*, que vamos a llamar f.

[1] El lector avisado se habrá percatado de que estamos utilizando el vector *normal* al plano, pero esto es irrelevante y solo útil por la sencillez.

2. No es tan sencillo pero uno puede darse cuenta (utilizando un razonamiento del tipo de Thales) de que si se conoce la imagen de un vector v, que designamos como $f(v)$, la imagen de cualquier "múltiplo" de v corresponde a *multiplicar* $f(v)$ por el mismo coeficiente. Es decir, $f(\lambda \cdot v) = \lambda \cdot f(v)$.
3. Otro ejercicio nada complicado si se sabe algo de geometría sintética es comprobar que la imagen de la suma es la suma de las imágenes: $f(u+v) = f(u) + f(v)$.
4. Es sencillo comprobar que la imagen de esta aplicación *no es* todo el espacio, sino que solo es el plano Π (el plano sobre el que se proyecta) y que cualquier punto de este plano tiene alguna contraimagen.
5. Si se toma un vector v y se considera su imagen $f(v)$, el conjunto de los vectores que tienen por imagen la misma $f(v)$ es (claramente, tal como está definida la proyección) el conjunto de vectores de la forma $v + \lambda \bar{d}$.
6. Obsérvese que la suma de las dimensiones del plano "al que se llega" y la recta "de dirección" es 3, la dimensión de... (del espacio de partida, pero en este caso el espacio de llegada también es $\mathbb{R}^3$, así que esta propiedad no es tan clara).

Todas esas propiedades son generales y son el contenido de todo el estudio de aplicaciones lineales, además de la expresión matricial.

Las propiedades 2 y 3 son clave porque *si se conoce la imagen de una base, se conoce toda la aplicación lineal*. En concreto, obviamente, si se conoce la imagen de cada vector de $\mathcal{B}_1$, $f(e_1), f(e_2), f(e_3)$, la imagen de $v = \lambda e_1 + \mu e_2 + \gamma e_3$ es:

$$f((\lambda, \mu, \gamma)_{\mathcal{B}_1} = \lambda f(e_1) + \mu f(e_2) + \gamma f(e_3).$$

De modo que a partir de esa información podemos calcular la imagen de cualquier vector. Es lo que vamos a hacer.

La imagen de e_1 se calcula así. La recta que "pasa por e_1" y tiene dirección $\bar{d}$ es

$$r \equiv (1,0,0) + \lambda(1,1,-1)$$

que, corta al plano $Pi \equiv x + y - z = 0$ en $(2/3, -1/3, 1/3)_{\mathcal{B}_2}$ (nótese que las coordenadas en la imagen están en la base $\mathcal{B}_2$). Para e_2 la imagen (del mismo modo) es $f(e_2) = (-1/3, 2/3, 1/3)_{\mathcal{B}_2}$ y para e_3 se tiene $f(e_3) = (1/3, 1/3, 2/3)_{\mathcal{B}_2}$.

Una vez calculados estos valores, aplicando las propiedades 2 y 3 se tiene que la imagen de $v = (x, y, z)_{\mathcal{B}_1}$ se puede calcular así

$$f(v) = f((x,y,z)_{\mathcal{B}_1}) = f(x \cdot e_1 + y \cdot e_2 + z \cdot e_3) = x \cdot f(e_1) + y \cdot f(e_2) + z \cdot f(e_3)$$

que, expresado en forma de vectores *columna* queda

$$f(v) = x \cdot \begin{pmatrix} 2/3 \\ -1/3 \\ 1/3 \end{pmatrix} + y \cdot \begin{pmatrix} -1/3 \\ 2/3 \\ 1/3 \end{pmatrix} + z \cdot \begin{pmatrix} 1/3 \\ 1/3 \\ 2/3 \end{pmatrix} = \begin{pmatrix} 2/3x - 1/3y + 1/3z \\ -1/3x + 2/3y + 1/3z \\ 1/3x + 1/3y + 2/3z \end{pmatrix}$$

y, para simplificar, introducimos la *notación matricial*:

$$f(v) = \begin{pmatrix} 2/3 & -1/3 & 1/3 \\ -1/3 & 2/3 & 1/3 \\ 1/3 & 1/3 & 2/3 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

que significa exactamente lo que se ha puesto antes y se llama *producto de matrices*: se multiplica cada fila por la columa que corresponda. Como solo hay una columa, queda solo una columna como resultado. Si hubiera varias columnas, se pondría una después de otra.

Esta notación es *extraordinaria*, por no decir *genial*.

Pero nótese que para escribir eso, hemos necesitado tanto una base de E como de F. Así que, realmente, habríamos debido escribir

$$f(v)_{\mathcal{B}_2} = A_{\mathcal{B}_1,\mathcal{B}_2} v_{\mathcal{B}_1}.$$

Como se ve, la matriz *depende de las dos bases*, el vector original solo de la primera y el vector final solo de la segunda.

Nótese también (y no es casualidad, claro) que la matriz $A_{\mathcal{B}_1,\mathcal{B}_2}$ se construye *escribiendo en columnas las imágenes de los vectores de la base* $\mathcal{B}_1$:

$$\begin{pmatrix} f(e_1) & f(e_2) & f(e_3) \\ \downarrow & \downarrow & \downarrow \end{pmatrix}.$$

Esto pasa independientemente de la base. Cambiemos la de F para ver qué ocurre. Tomemos una algo más conveniente para el problema. Puesto que estamos proyectando sobre el plano $\Pi \equiv x + y - z = 0$, podemos pensar que es más cómodo utiliza $\mathcal{B}_3 = \{\bar{e}_1, \bar{e}_2, \bar{e}_3\}$, con

$$\bar{e}_1 = (1,0,1),\ \ \bar{e}_2 = (0,1,1),\ \ \bar{e}_3 = (0,0,1).$$

De este modo, (tras hacer las operaciones correspondientes de cambio de base que no vamos a detallar)

$$\begin{aligned} f(e_1)_{\mathcal{B}_3} &= (2/3, -1/3, 0)_{\mathcal{B}_3} \\ f(e_2)_{\mathcal{B}_3} &= (-1/3, 2/3, 0)_{\mathcal{B}_3} \\ f(e_3)_{\mathcal{B}_3} &= (1/3, 1/3, 0)_{\mathcal{B}_3} \end{aligned}$$

de manera que *la matriz de la aplicación lineal f en las bases $\mathcal{B}_1$ y $\mathcal{B}_3$* es:

$$A_{\mathcal{B}_1,\mathcal{B}_3} = \begin{pmatrix} 2/3 & -1/3 & 1/3 \\ -1/3 & 2/3 & 1/3 \\ 0 & 0 & 0 \end{pmatrix}.$$

Si cambiáramos ahora también la base $\mathcal{B}_1$ por otra, digamos $\mathcal{B}_4 = (\bar{e}_1, \bar{e}_2, \bar{e}_3)$, quedaría (haciendo las operaciones correspondientes, sin detallarlas):

$$A_{\mathcal{B}_4,\mathcal{B}_3} = \begin{pmatrix} 1 & 0 & 1/3 \\ 0 & 1 & 1/3 \\ 0 & 0 & 0 \end{pmatrix}.$$

Está claro, vista la construcción, que el plano imagen está generado por $f(e_1), f(e_2), f(e_3)$ (por las imágenes de los vectores de una base cualquiera). A este conjunto se le llama $\mathrm{Im}(f)$, y resulta ser un subespacio vectorial:

$$\mathrm{Im}(f) = \langle f(E) \rangle$$

es un subespacio vectorial de F. En este caso concreto, está generado por los vectores (en la base $\mathcal{B}_2$, que es la estándar):

$$\mathrm{Im}(f) = \langle (2/3, -1/3, 1/3), (-1/3, 2/3, 1/3), (1/3, 1/3, 2/3) \rangle$$

que, de hecho, *no* son una base de la imagen, pues son linealmente dependientes. Para una base de la imagen basta tomar los dos primeros. Obviamente, de aquí se deduce que el rango de $A_{\mathcal{B}_1,\mathcal{B}_2}$ es justamente la dimensión del espacio "imagen"; el rango da una idea de "cuánto llena la aplicación": en este caso particular, un plano, dimensión 2.

Por otro lado, un conjunto de especial interés es el de los vectores "que van a parar al $\bar{0}$", por un motivo que veremos en seguida. Este conjunto es, lógicamente, un subconjunto de E (del espacio origen). En nuestro ejemplo se corresponde (fácilmente) con la recta de dirección $\bar{d}$. Pero se puede calcular también en coordenadas: es el conjunto que "cae" en el $(0,0,0)_{\mathcal{B}_2}$. Es decir, son los $(x,y,z)_{\mathcal{B}_1}$ cuya imagen es $(0,0,0)_{\mathcal{B}_2}$. Escrito en forma matricial:

$$A_{\mathcal{B}_1,\mathcal{B}_2} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix},$$

donde

$$A_{\mathcal{B}_1,\mathcal{B}_2} = \begin{pmatrix} 2/3 & -1/3 & 1/3 \\ -1/3 & 2/3 & 1/3 \\ 1/3 & 1/3 & 2/3 \end{pmatrix}$$

Esta misma forma matricial ya basta para comprobar que el conjunto que estamos estudiando ($f^{-1}(\bar{0})$) es un subespacio vectorial de E, pues es el conjunto de soluciones de sistema lineal homogéneo de matriz asociada

$$\left(\begin{array}{ccc|c} 2/3 & -1/3 & 1/3 & 0 \\ -1/3 & 2/3 & 1/3 & 0 \\ 1/3 & 1/3 & 2/3 & 0 \end{array} \right).$$

Este conjunto se denomina el *núcleo* (Kernel) de la aplicación y se denota $\mathrm{Ker}(f)$.

Como se ve, dada la expresión matricial de la aplicación, se obtienen ecuaciones *paramétricas de la imagen* muy fácilmente (pues se tiene un sistema de generadores de la imagen en las columnas de la matriz) y ecuaciones *implícitas del kernel* muy fácilmente (pues la matriz de f es la matriz de coeficientes del sistema homogéneo que describe el Kernel).

Finalmente, el Kernel es justamente lo que hay que sumar a cualquier vector para calcular lo que "va a parar al mismo sitio": esto es obvio. Si se tiene el vector $v = (2,2,4)_{\mathcal{B}_2}$ y se quiere saber qué vectores van a parar a él, no hay más que resolver el sistema de matriz asociada (en las base $\mathcal{B}_1$):

$$\left(\begin{array}{ccc|c} 2/3 & -1/3 & 1/3 & 2 \\ -1/3 & 2/3 & 1/3 & 2 \\ 1/3 & 1/3 & 2/3 & 4 \end{array}\right)$$

y, como sabemos por la sección 3 del capítulo 2, cualquier solución de un sistema de ecuaciones es una solución particular más las del homogéneo (y el homogéneo corresponde justamente al Kernel de f).

2. El caso general. Definición y ejemplos

El caso general no es *esencialmente* diferente del ejemplo que acabamos de estudiar: una aplicación lineal general es una "transformación" de un espacio "en otro" *que conserva las propiedades de espacio vectorial*: es decir, que conserva la adición de vectores y el producto por escalares. Esta característica hace que sean especialmente fáciles de estudiar, pues poseen una colección de propiedades interesantes.

DEFINICIÓN 2.1. Una *aplicación lineal* de un espacio vectorial E en otro F es una aplicación

$$E \xrightarrow{f} F$$

que *conserva la linealidad*. Es decir, que para todos $u, v \in E$ y $\lambda, \mu \in \mathbb{R}$, verifica que

$$f(\lambda u + \mu v) = \lambda f(u) + \mu f(v).$$

Las aplicaciones lineales se llaman también *homomorfismos* de espacios vectoriales (*homos*: "misma", *morfé*: "forma").

Lo que se trata de expresar es que "la estructura lineal" (que viene dada por la existencia de una suma y de un producto por escalares" no se *estropea* al transformar los vectores de E en vectores de F mediante la aplicación f. Es muy sencillo comprobar el siguiente

LEMA 2.2. *Una aplicación f de E en F es lineal si y solo si cumple las dos siguientes propiedades:*

- *Para cada par $u, v \in E$, $f(u+v) = f(u) + f(v)$,*
- *Para cada $u \in E$ y $\lambda \in \mathbb{R}$, $f(\lambda u) = \lambda f(u)$.*

Las aplicaciones (como todo el mundo debería saber) pueden ser *inyectivas* (si no hay dos elementos que tengan la misma imagen), *sobreyectivas* (si cualquier elemento de F tiene una contraimagen) y, cuando se dan las dos propiedades, *biyectivas*. Si una aplicación es lineal, se usan diferentes nombres:

DEFINICIÓN 2.3. Un *epimorfismo* es una aplicación lineal sobreyectiva. Un *monomorfismo* es una aplicación lineal inyectiva. Un *isomorfismo* es una aplicación lineal inyectiva y sobreyectiva.

Un *isomorfismo* se corresponde con la idea que todos tenemos en la cabeza de que "estamos hablando de lo mismo pero con distintas palabras". Dentro de unos días veremos cómo dos espacios vectoriales que tienen la misma dimensión son isomorfos: son, a efectos prácticos, *la misma cosa*. Las matrices cuadradas 2×2 son "lo mismo" que los vectores de longitud 4 ($\mathbb{R}^4$), pues al fin y al cabo la única diferencia es cómo se escriben. Pasa lo mismo con los polinomios en una variable de grado menor o igual que 7 y $\mathbb{R}^8$.

Como en el ejemplo de la proyección, una aplicación lineal tiene asociados siempre dos *espacios vectoriales* que dan información relevante. Se llaman la *imagen* y el *núcleo*:

Definición 2.4. La *imagen* de una aplicación lineal $f : E \to F$ es el conjunto de vectores v de F para los cuales existe un vector $u \in E$ tal que $f(u) = v$. El *núcleo* de f es el conjunto de vectores u de E tales que $f(u) = \bar{0}$. Se denotan, respectivamente, $\mathrm{Im}(f)$ y $\mathrm{Ker}(f)$.

Proposición 2.5. *Dada una aplicación lineal $f : E \to F$, se tiene que $\mathrm{Im}(f)$ es un subespacio vectorial de F y que $\mathrm{Ker}(f)$ es un subespacio vectorial de E.*

Demostración. Ambos casos son muy sencillos, utilizando la Proposición 3.2.

Comenzamos por $\mathrm{Im}(f)$. Tomemos dos vectores $v_1, v_2 \in \mathrm{Im}(f)$ y $\lambda \in \mathbb{R}$. Hemos de comprobar que $v_1 + v_2 \in Im(f)$ y que $\lambda v_1 \in \mathrm{Im}(f)$. Como $v_1, v_2 \in \mathrm{Im}(f)$, existen $u_1, u_2 \in E$ tales que $f(u_1) = v_1$ y $f(u_2) = v_2$. Pero entonces

$$f(u_1 + u_2) = f(u_1) + f(u_2) = v_1 + v_2,$$

así que $v_1 + v_2 \in \mathrm{Im}(f)$ (pues el vector $u_1 + u_2$ "cae" en él por f). Por otro lado,

$$f(\lambda u_1) = \lambda f(u_1) = \lambda v_1,$$

con lo que λv_1 también está en $\mathrm{Im}(f)$. Por tanto, $\mathrm{Im}(f)$ es un subespacio vectorial de F.

El caso del núcleo es análogo. Tomemos $u_1, u_2 \in \mathrm{Ker}(f)$. Hemos de comprobar que $f(u_1 + u_2) = \bar{0}$ (es decir, que la suma está en el núcleo) y que $f(\lambda u_1) = \bar{0}$ (el producto por escalares está en el núcleo). Pero esto es obvio, utilizando las propiedades de linealidad:

$$f(u_1 + u_2) = f(u_1) + f(u_2) = \bar{0} + \bar{0} = \bar{0}$$

y

$$f(\lambda u_1) = \lambda f(u_1) = \lambda \times \bar{0} = \bar{0}.$$

Por tanto el núcleo es un subespacio vectorial de E. □

Así que una aplicación lineal tiene asociados dos subespacios que, como veremos más adelante, prácticamente la describen del todo, desde el punto de vista geométrico.

La *imagen* de una aplicación lienal ha de verse como "el espacio que recubre" la aplicación: en realidad, ha de verse como "la imagen" que produce el espacio E al "ser enviado" por la aplicación f en el espacio F. El núcleo no es más que la solución de la ecuación $f(u) = \bar{0}$, que es precisamente una ecuación lineal homogénea (esto lo veremos en breve). Si tomamos un vector de la imagen, $v \in \mathrm{Im}(f)$, la ecuación

$$f(u) = v$$

significa "¿qué vectores de E caen en v?" La solución es: basta calcular un vector que caiga en v (llamémoslo u_0) y el conjunto de soluciones de

$$f(u) = v$$

es exactamente el conjunto de vectores $u_0 + \tilde{u}$, donde $\tilde{u}$ es cualquier vector de $\mathrm{Ker}(f)$, del núcleo. Es decir, la soluciónde $f(u) = v$ es *la suma de una solución particular con las de la homogénea*, pues $\mathrm{Ker}(f)$ es el conjunto de vectores $u \in E$ tales que

$$f(u) = 0$$

(ecuación homogénea asociada a $f(u) = v$). Esto se verá muy fácilmente cuando utilicemos coordenadas.

2.1. Composición de aplicaciones. Las aplicaciones lineales (como cualesquiera otras aplicaciones) se pueden *componer*: es decir, se puede "realizar primero una transformación y luego otra" (recuérdese que una aplicación no es más que una *transformación* de un espacio en otro). En concreto,

DEFINICIÓN 2.6. Dadas $f : E \to F$ y $g : F \to G$ dos aplicaciones lineales, la *composición* $g \circ f$ es la aplicación

$$E \xrightarrow{g \circ f} G$$

definida como

$$\begin{aligned} E &\xrightarrow{g\circ f} G \\ u &\longmapsto g(f(u)) \end{aligned}$$

es decir, "aplicar primero f y después g".

LEMA 2.7. *La composición de dos aplicaciones lineales es una aplicación lineal.*

DEMOSTRACIÓN. Esto es poco menos que evidente:

Suma: $g \circ f(u+v) = g(f(u+v)) = g(f(u)+f(v)) = g(f(u))+g(f(v)) = g \circ f(u) + g \circ f(v)$.

Producto exterior: $g \circ f(\lambda u) = g(f(\lambda u)) = g(\lambda f(u)) = \lambda g(f(u)) = \lambda g \circ f(u)$.

Lo que se quiere decir es que transformar linealmente lo transformado linealmente es lineal. □

Proposición 2.8. *Las siguientes aplicaciones son lineales:*

1. *La aplicación* identidad *en cualquier espacio vectorial E, que se denota Id_E, dada por $Id_E(u) = u$.*
2. *La operación "suma" es lineal. Si E es un esapcio vectorial y $E \times E$ es el producto cartesiano de E consigo mismo, la aplicación $f : E \times E \to E$ que envía un par (u, v) a $u + v$ es lineal.*
3. *Fijado un escalar $\lambda \in \mathbb{R}$, la aplicación "multiplicar por λ", de cualquier espacio vectorial en sí mismo: $f : E \to E$, $f(u) = \lambda \cdot u$. Esta aplicación se llama* homotecia *de razón λ.*
4. *Dado un subespacio vectorial $F \subset E$, la aplicación "inclusión", que no es más que mandar cada vector en sí mismo, pero de un espacio pequeño en uno grande: $i : F \to E$ $i(u) = u$.*
5. *Si $F = E_1 \times E_2$ (el producto cartesiano de dos espacios vectoriales), cada* proyección *$\pi_1 : F \to E_1$ y $\pi_2 : F \to E_2$ son lineales.*
6. Proyección en una dirección. *Si E es un espacio vectorial y F_1 y F_2 son dos subespacios tales que $E = [F_1, F_2]$ (entre ambos generan todo el espacio) y $F_1 \cap F_2 = \{\bar{0}\}$, entonces cualquier vector $v \in E$ se escribe de manera única como $v = v_1 + v_2$, con $v_1 \in F_1$ y $v_2 \in F_2$ (esto es un sencillo f ejercicio). Se llama* proyección sobre F_1 en la dirección F_2 *a la aplicación $\pi : E \to F_1$ definida como $\pi(u) = v_1$ (donde v_1 es la "componente de u en F_1). Este es el ejemplo de la Sección 1.1.*

Nota 2.9. El producto cartesiano de dos espacios vectoriales E_1, E_2 es el producto cartesiano $E_1 \times E_2$, el conjunto de pares (u, v) donde $u \in E_1$ y $v \in E_2$. Es prácticamente evidente que $E_1 \times E_2$ es un espacio vectorial.

De todo lo dicho anteriormente, aunque no lo hayamos explicitado, se deduce claramente el siguiente resultado:

Lema 2.10. *Si E y F son espacios vectoriales de dimensión finita, con $\dim(E) = m$ y $\mathcal{B}_E = (e_1, \dots, e_m)$ es una base de E y $v_1, \dots, v_m$ son m vectores de F, existe una única aplicación lineal $f : E \to F$ tal que $f(e_1) = v_1, \dots, f(e_m) = v_m$.*

Cuando se utilice este lema, se dirá que la aplicación "se extiende por linealidad".

Como es bien sabido, una aplicación $f : E \to F$ es biyectiva si y solo si existe una aplicación f^{-1} (la *inversa* de f) tal que $f \circ f^{-1} = Id_F$ y $f^{-1} \circ f = Id_E$.

Proposición 2.11. *Si $f : E \to F$ es inversible (biyectiva) y f^{-1} es su inversa, entonces f^{-1} es lineal.*

Demostración. Se trata de comprobar, como antes, que $f^{-1}(u+v) = f^{-1}(u) + f^{-1}(v)$ y que $f^{-1}(\lambda u) = \lambda f^{-1}(u)$, para $u, v \in F$ y $\lambda \in \mathbb{R}$.

Para la suma, sean $\bar{u}, \bar{v} \in E$ tales que $f(\bar{u}) = u$ y $f(\bar{v}) = v$. Como f es lineal, resulta que $f(\bar{u} + \bar{v}) = u + v$. Pero esto significa, al ser f biyectiva, que $f^{-1}(u + v) = \bar{u} + \bar{v}$, que no es más que $f^{-1}(u) + f^{-1}(v)$.

Para el producto por escalares se procede exactamente igual. Se sabe que $f(\lambda\bar{u}) = \lambda f(\bar{(u)}) = \lambda u$. Por tanto, como f es biyectiva, ha de ser $f^{-1}(\lambda u) = \lambda\bar{u} = \lambda f^{-1}(u)$. □

De esta manera podemos construir una buena cantidad de aplicaciones lineales, si no todas.

Terminamos con un resultado relativo a la dependencia e independencia lineal:

LEMA 2.12. *Sea $f : E \to F$ una aplicación lineal. Si un subconjunto $S \subset E$ es ligado, entonces su imagen $f(S)$ es un subconjunto ligado de F. Por tanto, si $f(S)$ es libre, entonces S era libre.*

DEMOSTRACIÓN. Es obvio: si $\lambda_1 u_1 + \cdots + \lambda_k u_k = 0$ con algún $\lambda_i \neq 0$, entonces $\lambda_1 f(u_1) + \cdots + \lambda_k f(u_k) = 0$ y el mismo λ_i es $\neq 0$. □

2.2. Ejemplos construidos a partir de lo anterior.

Está sección se rellenará en el futuro.

3. Expresión matricial de las aplicaciones lineales

Las aplicaciones lineales se reducen prácticamente a composiciones de homotecias, sumas, productos por escalares, inclusiones y proyecciones en diferentes direcciones. Pero esta manera de utilizarlas es poco útil computacionalmente hablando. La expresión de vectores en coordenadas, utilizando *bases* permite escribir las aplicaciones lineales en forma más útil (desde el punto de vista computacional): como matrices.

Partimos de una aplicación lineal $f : E \to F$, entre espacios vectoriales de dimensión finita (esto es esencial para la escritura en forma matricial), digamos que $\dim(E) = m$ y $\dim(F) = n$.

Fijamos una base de E y una de F.

Repetimos por si no ha quedado claro:

Fijamos una base de E y una de F.

Y por tercera vez:

Fijamos una base de E y una de F.

Finalmente:

Fijamos una base de E y una de F.

Es decir, fijamos una base $\mathcal{B}_E = \{u_1, \dots, u_m\}$ de E y una base $\mathcal{B}_F = \{v_1, \dots, v_n\}$ de F (recordamos que F tiene dimensión n y E tiene dimensión m).

Una vez que hemos **fijado una base de E y una de F**, como hemos hecho, cada vector de $\mathcal{B}_E$ tendrá una imagen:

$$\begin{aligned} f(u_1) &= a_{11}v_1 + a_{21}v_2 + \cdots + a_{n1}v_n \\ f(u_2) &= a_{12}v_1 + a_{22}v_2 + \cdots + a_{n2}v_n \\ &\vdots \\ f(u_m) &= a_{1m}v_1 + a_{2m}v_2 + \cdots + a_{nm}v_n \end{aligned} \tag{18}$$

Cualquier vector $u \in E$ puede expresarse como combinación lineal de los vectores de la base $\mathcal{B}_E$:

$$u = \lambda_1 u_1 + \cdots + \lambda_m u_m,$$

y, como f es lineal, resulta que

$$f(u) = \lambda_1 f(u_1) + \cdots + \lambda_m f(u_m),$$

que, utilizando las igualdades (18), queda

$$\begin{aligned} f(u) = (\lambda_1 a_{11} + \lambda_2 a_{12} + \cdots + \lambda_m a_{1m})v_1 + \\ (\lambda_1 a_{21} + \lambda_2 a_{22} + \cdots + \lambda_m a_{2m})v_2 + \\ + \cdots + \\ + (\lambda_1 a_{n1} + \lambda_2 a_{n2} + \cdots + \lambda_m a_{nm})v_n. \end{aligned}$$

Que, **utilizando las coordenadas que proporcionan las bases $\mathcal{B}_E$ y $\mathcal{B}_F$**, puede escribirse:

$$f(u)_{\mathcal{B}_F} = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1m} \\ a_{21} & a_{22} & \dots & a_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nm} \end{pmatrix} \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_m \end{pmatrix}_{\mathcal{B}_E} \tag{19}$$

Es decir, **las coordenadas del vector $f(u)$ en la base $\mathcal{B}_F$ se calculan "multiplicando" la matriz (a_{ij}) por el vector de las coordenadas de u en la base $\mathcal{B}_E$**. La matriz se denomina *matriz asociada a la aplicación lineal f en las bases $\mathcal{B}_E$ y $\mathcal{B}_F$* y a veces escribiremos $A(f, \mathcal{B}_E, \mathcal{B}_F)$, pero por lo general diremos A, porque sabremos a qué aplicación y bases nos estaremos refiriendo.

No se hará nunca demasiado hincapié en que *todo lo que se ha escrito arriba depende esencialmente de las bases que se utilizan.*

Como se ve, la matriz $A(f, \mathcal{B}_E, \mathcal{B}_F)$ se "calcula" escribiendo las coordenadas de cada vector $f(u_1), \dots, f(u_m)$ en la base $\mathcal{B}_F$ *en columnas*, de izquierda a derecha:

$$A(f,\mathcal{B}_E,\mathcal{B}_F)=\begin{pmatrix} | & | & & | \\ f(u_1) & f(u_2) & \dots & f(u_m) \\ \downarrow & \downarrow & & \downarrow \end{pmatrix}$$

de manera que tiene m columnas y n filas.

Si $u=(x_1,\dots,x_m)_{\mathcal{B}_E}$ es un vector de E y $f(u)=(y_1,\dots,y_n)_{\mathcal{B}_F}$ es su imagen por la aplicación lineal f, entonces, si $A(f,\mathcal{B}_E,\mathcal{B}_F)=(a_{ij})$ es la matriz de f en las bases $\mathcal{B}_E$ y $\mathcal{B}_F$, se tiene que (esto ya lo hemos dicho arriba en (19))

$$\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}=\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1m} \\ a_{21} & a_{22} & \dots & a_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nm} \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{pmatrix} \tag{20}$$

3.1. La imagen y el núcleo, vistos en la matriz. La expresión matricial de una aplicación lineal permite conocer de manera inmediata tanto la imagen como el núcleo de la aplicación (en concreto, un sistema de generadores de la imagen y un sistema de ecuaciones que define el núcleo).

Partimos de una aplicación lineal $f:E\to F$ y de bases $\mathcal{B}_E$ y $\mathcal{B}_F$. Se considera la matriz A asociada a f en esas bases.

PROPOSICIÓN 3.1. *Los vectores columna de la matriz, en la base $\mathcal{B}_F$ son un sistema de generadores de* $\mathrm{Im}(f)$*.*

Las filas de A, entendidas como ecuaciones de un sistema lineal homogéneo definen el $\mathrm{Ker}(f)$ en la base $\mathcal{B}_E$*.*

DEMOSTRACIÓN. La primera afirmación es evidente, pues al ser la aplicación lineal y la imagen un subespacio, está generada por la imagen de una base. Y las coordenadas en $\mathcal{B}_F$ de la imagen de la base $\mathcal{B}_E$ son precisamente las columnas de A.

La segunda afirmación es exactamente la expresión en coordenadas de la definición de $\mathrm{Ker}(f)$: los vectores que "van a parar al $\bar{0}$". En las bases que se tienen, se escribe

$$\begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}=\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1m} \\ a_{21} & a_{22} & \dots & a_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nm} \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{pmatrix}_{\mathcal{B}_E}$$

(el vector $\bar{0}$ siempre tiene coordenadas $(0,\dots,0)$). Y esta expresión es exactamente el sistema de ecuaciones homogéneo cuya matriz asociada es A. □

3.2. Ejemplos.

Ejemplo 33. Si E es un espacio vectorial de dimensión n y fijamos una base $\mathcal{B}_E$ como "espacio de partida" **y la misma base** como "espacio de llegada", entonces la matriz de la aplicación identidad es la matriz "identidad" (matriz diagonal con unos en la diagonal):

$$A(Id_E, \mathcal{B}_E, \mathcal{B}_E) = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{pmatrix}$$

Nótese que *aunque el espacio de partida y el de llegada sean el mismo*, se han de fijar dos bases. La matriz identidad de tamaño n se denota Id_n.

Sin embargo, si no se utiliza la misma base, la matriz no es necesariamente esa. Por ejemplo, $E = \mathbb{R}^2$, la base de partida $\mathcal{B}_1$ la estándar, la base de llegada $\mathcal{B}_2$ dada por $v_1 = (1,1)_{\mathcal{B}_1}, v_2 = (1,0)_{\mathcal{B}_1}$. Entonces

$$\begin{aligned} f(e_1) &= (1,0)_{\mathcal{B}_1} = \lambda_1(1,1)_{\mathcal{B}_1} + \mu_1(1,0)_{\mathcal{B}_1} \\ f(e_2) &= (0,1)_{\mathcal{B}_1} = \lambda_2(1,1)_{\mathcal{B}_1} + \mu_2(1,0)_{\mathcal{B}_1}. \end{aligned}$$

que tienen como solución $\lambda_1 = 0, \mu_1 = 1$ y $\lambda_2 = 1, \mu_2 = -1$. Así que

$$A(Id, \mathcal{B}_1, \mathcal{B}_2) = \begin{pmatrix} 0 & 1 \\ 1 & -1 \end{pmatrix}$$

que no se parece a la matriz identidad. En resumen: la aplicación identidad tiene como matriz asociada la identidad *si se toma la misma base de origen y de partida.*

Ejemplo 34. Homotecias. Si $E \xrightarrow{f} E$ es la homotecia de razón λ (es decir, $f(v) = \lambda v$), *y se toma la misma base de partida y de llegada*, entonces

$$A(f, \mathcal{B}, \mathcal{B}) = \lambda Id_n.$$

Es decir, la matriz diagonal que tiene en la diagonal solo el valor λ. Comprobar esto es inmediato.

Por la misma razón que en el ejemplo anterior, si la base de llegada no es la misma que la de partida, la matriz no tiene por que ser esa, y de hecho no lo será.

3.3. Resultados inmediatos.

De la expresión matricial de una aplicación lineal en unas bases se deducen resultados importantes de manera elemental.

TEOREMA 3.2. *[Fórmula de las dimensiones] La dimensión de la imagen más la dimensión del núcleo es la dimensión del espacio original, si los espacios son de dimensión finita.*

Demostración. Si $f : E \to F$ es una aplicación lineal entre espacios de dimensión finita, digamos m y n respectivamente, y $A = A(f, \mathcal{B}_E, \mathcal{B}_F)$ la matriz en unas bases, por construcción de la matriz, la dimensión de la imagen es el rango de A y la dimensión de E es el número de columnas de A.

El núcleo de la aplicación es precisamente el subespacio vectorial de E definido por

$$\operatorname{Ker}(f) = A \cdot \begin{pmatrix} x_1 \\ \vdots \\ x_m \end{pmatrix}_{\mathcal{B}_E} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}.$$

Precisamente la dimensión de $\operatorname{Ker}(f)$ es $m - \operatorname{rg}(A)$. Así que $\dim(\operatorname{Im}(f)) + \dim(\operatorname{Ker}(f)) = \dim(E)$. $\square$

De donde, si E tiene dimensión m y F dimensión n y A es cualquier matriz de una aplicación lineal f en ciertas bases:

Corolario 7. Se tiene que:

- La aplicación f es inyectiva si y solo si $\operatorname{rg}(A) = \dim(E)$.
- La aplicación f es sobreyectiva si y solo si $\operatorname{rg}(A) = \dim(F)$.
- La aplicación f es un isomorfismo si y solo si $\operatorname{rg}(A) = \dim(E) = \dim(F)$.

Por tanto, dos espacios vectoriales son isomorfos si y solo si tienen la misma dimensión.

El último resultado es también elemental de comprobar:

Teorema 3.3. *Dos espacios vectoriales de dimensión finita son isomorfos si y solo si su dimensión es la misma.*

Demostración. $\Rightarrow$) Como son isomorfos, existe una aplicación lineal $f : E \to F$ que es inyectiva y sobre. De la fórmula de las dimensiones resulta que $\dim(Ker(f)) + \dim(\operatorname{Im}(f)) = \dim(E)$. Pero como f es inyectiva, $\operatorname{Ker}(f) = \{\bar{0}\}$, y como es sobreyectiva, $\operatorname{Im}(f) = F$. Por tanto, $\dim(F) = \dim(E)$.

$\Leftarrow$) Si tienen la misma dimensión, E tiene una base $\mathcal{B}_E = (e_1, \ldots, e_m)$ de m elementos y F otra $\mathcal{B}_F = (v_1, \ldots, v_m)$ de también m elementos. Definamos la aplicación lineal

$$E \xrightarrow{f} F$$

como $f(e_1) = v_1, \ldots, f(e_m) = v_m$. Sabemos que hay una única aplicación lineal f que cumple esas propiedades. La matriz de esa aplicación en $\mathcal{B}_E$ y $\mathcal{B}_F$ es la identidad (¿por qué?), que tiene rango m. Así que f es inyectiva y sobreyectiva, usando la fórmula de las dimensiones. $\square$

La importancia de los espacios isomorfos (del concepto de isomorfía) es que nos permite "identificarlos", trabajar siempre con un espacio de dimensión n *como si fuera* $\mathbb{R}^n$.

Por otro lado, es importante saber que *cualquier matriz es la matriz asociada a una aplicación lineal.*

TEOREMA 3.4. *Sea A una matriz $n \times m$ (n filas y m columnas) y sean E y F espacios vectoriales de dimensiones m y n respectivamente. Existe una aplicación lineal $f : E \to F$ y bases $\mathcal{B}_E$ de E y $\mathcal{B}_F$ de F tales que $A(f, \mathcal{B}_E, \mathcal{B}_F) = A$.*

DEMOSTRACIÓN. Es evidente. Fijemos dos bases cualesquiera $\mathcal{B}_E = (e_1, \dots, e_m)$ y $\mathcal{B}_F = (v_1, \dots, v_n)$ de E y F respectivamente. Si la matriz es

$$A = \begin{pmatrix} a_{11} & \dots & a_{1m} \\ \vdots & \ddots & \vdots \\ a_{n1} & \dots & a_{nm} \end{pmatrix}$$

basta con definir las imágenes de cada e_i:

$$f(e_i) = a_{1i}v_1 + \dots + a_{ni}v_n$$

y "extender" por linealidad. □

3.4. Composición y producto. Supongamos que E, F y G son espacios vectoriales de dimensiones m, n y p respectivamente y que $f : E \to F$ y $g : E \to F$ son aplicaciones lineales. Sabemos que $g \circ f$ (la composición) es una aplicación lineal.

Fijemos bases $\mathcal{B}_E = (e_1, \dots, e_m)$, $\mathcal{B}_F = (f_1, \dots, f_n)$ y $\mathcal{B}_G = (g_1, \dots, g_p)$. Una vez hecho esto, la aplicación f tiene una matriz asociada que es $A = (a_{ij})$ y la aplicación g otra, $B = (b_{ij})$, la primera de tamaño $n \times m$ y la segunda $p \times n$.

Nos preguntamos cuál es la matriz asociada a $g \circ f$.

TEOREMA 3.5. *La matriz asociada a $g \circ f$ es el producto de las matrices A y B, de la forma $B \times A$ (en el mismo orden en que se escribe la composición).*

DEMOSTRACIÓN. Esto no es más que un lío de cuentas.

Hasta ahora no hemos definido el *producto* de matrices, pero está claro que la matriz B "se multiplica" por vectores de longitud n (en la base $\mathcal{B}_F$) para producir vectores de longitud p (en la base $\mathcal{B}_G$). Ahora bien, A no es más que *una lista de n vectores de longitud m.* Se define el *producto* $B \times A$ como la lista de n vectores que se genera al multiplicar B por cada vector columna de A.

La comprobación de que esa es la matriz de $g \circ f$ no es más que verificar, para cada i, que la imagen de e_i por $g \circ f$ en la base $\mathcal{B}_G$ es la columna i de B. Pero esto es evidente, pues $(g \circ f)(e_i) = g(f(e_i))$, que es la imagen en la base $\mathcal{B}_G$ del vector columna i de la matriz A. (Es decir, estamos diciendo lo obvio). □

NOTA 3.6. Téngase siempre mucho cuidado al multiplicar con el orden y los tamaños de las matrices: no se puede multiplicar cualquier cosa por cualquier cosa.

NOTA 3.7. Compruébese qué ocurre si se multiplica una matriz por la identidad (a cada lado). Lo mismo por una matriz diagonal (a cada lado).

NOTA 3.8. Compruébese que "las operaciones elementales por filas o por columnas" se pueden entender como multiplicaciones por matrices "sencillas" (bien a la derecha, bien a la izquierda).

3.5. Matriz y aplicación inversa. El caso especial de las aplicaciones entre el mismo espacio vectorial (o entre espacios de la misma dimensión) requiere un estudio especial.

DEFINICIÓN 3.9. Una aplicación lineal $f : E \to E$ de un espacio en sí mismo se llama *endomorfismo.*

En este caso, se tiene:

TEOREMA 3.10. *Dado un endomorfismo $f : E \to E$, son equivalentes:*

- *El endomorfismo f es inyectivo.*
- *El endomorfismo f es epiyectivo.*
- *El endomorfismo f es biyectivo.*

DEMOSTRACIÓN. Esto se deduce de la fórmula de las dimensiones:

$$\dim(\mathrm{Ker}(f)) + \dim(\mathrm{Im}(f)) = \dim(E)$$

Si es inyectivo, el núcleo es de dimensión 0, con lo que la imagen es de dimensión total, así que es epiyectivo. Si es epiyectivo, el núcleo es nulo, con lo que es biyectivo. □

En el caso en que un endomorfismo $f : E \to E$ es biyectivo, se dice que f es un *automorfismo*. En este caso especial, existe una aplicación lineal f^{-1} que es la *inversa* de f y que es también biyectiva, y cumple:

$$f \circ f^{-1} = f^{-1} \circ f = Id_E.$$

Si fijamos *la misma base* en la salida que en la llegada (esto no es obligatorio, pero es lo común), f tendrá una matriz asociada A y f^{-1} otra, B. Pero como estamos en la misma base, se tiene que $A(Id_E, \mathcal{B}_E, \mathcal{B}_E) = Id_m$, es decir, la matriz de $f \circ f^{-1}$ y la de $f^{-1} \circ f$ son la matriz identidad de tamaño m. Es decir,

$$A \times B = B \times A = Id_n.$$

DEFINICIÓN 3.11. Dada una matriz cuadrada A, se llama *inversa de A* a la matriz A^{-1} (si existe) tal que $A \times A^{-1} = A^{-1} \times A = Id_n$.

Está claro que

Lema 3.12. *Una matriz A cuadrada tiene inversa si y solo si su rango es máximo, si y solo si su determinante es no nulo. En este caso se dice que A es inversible.*

Lema 3.13. *[Algoritmo para calcular la inversa] Dada una matriz cuadrada A de n filas, para calcular la inversa se opera como sigue:*

1. *Se escriben al lado Id_n y A (a la derecha y a la izquierda, respectivamente).*
2. *Se hacen operaciones elementales en la izquierda para transformarla en la Id_n (Gauss, luego se divide cada fila por el pivote y se hace Gauss de abajo arriba para hacer ceros —esta es la parte de "Jordan" de Gauss-Jordan—).*
3. Cada una *de esas operaciones elementales se realiza a la vez en la matriz de la izquierda.*

Al final, la matriz de la izquierda es la inversa de A (y a la derecha ha quedado, claro Id_n).

3.6. Cambio de base. Como ya hemos visto (y espero que hayamos insistido lo suficiente), la expresión matricial de una aplicación lineal depende *esencialmente* de las bases fijadas en cada espacio vectorial. En esta parte vamos a explicar cómo se "cambia de base" una matriz. Es decir, dadas nuevas bases, ¿cómo es la matriz de una aplicación lineal en estas nuevas bases, en relación con la anterior matriz?

Suponemos para comenzar que se tiene una aplicación lineal $f : E \to F$ y bases $\mathcal{B}_E^1$ y $\mathcal{B}_F^1$ en E y F. Llamemos A a la matriz de f en estas bases: $A = A(f, \mathcal{B}_E^1, \mathcal{B}_F^1)$. Digamos que la base $\mathcal{B}_E^1$ está compuesta por los vectores $(e_1, \dots, e_m)$ y la $\mathcal{B}_F^1 = (f_1, \dots, f_n)$ (E tiene dimensión m y F dimensión n, como siempre).

Ahora fijamos otras bases, $\mathcal{B}_E^2$ y $\mathcal{B}_F^2$. La manera de "darlas" es siempre la misma: se supone que se *sabe escribir los vectores nuevos en función de los antiguos*. Es decir, si las nuevas bases son $\mathcal{B}_E^2 = (\bar{e}_1, \dots, \bar{e}_m)$ y $\mathcal{B}_F^2 = (\bar{f}_1, \dots, \bar{f}_n)$:

$$\begin{array}{ll} \bar{e}_1 = p_{11}e_1 + \dots + p_{m1}e_m & \bar{f}_1 = q_{11}f_1 + \dots + q_{n1}f_n \\ \vdots & \vdots \\ \bar{e}_m = p_{1m}e_1 + \dots + p_{mm}e_m & \bar{f}_n = q_{1n}f_1 + \dots + q_{nn}f_n \end{array} \tag{21}$$

que no es más que decir que la aplicación Id_E tiene como matriz asociada $P = (p_{ij})$ en las bases $(\mathcal{B}_E^2, \mathcal{B}_E^1)$ (en ese orden) y que la aplicación Id_F tiene como matriz asociada $Q = (q_{ij})$ en las bases $(\mathcal{B}_F^2, \mathcal{B}_F^1)$ (en ese orden).

Definición 3.14. La matriz P se llama *matriz de cambio de base* de $\mathcal{B}_E^1$ a $\mathcal{B}_E^2$, y la Q *matriz de cambio de base* de $\mathcal{B}_F^1$ a $\mathcal{B}_F^2$.

NOTA 3.15. Es muy sencillo, visto el comentario anterior, comprobar que *la matriz de cambio de base de* $\mathcal{B}_E^2$ *a* $\mathcal{B}_E^1$ *es* P^{-1}*, la inversa.*

Escribir la matriz de f en las bases nuevas $\mathcal{B}_E^2$ y $\mathcal{B}_F^2$ no es más que escribir la matriz que resulta de componer las siguientes aplicaciones, escrita en cada base:

$$(E, \mathcal{B}_E^2) \xrightarrow{Id_E} (E\mathcal{B}_E^1) \xrightarrow{f} (F, \mathcal{B}_F^1) \xrightarrow{Id_F} (F, \mathcal{B}_F^2)$$

pero la primera Id_E, como hemos visto, tiene como matriz asociada en las bases $\mathcal{B}_E^2$, $\mathcal{B}_E^1$ la matriz P, mientras que la Id_F en las bases $\mathcal{B}_F^1$ y $\mathcal{B}_F^2$ tiene como matriz asociada Q^{-1}, la inversa de Q.

Por tanto:

TEOREMA 3.16. *En las condiciones anteriores, la matriz de* f *en las bases nuevas* $\mathcal{B}_E^2$ *y* $\mathcal{B}_F^2$ *es*

$$A(f, \mathcal{B}_E^2, \mathcal{B}_F^2) = Q^{-1} \times A \times P,$$

donde A *es la matriz en las bases originales, y* P *y* Q *las matrices de cambio de base correspondientes a* E *y* F*, respectivamente.*

Como es obvio, si E y F son el mismo espacio (es decir, si f es un endomorfismo) *y se usa la misma base en el origen y en la llegada*, entonces se tiene un resultado más simple:

TEOREMA 3.17. *Si* $f : E \to E$ *es un endomorfismo y* $\mathcal{B}_E$ *es una base en* E *y* $A(f, \mathcal{B}_E) = A$ *y* $\mathcal{B}'_E$ *es otra base cuya matriz de cambio es* P*, entonces*

$$A(\mathcal{B}'_E) = P^{-1} \times A \times P.$$

NOTA 3.18. Insistimos en que la matriz de cambio de $\mathcal{B}_E$ a $\mathcal{B}'_E$ expresa *cómo se escribe cada vector de la segunda base en función de los de la primera.*

3.6.1. Suma y producto por escalares de aplicaciones lineales. Dadas dos aplicaciones lineales $f : E \to F$ y $g : E \to F$ entre los mismos espacios vectoriales y un número real λ, se pueden definir la aplicación *suma* y la aplicación *producto por un escalar*:

$$E \xrightarrow{f+g} F, \quad E \xrightarrow{\lambda f} F$$

La suma se define como $(f + g)(u) = f(u) + g(u)$ y el producto por escalar como $(\lambda f)(u) = \lambda f(u)$. Se tiene el siguiente resultado:

LEMA 3.19. *Sean* $\mathcal{B}_E$ *y* $\mathcal{B}_F$ *bases de* E *y* F *respectivamente. Si* f *tiene como matriz asociada en esas bases la matriz* A *y* g *la matriz* B*, entonces:*

- $f + g$ *tiene como matriz asociada* $A + B$ *(en esas bases)*
- λf *tiene como matriz asociada* λA *(ibid.).*

Como es fácilmente comprobable, $h \circ (f + g) = h \circ f + h \circ g$, si $g : F \to G$ es otra aplicación lineal. Así que, escribiendo todo en las bases correspondientes, es fácil probar que

Lema 3.20. *Dadas tres matrices A, B y C, si A y B son $n \times m$ y C es $p \times n$, entonces*

$$C(A+B) = CA + CB.$$

Además, como $h \circ (g \circ (f)) = (h \circ (g)) \circ f$ (esto es nada más que la composición de aplicaciones) se tiene que

Lema 3.21. *Dadas matrices A, B y C* que se puedan multiplicar *(es decir, A es $n \times m$, B es $m \times p$ y C es $p \times r$), se tiene que*

$$A \times (B \times C) = (A \times B) \times C.$$

Para terminar:

Lema 3.22. *Supongamos que A y B son matrices cuadradas $n \times n$. Entonces*

$$\det(A \times B) = \det(A)\det(B)$$

el producto de los determinantes es el determinante del producto.

4. Vectores y valores propios

Dado un endomorfismo, $f : E \to E$, hay ciertos vectores especialmente interesantes, que son aquellos para los que f se comporta de manera "sencilla".

Definición 4.1. Se dice que v es un autovector de f si existe algún número real λ tal que $f(v) = \lambda v$.

Un autovector es un vector para el que f funciona "como una mera homotecia". El factor se llama autovalor:

Definición 4.2. Se dice que λ es un autovalor de f si existe algún vector $v \in E$ tal que $f(v) = \lambda v$.

Es decir, un autovector *está asociado* a un autovalor y viceversa.

Nota 4.3. Nótese que si v es un autovector, entonces cualquier vector que esté en "la recta de dirección v" también es un autovector, pues:

$$f(\mu v) = \mu f(v) = \mu\lambda v = \lambda(\mu v).$$

Por tanto, dado un autovalor (como λ), hay siempre al menos una *recta* de autovectores asociados.

De hecho, el resultado es más general:

Lema 4.4. *Sea λ un autovalor de un endomorfismo E. El conjunto de vectores de E que son autovectores de λ es un subespacio vectorial de E, que llamaremos F_λ. Además, F_λ es* invariante *por f.*

DEMOSTRACIÓN. Es igual de sencillo que antes. De hecho, por la nota 4.3 sabemos que si $v \in F_\lambda$, entonces $\mu v \in F_\lambda$. Solo nos queda probar que la suma de autovectores asociados a λ es un autovector asociado a λ. Tomemos dos u, v:

$$f(u+v) = f(u) + f(v) = \lambda u + \lambda v = \lambda(u+v).$$

Así que $u+v$ es un autovector asociado a λ.

Por otro lado, la "invarianza" de F_λ significa que $f(F_\lambda) \subset F_\lambda$ (la imagen del subespacio está en el subespacio). Pero esto es obvio, pues la imagen de cualquier autovector v asociado a λ es un múltiplo de v, con lo cual $f(v) \in [v]$, por lo que $f(v) \in [F_\lambda]$. □

El subespacio F_λ se denomina *subespacio propio* asociado a λ.

Si sabemos que λ es un autovalor de f, es sencillo calcular el subespacio propio F_λ:

LEMA 4.5. *El subespacio F_λ es el núcleo del endomorfismo $f - \lambda Id$.*

DEMOSTRACIÓN. Esto es obvio, por definición de núcleo y de autovalor. □

4.1. Diagonalización. Hay casos especiales de endomorfismos cuya estructura conviene conocer. Repetimos que *siempre que estudiemos endomorfismos, suponemos que la base inicial es la misma que la final.*

DEFINICIÓN 4.6. Un endomorfismo $f : E \to E$ se dice *diagonalizable* si existe una base $\mathcal{B}_E$ tal que la matriz asociada a f en esa base, $A(f, \mathcal{B}_E)$, es diagonal.

Es claro que conviene conocer si un endomorfismo es diagonalizable, porque si se expresa en la base adecuada, su matriz es una matriz diagonal. ¿Cómo se puede saber si esto ocurre?

DEFINICIÓN 4.7. Sea f un endomorfismo de E cuya matriz en determinada base $\mathcal{B}_E$ es $A = (a_{ij})$. El polinomio que resulta de calcular el siguiente determinante:

$$\begin{vmatrix} a_{11} - X & a_{12} & \dots & a_{1m} \\ a_{21} & a_{22} - X & \dots & a_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mm} - X \end{vmatrix}$$

se denomina *polinomio característico* de f.

LEMA 4.8. *El polinomio característico* no depende de la base.

LEMA 4.9. *Un número λ es un autovalor de f si y solo si es raíz del polinomio caracterísitco de f.*

Una propiedad que asegura la diagonalización es la siguiente:

TEOREMA 4.10. *Si el polinomio característico de un endomorfismo f tiene todas sus raíces simples, entonces f es diagonalizable.*

CAPíTULO 7

Ecuaciones Diferenciales...

1. Introducción

Este capítulo es más bien una broma que una introducción a las ecuaciones diferenciales, pues *no tenemos tiempo de decir absolutamente nada*. El caso es que "hay que hablar de ellas". Vamos simplemente a introducir los tipos más simples y a resolver un par de ellas.

Una *ecuación diferencial* es una ecuación en la que una de las incógnitas es la derivada de una función.

Es mejor entender una ecuación diferencial como la ecuación de un sistema en el que una de las variables estudiadas depende de la otra de manera infinitesimal. Realmente, es mejor entender una ecuación diferencial como una ecuación que *describe un sistema no estático*. Se podría pensar también que una ecuación diferencial describe un sistema *que depende del tiempo*, aunque esto no es realmente muy correcto.

Un misil que se desplaza en horizontal quema el combustible de manera constante. ¿Cómo es su movimiento mientras le queda combustible? El asunto clave aquí es que *el peso del cohete decrece según se va quemando combustible*. Como se quema de modo constante, la fuerza que se ejerece es la misma (mientras quede combustible), pero *la masa decrece*. Así que, por una parte, tenemos la fórmula de Newton en cada instante:

$$F(t) = m(t) \cdot a(t)$$

y nos están diciendo que la fuerza es constante (tanto combustible se quema, tanta fuerza se ejerce). Pero la masa decrece (y lo hace de manera constante), así que

$$F(t) = c_1, \ \ \frac{\mathrm{ds(t)}}{\mathrm{dt}} = c_2,$$

con lo que

$$m(t) = c_2 t + m_0$$

donde m_0 es la masa incial. Es obvio que c_2 es negativo (la masa decrece con el tiempo). Por tanto, la ecuación original queda

$$a(t)(c_2 t + m_0) = c_1.$$

Claro que $a(t)$ es la *aceleración*), que es la derivada segunda de la posición (que es justamente lo que queremos conocer):

$$\frac{\mathrm{d^2s(t)}}{\mathrm{dt^2}} \cdot (c_2 t + m_0) = c_1.$$

Es decir, podemos "conocer la posición", $s(t)$ si sabemos resolver esa ecuación, en la que aparece la posición *derivada dos veces*. De hecho, podemos calcular la aceleración instantánea, despejando:

$$\frac{\mathrm{d}^2\mathrm{s(t)}}{\mathrm{dt}^2} = \frac{c_1}{c_2 + m_0}$$

pero esto solo nos dice *cuánto acelera el misil en cada momento*. De aquí ?se puede deducir la trayectoria?

La respuesta es que sí, *salvo* que necesitamos conocer dos datos: la posición en el instante $t = 0$ y la velocidad inicial. Estos dos datos son necesarios. El hecho de conocer la aceleración *no determina el movimiento*: cualquier cuerpo *cae* con aceleración constante. No es lo mismo decir: el cuerpo cae desde 10 metros de altura que cae desde 100 metros. Y, aunque digamos que cae desde 10 metros, no es lo mismo decir que "se le deja caer" que decir que "viene a 100 Km/h". El movimiento es obviamente diferente. Lo mismo para las ecuaciones diferenciales: cuando se resuelve una ecuación diferencial en que la variable aparece derivada k veces, hace falta también dar k "datos iniciales" si se quiere conocer la solución explícitamente. Si no, se obtiene una solución "genérica", con k parámetros.

En el caso que nos ocupa, la trayectoria es, si suponemos que el misil parte de $s(0) = 0$ y con velocidad cero, para valores de los parámetros $c_1 = 1, c_2 = -1/3, c_3 = 3$:

$$s(t) = -3(x-9)\log(|x-9|) + 3x\log(9) + 3x - 27\log(9)$$

Como se ve en la figura, la velocidad disminuye a partir de un momento:

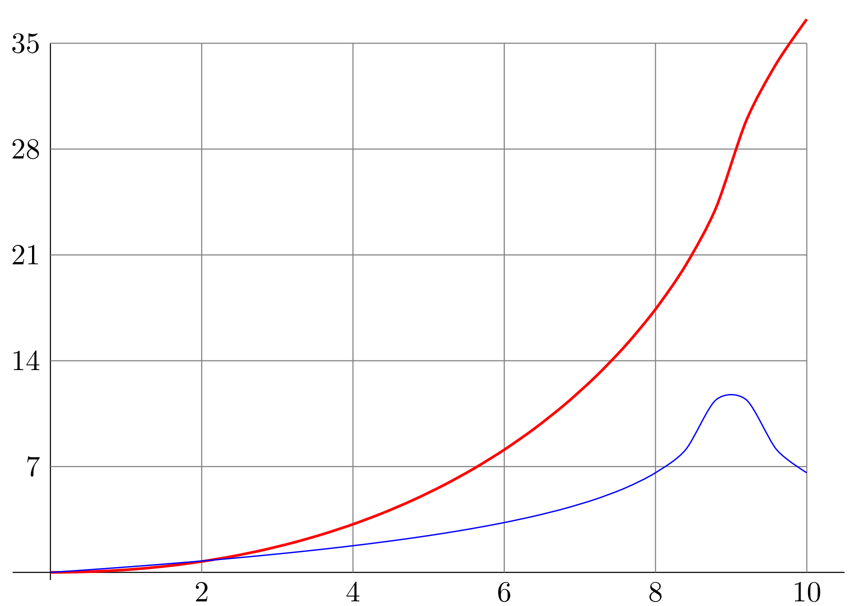

Figura 1. Trayectoria de un misil y velocidad

cuando la masa se ha terminado[1] (el modelo que hemos diseñado *no tiene en cuenta* que la masa es finita, así que llegará un momento en que la masa desaparezca...). Hay que tener cuidado con los modelos: solo son válidos mientras las hipótesis que se utilizan en el diseño lo son. Lo más difícil habitualmente es *darse cuenta de qué hipótesis se están utilizando*, sea implícita o explícitamente.

2. Ecuaciones diferenciales de orden 1. Recetario.

A continuación vamos a describir somera y rápidamente los tipos más comunes y solubles de ecuaciones diferenciales ordinarias de orden 1, dando una idea de cómo se pueden "resolver", cuando se puede.

Definición 2.1. Una *ecuación diferencial ordinaria* (EDO a partir de ahora) es una ecuación en la que aparecen las variables x, y y una variable y', que representa la derivada de y respecto a x.

Definición 2.2. El *orden* de una ecuación diferencial ordinaria es el mayor orden de derivación que aparece. Es decir, 1 si solo hay derivadas y', 2 si hay derivadas segundas y'' como mucho, etc...

Nosotros vamos a estudiar *exclusivamente* ecuaciones de orden 1.

Definición 2.3. Una *solución* de una EDO es una función $f : (a, b) \to \mathbb{R}$ de variable real tal que al sustituir y por f y y' por $f'(x)$, la ecuación se satisface.

Definición 2.4. Un *problema de condiciones iniciales* es una ecuación diferencial ordinaria E y una ecuación de la forma $y(x_0) = y_0$.

Definición 2.5. Una *solución* de un problema de condiciones iniciales es una solución $f(x)$ de la EDO E tal que $f(x_0) = y_0$ (es decir, que cumple las condiciones iniciales).

Como dijimos en la introducción, una ecuación diferencial por lo general tiene *demasiadas soluciones* porque para describir un problema diferencial con precisión *hay que imponer una condición de partida* (lo que se llama la *condición inicial*).

Ejemplo 35. Se considera la ecuación diferencial

$$y' = 1.$$

Las soluciones, como todo el mundo sabe, son las funciones $f(x) = x + c$, donde c es cualquier número real.

En términos físicos, la ecuación de arriba describe las trayectorias de los cuerpos que van a velocidad constante 1. Como es obvio, para describir *completamente* una trayectoria, hay que *decir por dónde pasa en un momento dado*: esto es la condición inicial.

[1]En el caso concreto de la gráfica, al cabo de nueve segundos no queda masa.

El problema de condiciones iniciales siguiente:

$$y' = 1,\;\; y(0) = 3,$$

describe la trayectoria de un cuerpo que va a velocidad constante 1 *y que en el instante* 0 *está en el punto* 3. Ahora sí que solo hay una solución, $f(x) = x + 3$. Como se ve, la condición inicial sirve para "saber cuál es la constante".

Ejemplo 36. Se considera la ecuación diferencial

$$y' = y.$$

¿Tiene solución?

Es bien sabido que una función $f(x)$ es igual a su derivada si y solo si es $f(x) = c \cdot e^x$, donde c es cualquier constante real. Así que $f(x) = c \cdot e^x$ es la solución de dicha ecuación. Como se ve, hay una constante libre: para cada problema de condiciones iniciales tomará un valor.

El problema de condiciones iniciales

$$y' = y,\;\; y(1) = 1$$

tiene como solución $f(x) = \frac{1}{e}e^x$. Se resuelve sin más planteando la igualdad $c \cdot e^1 = 1$ y resolviendo para c.

Ejemplo 37. La ecuación diferencial

$$y' = \cos(x)$$

tiene como soluciones las funciones $f(x) = \operatorname{sen}(x) + c$. Dependiendo de la condición inicial, c será una u otra cosa.

De ahora en adelante, escribiremos indistintamente y' y $\frac{dy}{dx}$, dependiendo de qué sea más cómodo en cada momento. También operaremos con dx y dy como si fueran valores "normales", como se verá. Todo lo que se haga tiene sentido y es correcto, pero no hace falta explicar por qué.

Pasamos a exponer los tipos más comunes de ecuaciones diferenciales ordinarias y su "solución". Entiéndase lo que sigue como un mero *recetario*.

2.1. Ecuaciones lineales.

DEFINICIÓN 2.6. Una EDO se dice *lineal* si se puede escribir como sigue:

$$y' + p(x)y = q(x).$$

Se llama "ecuación homogénea asociada" a la ecuación

$$y' + p(x)y = 0.$$

NOTA 2.7. Si $f(x)$ y $g(x)$ son una soluciones de la misma EDO lineal, entonces $f(x) - g(x)$ es solución de la ecuación homogénea asociadda.

Esto es obvio, pues la derivada es una aplicación lineal, de donde $(f(x) - g(x))' = f'(x) - g'(x)$.

2.1.1. Receta para resolver ecuaciones lineales. Partimos de una ecuación diferencial lineal

$$y' + p(x)y = q(x). \tag{22}$$

Hagamos una trampa. Multiplicamos todo por una función $m(x)$

$$m(x)y' + m(x)p(x)y = m(x)q(x)$$

y supongamos (que, como veremos, no es mucho suponer), que $m(x)p(x) = m'(x)$. Resulta que la ecuación queda

$$m(x)y' + m'(x)y = m(x)q(x),$$

cuyo miembro izquierdo es $(m(x) \cdot y)'$. Es decir, nos ha quedado la siguiente ecuación:

$$(y \cdot m(x))' = m(x)q(x).$$

De esta podemos eliminar la derivada integrando en ambos miembros, así que nos queda:

$$y \cdot m(x) = \int m(x)q(x)\,dx + c.$$

(La integración deja una constante "fuera"). Y de aquí podemos resolver ya y:

$$y = \frac{\int m(x)q(x)\,dx + c}{m(x)}, \tag{23}$$

siempre y cuando conozcamos $m(x)$.

DEFINICIÓN 2.8. Dada una EDO lineal como (22), se llama *factor integrante* a una función $m(x)$ tal que $m(x)p(x) = m'(x)$.

Es decir, si podemos encontrar un factor integrante, sabemos resolver cualquier ecuación diferencial lineal. Pero para calcular el factor integrante no hay más que resolver:

$$m'(x) = m(x)p(x),$$

es decir

$$\frac{m'(x)}{m(x)} = p(x).$$

Pero el miembro izquierdo es la derivada de $\log(m(x))$, así que es

$$(\log(m(x)))' = p(x),$$

es decir

$$\log(m(x)) = \int p(x)\,dx + c$$

o lo que es lo mismo,

$$m(x) = e^{\int p(x)+c} = ke^{\int p(x)\,dx}.$$

Es decir, *siempre se puede encontrar un factor integrante para una ecuación lineal.* Y una vez que se conoce, se resuelve la ecuación utilizando (23). Como vale cualquiera, se pone cualquier constante $k \neq 0$ y basta.

En realidad es mucho más fácil acordarse del proceso que acordarse de las fórmulas. Solo hay que "intentar escribir el miembro izquierdo como la derivada de un producto", multiplicando todo por $m(x)$.

Ejemplo 38. Encuéntrese la solución al siguiente problema de condiciones iniciales:

$$xy' + 2y = x^2 - x + 1, \;\; y(1) = \frac{1}{2}.$$

Los problemas de condiciones iniciales se resuelven calculando primero la solución "general" (que tendrá un parámetro) y luego calculando el valor del parámetro aplicando la condición inicial.

La ecuación tal como está *no* es lineal, pero se puede convertir en una lineal fácilmente, dividiendo por x:

$$y' + 2\frac{y}{x} = x - 1 + \frac{1}{x}.$$

Vamos por partes:

Factor integrante: Será una función $m(x)$. Escribamos

$$m(x)y' + m(x)\frac{2}{x}y = m(x)(x - 1 + \frac{1}{x}),$$

y se impone que

$$m(x)\frac{2}{x} = m'(x). \tag{24}$$

De aquí queda que

$$(m(x)y)' = m(x)(x - 1 + \frac{1}{x}).$$

Se resuelve con el factor: Esta última ecuación se resuelve muy fácilmente, llamando $u = m(x)y$:

$$u' = m(x)(x - 1 + \frac{1}{x}),$$

así que

$$u(x) = \int m(x)(x - 1 + \frac{1}{x})\,dx + c$$

y por tanto

$$y(x) = \frac{\int m(x)(x - 1 + \frac{1}{x})\,dx + c}{m(x)}.$$

Cálculo del f.i.: Para calcular el factor integrante se ha de resolver la ecuación (24)

$$m(x)\frac{2}{x} = m'(x).$$

Que (estas son siempre iguales) es

$$(\log(m(x)))' = \frac{2}{x},$$

es decir,

$$\log(m(x)) = 2\log(|x|)$$

de donde

$$m(x) = e^{2\log(|x|)} = |x|^2 = x^2.$$

Nótese que se pueden quitar los valores absolutos porque se está elevando al cuadrado.

Paso final: sustituir $m(x)$: Se susituye $m(x)$ por su valor en $y(x)$:

$$y(x) = \frac{\int m(x)(x-1+\frac{1}{x})\,dx + c}{m(x)},$$

así que queda

$$y(x) = \frac{\int x^2(x-1+\frac{1}{x})\,dx + c}{x^2},$$

es decir,

$$y(x) = \frac{\frac{1}{4}x^4 - \frac{1}{3}x^3 + \frac{1}{2}x^2 + c}{x^2},$$

que simplificando queda

$$y(x) = \frac{1}{4}x^2 - \frac{1}{3}x + \frac{1}{2} + \frac{c}{x^2}$$

La condición inicial $y(1) = 1/2$ implica que $1/4 - 1/3 + 1/2 + c = 1/2$, de donde $c = 1/12$.

Es decir, *la solución al problema de condiciones iniciales propuesto es*

$$y(x) = \frac{1}{4}x^2 - \frac{1}{3}x + \frac{1}{2} + \frac{1}{12x^2}.$$

La gráfica correspondiente puede verse en la figura 3.

2.2. Ecuaciones de variables separadas.

Definición 2.9. Una ecuación diferencial se dice que es *de variables separadas* si se puede escribir de la siguiente forma:

$$n(y)\frac{dy}{dx} = m(x) \tag{25}$$

donde n y m son dos funciones de variable real.

Estas ecuaciones se resuelven *implícitamente*: no se puede, por lo general, encontrar una solución de la forma $y(x)$ que, al sustituirla en (25) haga que la igualdad se cumpla. Lo que se consigue es expresar una relación *implícita* entre una función con esas características y x. En concreto

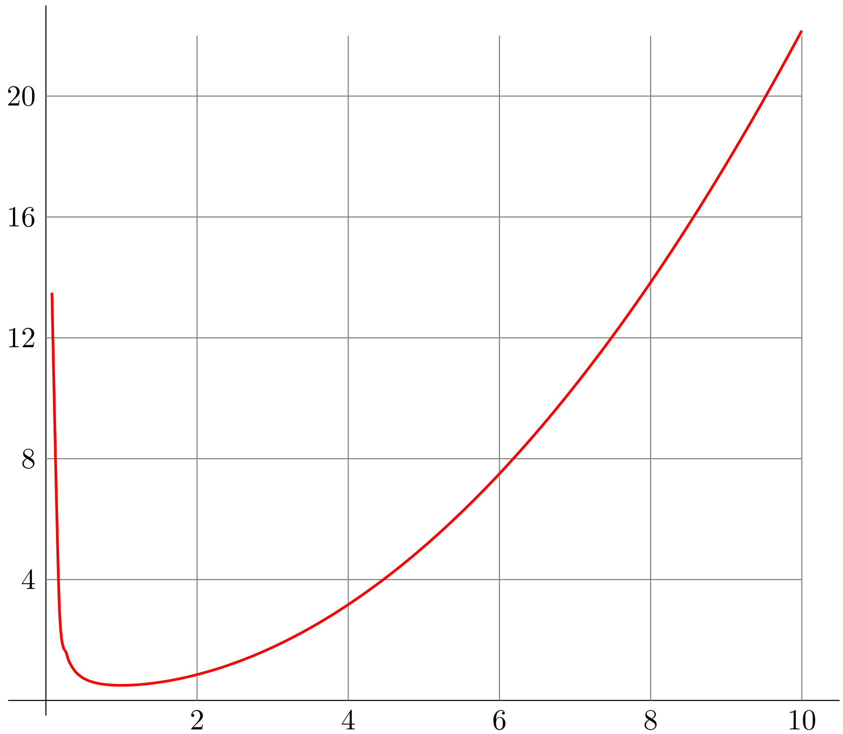

FIGURA 2. Solución del problema de c.i. propuesto

2.2.1. Receta para "resolver" las ecuaciones de variables separadas. Una vez que se ha escrito la ecuación en la forma (25), se transforma en

$$n(y)y' = m(x)dx$$

y *se integran ambos miembros de la igualdad*, el de la izquierda con respecto a y y el de la derecha con respecto a x. El resultado de esa integración

$$\int n(y)\, dy = \int m(x)\, dx$$

se llama "solución implícita" de la ecuación diferencial.

Por lo general *no se puede despejar* la y en función de la x. Cuando se pueda, se tendrán que tener en cuenta los intervalos de validez (por ejemplo, si sale una raíz cuadrada, habrá que ver cuándo el radical es positivo, etc).

Ejemplo 39. Se considera la ecuación diferencial

$$\frac{\cos(y)}{\operatorname{sen}(x)} y' = 1.$$

Calcular la solución.

Tal como está enunciado, no tiene la forma (25), pero es fácil transformarla en una de variables separadas:

$$\cos(y)y' = \operatorname{sen}(x).$$

Esta es “sencilla” de resolver implícitamente:

$$\int \cos(y)\,dy = \int \operatorname{sen}(x)\,dx$$

es decir,

$$\operatorname{sen}(y) = -\cos(x) + k.$$

Si el problema tuviera condiciones inciales, podríamos calcular la k, pero no es el caso.

Ejemplo 40. Resuélvase el siguiente problema de condiciones iniciales:

$$y' = e^{-y}(2x-4) \;\; y(5) = 0.$$

Igual que antes, basta hacer una simple transformación para expresar la ecuación como una de variables separadas:

$$e^y y' = 2x - 4.$$

Esta es sencilla de resolver:

$$\int e^y\,dy = \int 2x - 4\,dx,$$

es decir

$$e^y = x^2 - 4x + k.$$

La k se calcula con la condición inicial: $y(5) = 0$. Es decir,

$$e^0 = 25 - 20 + k,$$

de donde

$$k = -4.$$

Así pues, la solución implícita es

$$e^y = x^2 - 4x - 4.$$

Pero en este caso se puede despejar utilizando logaritmos (*habría que tener mucho cuidado de estudiar cuándo es válido lo que sigue*):

$$y = \ln(x^2 - 4x - 4).$$

2.3. Ecuaciones exactas. Estas ecuaciones son las que corresponden a la existencia de una función *potencial*: son las ecuaciones que definen las curvas de nivel de dicho potencial.

DEFINICIÓN 2.10. Una ecuación diferencial se llama *exacta* si es de la forma siguiente:

$$n(x,y)y' + m(x,y) = 0$$

donde

$$\frac{\partial n(x,y)}{\partial x} = \frac{\partial m(x,y)}{\partial y}.$$

El resultado clave para estas ecuaciones es:

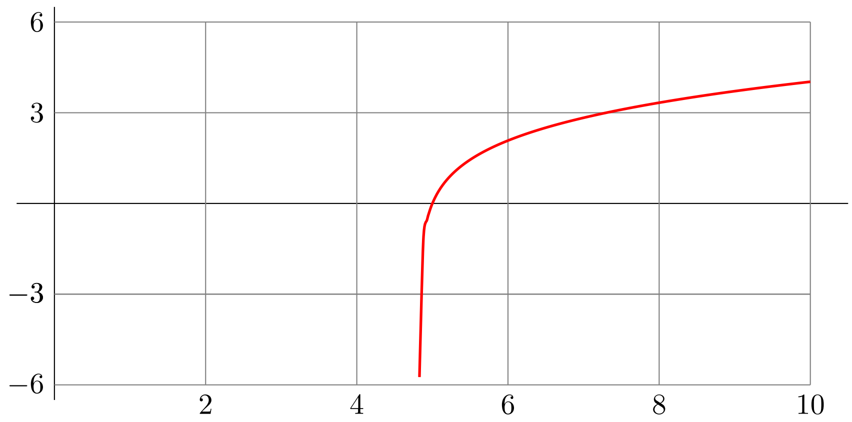

FIGURA 3. Solución de la ecuación de var. sep. propuesta

TEOREMA 2.11. *En condiciones generales (que no vamos a explicitar) dada una ecuación diferencial exacta $n(x,y)y' + m(x,y) = 0$, existe una función $V(x,y)$, que se denomina* potencial*, tal que*

$$\frac{\partial V(x,y)}{\partial x} = m(x,y), \quad \frac{\partial V(x,y)}{\partial y} = n(x,y).$$

2.3.1. Receta para resolver las ecuaciones exactas. Resolver una ecuación diferencial exacta consiste en *calcular la función potencial*, nada más. Con ella ya se tiene toda la información.

Para calcular la función potencial se procede como sigue:

1. Se integra $m(x,y)$ con respecto a x "como si la y fuera una constante". La constante de integración se escribe como una función de y. Llámese $f_1(x,y)$ al resultado de esta operación.
2. Se deriva $f_1(x,y)$ con respeto a y "como si la x fuera una constante" y se iguala el resultado a $n(x,y)$. De aquí (en concreto, de la "constante" de integración anterior) sale una ecuación diferencial en la y que se resulve y da el valor de la constante de integración que hay que utilizar.

Ejemplo 41. Resolver la siguiente ecuación diferencial:

$$2x + 3y + (3x - 3y^2)y' = 0.$$

Como se puede comprobar, la ecuación es *exacta*: si se toma $m(x,y) = 2x + 3y$ y $n(x,y) = 3x - 3y^2$, se tiene que

$$\frac{\partial m(x,y)}{\partial y} = 3 = \frac{\partial n(x,y)}{\partial y}.$$

Para resolverla, sea

$$V(x,y) = \int m(x,y)\,dx = \int 2x + 3y\,dx = x^2 + 3yx + h(y),$$

la integral de $m(x,y)$ con respecto a x y con "las constantes dependiendo de y". Ahora derivamos respecto a y como si x fuera constante:

$$\frac{\partial V(x,y)}{\partial y} = 3x + h'(y)$$

y esto debe ser igual a $n(x,y)$. Es decir,

$$3x + h'(y) = 3x - 3y^2,$$

de donde (obviamente) $h'(y) = -3y^2$, es decir

$$h(y) = -y^3 + k.$$

Por lo que

$$V(x,y) = x^2 + 3xy - y^3 + k.$$

La constante es irrelevante, por lo general. O se calcula utilizando las condiciones iniciales, si se dan.

Nota: por lo general, en las ecuaciones exactas no se pide que se *resuelva y en función de x*, sino que interesa calcular la función potencial $V(x,y)$.

2.4. Ecuaciones homogéneas.

Definición 2.12. Una ecuación diferencial se llama homogénea si se puede escribir

$$y' = F\left(\frac{y}{x}\right).$$

Las ecuaciones homogéneas se resuelven transformándolas en ecuaciones de variables separadas.

2.4.1. Receta para las ecuaciones homogéneas. Dada una ecuación homogénea

$$y' = F\left(\frac{y}{y}\right)$$

se realiza *siempre* la sustitución

$$v(x) = \frac{y}{x}.$$

teniendo en cuenta que de ahí sale

$$y(x) = x \cdot v(x),$$

es decir

$$y' = v(x) + x \cdot v'.$$

Por tanto, la ecuación se transforma en

$$v + xv' = F(v),$$

que queda

$$\frac{v'}{F(v) - v} = \frac{1}{x},$$

que se resuelve como una de variables separadas.

Ejemplo 42. Resolver el siguiente problema de condiciones iniciales

$$xyy' + 4x^2 + y^2 = 0, \;\; y(2) = -7, \;\; x > 0.$$

Nótese que, a primera vista, la ecuación *no parece* homogénea, pero es porque ha de ponerse la x dividiendo. Como $x > 0$ y el miembro de la derecha es 0, podemos dividir todo por x^2 y queda

$$\frac{y}{x}y' + 4 + \frac{y^2}{x^2} = 0,$$

que, despejando, queda

$$y' = -4\frac{x}{y} - \frac{y}{x}.$$

(nótese que $x/y = 1/(y/x)$, así que esta ecuación es homogénea). Ahora se hace el cambio

$$v = \frac{y}{x}, \;\; y' = v + xv',$$

es decir

$$v + xv' = -\frac{4}{v} - v.$$

Esta ecuación es de variables separadas (hemos ya agrupado y sacado factor común, etc...):

$$\frac{vv'}{-4-2v^2} = \frac{1}{x}.$$

Esta ecuación se integra muy fácilmente:

$$\frac{1}{4}\ln(4+2v^2) = -\ln(x) + c,$$

de donde

$$(4+2v^2)^{\frac{1}{4}} = \frac{c}{x}.$$

Y deshaciendo el cambio $v = y/x$ y teniendo en cuenta que las constantes da igual que estén elevadas o no:

$$(4 + 2\frac{y^2}{x^2}) = \frac{c}{x^4}.$$

Con la condición inicial $y(2) = -7$ sale

$$(4 + 2\frac{49}{4}) = \frac{c}{16},$$

de donde $c = 456$. La solución implícita queda

$$(4 + 2\frac{y^2}{x^2}) = \frac{456}{x^4}.$$

CAPíTULO 8

Problemas, ejemplos...

La mayoría de todos estos ejemplos están tomados del libro de Hefferon [**1**] que utilizamos como guía en todo este curso. Otros serán elementales.

1. Problemas de Gauss-Jordan

NOTA 1.1. Cuando en un sistema de ecuaciones se habla de *la forma de escalón*, se debería indicar previamente *el orden en que se consideran las variables* (pues no es lo mismo considerar la variable x "antes que" la y ó al revés. Este detalle vamos a ignorarlo. Basta con que se ordenen de la misma manera todas las variables en las ecuaciones.

Problema 1
Se considera el sistema de ecuaciones lineales siguiente:

$$\begin{gathered} 3x + 2z - 27 = y \\ z - 2y + 7x = 8x - 3 \\ 3x + 4z = 28x + y \end{gathered}$$

escribirlo en forma matricial. Calcular la forma de escalón *indicando todos los pasos intermedios.*

Problema 2
Calcular la forma de escalón del siguiente sistema de ecuaciones, *indicando todos los pasos intermedios*:

$$\begin{gathered} a + b + c + d = 1 \\ 1 = a - b + c - d \\ b + d = a + c \end{gathered}$$

Problema 3
Calcular la forma de escalón del siguiente sistema de ecuaciones *indicando todos los pasos intermedios*:

$$\begin{gathered} c_1 - 2c_2 = c_4 - 5c_7 \\ 2c_3 + c_6 = c_5 \\ c_7 - 3c_2 = 2c_1 \\ c_6 + c_2 - c_1 = c_2 \end{gathered}$$

Problema 4
¿Es el sistema cuya matriz asociada es la siguiente compatible? ¿Es determinado, si es que es compatible? ¿Por qué?

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 0 \\ 0 & 0 & 2 & 1 & 0 \\ 0 & 0 & 2 & 1 & 1 \end{pmatrix}$$

Problema 5
Mismas preguntas que en el ejercicio anterior para el sistema con matriz asociada:

$$\begin{pmatrix} 0 & 0 & 2 & 1 & 0 \\ 0 & 0 & 2 & 1 & 0 \\ 1 & 2 & 3 & 4 & 0 \end{pmatrix}$$

Problema 6
Indicar los pivotes del sistema cuya matriz asociada es la siguiente. ¿Tiene el sistema solución? Si la tiene, ¿es única? ¿Por qué? Si no es única, ¿cuántos parámetros hay? ¿cuáles son?

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

Problema 7
Discutir (es decir, explicar si el sistema es compatible, determinado o indeterminado ó incompatible) el sistema de ecuaciones cuya matriz asociada es:

$$\begin{pmatrix} 1 & 2 & 2 & 1 & 3 \\ 2 & 0 & 2 & 1 & 0 \\ 0 & 4 & 0 & 1 & 1 \end{pmatrix}$$

2. Determinantes

Problema 8
Calcular el siguiente determinante, explicando el camino utilizado para llegar al resultado:

$$\begin{vmatrix} 2 & 1 & 0 & 3 & 4 \\ 0 & 3 & 1 & 4 & 2 \\ 0 & 0 & 1 & 3 & 1 \\ 0 & 0 & 0 & -1 & 2 \\ 0 & 0 & 0 & -1 & 3 \end{vmatrix}$$

Problema 9
Calcular el siguiente determinante, explicando el camino utilizado para llegar al resultado:

$$\begin{vmatrix} 1 & 0 & 0 & 0 \\ 2 & 7 & 0 & 0 \\ 3 & 23 & -1 & 0 \\ 43 & 237 & -32 & 1/7 \end{vmatrix}$$

Problema 10
Calcular el siguiente determinante, explicando el camino utilizado para llegar al resultado:

$$\begin{vmatrix} 1024 & 1025 & 2049 \\ 2 & 3 & 5 \\ -3 & -6 & -9 \end{vmatrix}$$

Problema 11
¿Cuánto vale el siguiente determinante? ¿por qué?

$$\begin{vmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 0 & 23 & 32 & 28 & 29 & 7 \\ -1 & 2 & -2 & 4 & -5 & 3 \\ 7 & 8 & 9 & 10 & 11 & 12 \\ 1 & 2 & 3 & 4 & 5 & 6 \\ -2 & 3 & -4 & 3 & 2 & 0 \end{vmatrix}$$

Problema 12
Calcular el siguiente determinante:

$$\begin{vmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 1 & 3 & 5 & 9 & 10 & 11 \\ 1 & 2 & 4 & 19 & -1 & 2 \\ 1 & 2 & 3 & 5 & 9 & 12 \\ 1 & 2 & 3 & 4 & 6 & 8 \\ 1 & 2 & 3 & 4 & 5 & 7 \end{vmatrix}$$

Problema 13
Calcular el siguiente determinante:

$$\begin{vmatrix} 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 \end{vmatrix}$$

Problema 14
Calcular el siguiente determinante:

$$\begin{vmatrix} 1 & 9 & 4 & 2 \\ 1 & 5 & 4 & 2 \\ 1 & 9 & 3 & 2 \\ 1 & 9 & 4 & 3 \end{vmatrix}$$

Problema 15
Calcular el siguiente determinante, explicando cómo se llega al resultado:

$$\begin{vmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 0 \\ 8 & 1 & 0 & 0 \\ 3 & 0 & 0 & 0 \end{vmatrix}$$

Problema 16
Calcular el siguiente determinante, explicando los pasos realizados:

$$\begin{vmatrix} 0 & 1 & 2 & 3 & 4 \\ 3 & 2 & 1 & 5 & 0 \\ 1 & 2 & 1 & 0 & 0 \\ 3 & 1 & 0 & 0 & 0 \\ 5 & 0 & 0 & 0 & 0 \end{vmatrix}$$

Problema 17
Calcular el siguiente determinante, explicando los pasos realizados:

$$\begin{vmatrix} 100 & 98 & 32 & 1 & 428 & 1 \\ 382 & 12 & 33 & 213 & 7 & 0 \\ 82 & 123 & 33 & 6 & 0 & 0 \\ 893 & 1 & 2 & 0 & 0 & 0 \\ 23 & 2 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \end{vmatrix}$$

Problema 18
¿Cuánto vale el siguiente determinante? ¿por qué?

$$\begin{vmatrix} 2 & 3 & 4 & 5 \\ 5 & 3 & 2 & 0 \\ 6 & 1 & 0 & 0 \end{vmatrix}$$

Problema 19
¿Cuánto vale el siguiente determinante? ¿por qué?

$$\begin{vmatrix} 1 & 1 & 1 & 1 & 1 \\ 2 & 3 & 4 & 5 & 0 \\ 2 & 2 & 2 & 2 & 2 \\ 4 & 8 & 9 & 0 & 0 \\ 7 & 3 & 2 & 1 & 0 \\ 0 & 1 & 2 & 4 & 33 \end{vmatrix}$$

3. Espacios vectoriales

Problema 20
Expresar unas ecuaciones paramétricas del subespacio F de $\mathbb{R}^5$ generado por los siguientes vectores:

$$v_1 = \begin{pmatrix} 1 \\ 2 \\ 0 \\ 0 \\ 1 \end{pmatrix}, v_2 = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 2 \end{pmatrix}, v_3 = \begin{pmatrix} 2 \\ 4 \\ 1 \\ 0 \\ 4 \end{pmatrix}.$$

¿Qué dimensión tiene F? Calcular una base de F. Dar una base de $\mathbb{R}^5$ que contenga la base de F calculada.

Problema 21
Se considera el sistema de ecuaciones siguiente. Decir si (y por qué) define un subespacio vectorial de $\mathbb{R}^7$. Calcular su dimensión y una base, si es que define un subespacio vectorial.

$$\begin{aligned} x_1 + 2x_2 - x_7 &= x_4 \\ 3x_1 + 4x_3 - x_5 &= x_6 - x_3 \\ x_1 - x_7 &= x_2 \end{aligned}$$

Problema 22
Calcular la dimensión del subespacio vectorial de $\mathbb{R}^6$ generado por los siguientes vectores:

$$v_1 = \begin{pmatrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \end{pmatrix}, v_2 = \begin{pmatrix} 2 \\ 1 \\ 2 \\ 3 \\ 4 \\ 5 \end{pmatrix}, v_3 = \begin{pmatrix} 5 \\ 1 \\ 2 \\ 3 \\ 2 \\ 1 \end{pmatrix}, v_4 = \begin{pmatrix} 5 \\ -2 \\ -2 \\ -2 \\ -4 \\ -6 \end{pmatrix}$$

Una vez calculada la dimensión, dar una base de dicho subespacio.

Problema 23
Se considera el subespacio vectorial de $\mathbb{R}^4$ dado por las *ecuaciones paramétricas*

$$\begin{aligned} x &= \lambda + \mu + \delta \\ y &= \lambda - \mu \\ z &= \lambda + 3\mu + 2\delta \\ u &= 2\lambda + \delta \end{aligned}$$

Se pide: calcular su dimensión. Dar unas coordenadas implícitas de dicho subespacio. Dar una base.

Problema 24
Se considera el subespacio de $\mathbb{R}^4$ dado por las ecuaciones implícitas

$$\begin{aligned} x + z &= 0 \\ 2y - z &= 0 \\ x - z + u &= 0 \\ 3x + 2y - 2z + 2u &= 0 \end{aligned}$$

Se pide: calcular su dimensión y una base del espacio. *Pasar* dichas ecuaciones implícitas a *paramétricas*

Problema 25
Calcular una base del subespacio F de los polinomios de grado menor o igual que 4 definido como sigue:

$$F = \{P(X) | P(0) = P(1) \text{ y } P(3) = 0\}.$$

¿Cuál es la dimensión de F? Dar unas ecuaciones implícitas y paramétricas (recuérdese que *para tener coordenadas y por tanto ecuaciones, hace falta una base*).

Problema 26
Calcular una base del subespacio F de $\mathbb{R}^5$ dado por los vectores

$$\begin{pmatrix}1\\2\\0\\1\\3\end{pmatrix}, \begin{pmatrix}-1\\3\\1\\2\\-1\end{pmatrix}, \begin{pmatrix}5\\0\\-2\\-1\\11\end{pmatrix}.$$

Dar unas ecuaciones paramétricas y unas implícitas de F. ¿Cuál es la dimensión de F?

Problema 27
Discutir el siguiente sistema de ecuaciones lineales. Es decir, dependiendo de los valores del parámetro a, decir si el sistema es compatible o no, indeterminado o no.

$$\begin{aligned} x + 2y + az &= 1 \\ ay + \ \ z &= 3 \\ x + ay + 2z &= 0 \end{aligned}$$

Problema 28
Discutir el siguiente sistema en función de los valores del parámetro a:

$$\begin{array}{rcl} x + 2y + az + \ \ t &=& 1 \\ ax \qquad\ \ + \ \ z \qquad &=& 0 \\ x \qquad\ \ + az \qquad &=& 1 \\ 2x + 4y \qquad\ \ + at &=& 3 \end{array}$$

Problema 29 (Hefferon pg. 129)
Dar una base del espacio de columnas de la siguiente matriz:

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

Problema 30 (Hefferon pg. 129)
Calcular el rango de cada matriz:

$$\begin{pmatrix} 2 & 1 & 3 \\ 1 & -1 & 2 \\ 1 & 0 & 3 \end{pmatrix}, \begin{pmatrix} 1 & -1 & 2 \\ 3 & -3 & 6 \\ -2 & 2 & -4 \end{pmatrix}, \begin{pmatrix} 1 & 3 & 2 \\ 5 & 1 & 1 \\ 6 & 4 & 3 \end{pmatrix}$$

Problema 31 (Hefferon pg. 129)
Hallar el rango por columnas de la siguiente matriz:

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

¿hace falta realmente "hacer algo"?

Problema 32 (Hefferon pg. 130)
Mostrar que el subespacio de $\mathbb{R}^4$ generado por los vectores $(1,-1,2,-3)$, $(1,1,2,0)$, $(3,-1,6,-6)$ es el mismo que el generado por los dos vectores $(1,0,1,0)$, $(0,2,0,3)$. Dar unas ecuaciones paramétricas y unas implícitas de dicho subespacio.

Problema 33
Sea F el subespacio de $\mathbb{R}^4$ generado por los vectores $(0,1,3,2)$, $(1,2,1,3)$ y $(-3,-4,3,-5)$. *Utilizar el algoritmo de Gauss* para encontrar una base de dicho subespacio vectorial. Dar ecuaciones implícitas y paramétricas.

Problema 34 (Hefferon pg. 126)
En el espacio de polinomios en una variable de grado menor o igual que 4, se considera el subespacio F generado por los polinomios x^2+x^4, $2x^2+3x^4$, $-x^2-3x^4$. Se pide calcular una base de dicho subespacio.

Problema 35 (Hefferon pg. 129)
Dígase si el vector está en el espacio por filas de la matriz correspondiente:

$$(\mathbf{a}) \begin{pmatrix} 2 & 1 \\ 3 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 \end{pmatrix} \quad (\mathbf{b}) \begin{pmatrix} 0 & 1 & 3 \\ -1 & 0 & 1 \\ -1 & 2 & 7 \end{pmatrix}, \begin{pmatrix} 1 & 1 & 1 \end{pmatrix}$$

Problema 36 (Hefferon pg. 122)
Calcúlese una base (y la dimensión) del espacio de soluciones del sistema

$$\begin{aligned} x_1 - 4x_2 + 3x_3 - x_4 &= 0 \\ 2x_1 - 8x_2 + 6x_3 - 2x_4 &= 0 \end{aligned}$$

Problema 37 (Hefferon pg. 122)
Dar bases de los siguientes subespacios (y sus dimensiones)

- El espacio de polinomios $p(x)$ de grado menor o igual a 3 tales que $p(7)=0$.

- El espacio de polinomios $p(x)$ de grado menor o igual a 3 tales que $p(7) = 0$ y $p(5) = 0$.
- El espacio de polinomios $p(x)$ de grado menor o igual a 3 tales que $p(7) = 0$, $p(5) = 0$ y $p(3) = 0$.
- El espacio de polinomios $p(x)$ de grado menor o igual a 3 tales que $p(7) = 0$, $p(5) = 0$, $p(3) = 0$ y $p(1) = 0$.

Problema 38 (Hefferon pg. 118)
Una matriz cuadrada se llama simétrica si para todos los índices i, j el elemento de lugar i, j es igual al elemento de lugar j, i. Se pide:

- Encontrar una base del espacio de matrices simétricas de tamaño 2×2.
- Encontrar una base del espacio de matrices simétricas de tamaño 3×3.

Problema 39 (Hefferon pg. 117)
Encontrar una base de los siguientes espacios:

- El espacio de vectores de longitud 3 cuyas primera y tercera coordenadas suman 0.
- El subespacio siguiente de las matrices 2×2:

$$\left\{ \begin{pmatrix} a & b \\ 0 & c \end{pmatrix} \middle| c - 2b = 0 \right\}$$

4. Aplicaciones lineales y matrices

Problema 40
Se considera la aplicación lineal $f : \mathbb{R}^3 \to \mathbb{R}^4$ dada en coordenadas en las bases estándar por las ecuaciones:

$$(x, y, z) \mapsto (x + y + z, x - y - z, x + z, y + z).$$

Se pide:

- Calcular dimensiones y bases de la imagen y el núcleo.
- Dar ecuaciones implícitas de la imagen y del núcleo.
- Calcular la matriz de f en las bases estándar.
- Calcular la matriz de f utilizando, en $\mathbb{R}^3$, la siguiente nueva base:

$$\begin{aligned} \bar{e}_1 &= (1, 0, 0)_{st} \\ \bar{e}_2 &= (1, 1, 0)_{st} \\ \bar{e}_3 &= (1, 0, -1)_{st} \end{aligned}$$

(donde $_{st}$ indica que las coordenadas son en la base estándar).

- Calcular la matriz de f utilizando la nueva base de $\mathbb{R}^3$ y la siguiente nueva base en $\mathbb{R}^4$:

$$\begin{aligned}\bar{f}_1 &= (0,1,0,0)_{st}\\ \bar{f}_2 &= (1,0,0,0)_{st}\\ \bar{f}_3 &= (0,0,1,-1)_{st}\\ \bar{f}_4 &= (0,0,0,1)_{st}\end{aligned}$$

Problema 41
Se considera la aplicación lineal f del espacio de polinomios de grado como mucho 3 al espacio de polinomios de grado como mucho 2 dada por:

$$P \mapsto P' + P(0),$$

es decir, la imagen de un polinomio es la suma de su derivada con su valor en 0. Se pide:

- Expresar f usando las siguientes bases:

$$\left\{1, T, T^2, T^3\right\} \text{ en } \mathcal{P}_3, \quad \left\{1, T, T^2\right\} \text{ en } \mathcal{P}_2.$$

- Calcular las dimensiones y dar bases de $\mathrm{Im}(f)$ y $\ker(f)$. Calcular ecuaciones implícitas de ambos subespacios.
- Cambiar de base la matriz calculada, para las nuevas bases:

$$\left\{1 + \frac{1}{2}T + \frac{1}{3}T^3, T + \frac{1}{2}T^2, \frac{1}{2}T^2, 1 + T + \frac{1}{2}T^2 + \frac{1}{3}T^3\right\} \text{ en } \mathcal{P}_3$$
$$\left\{1 + T + T^2, 1 + T, T\right\} \text{ en } \mathcal{P}_2.$$

Problema 42
Se considera el endomorfismo f de $\mathbb{R}^3$ dado por la proyección sobre el plano generado por los vectores $(1,1,1)$ y $(0,1,1)$ de la base estándar en la dirección $(1,1,0)$. Se pide:

- Calcular las ecuaciones de f en la base estándar.
- Calcular ecuaciones implícitas del núcleo y de la imagen. Dar bases de ambos subespacios.
- Calcular $f^{-1}(1,2,3)$. Calcular $f^{-1}(1,2,2)$.
- Cambiar de base utilizando como nueva base los vectores $(1,1,1)$, $(0,1,1)$ y $(1,1,0)$.
- ¿Es el endomorfismo diagonalizable? ¿Cuáles son los autovalores? ¿Por qué?

Problema 43
Se considera la aplicación lineal f de los polinomios de grado como mucho

5 a los polinomios de grado como mucho 3 dada por:

$$P(T) \overset{f}{\mapsto} P''(T)$$

(es decir, a un polinomio se le envía a su derivada segunda). Se pide:

- Calcular la matriz de f en las bases
$$\left\{1, T, T^2, T^3, T^4, T^5\right\},$$
$$\left\{1, T, T^2, T^3\right\}.$$
- Calcular dimensiones de la imagen y del núcleo, bases y ecuaciones implícitas de ambos subespacios.
- Calcular la contraimagen de $T + T^3$.
- ¿Qué quiere decir que el núcleo tenga dimensión 2?

Problema 44
Calcular la inversa de la siguiente matriz:

$$\begin{pmatrix} 1 & 1 & 2 \\ 1 & 1 & 3 \\ 1 & 2 & 2 \end{pmatrix}.$$

Problema 45
Decir si las siguientes matrices son diagonalizables:

$$\begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 0 \end{pmatrix}, \quad \begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & 1 \\ 3 & 1 & 0 \end{pmatrix}, \quad \begin{pmatrix} 1 & 1 & 2 & 0 \\ 1 & 1 & 2 & 0 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 2 & 1 \end{pmatrix}.$$

Si lo son, calcular los autovalores y una base de vectores propios.

Problema 46
¿Se pueden multiplicar las siguientes matrices? ¿Qué hace falta para poder multiplicar dos matrices?

$$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}, \begin{pmatrix} 7 & 8 & 9 \\ 10 & 11 & 12 \end{pmatrix}.$$

Problema 47
Se considera la aplicación f de $\mathcal{P}_2$ a $\mathcal{P}_3$ (los polinomios en una variable de grados como mucho 2 y 3, respectivamente) dada por:

$$f(a + bX + cX^2) = aX + \frac{b}{2}X^2 + \frac{c}{3}X^3.$$

Se pide:

- Decir qué aplicación es, en palabras.
- Comprobar que es lineal.
- Calcular la matriz asociada en las bases $\mathcal{B}_2 = (1, X, X^2)$ y $\mathcal{B}_3 = (1, X, X^2, X^3)$.
- Calcular las dimensiones, bases y ecuaciones implícitas de $\mathrm{Im}(f)$ y de $\ker(f)$.
- Cambiar de base a las nuevas:

$$\bar{\mathcal{B}}_2 = (1+X, 1+X^2, X), \;\; \bar{\mathcal{B}}_3 = (X^3, X^2, X, 1-X).$$

Problema 48

Calcular las inversas de la siguientes matrices (o explicar por qué no existen):

$$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}, \begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{pmatrix}, \begin{pmatrix} 0 & 1 & 2 & 0 \\ 1 & 0 & 1 & 2 \\ 0 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 \\ 2 & 0 & 0 & 0 & 1 \end{pmatrix}$$

Problema 49

Se consideran endomorfismos con las siguientes matrices asociadas (fijadas unas bases en los espacios que correspondan). Se pide: saber si cada endomorfismo es diagonalizable, decir qué valores propios tiene y encontrar una base de vectores propios, si es el caso.

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

Problema 50

Explicar por qué un endomorfismo $f : E \to E$ que en alguna base tiene la siguiente matriz asociada *no es diagonalizable*:

$$\begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix}.$$

Generalizar: si existen dos vectores v_1 y v_2 tales que $f(v_1) = \lambda v_1$ (i.e. v_1 es un vector propio de valor propio λ) y $f(v_2) = v_1 + \lambda v_2$, entonces f *no es diagonalizable*.

Problema 51
Dar ejemplos de que el rango de la suma de dos matrices puede ser: la suma de los rangos, menor que el menor de los rangos, estar entre el menor y el mayor... En fin, que puede ser cualquier cosa.

Problema 52
Dar ejemplos de endomorfismos en que la dimensión del núcleo es mayor que la de la imagen, en que es al revés y en que sean iguales.

Problema 53
Si $f : E \to E$ es un endomorfismo de un espacio vectorial de dimensión impar, comprobar que el núcleo no puede ser igual a la imagen.

Problema 54
Calcular la imagen y el núcleo (es decir, dimensiones y dar bases) del endomorfismo $f : \mathbb{R}^4 \to \mathbb{R}^4$ que, en las bases estándar tiene como matriz asociada la siguiente:

$$\begin{pmatrix} -1 & 0 & 0 & 1 \\ 1 & 0 & \frac{1}{2} & -1 \\ 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 1 \end{pmatrix}$$

Problema 55

Resolver las siguientes ecuaciones diferenciales. Indicar de qué tipo son.

$$y' + 2xy = \frac{1}{x} \tag{26}$$

$$y' - 2xy = \frac{1}{x} \tag{27}$$

$$y' + \frac{y}{x} = \cos(x) \tag{28}$$

$$y' - \frac{2}{x+1}y = (x+1)^3 \tag{29}$$

$$(1+x)y + (1-y)y' = 0 \tag{30}$$

$$y' = \frac{xy}{x^2 - y^2} \text{ (homog.)} \tag{31}$$

$$\frac{2x}{y^3} + \frac{y^2 - 3x^2}{y^4}y' = 0 \text{ (exacta)} \tag{32}$$

$$ydx - xdy = 0 \tag{33}$$

$$(y-x) + (y+x)y' = 0 \text{ (homog.)} \tag{34}$$

$$y' - a\frac{y}{x} = \frac{x+1}{x} \tag{35}$$

Bibliografía

1. J. Hefferon, *Linear algebra*, J. Hefferon, Available on the Internet, 2008, GPL licence.