Geometría I: Tema 1
Matrices, determinantes y sistemas de ecuaciones lineales

Juan de Dios Pérez

1. Ecuaciones lineales y sistemas.

A lo largo de todo el tema K va a denotar un cuerpo en general, aunque a partir de ahora cuando apliquemos nuestros resultados consideremos que o bien $K = \mathbb{R}$ o bien $K = \mathbb{C}$.
Llamaremos ecuación lineal en K a cualquier expresión de la forma

$$a_1x_1 + a_2x_2 + \cdots + a_nx_n = b, \tag{1.1}$$

donde $a_1, a_2, \ldots, a_n$ son elementos (conocidos y, por tanto, fijos) de K que llamaremos coeficientes de la ecuación lineal, b es también un elemento conocido de K, llamado el término independiente de la ecuación, mientras que $x_1, x_2, \ldots, x_n$ son símbolos que vamos a llamar las incógnitas de dicha ecuación lineal. Podemos escribir la ecuación anterior en forma compacta utilizando el símbolo del sumatorio, $\sum$, de la siguiente forma:

$$\sum_{i=1}^{n} a_ix_i = b.$$

Nótese que una ecuación lineal viene dada de la forma anterior, por lo que no pueden aparecer términos como x_i^n, siendo $n \in \mathbb{N}$, $n > 1$, ni x_ix_j, siendo $i, j \in \{1, \ldots, n\}$, ni ninguna aplicación de las incógnitas distinta de la identidad (como $\cos(x_i)$, $x_i^2 - x_j$, etc.). Así $3x_1 + ix_2 + \pi x_3 = 2$ es una ecuación lineal con coeficientes complejos (ó ecuación lineal en (ó sobre) $\mathbb{C}$), mientras que $3x_1 - ex_2 + \pi x_3 = 2$ es una ecuación lineal con coeficientes reales. Sin embargo $3e^{x_1} - ix_2 + \pi\cos(x_3) = 0$ no es una ecuación lineal (ni sobre $\mathbb{C}$ ni sobre $\mathbb{R}$).
Una solución de la ecuación lineal (1.1) es una elección de valores para las incógnitas (en el cuerpo que estemos considerando) que haga que en (1.1) se dé la igualdad; es decir, sería tomar n elemento en K, $c_1, c_2, \ldots, c_n$ tales que $a_1c_1 + a_2c_2 + \cdots + a_nc_n = b$.
Por ejemplo, si consideramos la ecuación anterior $3x_1 - ex_2 + \pi x_3 = 2$, los valores $x_1 = 1, x_2 = 0, x_3 = -\frac{1}{\pi}$ nos dan una solución para la ecuación. Asimismo, $x_1 = \frac{1}{3}, x_2 = -\frac{1}{e}, x_3 = 0$, nos dan otra solución de la misma ecuación.

Si consideramos un conjunto de m ecuaciones lineales todas con las mismas incógnitas sobre K:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \\ \quad \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m \end{cases}$$

lo vamos a llamar un sistema de m ecuaciones lineales con n incógnitas sobre el cuerpo K (a partir de ahora vamos a suponer que está claro sobre qué cuerpo consideramos las ecuaciones y no lo mencionaremos más).

Una solución del sistema de ecuaciones lineales anterior consiste en encontrar n elementos concretos del cuerpo K, $c_1, c_2, \ldots, c_n$ tales que al sustituir sus valores por las incógnitas de las ecuaciones del sistema nos den una solución de todas y cada una de ellas.

Llamaremos solución general del sistema al conjunto formado por todas las soluciones del sistema.

Dos sistemas de ecuaciones lineales con las mismas incógnitas se llamarán equivalentes si tienen la misma solución general, es decir, cuando tienen, exactamente, las mismas soluciones.

Un sistema de ecuaciones lineales puede tener una única solución, más de una solución o ninguna solución.

Ejemplos:

1. Si consideramos el sistema con coeficientes en $\mathbb{R}$

$$\begin{cases} x + y = 1 \\ x - y = 1 \end{cases}$$

   Tomando $x = 1$, $y = 0$, claramente obtenemos una solución. Si suponemos que c_1 y c_2 es otra solución tendríamos que $c_1 + c_2 = 1$ y $c_1 - c_2 = 1$. Si sumamos ambas expresiones tendríamos que $2c_1 = 2$ y, por tanto, $c_1 = 1$. Entonces tendríamos que $1 + c_2 = 1$ y, así, $c_2 = 0$. Es decir, este sistema admite una única solución.

2. Si ahora consideramos el sistema

$$\begin{cases} x_1 + x_2 + x_3 = 5 \\ x_1 + x_2 + x_3 = 3 \end{cases}$$

   podemos ver que no tiene solución. Si c_1, c_2, c_3 nos dieran una solución del sistema tendríamos que tener, a la vez, que $c_1 + c_2 + c_3 = 5$ y $c_1 + c_2 + c_3 = 3$. Restando ambas expresiones llegaríamos a que $0 = 2$, que es imposible.

3. El sistema de ecuaciones lineales

$$\begin{cases} x_1 + x_2 + x_3 = 2 \\ 2x_1 + 2x_2 + 2x_3 = 4 \end{cases}$$

   admite a $x_1 = 1$, $x_2 = 1$, $x_3 = 0$ como solución, pero también a $x_1 = 1$, $x_2 = 0$, $x_3 = 1$, es decir, tiene más de una solución. De hecho, $\{(\lambda, \mu, 2 - \lambda - \mu) \,/\, \lambda, \mu \in \mathbb{R}\}$ es la solución general de dicho sistema. (Nótese que como λ y μ pueden tomar cualquier valor real, de hecho aparecen infinitas soluciones).

Según el número de soluciones que admita un sistema de ecuaciones lineales podemos clasificar los sistemas de ecuaciones lineales en tres tipos: si el sistema no tiene ninguna solución, diremos que es incompatible. Por tanto, a un sistema que admita soluciones lo llamaremos compatible. Si, siendo compatible, el sistema solo tiene una solución, diremos que es compatible determinado, mientras que si admite más de una solución, lo llamaremos compatible indeterminado.
Así, el sistema del ejemplo 1 anterior sería compatible determinado, el del ejemplo 2 sería incompatible, y el del ejemplo 3 sería compatible indeterminado.

2. Método de Gauss-Jordan

Este método nos va a permitir resolver un sistema de ecuaciones lineales, transformándolo en otro equivalente (por tanto, con las mismas soluciones) pero más sencillo. Repitiendo el proceso, llegaremos a un sistema equivalente al primitivo y cuyas soluciones serán evidentes (si las tiene).
Para ello veremos que si dado un sistema de ecuaciones lineales, multiplicamos una de sus ecuaciones por un elemento no nulo del cuerpo, ó se intercambian dos ecuaciones entre si ó le sumamos a una de las ecuaciones otra multiplicada por un elemento del cuerpo, vamos a obtener un sistema de ecuaciones lineales equivalente al primitivo.

Claramente, si intercambiamos dos ecuaciones en un sistema vamos a obtener un sistema equivalente, pues recodemos que una solución del sistema ha de ser solución de todas y cada una de las ecuaciones que lo forman.

Que los sistemas

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ \quad\vdots \\ a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{in}x_n = b_i \\ \quad\vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m \end{cases} \tag{2.1}$$

y

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ \quad\vdots \\ ka_{i1}x_1 + ka_{i2}x_2 + \cdots + ka_{in}x_n = kb_i \\ \quad\vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m \end{cases} \tag{2.2}$$

donde $i \in \{1, \ldots, m\}$ y $k \in K \setminus \{0\}$, son equivalentes es inmediato: si $c_1, \ldots, c_n$ nos da una solución de (2.1) como, excepto la ecuación i-ésima, el resto de ecuaciones en (2.1) y (2.2) coinciden solo hemos de ver que $c_1, \ldots, c_n$ es solución de la i-ésima ecuación de (2.2). Por ser solución de (2.1) tenemos que $a_{i1}c_1 + a_{12}c_2 + \cdots + a_{in}c_n = b_i$. Pero como $k \neq 0$ este hecho es equivalente a que $ka_{i1}c_1 + ka_{i2}c_2 + \cdots + ka_{in}c_n = kb_i$ (recordad que todos los elementos que intervienen pertenecen al cuerpo K y en el cuerpo se verifican las propiedades distributivas del producto con respecto a la suma).

Finalmente consideremos los sistemas

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ \quad \vdots \\ a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{in}x_n = b_i \\ \quad \vdots \\ a_{j1}x_1 + a_{j2}x_2 + \cdots + a_{jn}x_n = b_j \\ \quad \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m \end{cases} \tag{2.3}$$

y

$$\begin{cases} \begin{array}{ccccc} a_{11}x_1 & +a_{12}x_2 & \ldots & +a_{1n}x_n & = b_1 \\ & \vdots & & & \\ (a_{i1} + ka_{j1})x_1 & +(a_{i2} + ka_{j2})x_2 & \ldots & +(a_{in} + ka_{jn})x_n & = b_i + kb_j \\ & \vdots & & & \\ a_{j1}x_1 & +a_{j2}x_2 & \ldots & +a_{jn}x_n & = b_j \\ & \vdots & & & \\ a_{m1}x_1 & +a_{m2}x_2 & \ldots & +a_{mn}x_n & = b_m \end{array} \end{cases} \tag{2.4}$$

donde $i \neq j$, $i, j \in \{1, \ldots, m\}$ y $k \in K$. Queremos ver que (2.3) y (2.4) son equivalentes. Supongamos, pues, que $x_1 = c_1, x_2 = c_2, \ldots, x_n = c_n$ es una solución de (2.3). Como todas las ecuaciones, excepto la i-ésima, de (2.3) y (2.4) son iguales, dicha solución es una solución para todas las ecuaciones de (2.4) excepto la i-ésima. Veamos que también lo es para esta. Sabemos que $a_{i1}c_1 + \cdots + a_{in}c_n = b_i$ y que $a_{j1}c_1 + \cdots + a_{jn}c_n = b_j$ (por ser solución de (2.3)). Entonces multiplicando esta segunda ecuación por k tendremos que $ka_{j1}c_1 + \ldots ka_{jn}c_n = kb_j$ y si sumamos ambas ecuaciones y aplicamos la conmutatividad de la suma en K tendremos que $a_{i1}c_1 + ka_{j1}c_1 + \cdots + a_{in}c_n + ka_{jn}c_n = b_i + kb_j$. Pero si ahora aplicamos n veces la distributividad del producto con respecto de la suma en K tendremos que $(a_{i1} + ka_{j1})c_1 + \cdots + (a_{in} + ka_{jn})c_n = b_i + kb_j$ y así, nuestra solución de (2.3) es también una solución de (2.4).
Recíprocamente, sea $x_1 = d_1, x_2 = d_2, \ldots, x_n = d_n$ una solución de (2.4). Tendremos que ver que es también una solución de (2.3). Como todas las ecuaciones de (2.4), excepto la i-ésima, son las mismas que las de (2.3) solo hemos de ver que nuestra solución es una solución para la i-ésima ecuación de (2.3).

Tenemos entonces que $d_1, d_2, \ldots, d_n$ verifican

$$(a_{i1} + ka_{j1})d_1 + \cdots + (a_{in} + ka_{jn})d_n = b_i + kb_j, \tag{2.5}$$

y que

$$a_{j1}d_1 + \cdots + a_{jn}d_n = b_j.$$

Esta segunda ecuación nos da (por la distributividad del producto con respecto a la suma en K) que $ka_{j1}d_1 + \cdots + ka_{jn}d_n = kb_j$. Si a (2.5) le restamos esta ecuación (teniendo en cuenta la conmutatividad de la suma de elementos en K) obtenemos $a_{i1}d_1 + \cdots + a_{in}d_n = b_i$, por lo cual nuestra solución también lo es para la ecuación i-ésima de (2.3) y terminamos la demostración.

Veamos cómo, a partir de lo que acabamos de ver podemos transformar un sistema de ecuaciones lineales en otro que va a ser el más simple posible y que nos permitirá saber de qué tipo es el sistema de partida. Lo haremos mediante un finito de pasos:

- ★ Paso 1: Ponemos en primer lugar una ecuación cuyo coeficiente para la incógnita x_1 sea distinto de cero.

- ★ Paso 2: Si el coeficiente de la incógnita x_1 en la primera ecuación es 1 pasamos al siguiente paso. Si es $\neq 1$, dividimos la ecuación por dicho coeficiente.

- ★ Paso 3: Como en la primera ecuación el coeficiente de x_1 es 1, eliminamos esta incógnita del resto de ecuaciones restándole a cada una de ellas el producto de la primera ecuación por los correspondientes coeficientes a_{j1}, para $j \in \{2, \ldots, m\}$. Conseguimos así que la incógnita x_1 solo aparezca en la primera ecuación, con lo que obtenemos, por lo visto anteriormente, un sistema equivalente al primitivo y de la forma:

$$\begin{cases} x_1 & +a_{12}x_2 & + \cdots + a_{1n}x_n = b_1 \\ & a_{22}x_2 & + \cdots + a_{2n}x_n = b_2 \\ & & \cdots \\ & a_{m2}x_2 & + \cdots + a_{mn}x_n = b_m \end{cases}$$

  (Tened en cuenta que escribimos los mismos a_{ij} y b_i que en el sistema primitivo simplemente por no complicar la notación aunque, en general, hayan cambiado).

A partir de ahora dejamos fija la primera ecuación y repetimos los pasos 1, 2 y 3 para el resto de ecuaciones y la incógnita x_2, obteniendo un sistema equivalente al anterior y de la forma

$$\begin{cases} x_1 & +a_{12}x_2 & +a_{13}x_3 & + \cdots + a_{1n}x_n = b_1 \\ & x_2 & +a_{23}x_3 & + \cdots + a_{2n}x_n = b_2 \\ & & \cdots & \\ & & +a_{m3}x_3 & + \cdots + a_{mn}x_n = b_m \end{cases}$$

Repetiremos el proceso mientras sea posible y obtendremos un sistema escalonado (la primera incógnita de cada ecuación tiene coeficiente 1 y no aparece en el resto e ecuaciones). Si en el proceso nos aparece una ecuación del tipo $0 = 0$, cualquier conjunto de elementos de K la va a verificar y, por tanto, podemos eliminarla. Es decir, llegaremos a un sistema de ecuaciones de la forma:

$$\begin{cases} x_1 & +a_{12}x_2 & +a_{13}x_3 & + \cdots + a_{1n}x_n = b_1 \\ & x_2 & +a_{23}x_3 & + \cdots + a_{2n}x_n = b_2 \\ & & \cdots & \\ & & x_r & + \cdots + a_{rn}x_n = b_r \end{cases}$$

Podemos hacer más simplificaciones. Vamos a llamar incógnitas principales a las que aparecen como primera incógnita de alguna de las ecuaciones de este último sistema e incógnitas secundarias a las restantes (si hay alguna). Ahora, aplicando de nuevo los pasos 1, 2 y 3 podemos eliminar la incógnita principal de una ecuación de las restantes y obtendremos así un sistema de ecuaciones escalonado reducido en el que cada incógnita principal no aparece en el resto de ecuaciones.

Una vez llegados a este punto podemos determinar, analizando cómo es este sistema escalonado reducido, de qué tipo es nuestro sistema original y, en el caso de que sea compatible, encontrar sus soluciones:

1. Puede ocurrir que en nuestro sistema compatible reducido nos aparezca una ecuación del tipo $0 = b$, para algún elemento del cuerpo K tal que $b \neq 0$. Como esta ecuación no admite solución, tampoco la tendrá el sistema de ecuaciones primitivo, por lo que será incompatible.

2. Si no aparece ninguna ecuación de tipo $0 = b$ con $b \in K \setminus \{0\}$, tendremos dos posibilidades:

   a) Todas las incógnitas son principales, con lo que nuestro sistema escalonado reducido será de la forma

   $$\begin{cases} x_1 & & & = b_1 \\ & x_2 & & = b_2 \\ & & \ddots & \\ & & & x_n = b_n \end{cases}$$

   y, por tanto, el sistema es compatible determinado y su única solución es, precisamente, $x_1 = b_1, x_2 = b_2, \ldots, x_n = b_n$.

   b) Si existen incógnitas secundarias el sistema tendrá la forma:

   $$\begin{cases} x_1 & & & +a_{1r+1}x_{r+1} + \cdots + a_{1n}x_n = b_1 \\ & x_2 & & +a_{2r+1}x_{r+1} + \cdots + a_{2n}x_n = b_2 \\ & & \ddots & \\ & & x_r & +a_{rr+1}x_{r+1} + \cdots + a_{rn}x_n = b_r \end{cases}$$

   En este caso, despejamos las incógnitas principales y obtenemos:

   $$\begin{cases} x_1 & & & = b_1 - a_{1r+1}x_{r+1} - \cdots - a_{1n}x_n \\ & x_2 & & = b_2 - a_{2r+1}x_{r+1} - \cdots - a_{2n}x_n \\ & & \ddots & \\ & & x_r & = b_r - a_{rr+1}x_{r+1} - \cdots - a_{rn}x_n \end{cases}$$

Fijaos en que cada vez que demos un valor arbitrario en K a las incógnitas secundarias $x_{r+1}, \ldots, x_n$, obtenemos un valor para cada una de las incógnitas principales y, por tanto, una solución del sistema. Así pues, en este caso nuestro sistema será compatible indeterminado y tendremos la solución general si le asignamos un parámetro (variará en todo K) a cada una de las incógnitas secundarias.

Nota 1. Si tenemos un sistema de ecuaciones lineales en el que todos los términos independientes son nulos, diremos que el sistema es homogéneo. Un sistema homogéneo en las variables $x_1, x_2, \ldots, x_n$ siempre es compatible, ya que $x_1 = 0, x_2 = 0, \ldots, x_n = 0$ nos proporciona una solución del mismo. En cada caso lo que tendremos que determinar es si es determinado o indeterminado.

Ejemplo 1.
Vamos a aplicar el método de Gauss-Jordan al siguiente sistema de ecuaciones lineales:

$$\begin{cases} x_1 & -2x_2 & +x_3 & = 7 \\ 2x_1 & -5x_2 & +2x_3 & = 6 \\ 3x_1 & +2x_2 & -x_3 & = 1 \end{cases}$$

Como en la primera ecuación el coeficiente de x_1 es 1, eliminaremos x_1 en las otras dos ecuaciones, sumándole a la 2ª el producto de -2 por la 1ª y sumándole a la 3ª el producto de -3 por la 1ª. Así, nuestro sistema será equivalente al siguiente:

$$\begin{cases} x_1 & -2x_2 & +x_3 & = 7 \\ & -x_2 & & = -8 \\ & 8x_2 & -4x_3 & = -20 \end{cases}$$

Como el coeficiente de x_2 en la 2ª ecuación es -1, multiplicamos esa segunda ecuación por -1 y nuestro sistema será equivalente al siguiente sistema:

$$\begin{cases} x_1 & -2x_2 & +x_3 & = 7 \\ & x_2 & & = 8 \\ & 8x_2 & -4x_3 & = -20 \end{cases}$$

Utilizamos ahora 2ª ecuación para eliminar x_2 de la 3ª ecuación; para ello, multiplicamos la 2ª ecuación por -8 y se la sumamos a la 3ª. Obtenemos como sistema equivalente:

$$\begin{cases} x_1 & -2x_2 & +x_3 & = 7 \\ & x_2 & & = 8 \\ & & -4x_3 & = -84 \end{cases}$$

Si dividimos la última ecuación entre -4, nuestro sistema es equivalente al sistema

$$\begin{cases} x_1 & -2x_2 & +x_3 & = 7 \\ & x_2 & & = 8 \\ & & x_3 & = 21 \end{cases}$$

Veamos ahora cómo hacer que este sistema, que ya es escalonado, sea escalonado reducido. Si a la 1ª ecuación le restamos la 3ª pasamos al sistema equivalente

$$\begin{cases} x_1 & -2x_2 & & = -14 \\ & x_2 & & = 8 \\ & & x_3 & = 21 \end{cases}$$

Y si ahora le sumamos a la 1ª ecuación el producto de la 2ª por 2 llegamos a

$$\begin{cases} x_1 & & & = 2 \\ & x_2 & & = 8 \\ & & x_3 & = 21 \end{cases}$$

Por lo tanto nuestro sistema es compatible determinado y su única solución es $x_1 = 2, x_2 = 8, x_3 = 21$.

Ejercicio 1.
Determinar por el método de Gauss-Jordan de qué tipo es el siguiente sistema de ecuaciones lineales y, si es posible, encontrar sus soluciones:

$$\begin{cases} x_1 & +x_2 & +x_2 & +x_4 & = 2 \\ x_1 & -x_2 & +x_3 & -x_4 & = -2 \\ 3x_1 & -x_2 & +3x_3 & -x_4 & = 2 \\ 7x_1 & -5x_2 & +7x_3 & -5x_4 & = -6 \end{cases}$$

3. Matrices y transformaciones elementales.

Consideremos un cuerpo K (como dijimos anteriormente, aunque lo que vamos a desarrollar servirá para un cuerpo arbitrario, consideraremos fundamentalmente los casos $K = \mathbb{C}$ ó $K = \mathbb{R}$). Una matriz de orden $m \times n$ con coeficiente en K (ó sobre K) es un conjunto de $m \cdot n$ elemento de K dispuestos en forma de caja con m filas y n columnas de la siguiente manera:

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}$$

Los elementos $a_{i1}, a_{i2}, \dots, a_{in}$, $\forall\, 1 \leq i \leq m$, constituyen la fila i-ésima de la matriz A y los elementos $a_{1j}, a_{2j}, \dots, a_{mj}$, $\forall\, 1 \leq j \leq n$, diremos que forman la columna j-ésima de dicha matriz. Por tanto, a_{ij}, $1 \leq i \leq m$, $1 \leq j \leq n$, denota el elemento que está en la fila i y la columna j de la matriz A. A veces, para simplificar la notación, escribiremos $A = (a_{ij})_{i,j}$, donde debemos entender que i varía en el conjunto de filas de A y j en el de columnas, es decir, $1 \leq i \leq m$, $1 \leq j \leq n$. Si está claro, no haremos esta puntualización.

Ejemplo 2 (Ejemplos de matrices).

- La matriz
$$\begin{pmatrix} 1 & -2 & -3 & 4 \\ 2 & 0 & 1 & 5 \\ 0 & 0 & -3 & 2 \end{pmatrix}$$
es una matriz de orden 3×4 con coeficientes reales.

- La matriz
$$\begin{pmatrix} i & 2+i \\ 0 & -3i \\ 4 & 5 \end{pmatrix}$$
es una matriz de orden 3×2 con coeficientes complejos.

En el caso de que una matriz tenga solo una fila, la llamaremos matriz fila y será de la forma
$$\begin{pmatrix} a_1 & a_2 & \dots & a_n \end{pmatrix}$$
Si una matriz tiene solo una columna, la llamaremos matriz columna y será de la forma
$$\begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}$$
Dos matrices serán iguales cuando tengan el mismo orden y en cada una de sus posiciones los mismos elementos. Así,
$$\begin{pmatrix} 3 & -2 & 0 \\ 1 & 5 & 2 \end{pmatrix} \text{ y } \begin{pmatrix} 3 & -2 \\ 1 & 5 \end{pmatrix}$$
no pueden ser iguales porque no tienen el mismo orden. Pero, aunque
$$\begin{pmatrix} 3 & -2 & 0 \\ 1 & 5 & 2 \end{pmatrix} \text{ y } \begin{pmatrix} 3 & -2 & 1 \\ 1 & 5 & 2 \end{pmatrix}$$
tienen el mismo orden no son iguales pues en el lugar $1,3$ de la primera aparece un 0 y en el mismo lugar de la segunda tenemos un 1.

Una matriz con el mismo número de filas que columnas, es decir, de orden $n \times n$, para algún número natural n, diremos que es una matriz cuadrada de orden n.

Vamos a denotar por $\mathcal{M}_{m\times n}(K) = \{A \,/\, A$ es una matriz de orden $m \times n$ con coeficientes en $K\}$. En el caso en que consideremos matrices cuadradas, es decir, cuando $m = n$, al conjunto de matrices cuadradas de orden m con coeficientes en K, lo vamos a notar $\mathcal{M}_m(K)$.

Si tenemos una matriz A y en ella suprimimos un cierto número de filas y/o columnas, la matriz que obtendremos la llamaremos submatriz de A. Así,
$$\begin{pmatrix} 1 & 2 \\ 0 & -1 \end{pmatrix}$$

es una submatriz de

$$A = \begin{pmatrix} 1 & 1 & 2 & 1 \\ 0 & -1 & -1 & 5 \\ -3 & 1 & 4 & 3 \end{pmatrix}$$

obtenida eliminando en A la 3ª fila y las columnas 2ª y 4ª. Asimismo,

$$\begin{pmatrix} -3 & 1 & 4 & 3 \end{pmatrix}$$

es una submatriz de A.

Supongamos que tenemos una matriz cuadrada $A = (a_{ij})_{i,j} \in \mathcal{M}_n(K)$ (en este caso, $1 \leq i, j \leq n$). Los elementos para los que $i = j$, es decir, $a_{11}, a_{22}, \ldots, a_{nn}$ forman la diagonal principal de A.

Diremos que una matriz cuadrada A es diagonal si todos sus elementos, excepto los de la diagonal principal, son nulos, es decir si $a_{ij} = 0$, $\forall\, 1 \leq i, j \leq n$, $i \neq j$.

Una matriz cuadrada A es triangular superior si todos los elementos que están debajo de los de su diagonal principal son nulos, es decir, si $a_{ij} = 0$, $\forall\, 1 \leq i, j \leq n$ con $j < i$.

Diremos que es triangular inferior si todos los elementos que están por encima de los de su diagonal principal son nulos; es decir, si $a_{ij} = 0$, $\forall\, 1 \leq i, j \leq n$, con $i < j$.

Ejemplo 3.

La matriz cuadrada de orden 3

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

es diagonal. Por otro lado, la siguiente matriz

$$B = \begin{pmatrix} 1 & 3 & 0 \\ 0 & -1 & 4 \\ 0 & 0 & 8 \end{pmatrix}$$

es triangular superior. Por último, la siguiente matriz es triangular inferior:

$$C = \begin{pmatrix} 4 & 0 & 0 \\ 5 & 1 & 0 \\ 7 & 8 & -2 \end{pmatrix}$$

La matriz identidad de orden n, que notaremos I_n, es la matriz diagonal tal que todos los elementos de su diagonal principal son iguales a 1. Esto es:

$$I_n = \begin{pmatrix} 1 & 0 & \ldots & 0 \\ 0 & 1 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & 1 \end{pmatrix}$$

Para $n = 1, 2, 3, 4$ la matriz identidad es, en cada caso:

$$I_1 = \begin{pmatrix} 1 \end{pmatrix}, \quad I_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad I_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad I_4 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Si utilizamos la función Delta de Kronecker, $\delta : \{1, \ldots, n\} \times \{1, \ldots, n\} \longrightarrow \{0, 1\}$ dada por

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

podemos escribir la matriz identidad $I_n = (\delta_{ij})_{i,j}$ (entendiendo que $1 \leq i, j \leq n$).
Si tenemos una matriz cuadrada $A = (a_{ij})_{i,j}$, definimos su traza como la suma de los elementos de su diagonal principal. Lo notaremos $\mathrm{tr}(A)$ y, así, $\mathrm{tr}(A) = a_{11} + a_{22} + \cdots + a_{nn}$.

Sea $A \in \mathcal{M}_{m \times n}(K)$. Llamaremos el pivote de una fila (ó una columna) de A al primer elemento no nulo de dicha fila (ó dicha columna), siempre que tal fila (ó columna) no esté compuesta únicamente de ceros. Se dice que la matriz A es escalonada por filas si verifica las siguientes condiciones:

1. Si A tiene filas cuyos elementos son todos nulos (filas nulas), dichas filas están todas en la parte inferior de la matriz.

2. El pivote de cada fila no nula es 1.

3. El pivote de cada fila no nula está a la derecha del pivote de la fila anterior.

4. Los elementos que aparecen en la misma columna que el pivote de una fila y debajo de él son todos nulos.

Diremos que A es escalonada reducida por filas si es escalonada por filas y, además, verifica

5. Los elementos que aparecen en la misma columna que el pivote de una fila son todos nulos.

Ejemplo 4.
La matriz

$$A = \begin{pmatrix} 3 & 1 & 0 & 5 \\ 0 & 1 & 1 & -2 \\ 0 & 0 & 1 & 3 \end{pmatrix}$$

no es escalonada por filas, puesto que el pivote de la 1ª fila no es 1. Por otro lado,

$$B = \begin{pmatrix} 1 & 1 & 0 & 5 \\ 0 & 1 & 1 & -2 \\ 0 & 1 & 1 & 3 \end{pmatrix}$$

no es escalonada por filas, ya que el elemento por debajo del pivote de la 2ª fila no es nula;

$$C = \begin{pmatrix} 1 & 1 & 0 & 5 \\ 0 & 1 & 1 & -2 \\ 0 & 0 & 1 & 3 \end{pmatrix}$$

es escalonada por filas, pero no es escalonada reducida por filas, porque los elementos encima de los pivotes de la 2ª y 3ª filas no son nulos. Finalmente,

$$D = \begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \end{pmatrix}$$

sí es escalonada reducida por filas.

Análogamente, podemos definir los conceptos de matriz escalonada por columnas y de matriz escalonada reducida por columnas: una matriz A es escalonada por columnas si verifica:

1'. Si A tiene columnas nulas (todos sus elementos son ceros), están a la derecha de la matriz.

2'. El pivote de cada columna no nula es 1.

3'. El pivote de cada columna no nula está más abajo que el de la columna anterior.

4'. Los elementos que aparecen en la misma fila en que se encuentra el pivote de una columna y a la derecha del mismo son todos nulos.

Si, además, se verifica que

5'. Los elementos en la misma fila en que se encuentra el pivote de una columna y distintos de él son todos nulos.

Entonces diremos que A es una matriz escalonada reducida por columnas.
Así, la matriz

$$E = \begin{pmatrix} 1 & 1 & 0 \\ -1 & 1 & 1 \\ 1 & 0 & 0 \\ 1 & 0 & 1 \end{pmatrix}$$

no es escalonada por columnas, pues el pivote de la segunda columna no está más abajo que el de la primera.

$$F = \begin{pmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 1 & 1 & 1 \\ 0 & 0 & 0 \end{pmatrix}$$

es escalonada por columnas, pero no es escalonada reducida por columnas, pues en las filas donde aparecen los pivotes de las columnas 2ª y 3ª aparecen elementos distintos de ellos no nulos. Sin embargo,

$$G = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

sí es escalonada reducida por columnas.

4. Forma normal de Hermite de una matriz

Intentaremos ahora obtener una matriz escalonada reducida a partir de una matriz arbitraria, utilizando cierto tipo de transformaciones. Para simplificar el lenguaje a cada elemento del cuerpo K que consideremos lo vamos a llamar un escalar.
Vamos a llamar transformaciones elementales por filas a cada una de los tipos siguientes:

1. Intercambiar la posición de dos filas en la matriz.

2. Multiplicar todos los elementos de una fila por un escalar no nulo.

3. Sumarle a una fila el producto de un escalar por otra fila (esto lo haremos sumando a cada elemento de la primera fila que consideremos el elemento de la segunda fila que esté en el mismo lugar multiplicado por el escalar).

Si $A, B \in \mathcal{M}_{m\times n}(K)$ diremos que A y B son equivalentes por filas, y lo escribiremos $A \sim_f B$ si podemos obtener B a partir de A mediante una sucesión (finita) de transformaciones elementales por filas. Notemos que el proceso inverso de cada transformación elemental vuelve a ser una transformación elemental (del mismo tipo), es fácil ver que esta relación es una relación de equivalencia sobre $\mathcal{M}_{m\times n}$ puesto que:

- $A \sim_f A$,

- $A \sim_f B$ si y solo si $B \sim_f A$,

- si $A \sim_f B$ y $B \sim_f C$ entonces $A \sim_f C$.

Tomemos entonces la situación siguiente:

Proposición 1.
Si A y $B \in \mathcal{M}_{m\times n}(K)$ son dos matrices escalonadas por filas, entonces si $A \sim_f B$ se tiene $A = B$.

Demostración. Para demostrar este hecho, dado que n es siempre un número natural, vamos a aplicar un proceso de inducción sobre n. (En este proceso, veremos que lo que queremos demostrar es cierto si $n = 1$, lo supondremos cierto para $n - 1 \in \mathbb{N}$, siendo $n \geq 2$, y a partir de este hecho lo demostraremos para el siguiente número natural n. Si el campo de variación de n no fuera el de los números naturales (por ejemplo, si n variara en $\mathbb{Q}$ ó en $\mathbb{R}$) no podríamos aplicar este método en nuestra demostración).

Así, en el caso de que $n = 1$, si $A = 0$, como A y B son equivalentes por filas, necesariamente $B = 0$. Si $A, B \neq 0$, como ambas son escalonadas reducidas por filas, tenemos una única posibilidad:

$$A = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} = B.$$

Así, vemos que si $n = 1$, nuestro resultado es cierto.
Supongamos, pues, que dicho resultado es cierto si consideramos $n-1$, siendo $n \in \mathbb{N}$, $n \geq 2$ y veamos que sigue siendo cierto para n. Sean A_1 y B_1 las submatrices de A y B que se obtienen quitándoles la última columna a A y B. Como $A \sim_f B$, es inmediato que $A_1 \sim_f B_1$. Además, como A_1 y B_1 siguen siendo escalonadas reducidas por filas y ahora el número de columnas tanto de A_1 como de B_1 es $n - 1$, por lo que hemos supuesto, tendríamos que $A_1 = B_1$. Por tanto, para ver que $A = B$ bastará con comprobar que sus últimas columnas son iguales.
Tenemos, entonces, dos posibilidades: la primera es que la última columna de A contenga un pivote y, como A es escalonada reducida por filas, necesariamente esa columna será

$$\begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix}$$

Pero como B es también escalonada reducida por filas, necesariamente su última columna tendrá también un pivote y, al ser $A \sim_f B$, esa última columna también ha de ser

$$\begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix}$$

y así, $A = B$.
La otra posibilidad es que las últimas columnas de A y B no contengan ningún pivote. Como $A \sim_f B$, A y B tendrán el mismo número de pivotes $r \leq n - 1$ y, por tanto, como $A_1 = B_1$, tendrán que estar en las mismas posiciones de las r primeras filas. Como, además, en la última columna de cada matriz solo pueden aparecer elementos distintos de 0 en las filas que contengan pivotes y estos están en las r primeras,

tendremos que las últimas columnas de A y B serán, respectivamente,

$$\begin{pmatrix} a_{1n} \\ \vdots \\ a_{rn} \\ \vdots \\ 0 \end{pmatrix} \quad y \quad \begin{pmatrix} b_{1n} \\ \vdots \\ b_{rn} \\ \vdots \\ 0 \end{pmatrix}$$

Como obtenemos B a partir de A por transformaciones elementales de filas, cada fila no nula de B será igual a una suma de las filas no nulas de A multiplicada cada una por un escalar (alguno de estos se puede anular); por tanto $\forall\, i \,/\, 1 \leq i \leq r$, tendremos

$$b_{ij} = k_1 a_{1j} + \cdots + k_r a_{rj}, \quad j \in \{1, \ldots, n\} \tag{4.1}$$

Si j_i es la columna que contiene al pivote de la fila i, entonces todos los elementos en esa columna, excepto el pivote $a_{ij_i} = b_{ij_i} = 1$, serán nulos y así, de la igualdad anterior, deducimos que

$$1 = b_{ij_i} = k_1 \cdot 0 + \cdots + k_i \cdot a_{ij_i} + \cdots + k_r \cdot 0 = k_i \cdot 1 = k_i$$

Además, para cada columna $j \neq j_i$ en la que se encuentre el pivote de alguna fila l $(1 \leq l \leq r)$ tendremos que a_{lj} y b_{lj} serán 1 (los pivotes) y el resto de elementos se anularán. Luego $\forall\, l \neq i$ tendremos que

$$0 = b_{ij} = k_1 \cdot 0 + \cdots + k_l \cdot a_{lj} + \cdots + k_r \cdot 0 = k_l \cdot 1 = k_l.$$

Así, hemos obtenido que $k_i = 1$ y que $k_l = 0$, $\forall\, l \neq i$, $1 \leq l \leq r$.
Si sustituimos esto es (4.1) obtenemos que $b_{in} = 0 \cdot a_{1n} + \cdots + 1 \cdot a_{in} + \cdots + 0 \cdot a_{rn} = a_{in}$, $\forall\, i \;/\, 1 \leq i \leq r$ y por tanto $A = B$. $\square$

A partir de este resultado podemos demostrar que cada matriz va a ser equivalente por filas a una única matriz escalonada reducida por filas.

Demostración. Para ello, ponemos como primera fila una fila cuyo primer elemento no sea nulo (si los primeros elementos de todas las filas son nulos, consideramos los segundos elementos y así sucesivamente). Si la nueva fila tiene como pivote el escalar a (que no será 0), la multiplicamos por a^{-1} y así conseguimos que el pivote sea ahora 1.
Entonces hacemos que los correspondientes elementos del resto de filas sean 0, sumándole a cada una de dichas filas el producto del opuesto de su primer elemento por la primera fila.
En el siguiente paso nos olvidamos de la primera fila y aplicamos lo anterior al resto de las filas de A tantas veces como sea necesario para obtener una matriz escalonada por filas. Finalmente, como el pivote de cada fila no nula es ahora 1, lo utilizamos para hacer 0 cada elemento de su columna que esté por encima de él y conseguir una matriz escalonada reducida por filas.
Como queremos que esta matriz sea única, si suponemos que A es equivalente por filas a dos matrices escalonadas reducidas por filas, digamos $A \sim_f H_1$ y $A \sim_f H_2$, como $\sim_f$ es una relación de equivalencia,

tendremos que $H_1 \sim_f A$ (hemos aplicado la propiedad simétrica) y $A \sim_f H_2$. Por tanto, la propiedad transitiva nos da $H_1 \sim_f H_2$ y, como ambas son escalonadas reducidas por filas, según lo visto anteriormente, $H_1 = H_2$, como queríamos demostrar. □

Si $A \in \mathcal{M}_{m\times n}(K)$, llamaremos forma normal de Hermite por filas de A a la única matriz escalonada reducida por filas que obtenemos a partir de A mediante transformaciones elementales por filas. (Pensad que podemos definir de manera análoga la forma normal de Hermite por columas de A, que existe y también es única).

Definición 1 (Rango de una matriz)**.**
Si $A \in \mathcal{M}_{m\times n}(K)$, vamos a llamar rango de A y lo vamos a notar $rg(A)$ al número de filas no nulas de la forma normal de Hermite por filas de A. Entonces tendremos que $rg(A) \leq \min\{m, n\}$, ya que por su propia definición $rg(A) \leq m$. Además, si $rg(A) = r$, A tiene r pivotes cada uno de los cuales ha de estar en una columna distinta de A y así, $rg(A) \leq n$.

Veamos, mediante un ejemplo, cómo calcular la forma normal de Hermite por filas de una matriz y su rango.
Sea

$$A = \begin{pmatrix} 2 & 1 & 3 & -2 \\ 2 & -1 & 5 & 2 \\ 1 & 1 & 1 & 1 \end{pmatrix}$$

Lo primero que hacemos es colocar como primera fila la tercera de A, pues su pivote es 1. Es decir

$$A \sim_f \begin{pmatrix} 1 & 1 & 1 & 1 \\ 2 & 1 & 3 & -2 \\ 2 & -1 & 5 & 2 \end{pmatrix}$$

y ahora con ese 1 hacemos nulos los primeros elementos de las filas 2ª y 3ª, sumándoles a cada una el producto de la primera fila y -2. Luego

$$A \sim_f \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & -1 & 1 & -4 \\ 0 & -3 & 3 & 0 \end{pmatrix}$$

Ahora mantendremos fija la 1ª fila y multiplicamos la 2ª por -1 para conseguir que su pivote sea 1. Por tanto,

$$A \sim_f \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & -1 & 4 \\ 0 & -3 & 3 & 0 \end{pmatrix}$$

A partir de esta matriz, haremos 0 el segundo elemento de la 3ª fila, sumándole el producto de la 2ª fila y 3, es decir, tendremos

$$A \sim_f \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & -1 & 4 \\ 0 & 0 & 0 & 12 \end{pmatrix}$$

Hacemos que el pivote de la 3ª fila sea 1, dividiéndola por 12 de modo que:

$$A \sim_f \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & -1 & 4 \\ 0 & 0 & 0 & 1 \end{pmatrix} (*)$$

y obtenemos una matriz escalonada por filas. Para obtener la forma normal de Hermite por filas utilizamos cada pivote para hacer nulos los elementos de su correspondiente columna y

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

Esta última matriz es la forma normal de Hermite por filas de A y, como el número de filas con pivote es 3, concluimos que $\mathrm{rg}(A) = 3$. Como el rango de una matriz es el número de filas no nulas de su forma normal de Hermite, no habría sido necesario llegar a esta última matriz, sino a una escalonada por filas equivalente por filas a la matriz original, es decir, la matriz $(*)$ (o la anterior) ya nos diría que el rango es 3.

5. Matrices y sistemas de ecuaciones.

Si tenemos un sistema de m ecuaciones lineales con n incógnitas sobre el cuerpo K

$$\begin{cases} a_{11}x_1 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + \cdots + a_{2n}x_n = b_2 \\ \quad \vdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n = b_m \end{cases}$$

llamaremos matriz de coeficientes de dicho sistema a la matriz $A \in \mathcal{M}_{m\times n}(K)$ dada por

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}$$

llamaremos matriz de términos independientes $B \in \mathcal{M}_{m\times 1}(K)$ a la matriz columna

$$B = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}$$

y llamaremos la matriz ampliada del sistema a la matriz $(A|B) \in \mathcal{M}_{m\times(n+1)}(K)$ dada por

$$(A|B) = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} & b_1 \\ a_{21} & a_{22} & \dots & a_{2n} & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} & b_m \end{pmatrix}$$

Por ejemplo, el sistema

$$\begin{cases} x_1 - 3x_2 - 2x_3 + 4x_4 = 5 \\ 3x_1 - 8x_2 - 3x_3 + 8x_4 = 18 \\ 2x_1 - 3x_2 + 5x_3 - 4x_4 = 19 \end{cases}$$

tiene a

$$A = \begin{pmatrix} 1 & -3 & -2 & 4 \\ 3 & -8 & -3 & 8 \\ 2 & -3 & 5 & -4 \end{pmatrix}$$

como matriz de coeficientes, a

$$B = \begin{pmatrix} 5 \\ 18 \\ 19 \end{pmatrix}$$

como matriz de términos independientes, y a

$$(A|B) = \begin{pmatrix} 1 & -3 & -2 & 4 & 5 \\ 3 & -8 & -3 & 8 & 18 \\ 2 & -3 & 5 & -4 & 19 \end{pmatrix}$$

como matriz ampliada.

Tenemos entonces que si consideramos un sistema de ecuaciones lineales cuya matriz ampliada es $(A|B)$, si H es la forma normal de Hermite por filas de esta matriz $(A|B)$, entonces el sistema de ecuaciones lineales cuya matriz ampliada es H es un sistema escalonado reducido equivalente al de partida. Para ver esto, como H se obtiene de $(A|B)$ mediante transformaciones elementales por filas, tendremos que ver que las transformaciones elementales por filas no afectan a la solución general del sistema, pero si aplicamos una transformación elemental de tipo 1 lo que hacemos es intercambiar dos ecuaciones (o varias), si aplicamos una transformación elemental del tipo 2, multiplicamos una ecuación por un escalar no nulo, y si lo que aplicamos es una transformación elemental de tipo 3, le estamos sumando a una ecuación el producto de un escalar por otra ecuación. Como en cada caso obtenemos un sistema de ecuaciones lineales equivalente al primitivo, el resultado es cierto.

Por ejemplo, para nuestro sistema anterior, cuya matriz ampliada es

$$A|B = \begin{pmatrix} 1 & -3 & -2 & 4 & 5 \\ 3 & -8 & -3 & 8 & 18 \\ 2 & -3 & 5 & -4 & 19 \end{pmatrix}$$

Su forma normal de Hermite por filas (calculadla siguiendo los pasos del ejemplo anterior) es:

$$H = \begin{pmatrix} 1 & 0 & 7 & -8 & 14 \\ 0 & 1 & 3 & -4 & 3 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

esto nos da que el sistema que consideramos es equivalente al sistema:

$$\begin{cases} x_1 & & +7x_3 & -8x_4 & = 14 \\ & x_2 & +3x_3 & -4x_4 & = 3 \end{cases}$$

Por tanto, es un sistema compatible indeterminado y su solución general será $x_1 = -7x_3 + 8x_4 + 14$, $x_2 = -3x_3 + 4x_4 + 3$ (así, cada vez que x_3 y x_4 tomen valores reales arbitrarios, tendremos una solución particular del sistema; por ejemplo, si $x_3 = x_4 = 0$, tendríamos que $x_1 = 14$, $x_2 = 3$; si $x_3 = 1$, $x_4 = 0$ entonces $x_1 = 7$ y $x_2 = 0$ serían otra solución).

Teorema 1 (Rouché-Frobenius)**.**
Dado un sistema de m ecuaciones lineales con n incógnitas sobre el cuerpo K, con matriz de coeficientes A y matriz ampliada $(A|B)$, tenemos:

1. *El sistema es compatible si y solo si $rg(A) = rg(A|B)$.*

2. *El sistema es compatible determinado si y solo si $rg(A) = rg(A|B) = n$.*

Demostración. Supongamos que H es la forma normal de Hermite por filas de $(A|B)$. Entonces la forma normal de Hermite por filas de A es la submatriz de H, H', que obtenemos eliminando la última columna de H. Sabemos que el sistema es compatible si y solo si en su forma escalonada reducida no aparece una ecuación como $0 = b$, con $b \in K \setminus \{0\}$, es decir, si H y H' tienen el mismo número de filas no nulas, lo que es equivalente a que $\mathrm{rg}(A) = \mathrm{rg}(A|B)$, lo que demuestra 1.
Si suponemos que $\mathrm{rg}(A) = \mathrm{rg}(A|B) = r$ (ya partimos de que el sistema es compatible), tendremos r incógnitas principales y sabemos que será compatible determinado cuando no haya incógnitas secundarias, esto es, si y solo si $r = n$, lo que demuestra 2. $\square$

6. Operaciones con matrices.

6.1. Suma de matrices.

Supongamos que $A, B \in \mathcal{M}_{m\times n}(K)$, $A = (a_{ij})_{i,j}$, $B = (b_{ij})_{i,j}$. Definimos su suma como la matriz $A + B \in \mathcal{M}_{m\times n}(K)$ dada por

$$A + B = (a_{ij} + b_{ij})_{i,j} = \begin{pmatrix} a_{11} + b_{11} & a_{12} + b_{12} & \dots & a_{1n} + b_{1n} \\ a_{21} + b_{21} & a_{22} + b_{22} & \dots & a_{2n} + b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} + b_{m1} & a_{m2} + b_{m2} & \dots & a_{mn} + b_{mn} \end{pmatrix}$$

Es decir, el elemento de $A+B$ que está en el lugar ij lo obtenemos sumando los elementos de A y de B que están en dicho lugar (tened en cuenta que si las matrices tienen órdenes distintos no las podremos sumar). Por ejemplo:

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

Hemos de señalar, de nuevo, que tenemos un abuso de notación: la suma de matrices es una operación distinta de la suma de los escalares que aparecen en cada matriz, aunque para simplificar la notación, ambas las vamos a escribir con el mismo signo $+$. Precisamente, a partir de las propiedades que tiene la suma en K, obtendremos las siguientes propiedades para la suma de matrices:

1. Asociativa: $\forall\, A, B, C \in \mathcal{M}_{m\times n}(K)$, $(A+B)+C = A+(B+C)$. Si notamos $A = (a_{ij})_{i,j}$, $B = (b_{ij})_{i,j}$, $C = (c_{ij})_{i,j}$ (recordad que $1 \leq i \leq m$, $1 \leq j \leq n$), tendremos que $(A+B)+C = (a_{ij}+b_{ij})_{i,j}+(c_{ij})_{i,j} = ((a_{ij}+b_{ij})+c_{ij})_{i,j} \overset{(*)}{=} (a_{ij}+(b_{ij}+c_{ij}))_{i,j} = (a_{ij})_{i,j}+(b_{ij}+c_{ij})_{i,j} = A+(B+C)$, donde en $*$ aplicamos la asociatividad de la suma en K.

2. Conmutativa: $\forall\, A, B \in \mathcal{M}_{m\times n}(K)$, $A+B = B+A$. Con la notación anterior, $A+B = (a_{ij}+b_{ij})_{i,j} \overset{(**)}{=} (b_{ij}+a_{ij})_{i,j} = B+A$; ahora $**$ es simplemente aplicar la propiedad conmutativa de la suma de K.

3. Existencia de elemento neutro: llamemos $0 \in \mathcal{M}_{m\times n}(K)$ tal que todos sus elementos son el $0 \in K$. Entonces $\forall\, A \in \mathcal{M}_{m\times n}(K)$, $A+0 = 0+A = A$. (Volvemos a abusar de la notación; la matriz 0 y el cero de K son elementos no comparables, aunque los notemos de la misma manera). Es claro que $A+0 = (a_{ij}+0)_{i,j} = (a_{ij})_{i,j} = A$. Como sabemos que la suma de matrices es conmutativa, no tenemos que demostrar que $0+A = A$.

4. Existencia de elementos simétricos: sea $A = (a_{ij})_{i,j} \in \mathcal{M}_{m\times n}(K)$. Llamemos $-A \in \mathcal{M}_{m\times n}(K)$ a la matriz dada por $-A = (-a_{ij})_{i,j}$; es decir, cada elemento de $-A$ es el opuesto en K del correspondiente elemento de A. Entonces $A+(-A) = (-A)+A = 0$. Como sabemos que la suma de matrices es conmutativa, nos bastará con ver una de las igualdades. Así, $A+(-A) = (a_{ij})_{i,j}+(-a_{ij})_{i,j} = (a_{ij}+(-a_{ij}))_{i,j} = (0)_{i,j} = 0$.

Hemos demostrado, pues, que $(\mathcal{M}_{m\times n}(K), +)$ es un grupo abeliano.

6.2. Producto de un escalar por una matriz.

Definición 2 (Producto de un escalar por una matriz)**.**
Sea $A = (a_{ij})_{i,j} \in \mathcal{M}_{m\times n}(K)$ y sea $\alpha \in K$ un escalar. Definimos el producto de α y A y lo notaremos $\alpha A \in \mathcal{M}_{m\times n}(K)$ como la matriz dada por

$$\alpha A = (\alpha a_{ij})_{i,j} = \begin{pmatrix} \alpha a_{11} & \alpha a_{12} & \dots & \alpha a_{1n} \\ \alpha a_{21} & \alpha a_{22} & \dots & \alpha a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \alpha a_{m1} & \alpha a_{m2} & \dots & \alpha a_{mn} \end{pmatrix}.$$

Es decir, en cada lugar ij el elemento de αA que ocupa dicho lugar es el producto por α del elemento que ocupa el lugar ij en A. Como estamos suponiendo que K es un cuerpo conmutativo ($\mathbb{R}$ ó $\mathbb{C}$), tendremos que $\alpha a_{ij} = a_{ij}\alpha$ y, por tanto, $\alpha A = A\alpha$.

Ejemplo 5 (Producto de un escalar por un matriz)**.**

$$3\begin{pmatrix} 1 & -2 & 5 \\ 0 & 1 & 3 \\ 1 & 1 & 1 \end{pmatrix} = \begin{pmatrix} 3 & -6 & 15 \\ 0 & 3 & 9 \\ 3 & 3 & 3 \end{pmatrix}.$$

De nuevo, teniendo en cuenta las propiedades del producto en K, este producto de escalares por matrices cumple las siguientes propiedades

1. Distributiva respecto de la suma de escalares: es decir, $\forall\, \alpha, \beta \in K$ y $\forall\, A = (a_{ij})_{i,j} \in \mathcal{M}_{m\times n}(K)$, tendremos que $(\alpha+\beta)A = \alpha A + \beta A$, dado que $(\alpha+\beta)A = (\alpha+\beta)(a_{ij})_{i,j} = ((\alpha+\beta)a_{ij})_{i,j} \overset{(*)}{=} (\alpha a_{ij} + \beta a_{ij})_{i,j} = (\alpha a_{ij})_{i,j} + (\beta a_{ij})_{i,j} = \alpha A + \beta A$, donde en $*$ hemos aplicado una de las propiedades distributivas del producto respecto de la suma en K.

2. Distributiva respecto de la suma de matrices: $\forall\, \alpha \in K$, $\forall\, A, B \in \mathcal{M}_{m\times n}(K)$, $\alpha(A+B) = \alpha A + \alpha B$: notemos $A = (a_{ij})_{i,j}$ y $B = (b_{ij})_{i,j}$. Entonces $\alpha(A+B) = \alpha(a_{ij}+b_{ij})_{i,j} = (\alpha(a_{ij}+b_{ij}))_{i,j} \overset{(**)}{=} (\alpha a_{ij} + \alpha b_{ij})_{i,j} = (\alpha a_{ij})_{i,j} + (\alpha b_{ij})_{i,j} = \alpha A + \alpha B$, donde en $**$ hemos aplicado la otra propiedad distributiva del producto respecto de la suma en K.

3. Pseudoasociativa: $\forall\, \alpha, \beta \in K$, $\forall\, A = (a_{ij})_{i,j} \in \mathcal{M}_{m\times n}(K)$, $(\alpha\beta)A = \alpha(\beta A)$, ya que $(\alpha\beta)A = ((\alpha\beta)a_{ij})_{i,j} \overset{(*)}{=} (\alpha(\beta a_{ij}))_{i,j} = \alpha(\beta a_{ij})_{i,j} = \alpha(\beta A)$, donde en $*$ aplicamos la asociatividad del producto en K.

4. Ley de identidad: $\forall\, A = (a_{ij})_{i,j} \in \mathcal{M}_{m\times n}(K)$, $1 \cdot A = 1 \cdot (a_{ij})_{i,j} = (1 \cdot a_{ij})_{i,j} = (a_{ij})_{i,j} = A$.

6.3. Producto de matrices.

Vamos a introducir un producto de matrices que no siempre podremos realizar. Dadas dos matrices $A \in \mathcal{M} \times \backslash(K)$, $B \in \mathcal{M}_{p\times q}(K)$, únicamente podremos definir AB si el número de columnas de A, $\underline{n}$, coincide con el número de filas de B, $\underline{p}$.
Sean, pues, $A = (a_{ik})_{i,k} \in \mathcal{M}_{m\times p}(K)$ y $B = (b_{kj})_{k,j} \in \mathcal{M}_{p\times n}(K)$. Definimos el producto de A y B, $AB = C = (c_{ij})_{i,j} \in \mathcal{M}_{m\times n}(K)$ tal que

$$c_{ij} = \sum_{k=1}^{p} a_{ik}b_{kj} = a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{ip}b_{pj}.$$

Observad que si A es de orden $m \times p$ y B es de orden $p \times n$, entonces AB es de orden $m \times n$. También que, aunque podamos definir AB, si m y n no coinciden, no podremos definir BA lo que implica que, en general, el producto de matrices no será conmutativo. Si $m = p = n$, es decir, si partimos de dos matrices cuadradas del mismo orden, $A, B \in \mathcal{M}_n(K)$, sí podremos considerar AB y BA, que volverán a ser matrices cuadradas de orden, pero incluso así, no podremos afirmar que, en general, $AB = BA$.

Ejemplo 6 (Producto de matrices)**.**
Si

$$A = \begin{pmatrix} 1 & -1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \text{y} \quad B = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 3 & 5 \\ 0 & 2 & 8 \end{pmatrix}$$

entonces

$$AB = \begin{pmatrix} 1 & -1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 3 & 5 \\ 0 & 2 & 8 \end{pmatrix} = \begin{pmatrix} 1 & -3 & -5 \\ 1 & 3 & 5 \\ 0 & 2 & 8 \end{pmatrix}$$

donde el elemento 11 del producto, rodeado por un círculo lo hemos obtenido considerando los elementos de la fila 1 de A y de la columna 1 de B (aquí es donde necesitamos que $n = p$ en nuestra definición) multiplicándolos uno a uno y sumando los resultados: $1 \cdot 1 + (-1) \cdot 0 + 0 \cdot 0 = 1$.
El 5 redondeado de la matriz producto, que está en el lugar 23, lo obtendríamos considerando la 2^{a} fila de A y la 3^{a} columna de B y repitiendo lo anterior: $1 \cdot 0 + 1 \cdot 5 + 0 \cdot 8 = 5$.
Si ahora hacemos BA obtenemos:

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 3 & 5 \\ 0 & 2 & 8 \end{pmatrix} \begin{pmatrix} 1 & -1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & -1 & 0 \\ 3 & 3 & 5 \\ 2 & 2 & 8 \end{pmatrix}$$

y comprobamos fácilmente que $AB \neq BA$.

Basándonos ahora en las propiedades del producto en K, tenemos las siguientes propiedades del producto de matrices:

1. Si $A \in \mathcal{M}_{m\times p}(K)$, $B \in \mathcal{M}_{p\times q}(K)$, $C \in \mathcal{M}_{q\times n}(K)$, entonces $A(BC) = (AB)C$.
   Para verlo, supongamos que $A = (a_{ik})_{i,k}$, $B = (b_{kl})_{k,l}$, $C = (c_{lj})_{l,j}$ con $1 \leq i \leq m$, $1 \leq k \leq p$, $1 \leq l \leq q$, $1 \leq j \leq n$.
   Entonces

$$A(BC) = (a_{ik})_{i,k} \left(\sum_l b_{kl} c_{lj} \right)_{k,j} = \left(\sum_k a_{ik} \left(\sum_l b_{kl} c_{lj} \right) \right)_{i,j} = \left(\sum_{k,l} a_{ik} b_{kl} c_{lj} \right)_{i,j} =$$

$$= \left(\sum_l \left(\sum_k a_{ik} b_{kl} \right) c_{lj} \right)_{i,j} = \left(\sum_k a_{ik} b_{kl} \right)_{i,l} (c_{lj})_{l,j} = (AB)C,$$

   (donde hemos aplicado la conmutatividad de la suma en K y la distributividad del producto con respecto de la suma).

2. Si $A \in \mathcal{M}_{m\times n}(K)$, $I_m A = A$, $AI_n = I_n$. Recordad que $I_m = (\delta_{ki})_{k,i}$, donde $\delta_{lm} = 0$ si $l \neq m$ y $\delta_{ll} = 1$ (la Delta de Kronecker que ya definimos) tendremos que el elemento en el lugar kj de $I_m A$

es

$$\sum_i \delta_{ki} a_{ij} = \delta_{k1} a_{ij} + \cdots + \delta_{kk} a_{kj} + \ldots \delta_{km} a_{mj} = a_{kj},$$

pues, excepto $\delta_{kk} = 1$, los demás δ_{ab} se anulan. Luego $I_m A = A$. La otra igualdad se demuestra de forma análoga.

3. Si $A \in \mathcal{M}_{m\times p}(K)$, $B, C \in \mathcal{M}_{p\times n}(K)$, tenemos que $A(B + C) = AB + AC$.
   Supongamos, pues, que $A = (a_{ik})_{i,k}$, $B = (b_{kj})_{k,j}$, $C = (c_{kj})_{k,j}$. Entonces

$$A(B+C) = (a_{ik})_{i,k}(b_{kj} + c_{kj})_{k,j} = \left(\sum_{k=1}^{p} a_{ik}(b_{kj} + c_{kj})\right)_{i,j} \overset{(*)}{=} \left(\sum_{k=1}^{p} (a_{ik}b_{kj} + a_{ik}c_{kj})\right)_{i,j} \overset{(**)}{=}$$

$$\overset{(**)}{=} \left(\left(\sum_{k=1}^{p} a_{ik}b_{kj}\right) + \left(\sum_{k=1}^{p} a_{ik}c_{kj}\right)\right)_{i,j} = \left(\sum_{k=1}^{p} a_{ik}b_{kj}\right)_{i,j} + \left(\sum_{k=1}^{p} a_{ik}c_{kj}\right)_{i,j} = AB + AC,$$

   donde en $*$ aplicamos la distributividad del producto con respecto de la suma en K y en $**$ la conmutatividad de la suma en K.

4. Si $A, B \in \mathcal{M}_{m\times p}(K)$, $C \in \mathcal{M}_{p\times n}(K)$, entonces $(A + B)C = AC + BC$. La demostración de esta propiedad se deja como **ejercicio** pues es completamente similar a la anterior.

5. $\forall\, A \in \mathcal{M}_{m\times p}(K)$, $\forall\, B \in \mathcal{M}_{p\times n}(K)$ y $\forall\, \alpha \in K$, se tiene que $\alpha(AB) = (\alpha A)B$. (Por lo que dijimos al introducir el producto de escalares por matrices también $\alpha(AB) = A(\alpha B)$, dado que suponemos que K es un cuerpo conmutativo).
   Si suponemos $A = (a_{ik})_{i,k}$, $B = (b_{kj})_{k,j}$, tendremos que

$$(\alpha A)B = (\alpha a_{ik})_{i,k}(b_{kj})_{k,j} = \left(\sum_{k=1}^{p} (\alpha a_{ik})b_{kj}\right)_{i,j} = \left(\sum_{k=1}^{p} \alpha(a_{ik}b_{kj})\right)_{i,j} = \left(\alpha \sum_{k=1}^{p} a_{ik}b_{kj}\right)_{i,j} = \alpha(AB).$$

6.4. Matriz traspuesta.

Sea $A = (a_{ij})_{i,j} \in \mathcal{M}_{m\times n}(K)$. Se llama matriz traspuesta de A, que denotamos A^t, a la matriz $A^t = (a_{ij})_{j,i}$, es decir, a la matriz que obtenemos a partir de A cuyas filas son las columnas de A y cuyas columnas son las filas de A. Así, si

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix} \in \mathcal{M}_{m\times n}(K),$$

entonces

$$A^t = \begin{pmatrix} a_{11} & a_{21} & \dots & a_{m1} \\ a_{12} & a_{22} & \dots & a_{m2} \\ \vdots & \vdots & \ddots & \vdots \\ a_{1n} & a_{2n} & \dots & a_{mn} \end{pmatrix} \in \mathcal{M}_{n\times m}(K).$$

Ejemplo 7 (Matriz traspuesta)**.**
Si

$$A = \begin{pmatrix} 1 & 2 & 3 & 0 \\ -1 & 5 & 1 & 1 \\ 0 & -3 & 0 & 4 \end{pmatrix},$$

entonces

$$A^t = \begin{pmatrix} 1 & -1 & 0 \\ 2 & 5 & -3 \\ 3 & 1 & 0 \\ 0 & 1 & 4 \end{pmatrix}.$$

Propiedades:

1. $(A + B)^t = A^t + B^t$, $\forall\, A, B \in \mathcal{M}_{m\times n}(K)$.

   Demostración. Si suponemos que $A = (a_{ij})_{i,j}$, $B = (b_{ij})_{i,j}$ con $1 \le i \le m$, $1 \le j \le n$ entonces:

   $$(A + B)^t = [(a_{ij} + b_{ij})_{i,j}]^t = (a_{ij} + b_{ij})_{j,i} = (a_{ij})_{j,i} + (b_{ij})_{j,i} = A^t + B^t.$$

   □

2. $(AB)^t = B^t A^t$, $\forall\, A \in \mathcal{M}_{m\times p}(K)$, $\forall\, B \in \mathcal{M}_{p\times n}(K)$.

   Demostración.

   $$B^t A^t = (b_{kj})_{j,k}(a_{ik})_{k,i} = \left(\sum_{k=1}^{n} b_{kj} a_{ik}\right)_{j,i} = \left(\sum_{k=1}^{n} a_{ik} b_{kj}\right)_{j,i} = \left(\left(\sum_{k=1}^{n} a_{ik} b_{kj}\right)_{i,j}\right)^t = (AB)^t.$$

   □

3. $(\alpha A)^t = \alpha A^t$, $\forall\, \alpha \in K$, $\forall\, A \in \mathcal{M}_{m\times n}(K)$.

Demostración.

$$(\alpha A)^t = ((\alpha a_{ij})_{i,j})^t = (\alpha a_{ij})_{j,i} = \alpha (a_{ij})_{j,i} = \alpha A^t.$$

□

Definición 3 (Matriz simétrica)**.**
Sea $A \in \mathcal{M}_n(K)$, diremos que es simétrica si $A^t = A$, luego $\forall\, i,j \in \{1,\ldots,n\}$ se verifica $a_{ij} = a_{ji}$.

Por ejemplo, la matriz

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

es simétrica mientras que la matriz $\begin{pmatrix} 1 & 1 \\ -1 & 0 \end{pmatrix}$ no lo es.

Diremos que A es antisimétrica si $A^t = -A$, es decir, si $\forall\, i,j \in \{1,\ldots,n\}$ se verifica $a_{ji} = -a_{ij}$. Nótese que, en particular, si $i = j$ tendremos que $a_{ii} = -a_{ii}$. Por tanto $a_{ii} = 0$, $\forall\, i \in \{1,\ldots,n\}$. Esto nos da una condición <u>necesaria</u> para que la matriz A sea antisimétrica, aunque no suficiente: la matriz

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

no es antisimétrica, aunque la matriz

$$B = \begin{pmatrix} 0 & -1 & 2 \\ 5 & 0 & 1 \\ -2 & -1 & 0 \end{pmatrix}$$

tampoco lo es. La matriz

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

sí es antisimétrica.

Sean $A, B \in \mathcal{M}_n(K)$. Entonces $\mathrm{tr}(A+B) = \mathrm{tr}(A) + \mathrm{tr}(B)$, ya que si $a_{11},\ldots,a_{nn}, b_{11},\ldots,b_{nn}$ son, respectivamente, los elementos de las diagonales principales de A y B, los elementos de la diagonal principal de $A+B$ son $a_{11}+b_{11}, a_{22}+b_{22},\ldots,a_{nn}+b_{nn}$. Luego $\mathrm{tr}(A+B) = a_{11}+b_{11}+a_{22}+b_{22}+\cdots+a_{nn}+b_{nn} \overset{(*)}{=} (a_{11}+a_{22}+\cdots+a_{nn})+(b_{11}+b_{22}+\cdots+b_{nn}) = \mathrm{tr}(A)+\mathrm{tr}(B)$, donde en $*$ hemos aplicado la conmutatividad y asociatividad de la suma en K.

Si, además, $k \in K$, $\mathrm{tr}(kA) = k\mathrm{tr}(A)$, ya que los elementos de la diagonal principal de kA son $ka_{11}, ka_{22},\ldots$, ka_{nn}. Así, $\mathrm{tr}(kA) = ka_{11}+ka_{22}+\cdots+ka_{nn} \overset{(**)}{=} k(a_{11}+a_{22}+\cdots+a_{nn}) = k\mathrm{tr}(A)$, donde en $**$ aplicamos la distributividad del producto respecto de la suma en K.

Además, si $A = (a_{ij})_{i,j}$ y $B = (b_{ij})_{i,j}$ son matrices cuadradas de orden n sobre K $(i, j \in \{1, \ldots, n\})$, se verifica que $\operatorname{tr}(AB) = \operatorname{tr}(BA)$.

Demostración. Para esta demostración, llamemos $C = BA$. Los elementos de la diagonal principal de C son:

$$c_{11} = a_{11}b_{11} + a_{12}b_{21} + \cdots + a_{1n}b_{n1},$$
$$\vdots$$
$$c_{nn} = a_{n1}b_{1n} + a_{n2}b_{2n} + \cdots + a_{nn}b_{nn}.$$

Por tanto,

$$\operatorname{tr}(AB) = \sum_{i=1}^{n} c_{ii} = \sum_{i,j=1}^{n} a_{ij}b_{ji} \overset{(*)}{=} \sum_{j,i=1}^{n} b_{ji}a_{ij} = \operatorname{tr}(BA).$$

Fijaos que en $*$ necesitamos la hipótesis de que el cuerpo sea conmutativo. □

7. Matrices regulares.

Vamos a ver cómo aplicar las transformaciones elementales a una matriz se puede considerar como el resultado de multiplicar dicha matriz por ciertas matrices. Llamaremos matrices elementales de orden n a las matrices que obtenemos al aplicarle una <u>única</u> transformación elemental por filas a la matriz identidad de orden n, I_n. Como hay tres tipos de transformaciones elementales por filas, tendremos tres tipos de matrices elementales:

- <u>Matrices elementales de tipo I:</u> llamaremos E_{ij} a la matriz que obtenemos de I_n intercambiando las filas i-ésima y j-ésima, $i, j \in \{1, \ldots, n\}$, $i \neq j$. Así,

$$E_{ij} = \begin{array}{r} \\ \\ \text{fila } i \\ \\ \text{fila } j \\ \\ \\ \end{array} \begin{pmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & 0 & \ldots & 1 & & \\ & & \vdots & \ddots & \vdots & & \\ & & 1 & \ldots & 0 & & \\ & & & & & \ddots & \\ & & & & & & 1 \end{pmatrix}.$$

- <u>Matrices elementales de tipo II:</u> llamamos $E_i(k)$ a la matriz que se obtiene de la matriz I_n al multiplicar por $k \in K \setminus \{0\}$ la fila i-ésima. Luego:

$$E_i(k) = \begin{array}{r} \\ \\ \text{fila } i \\ \\ \\ \end{array} \begin{pmatrix} 1 & & & & \\ & \ddots & & & \\ & & k & & \\ & & & \ddots & \\ & & & & 1 \end{pmatrix}.$$

- Matrices elementales de tipo III: notaremos $E_{ij}(k)$ a la matriz que se obtiene de I_n al sumarle a la fila i-ésima el producto por k de la fila j-ésima, $i, j \in \{1, \dots, n\}$, $i \neq j$. Por tanto,

$$E_{ij}(k) = \begin{array}{r} \\ \\ \text{fila } i \\ \\ \text{fila } j \\ \\ \\ \end{array} \begin{pmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & 1 & \dots & k & & \\ & & & \ddots & \vdots & & \\ & & & & 1 & & \\ & & & & & \ddots & \\ & & & & & & 1 \end{pmatrix}.$$

Nótese que estas matrices también se pueden obtener de I_n aplicándole transformaciones elementales por columnas. De hecho, si en I_n se intercambian las columnas i-ésima y j-ésima se obtiene E_{ij}, si se multiplica por $k \in K \setminus \{0\}$ la columna i-ésima se obtiene $E_i(k)$ y si se le suma a la columna j-ésima el producto por k de la columna i-ésima se obtiene $E_{ij}(k)$.
Así, en $\mathcal{M}_4(\mathbb{R})$ tenemos:

$$E_{14} = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}, \quad E_2(5) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 5 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \quad E_{13}(2) = \begin{pmatrix} 1 & 0 & 2 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Es inmediato ver que si $A \in \mathcal{M}_{m \times n}(K)$, E es una matriz elemental de orden m y F es una matriz elemental de orden n, entonces:

1. EA es la matriz que se obtiene de A aplicándole a sus filas la misma transformación elemental que nos permite obtener E a partir de I_m.

2. AF es la matriz que se obtiene de A al aplicarle a sus columnas la misma transformación elemental que nos da F a partir de I_n.

Por tanto si $A \in \mathcal{M}_{m \times n}(K)$ y H es la forma normal de Hermite por filas de A, mientras que C es la forma normal de Hermite por columnas de A, tendremos:

1. $H = E_k E_{k-1} \dots E_1 A$ para ciertas matrices elementales de orden m, $E_1, E_2, \dots, E_k$.

2. $C = A F_1 F_2 \dots F_s$, para ciertas matrices elementales de orden n, $F_1, F_2, \dots, F_s$.

8. Matriz inversa.

Sean dos matrices $A, B \in \mathcal{M}_n(K)$. Diremos que B es una matriz inversa de A si $AB = BA = I_n$. Si nos fijamos en la matriz $A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ no tiene una matriz inversa, porque sea cual sea la matriz

$B = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathcal{M}_2(K)$, el producto

$$\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} c & d \\ 0 & 0 \end{pmatrix}$$

y el 0 que aparece redondeado nos da que esta matriz nunca es la matriz identidad.

Así, <u>no toda matriz cuadrada admite una matriz inversa</u>. Diremos que la matriz $A \in \mathcal{M}_n(K)$ es invertible si existe una matriz inversa de A. Veamos que si A es invertible solo podemos encontrar una única matriz inversa para A, pues si suponemos que B y C son matrices inversas de A, se verificaría que $AB = BA = I_n$ y que $AC = CA = I_n$. Entonces $B = BI_n = B(AC) = (BA)C = I_nC = C$.

Luego si $A \in \mathcal{M}_n(K)$ es una matriz invertible, tiene una única matriz inversa que notaremos A^{-1} (nótese que A^{-1} también será invertible con $(A^{-1})^{-1} = A$).

Por ejemplo, la matriz $\begin{pmatrix} 1 & 1 \\ 2 & 3 \end{pmatrix}$ es invertible y su inversa es $A^{-1} = \begin{pmatrix} 3 & -1 \\ -2 & 1 \end{pmatrix}$ pues

$$\begin{pmatrix} 1 & 1 \\ 2 & 3 \end{pmatrix} \begin{pmatrix} 3 & -1 \\ -2 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad \text{y} \quad \begin{pmatrix} 3 & -1 \\ -2 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 2 & 3 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Sean $A, B \in \mathcal{M}_n(K)$. Se va a verificar que:

1. Si A y B son invertibles, AB es invertible y $(AB)^{-1} = B^{-1}A^{-1}$.
2. Si $A_1, \ldots, A_m \in \mathcal{M}_n(K)$ son invertibles, $A_1 \ldots A_m$ es invertible y $(A_1 \ldots A_m)^{-1} = A_m^{-1} \ldots A_1^{-1}$.
3. Si A es invertible, entonces A^t es invertible y $(A^t)^{-1} = (A^{-1})^t$.

Demostración.

1. $(AB)(B^{-1}A^{-1}) = A(BB^{-1})A^{-1} = (AI)A^{-1} = AA^{-1} = I_n$ y $(B^{-1}A^{-1})(AB) = B^{-1}(A^{-1}A)B = B^{-1}(I_nB) = B^{-1}B = I_n$.
2. Para demostrar esta propiedad utilizaremos una inducción sobre el índice m. El caso $m = 2$ lo hemos visto en el punto 1. Supongamos que el resultado es cierto para $m - 1 \in \mathbb{N}$ con $m \geq 4$, es decir que $A_1 \ldots A_{m-1}$ es invertible y $(A_1 \ldots A_{m-1})^{-1} = A_{m-1}^{-1} \ldots A_1^{-1}$. Sean entonces $A_1, \ldots, A_m \in \mathcal{M}_n(K)$ invertibles. Como, por hipótesis de inducción, $A_1 \ldots A_{m-1}$ es invertible y A_m también lo es entonces $(A_1 \ldots A_{m-1})A_m = A_1 \ldots A_{m-1}A_m$ es invertible. Además, por el punto 1, $(A_1 \ldots A_{m-1}A_m)^{-1} = ((A_1 \ldots A_{m-1})A_m)^{-1} = A_m^{-1}(A_1 \ldots A_{m-1})^{-1} \overset{(*)}{=} A_m^{-1}A_{m-1}^{-1} \ldots A_1^{-1}$, donde en $*$ hemos aplicado la hipótesis de inducción.
3. Para demostrar esta propiedad, si suponemos que A es invertible y tomamos $A^t(A^{-1})^t = (A^{-1}A)^t = I_n^t = I_n$ y $(A^{-1})^tA^t = (AA^{-1})^t = I_n^t = I_n$, lo que nos permite afirmar que A^t es invertible y tiene como inversa a $(A^{-1})^t$.

□

Es inmediato ver que cada matriz elemental es invertible y que su inversa es otra matriz elemental, pues $E_{ij}E_{ij} = I_n$, y así $E_{ij}^{-1} = E_{ij}$, $E_i(k)E_i\left(\frac{1}{k}\right) = E_i\left(\frac{1}{k}\right)E_i(k) = I_n$, luego $\forall\, k \in K \setminus \{0\}$ se tiene $(E_i(k))^{-1} = E_i(k^{-1})$ y también se tiene que $(E_{ij}(k))^{-1} = E_{ij}(-k)$, $\forall\, k \in K$.
A partir de estos hechos tenemos el siguiente resultado:

Teorema 2.
Sea $A \in \mathcal{M}_n(K)$. Entonces los siguientes hechos son equivalentes entre sí:

a) *A es invertible.*

b) *A es regular a la derecha (si $B \in \mathcal{M}_{m\times n}(K)$ tal que $BA = 0$, entonces $B = 0$).*

b') *A es regular a la izquierda (si $B \in \mathcal{M}_{n\times p}(K)$ tal que $AB = 0$, entonces $B = 0$).*

c) *$rg(A) = n$.*

d) *La forma normal de Hermite por filas de A es I_n.*

d') *La forma normal de Hermite por columnas de A es I_n.*

e) *A es un producto de matrices elementales.*

Demostración. Para ver que todas estas condiciones son equivalentes veremos que $a) \Longrightarrow b) \Longrightarrow c) \Longrightarrow d) \Longrightarrow e) \Longrightarrow a)$ con lo que si se da alguna de ellas se da cualquiera de ellas.
Como el hecho de que A es invertible es equivalente a que lo sea A^t, a partir de lo que veremos en las implicaciones anteriores se obtiene que $a) \Longrightarrow b') \Longrightarrow d') \Longrightarrow e)$.
Empezamos, por tanto:

- $a) \Longrightarrow b)$: si A es invertible y $BA = 0$, multiplicamos esa igualdad por A^{-1} por la derecha y tendremos que $0A^{-1} = 0 = (BA)A^{-1} = B(AA^{-1}) = BI_n = B$.

- $b) \Longrightarrow c)$: lo haremos demostrando el contrarrecíproco (es decir, supondremos que $\mathrm{rg}(A) < n$ y entonces encontraremos una matriz $D \neq 0$ tal que $DA = 0$).
  Si $\mathrm{rg}(A) < n$ y H es la forma normal de Hermite por filas de A, nos aseguramos que la última fila de H es nula. Sea entonces $D \in \mathcal{M}_{m\times n}(K)$ dada por:
  $$D = \begin{pmatrix} 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & \dots & 0 & 1 \end{pmatrix}.$$
  El único elemento no nulo de D es el 1 en la posición mn. Como la última fila de H es nula claramente $DH = 0$. Como $H = E_k \dots E_1 A$ para ciertas matrices elementales $E_1, \dots, E_k$, tendremos $0 = DH = D(E_k \dots E_1 A) = (DE_k \dots E_1)A$ y así, la matriz $DE_k \dots E_1$ no es nula (al no serlo D, si suponemos que $DE_k \dots E_1 = 0$ entonces $DE_k \dots E_1 E_1^{-1} \dots E_k^{-1} = D = 0E_1^{-1} \dots E_k^{-1} = 0$), lo que nos da una contradicción.

- $c) \implies d)$: Si el rango de A es n, la forma normal de Hermite por filas H de A es una matriz escalonada reducida, de orden $n \times n$ y con n pivotes que son 1 y cada uno situado a la derecha del anterior. Así, necesariamente, $H = I_n$.
- $d) \implies e)$: si $H = I_n$, sabemos que $\exists\, E_1, \dots, E_k$ matrices elementales, tales que $H = I_n = E_k \dots E_1 A$. Como cada matriz elemental es invertible $E_1^{-1} \dots E_k^{-1} = E_1^{-1} \dots E_k^{-1} I_n = E_1^{-1} \dots E_k^{-1} E_k \dots E_1 A = A$, y como la inversa de cada matriz elemental es otra matriz elemental, tenemos el resultado.
- $e) \implies a)$: es inmediato, pues si A es un producto de matrices elementales y cada una de ellas es invertible, lo es A.

□

Por este Teorema a una matriz que sea invertible la vamos a llamar también matriz regular. Si una matriz no es regular, la llamaremos matriz singular.
Recordad que para que B sea inversa de A exigíamos que $AB = BA = I_n$. Veamos que basta con una de las igualdades para que A sea regular. Es decir, si $A, B \in \mathcal{M}_n(K)$ verifican que $AB = I_n$, entonces A es invertible y $B = A^{-1}$.
Por el Teorema anterior, bastará ver que se verifica la condición $b)$ (que es equivalente a la nuestra). Supongamos, entonces que $X \in \mathcal{M}_{m\times n}(K)$ tal que $XA = 0$. Entonces, $X = XI_n = X(AB) = (XA)B = 0B = 0$, con lo que A es invertible. Supongamos que A^{-1} es su inversa. Entonces como $AB = I_n$, $A^{-1}(AB) = A^{-1}I_n = A^{-1} = (A^{-1}A)B = I_nB = B$. Luego, efectivamente $B = A^{-1}$.

Otra consecuencia del Teorema anterior es que si $A \in \mathcal{M}_{m\times n}(K)$ y H es la forma normal de Hermite por filas de A, podemos encontrar una matriz regular $Q \in \mathcal{M}_m(K)$ tal que $H = QA$, ya que hemos visto que $H = E_k \dots E_1 A$ para ciertas matrices elementales $E_1, \dots, E_k$. Si llamamos $Q = E_k \dots E_1$, como cada E_i, $1 \le i \le k$, es regular, Q será una matriz regular y, efectivamente, $H = QA$.
La matriz Q es fácil de calcular, como $Q = E_k \dots E_1 = E_k \dots E_1 I_m$, Q se obtiene aplicando a I_m las mismas operaciones elementales que hemos de aplicar a A para obtener H. Así, tomamos la matriz

$$(A|I_m) = \left(\begin{array}{cccc|cccc} a_{11} & a_{12} & \dots & a_{1n} & 1 & 0 & \dots & 0 \\ a_{21} & a_{22} & \dots & a_{2n} & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} & 0 & 0 & \dots & 1 \end{array}\right)$$

y aplicando transformaciones elementales por filas a esta matriz, llegaremos a la matriz $(H|Q)$.
En particular, si A es regular, sabemos que su forma normal de Hermite por filas es I_n. En este caso, la matriz Q anterior verificará que $I_n = QA$, con lo que $Q = A^{-1}$. Si aplicamos lo anterior a $(A|I_n)$ llegaremos a $(I_n|Q = A^{-1})$ y podremos calcular la inversa de A. Apliquémoslo a la matriz

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 4 \\ 3 & 4 & 6 \end{pmatrix}.$$

Tenemos

$$(A|I) = \left(\begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 2 & 3 & 4 & 0 & 1 & 0 \\ 3 & 4 & 6 & 0 & 0 & 1 \end{array}\right)$$

Utilizando el pivote 1 de la 1ª fila haremos nulos los primeros elementos de la 2ª y 3ª filas, obteniendo:

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

Cambiamos de signo las filas 2ª y 3ª y tenemos la matriz:

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

Ahora con el pivote de la 2ª fila hacemos cero el segundo elemento de la tercera fila y obtenemos:

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

Si cambiamos de signo la última fila pasamos a

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

(Ya sabemos que el $\mathrm{rg}(A) = 3$ y que es regular). Con el pivote de la 3ª fila hacemos nulos los elementos de la 3ª columna que está sobre él

$$\left(\begin{array}{ccc|ccc} 1 & 2 & 0 & -2 & 6 & -3 \\ 0 & 1 & 0 & 0 & 3 & -2 \\ 0 & 0 & 1 & 1 & -2 & 1 \end{array}\right)$$

y, finalmente, con el pivote de la 2ª fila anulamos el 2 que hay sobre él obteniendo:

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

de modo que concluimos $A^{-1} = \begin{pmatrix} -2 & 0 & 1 \\ 0 & 3 & -2 \\ 1 & -2 & 1 \end{pmatrix}$.

9. Matrices equivalentes.

Recordad que dos matrices $A, B \in \mathcal{M}_{m\times n}(K)$ son equivalentes por filas ($A \sim_f B$) si podemos pasar de A a B mediante una sucesión finita de transformaciones elementales por filas y que son equivalentes por columnas ($A \sim_c B$) si podemos pasar de A a B mediante una sucesión finita de transformaciones elementales por columnas. Para la equivalencia por filas, tenemos que si $A, B \in \mathcal{M}_{m\times n}(K)$, los siguientes hechos son equivalentes:

1. $A \sim_f B$,
2. A y B tienen la misma forma normal de Hermite por filas,
3. $\exists\, Q \in \mathcal{M}_m(K)$, <u>regular</u>, tal que $B = QA$.

Demostración. Si llamamos H_1 a la forma normal de Hermite por filas de A y H_2 a la de B, sabemos que $A \sim_f H_1$ y que $B \sim_f H_2$.
Vamos a demostrar que 1.) $\Longleftrightarrow$ 2.): tenemos que $A \sim_f B$ si y solo si $H_1 \sim_f H_2$ (la relación $\sim_f$ es de equivalencia) lo que es equivalente a que $H_1 = H_2$, ya que ambas matrices son escalonadas reducidas por filas y la igualdad se da por un resultado ya visto.
Veamos ahora que 1.) $\Longleftrightarrow$ 3.) : tenemos que $A \sim_f B$ si y solo si obtenemos B a partir de A mediante una sucesión finita de transformaciones elementales por filas, lo que es equivalente a que existan matrices elementales $E_1, \dots, E_k$, tales que $B = E_k \dots E_1 A$, y esto equivale a que $\exists\, Q = E_k \dots E_1$ (y, por tanto, regular) tal que $B = QA$. $\square$

Cuando consideremos la equivalencia por columnas tendremos análogamente la equivalencia de las siguientes afirmaciones:

1. $A \sim_c B$,
2. A y B tienen la misma forma normal de Hermite por columnas,
3. $\exists\, P \in \mathcal{M}_n(K)$, regular, tal que $B = AP$.

A partir de estos resultados introducimos una nueva definición: sean $A, B \in \mathcal{M}_{m\times n}(K)$. Decimos que A y B son equivalentes y lo notamos $A \sim B$ si podemos obtener B a partir de A con una sucesión finita de transformaciones elementales por filas y por columnas. Así, si $A \sim_f B$, entonces $A \sim B$, pero el recíproco no es cierto, es decir, $A \sim B \not\Longrightarrow A \sim_f B$, pues para pasar de A a B puede que necesitemos una transformación elemental por columnas y, si este es el caso, no podremos afirmar que $A \sim_f B$.
Uniendo los resultados que hemos visto anteriormente tenemos que $A, B \in \mathcal{M}_{m\times n}(K)$ son equivalentes si y solo si existen matrices regulares $Q \in \mathcal{M}_m(K)$ y $P \in \mathcal{M}_n(K)$ tales que $B = QAP$, donde Q será el producto de las matrices elementales que corresponden a las transformaciones elementales por filas que hay que aplicarle a A, mientras que P será el producto de las matrices elementales correspondientes a las

transformaciones elementales por columnas que debemos aplicarle a A.
Tenemos entonces que dada $A \in \mathcal{M}_{m\times n}(K)$, $\mathrm{rg}(A) = r$ si y solo si $A \sim J$, donde

$$J = \left(\begin{array}{c|c} I_r & 0 \\ \hline 0 & 0 \end{array}\right)$$

pues si H es la forma de Hermite por filas de A, si suponemos que $\mathrm{rg}(A) = r$, H solo tiene r filas no nulas y ahora aplicándole a H transformaciones elementales por columnas obtenemos J, pues el pivote de cada fila permite anular todos los demás elementos de dicha fila.
Recíprocamente, si $A \sim J$, podemos obtener A de J aplicando primero transformaciones elementales por columnas y luego transformaciones elementales por filas. El rango de J es claramente r y cada matriz que obtengamos de ella mediante transformaciones elementales por columnas seguirá teniendo rango r (el número de filas no nulas de una matriz no queda afectado por estas transformaciones). Como sabemos que las transformaciones elementales por filas no afectan al tango, concluimos que $\mathrm{rg}(A) = r$.
Podemos demostrar ahora que $A, B \in \mathcal{M}_{m\times n}(K)$ son equivalentes si y solo si tienen el mismo rango: si $\mathrm{rg}(A) = r$, y $\mathrm{rg}(B) = s$, $A \sim \left(\begin{array}{c|c} I_r & 0 \\ \hline 0 & 0 \end{array}\right)$, $B \sim \left(\begin{array}{c|c} I_s & 0 \\ \hline 0 & 0 \end{array}\right)$. Así,

$$A \sim B \Longleftrightarrow \left(\begin{array}{c|c} I_r & 0 \\ \hline 0 & 0 \end{array}\right) \sim \left(\begin{array}{c|c} I_s & 0 \\ \hline 0 & 0 \end{array}\right) \Longleftrightarrow r = s.$$

Concluimos probando que $\forall\, A \in \mathcal{M}_{m\times n}(K)$ se tiene que $\mathrm{rg}(A) = \mathrm{rg}(A^t)$.

Demostración. Si $\mathrm{rg}(A) = r$, $A \sim \left(\begin{array}{c|c} I_r & 0 \\ \hline 0 & 0 \end{array}\right)$ y $\exists\, Q \in \mathcal{M}_m(K)$, $P \in \mathcal{M}_n(K)$ <u>regulares</u> tales que $J = QAP$. Si transponemos en esa igualdad, tenemos $J^t = (QAP)^t = P^t A^t Q^t$. Ahora $J^t \in \mathcal{M}_{n\times m}(K)$ es

$$J^t = \left(\begin{array}{c|c} I_r & 0 \\ \hline 0 & 0 \end{array}\right)$$

(no tiene por qué coincidir con J si $m \neq n$). Como además sabemos que P^t y Q^t también son regulares, obtenemos que J^t y A^t son equivalentes. Así, $\mathrm{rg}(A^t) = \mathrm{rg}(J^t) = r = \mathrm{rg}(A)$, con lo que podemos afirmar que el rango de una matriz coincide también con el número de columnas no nulas de su forma normal de Hermite por columnas. $\square$

10. Determinante de una matriz cuadrada.

Este concepto lo definiremos mediante la siguiente inducción sobre el orden de la matriz (que, recordad, es un número natural):
Si la matriz es cuadrada de orden 1, $A = (a_{11})$, siendo $a_{11} \in K$. Entonces definimos el determinante de A como $\det(A) = a_{11}$.

Supongamos conocido el valor del determinante de cualquier matriz $B \in \mathcal{M}_{n-1}(K)$. Sea entonces $A \in \mathcal{M}_n(K)$ dada por

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix}.$$

Llamamos el ij-ésimo menor adjunto de A a $\alpha_{ij} = (-1)^{i+j}\det(A_{ij})$, siendo A_{ij} la submatriz de A obtenida al eliminar en A la fila i-ésima y la columna j-ésima, $\forall\, i, j \in \{1, \dots, n\}$, A_{ij} es una matriz cuadrada de orden $n-1$ y, por la hipótesis de inducción, conocemos $\det(A_{ij})$ y, por tanto, α_{ij}. Entonces definimos:

$$\det(A) = a_{11}\alpha_{11} + a_{21}\alpha_{21} + \dots + a_{n1}\alpha_{n1} = \sum_{i=1}^{n} a_{i1}\alpha_{i1}.$$

Esta fórmula se llama el desarrollo de Laplace del determinante de A por su primera columna.

Ejemplo 8.

Consideremos I_n, la matriz identidad de orden n. Veamos que $\det(I_n) = 1$. Como $I_1 = (1)$, claramente $\det(I_1) = 1$. Supongamos ahora que $\forall\, n-1 > 1$, $n \in \mathbb{N}$, $\det(I_{n-1}) = 1$. Entonces como podemos escribir $I_n = (\delta_{ij})_{i,j}$, si aplicamos el desarrollo de Laplace por la primera columna al $\det(I_n)$ obtenemos que

$$\det(I_n) = \delta_{11}\alpha_{11} + \dots + \delta_{n1}\alpha_{n1}.$$

Pero $\delta_{j1} = 0$, $\forall\, j \in \{2, \dots, n\}$. Así, $\det(I_n) = \delta_{11}\alpha_{11} = \alpha_{11}$ y $\alpha_{11} = (-1)^{1+1}\det(A_{11}) = \det(A_{11})$. Pero A_{11} es la matriz obtenida de I_n al quitarle la primera fila y la primera columna. Es decir, $A_{11} = I_{n-1}$ y, por la hipótesis de inducción, $\det(I_{n-1}) = 1$, con lo que, efectivamente, $\det(I_n) = 1$.

También escribiremos el determinante de la matriz A de la forma

$$\det(A) = \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix} = \left|A\right|.$$

Así, si $A \in \mathcal{M}_2(K)$ tendremos

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}(-1)^{1+1}\det(A_{11}) + a_{12}(-1)^{2+1}\det(A_{21}) = a_{11}a_{22} - a_{12}a_{21}.$$

Si $A \in \mathcal{M}_3(K)$ se tiene

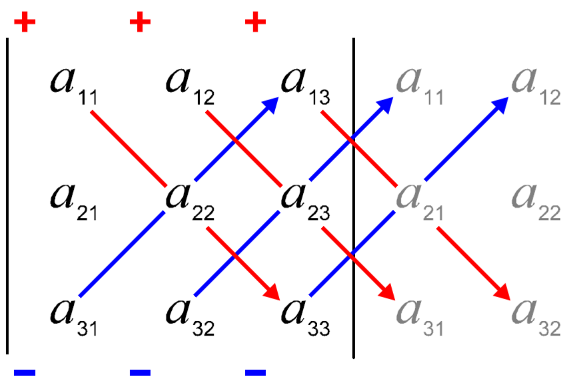
$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{11}(-1)^{1+1}\det(A_{11}) + a_{21}(-1)^{2+1}\det(A_{21}) + a_{31}(-1)^{3+1}\det(A_{31}) =$$

$$= a_{11}\begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{21}\begin{vmatrix} a_{12} & a_{13} \\ a_{32} & a_{33} \end{vmatrix} + a_{31}\begin{vmatrix} a_{12} & a_{13} \\ a_{22} & a_{23} \end{vmatrix} =$$

$$= a_{11}(a_{22}a_{33} - a_{32}a_{23}) - a_{21}(a_{12}a_{33} - a_{32}a_{13}) + a_{31}(a_{12}a_{23} - a_{22}a_{13}) =$$

$$= a_{11}a_{22}a_{33} + a_{13}a_{21}a_{32} + a_{12}a_{23}a_{31} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} - a_{13}a_{22}a_{31},$$

donde hemos usado que el cuerpo K es conmutativo. Se obtiene la llamada regla de Sarrus que nos da gráficamente dos sumando positivos y negativos para el determinante de una matriz de orden 3



Así, $\begin{vmatrix} 1 & 4 \\ -3 & 3 \end{vmatrix} = 3 + 12 = 15$, y $\begin{vmatrix} 1 & 2 & 3 \\ 0 & 1 & 4 \\ 2 & 1 & 2 \end{vmatrix} = 2 + 16 - 6 - 4 = 18 - 10 = 8.$

Hay que tener cuidado, pues para una matriz de orden $n \geq 4$ no tenemos reglas como la de Sarrus para poder obtener fácilmente su determinante.

Sin embargo, las propiedades de los determinantes que vamos a ver a continuación nos permitirán ir reduciendo el cálculo del determinante de una matriz cuadrada a los de matrices de orden inferior y, repitiendo el proceso, al cálculo de los determinantes de matrices de orden 3, que serán muy fáciles de obtener.

Propiedades de los determinantes:

1. Si tres matrices cuadradas A, A' y A'' son idénticas, excepto en que la i-ésima fila (o columna) de A es la suma de las correspondientes filas (o columnas) de A' y A'', entonces $\det(A) = \det(A') + \det(A'')$. Es decir:

$$\begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ x_1 + y_1 & x_2 + y_2 & \dots & x_n + y_n \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix} = \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ x_1 & x_2 & \dots & x_n \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix} + \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ y_1 & y_2 & \dots & y_n \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix}.$$

Demostración. Solo demostraremos esta propiedad y las siguientes para el saco en que consideramos filas. Más adelante veremos por qué. Lo que hacemos es una inducción sobre el orden n de las matrices:

Si $n = 1$ es evidente pues $|x_1 + y_1| = x_1 + y_1 = |x_1| + |y_1|$.

Supongamos que la propiedad es cierta para $n - 1 > 1$, $n \in \mathbb{N}$. Demostrémosla para n. Notaremos

$\alpha_{ij}, \alpha'_{ij}, \alpha''_{ij}$, respectivamente, a los adjuntos del elemento que ocupa el lugar ij en A, A' y A'', respectivamente. Así,

$$\det(A) = a_{11}\alpha_{11} + \cdots + (x_1 + y_1)\alpha_{i1} + \cdots + a_{n1}\alpha_{n1}.$$

Además, $\alpha_{i1} = \alpha'_{i1} = \alpha''_{i1}$ pues para su cálculo eliminamos la fila i-ésima en las matrices A, A' y A'' y esa fila es la única en la que difieren dichas matrices. Además, si $t \neq i$, $t \in \{1, \ldots, n\}$, la hipótesis de inducción nos dice que $\alpha_{t1} = \alpha'_{t1} + \alpha''_{t1}$. Trasladando esto a la igualdad anterior tendremos

$$\det(A) = a_{11}(\alpha'_{11} + \alpha''_{11}) + \cdots + (x_1 + y_1) + \cdots + a_{n1}(\alpha'_{n1} + \alpha''_{n1}) =$$

$$= (a_{11}\alpha'_{11} + \cdots + a_1\alpha'_{i1} + \cdots + a_{n1}\alpha'_{n1}) + (a_{11}\alpha''_{11} + \cdots + y_1\alpha''_{i1} + \cdots + a_{n1}\alpha''_{n1}) = \det(A') + \det(A'').$$

$\square$

2. Si una matriz $A \in \mathcal{M}_n(K)$ tiene dos filas (ó dos columnas) iguales, entonces su determinante se anula, $\det(A) = 0$.

$$\begin{vmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ x_1 & x_2 & \ldots & x_n \\ \vdots & \vdots & \ddots & \vdots \\ x_1 & x_2 & \ldots & x_n \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{vmatrix} = 0.$$

3. Si se intercambian dos filas (ó dos columnas) de la matriz A su determinante cambia de signo:

$$\begin{vmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ x_1 & x_2 & \ldots & x_n \\ \vdots & \vdots & \ddots & \vdots \\ y_1 & y_2 & \ldots & y_n \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{vmatrix} = - \begin{vmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ y_1 & y_2 & \ldots & y_n \\ \vdots & \vdots & \ddots & \vdots \\ x_1 & x_2 & \ldots & x_n \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{vmatrix}.$$

Demostración. Vamos a demostrar las dos propiedades anteriores a la vez: primero en el caso particular de dos filas consecutivas. Vamos a ver que si una matriz tiene dos filas consecutivas iguales su determinante se anula. De nuevo lo haremos por inducción sobre el orden de la matriz, n. Tenemos que empezar en $n = 2$ (para $n = 1$, estas propiedades no tienen sentido). En ese caso

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{11} & a_{12} \end{vmatrix} = a_{11}a_{12} - a_{11}a_{12} = 0.$$

Vamos a suponer que la propiedad es cierta para un número natural $n-1 \geq 2$, y lo demostraremos para n. Si las filas consecutivas que son iguales son la i-ésima y la $(i+1)$-ésima, tenemos que:

$$\det(A) = a_{11}\alpha_{11} + \cdots + x_1\alpha_{i1} + x_1\alpha_{i+11} + \cdots + a_{n1}\alpha_{n1}.$$

Supongamos que $t \neq i, i+1$. Por la hipótesis de inducción $\alpha_{t1} = (-1)^{t+1}\det(A_{t1}) = 0$, pues la matriz A_{t1} es de orden $n-1$ y tiene dos filas consecutivas iguales. Luego

$$\det(A) = x_1\alpha_{i1} + x_1\alpha_{i+11}.$$

Como $A_{i1} = A_{i+11}$, pues las obtenemos eliminando en A la primera columna y las filas i-ésima e $(i+1)$-ésima (respectivamente), que son iguales, tendremos que $|A_{i1}| = |A_{i+11}|$. Entonces $\alpha_{i1} = (-1)^{i+1}\det(A_{i1})$ mientras que $\alpha_{i+11} = (-1)^{i+2}\det(A_{i1})$ son claramente opuesto entre sí. Por tanto, efectivamente, $\det(A) = 0$.

A continuación vamos a demostrar que si en una matriz intercambiamos entre sí dos filas consecutivas, el determinante cambia de signo. Esto es consecuencia de la Propiedad 1 y la que acabamos de ver, pues

$$0 = \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ x_1+y_1 & x_2+y_2 & \dots & x_n+y_n \\ x_1+y_1 & x_2+y_2 & \dots & x_n+y_n \\ \vdots & \vdots & \ddots & \vdots \\ a_{1n} & a_{2n} & \dots & a_{nn} \end{vmatrix} = \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ x_1 & x_2 & \dots & x_n \\ x_1+y_1 & x_2+y_2 & \dots & x_n+y_n \\ \vdots & \vdots & \ddots & \vdots \\ a_{1n} & a_{2n} & \dots & a_{nn} \end{vmatrix} +$$

$$+ \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ y_1 & y_2 & \dots & y_n \\ x_1+y_1 & x_2+y_2 & \dots & x_n+y_n \\ \vdots & \vdots & \ddots & \vdots \\ a_{1n} & a_{2n} & \dots & a_{nn} \end{vmatrix} =$$

$$= \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ x_1 & x_2 & \dots & x_n \\ x_1 & x_2 & \dots & x_n \\ \vdots & \vdots & \ddots & \vdots \\ a_{1n} & a_{2n} & \dots & a_{nn} \end{vmatrix} + \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ x_1 & x_2 & \dots & x_n \\ y_1 & y_2 & \dots & y_n \\ \vdots & \vdots & \ddots & \vdots \\ a_{1n} & a_{2n} & \dots & a_{nn} \end{vmatrix} + \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ y_1 & y_2 & \dots & y_n \\ x_1 & x_2 & \dots & x_n \\ \vdots & \vdots & \ddots & \vdots \\ a_{1n} & a_{2n} & \dots & a_{nn} \end{vmatrix} + \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ y_1 & y_2 & \dots & y_n \\ y_1 & y_2 & \dots & y_n \\ \vdots & \vdots & \ddots & \vdots \\ a_{1n} & a_{2n} & \dots & a_{nn} \end{vmatrix} (*).$$

Pero los determinantes de las matrices 1ª y 4ª se anulan ya que en ambos casos hay dos filas contiguas iguales. Así,

$$\begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ x_1 & x_2 & \dots & x_n \\ y_1 & y_2 & \dots & y_n \\ \vdots & \vdots & \ddots & \vdots \\ a_{1n} & a_{2n} & \dots & a_{nn} \end{vmatrix} = - \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ y_1 & y_2 & \dots & y_n \\ x_1 & x_2 & \dots & x_n \\ \vdots & \vdots & \ddots & \vdots \\ a_{1n} & a_{2n} & \dots & a_{nn} \end{vmatrix}.$$

Vamos ya al caso general:

Podemos deducir la Propiedad 2 de los resultados anteriores: si A tiene dos filas iguales no consecutivas (la i-ésima y la j-ésima), intercambiando la fila j-ésima con cada una de las filas anteriores a ella y posteriores a la i-ésima, cambiamos el signo un cierto número de veces, pero llegamos a una matriz con dos filas consecutivas iguales y sabemos que su determinante es 0, que no se ve afectado por los cambios de signo que hayamos hecho.

Ahora la Propiedad 3 se deduce de la Propiedad 2 por el mismo método seguido en $(*)$. □

4. Si se multiplican todos los elementos de una fila (ó de una columna) de la matriz A por un escalar k, el determinante de la nueva matriz es $k\det(A)$.

Demostración. Aplicaremos, de nuevo, una inducción sobre el orden n de la matriz: si $n = 1$ entonces $\det(ka_{11}) = ka_{11} = k\det(a_{11})$. Supongamos cierto el resultado para $n - 1 \geq 1$ y consideremos que $A \in \mathcal{M}_n(K)$. Llamaremos A' a la matriz que obtenemos multiplicando una fila de A por el escalar k y α_{ij} y α'_{ij}, respectivamente, a los menores adjuntos de los elementos que están en el lugar ij en A y A' (respectivamente). Entonces $\det(A') = a_{11}\alpha'_{11} + \cdots + kx_1\alpha'_{i1} + \cdots + a_{n1}\alpha'_{n1}$. Como $\alpha'_{i1} = \alpha_{i1}$, pues en ambos casos quitamos la fila i-ésima y la columna primera en A y las matrices que nos quedan son iguales, y como, por la hipótesis de inducción cada α'_{t1} con $t \neq i$ verifica $\alpha'_{t1} = k\alpha_{t1}$, ya que en las matrices A'_{t1} con $t \neq 1$ una fila es el resultado de multiplicar por k la correspondiente fila de A_{t1}, tendremos

$$\det(A') = a_{11}k\alpha_{11} + \cdots + kx_1\alpha_{i1} + \cdots + a_{n1}k\alpha_{n1} =$$

$$= k(a_{11}\alpha_{11} + \cdots + x_1\alpha_{i1} + \cdots + a_{n1}\alpha_{n1}) = k\det(A).$$

□

En particular, esta propiedad nos dice que si A tiene una fila (ó una columna) cuyos elementos son todos nulos, entonces su determinante es cero.

5. Si a una fila (ó una columna) de una matriz cuadrada A le sumamos el producto por un escalar de

otra fila, entonces su determinante no varía. Es decir,

$$(*)\begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ x_1 + ky_1 & x_2 + ky_2 & \dots & x_n + ky_n \\ \vdots & \vdots & \ddots & \vdots \\ y_1 & y_2 & \dots & y_n \\ \vdots & \vdots & \ddots & \vdots \\ a_{1n} & a_{2n} & \dots & a_{nn} \end{vmatrix} = \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ x_1 & x_2 & \dots & x_n \\ \vdots & \vdots & \ddots & \vdots \\ y_1 & y_2 & \dots & y_n \\ \vdots & \vdots & \ddots & \vdots \\ a_{1n} & a_{2n} & \dots & a_{nn} \end{vmatrix} (**).$$

Demostración. Esta propiedad se deduce fácilmente de las Propiedades 1 y 4 pues

$$(*) = \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ x_1 & x_2 & \dots & x_n \\ \vdots & \vdots & \ddots & \vdots \\ y_1 & y_2 & \dots & y_n \\ \vdots & \vdots & \ddots & \vdots \\ a_{1n} & a_{2n} & \dots & a_{nn} \end{vmatrix} + k \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ y_1 & y_2 & \dots & y_n \\ \vdots & \vdots & \ddots & \vdots \\ y_1 & y_2 & \dots & y_n \\ \vdots & \vdots & \ddots & \vdots \\ a_{1n} & a_{2n} & \dots & a_{nn} \end{vmatrix} = (**).$$

Ya que, por la Propiedad 2 el segundo determinante de la expresión anterior se anula. □

6. Una matriz cuadrada A es regular si y solo si $\det(A) \neq 0$.

   Demostración. Si le aplicamos las propiedades 3, 4 y 5 anteriores a la matriz identidad I_n, deducimos que los determinantes de las matrices elementales vienen dados por $\det(E_{ij}) = -1$, $\det(E_i(k)) = k$ ($k \in K \setminus \{0\}$) y $\det(E_{ij}(k)) = \det(I_n) = 1$.

   Por tanto, esas propiedades 3, 4 y 5 se resumen en que si ahora E es una matriz elemental entonces $\det(EA) = \det(E)\det(A)$. Así, es inmediato comprobar que si $E_1, \dots, E_k$ son matrices elementales, entonces $\det(E_k \dots E_1 A) = \det(E_k) \dots \det(E_1)\det(A)$. Entonces, sabemos que si A es una matriz regular, es un producto de matrices elementales, $A = E_k \dots E_1$ y, entonces $\det(A) = \det(E_k) \dots \det(E_1) \neq 0$ pues cada uno de ellos es no nulo.

   Recíprocamente, supongamos que A no es regular. Sabemos que $\mathrm{rg}(A) < n$ y, por tanto, su forma normal de Hermite por filas H tiene, al menos, una fila de ceros. Luego $\det(H) = 0$ por la Propiedad 4. Sabemos que $H = E_k \dots E_1 A$ para ciertas matrices elementales E_i, $i \in \{1, \dots, k\}$. Así, $0 = \det(H) = \det(E_k \dots E_1 A) = \det(E_k) \dots \det(E_1)\det(A)$ y como $\det(E_i) \neq 0$, $\forall\, i \in \{1, \dots, k\}$, necesariamente $\det(A) = 0$. □

7. El determinante de un producto de matrices $A, B \in \mathcal{M}_n(K)$ verifica $\det(AB) = \det(A)\det(B)$.

Demostración. Supongamos, en primer lugar, que $\det(A) = 0$ ó $\det(B) = 0$. Entonces A ó B es singular, con lo que AB es singular. Si ahora $\det(A) \neq 0$ y $\det(B) \neq 0$; entonces, por la Propiedad 6 ambas son regulares y, por tanto son producto de matrices elementales: $A = E_1 \dots E_k$, $B = E'_1 \dots E'_l$. Entonces

$$\det(AB) = \det((E_1 \dots E_k)(E'_1 \dots E'_l)) = \det(E_1 \dots E_k)\det(E'_1 \dots E'_l) = \det(A)\det(B).$$

□

8. El determinante de una matriz cuadrada coincide con el de su traspuesta.

   Demostración. De nuevo consideramos dos casos: si $\det(A) = 0$, entonces es singular. Sabemos que, entonces, A^t también es singular y, por tanto, $\det(A^t)) = 0$. Si $\det(A) \neq 0$, A es regular y podemos encontrar matrices elementales $E_1, \dots, E_k$ tales que $A = E_1 \dots E_k$. Entonces $A^t = E_k^t \dots E_1^t$ y para cada matriz elemental E, $\det(E) = \det(E^t)$. Así, $\det(A^t) = \det(E_k^t) \dots \det(E_1^t) = \det(E_k) \dots \det(E_1) = \det(A)$. □

   A partir de aquí, todas las Propiedades de los determinantes referidas a filas son válidas también para columnas.

9. El determinante de una matriz se puede obtener desarrollando por cualquier de sus filas ó columnas, es decir,

   $$\det(A) = a_{i1}\alpha_{i1} + a_{i2}\alpha_{i2} + \dots + a_{in}\alpha_{in} = a_{1j}\alpha_{1j} + a_{2j}\alpha_{2j} + \dots + a_{nj}\alpha_{nj},$$

   $\forall\, i, j \in \{1, \dots, n\}$.

   Demostración. La propiedad se obtiene como consecuencia de las Propiedades 3 y 8. □

Veamos, con un ejemplo, cómo calcular el determinante de una matriz:

Ejemplo 9.
Sea la matriz

$$A = \begin{pmatrix} 1 & 12 & 123 & 1234 \\ 2 & 23 & 234 & 2341 \\ 3 & 34 & 341 & 3412 \\ 4 & 41 & 412 & 4123 \end{pmatrix}$$

Para calcular su determinante multiplicamos la 1ª columna por 12 y se la restamos a la 2ª y así,

$$\det(A) = \begin{vmatrix} 1 & 0 & 123 & 1234 \\ 2 & -1 & 234 & 2341 \\ 3 & -2 & 341 & 3412 \\ 4 & -7 & 412 & 4123 \end{vmatrix}.$$

Ahora multiplicamos la 1ª columna por 123 y se la restamos a la 3ª y luego por 1234 y se la restamos a la cuarta, obteniendo:

$$\det(A) = \begin{vmatrix} 1 & 0 & 0 & 0 \\ 2 & -1 & -12 & -127 \\ 3 & -2 & -28 & -290 \\ 4 & -7 & -80 & -813 \end{vmatrix} = - \begin{vmatrix} 1 & 0 & 0 & 0 \\ 2 & 1 & -12 & -127 \\ 3 & 2 & -28 & -290 \\ 4 & 7 & -80 & -813 \end{vmatrix} = - \begin{vmatrix} 1 & 0 & 0 & 0 \\ 2 & 1 & 12 & 127 \\ 3 & 2 & 28 & 290 \\ 4 & 7 & 80 & 813 \end{vmatrix}$$

donde hemos aplicado 3 veces la Propiedad 4. Si ahora desarrollamos por la 1ª fila obtenemos que:

$$\det(A) = - \begin{vmatrix} 1 & 12 & 127 \\ 2 & 28 & 290 \\ 7 & 80 & 813 \end{vmatrix}.$$

Multiplicando ahora la 1ª fila por 2 y restándola a la 2ª y luego por 7 y restándosela a la 3ª tenemos

$$\det(A) = - \begin{vmatrix} 1 & 12 & 127 \\ 0 & 4 & 36 \\ 0 & -4 & -76 \end{vmatrix}.$$

Aplicando de nuevo la Propiedad 4 en la 3ª fila tenemos

$$\det(A) = \begin{vmatrix} 1 & 12 & 127 \\ 0 & 4 & 36 \\ 0 & 4 & 76 \end{vmatrix}.$$

Desarrollando por la 1ª columna tenemos $\det(A) = \begin{vmatrix} 4 & 36 \\ 4 & 76 \end{vmatrix} = 8 \begin{vmatrix} 1 & 18 \\ 1 & 38 \end{vmatrix} = 8(38-19) = 8 \cdot 20 = 160.$

11. Matriz inversa y determinantes.

Veamos cómo podemos utilizar los determinantes para el cálculo de la matriz inversa. Sea $A \in \mathcal{M}_n(K)$, llamaremos matriz adjunta de $A = (a_{ij})_{i,j}$ a la matriz $A^* = (\alpha_{ij})_{i,j}$, que en cada lugar tiene el correspondiente menor adjunto de A.

Así, si $A = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 4 \\ 3 & 4 & 6 \end{pmatrix}$ entonces $A^* = \begin{pmatrix} 2 & 0 & -1 \\ 0 & -3 & 2 \\ -1 & 2 & -1 \end{pmatrix}$.

Veamos que si $A \in \mathcal{M}_n(K)$ y A^* es su matriz adjunta, se verifica que $AA^{*t} = \det(A) I_n$.

Para ello, llamemos $C = (c_{ij})_{i,j}$ a la matriz $C = AA^{*t}$. Por tanto, $\forall\, i, j \in \{1, \ldots, n\}$ tendremos que

$$c_{ij} = a_{i1}\alpha_{j1} + a_{i2}\alpha_{j2} + \cdots + a_{in}\alpha_{jn}.$$

Por tanto, si $i = j$,

$$c_{ii} = a_{i1}\alpha_{i1} + a_{i2}\alpha_{i2} + \cdots + a_{in}\alpha_{in} = \det(A),$$

ya que la expresión para c_{ii} es el desarrollo del $\det(A)$ por la fila i-ésima.
Si $i \neq j$ entonces

$$c_{ij} = a_{i1}\alpha_{j1} + \cdots + a_{in}\alpha_{jn} \underset{(*)}{=} \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{i1} & a_{i2} & \dots & a_{in} \\ \vdots & \vdots & \ddots & \vdots \\ a_{i1} & a_{i2} & \dots & a_{in} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix} \begin{array}{l} \\ \\ \text{fila } i) \\ \\ \text{fila } j) \\ \\ \\ \end{array}$$

$*$ nos da el desarrollo del determinante de la derecha por la fila j-ésima y, como la matriz a la que calculamos el determinante tiene dos filas iguales, $c_{ii} = 0$.
Hemos obtenido que

$$AA^{*t} = \begin{pmatrix} \det(A) & 0 & \dots & 0 \\ 0 & \det(A) & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \det(A) \end{pmatrix} = \det(A) I_n.$$

Si suponemos, entonces que la matriz $A \in \mathcal{M}_n(K)$ es regular, tenemos que $\det(A) \neq 0$. Si en la igualdad anterior dividimos por $\det(A)$, ya que no anula, tendremos que

$$A\left(\frac{1}{\det(A)}A^{*t}\right) = I_n, \quad \text{luego} \qquad A^{-1} = \frac{1}{\det(A)}A^{*t}.$$

Por ejemplo, para la matriz que hemos considerado antes, $A = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 4 \\ 3 & 4 & 6 \end{pmatrix}$, tiene determinante

$$|A| = \begin{vmatrix} 1 & 0 & 0 \\ 2 & -1 & -2 \\ 3 & -2 & -3 \end{vmatrix} = \begin{vmatrix} 1 & 0 & 0 \\ 2 & 1 & 2 \\ 3 & 2 & 3 \end{vmatrix} = \begin{vmatrix} 1 & 2 \\ 2 & 3 \end{vmatrix} = -1,$$

mientras que $A^* = \begin{pmatrix} 2 & 0 & -1 \\ 0 & -3 & 2 \\ -1 & 2 & -1 \end{pmatrix}$, A es regular y $A^{-1} = -\begin{pmatrix} 2 & 0 & -1 \\ 0 & -3 & 2 \\ -1 & 2 & -1 \end{pmatrix} = \begin{pmatrix} -2 & 0 & 1 \\ 0 & 3 & -2 \\ 1 & -2 & 1 \end{pmatrix}$.

12. Rango y determinantes.

A partir de la definición de determinante obtendremos un nuevo método para calcular el rango de una matriz: sea $A \in \mathcal{M}_{m\times n}(K)$. Entonces, $\mathrm{rg}(A)$ es el mayor orden de una submatriz cuadrada y regular de A.
Para demostrar esto, sea $r = \mathrm{rg}(A)$ y sea H la forma normal de Hermite por filas de A. Por tanto, H contiene una submatriz regular de orden r (la que contiene los r pivotes) y como, además, en H hay

exactamente r filas no nulas, el determinante de cualquier submatriz cuadrada de orden $r+1$ de H se anula, pues tendrá una fila de ceros. Así, todas las submatrices cuadradas de orden $r+1$ de H son singulares. Como esta propiedad se mantiene por cada transformación elemental por filas, se verificará que A tiene una submatriz cuadrada regular de r y todas sus submatrices cuadradas de orden $r+1$ serán singulares. Esto nos proporcionará un método para calcular el rango de una matriz.

Si $A = \begin{pmatrix} 1 & 2 & 1 & 0 \\ 2 & 4 & -2 & 8 \\ -1 & 0 & 1 & -2 \end{pmatrix}$, el elemento a_{11} de A es $1 \neq 0$, luego $\mathrm{rg}(A) \geq 1$. Si consideramos la submatriz que contiene a a_{11} obtenida eliminando en A la 2ª fila y la 3ª y 4ª columnas, $\begin{vmatrix} 1 & 2 \\ -1 & 0 \end{vmatrix} = 2 \neq 0$. Esa submatriz es regular y, por tanto, $\mathrm{rg}(A) \geq 2$. Para obtener submatrices cuadradas de A a partir de esa submatriz añadiremos la 2ª fila y, ó bien la 3ª columna y la 4ª. Si alguno de los determinantes de las submatrices así obtenidas no es nulo, el $\mathrm{rg}(A) = 3$, y si ambos se anulan $\mathrm{rg}(A) = 2$.

Como $\begin{vmatrix} 1 & 2 & 1 \\ 2 & 4 & -2 \\ -1 & 0 & 1 \end{vmatrix} = 4 + 4 + 4 - 4 = 8 \neq 0$, podemos asegurar que $\mathrm{rg}(A) = 3$.

13. Regla de Cramer.

Dado un sistema de ecuaciones lineales sobre K

$$\begin{cases} a_{11}x_1 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + \cdots + a_{2n}x_n = b_2 \\ \quad \vdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n = b_m \end{cases}$$

Denotamos $A = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ a_{21} & \dots & a_{2n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix} \in \mathcal{M}_{m \times n}(K)$ y la llamamos matriz de coeficientes del sistema. Notamos asimismo $X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathcal{M}_{n \times 1}(K)$ y la llamamos matrix incógnita y $B = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix} \in \mathcal{M}_{m \times 1}(K)$ que llamaremos matriz de términos independientes. A la matriz $(A|B)$ la llamaremos matriz ampliada del sistema.

Con estas notaciones el sistema anterior lo podemos escribir matricialmente como $AX = B$. Dicho sistema será de Cramer si A es una matriz cuadrada (hay, por tanto, tantas ecuaciones como incógnitas) y regular. Si aplicamos el Teorema de Rouché-Frobenius todo sistema de Cramer es compatible determinado (ya que, al ser A regular, su forma normal de Hermite por filas es I_n). La regla de Cramer nos permite obtener la solución para los sistemas de Cramer:

Si el sistema

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \\ \quad \vdots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n = b_n \end{cases}$$

es de Cramer su solución (única) viene dada por:

$$x_1 = \frac{1}{\det(A)} \begin{vmatrix} b_1 & a_{12} & \dots & a_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ b_n & a_{n2} & \dots & a_{nn} \end{vmatrix}, \dots, x_j = \frac{1}{\det(A)} \begin{vmatrix} a_{11} & b_1 & \dots & a_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & b_n & \dots & a_{nn} \end{vmatrix}, \dots,$$

$$x_n = \frac{1}{\det(A)} \begin{vmatrix} a_{11} & a_{12} & \dots & b_1 \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & b_n \end{vmatrix},$$

ya que, matricialmente, escribiremos el sistema $AX = B$ y, como A es regular, $\exists\, A^{-1}$ y $A^{-1}(AX) = (A^{-1}A)X = I_n X = X = A^{-1}B$. Luego

$$\begin{pmatrix} x_1 \\ \vdots \\ x_i \\ \vdots \\ x_n \end{pmatrix} = \frac{1}{\det(A)} \begin{pmatrix} \alpha_{11} & \alpha_{21} & \dots & \alpha_{n1} \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_{1i} & \alpha_{2i} & \dots & \alpha_{ni} \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_{1n} & \alpha_{2n} & \dots & \alpha_{nn} \end{pmatrix} \begin{pmatrix} b_1 \\ \vdots \\ b_i \\ \vdots \\ b_n \end{pmatrix}.$$

Así, $\forall\, 1 \leq i \leq n$, tendremos que

$$x_i = \frac{1}{\det(A)}(b_1\alpha_{1i} + \cdots + b_n\alpha_{ni}) = \frac{1}{\det(A)} \begin{vmatrix} a_{11} & \dots & b_1 & \dots & a_{1n} \\ a_{21} & \dots & b_2 & \dots & a_{2n} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ a_{n1} & \dots & b_n & \dots & a_{nn} \end{vmatrix}.$$

Veamos, finalmente, cómo aplicar la regla de Cramer para resolver cualquier sistema compatible:
En primer lugar nos aseguramos de que para nuestro sistema se verifica que $\mathrm{rg}(A) = \mathrm{rg}(A|B)$, pues en caso contrario el sistema es incompatible.
Si $\mathrm{rg}(A) = \mathrm{rg}(A|B) = n =$ número de incógnitas, nuestro sistema es compatible determinado. Si $m > n$, y suponemos que el $\mathrm{rg}(A) = \mathrm{rg}(A|B)$ lo alcanzamos para las primeras n filas de A (y, por tanto, de $(A|B)$), el resto de filas las podemos obtener multiplicando cada una de las primeras n filas por un escalar, luego restándole a cada una de las filas de la $n+1$ a la m, la correspondiente suma de escalar multiplicados por las n primeras filas, logramos que en $(A|B)$ y, por lo tanto, en A, esas últimas filas sean nulas.
Esto se traduce en que en nuestro sistema hemos conseguido que las ecuaciones desde la $n+1$ a la m sean ahora $0 = 0$, y las podemos quitar. Ahora tenemos un sistema de n ecuaciones y n incógnitas que es de

Cramer, pues el rango de su matriz de coeficientes es n y lo resolvemos.

Si $\mathrm{rg}(A) = \mathrm{rg}(A|B) = r < n$, en A podemos encontrar una submatriz cuadrada de orden r cuyo determinante es distinto de cero, mientras que el determinante de cualquier submatriz de A de orden superior a r se anula. Con esta submatriz alcanzamos también el rango de $(A|B)$. Por notación consideremos que tal submatriz de A se obtiene considerando las r primeras filas y columnas de A (si no, aplicaríamos transformaciones elementales de Tipo I al sistema hasta conseguir poner las filas correspondientes a la submatriz como las primeras ecuaciones del sistema y cambiaríamos en cada ecuación resultante el orden de las incógnitas para que las r primeras correspondan a las columnas de la submatriz de A). Por el mismo procedimiento de antes, podremos eliminar en nuestro sistema las $n - r$ últimas ecuaciones y escribirlo de la forma

$$\begin{cases} a_{11}x_1 + \cdots + a_{1r}x_r + a_{1r+1}x_{r+1} + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + \cdots + a_{2r}x_r + a_{2r+1}x_{r+1} + \cdots + a_{2n}x_n = b_2 \\ \quad \vdots \\ a_{r1}x_1 + \cdots + a_{rr}x_r + a_{rr+1}x_{r+1} + \cdots + a_{rn}x_n = b_r \end{cases}$$

donde $\begin{vmatrix} a_{11} & \dots & a_{1r} \\ \vdots & \ddots & \vdots \\ a_{r1} & \dots & a_{rr} \end{vmatrix} \neq 0.$

Nuestro sistema es equivalente al siguiente:

$$\begin{cases} a_{11}x_1 + \cdots + a_{1r}x_r = b_1 - a_{1r+1}x_{r+1} - \cdots - a_{1n}x_n \\ a_{21}x_1 + \cdots + a_{2r}x_r = b_2 - a_{2r+1}x_{r+1} - \cdots - a_{2n}x_n \\ \quad \vdots \\ a_{r1}x_1 + \cdots + a_{rr}x_r = b_r - a_{rr+1}x_{r+1} - \cdots - a_{rn}x_n \end{cases}$$

que ahora es de Cramer y podemos resolver por la regla de Cramer. (Nótese ahora que cada variable $x_1, \dots, x_r$, dependerá de los valores de $x_{r+1}, \dots, x_n$. Si cada una de estas $n - r$ incógnitas toma un valor concreto, tendremos valores concretos para $x_1, \dots, x_r$ y, por tanto, una solución de nuestro sistema primitivo).

Como ejemplo, consideremos el sistema

$$\begin{cases} x_1 + 2x_2 - 3x_3 + 4x_4 = 2 \\ 2x_1 + 5x_2 - 2x_3 + x_4 = 1 \\ 5x_1 + 12x_2 - 7x_3 + 6x_4 = 7 \end{cases}$$

que tiene por matriz

$$A = \begin{pmatrix} 1 & 2 & -3 & 4 \\ 2 & 5 & -2 & 1 \\ 5 & 12 & -7 & 6 \end{pmatrix}$$

y por matriz aumentada

$$(A|B) = \begin{pmatrix} 1 & 2 & -3 & 4 & 2 \\ 2 & 5 & -2 & 1 & 1 \\ 5 & 12 & -7 & 6 & 7 \end{pmatrix}.$$

Calculemos $\mathrm{rg}(A)$. Como la submatriz $\begin{pmatrix} 1 & 2 \\ 2 & 5 \end{pmatrix}$ tiene determinante igual a 1, $\mathrm{rg}(A) \geq 2$. Para ver si es 3 calculamos

$$\begin{vmatrix} 1 & 2 & -3 \\ 2 & 5 & -2 \\ 5 & 12 & -7 \end{vmatrix} = -35 - 72 - 20 + 75 + 24 + 28 = -127 + 127 = 0.$$

Tomamos entonces

$$\begin{vmatrix} 1 & 2 & 4 \\ 2 & 5 & 1 \\ 5 & 12 & 6 \end{vmatrix} = 30 + 96 + 10 - 100 - 12 - 24 = 136 - 136 = 0.$$

Como estos dos determinantes se anulan entonces $\mathrm{rg}(A) = 2$. Para calcular el rango de la matriz aumentada tenemos que la submatriz $\begin{pmatrix} 1 & 2 \\ 2 & 5 \end{pmatrix}$ tiene determinante $\neq 0$, pero si le añadimos a esa matriz la 3ª fila y la 3ª o 4ª columnas, hemos visto que los correspondientes determinantes se anulan. Entonces $\mathrm{rg}(A|B) = 3$ si y solo si $\begin{vmatrix} 1 & 2 & 2 \\ 2 & 5 & 1 \\ 5 & 12 & 7 \end{vmatrix} \neq 0$. Como

$$\begin{vmatrix} 1 & 2 & 2 \\ 2 & 5 & 1 \\ 5 & 12 & 7 \end{vmatrix} = 35 + 48 + 10 - 50 - 12 - 28 = 93 - 90 = 3,$$

obtenemos que $\mathrm{rg}(A|B) = 3$ y el sistema es, por tanto, incompatible.

Geometría I: Tema 2
Espacios vectoriales

Juan de Dios Pérez

Índice

1. Introducción. 2

2. Dependencia e independencia lineal. 4

3. Sistemas de generadores de un espacio vectorial. 7

4. Bases de un espacio vectorial. 9

5. Coordenadas de un vector respecto de una base. 11

6. Coordenadas y dependencia lineal. 11

7. Cambio de base 12

8. Subespacios vectoriales. 14

9. Ecuaciones cartesianas y paramétricas de un subespacio. 16

10. Ecuaciones cartesianas y dimensión de un subespacio. 18

11. Intersección de subespacios. 19

12. Suma de subespacios. 20

13. Suma directa de subespacios. 22

14. Fórmula de las dimensiones. 23

15. Espacio vectorial cociente. 24

1. Introducción.

Sea K un cuerpo conmutativo y V un conjunto no vacío. Diremos que V es un espacio vectorial sobre el cuerpo K y lo notaremos $V(K)$ si se verifican las siguientes condiciones:

1. En V hay definida una operación que llamaremos suma y notaremos $+$ tal que $(V, +)$ es un grupo abeliano. Por tanto, se han de verificar las propiedades siguientes:

   a) Asociativa: $\forall\, u, v, w \in V$ se cumple $(u + v) + w = u + (v + w)$.

   b) Conmutativa: $\forall\, u, v \in V$ se verifica $u + v = v + u$.

   c) Existencia de elemento neutro: $\exists\, 0 \in V$ tal que $0 + v = v + 0 = v$, $\forall\, v \in V$.

   d) Existencia de opuestos: $\forall\, v \in V$, $\exists - v$ tal que $v + (-v) = (-v) + v = 0$.

   Nota 1. De nuevo tenemos abusos de notación: la suma que tenemos definida en V no tiene nada que ver con la suma de escalares en K. Asimismo el 0 de V (neutro para la suma en V) es completamente distinto del neutro para la suma en K aunque estén representados por el mismo símbolo "0". En cada situación hemos de estar seguros de qué operaciones o qué elemento neutro estamos manejando.

2. En V hay definida una ley de composición externa a partir de K que llamaremos producto por escalares; es decir $\cdot : K \times V \longrightarrow V$ tal que $\cdot(\alpha, v) = \alpha v$, que tiene que verificar:

   a) Distributividad: $\forall\, a \in K$, $\forall\, u, v \in V$ se tiene $a(u + v) = au + av$.

   b) Distributividad: $\forall\, a, b \in K$, $\forall\, u \in V$ se tiene $(a + b)u = au + bu$.

   c) Pseudoasociatividad: $\forall\, a, b \in K$, $\forall\, u \in V$ se tiene $(ab)v = a(bv)$. (Considerad que, de nuevo, aparece un abuso de notación).

   d) Propiedad modular: $\forall\, u \in V$, $1u = u$, siendo 1 el elemento neutro para el producto de K.

A los elementos de $V(K)$ los llamaremos vectores. Cuando se conoce el cuerpo K con el que estamos trabajando, es habitual omitirlo de la notación y notar simplemente por V al espacio vectorial.

Ejemplo 1. 1. Por lo que vimos en el tema anterior, $\mathcal{M}_{m\times n}(K)$ es un espacio vectorial sobre K para la suma de matrices y el producto por escalares allí definidos.

2. Un cuerpo K siempre es un espacio vectorial sobre sí mismo considerando el producto por escalares simplemente como el producto de elementos de K.

3. En general, dado un cuerpo K, si consideramos $K^n = K \times \overset{n}{\cdots} \times K = \{(x_1, \ldots, x_n) \,/\, x_i \in K\ \forall\, i \in \{1, \ldots, n\}\}$, siendo n un número natural, K^n va a ser un espacio vectorial sobre K para las siguientes operaciones:

$$(x_1, x_2, \ldots, x_n) + (y_1, y_2, \ldots, y_n) := (x_1 + y_1, x_2 + y_2, \ldots, x_n + y_n),$$

$$k(x_1, x_2, \ldots, x_n) := (kx_1, kx_2, \ldots, kx_n),$$

   $\forall\, k \in K$. Por ejemplo, $\mathbb{R}^n$ y $\mathbb{C}^n$ son espacios vectoriales sobre $\mathbb{R}$ y $\mathbb{C}$ (resp.) para $n \in \mathbb{N}$.

4. Sea $\mathcal{P}(K)$ el conjunto de polinomios en una indeterminada con coeficientes en K. Si definimos la suma usual de polinomios (sumando los monomios de igual grado) y el producto de un polinomio por una constante, $\mathcal{P}(K)$ es un espacio vectorial sobre K.

5. Sea $\mathcal{F}(I,\mathbb{R}) = \{f : I \longrightarrow \mathbb{R} \,/\, \text{I es un intervalo de } \mathbb{R} \text{ y f una aplicación}\}$. Si dadas $f, g \in \mathcal{F}(I,\mathbb{R})$ definimos $f + g : I \longrightarrow \mathbb{R}$ mediante $(f + g)(x) = f(x) + g(x)\ \forall\, x \in I$ y $\forall\, \alpha \in \mathbb{R}$ definimos $\alpha f : I \longrightarrow \mathbb{R}$ mediante $(\alpha f)(x) = \alpha f(x)$, tenemos que $\mathcal{F}(I,\mathbb{R})$ es un espacio vectorial sobre $\mathbb{R}$ o <u>espacio vectorial real</u>.

6. Existe un espacio vectorial con un único vector. Este vector necesariamente será el elemento neutro para la suma 0. Así, tendremos que $0 + 0 = 0$ y $\forall\, \alpha \in K$ (K es un cuerpo conmutativo arbitrario) se tiene $\alpha 0 = 0$. Este espacio vectorial (que lo es para cualquier cuerpo conmutativo) se llama <u>espacio vectorial cero</u> o espacio trivial y lo notaremos $\{0\}$.

7. Como hemos visto en el ejemplo anterior, algunos conjuntos pueden admitir estructura de espacio vectorial sobre diferentes cuerpos. Así, si consideramos el conjunto de los números complejos $\mathbb{C}$, por lo visto en el ejemplo 2 es un espacio vectorial sobre sí mismo (diremos que es un <u>espacio vectorial complejo</u>), pero es también un espacio vectorial real, pues el producto de un número real y uno complejo nos da un número complejo y, como cada número real es un número complejo, ese producto verificará todas las propiedades requeridas. Más adelante veremos que estos dos espacios vectoriales ($\mathbb{C}(\mathbb{C})$ y $\mathbb{C}(\mathbb{R})$) son completamente distintos (verificarán propiedades distintas), con lo cual, cuando trabajemos con un conjunto que puede ser un espacio vectorial sobre varios cuerpos, hemos de tener en mente, en cada caso, cuál es el cuerpo que estamos considerando.

De la definición de espacio vectorial V sobre un cuerpo K se deducen las siguientes propiedades:

1. $0u = 0,\ \forall\, u \in V$.

   Demostración. Puesto que $0u = (0+0)u = 0u+0u$. Si ahora le sumamos a cada miembro el opuesto de $0u$, nos quedará $0 = 0u$. □

2. $a0 = 0,\ \ \forall\, a \in K$.

   Demostración. La prueba es análoga a la anterior. □

3. Sean $a \in K$, $u \in V$. Si $au = 0$, entonces o bien $a = 0$ o bien $u = 0$.

   Demostración. Si $a = 0$ no hay nada que demostrar, pues ya se ha visto. Si, por el contrario, $a \neq 0$, como K es un cuerpo, $\exists\, a^{-1} \in K$. Entonces $a^{-1}(au) = a^{-1}0 = 0 = (a^{-1}a)u = 1u = u$, lo cual, si $a \neq 0$, necesariamente $u = 0$. □

4. $-(au) = (-a)u = a(-u)$.

Demostración. Como $au + (-a)u = (a + (-a))u = 0u = 0$, tenemos que $(-a)u = -(au)$. Análogamente $au + a(-u) = a(u + (-u)) = a0 = 0$, con lo que $a(-u) = -(au)$. □

A partir de ahora a cualquier de esos términos lo notaremos simplemente $-au$.

5. $\forall\, a \in K$, $\forall\, u, v \in V$ se tiene $a(u - v) = au + av$, aplicando directamente la propiedad anterior.

6. Análogamente, $\forall\, a, b \in K$, $\forall\, u \in U$ se tiene $(a - b)u = au - bu$.

7. Si $a, b \in K$, $u \in U$, siendo $u \neq 0$, y $au = bu$, entonces $a = b$.

   Demostración. Como $au = bu$ entonces $au - bu = 0 = (a - b)u$ y, como suponemos que $u \neq 0$, entonces $a - b = 0$ de donde $a = b$. □

8. $\forall\, a \in K \setminus \{0\}$, $\forall\, u, v \in V$ si $au = av$, entonces $u = v$, similarmente a lo anterior.

2. Dependencia e independencia lineal.

Dado un conjunto finito de vectores $v_1, \ldots, v_n$ en un espacio vectorial V sobre el cuerpo K, llamaremos una combinación lineal de estos vectores a cualquier vector de la forma $a_1v_1 + a_2v_2 + \cdots + a_nv_n$, donde $a_1, \ldots, a_n \in K$.

Por ejemplo, si en $\mathbb{R}^3$ consideramos los vectores $(1, 1, 0)$ y $(0, 2, 1)$, el vector $(3, -1, -2)$ es combinación de dichos vectores, pues $3(1, 1, 0) - 2(0, 2, 1) = (3, -1, -2)$. Sin embargo el vector $(0, 0, 1)$ no es combinación lineal de tales vectores, pues necesitaríamos que existieran $a, b \in \mathbb{R}$ tales que $a(1, 1, 0) + b(0, 2, 1) = (0, 0, 1)$ lo que nos llevaría a que $a = 0, a + 2b = 0, b = 1$, pero siendo $a = 0$, la segunda ecuación nos daría $b = 0$ y, por tanto, una contradicción con el hecho de que $b = 1$.

Diremos que el conjunto $\{v_1, \ldots, v_n\}$ es linealmente dependiente o que los vectores $v_1, \ldots, v_n$ son linealmente dependientes si podemos escribir el vector 0 como combinación lineal de ellos con no todos los escalares nulos; es decir, si $\exists\, a_1, \ldots, a_n \in K$ no todos nulos tales que

$$0 = a_1v_1 + a_2v_2 + \cdots + a_nv_n.$$

Diremos que el conjunto anterior es linealmente independiente o que los vectores $v_1, \ldots, v_n$ son linealmente independientes si no son linealmente dependientes, esto es, si cada combinación lineal

$$a_1v_1 + a_2v_2 + \cdots + a_nv_n = 0$$

nos lleva a que $a_1 = a_2 = \cdots = a_n = 0$.

Notad que demostrar la dependencia o independencia lineal de un cierto conjunto de vectores será equivalente a decidir si un cierto sistema de ecuaciones lineales homogéneo es indeterminado o determinado.

Por ejemplo, consideremos en $\mathbb{R}^3$ el siguiente conjunto de vectores: $\{(1, 2, 0), (1, 1, 0), (1, 1, 1), (3, 0, 0)\}$.

Para ello, si consideramos la combinación lineal

$$a(1, 2, 0) + b(1, 1, 0) + c(1, 1, 1) + d(3, 0, 0) = (0, 0, 0),$$

obtenemos el siguiente sistema homogéneo con incógnitas a, b, c y d:

$$\begin{cases} a & +b & +c & +3d & = 0 \\ 2a & +b & +c & & = 0 \\ & & +c & & = 0 \end{cases}$$

cuya matriz de coeficientes es

$$\begin{pmatrix} 1 & 1 & 1 & 3 \\ 2 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}$$

(Fijaos en que las columnas de esta matriz son los vectores del conjunto). Como tenemos un sistema homogéneo, para que fuera determinado (y entonces su única solución sería la trivial $a = b = c = d = 0$, lo que nos diría que los vectores son linealmente independientes) el rango de dicha matriz habría de coincidir con el número de incógnitas (4). Pero esa matriz es de orden 3×4 y su rango máximo sería 3. Como consecuencia, nuestro conjunto de vectores es linealmente dependiente.

Si nos fijamos en la submatriz que obtenemos al eliminar la 4ª columna

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

podemos ver que su determinante es -1, con lo cual la matriz es regular, así el sistema

$$\begin{cases} a & +b & +c & = 0 \\ 2a & +b & +c & = 0 \\ & & +c & = 0 \end{cases}$$

es de Cramer y, por tanto $a = b = c = 0$ es su única solución. Hemos demostrado entonces que el subconjunto $\{(1, 2, 0), (1, 1, 0), (1, 1, 1)\}$ del conjunto anterior sí es linealmente independiente.

Sea ahora el espacio vectorial $\mathcal{P}(\mathbb{R})$ y en él consideramos los polinomio $p(x) = x^2 + x$, $q(x) = 3x - 2$, $r(x) = x^2 + 1$. Tomamos una combinación lineal de ellos igualada al polinomio 0

$$ap(x) + bq(x) + cr(x) = 0.$$

Tendremos entonces $0 = a(x^2 + x) + b(3x - 2) + c(x^2 + 1) = (a + c)x^2 + (a + 3b)x - 2b + c$. Para que este polinomio sea el polinomio 0, hemos de tener

$$\begin{cases} a & & +c & = 0 \\ a & +3b & & = 0 \\ & -2b & +c & = 0 \end{cases}$$

Como la matriz de coeficientes de este sistema es

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

para calcular su determinante le restamos a la 2ª fila la 1ª: $\det(A) = \begin{vmatrix} 1 & 0 & 1 \\ 0 & 3 & -1 \\ 0 & -2 & 1 \end{vmatrix} = \begin{vmatrix} 3 & -1 \\ -2 & 1 \end{vmatrix} = 3-2 = 1$, luego el sistema es de Cramer y tendremos $a = b = c$ por lo que nuestro conjunto de vectores es linealmente independiente.

Proposición 1 (Propiedades relacionadas con la dependencia e independencia lineal)**.**
Sea $V(K)$ un espacio vectorial. Se tiene entonces:

1. *Si $0 \in \{v_1, \ldots, v_n\}$, entonces $\{v_1, \ldots, v_n\}$ es linealmente dependiente.*

2. *$\{v\}$ es linealmente independiente si y solo si $v \neq 0$.*

3. *Si $\{v_1, \ldots, v_n\}$ es linealmente dependiente, cualquier conjunto que lo contenta $\{v_1, \ldots, v_n, v_{n+1}, \ldots, v_{n+r}\}$ también es linealmente dependiente.*

4. *Si $\{v_1, \ldots, v_n, v_{n+1}, \ldots, v_{n+r}\}$ es linealmente independiente, cualquier subconjunto suyo no vacío $\{v_1, \ldots, v_n\}$ sigue siendo linealmente independiente.*

Demostración.

1. Para demostrar esta propiedad, basta tomar una combinación lineal tal que el escalar que multiplica al vector 0 sea 1 y el resto de escalares sean nulos. Esta combinación lineal nos dará el vector 0 pero uno de los escalares que intervienen en $\neq 0$.

2. Para esta propiedad, sabemos, por lo anterior, que $\{0\}$ es linealmente dependiente. Así solo hemos de ver que si $v \neq 0$, $\{v\}$ es linealmente independiente. Si para algún escalar a, $av = 0$, como suponemos $v \neq 0$, tendremos que $a = 0$ y tenemos lo que queríamos.

3. Sea $a_1v_1 + a_2v_2 + \cdots + a_nv_n = 0$ una combinación lineal tal que no todos los escalares que intervienen sean nulos. Supongamos entonces que $\exists\, i \in \{1, \ldots, n\} \,/\, a_i \neq 0$. Tenemos entonces que $a_1v_1 + \cdots + a_iv_i + \cdots + a_nv_n + 0v_{n+1} + \cdots + 0v_{n+r} = 0$ y como $a_i \neq 0$, el conjunto $\{v_1, \ldots, v_n, v_{n+1}, \ldots, v_{n+r}\}$ es linealmente dependiente.

4. Por último, para la cuarta propiedad, supongamos que tenemos una combinación lineal $a_1v_1 + \cdots + a_nv_n = 0$. Como queremos ver que $\{v_1, \ldots, v_n\}$ es linealmente independiente, tendremos que llegar a que $a_1 = \cdots = a_n = 0$. Ahora bien, de lo anterior tenemos que $a_1v_1 + \cdots + a_nv_n + 0v_{n+1} + \cdots + 0v_{n+r} = 0$ y, como estos vectores sí sabemos que son linealmente independientes, efectivamente $a_1 = \cdots = a_n = 0$.

□

Tenemos, además, el siguiente importante resultado:

Teorema 1.
Un conjunto de vectores $\{v_1, \ldots, v_n\}$ es linealmente dependiente si y solo si alguno de dichos vectores es combinación lineal de los restantes.

Demostración. $\Longrightarrow$) Para ver esta implicación, supongamos que $\{v_1, \ldots, v_n\}$ es linealmente dependiente. Entonces podremos encontrar escalares $a_1, \ldots, a_n$ no todos nulos (vamos a suponer, por comodidad que $a_1 \neq 0$, esto no nos quita generalidad a la demostración) tales que $a_1v_1 + a_2v_2 + \cdots + a_nv_n = 0$. Como $a_1 \in K \setminus \{0\}$, $\exists\, a^{-1} \in K$ y entonces podemos escribir

$$v_1 = -(a_1^{-1}a_2)v_2 - (a_1^{-1}a_3)v_3 - \cdots - (a_1^{-1}a_n)v_n$$

con lo que, efectivamente, v_1 será combinación lineal de los demás. (Fijaos que esto ocurrirá para cualquier vector del conjunto cuyo coeficiente en la anterior combinación lineal sea $\neq 0$ y no para un vector arbitrario: en $\mathbb{R}^2$ los vectores $\{(1,-1),(2,0),(3,-3)\}$ son linealmente dependientes pues $-3(1,-1)+0(2,0)+1(3,-3) = 0$; sin embargo, si $(2,0) = a(1,-1)+b(3,-3)$ lo que nos da un sistema

$$\begin{cases} 2 = a + 3b \\ 0 = -a - 3b \end{cases}$$

que es incompatible, al ser $2 \neq 0$).
$\Longleftarrow$) Recíprocamente, si suponemos, por ejemplo, que $v_1 = b_2v_2 + \cdots + b_nv_n$ con $b_2, \ldots, b_n \in K$, tendremos que $(-1)v_1 + b_2v_2 + \cdots + b_nv_n = 0$ y, como $-1 \neq 0$, los vectores son linealmente dependientes. $\square$

Antes hablamos de que un mismo conjunto que sea espacio vectorial, a la vez, sobre dos cuerpos distintos, tendrá propiedades distintas. Vamos a ver ahora una de ellas: sea $\mathbb{C}$ que sabemos que es un espacio vectorial real $\mathbb{C}(\mathbb{R})$ y también un espacio vectorial complejo $\mathbb{C}(\mathbb{C})$. Consideremos el subconjunto de $\mathbb{C}$, $\{1, i\}$. Veamos que este conjunto es linealmente dependiente en $\mathbb{C}(\mathbb{C})$: $1, i \in \mathbb{C}$ son escalares no nulos y tenemos que $1 \cdot 1 + i \cdot i = 1 - 1 = 0$. Sin embargo, en $\mathbb{C}(\mathbb{R})$ es linealmente independiente: si $a, b \in \mathbb{R}$ tales que $a \cdot 1 + bi = 0$, tenemos el número complejo $a + bi$ y, para que se anule su parte real, a, y su parte imaginaria, b, han de ser $a = b = 0$.
Hemos definido la dependencia e independencia lineales para subconjuntos finitos de V. Si S es un conjunto arbitrario de vectores (puede, por tanto, ser infinito), diremos que es linealmente dependiente si podemos encontrar un subconjunto finito suyo que sea linealmente dependiente. Si cualquier subconjunto finito de S es linealmente independiente, diremos que S también lo es.

3. Sistemas de generadores de un espacio vectorial.

Un conjunto de vectores S de $V(K)$ es un sistema de generadores de dicho espacio vectorial si cualquier vector de V se puede expresar como combinación lineal de los vectores de S (en el caso de que S sea infinito, consideraremos combinaciones lineales de un número finito de vectores de S).
Por ejemplo, en $\mathbb{R}^2$, el conjunto $\{(1,-1),(1,0),(2,1)\}$ es un sistema de generadores: sea $(x,y) \in \mathbb{R}^2$ un vector cualquiera. Tendremos que ver que $\exists\, a, b, c \in \mathbb{R}$ tales que $a(1,-1)+b(1,0)+c(2,1) = (x,y)$. Luego $(a+b+2c, -a+c) = (x,y)$. Tenemos entonces el sistema de ecuaciones lineales en las incógnitas a, b y c:

$$\begin{cases} a \quad +b \quad +2c \quad = x \\ -a \qquad\quad +c \quad = y \end{cases}$$

Como $\begin{vmatrix} 1 & 1 \\ -1 & 0 \end{vmatrix} = 1 \neq 0$, a y b son las incógnitas principales del sistema y $a = c - y$, $b = x - 2c - a = x - 2c - c + y = x + y - 3c$, nos da la solución general del sistema. Si, por ejemplo, tomamos $c = 0$, $a = -y$, $b = x + y$ sería una solución y hemos visto lo que queríamos.

Si $\{u_1, u_2, \dots, u_n\}$ es un sistema de generadores de $V(K)$ y uno de ellos, por ejemplo u_i, es combinación lineal de los restantes, el conjunto de vectores que se obtiene al eliminar u_i, $\{u_1, \dots, u_{i-1}, u_{i+1}, \dots, u_n\}$ sigue siendo un sistema de generadores de $V(K)$.

Para ello, supongamos que $u_i = b_1u_1 + \cdots + b_{i-1}u_{i-1} + b_{i+1}u_{i+1} + \cdots + b_nu_n$. Sea $v \in V$ un vector arbitrario. Luego encontramos $a_1, \dots, a_{i-1}, a_i, a_{i+1}, \dots, a_n \in K$ tales que

$$x = a_1u_1 + \cdots + a_{i-1}u_{i-1} + a_iu_i + a_{i+1}u_{i+1} + \cdots + a_nu_n =$$

$$= a_1u_1 + \cdots + a_{i-1}u_{i-1} + a_i(b_1u_1 + \cdots + b_{i-1}u_{i-1} + b_{i+1}u_{i+1} + \cdots + b_nu_n) + a_{i+1}u_{i+1} + \cdots + a_nu_n =$$

$$= (a_1 + a_ib_1)u_1 + \cdots + (a_{i-1} + a_ib_{i-1})u_{i-1} + (a_{i+1} + a_ib_{i+1})u_{i+1} + \cdots + (a_n + a_ib_n)u_n.$$

Como cada $a_j + a_ib_j \in K$, $\forall\, j \in \{1, \dots, n\} \setminus \{i\}$, tenemos que el subconjunto es un sistema de generadores de V.

En el primer ejemplo que dimos, $\{(1,-1), (1,0), (2,1)\}$ es un sistema de generadores de $\mathbb{R}^2$.

Como $(1,-1) = 3(1,0) + (-1)(2,1)$, $(-1,1)$ es combinación lineal de los otros dos vectores. Entonces $\{(1,0), (2,1)\}$ sigue siendo un sistema de generadores de $\mathbb{R}^2$.

Para lo que sigue, el siguiente resultado es fundamental:

Proposición 2.

Sea $\{v_1, v_2, \dots, v_m\}$ un conjunto linealmente independiente en $V(K)$ y $\{u_1, u_2, \dots, u_s\}$ un sistema de generadores del mismo espacio. Entonces $m \leq s$.

Demostración. Como $\{u_1, u_2, \dots, u_s\}$ es un sistema de generadores de V, también lo será $\{v_1, u_1, u_2, \dots, u_s\}$. Por otra parte, podremos escribir v_1 como combinación lineal de los vectores $\{u_1, \dots, u_s\}$, $v_1 = a_1u_1 + a_2u_2 + \cdots + a_su_s$ y como $v_1 \neq 0$ ya que forma parte de un conjunto linealmente independiente, algunos de los escalares anteriores habrá de ser no nulo. Por simplicidad supongamos que $a_1 \neq 0$. Así, $u_1 = a_1^{-1}v_1 - a_1^{-1}a_2u_2 - \cdots - a_1^{-1}a_su_s$, es combinación lineal de los demás y en $\{v_1, u_1, u_2, \dots, u_s\}$ lo podemos eliminar, de manera que $\{v_1, u_2, \dots, u_s\}$ es otro sistema de generadores de V.

Ahora repetimos el proceso con v_2, $\{v_1, v_2, u_2, \dots, u_s\}$ será un sistema de generadores y escribiremos $v_2 = b_1v_1 + b_2u_2 + \cdots + b_su_s$ y, además, alguno de los escalares $b_2, \dots, b_s$ ha de ser no nulo. En caso contrario $v_2 = b_1v_1$ nos diría que $\{v_1, v_2\}$ sería linealmente dependiente, en contra de la hipótesis. Supongamos pues que $b_2 \neq 0$, con lo que, como antes, u_2 sería una combinación lineal de $\{v_1, v_2, u_3, \dots, u_s\}$. Por tanto lo eliminamos, obteniendo que $\{v_1, v_2, u_3, \dots, u_s\}$ es un sistema de generadores. Podemos repetir este proceso hasta agotar alguno de los conjuntos, pero si en el conjunto linealmente independiente tuviéramos más de s vectores, v_{s+1} sería una combinación lineal de $v_1, v_2, \dots, v_s$, en contra de que el conjunto que los contiene a todos es linealmente independiente. $\square$

4. Bases de un espacio vectorial.

Definición 1 (Base de un espacio vectorial)**.**
Sea $V(K)$ un espacio vectorial y $B \subseteq V$. Diremos que $\mathcal{B}$ es una base de V si:

1. *$\mathcal{B}$ es linealmente independiente, y*
2. *$\mathcal{B}$ es un sistema de generadores de V.*

Teorema 2 (de la Base)**.**
Si un espacio vectorial V tiene una base formada por un número finito de vectores, entonces todas las bases de V son finitas y tienen el mismo número de vectores.

Demostración. Sea $\mathcal{B} = \{v_1, v_2, \ldots, v_n\}$ una base de V y $\mathcal{B}'$ otra base de V. Como $\mathcal{B}'$ es linealmente independiente y $\mathcal{B}$ es un sistema de generadores de V, el número de elementos de $\mathcal{B}'$ es $\leq n$, y por tanto, finito. Como además $\mathcal{B}'$ es un sistema de generadores de V y $\mathcal{B}$ es linealmente independiente, si m, finito, es el número de elementos de $\mathcal{B}'$, $n \leq m$. Luego $n = m$, demostrando el enunciado. $\square$

Si un espacio vectorial $V(K)$ admite una base finita diremos que es un espacio vectorial de dimensión finita, siendo la dimensión de V el número de vectores de una de sus bases, y lo denotaremos $\dim_K(V)$ (vemos por el ejemplo anterior que $\dim_{\mathbb{C}}(\mathbb{C}) = 1$, pero $\dim_{\mathbb{R}}(\mathbb{C}) = 2$. En el primer caso $\{1\}$ es una base de $\mathbb{C}(\mathbb{C})$, mientras que en el caso de $\mathbb{C}(\mathbb{R})$, $\{1, i\}$ sería una base) ó, si no hay posibilidad de equivocación sobre el cuerpo sobre el que V es espacio vectorial, $\dim(V)$.

Ejemplo 2. 1. En el espacio vectorial K^n, para un cuerpo conmutativo K,

$$\mathcal{B}_U = \{(1,0,0,\ldots,0), (0,1,0,\ldots,0), \ldots, (0,0,0,\ldots,1)\}$$

es una base que llamaremos la base canónica o usual[1] de K^n. Veamos, primero, que es un sistema de generadores: cuando consideremos cualquier columna de términos independientes, el sistema cuya matriz de coeficientes es I_n (sus columnas son los vectores de $\mathcal{B}_U$) es un sistema de Cramer con solución única.
Pero, además, son linealmente independientes: si tenemos

$$a_1(1,0,0,\ldots,0) + a_2(0,1,0,\ldots,0) + \cdots + a_n(0,0,0,\ldots,1) = (0,0,0,\ldots,0)$$

obtenemos el siguiente sistema

$$\begin{cases} a_1 & & & & = 0 \\ & a_2 & & & = 0 \\ & & \ddots & & = 0 \\ & & & a_n & = 0 \end{cases}$$

que, por tanto, solo tiene la solución trivial. Así, $\dim_K(K^n) = n$.

[1] De ahí el subíndice U.

2. $\mathcal{P}(K)$ no tiene dimensión finita: para ello supongamos que $\dim_K \mathcal{P}(K) = m$ y sea $\{p_1(x), \ldots, p_m(x)\}$ una base de $\mathcal{P}(K)$. Llamemos $g_i = \text{grado}(p_i(x))$, $i \in \{1, \ldots, m\}$ y $r = \max\limits_{1 \le i \le m} \{g_i\}$. Luego los grados de todos los polinomios son, como máximo r. Si ahora consideramos el polinomio x^{r+1} habría de ser una combinación lineal de $p_1(x), \ldots, p_m(x)$. Pero el grado del polinomio que obtenemos mediante dicha combinación lineal es menor o igual que r, llegando a una contradicción.

3. Sea $\mathcal{M}_{m\times n}(K)$. $\forall\, i \in \{1, \ldots, m\}, j \in \{1, \ldots, n\}$, llamemos A_{ij} a la matriz que tiene un 1 en la posición ij y 0 en el resto de posiciones. Entonces $\mathcal{B} = \{A_{ij} \,/\, 1 \le i \le m, 1 \le j \le n\}$ es una base, que llamaremos la base estándar de $\mathcal{M}_{m\times n}(K)$. Así, $\dim_K(\mathcal{M}_{m\times n}(K)) = mn$.

4. El espacio vectorial trivial $\{0\}$ tiene dimensión 0, puesto que, aunque tiene un sistema de generadores formado por un número finito de vectores, $\{0\}$, este conjunto no es linealmente independiente.

Aunque ya hemos visto que existen espacios vectoriales con dimensión no finita, a partir de ahora solo consideraremos espacios vectoriales con dimensión finita.

Si consideramos un espacio vectorial no nulo, de cada sistema de generadores finito S podremos extraer una base: si el sistema de generadores es linealmente independiente ya es una base y no tendremos que hacer nada. Si no, uno de los vectores de S se podrá expresar como combinación lineal de los restantes, y el conjunto S' que obtenemos al eliminarlo de S seguirá siendo un sistema de generadores. Si S' es linealmente independiente, ya tenemos la base. Si no, repetimos el proceso. Eventualmente, llegaríamos a un sistema de generadores con un solo elemento $\{v\}$ de forma que $v \neq 0$, pues hemos supuesto que $V \neq \{0\}$. Así, $\{v\}$ sería linealmente independiente y, por tanto, una base.

El siguiente resultado nos da otro método para encontrar bases de un espacio vectorial:

Teorema 3 (de ampliación de la Base)**.**

Sea $V(K)$ un espacio vectorial de dimensión n y $\{v_1, v_2, \ldots, v_r\}$ un conjunto linealmente independiente. Entonces podemos encontrar vectores $v_{r+1}, \ldots, v_n$ tales que $\{v_1, v_2, \ldots, v_r, v_{r+1}, \ldots, v_n\}$ sea una base de V.

Demostración. Si $\{v_1, v_2, \ldots, v_r\}$ es un sistema de generadores ya sería una base y $r = n$. Si no, $\exists\, v_{r+1} \in V$, $v_{r+1} \neq 0$ que no es combinación lineal de $\{v_1, v_2, \ldots, v_r\}$. Veamos que, entonces, $\{v_1, v_2, \ldots, v_r, v_{r+1}\}$ es también linealmente independiente. Si encontramos escalares $a_1, \ldots, a_{r+1}$ tales que $0 = a_1v_1 + a_2v_2 + \cdots + a_rv_r + a_{r+1}v_{r+1}$, hemos de tener que $a_{r+1} = 0$, pues, en caso contrario, podríamos expresar v_{r+1} como una combinación lineal de los demás vectores, pero hemos supuesto que eso no ocurre. El hecho de ser $a_{r+1} = 0$ nos lleva a que la anterior combinación lineal se reduce a $0 = a_1v_1 + a_2v_2 + \cdots + a_rv_r$, y, como $\{v_1, \ldots, v_r\}$ es linealmente independiente, $a_1 = a_2 = \cdots = a_r = 0$. Así, $\{v_1, \ldots, v_r, v_{r+1}\}$ es linealmente independiente. Si fuese, además, un sistema de generadores, habríamos terminado. Si no, repetiríamos el proceso para un vector v_{r+2} que no fuese combinación lineal del anterior conjunto de vectores. Finalmente, llegaríamos a un conjunto $\{v_1, v_2, \ldots, v_s, v_{s+1}, \ldots, v_n\}$ que sería linealmente independiente y sistema de generadores, siendo la base buscada. $\square$

Si sabemos, entonces, que $\dim_K(V) = n$ y en V tomamos un conjunto de n vectores $S = \{v_1, v_2, \ldots, v_n\}$, las siguiente afirmaciones son equivalentes:

1. S es linealmente independiente,
2. S es un sistema de generadores de V,
3. S es una base de V.

5. Coordenadas de un vector respecto de una base.

El hecho que presentamos a continuación nos va a permitir trabajar en cada espacio vectorial de dimensión n como en el espacio K^n:

Proposición 3.
Sea $V(K)$ un espacio vectorial. Si $\mathcal{B} = \{v_1, v_2, \ldots, v_n\}$ es una base de V, entonces todo vector $x \in V$ admite una única expresión como combinación lineal de los vectores de $\mathcal{B}$.

Demostración. Ya sabemos que, al ser $\mathcal{B}$ un sistema de generadores de V, dado $x \in V$, podremos encontrar escalares $a_1, \ldots, a_n \in K$ tales que $x = a_1v_1 + a_2v_2 + \cdots + a_nv_n$. Lo que tenemos que ver es que $a_1, \ldots, a_n$ son únicos. Supongamos, por contra, que encontramos $b_1, \ldots, b_n \in K$ tales que $x = b_1v_1 + b_2v_2 + \cdots + b_nv_n$. Restando ambas expresiones llegamos a que $0 = (a_1 - b_1)v_1 + (a_2 - b_2)v_2 + \cdots + (a_n - b_n)v_n$ y, como $\{v_1, \ldots, v_n\}$ es linealmente independiente, $a_1 - b_1 = a_2 - b_2 = \cdots = a_n - b_n = 0$. Esto nos dice que $\forall\, i \in \{1, \ldots, n\}$, $a_i = b_i$, como queríamos. □

Así, si dado $x \in V$, $x = x_1v_1 + x_2v_2 + \cdots + x_nv_n$ es la expresión única del vector x como combinación lineal de los vectores de la base $\mathcal{B}$, diremos que $(x_1, x_2, \ldots, x_n)$ son las coordenadas de x respecto de la base $\mathcal{B}$ y lo escribiremos $x = (x_1, x_2, \ldots, x_n)_{\mathcal{B}}$. Así, si otro vector $y \in V$ tiene por coordenadas respecto de $\mathcal{B}$ $y = (y_1, y_2, \ldots, y_n)_{\mathcal{B}}$, tendremos $x = x_1v_1 + x_2v_2 + \cdots + x_nv_n$, $y = y_1v_1 + y_2v_2 + \cdots + y_nv_n$. Así, $x + y = (x_1 + y_1)v_1 + (x_2 + y_2)v_2 + \cdots + (x_n + y_n)v_n$. Es decir, $x + y = (x_1 + y_1, x_2 + y_2, \ldots, x_n + y_n)_{\mathcal{B}}$. Además, $kx = kx_1v_1 + kx_2v_2 + \cdots + kx_nv_n$. Por tanto, $kx = (kx_1, kx_2, \ldots, kx_n)_{\mathcal{B}}$.
Representando así cada vector de V mediante sus coordenadas respecto de la base $\mathcal{B}$ (que no son más que un elemento de K^n), las coordenadas de la suma de dos vectores vienen dadas por la suma de los correspondientes elementos de K^n, mientras que las del producto de un vector por un escalar, vienen dadas del producto por ese escalar del correspondiente elemento de K^n.

6. Coordenadas y dependencia lineal.

Sea $V(K)$ un espacio vectorial y $\mathcal{B}$ una base de V. Entonces un conjunto de n vectores $\{v_1, \ldots, v_n\}$ en V es linealmente independiente si y solo si la matriz cuyas columnas (o filas) son las coordenadas de dichos vectores respecto de $\mathcal{B}$ tiene rango n.
Como una combinación lineal de los vectores $v_1, \ldots, v_n$ igualada a 0 es de la forma $x_1v_1 + x_2v_2 + \cdots + x_nv_n = 0$ con $x_1, \ldots, x_n \in K$, si usamos las coordenadas respecto de la base $\mathcal{B}$ de $v_1, v_2, \ldots, v_n$, obtendremos un

sistema de ecuaciones lineales homogéneo con incógnitas $x_1, x_2, \ldots, x_n$ y cuya matriz de coeficientes A es precisamente la matriz cuyas columnas son las coordenadas de los vectores $v_1, \ldots, v_n$ respecto de $\mathcal{B}$. Los vectores serán, pues, linealmente dependientes si y solo si dicho sistema admite alguna solución distinta de la trivial lo que equivale por el Teorema de Rouché-Frobenius a que $\mathrm{rg}(A) < n$. Como el $\mathrm{rg}(A) = \mathrm{rg}(A^t)$, el resultado sigue siendo el mismo si consideramos las filas de A en lugar de las columnas.

Ejemplo 3.
Supongamos en $\mathbb{R}^4$ los vectores $v_1 = (1, -1, 0, 1)$, $v_2 = (0, 1, 1, 1)$ y $v_3 = (3, 0, 3, 6)$. En la base usual de $\mathbb{R}^4$, las coordenadas de cada uno de esos vectores son precisamente sus componentes. Así, obtenemos la matriz A cuyas columnas son esas coordenadas,

$$A = \begin{pmatrix} 1 & 0 & 3 \\ -1 & 1 & 0 \\ 0 & 1 & 3 \\ 1 & 1 & 6 \end{pmatrix}.$$

Como en A la submatriz $\begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix}$ tiene determinante $1 \neq 0$, el rango de A es, al menos, 2 y por tanto, los vectores v_1 y v_2 son linealmente independientes. Si ahora calculamos

$$\begin{vmatrix} 1 & 0 & 3 \\ -1 & 1 & 0 \\ 0 & 1 & 3 \end{vmatrix} = 3 - 3 = 0 \quad \text{y} \quad \begin{vmatrix} 1 & 0 & 3 \\ -1 & 1 & 0 \\ 1 & 1 & 6 \end{vmatrix} = 6 - 3 - 3 = 0$$

por lo que concluimos que $\mathrm{rg}(A) = 2$ y $\{v_1, v_2, v_3\}$ es linealmente dependiente.

7. Cambio de base

Supongamos que tenemos un espacio vectorial $V(K)$ de dimensión n y que consideramos en V dos bases distintas

$$\mathcal{B} = \{u_1, u_2, \ldots, u_n\}, \quad \mathcal{B}' = \{v_1, v_2, \ldots, v_n\}.$$

Así, un vector dado de V tendrá unas coordenadas respecto de $\mathcal{B}$ y otras (distintas) respecto de $\mathcal{B}'$. Sea $v \in V$ y supongamos que $(x_1, x_2, \ldots, x_n)$ son las coordenadas de x en $\mathcal{B}$ mediante que $(y_1, y_2, \ldots, y_n)$ son sus coordenadas respecto de $\mathcal{B}'$. Como $\mathcal{B}$ es una base de V y los vectores de $\mathcal{B}'$ son de V, supongamos que conocemos las coordenadas de cada vector de $\mathcal{B}'$ respecto de $\mathcal{B}$; esto es, supongamos que

$$\begin{aligned} v_1 &= a_{11}u_1 + a_{21}u_2 + \cdots + a_{n1}u_n \\ v_2 &= a_{12}u_1 + a_{22}u_2 + \cdots + a_{n2}u_n \\ &\vdots \\ v_n &= a_{1n}u_1 + a_{2n}u_2 + \cdots + a_{nn}u_n \end{aligned}$$

Entonces,

$$x = y_1v_1 + y_2v_2 + \cdots + y_nv_n =$$

$$= y_1(a_{11}u_1 + a_{21}u_2 + \cdots + a_{n1}u_n) + y_2(a_{12}u_1 + a_{22}u_2 + \cdots + a_{n2}u_n) + \cdots + y_n(a_{1n}u_1 + a_{2n}u_2 + \cdots + a_{nn}u_n) =$$

$$= (a_{11}y_1 + a_{12}y_2 + \cdots + a_{1n}y_n)u_1 + (a_{21}y_1 + a_{22}y_2 + \cdots + a_{2n}y_n)u_2 + \cdots + (a_{n1}y_1 + a_{n2}y_2 + \cdots + a_{nn}y_n)u_n =$$

$$= x_1u_1 + x_2u_2 + \cdots + x_nu_n.$$

Tenemos, pues, dos expresiones del vector x respecto de la base $\mathcal{B}$ y, siendo dicha expresión única, obtenemos que

$$\begin{aligned} x_1 &= a_{11}y_1 + a_{12}y_2 + \cdots + a_{1n}y_n \\ x_2 &= a_{21}y_1 + a_{22}y_2 + \cdots + a_{2n}y_n \\ &\vdots \\ x_n &= a_{n1}y_1 + a_{n2}y_2 + \cdots + a_{nn}y_n. \end{aligned}$$

O bien, en forma matricial, $X = PY$, donde X e Y son las matrices columna que nos dan las coordenadas de un vector arbitrario de V en las bases $\mathcal{B}$ y $\mathcal{B}'$, respectivamente, y P es la matriz cuyas columnas son las coordenadas de cada vector de $\mathcal{B}'$ respecto de la base $\mathcal{B}$. Esta matriz P se llama la matriz del cambio de base de $\mathcal{B}'$ a $\mathcal{B}$ y es una matriz regular pues sus columnas son las coordenadas en $\mathcal{B}$ de los vectores de $\mathcal{B}'$ que son linealmente independientes, por lo que su rango es n. Así P tiene matriz inversa, y la anterior ecuación matricial nos da $Y = P^{-1}X$ que nos da las coordenadas de un vector de V en la base $\mathcal{B}'$, conocidas sus coordenadas en $\mathcal{B}$. Así, P^{-1} es la matriz de cambio de base de $\mathcal{B}$ a $\mathcal{B}'$.
Así, cada matriz de cambio de base es regular. El recíproco también es cierto: toda matriz regular es una matriz de cambio de base, ya que si $Q \in \mathcal{M}_n(K)$ y es regular, podemos considerar sus columnas como n vectores de K^n que, siendo Q regular, han de ser linealmente indpendientes. Como $\dim_K(K^n) = n$, estos n vectores de K^n forman una base de K^n, que llamamos $\mathcal{B}'$. Si consideramos las coordenadas de cada vector de $\mathcal{B}'$ en $\mathcal{B}_U$ de K y las escribimos como una matriz columna, todas juntas nos dan Q, que, por tanto, es la matriz de cambio de base de $\mathcal{B}'$ a $\mathcal{B}_U$.

Ejemplo 4.
En $\mathbb{R}^2$ consideramos $\mathcal{B} = \{(1,0),(1,1)\}$ y $\mathcal{B}' = \{(1,2),(1,3)\}$. Como $\begin{vmatrix} 1 & 1 \\ 0 & 1 \end{vmatrix} = 1$ y $\begin{vmatrix} 1 & 1 \\ 2 & 3 \end{vmatrix} = 1$, ambas nos proporcionan bases de $\mathbb{R}^2$.
Además, $(1,2) = a(1,0) + b(1,1) = (a+b,b)$ implica $a+b = 1, b = 2$. Así, $a = -1$. Luego las coordenadas de $(1,2)$ en $\mathcal{B}$ son $(-1,2)$.
Por otro lado, $(1,3) = a(1,0) + b(1,1) = (a+b,b)$ implica $b = 3, a+b = 1$, luego $a = -2$. Por tanto las coordenadas de $(1,3)$ en $\mathcal{B}$ son $(-2,3)$.
Así, la matriz del cambio de base de $\mathcal{B}'$ a $\mathcal{B}$ es

$$P = \begin{pmatrix} -1 & -2 \\ 2 & 3 \end{pmatrix}.$$

A partir de ahora esta matriz la escribiremos $\mathcal{M}_{\mathcal{B}\leftarrow\mathcal{B}'}$. Como $P^{-1} = \begin{pmatrix} 3 & 2 \\ -2 & -1 \end{pmatrix}$, esta será la matriz del cambio de base de $\mathcal{B}$ a $\mathcal{B}'$, $\mathcal{M}_{\mathcal{B}'\leftarrow\mathcal{B}}$.

8. Subespacios vectoriales.

Definición 2.
Sea $V(K)$ un espacio vectorial y U un subconjunto no vacío de V. Diremos que U es un subespacio vectorial de V si se verifican las siguientes condiciones:

1. *U es cerrado para la suma: $\forall\, u, v \in U$, $u + v \in U$.*
2. *U es cerrado para el producto por escalares: $\forall\, a \in K$, $\forall\, u \in U$, $au \in U$.*

Como en el caso de los subgrupos de un grupo, es fácil ver que si U es un subespacio vectorial de V, entonces U, con las operaciones que hacen que V sea un espacio vectorial sobre K, es también un espacio vectorial sobre K.
Así, todo espacio vectorial distinto del espacio vectorial trivial tiene, al menos, dos subespacios vectoriales: él mismo y el formado solo por el vector nulo, $\{0\}$. Estos subespacios vectoriales se llaman subespacios impropios.
Consideramos el espacio vectorial $\mathcal{M}_n(K)$. Los siguientes subconjunto de $\mathcal{M}_n(K)$ son subespacios vectoriales:

- Las matrices triangulares superiores, pues si sumamos dos matrices triangulares superiores obtenemos otra matriz triangular superior y si hacemos el producto por un escalar de una matriz triangular superior, igualmente obtenemos una matriz triangular superior.
- Análogamente, las matrices triangulares inferiores nos dan un subespacio vectorial de $\mathcal{M}_n(K)$.
- Las matrices diagonales (que son a la vez triangulares superiores e inferiores) forman un subespacio vectorial de $\mathcal{M}_n(K)$.
- Las matrices simétricas también constituyen un subespacio vectorial de $\mathcal{M}_n(K)$: si $A, B \in \mathcal{M}_n(K)$ y ambas son simetricas, $A = A^t$, $B = B^t$, entonces $(A + B)^t = A^t + B^t = A + B$, que demuestra que $A + B$ es también simétrica. Además, $\forall\, k \in K$, $(kA)^t = kA^t = kA$ y kA sigue siendo simétrica.
- Tenemos un caso análogo para las matrices antisimétricas.
- Supongamos ahora el subconjunto $\mathcal{P}_n(K)$ de $\mathcal{P}(K)$ dado por los polinomios que tienen grado n. En este caso, $\mathcal{P}_n(K)$ no es cerrado para la suma: si consideramos $p(x) = x^n - 3x$ y $q(x) = -x^n + 5$ (con $n > 1$), entonces $p(x) + q(x) = -3x + 5$, que no tiene grado n. Luego $\mathcal{P}_n(K)$ no es un subespacio vectorial de $\mathcal{P}(K)$. Sin embargo, si notamos por $\mathcal{P}^n(K)$ el subconjunto de $\mathcal{P}(K)$ formado por los polinomios de grado $\leq n$ con coeficientes en K, este sí que es un subespacio vectorial de $\mathcal{P}(K)$

y, además, cada elemento de $\mathcal{P}^n(K)$ es una combinación lineal de los polinomios $\{1, x, x^2, \ldots, x^n\}$. Además, es inmediato que si $a_0 \cdot 1 + a_1 x + a_2 x^2 + \cdots + a_n x^n = 0$, necesariamente $a_0 = a_1 = \cdots = a_n = 0$, lo que nos demuestra que $\{1, x, x^2, \ldots, x^n\}$ es una base de $\mathcal{P}^n(K)$. Así, $\dim_K(\mathcal{P}^n(K)) = n + 1$.

Las dos condiciones que hemos exigido para que un subconjunto de V sea un subespacio vectorial pueden agruparse en una sola, como vemos en el siguiente resultado:

Proposición 4.
Sea $\emptyset \neq U \subseteq V$. Entonces los siguientes hechos son equivalentes:

1. *U es un subespacio vectorial de V.*
2. *$\forall\, u, v \in U$, $\forall\, a, b \in K$, $au + bv \in U$.*

Demostración. Para ver que 1) $\Longrightarrow$ 2), dado que U es un subespacio vectorial de V, si $u, v \in U$ y $a, b \in K$, por la 2ª condición en la definición de subespacio vectorial, $au \in U$ y $bv \in U$. Ahora la primera condición implica que $au + bv \in U$, como queríamos.
Si para el recíproco suponemos que se verifica la condición 2), si queremos ver que U es, efectivamente, un subespacio vectorial de V hemos de probar que se verifican las condiciones de la definición. Para la 1ª, dados $u, v \in U$, tomando $a = b = 1$, de 2) tenemos que $1 \cdot u + 1 \cdot v = u + v \in U$. Y para la 2ª condición, $\forall\, u \in U$, $\forall\, a \in K$, tomamos $v = u$, $b = 0$ y tendremos que $au + 0u = au + 0 = au \in U$. $\square$

Nota 2. Cada subespacio vectorial de un espacio vectorial ha de contener al vector 0, ya que si $u \in U$, $-u = (-1)u \in U$ y, entonces, $u + (-u) = u - u = 0 \in U$. Así, si nos dan un subconjunto de un espacio vectorial que no contenga al vector 0, podemos afirmar directamente que no es un subespacio vectorial.

Sea de nuevo $V(K)$ un espacio vectorial sobre el cuerpo K y S un subconjunto arbitrario de V. Consideremos un nuevo subconjunto de V dado por

$$\mathcal{L}(S) = \{a_1 s_1 + \cdots + a_m s_m \,/\, m \in \mathbb{N},\ a_i \in K,\ s_i \in U,\ i \in \{1, \ldots, m\}\}$$

que, por tanto, consiste en tomar todas las posibles combinaciones lineales de elementos de S. Por ejemplo, si en $\mathbb{R}^3$ consideramos los vectores $s_1 = (1, -1, 0), s_2 = (0, 2, 2)$, sea $S = \{s_1, s_2\}$. Entonces $\mathcal{L}(S) = \{a(1, -1, 0) + b(0, 2, 2) \,/\, a, b \in \mathbb{R}\} = \{(a, -a + 2b, 2b) \,/\, a, b \in \mathbb{R}\}$.
Veamos que $\mathcal{L}(S)$ es el menor subespacio vectorial de V que contiene a S, con lo que queremos expresar que si W es otro subespacio vectorial de V tal que $S \subset W$, entonces $\mathcal{L}(S) \subset W$.
Para ello, hemos de probar, en primer lugar, que $\mathcal{L}(S)$ es un subespacio vectorial de V. Dados $a, b \in K$ y $a_1 s_1 + \cdots + a_m s_m$, $b_1 t_1 + \cdots + b_m t_m$ con $a_i, b_j \in K$, $1 \leq i \leq m$, $1 \leq j \leq n$, $s_1, \ldots, s_m, t_1, \ldots, t_n \in S$, entonces

$$a(a_1 s_1 + \cdots + a_m s_m) + b(b_1 t_1 + \cdots + b_n t_n) = a a_1 s_1 + \cdots + a a_m s_m + b b_1 t_1 + \cdots + b b_n t_n \in \mathcal{L}(S),$$

pues es una combinación lineal de elementos de S.
Además, $\forall\, s \in S$, $s \in \mathcal{L}(S)$ pues es la combinación lineal $1 \cdot s$; así, $S \subset \mathcal{L}(S)$.

Si ahora $S \subset W$, siendo W un subespacio vectorial de V, dado un elemento $a_1s_1 + \cdots + a_rs_r \in \mathcal{L}(S)$, tendremos que $s_1, \ldots, s_r \in S$ y, por tanto, $s_1, \ldots, s_r \in W$ y dados $a_1, \ldots, a_r \in K$, como W es un subespacio vectorial de V, $a_1s_1 + \cdots + a_rs_r \in W$, con lo que, efectivamente, $\mathcal{L}(S) \subset W$.

A $\mathcal{L}(S)$ lo llamaremos el subespacio generado por S, ya que S es, de hecho, un sistema de generadores de $\mathcal{L}(S)$. Como vimos en un ejemplo anterior, para un subespacio U de un espacio vectorial V, al ser él mismo un espacio vectorial, podremos hablar de bases de U y, por tanto, de dimensión de U. En general, tendremos que $\dim_K U \leq \dim_K V$ y se dará la igualdad si y solo si $U = V$.

9. Ecuaciones cartesianas y paramétricas de un subespacio.

Supongamos que tenemos un sistema homogéneo de m ecuaciones lineales con n incógnitas sobre el cuerpo K, $AX = 0$. Cada solución del sistema será un elemento de K^n. El conjunto formado por todas las soluciones de dicho sistema es un subespacio vectorial de K^n: en efecto, si $(s_1, \ldots, s_n)$ y $(t_1, \ldots, t_n)$ son dos soluciones de dicho sistema, llamemos S a la matriz columna $(s_1, \ldots, s_n)^t$ y T a la matriz columna $(t_1, \ldots, t_n)^t$. Por tanto $AS = 0$ y $AT = 0$. Así, $A(S + T) = AS + AT = 0$ y $\forall\, k \in K$, $A(kS) = k(AS) = k0 = 0$, como queríamos probar.
Veremos a continuación que cada subespacio vectorial de un espacio vectorial V de dimensión finita se puede considerar como el conjunto de soluciones de un sistema de ecuaciones. Para ello, elijamos una base $\mathcal{B} = \{v_1, v_2, \ldots, v_n\}$ respecto de la cual vamos a tomar las coordenadas de cada vector. Si U es un subespacio vectorial de V con dimensión $r \leq n$, podemos hallar una base $\mathcal{B}_U = \{u_1, u_2, \ldots, u_r\}$ de U.
Si conocemos las coordenadas de los u_i en función de $\mathcal{B}$,

$$u_i = (a_{1i}, a_{2i}, \ldots, a_{ni})_{\mathcal{B}}, \quad \forall\, 1 \leq i \leq r,$$

como cada vector de U será una combinación lineal de los vectores de $\mathcal{B}_U$, sea $x \in U$. Por un lado, conoceremos sus coordenadas en $\mathcal{B}$, $x = (x_1, x_2, \ldots, x_n)_{\mathcal{B}}$. Además, si $x \in U$, $\exists\, \lambda_1, \ldots, \lambda_r \in K$ tales que $x = \lambda_1 u_1 + \cdots + \lambda_r u_r$. Entonces,

$$x = (x_1, \ldots, x_n)_{\mathcal{B}} = \lambda_1(a_{11}, a_{21}, \ldots, a_{n1})_{\mathcal{B}} + \lambda_2(a_{12}, a_{22}, \ldots, a_{n2})_{\mathcal{B}} + \cdots + \lambda_r(a_{1r}, a_{2r}, \ldots, a_{nr})_{\mathcal{B}}.$$

Así, por la unicidad de las coordenadas de x respecto de $\mathcal{B}$, obtenemos unas ecuaciones paramétricas de U respecto de la base $\mathcal{B}$:

$$\begin{aligned} x_1 &= a_{11}\lambda_1 + a_{12}\lambda_2 + \cdots + a_{1r}\lambda_r \\ x_2 &= a_{21}\lambda_1 + a_{22}\lambda_2 + \cdots + a_{2r}\lambda_r \\ &\vdots \\ x_n &= a_{n1}\lambda_1 + a_{n2}\lambda_2 + \cdots + a_{nr}\lambda_r \end{aligned}$$

en los parámetros $\lambda_1, \lambda_2, \ldots, \lambda_r$, pues cada vez que cada uno de estos parámetros tiene un valor en K, obtendremos un vector de U. Las columnas que acompañan a cada parámetro son las coordenadas en

$\mathcal{B}$ de los vectores de la base $\mathcal{B}_U$ de U. Además, podemos contemplar estas igualdades como el conjunto de soluciones de un cierto sistema de ecuaciones lineales con n incógnitas sobre el cuerpo K. De hecho, como $0 \in U$, la solución trivial será una de las soluciones de dicho sistema que, por tanto, tendrá que ser homogéneo. Entonces, a cualquier sistema homogéneo cuya solución general sea la anterior, lo llamaremos unas ecuaciones implícitas (ó cartesianas) de U respecto de la base $\mathcal{B}$.

Ejemplo 5.
Supongamos el sistema homogéneo $x_1 + x_2 - x_3 = 0$. Sus soluciones nos van a proporcionar un subespacio vectorial U de $\mathbb{R}^3$, pues podemos escribir $x_1 = -x_2 + x_3$ y, entonces tomando $x_2 = \lambda$ como un parámetro y $x_3 = \mu$ como otro parámetro, tendremos que

$$\begin{aligned} x_1 &= -1 \cdot \lambda + 1 \cdot \mu \\ x_2 &= 1 \cdot \lambda + 0 \cdot \mu \\ x_3 &= 0 \cdot \lambda + 1 \cdot \mu \end{aligned}$$

Así, cada vector de dicho subespacio U será una combinación lineal de los vectores que, en la base usual, tienen como coordenadas los coeficientes de cada parámetro. Así, $\{(-1, 1, 0), (1, 0, 1)\}$ nos daría un sistema de generadores de dicho subespacio y, como claramente son linealmente independientes, forman de hecho una base de U. Esto es, $U = \mathcal{L}(\{(-1, 1, 0), (1, 0, 1)\})$.

Existen distintos métodos para calcular unas ecuaciones cartesianas de un subespacio, entre ellos el de eliminación de parámetros. Destacamos el siguiente:
Si $x = (x_1, x_2, \ldots, x_n)_{\mathcal{B}}$ es un vector de U, entonces x es combinación lineal de $\{u_1, \ldots, u_r\}$ y al considerar la matriz

$$B = \left(\begin{array}{cccc|c} a_{11} & a_{12} & \ldots & a_{1r} & x_1 \\ a_{21} & a_{22} & \ldots & a_{2r} & x_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{n1} & a_{n2} & \ldots & a_{nr} & x_n \end{array}\right)$$

como el último vector (columna) ha de ser combinación lineal de los r anteriores, el rango de esta matriz coincidirá con el de la submatriz obtenida al quitarle la última columna. Pero dicha submatriz tiene rango r. por tanto, todos los menores de orden $r + 1$ de la matriz B se han de anular.

Ejemplo 6.
Supongamos que en $\mathbb{R}^4$ consideramos el subespacio dado por $U = \mathcal{L}(\{(1, 1, 0, -1), (2, 0, 1, 0), (0, -2, 1, 2)\})$. Considerando $\mathcal{B}_U$ en $\mathbb{R}^4$ vamos a calcular unas ecuaciones implícitas para U. Como los 3 vectores que usamos para determinar U forman un sistema de generadores de U, vamos a obtener una base de U. Consideremos entonces la matriz

$$\begin{pmatrix} 1 & 2 & 0 \\ 1 & 0 & -2 \\ 0 & 1 & 1 \\ -1 & 0 & 2 \end{pmatrix}$$

La submatriz $\begin{pmatrix} 1 & 2 \\ 1 & 0 \end{pmatrix}$ tiene determinante no nulo, así que los dos primeros vectores son linealmente independientes. Por otro lado, como

$$\begin{vmatrix} 1 & 2 & 0 \\ 1 & 0 & -2 \\ 0 & 1 & 1 \end{vmatrix} = 2 - 2 = 0 \quad \text{y} \quad \begin{vmatrix} 1 & 2 & 0 \\ 1 & 0 & -2 \\ -1 & 0 & 2 \end{vmatrix} = -2 \begin{vmatrix} 1 & -2 \\ -1 & 2 \end{vmatrix} = 0,$$

el tercer vector depende linealmente de los dos primeros, con lo que $\mathcal{B}_U = \{(1,1,0,-1),(2,0,1,0)\}$. Así unas ecuaciones paramétricas de U serán

$$\begin{cases} x_1 = \lambda + & 2\mu \\ x_2 = \lambda & \\ x_3 = & \mu \\ x_4 = -\lambda & \end{cases}$$

Y considerando la matriz

$$\begin{pmatrix} 1 & 2 & x_1 \\ 1 & 0 & x_2 \\ 0 & 1 & x_3 \\ -1 & 0 & x_4 \end{pmatrix}$$

para que la 3ª columna sea linealmente dependiente de las dos primeras habrá de darse que

$$0 = \begin{vmatrix} 1 & 2 & x_1 \\ 1 & 0 & x_2 \\ 0 & 1 & x_3 \end{vmatrix} = x_1 - x_2 - 2x_3,$$

y

$$0 = \begin{vmatrix} 1 & 2 & x_1 \\ 1 & 0 & x_2 \\ -1 & 0 & x_4 \end{vmatrix} = -2 \begin{vmatrix} 1 & x_2 \\ -1 & x_4 \end{vmatrix} = -2(x_4 + x_2).$$

Así,

$$\begin{cases} x_1 & -x_2 & -2x_3 & & = 0 \\ & x_2 & & +x_4 & = 0 \end{cases}$$

son unas ecuaciones implícitas, respecto de $\mathcal{B}_U$, de U.

10. Ecuaciones cartesianas y dimensión de un subespacio.

Sea $V(K)$ un espacio vectorial con $\dim_K(V) = n$ y U un subespacio vectorial suyo con $\dim_K(U) = r$. Sabemos entonces que en unas ecuaciones paramétricas de U aparecerán r parámetros. Dado un sistema homogéneo con n incógnitas, para que la solución dependa de r parámetros, es necesario que la matriz de coeficientes tenga rango $n - r$, por lo que habrá, al menos, $n - r$ ecuaciones. Pero si hay más de $n - r$,

las que exceden se podrán hacer 0 mediante transformaciones elementales en el sistema. Así tenemos que el número de ecuaciones cartesianas es igual a $\dim_K(V) - \dim_K(U)$.
Esta relación será útil cuando queramos calcular las ecuaciones cartesianas de un subespacio. Por ejemplo, si en $\mathbb{R}^3$ consideramos el subespacio $U = \mathcal{L}(\{(1,1,0),(1,2,0)\})$, como ambos vectores son linealmente independientes ya que $\begin{vmatrix} 1 & 1 \\ 1 & 2 \end{vmatrix} = 1 \neq 0$, sabemos que U tiene dimensión 2, con lo cual para unas ecuaciones cartesianas de U solo habremos de buscar una ecuación. Como las ecuaciones paramétricas de U (en la base usual de $\mathbb{R}^3$) vienen dadas por

$$\begin{cases} x_1 = \lambda + & \mu \\ x_2 = \lambda + & 2\mu \\ x_3 = 0 \end{cases}$$

Esta última ecuación ya nos proporciona unas ecuaciones cartesianas de U.
Si en $\mathbb{R}^4$ consideramos ahora $U = \mathcal{L}(\{(1,2,1,0),(0,1,-1,1)\})$, tenemos de nuevo que $\dim_{\mathbb{R}}(U) = 2$, luego unas ecuaciones cartesianas comprenderán 2 ecuaciones. Las ecuaciones paramétricas de U serán (en $\mathcal{B}_U$ de $\mathbb{R}^4$)

$$\begin{cases} x_1 = \lambda \\ x_2 = 2\lambda + & \mu \\ x_3 = \lambda - & \mu \\ x_4 = & \mu \end{cases}$$

Si sustituimos los valores de x_1 y x_4 en la 2ª y 3ª ecuaciones tendremos $x_2 = 2x_1 + x_4, x_3 = x_1 - x_4$. Luego en $\mathcal{B}_U$ unas ecuaciones cartesianas de U serán

$$\begin{cases} 2x_1 & -x_2 & & +x_4 & = 0 \\ x_1 & & -x_3 & -x_4 & = 0 \end{cases}$$

11. Intersección de subespacios.

Sean U y W subespacios vectoriales del espacio vectorial $V(K)$. Como subconjuntos de V que son, podemos considerar su intersección $U \cap W$. La pregunta es si $U \cap W$ (que nunca será $\emptyset$, pues $0 \in U \cap W$) es otro subespacio vectorial de V. La respuesta es afirmativa, dado que $\forall\, a, b \in K$, $\forall\, x, y \in U \cap W$, tenemos que $x, y \in U$ y, como U es subespacio vectorial entonces $ax + by \in U$. Pero también $x, y \in W$ y, por la misma razón, $ax + by \in U$. Pero también $x, y \in W$ y, por la misma razón, $ax + by \in W$. Luego, efectivamente, $ax + by \in U \cap W$, que es un subespacio vectorial de V.
En general, dada una familia cualquiera $\{U_i \,/\, i \in I\}$ de subespacios vectoriales de V, su intersección $\bigcap_{i \in I} U_i$ es otro subespacio vectorial de V.
Si tenemos dos subespacios vectoriales U, W de V y queremos calcular su intersección, lo más fácil es considerar unas ecuaciones cartesianas de cada uno de ellos. Estas ecuaciones nos dan las condiciones que ha de verificar un vector de V para estar en U ó en W. Si queremos que dicho vector esté a la vez en

U y en W tendrá que verificar los dos conjuntos de ecuaciones cartesianas. Pero puede que al unirlas, algunas de ellas las obtengamos de las demás por transformaciones elementales, pudiéndolas eliminar del conjunto. Así tendríamos una ecuaciones cartesianas de $U \cap W$.

Ejemplo 7.
Consideremos en $\mathbb{R}^3$ los subespacios vectoriales

$$U = \{(x_1, x_2, x_3) \,/\, x_1 - x_2 - x_3 = 0\}, \quad W = \mathcal{L}(\{(1, -1, -1), (1, 1, 0), (2, 0, -1)\}).$$

Veamos cómo calcular $U \cap W$:
$x_1 - x_2 - x_3 = 0$ nos da la única ecuación cartesiana que tiene U. Calculemos ahora las de W. En primer lugar vemos si los vectores que generan W son una base de W:
Como $\begin{vmatrix} 1 & 1 & 2 \\ -1 & 1 & 0 \\ -1 & 0 & -1 \end{vmatrix} = -1 + 2 - 1 = 0$, pero $\begin{vmatrix} 1 & 2 \\ 1 & 0 \end{vmatrix} \neq 0$, sabemos que una base de W viene dada por los vectores $\{(1, 1, 0), (2, 0, -1)\}$. Como tiene dimensión 2, buscaremos una única ecuación cartesiana que será

$$\begin{vmatrix} 1 & 2 & x_1 \\ 1 & 0 & x_2 \\ 0 & -1 & x_3 \end{vmatrix} = 0,$$

es decir, $-x_1 + x_2 - 2x_3 = 0$. Luego, como esta ecuación no podemos obtenerla de la ecuación que define a U, concluimos que unas ecuaciones cartesianas de $U \cap W$ serían

$$\begin{cases} x_1 - x_2 - x_3 = 0 \\ -x_1 + x_2 - 2x_3 = 0 \end{cases}$$

Nota 3. En este caso $\mathbb{R}^3$ tiene la base usual, que es la más fácil a la hora de considerar los vectores de $\mathbb{R}^3$ por sus coordenadas. Cuando calculemos unas ecuaciones cartesianas de un subespacio de $\mathbb{R}^3$ respecto de dicha base no haremos mención de la misma. Lo mismo ocurrirá cuando tratemos con espacios vectoriales para los que dispongamos de bases "canónicas" o "naturales" como $\mathcal{M}_{m\times n}(K)$ o $\mathcal{P}^n(K)$.

12. Suma de subespacios.

Tras estudiar la intersección de subespacios, veamos que la unión de dos subespacios vectoriales U y W de un espacio vectorial $V(K)$ no es, en general, un subespacio vectorial. Para ello, consideramos en $\mathbb{R}^2$ los subespacios $U = \{(x, 0) \,/\, x \in \mathbb{R}\}$ y $W = \{(0, y) \,/\, y \in \mathbb{R}\}$. Tenemos que $(1, 0) \in U \subset U \cup W$ y que $(0, 1) \in W \subset U \cup W$. Luego ambos vectores estarían en $U \cup W$ y, si este fuera un subespacio vectorial de $\mathbb{R}^2$, entonces $(1, 0) + (0, 1) = (1, 1)$ tendría que pertenecer a $U \cup W$. Pero $(1, 1) \notin U$ y $(1, 1) \notin W$ con lo que $(1, 1) \notin U \cup W$, y así tendríamos una contradicción.
Entonces, como $U \cup W$ sigue siendo un subconjunto de V, llamaremos suma de U y W y lo notaremos $U + W$ al menor subespacio vectorial de V que contenga a $U \cup W$. Es decir, $U + W = \mathcal{L}(U \cup W)$. El por

qué del nombre de suma viene del siguiente hecho:

$$U + W = \{u + w \,/\, u \in U, w \in W\}.$$

Que se da esta igualdad es fácil de comprobar: en efecto, $\{u + w \,/\, u \in U, w \in W\}$ está contenido en $\mathcal{L}(U \cup W)$, puesto que $u, w \in U \cup W \subset \mathcal{L}(U \cup W)$ y por tanto $u + w \in \mathcal{L}(U \cup W)$. Por otro lado si $u_1 + w_1$, $u_2 + w_2$ son elementos del subconjunto y $a, b \in K$ entonces $a(u_1+w_1)+b(u_2+w_2) = (au_1+bu_2)+(aw_1+bw_2)$. Como U es un subespacio vectorial el primer sumando pertenece a U. Por la misma razón el segundo sumando pertenece a W y por tanto $\{u + w \,/\, u \in U, w \in W\}$ es un subespacio vectorial de V que, trivialmente, contiene a U y a W, por lo tanto a $U \cup W$ por lo cual contiene al subespacio que genera $U \cup W$, es decir, a $U + W$.

La forma más sencilla de conocer la suma de dos subespacios vectoriales es reunir bases de ambos, lo que nos da un sistema de generadores de la tal suma. Quedándonos con los vectores que sean linealmente independientes en ese sistema de generadores obtendremos una base de la suma:

Si $\{u_1, u_2, \ldots, u_r\}$ es un sistema de generadores de U y $\{w_1, w_2, \ldots, w_s\}$ es un sistema de generadores de W, sea $v = u + w \in U + W$. Entonces $u = a_1u_1 + \cdots + a_ru_r$ para ciertos escalares $a_1, \ldots, a_r$ y $w = b_1w_1+\cdots+b_sw_s$ para ciertos escalares $b_1, \ldots, b_s$. Entonces $u+w = a_1u_1+\cdots+a_ru_r+b_1w_1+\cdots+b_sw_s$. Como esto ocurre $\forall\, v \in U + W$, $\{u_1, \ldots, u_r, w_1, \ldots, w_s\}$ es un sistema de generadores de $U + W$.

La definición que hemos dado de suma de dos subespacios vectoriales se puede generalizar a una familia arbitraria de subespacios vectoriales $\{U_i \,/\, i \in I\}$ mediante

$$\sum_{i \in I} U_i = \mathcal{L}\left(\bigcup_{i \in I} U_i\right).$$

Ejemplo 8.

Consideremos los mismos subespacios del ejemplo 7, por un lado U dado por la ecuación cartesiana $x_1 - x_2 - x_3 = 0$ y por el otro $W = \mathcal{L}(\{(1, -1, -1), (1, 1, 0), (2, 0, -1)\})$ y calculamos $U + W$. En tal ejemplo calculamos una base de W, $\{(1, 1, 0), (2, 0, -1)\}$. Calculemos ahora una de U: como solo tenemos una ecuación cartesiana, sabemos que la dimensión de U es 2 (ya que es un subespacio de $\mathbb{R}^3$) y tendremos, por tanto, que encontrar dos vectores de U que sean linealmente independientes. Es fácil comprobar que $(1, 1, 0)$ y $(1, 0, 1)$ verifican esas condiciones y nos dan, por tanto, una base de U. Así, $\{(1, 1, 0), (2, 0, -1), (1, 1, 0), (1, 0, 1)\}$ formarían un sistema de generadores de $U + W$. Para encontrar una base de $U + W$, observamos que el vector $(1, 1, 0)$ se repite. Entonces eliminamos uno de ellos y tenemos que $\{(1, 1, 0), (2, 0, -1), (1, 0, 1)\}$ seguiría siendo un sistema de generadores de $U + W$. Para ver si forman una base calculamos

$$\begin{vmatrix} 1 & 2 & 1 \\ 1 & 0 & 0 \\ 0 & -1 & 1 \end{vmatrix} = -\begin{vmatrix} 2 & 1 \\ -1 & 1 \end{vmatrix} = -3 \neq 0,$$

con lo que afirmamos que, efectivamente, forman una base de $U + W$. Eso nos dice que $\dim_{\mathbb{R}}(U + W) = 3$, pero un subespacio de dimensión 3 en $\mathbb{R}^3$ ha de ser todo $\mathbb{R}^3$. Por consiguiente, $U + W = \mathbb{R}^3$.

13. Suma directa de subespacios.

Hemos visto que cualquier vector de una suma de subespacios es una suma de un vector de cada uno de dichos subespacios. Pero podría ocurrir que los vectores de cada subespacio no fueran únicos a la hora de obtener mediante su suma un determinado vector del subespacio suma.

Cuando para cada vector del espacio suma, la forma de expresarlo como suma de los vectores del subespacio es única, diremos que tenemos una suma directa de subespacios.

Dada una familia finita de subespacios $U_1, U_2, \ldots, U_m$ del espacio vectorial $V(K)$, diremos que es independiente si

$$\forall\, i \in \{1, \ldots, m\} \quad U_i \cap \left(\sum_{j=1, j\neq i}^{m} U_j \right) = \{0\}.$$

En el caso particular en que $m = 2$, la condición se reduce a tener $U_1 \cap U_2 = \{0\}$.

Veamos que la suma de subespacios $U_1 + U_2 + \cdots + U_m$ es directa si y solo si la familia $U_1, U_2, \ldots, U_m$ es independiente.

Demostración. $\Longrightarrow$) En primer lugar, suponemos que la suma es directa. Por tanto, cada vector de esa suma tendrá una expresión única como suma de vectores de cada uno de los U_i. Supongamos entonces, por reducción al absurdo, que

$$U_i \cap \left(\sum_{j\neq i} U_j \right) \neq \{0\}.$$

Así, $\exists\, v \neq 0$ tal que $v \in U_i \cup \left(\sum_{j\neq i} U_j \right)$. Como $v \in U_i$, $v = 0 + \cdots + \overset{i)}{v} + \cdots + 0$. Pero también $v \in \sum_{j\neq i} U_j$, y, por tanto, $v = u_1 + \cdots + u_{i-1} + 0 + u_{i+1} + \cdots + u_m$, con $u_j \in U_j$, $j \neq i$. Esto nos llevaría a que los dos elementos en la misma posición en la suma han de ser iguales y, en particular, $v = 0$, lo que nos lleva a una contradicción.

$\Longleftarrow$) Recíprocamente, supongamos que la familia es independiente. Para ver que la suma es directa, supongamos que un vector v de esa suma verifica $v = u_1 + u_2 + \cdots + u_m = v_1 + v_2 + \cdots + v_m$, donde $u_i, v_i \in U_i$, $i \in \{1, \ldots, m\}$. Entonces tendríamos

$$0 = (u_1 - v_1) + (u_2 - v_2) + \cdots + (u_m - v_m).$$

Esto es, $v_1 - u_1 = (u_2 - v_2) + \cdots + (u_m - v_m)$. Pero entonces $v_1 - u_1 \in U_1 \cap \left(\sum_{j=2}^{m} U_m \right)$. Entonces por hipótesis, $u_1 = v_1$. Este razonamiento lo podemos hacer con cada subíndice y, por tanto, la suma es directa. $\square$

A partir de ahora denotaremos una suma directa de subespacios como $U_1 \oplus U_2 \oplus \cdots \oplus U_m$.

Definición 3 (Subespacio complementario).
Sea U un subespacio vectorial de $V(K)$, llamaremos subespacio complementario (o suplementario) de U a cualquier subespacio W de V que verifique que $V = U \oplus W$.

Si el espacio vectorial es de dimensión finita, para calcular un complementario del subespacio U, tomamos una base suya, $\{u_1, \ldots, u_r\}$ y la ampliamos hasta tener una base de V, $\{u_1, \ldots, u_r, w_{r+1}, \ldots, w_n\}$. Entonces $W = \mathcal{L}(\{w_{r+1}, \ldots, w_n\})$ es un complementario de U en V: como al unir las bases de U y W obtenemos una base de V, $U + W = V$. Pero es inmediato ver que $U \cap W = \{0\}$.

Ejemplo 9.
Consideremos en $\mathcal{P}^3(\mathbb{R})$ el subespacio vectorial $U = \{p(x) \in \mathcal{P}^3(\mathbb{R}) \,/\, p(x) = p(-x)\}$.
En primer lugar vamos a calcular una base de U. Sabemos que $\{1, x, x^2, x^3\}$ es una base de $\mathcal{P}^3(\mathbb{R})$. Así un polinomio arbitrario de $\mathcal{P}^3(\mathbb{R})$ será de la forma $a + bx + cx^2 + dx^3$ (en la base elegida (a, b, c, d) serán las coordenadas de dicho polinomio). Para que dicho polinomio esté en U tendremos que tener que $a + bx + cx^2 + dx^3 = a - bx + cx^2 - dx^3$. Luego $2bx + 2dx^3 = 0$. Esto nos lleva a que $b = 0$, $d = 0$. Estas son, por tanto, unas ecuaciones paramétricas de U y hemos visto que un polinomio arbitrario de U será de la forma $a + cx^2$. Así, una base de U es $\{1, x^2\}$. Para calcular un complementario de U en $\mathcal{P}^3(\mathbb{R})$ bastaría añadir dos vectores a la base de U hasta tener una base de $\mathcal{P}^3(\mathbb{R})$. Si tomamos, por ejemplo, $\{x, x^3\}$, entonces $\mathcal{L}(\{x, x^3\})$ es un complementario de U.

Nota 4. Por otra parte, si en lugar de tomar x y x^3, tomamos otros dos polinomios arbitrarios que sean linealmente independientes con $\{1, x^2\}$, obtendremos un subespacio complementario de U distinto del anterior. No podemos hablar entonces <u>del</u> complementario de un subespacio vectorial, sino de <u>un</u> complementario.

14. Fórmula de las dimensiones.

En esta sección nos centramos en demostrar el siguiente resultado:

Proposición 5 (Fórmula de las dimensiones)**.**
Si U y W son subespacios vectoriales de un espacio vectorial $V(K)$ de dimensión finita, se verifica la siguiente igualdad:

$$\dim_K(U) + \dim_K(W) = \dim_K(U + W) + \dim_K(U \cap W).$$

Demostración. Supongamos que $\dim_K(U) = r$, $\dim_K(W) = S$ y $\dim_K(U \cap W) = m$. Tendremos que ver que $\dim_K(U + W) = r + s - m$.
Sea entonces $\{v_1, \ldots, v_m\}$ una base de $U \cap W$. Como $U \cap W$ es un subespacio vectorial de U, podremos añadir $r - m$ vectores $u_{m+1}, \ldots, u_r$ de manera que $\{v_1, \ldots, v_m, u_{m+1}, \ldots, u_r\}$ nos dé una base de U. Pero $U \cap W$ también es un subespacio vectorial de W. Podremos encontrar entonces $s - m$ vectores $w_{m+1}, \ldots, w_s$ tales que $\{v_1, \ldots, v_m, w_{m+1}, \ldots, w_s\}$ forman una base de W. Si vemos ahora que

$$\{v_1, \ldots, v_m, u_{m+1}, \ldots, u_r, w_{m+1}, \ldots, w_s\}$$

es una base de $U + W$, tendremos el resultado. Claramente ese conjunto es un sistema de generadores de $U + W$, pues lo obtenemos a partir de sendos sistemas de generadores de U y W.

Veamos que, además, son linealmente independientes. Para ello, si la siguiente combinación lineal se anula

$$0 = a_1v_1 + \cdots + a_mv_m + b_{m+1}u_{m+1} + \cdots + b_ru_r + c_{m+1}w_{m+1} + \cdots + c_sw_s \tag{14.1}$$

Entonces el vector

$$v = a_1v_1 + \cdots + a_mv_m + b_{m+1}u_{m+1} + \cdots + b_ru_r = -c_{m+1}w_{m+1} - \cdots - c_sw_s$$

La primera igualdad nos dice que $v \in U$ y la segunda que $v \in W$. Es decir, $v \in U \cap W$ y, como $\{v_1, \ldots, v_m\}$ es una base del subespacio

$$v = d_1v_1 + \cdots + d_mv_m = a_1v_1 + \cdots + a_mv_m + b_{m+1}u_{m+1} + \cdots + b_ru_r.$$

Ahora bien, la igualdad anterior nos da

$$(a_1 - d_1)v_1 + \cdots + (a_m - d_m)v_m + b_{m+1}u_{m+1} + \cdots + b_ru_r = 0.$$

Como los vectores que intervienen en dicha combinación lineal son linealmente independientes tendremos que $a_1 = d_1, \ldots, a_m = d_m, b_{m+1} = \cdots = b_r = 0$. Introduciendo esto en (14.1) tenemos que

$$0 = a_1v_1 + \cdots + a_mv_m + c_{m+1}w_{m+1} + \cdots + c_sw_s.$$

Pero los vectores que aparecen en esta combinación lineal forman una base de W. Por tanto, $a_1 = \cdots = a_m = c_m = \cdots = c_s = 0$ y tenemos lo que queríamos. $\square$

15. Espacio vectorial cociente.

Sea $V(K)$ un espacio vectorial y U un subespacio vectorial de V. Definimos en V la siguiente relación $\sim_U$: dados $v, w \in V$, diremos que $v \sim_U w$ si y solo si $v - w \in U$. Esta relación es de equivalencia ya que es reflexiva, pues $\forall\, v \in V$, $v - v = 0 \in U$ y así, $v \sim_U v$. También es simétrica, pues si $v, w \in V$ y $v \sim_U w$, tendremos que $v - w \in U$, pero como U es un subespacio vectorial, $-(v - w) = w - v \in U$ y, entonces, $w \sim_U v$. Finalmente, es transitiva, ya que si $v, w, z \in V$ y $v \sim_U w$, $w \sim_U z$, tendremos que $v - w \in U$ y que $w - z \in U$, pero al ser U un subespacio vectorial $(v - w) + (w - z) = v - z \in U$, luego $v \sim_U z$.

Sea $v \in V$. Vamos a notar su clase de equivalencia respecto de la relación de equivalencia anterior como $v + U$, pues todos los elementos relacionados con v se diferencian de él en un elemento de U. Como subconjunto de V, $v + U = \{v + u \,/\, u \in U\}$. Tendremos en particular que la clase del vector $0 \in V$ es U. Al conjunto cociente de V mediante esta clase de equivalencia lo notaremos $V/U = \{x + U \,/\, x \in V\}$. Lo importante es que podremos definir una suma de clases y un producto por escalares de forma que V/U va a ser también un espacio vectorial sobre K (el espacio vectorial cociente):
Dadas $v + U$, $w + U \in V/U$ definimos su suma de la siguiente forma: $(v + U) + (w + U) = (v + w) + U$. Como vimos en el Tema 0, para que esta suma esté bien definida hemos de ver que no depende de los representantes que elijamos en cada clase; es decir, si $v + U = v_1 + U$ y $w + U = w_1 + U$, tendremos que ver

que $(v+w)+U = (v_1+w_1)+U$. Pero si $v+U = v_1+U$, tendremos que $v-v_1 \in U$, y como $w+U = w_1+U$, $w - w_1 \in U$. Siendo U un subespacio vectorial, entonces $(v - v_1) + (w - w_1) = (v + w) - (v_1 + w_1) \in U$. Luego, efectivamente, $(v + w) + U = (v_1 + w_1) + U$.
Es muy fácil demostrar que $(V/U, +)$ es un grupo abeliano y lo dejamos como **ejercicio**.
Si ahora tenemos $a \in K$ y $v + U \in V/U$, definimos el producto $a(v + U) = (av) + U$. Como antes, hemos de ver que esta definición no depende del representante de la clase que elijamos: si $v + U = v_1 + U$, entonces $v - v_1 \in U$, pero al ser U un subespacio vectorial $a(v - v_1) = av - av_1 \in U$, $\forall\, a \in K$ y así, $(av) + U = (av_1) + U$.
Como antes, la comprobación de las cuatro propiedades que faltan para ver que V/U es un espacio vectorial es inmediata.

Veamos ahora que si $V(K)$ es un espacio vectorial con dimensión finita n y U es un subespacio vectorial suyo con dimensión r, entonces $\dim_K(V/U) = n-r$, ya que si $\{u_1, \ldots, u_r\}$ nos da una base de U, y podemos ampliarla a una base $\{u_1, \ldots, u_r, v_{r+1}, \ldots, v_n\}$ de V, entonces las clases $\{v_{r+1} + U, \ldots, v_n + U\}$ forman una base de V/U.
Para ello, si tomamos un vector arbitrario $x \in V$, lo podremos escribir como $x = a_1u_1 + \cdots + a_ru_r + b_{r+1}v_{r+1} + \cdots + b_nv_n$, para los escalares $a_1, \ldots, a_r, b_{r+1}, \ldots, b_n$. Si ahora tomamos clases de equivalencia, tenemos $x+U = a_1(u_1 +U) + \cdots + a_r(u_r +U) + b_{r+1}(v_{r+1} +U) + \cdots + b_n(v_n +U)$, donde hemos aplicado las definiciones de suma de clases de equivalencia y producto por escalares. Ahora bien, $\forall\, i \in \{1, \ldots, r\}$, $u_i \in U$, luego $u_i + U = 0 + U$. Por tanto, la expresión anterior se reduce a que

$$x + U = b_{r+1}(v_{r+1} + U) + \cdots + b_n(v_n + U),$$

que nos demuestra que $\{v_{r+1} + U, \ldots, v_n + U\}$ es un sistema de generadores de V/U.
Veamos que, además, es linealmente independiente: si tenemos escalares $c_{r+1}, \ldots, c_n$ tales que $c_{r+1}(v_{r+1}+U) + \cdots + c_n(v_n + U) = 0 + U$, entonces, $(c_{r+1}v_{r+1} + \cdots + c_nv_n) + U = 0 + U$. Esto nos indica que $c_{r+1}v_{r+1} + \cdots + c_nv_n \in U$. Como $\{u_1, \ldots, u_r\}$ es una base de U, tendremos que $c_{r+1}v_{r+1} + \cdots + c_nv_n = d_1u_1 + \cdots + d_ru_r$ para ciertos escalares $d_1, \ldots, d_r$. Pero lo anterior equivale a tener

$$d_1u_1 + \cdots + d_ru_r - c_{r+1}v_{r+1} - \cdots - c_nv_n = 0.$$

Como los vectores que intervienen en la combinación lineal nos dan una base de V, han de ser linealmente independientes, luego todos los escalares que aparecen se han de anular. En particular, $c_{r+1} = \cdots = c_n$, lo que nos demuestra que, efectivamente, $\{v_{r+1} + U, \ldots, v_n + U\}$ son linealmente independientes. Por tanto tenemos que es una base de V/U y por tanto su dimensión es $n - r$.

Ejemplo 10.
Consideremos el subespacio vectorial U de $\mathbb{R}^3$ que, respecto de la base usual, tiene como ecuaciones cartesianas

$$\begin{cases} x_1 & +x_2 & +2x_3 & = 0 \\ x_1 & & -x_3 & = 0 \end{cases}$$

Sabemos, pues, que $\dim_{\mathbb{R}}(U) = 1$. Busquemos una base de U. La segunda ecuación nos dice que $x_1 = x_3$. Si introducimos esto en la primera ecuación obtenemos que $x_2 = -3x_3$. Tomando x_3 como parámetro y dándole el valor 1, por ejemplo, obtenemos que $\{(1, -3, 1)\}$ es una base de U. Busquemos entonces un complementario de U en $\mathbb{R}^3$. Para ello basta con elegir dos vectores linealmente independientes con él. Así, si $v_2 = (1, 0, 0)$ y $v_3 = (0, 0, 1)$, como

$$\begin{vmatrix} 1 & 1 & 0 \\ -3 & 0 & 0 \\ 1 & 0 & 1 \end{vmatrix} = \begin{vmatrix} 1 & 1 \\ -3 & 0 \end{vmatrix} = 3 \neq 0,$$

entonces $W = \mathcal{L}(\{(1, 0, 0), (0, 0, 1)\})$ es un complementario de U en $\mathbb{R}^3$ y, por tanto, una base de $\mathbb{R}^3/U$ será la dada por $\{(1, 0, 0) + U, (0, 0, 1) + U\}$.

Geometría I: Tema 3
Aplicaciones lineales

Juan de Dios Pérez

Índice

1. Introducción. **2**

2. Núcleo e imagen de una aplicación lineal. **3**

3. Monomorfismos, epimorfismos e isomorfismos. **4**

4. Operaciones con aplicaciones lineales. **5**

5. Aplicaciones lineales y matrices. **7**

5.1. Matriz asociada a una aplicación lineal. 7

5.2. Matriz asociada, núcleo e imagen. 9

5.3. Matriz asociada y cambio de base. 10

5.4. Matriz asociada y operaciones con aplicaciones lineales. 12

1. Introducción.

Sean V y V' dos espacios vectoriales sobre el mismo cuerpo K y $f : V \longrightarrow V'$ una aplicación. Diremos que f es una aplicación lineal (u homomorfismo de espacios vectoriales) si f verifica las siguientes condiciones

1. $\forall\, u, v \in V,\ f(u+v) = f(u) + f(v)$.
2. $\forall\, a \in K\ \forall\, v \in V,\ f(av) = af(v)$.

Hemos de hacer notar que, de nuevo, abusamos de la notación, pues las sumas en V y en V' no tienen por qué coincidir, aunque para simplificar las representemos por el mismo símbolo. Lo mismo ocurre con el producto por escalares en ambos espacios.

Ejemplo 1.

1. Dado $V(K)$, la aplicación $1_V : V \longrightarrow V$ es claramente lineal, pues $\forall\, u, v \in V$ y $\forall\, a \in K$, $1_V(u+v) = u+v = 1_V(u) + 1_V(v)$ y $1_V(au) = au = a1_V(u)$.
   Análogamente, si U es un subespacio vectorial de V e $i_U : U \longrightarrow V$ es la aplicación de inclusión, i_U es lineal: $\forall\, u_1, u_2 \in U$, $\forall\, a \in K$, $i_U(u_1+u_2) = u_1+u_2 = i_U(u_1)+i_U(u_2)$ e $i_U(au_1) = au_1 = ai_U(u_1)$.

2. Sean V y V' espacios vectoriales arbitrarios y la aplicación constante 0; $0 : V \longrightarrow V'$ tal que $0(v) = 0$, $\forall\, v \in V$. Es inmediato comprobar que 0 es lineal. Pero si $v' \in V'$ es $v' \neq 0$, la aplicación constante $c_{v'} : V \longrightarrow V'$ dada por $c_{v'}(u) = v'$, $\forall\, u \in V$ no es lineal. Si lo fuese, $\forall\, u, v \in V$ tendríamos $c_{v'}(u+v) = v' = c_{v'}(u) + c_{v'}(v) = v' + v' = 2v'$, lo que nos llevaría a que $v' = 0$ en contra de nuestra hipótesis.

3. Sea $f : \mathbb{R}^3 \longrightarrow \mathbb{R}^2$ la aplicación dada por $f(x_1, x_2, x_3) = (x_1 + x_2, x_2 + x_3)$. Veamos que esta aplicación es lineal: sean $a \in K, (x_1, x_2, x_3), (x_1', x_2', x_3') \in \mathbb{R}^3$. Entonces $f((x_1, x_2, x_3)+(x_1', x_2', x_3')) = f(x_1+x_1', x_2+x_2', x_3+x_3') \overset{(*)}{=} (x_1+x_1'+x_2+x_2', x_2+x_2'+x_3+x_3') = (x_1+x_2, x_2+x_3)+(x_1'+x_2', x_2'+x_3') = f(x_1, x_2, x_3) + f(x_1', x_2', x_3')$.
   Además, $f(a(x_1, x_2, x_3)) = f(ax_1, ax_2, ax_3) \overset{(*)}{=} (ax_1 + ax_2, ax_2 + ax_3) = a(x_1 + x_2, x_2 + x_3) = af(x_1, x_2, x_3)$, donde en $*$ hemos aplicado la definición de f.

4. La aplicación $f : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ dada por $f(x_1, x_2) = (x_1^2, x_2)$ no es lineal. Si lo fuese, $f(2,2) = (4,2) = 2f(1,1) = 2(1,1) = (2,2)$, lo que es imposible ya que $4 \neq 2$.

5. La aplicación $t : \mathcal{M}_{m\times n}(K) \longrightarrow \mathcal{M}_{n\times m}(K)$ tal que $t(A) = A^t$ es lineal por las propiedades que ya conocemos.

Como en el caso de los subespacios vectoriales, las dos condiciones de la definición de aplicación lineal se pueden resumir en una única condición:

Proposición 1.
Una aplicación $f : V \longrightarrow V'$ donde V y V' son espacios vectoriales sobre el mismo cuerpo K es lineal si y solo si $\forall\, a, b \in K$ y $\forall\, u, v \in V$ se tiene $f(au + bv) = af(u) + bf(v)$.

Demostración. Para demostrar la primera implicación, supongamos que f es lineal. En las condiciones que tenemos $f(au+bv) \overset{(*)}{=} f(au)+f(bv) \overset{(**)}{=} af(u)+bf(v)$, donde en $*$ aplicamos la primera condición para que f sea lineal y en $**$ la segunda.
Para demostrar el recíproco, partimos de que se cumple la nueva condición $f(au+bv) = af(u)+bf(v)$ y hemos de ver que se dan las dos condiciones que aparecen en la definición de aplicación lineal: así, si tomamos $u, v \in V$ y $a = b = 1 \in K$, tendríamos que $f(u+v) = f(1\cdot u+1\cdot v) = 1f(u)+1f(v) = f(u)+f(v)$ y se da la primera condición y tomando $v = u, b = 0$, se tiene que $f(au) = f(au+0\cdot u) = af(u)+0f(u) = af(u)$, obteniendo la segunda condición. □

De la definición de aplicación lineal y la caracterización que acabamos de ver, tenemos las siguientes propiedades genéricas para una aplicación lineal $f : V \longrightarrow V'$ (a partir de ahora no mencionaremos el cuerpo K, salvo que tengamos que hacer elecciones de escalares):

1. $f(0) = 0$, que se sigue de que $f(0) = f(0+0) = f(0)+f(0)$, con lo que, restando $f(0)$ en los dos miembros, obtenemos el resultado.

2. $f(-u) = -f(u)$, pues $f(-u) = f((-1)\cdot u) = (-1)f(u) = -f(u)$.

3. $\forall\, a_1, \ldots, a_n \in K$, $\forall\, u_1, \ldots, u_n \in V$, $f(a_1u_1+a_2u_2+\cdots+a_nu_n) = a_1f(u_1)+a_2f(u_2)+\cdots+a_nf(u_n)$.

   Demostración. Para demostrar esta propiedad, haremos una inducción sobre n: si $n = 2$, tenemos la caracterización que hemos visto antes. Supongamos que para $n-1 \geq 2$, se da que

   $$f(a_1u_1+a_2u_2+\cdots+a_{n-1}u_{n-1}) = a_1f(u_1)+a_2f(u_2)+\cdots+a_{n-1}f(u_{n-1}).$$

   Entonces

   $$f(a_1u_1+\cdots+a_{n-1}u_{n-1}+a_nu_n) = f((a_1u_1+\cdots+a_{n-1}u_{n-1})+(a_nu_n)) \overset{(*)}{=}$$

   $$\overset{(*)}{=} f(a_1u_1+\cdots+a_{n-1}u_{n-1})+f(a_nu_n) \overset{(**)}{=} a_1f(u_1)+\cdots+a_{n-1}f(u_{n-1})+a_nf(u_n),$$

   donde en $*$ hemos aplicado la primera propiedad de la definición de aplicación lineal y en $**$ la hipótesis de inducción y la segunda propiedad. □

Fijaos en que la primera propiedad de las anteriores da una condición necesaria para que una aplicación entre espacios vectoriales sea lineal: la imagen del 0 de V ha de ser el 0 de V'. Esta condición, sin embargo, no es suficiente: en el ejemplo número 4 anterior $f(0) = 0$, pero f no es lineal.

2. Núcleo e imagen de una aplicación lineal.

Definición 1.
Dada la aplicación lineal $f : V \longrightarrow V'$, definimos su <u>núcleo</u> mediante $Ker(f) = \{v \in V \,/\, f(v) = 0\}$ y su <u>imagen</u> como $Im(f) = \{f(v) \,/\, v \in V\}$.

Tenemos, entonces, que $\text{Ker}(f)$ es un subconjunto de V, mientras que $\text{Im}(f)$ lo es de V'. Lo importante es que ambos van a ser subespacios vectoriales de V y de V', respectivamente.
Para ver que $\text{Ker}(f)$ es un subespacio vectorial de V, sean $a, b \in K$ y $u, v \in \text{Ker}(f)$. Entonces $f(u) = f(v) = 0$ y $f(au + bv) = af(u) + bf(v) = a \cdot 0 + b \cdot 0 = 0$, y así, $au + bv \in \text{Ker}(f)$.
Sean ahora $a, b \in K$, $u', v' \in \text{Im}(f)$. Entonces $\exists\, u \in V$ tal que $f(u) = u'$ y $\exists\, v \in V$ tal que $f(v) = v'$. El vector $au + bv \in V$ verifica que $f(au + bv) = af(u) + bf(v) = au' + bv'$, y, por tanto, $au' + bv' \in \text{Im}(f)$.

Tenemos, además, que si $\{v_1, \ldots, v_m\}$ es un sistema de generadores de V, $\{f(v_1), \ldots, f(v_m)\}$ es un sistema de generadores de $\text{Im}(f)$. Para demostrar esto, sea $u' \in \text{Im}(f)$. Entonces $\exists\, u \in V$ tal que $f(u) = u'$. Pero como $\{u_1, \ldots, u_m\}$ es un sistema de generadores de V, podremos encontrar escalares $a_1, \ldots, a_m \in K$ tales que $u = a_1u_1 + \cdots + a_mu_m$. Así, $u' = f(u) = f(a_1u_1 + \cdots + a_mu_m) = a_1f(u_1) + \cdots + a_mf(u_m)$. Luego cualquier vector de $\text{Im}(f)$ lo podemos escribir como combinación lineal del conjunto $\{f(u_1), \ldots, f(u_m)\}$ que es, por tanto, un sistema de generadores de $\text{Im}(f)$.
Consideremos el ejemplo 3 anterior. Como sistema de generadores de $\mathbb{R}^3$ podemos considerar la base usual $\mathcal{B}_U = \{(1,0,0), (0,1,0), (0,0,1)\}$. Un sistema de generadores de $\text{Im}(f)$ vendrá dado por $\{f(1,0,0) = (1,0), f(0,1,0) = (1,1), f(0,0,1) = (0,1)\}$. Como $(1,1) = (1,0) + (0,1)$ y estos dos últimos vectores nos dan la base usual de $\mathbb{R}^2$, podemos concluir que, en este caso, $\text{Im}(f) = \mathbb{R}^2$.
Si queremos calcular $\text{Ker}(f) = \{(x_1, x_2, x_3) \in \mathbb{R}^3 \,/\, f(x_1, x_2, x_3) = (x_1 + x_2, x_2 + x_3) = (0,0)\}$. Este núcleo, vendrá dado entonces por las condiciones $x_1 + x_2 = 0, x_2 + x_3 = 0$, que serán unas ecuaciones cartesianas suyas. Como son dos, la dimensión de $\text{Ker}(f)$ será 1 y el vector $(1, -1, 1) \in \text{Ker}(f)$ pues verifica las dos ecuaciones. Todo esto nos dice que $\text{Ker}(f) = \mathcal{L}(\{(1, -1, 1)\})$.

3. Monomorfismos, epimorfismos e isomorfismos.

Sea $f : V \longrightarrow V'$ una aplicación lineal. Tenemos las siguientes definiciones:

- Si f es, además, inyectiva, diremos que f es un monomorfismo de espacios vectoriales.
- Si f es sobreyectiva, diremos que f es un epimorfismo de espacios vectoriales.
- Por último, si f es biyectiva (a la vez monomorfismo y epimorfismo), diremos que f es un isomorfismo de espacios vectoriales.

Veamos cómo $\text{Ker}(f)$ e $\text{Im}(f)$ nos permiten caracterizar estos tipos de aplicaciones lineales:

Proposición 2.
Dada una aplicación lineal $f : V \longrightarrow V'$, se verifica que

1. *f es un monomorfismo si y solo si $Ker(f) = \{0\}$.*
2. *f es un epimorfismo si y solo si $Im(f) = V'$.*

Demostración. Para demostrar la primera condición, supongamos, en primer lugar, que f es inyectiva y que $v \in \text{Ker}(f)$ es un vector suyo arbitrario. Entonces $f(v) = 0$, pero sabemos que $f(0) = 0$ y, por tanto,

$f(v) = f(0)$. Como f es inyectiva esto implica que $v = 0$ y, así, $\mathrm{Ker}(f) = \{0\}$.
Recíprocamente, supongamos que $\mathrm{Ker}(f) = \{0\}$ y que $u, v \in V$ son tales que $f(u) = f(v)$. Entonces $f(u - v) = f(u) - f(v) = 0$, con lo cual $u - v \in \mathrm{Ker}(f)$. Como se reduce a $\{0\}$, tendremos que $u = v$ y, por tanto, f es inyectiva.

La segunda condición es la definición de que f es sobreyectiva. □

Tenemos a continuación sendas caracterizaciones para las definiciones de monomorfismo y epimorfismo:

Proposición 3.
Dada una aplicación lineal $f : V \longrightarrow V'$, tenemos que:

1. *f es un monomorfismo si y solo si para cada conjunto linealmente independiente en V, $\{v_1, v_2, \ldots, v_n\}$, el conjunto $\{f(v_1), f(v_2), \ldots, f(v_n)\}$ es linealmente independiente en V'.*

   Demostración. Para verlo, supongamos, en primer lugar, que f es inyectiva. Sea $\{v_1, v_2, \ldots, v_n\}$ linealmente independiente en V y suponemos que $\{f(v_1), f(v_2), \ldots, f(v_n)\}$ admite una combinación lineal
   $$a_1 f(v_1) + \cdots + a_n f(v_n) = 0.$$
   Como f es lineal lo anterior nos da que $f(a_1v_1+\cdots+a_nv_n) = 0 = f(0)$. Entonces $a_1v_1+\cdots+a_nv_n = 0$ (f es inyectiva) y, como $\{v_1, \ldots, v_n\}$ es linealmente independiente, $a_1 = \cdots = a_n = 0$, con lo que $\{f(v_1), \ldots, f(v_n)\}$ es también linealmente independiente.
   Recíprocamente, supongamos que se verifica el recíproco. Queremos ver que f es inyectiva para lo que será suficiente con ver que $\mathrm{Ker}(f) = \{0\}$. Sea entonces $v \in \mathrm{Ker}(f)$. Entonces $\{f(v)\} = \{0\}$ es linealmente dependiente. Pero por lo que hemos supuesto, $\{v\}$ ha de ser también linealmente dependiente y eso solo se puede dar si $v = 0$. □

2. *f es un epimorfismo si y solo si para cada sistema de generadores de V, $\{u_1, \ldots, u_m\}$, el conjunto $\{f(u_1), \ldots, f(u_m)\}$ es un sistema de generadores de V'.*

   Demostración. Como f es sobreyectiva si y solo si $\mathrm{Im}(f) = V'$, el resultado es inmediato. □

4. Operaciones con aplicaciones lineales.

Dados los espacios vectoriales V y V' sobre el cuerpo K, notaremos $\mathrm{Hom}_K(V, V') = \{f : V \longrightarrow V' \,/\, f \text{ es una aplicación lineal}\}$. En este conjunto definimos las siguientes operaciones:

- Suma de aplicaciones lineales: dadas $f, g \in \mathrm{Hom}_K(V, V')$ definimos $f + g : V \longrightarrow V'$ mediante $(f+g)(v) = f(v)+g(v), \forall\, v \in V$. Veamos que $f+g \in \mathrm{Hom}_K(V, V')$. Para ello, sean $a, b \in K,\ v, w \in V$. Entonces $(f + g)(av + bw) = f(av + bw) + g(av + bw) = af(v) + bf(w) + ag(v) + bg(w) \overset{(*)}{=} af(v) + ag(v) + bf(w) + bg(w) \overset{(**)}{=} a(f(v) + g(v)) + b(f(w) + g(w)) = a(f + g)(v) + b(f + g)(w)$, donde en $*$ aplicamos la conmutatividad de la suma en V' y en $**$ la distributividad del producto por escalares con respecto a la suma.

- Producto por escalares: dados $a \in K$ y $f \in \mathrm{Hom}_K(V, V')$, definimos $af : V \longrightarrow V'$ mediante $(af)(v) = af(v), \forall\, v \in V$. Podemos ver, de nuevo, que $af \in \mathrm{Hom}_K(V, V')$: sean $b, c \in K$ y $v, w \in V$, entonces $(af)(bv + cw) = af(bv + cw) = a(bf(v) + cf(w)) = abf(v) + acf(w) = baf(v) + caf(w) = b(af)(v) + c(af)(w)$.

Con estas operaciones podemos afirmar que $\mathrm{Hom}_K(V, V')$ es otro espacio vectorial sobre K. La demostración es inmediata a partir de las propiedades de V y V' y se deja como **ejercicio**.

Por otro lado, si ahora tenemos tres espacios vectoriales sobre el mismo cuerpo K y $f : V \longrightarrow V'$, $g : V' \longrightarrow V''$ son aplicaciones lineales, veamos que $g \circ f : V \longrightarrow V''$ sigue siendo lineal: $\forall\, a, b \in K$, $\forall\, v, w \in V$, tenemos $(g \circ f)(av + bw) = g(f(av + bw)) \overset{(*)}{=} g(af(v) + bf(w)) \overset{(**)}{=} ag(f(v)) + bg(f(w)) = a(g \circ f)(v) + b(g \circ f)(w)$, donde en $*$ aplicamos que f es lineal y en $**$ que lo es g.

Si una aplicación lineal tiene iguales el dominio y el codominio, es decir, $f : V \longrightarrow V$, diremos que f es un endomorfismo de V. Al conjunto de los endomorfismos de un espacio vectorial $V(K)$ lo notaremos $\mathrm{End}_K(V)$. Sabemos entonces que $\mathrm{End}_K(V) = \mathrm{Hom}_K(V, V)$ es un espacio vectorial sobre K. Pero además, la composición nos da una ley de composición interna que se ve fácilmente que verifica:

1. Es asociativa: $(h \circ g) \circ f = h \circ (g \circ f)$ (pues esto se da en general).

2. Admite como elemento neutro a 1_V: se tiene $1_V \circ f = f \circ 1_V = f$.

3. Se verifican las siguientes propiedades distributivas con respecto a la suma:

   - $f \circ (g + h) = f \circ g + f \circ h$.
   - $(f + g) \circ h = f \circ h + g \circ h$.

4. Compatibilidad: $\forall\, a \in K$ se tiene $a(g \circ f) = (ag) \circ f = g \circ (af)$.

Si $f : V \longrightarrow V'$ es un isomorfismo, al ser f biyectiva admite una aplicación inversa $f^{-1} : V' \longrightarrow V$ que también es biyectiva. Veamos que f^{-1} también es lineal y, por tanto, es otro isomorfismo de espacios vectoriales.

Demostración. Sean $a, b \in K$ y $v', w' \in V'$. Llamemos $v = f^{-1}(v'), w = f^{-1}(w')$ (esto equivale a que $f(v) = v'$ y $f(w) = w'$). Entonces $f(av + bw) = af(v) + bf(w) = av' + bw'$, con lo que tenemos que $f^{-1}(av' + bw') = av + bw = af^{-1}(v') + bf^{-1}(w')$, que es lo que necesitamos. $\square$

Dado un espacio vectorial V, un automorfismo de V es un endomorfismo biyectivo de V. Si $\mathrm{Aut}_K(V)$ designa el conjunto de todos los automorfismos de V, todas las propiedades anteriores se resumen en que $(\mathrm{Aut}_K(V), \circ)$ es un grupo (en general, no abeliano) que llamaremos el grupo general de V.

5. Aplicaciones lineales y matrices.

5.1. Matriz asociada a una aplicación lineal.

Sea $f : V \longrightarrow V'$ una aplicación lineal. Si elegimos una base $\mathcal{B} = \{v_1, v_2, \ldots, v_n\}$ de V, f está totalmente determinada por los vectores $f(v_1), f(v_2), \ldots, f(v_n)$, ya que $\forall\, x \in V$ con coordenadas $x = (x_1, \ldots, x_n)_{\mathcal{B}}$, será $x = x_1v_1 + x_2v_2 + \cdots + x_nv_n$. Por lo tanto $f(x) = f(x_1v_1 + x_2v_2 + \cdots + x_nv_n) = x_1f(v_1) + x_2f(v_2) + \cdots + x_nf(v_n)$.

Ejemplo 2.
Sea $f : \mathbb{R}^3 \longrightarrow \mathbb{R}^4$ y $f(1,0,0) = (2,-1,1,3), f(0,1,0) = (1,0,1,0), f(0,0,1) = (-2,1,0,-1)$, entonces $f(1,2,3) = f((1,0,0) + 2(0,1,0) + 3(0,0,1)) = f(1,0,0) + 2f(0,1,0) + 3f(0,0,1) = (2,-1,1,3) + 2(1,0,1,0) + 3(-2,1,0,-1) = (-2,2,3,0)$.
En general $\forall\,(x_1, x_2, x_3) \in \mathbb{R}^3$, $f(x_1, x_2, x_3) = x_1(2,-1,1,3) + x_2(1,0,1,0) + x_3(-2,1,0,1) = (2x_1 + x_2 - 2x_3, -x_1 + x_3, x_1 + x_2, 3x_1 + x_3)$

Tomemos ahora otra base $\mathcal{B}' = \{u_1, u_2, \ldots, u_m\}$ de V' y consideremos las coordenadas de $f(v_1)$, $f(v_2)$, $\ldots$, $f(v_n)$ en $\mathcal{B}'$:

$$\begin{aligned} f(v_1) &= (a_{11}, \ldots, a_{m1})_{\mathcal{B}'} \\ f(v_2) &= (a_{12}, \ldots, a_{m2})_{\mathcal{B}'} \\ &\vdots \\ f(v_n) &= (a_{1n}, \ldots, a_{mn})_{\mathcal{B}'} \end{aligned}$$

Dado un vector arbitrario $x \in V$ con coordenadas $x = (x_1, \ldots, x_n)_{\mathcal{B}'}$ como hemos visto será $f(x) = x_1f(v_1) + x_2f(v_2) + \cdots + x_nf(v_n)$. Así, las coordenadas de $f(x)$ en $\mathcal{B}'$ serán:

$$f(x) = (a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n, a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n, \ldots, a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n)_{\mathcal{B}'}.$$

Pero si notamos las coordenadas de $f(x)$ en $\mathcal{B}'$ como $f(x) = (y_1, y_2, \ldots, y_m)_{\mathcal{B}'}$, la unicidad de tales coordenadas nos da

$$\begin{cases} y_1 &= a_{11}x_1 + \cdots + a_{1n}x_n \\ y_2 &= a_{21}x_1 + \cdots + a_{2n}x_n \\ &\vdots \\ y_m &= a_{m1}x_1 + \cdots + a_{mn}x_n \end{cases}$$

que podemos escribir matricialmente como

$$\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}.$$

Esta expresión se llama la <u>ecuación matricial</u> de la aplicación lineal f respecto de las bases $\mathcal{B}$ de V y $\mathcal{B}'$ de V'. La matriz que aparece

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}$$

se llama <u>matriz asociada</u> a f respecto de las bases $\mathcal{B}$ y $\mathcal{B}'$ y la notaremos como $\mathcal{M}(f;\mathcal{B}' \leftarrow \mathcal{B})$. Hay que fijarse en que su número de filas es la dimensión de V' mientras que su número de columnas es la dimensión de V. Si ahora llamamos

$$Y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{pmatrix} \quad \text{y} \quad X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix},$$

la ecuación matricial se reduce a $Y = AX$.

En esta expresión matricial, cada vez que conocemos las coordenadas X de un vector $x \in V$ en la base $\mathcal{B}$, obtenemos las coordenadas Y del vector $f(x)$ en la base $\mathcal{B}'$.

Nótese que cada columna de A nos da las coordenadas de las imágenes de los vectores $f(v_i)$ en la base $\mathcal{B}'$, donde $\mathcal{B} = \{v_1, v_2, \dots, v_n\}$.

Un caso particularmente interesante es cuando consideramos un $f \in \mathrm{End}_K(V)$, pues, en este caso, podemos considerar una única base $\mathcal{B} = \{v_1, v_2, \dots, v_n\}$ en V. Llamaremos entonces simplemente $\mathcal{M}(f;\mathcal{B}) = \mathcal{M}(f;\mathcal{B} \leftarrow \mathcal{B})$.

Ejemplo 3.

Siguiendo con el ejemplo anterior y considerando en $\mathbb{R}^3$ su base usual $\mathcal{B}_U^3$ y en $\mathbb{R}^4$ la correspondiente base usual $\mathcal{B}_U^4$, tendremos que

$$\mathcal{M}(f;\mathcal{B}_U^4 \leftarrow \mathcal{B}_U^3) = \begin{pmatrix} 2 & 1 & -2 \\ -1 & 0 & 1 \\ 1 & 1 & 0 \\ 3 & 0 & -1 \end{pmatrix}.$$

Con lo cual, la expresión matricial de f en esas bases será

$$\begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{pmatrix} = \begin{pmatrix} 2 & 1 & -2 \\ -1 & 0 & 1 \\ 1 & 1 & 0 \\ 3 & 0 & -1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}.$$

Ejemplo 4.

Sea ahora $D : \mathcal{P}^2(\mathbb{R}) \longrightarrow \mathcal{P}^2(\mathbb{R})$ la aplicación lineal derivada. Consideremos en $\mathcal{P}^2(\mathbb{R})$ la base estándar $\mathcal{B} = \{1, x, x^2\}$. Tenemos entonces que $D(1) = 0 = (0,0,0)_{\mathcal{B}}$, $D(x) = 1 = (1,0,0)_{\mathcal{B}}$ y $D(x^2) = 2x =$

$(0,2,0)_{\mathcal{B}}$. Por tanto,

$$\mathcal{M}(D;\mathcal{B}) = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{pmatrix}.$$

Toda matriz es la matriz asociada a una aplicación lineal respecto de ciertas bases: dados dos espacios vectoriales sobre el mismo cuerpo K, $V(K)$ y $V'(K)$ con $\dim_K(V) = n$ y $\dim_K(V') = m$ y elegidas bases $\mathcal{B}$ de V y $\mathcal{B}'$ de V', $\forall\, A \in \mathcal{M}_{m\times n}(K)$, $\exists_1\, f : V \longrightarrow V'$, aplicación lineal, tal que $\mathcal{M}(f;\mathcal{B}' \leftarrow \mathcal{B}) = A$. Basta definir f como la aplicación que lleva cada vector x de V con coordenadas X respecto de V en el vector $y = f(x)$ de V' cuyas coordenadas en $\mathcal{B}'$, Y verifican $Y = AX$.

5.2. Matriz asociada, núcleo e imagen.

La matriz asociada a una aplicación lineal respecto de bases de V y V', respectivamente, nos va a permitir calcular el núcleo y la imagen de dicha aplicación lineal. En concreto, sea $f : V \longrightarrow V'$ una aplicación lineal con $\dim_K(V) = n$ y $\dim_K(V') = m$, sean $\mathcal{B}$ una base de V, $\mathcal{B}'$ una base de V' y $A = \mathcal{M}(f;\mathcal{B}' \leftarrow \mathcal{B})$. Supongamos que $r = \mathrm{rg}(A)$.

Como las columnas de A son las coordenadas respecto de $\mathcal{B}'$ de las imágenes de los vectores de una base $\mathcal{B}$ de V, estas columnas forman un sistema de generadores de $\mathrm{Im}(f)$, y, por tanto, como el rango de A es r, $\dim_K(\mathrm{Im}(f)) = r$.

Basta, por tanto, con quedarnos con r columnas linealmente independientes para tener una base de $\mathrm{Im}(f)$.

Por otra parte, si $x = (x_1, x_2, \ldots, x_n)_{\mathcal{B}}$ es un vector de V, será $x \in \mathrm{Ker}(f)$ si y solo si $f(x) = 0$. Matricialmente hablando si $X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$ tendremos que $AX = 0$ y, por tanto, unas ecuaciones cartesianas de $\mathrm{Ker}(f)$ respecto de la base $\mathcal{B}$ serán

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = 0 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = 0 \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = 0 \end{cases}$$

Como el $\mathrm{rg}(A) = r$, ese sistema será equivalente al que obtenemos con las r filas de A que nos dan el rango y, en consecuencia, $\dim_K(\mathrm{Ker}(f)) = n - r$. Tenemos, pues, el siguiente resultado:

Teorema 1 (Fórmula de las dimensiones para aplicaciones lineales.)**.**
Sea $f : V \longrightarrow V'$ una aplicación lineal. Entonces

$$\dim_K(Ker(f)) + \dim_K(Im(f)) = \dim_K(V).$$

Para una tal aplicación lineal f, llamaremos rango de f (notado $\text{rg}(f)$) a la dimensión de $\text{Im}(f)$ y nulidad de f (notado $\text{nul}(f)$) a la dimensión de $\text{Ker}(f)$.
Para la aplicación lineal del ejemplo que venimos considerando, cuya matriz es

$$\mathcal{M}(f; \mathcal{B}_U^4 \leftarrow \mathcal{B}_U^3) = \begin{pmatrix} 2 & 1 & -2 \\ -1 & 0 & 1 \\ 1 & 1 & 0 \\ 3 & 0 & -1 \end{pmatrix},$$

puesto que $\begin{vmatrix} -1 & 0 & 1 \\ 1 & 1 & 0 \\ 3 & 0 & -1 \end{vmatrix} = \begin{vmatrix} -1 & 1 \\ 3 & -1 \end{vmatrix} = 1 - 3 = -2 \neq 0$, entonces $\text{rg}(A) = 3$, y los vectores $\{(2, -1, 1, 3),$ $(1, 0, 1, 0), (-2, 1, 0, -1)\}$ nos proporcionan una base de $\text{Im}(f)$. Por la fórmula de las dimensiones, $\dim_{\mathbb{R}}(\text{Ker}(f)) = 3 - 3 = 0$ y así, $\text{Ker}(f) = \{0\}$.
A partir de lo que conocemos, el siguiente resultado es inmediato.

Proposición 4.
Sea $f : V \longrightarrow V'$ una aplicación lineal con $\dim_K(V) = n$ y $\dim_K(V') = m$ y sea A la matriz $A = \mathcal{M}(f; \mathcal{B}' \leftarrow \mathcal{B})$, para una base $\mathcal{B}$ de V y una base $\mathcal{B}'$ de V'. Entonces:

1. *f es un monomorfismo si y solo si $rg(A) = n$.*
2. *f es un epimorfismo si y solo si $rg(A) = m$.*
3. *f es un isomorfismo si y solo si A es cuadrada y regular.*

Dos espacios vectoriales V y V' sobre el mismo cuerpo K se dirán isomorfos si podemos encontrar un isomorfismo $f : V \longrightarrow V'$ y lo notaremos $V \cong V'$. Tenemos entonces que $V \cong V'$ si y solo si $\dim_K(V) = \dim_K(V')$.

Demostración. Supongamos que $V \cong V'$, que $\dim_K(V) = n$ y que $\dim_K(V') = m$. Sea $f : V \longrightarrow V'$ un isomorfismo y $A \in \mathcal{M}_{m \times n}(K)$, $A = \mathcal{M}(f; \mathcal{B}' \leftarrow \mathcal{B})$ para ciertas bases $\mathcal{B}$ de V y $\mathcal{B}'$ de V'. Por lo que acabamos de ver, A es una matriz cuadrada y regular y, por tanto, $m = n$; es decir, $\dim_K(V) = \dim_K(V')$.
Recíprocamente, supongamos que $\dim_K(V) = \dim_K(V') = n$. Elegimos $\mathcal{B} = \{v_1, v_2, \ldots, v_n\}$ una base de V y $\mathcal{B}' = \{w_1, w_2, \ldots, w_n\}$ una base de $\mathcal{B}'$. Sea $f : V \longrightarrow V'$ la aplicación lineal tal que $\mathcal{M}(f; \mathcal{B}' \leftarrow \mathcal{B}) = I_n$ (así, $f(v_i) = w_i$, $\forall\, i \in \{1, \ldots, n\}$). Como claramente I_n es cuadrada y regular, f es un isomorfismo. $\square$

En particular esto nos dice que si V es un espacio vectorial real de dimensión n, $V \cong \mathbb{R}^n$. En general, si $V(K)$ es un espacio vectorial de dimensión n, $V \cong K^n$.

5.3. Matriz asociada y cambio de base.

Nos proponemos ver qué relación hay entre las matrices asociadas a una aplicación lineal respecto de dos bases distintas en el primer espacio y otras dos en el segundo espacio. Si A y C son las matrices

asociadas a f respecto de dos pares de bases, tendremos que $\mathrm{rg}(A) = \mathrm{rg}(f) = \mathrm{rg}(C)$ y, por tanto, A y C serán equivalentes. Veamos qué matrices nos dan esa equivalencia:
Sean $\mathcal{B}$ y $\overline{\mathcal{B}}$ dos bases de V tales que $\mathcal{M}_{\mathcal{B}\leftarrow\overline{\mathcal{B}}} = P$. Es decir, suponemos que $X = P\overline{X}$, donde X es la matriz columna de las coordenadas de un vector $x \in V$ en $\mathcal{B}$ y $\overline{X}$ la matriz columna de las coordenadas de x en $\overline{\mathcal{B}}$.
Asimismo, sean $\mathcal{B}'$ y $\overline{\mathcal{B}'}$ dos bases de V' tales que $\mathcal{M}_{\mathcal{B}'\leftarrow\mathcal{B}} = Q$. Por tanto, tendremos $Y = Q\overline{Y}$, donde ahora Y será matriz columna de las coordenadas de un vector $y \in V'$ en $\mathcal{B}'$ e $\overline{Y}$ la matriz columna de las coordenadas de y en $\overline{\mathcal{B}'}$. Así, si $A = \mathcal{M}(f; \mathcal{B}' \leftarrow \mathcal{B})$, en estas bases la ecuación matricial de f será

$$Y = AX.$$

Sea, por otra parte, $C = \mathcal{M}(f; \overline{\mathcal{B}'} \leftarrow \overline{\mathcal{B}})$ con lo que

$$\overline{Y} = C\overline{X}.$$

Entonces $\overline{Y} = Q^{-1}Y = Q^{-1}AX = Q^{-1}AP\overline{X}$ y de estas dos expresiones concluimos que

$$C = Q^{-1}AP,$$

que, traduciendo, nos da

$$\mathcal{M}(f; \overline{\mathcal{B}'} \leftarrow \overline{\mathcal{B}}) = \mathcal{M}_{\overline{\mathcal{B}}\leftarrow\mathcal{B}'}\mathcal{M}(f; \mathcal{B}' \leftarrow \mathcal{B})\mathcal{M}_{\mathcal{B}\leftarrow\overline{\mathcal{B}'}}.$$

En el caso particular de que f sea un endomorfismo de V, tomando $A = \mathcal{M}(f; \mathcal{B})$, $C = \mathcal{M}(f; \overline{\mathcal{B}'})$ tendríamos

$$\mathcal{M}(f; \overline{\mathcal{B}}) = \mathcal{M}_{\overline{\mathcal{B}}\leftarrow\mathcal{B}}\mathcal{M}(f; \mathcal{B})\mathcal{M}_{\mathcal{B}\leftarrow\overline{\mathcal{B}}}.$$

Luego $C = P^{-1}AP$.
Recordemos que P es una matriz regular. Entonces las matrices $A, C \in \mathcal{M}_n(K)$ diremos que son <u>semejantes</u> si $\exists\, P \in \mathcal{M}_n(K)$ regular tal que $C = P^{-1}AP$. Tenemos entonces:

1. Dos matrices son equivalentes si y solo si son matrices asociadas a la misma aplicación lineal respecto de distintas bases.

2. Dos matrices cuadradas son semejantes si y solo si son matrices asociadas al mismo endomorfismo respecto de distintas bases.

Solo veremos la primera propiedad, pues la segunda es completamente similar.

Demostración. Hemos visto que si tenemos dos matrices asociadas a la misma aplicación lineal respecto de distintas bases son equivalentes.
Veamos el recíproco: si $A, C \in \mathcal{M}_{m\times n}(K)$ son equivalentes, $\exists\, P \in \mathcal{M}_n(K)$ regular y $\exists\, Q \in \mathcal{M}_m(K)$ también regular tales que $C = Q^{-1}AP$. Sea $f : K^n \longrightarrow K^m$ cuya matriz asociada respecto de $\mathcal{B}_U^n$ de K^n y la de $\mathcal{B}_U^m$ de K^m sea A ($A = \mathcal{M}(f; \mathcal{B}_U^m \leftarrow \mathcal{B}_U^n)$). Sea $\overline{\mathcal{B}}$ la base de K^n tal que $P = \mathcal{M}_{\mathcal{B}_U^n\leftarrow\overline{\mathcal{B}}}$ y $\overline{\mathcal{B}'}$ la base de K^m tal que $Q = \mathcal{M}_{\mathcal{B}_U^m\leftarrow\overline{\mathcal{B}'}}$. Por lo que hemos visto antes tenemos que $C = \mathcal{M}(f; \overline{\mathcal{B}'} \leftarrow \overline{\mathcal{B}})$. $\square$

5.4. Matriz asociada y operaciones con aplicaciones lineales.

Sean $f, g : V \longrightarrow V'$ dos aplicaciones lineales. Sea $\mathcal{B}$ una base de V, $\mathcal{B}'$ una base de V' y supongamos que $A = \mathcal{M}(f; \mathcal{B}' \leftarrow \mathcal{B})$ y que $C = \mathcal{M}(g; \mathcal{B}' \leftarrow \mathcal{B})$. Tenemos entonces que $Y = AX$ e $\overline{Y} = CX$, siendo X la matriz columna de las coordenadas de un vector $x \in V$ en la base $\mathcal{B}$, Y la matriz columna de las coordenadas de $f(x)$ en $\mathcal{B}'$ e $\overline{Y}$ la matriz columna de las coordenadas de $g(x)$ en $\mathcal{B}'$.
Como $(f+g)(x) = f(x) + g(x)$, las coordenadas de $(f+g)(x)$ en $\mathcal{B}'$ serán $Y + \overline{Y}$ y, por tanto,

$$Y + \overline{Y} = AX + CX = (A + C)X,$$

lo que prueba que $\mathcal{M}(f+g; \mathcal{B}' \leftarrow \mathcal{B}) = A + C = \mathcal{M}(f; \mathcal{B}' \leftarrow \mathcal{B}) + \mathcal{M}(g; \mathcal{B}' \leftarrow \mathcal{B})$.
Análogamente, sabemos que $\forall\, k \in K$, $(kf)(x) = kf(x)$ y las coordenadas de este vector en $\mathcal{B}'$ son kY, luego

$$kY = k(AX) = (kA)X$$

y por tanto, $\mathcal{M}(kf; \mathcal{B}' \leftarrow \mathcal{B}) = k\mathcal{M}(f; \mathcal{B}' \leftarrow \mathcal{B})$.
Veamos ahora qué ocurre con la composición: sea f como antes y $h : V' \longrightarrow V''$ otra aplicación lineal. Sea $\mathcal{B}''$ una base de V'' tal que $\mathcal{M}(h; \mathcal{B}'' \leftarrow \mathcal{B}') = D$. La ecuación matricial de h respecto de $\mathcal{B}'$ y $\mathcal{B}''$ será $Z = DY$.
Si X es la matriz columna de las coordenadas de $x \in V$ en $\mathcal{B}$, $Y = AX$ es la matriz columna de las coordenadas de $f(x)$ en $\mathcal{B}'$ y $Z = DY$ será la matriz columna de las coordenadas de $h(f(x))$ en $\mathcal{B}''$. Por tanto, $Z = DY = D(AX) = (DA)X$ nos dice que

$$\mathcal{M}(h \circ f; \mathcal{B}'' \leftarrow \mathcal{B}) = DA = \mathcal{M}(h; \mathcal{B}'' \leftarrow \mathcal{B}')\mathcal{M}(f; \mathcal{B}' \leftarrow \mathcal{B}).$$

Estas propiedades nos dan también lo siguiente:

Teorema 2.
Sean V y V' espacios vectoriales sobre el cuerpo K con $\dim_K(V) = n$ y $\dim_K(V') = m$. Dadas una base $\mathcal{B}$ de V y una base $\mathcal{B}'$ de V' la aplicación

$$T : Hom_K(V, V') \longrightarrow \mathcal{M}_{m\times n}(K)$$

tal que $T(f) = \mathcal{M}(f; \mathcal{B}' \leftarrow \mathcal{B})$ es un isomorfismo de espacios vectoriales. En particular,

$$\dim_K(Hom_K(V, V')) = m \cdot n.$$

Demostración. Las dos primeras propiedades anteriores nos dan que T es lineal. Es inyectiva puesto que, fijadas las bases, para cada f hay una única matriz que la representa en esas bases. Pero antes vimos que esta aplicación también es sobreyectiva. Por lo tanto es biyectiva. $\square$

Geometría I: Tema 4
Espacio dual de un espacio vectorial

Juan de Dios Pérez

Índice

1. Introducción. **2**

2. Bases duales. **2**

3. Teorema de Reflexividad. **5**

4. Anulador de un subespacio. **6**

5. Aplicación lineal traspuesta. **7**

1. Introducción.

Como sabemos que un cuerpo conmutativo K es un espacio vectorial sobre sí mismo, si V es un espacio vectorial sobre K podemos considerar aplicaciones lineales $f : V \longrightarrow K$, a cada una de las cuales la llamaremos una forma lineal sobre V.
Al conjunto de todas las formas lineales $\{f : V \longrightarrow K \,/\, f \text{ es lineal}\}$ lo notaremos V^* y lo llamaremos el espacio vectorial dual de V. Es un espacio vectorial, pues por lo que hemos visto $V^* = \mathrm{Hom}_K(V, K)$ y, además, si $\dim_K(V) = n$, entonces $\dim_K(V^*) = n$.
Si $f : V \longrightarrow K$ es una forma lineal, podemos suponer en K la base $\{1\}$. Entonces dada $\mathcal{B}$ base de V, podemos considerar $\mathcal{M}(f; \{1\} \leftarrow \mathcal{B})$ que notaremos, simplemente, $\mathcal{M}(f; \mathcal{B})$.

2. Bases duales.

Sean $\mathcal{B} = \{v_1, v_2, \dots, v_n\}$ una base de V y $\mathcal{B}^* = \{\varphi_1, \varphi_2, \dots, \varphi_n\}$ una base de V^*. Diremos que $\mathcal{B}^*$ es la base dual de $\mathcal{B}$ si $\forall\, i \in \{1, \dots, n\}$ se verifica que $\varphi_i(v_i) = 1$, pero $\forall\, j \in \{1, \dots, n\}$, $j \neq i$, se tiene $\varphi_i(v_j) = 0$.

Proposición 1 (1ª propiedad de las bases duales)**.**
Si $\mathcal{B}^*$ es la base dual de $\mathcal{B}$, entonces $\forall\, \varphi \in V^*$, los elementos de su matriz asociada en la base $\mathcal{B}$ coinciden con sus coordenadas en la base $\mathcal{B}^*$.

Demostración. Para verlo, llamemos $A = \begin{pmatrix} a_1 & a_2 & \dots & a_n \end{pmatrix} = \mathcal{M}(f; \mathcal{B})$. Entonces $a_i = \varphi(v_i)$ $\forall\, i \in \{1, \dots, n\}$. Por otra parte, si $\varphi = (b_1, b_2, \dots, b_n)_{\mathcal{B}^*}$, tendremos que $\varphi = b_1\varphi_1 + b_2\varphi_2 + \cdots + b_n\varphi_n$ y así,

$$\varphi(v_i) = a_i = (b_1\varphi_1 + b_2\varphi_2 + \cdots + b_n\varphi_n)(v_i) = b_1\varphi_1(v_i) + \cdots + b_i\varphi_i(v_i) + \cdots + b_n\varphi_n(v_i) = b_i\varphi_i(v_i) = b_i \cdot 1 = b_i,$$

como queríamos. $\square$

La pregunta que nos hacemos ahora es que si dada una base $\mathcal{B}$ de V, podemos calcular su base dual $\mathcal{B}^*$. Sabemos que si $\mathcal{B} = \{v_1, \dots, v_n\}$, una forma lineal está completamente determinada conociendo sus imágenes de los vectores de $\mathcal{B}$. Así, $\forall\, i \in \{1, \dots, n\}$, $\exists_1\, \varphi_i \in V^*$ que verifica $\varphi_i(v_i) = 1$ y $\varphi_i(v_j) = 0$, $\forall\, j \in \{1, \dots, n\}, i \neq j$.
Tenemos así definidas n formas lineales sobre V, $\{\varphi_1, \varphi_2, \dots, \varphi_n\}$. Queremos ver que, efectivamente, forman una base de V^*. Como sabemos que $\dim_K(V^*) = n$, bastará ver que tales formas lineales son linealmente independientes. Supongamos entonces que $\exists\, a_1, \dots, a_n \in K$ tales que $a_1\varphi_1 + a_2\varphi_2 + \cdots + a_n\varphi_n = \varphi_0$ (esta es la forma lineal nula (neutro de V^* para la suma) dada por $\varphi_0(v) = 0$, $\forall\, v \in V$). Entonces, $\forall\, i \in \{1, \dots, n\}$, $\varphi_0(v_i) = 0 = (a_1\varphi_1 + \cdots + a_i\varphi_i + \cdots + a_n\varphi_n)(v_i) = a_1\varphi_1(v_i) + \cdots + a_i\varphi_i(v_i) + \cdots + a_n\varphi_n(v_i) = a_i \cdot 1 = a_i$, y hemos demostrado el siguiente resultado:

Proposición 2.
Para cada base $\mathcal{B}$ de un espacio vectorial V existe una base $\mathcal{B}^*$ de V^* que es dual de $\mathcal{B}$.

Ejemplo 1.

Consideremos en $\mathbb{R}^3$ la base $\mathcal{B} = \{(1,-1,1),(1,2,-1),(-1,1,0)\}$ y calculemos su base dual.

Consideremos $\mathcal{B}_U = \{e_1 = (1,0,0), e_2 = (0,1,0), e_3 = (0,0,1)\}$ la base usual de $\mathbb{R}^3$ y veamos como es su dual $\mathcal{B}_U^* = \{\varphi_1, \varphi_2, \varphi_3\}$. Como $\varphi_i(e_j) = \delta_{ij}\ \forall\, i,j \in \{1,2,3\}$, dando un vector arbitrario $(x_1,x_2,x_3) \in \mathbb{R}^3$ se tiene

$$\begin{aligned}
\varphi_1(x_1,x_2,x_3) &= \varphi_1(x_1e_1 + x_2e_2 + x_3e_3) = x_1,\\
\varphi_2(x_1,x_2,x_3) &= \varphi_2(x_1e_1 + x_2e_2 + x_3e_3) = x_2,\\
\varphi_3(x_1,x_2,x_3) &= \varphi_3(x_1e_1 + x_2e_2 + x_3e_3) = x_3.
\end{aligned}$$

Llamemos ahora $\mathcal{B}^* = \{\psi_1, \psi_2, \psi_3\}$; $\psi_1 = a_1\varphi_1 + a_2\varphi_2 + a_3\varphi_3$, y como $\mathcal{B}^*$ ha de ser la base dual de $\mathcal{B}$, tendremos que

$$\begin{aligned}
\psi_1(1,-1,1) = 1 &= (a_1\varphi_1 + a_2\varphi_2 + a_3\varphi_3)(1,-1,1) = a_1 - a_2 + a_3\\
\psi_1(1,2,-1) = 0 &= (a_1\varphi_1 + a_2\varphi_2 + a_3\varphi_3)(1,2,-1) = a_1 + 2a_2 - a_3\\
\psi_1(-1,1,0) = 0 &= (a_1\varphi_1 + a_2\varphi_2 + a_3\varphi_3)(-1,1,0) = -a_1 + a_2
\end{aligned}$$

De manera que cuando conozcamos a_1, a_2 y a_3 tendremos determinada ψ_1, así que, en definitiva, hemos de resolver el siguiente sistema de Cramer:

$$\begin{cases}
a_1 & -a_2 & +a_3 & = 1\\
a_1 & +2a_2 & -a_3 & = 0\\
-a_1 & +a_2 & & = 0
\end{cases}$$

Como $\begin{vmatrix} 1 & -1 & 1\\ 1 & 2 & -1\\ -1 & 1 & 0\end{vmatrix} = \begin{vmatrix} 1 & -1 & 1\\ 2 & 1 & 0\\ -1 & 1 & 0\end{vmatrix} = \begin{vmatrix} 2 & 1\\ -1 & 1\end{vmatrix} = 3$, tendremos:

$$a_1 = \frac{\begin{vmatrix} 1 & -1 & 1\\ 0 & 2 & -1\\ 0 & 1 & 0\end{vmatrix}}{3} = \frac{1}{3}, \quad a_2 = \frac{\begin{vmatrix} 1 & 1 & 1\\ 1 & 0 & -1\\ -1 & 0 & 0\end{vmatrix}}{3} = \frac{-\begin{vmatrix} 1 & -1\\ -1 & 0\end{vmatrix}}{3} = \frac{1}{3}, \quad a_3 = \frac{\begin{vmatrix} 1 & -1 & 1\\ 1 & 2 & 0\\ -1 & 1 & 0\end{vmatrix}}{3} = 1.$$

Así, $\psi_1 = \frac{1}{3}\varphi_1 + \frac{1}{3}\varphi_2 + \varphi_3$ lo que nos dice que $\forall\,(x_1,x_2,x_3) \in \mathbb{R}^3$ se tiene $\psi_1(x_1,x_2,x_3) = \frac{1}{3}x_1 + \frac{1}{3}x_2 + x_3$.

Si suponemos ahora que $\psi_2 = b_1\varphi_1 + b_2\varphi_2 + b_3\varphi_3$, tendremos, repitiendo el proceso, que

$$\begin{aligned}
\psi_2(v_1) = 0 &= (b_1\varphi_1 + b_2\varphi_2 + b_3\varphi_3)(1,-1,1) = b_1 - b_2 + b_3\\
\psi_2(v_2) = 1 &= (b_1\varphi_1 + b_2\varphi_2 + b_3\varphi_3)(1,2,-1) = b_1 + 2b_2 - b_3\\
\psi_2(v_3) = 0 &= (b_1\varphi_1 + b_2\varphi_2 + b_3\varphi_3)(-1,1,0) = -b_1 + b_2
\end{aligned}$$

Su solución es ahora:

$$b_1 = \frac{\begin{vmatrix} 0 & -1 & 1\\ 1 & 2 & -1\\ 0 & 1 & 0\end{vmatrix}}{3} = \frac{-\begin{vmatrix} -1 & 1\\ 1 & 0\end{vmatrix}}{3} = \frac{1}{3}, \quad b_2 = \frac{\begin{vmatrix} 1 & 0 & 1\\ 1 & 1 & -1\\ -1 & 0 & 0\end{vmatrix}}{3} = \frac{1}{3}, \quad b_3 = \frac{\begin{vmatrix} 1 & -1 & 0\\ 1 & 2 & 1\\ -1 & 1 & 0\end{vmatrix}}{3} = \frac{-\begin{vmatrix} 1 & -1\\ -1 & 1\end{vmatrix}}{3} = 0.$$

Luego $\psi_2 = \frac{1}{3}\varphi_1 + \frac{1}{3}\varphi_2$, lo que nos dice que $\forall\,(x_1,x_2,x_3) \in \mathbb{R}^3$, $\psi_2(x_1,x_2,x_3) = \frac{1}{3}x_1 + \frac{1}{3}x_2$.
Y si $\psi_3 = c_1\varphi_1 + c_2\varphi_2 + c_3\varphi_3$, análogamente nos queda el sistema

$$\begin{cases} c_1 & -c_2 & +c_3 & = 0 \\ c_1 & +2c_2 & -c_3 & = 0 \\ -c_1 & +c_2 & & = 1 \end{cases}$$

cuya solución viene dada por:

$$c_1 = \frac{\begin{vmatrix} 0 & -1 & 1 \\ 0 & 2 & -1 \\ 1 & 1 & 0 \end{vmatrix}}{3} = \frac{\begin{vmatrix} -1 & 1 \\ 2 & -1 \end{vmatrix}}{3} = -\frac{1}{3}, \quad c_2 = \frac{\begin{vmatrix} 1 & 0 & 1 \\ 1 & 0 & -1 \\ -1 & 1 & 0 \end{vmatrix}}{3} = \frac{-\begin{vmatrix} 1 & 1 \\ 1 & -1 \end{vmatrix}}{3} = \frac{2}{3}, \quad c_3 = \frac{\begin{vmatrix} 1 & -1 & 0 \\ 1 & 2 & 0 \\ -1 & 1 & 1 \end{vmatrix}}{3} = 1.$$

Ahora $\psi_3 = -\frac{1}{3}\varphi_1 + \frac{2}{3}\varphi_2 + \varphi_3$, con lo que $\psi_3(x_1,x_2,x_3) = -\frac{1}{3}x_1 + \frac{2}{3}x_2 + x_3$.
Lo importante aquí, es que hemos fijado las bases $\mathcal{B}_U$ de $\mathbb{R}^3$ y $\mathcal{B}_U^*$ de $(\mathbb{R}^3)^*$ de manera que hemos expresado en $\mathcal{B}_U$ los vectores de $\mathcal{B}$ y en $\mathcal{B}_U^*$ las formas lineales de $\mathcal{B}^*$. Lo que hemos hecho, en realidad es imponer que

$$\begin{pmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{pmatrix} \begin{pmatrix} 1 & -1 & 1 \\ 1 & 2 & -1 \\ -1 & 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Así, concluimos que las coordenadas en $\mathcal{B}_U^*$ de las formas de $\mathcal{B}^*$ vienen dadas por las filas de A^{-1}, siendo A la matriz cuyas columnas son las coordenadas de los vectores de $\mathcal{B}$ en $\mathcal{B}_U$.

Este procedimiento es general, y nos sirve también para resolver el problema inverso: dada una base de V^* calcular la base de V de la que es dual, como veremos en el siguiente ejemplo:

Ejemplo 2.
Sean ψ_1, ψ_2, ψ_3 formas lineales sobre $\mathbb{R}^3$ dadas por:

$$\begin{aligned} \psi_1(x_1,x_2,x_3) &= x_1 + 2x_2 + 2x_3 \\ \psi_2(x_1,x_2,x_3) &= x_1 - x_2 \\ \psi_3(x_1,x_2,x_3) &= x_1 \end{aligned}$$

Veamos que $\{\psi_1,\psi_2,\psi_3\} = \overline{\mathcal{B}}$ nos da una base de $\mathbb{R}^3$. Si de nuevo consideramos $\mathcal{B}_U$ y $\mathcal{B}_U^*$, las coordenadas en $\mathcal{B}_U^*$ de ψ_1 son $(1,2,2)$, las de ψ_2 son $(1,-1,0)$ y las de ψ_3 son $(1,0,0)$. Como $\begin{vmatrix} 1 & 2 & 2 \\ 1 & -1 & 0 \\ 1 & 0 & 0 \end{vmatrix} = \begin{vmatrix} 2 & 2 \\ -1 & 0 \end{vmatrix} = 2 \neq 0$, son linealmente independientes y forman una base de $(\mathbb{R}^3)^*$. Si queremos calcular la base $\mathcal{B}$ de $\mathbb{R}^3$ tal que $\overline{\mathcal{B}} = \mathcal{B}^*$, $\mathcal{B} = \{(x_1,x_2,x_3),(y_1,y_2,y_3),(z_1,z_2,z_3)\}$ ha de verificar que

$$\begin{pmatrix} 1 & 2 & 2 \\ 1 & -1 & 0 \\ 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} x_1 & y_1 & z_1 \\ x_2 & y_2 & z_2 \\ x_3 & y_3 & z_3 \end{pmatrix} = I_3.$$

Luego los vectores de $\mathcal{B}$ son las columnas de $\begin{pmatrix} 1 & 2 & 2 \\ 1 & -1 & 0 \\ 1 & 0 & 0 \end{pmatrix}^{-1}$. Así,

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

Así, $\mathcal{B} = \left\{ \left(0, 0, \frac{1}{2}\right), (0, -1, 1), \left(1, 1, \frac{-3}{2}\right) \right\}$.

Proposición 3 (2^{a} propiedad de las bases duales)**.**
Si $\mathcal{B}^* = \{\varphi_1, \varphi_2, \ldots, \varphi_n\}$ es la base dual de $\mathcal{B}$, entonces si para $x \in V$, $x = (x_1, x_2, \ldots, x_n)_{\mathcal{B}}$, tenemos que $x_i = \varphi_i(x)$, $\forall\, i \in \{1, \ldots, n\}$, ya que si $x = x_1v_1 + x_2v_2 + \cdots + x_nv_n$, $\varphi_i(x) = x_1\varphi_i(v_1) + \cdots + x_i\varphi_i(v_i) + \cdots + x_n\varphi_i(v_n) = x_i \cdot 1 = x_i$.

3. Teorema de Reflexividad.

Supongamos que V es un espacio vectorial sobre el cuerpo K con $\dim_K(V) = n$. Sabemos entonces que $\dim_K(V^*) = n$ y, análogamente, si consideramos el espacio bidual de V, $(V^*)^*$, también tendrá dimensión n. Por tanto, V y $(V^*)^*$ serán isomorfos. Como hemos visto, en general, un isomorfismo entre V y $(V^*)^*$ dependerá de bases que escojamos en V y en su bidual. El Teorema de Reflexividad, sin embargo, nos va a dar un isomorfismo intrínseco entre ambos espacios, que no dependerá de bases de los espacios.

Teorema 1 (Teorema de Reflexividad)**.**
La aplicación $\Phi : V \longrightarrow (V^*)^*$, dada por $\Phi(v) = \Phi_v : V^* \longrightarrow K$ tal que $\forall\, \varphi \in V^*$, $\Phi_v(\varphi) = \varphi(v)$ es un isomorfismo de espacios vectoriales.

Demostración. En primer lugar, hemos de ver que Φ está bien definida; es decir, que $\forall\, v \in V$, $\Phi_v : V^* \longrightarrow K$ es lineal. Para ello, sean $a, b \in K$ y $\varphi_1, \varphi_2 \in V^*$. $\Phi_v(a\varphi_1 + b\varphi_2) = (a\varphi_1 + b\varphi_2)(v) = a\varphi_1(v) + b\varphi_2(v) = a\Phi_v(\varphi_1) + b\Phi_v(\varphi_2)$ y, por tanto, $\Phi : V \longrightarrow (V^*)^*$ es una aplicación bien definida.
A continuación, veamos que Φ es lineal, es decir, que $\forall\, c, d \in K$, $\forall\, v_1, v_2 \in V$ se tiene $\Phi(cv_1 + dv_2) = c\Phi(v_1) + d\Phi(v_2)$ o, análogamente, que $\Phi_{cv_1+dv_2} = c\Phi_{v_1} + d\Phi_{v_2}$. Tomemos $\varphi \in V^*$; entonces $\Phi_{cv_1+dv_2}(\varphi) = \varphi(cv_1 + dv_2) = c\varphi(v_1) + d\varphi(v_2) = c\Phi_{v_1}(\varphi) + d\Phi_{v_2}(\varphi) = (c\Phi_{v_1} + d\Phi_{v_2})(\varphi)$, lo que nos proporciona la igualdad deseada.
Finalmente, hemos de ver que Φ es un isomorfismo. Como sabemos que ambos espacios vectoriales tienen la misma dimensión, bastará con ver que Φ es un monomorfismo: sea entonces $v \in \mathrm{Ker}(\Phi)$. Esto significa que $\Phi_v : V^* \longrightarrow K$ es la aplicación lineal cero. Entonces $\forall\, \varphi \in V^*$, $\Phi_v(\varphi) = \varphi(v) = 0$. Veamos que si $v \in V$ verifica que $\forall\, \varphi \in V^*$, $\varphi(v) = 0$, necesariamente $v = 0$: si no fuera asi, $v \neq 0$ y podemos añadir vectores para que $\mathcal{B} = \{v_1, v_2, \ldots, v_n\}$ nos dé una base de V. Sea entonces $\{\varphi_1, \varphi_2, \ldots, \varphi_n\} = \mathcal{B}^*$ su base dual. Claramente $\varphi_1(v) = 1 \neq 0$, en contra de nuestra hipótesis. Así que si $\Phi_v = 0$ entonces $v = 0$ y Φ es, efectivamente, un isomorfismo. $\square$

El Teorema de Reflexividad nos permite identificar V y $(V^*)^*$ como espacios vectoriales haciendo que $v \equiv \Phi_v$. A partir de ahora, consideraremos V y $(V^*)^*$ como el mismo espacio vectorial.

4. Anulador de un subespacio.

Definición 1.
Sea $V(K)$ un espacio vectorial y S un subconjunto de V. Definimos el anulador de S como el conjunto $an(S) = \{\varphi \in V^* \,/\, \phi(v) = 0 \;\forall\, v \in S\}$.

Tenemos las siguientes propiedades para el anulador:

1. $\mathrm{an}(S)$ es un subespacio vectorial de V^*.
2. $\mathrm{an}(S) = \mathrm{an}(\mathcal{L}(S))$.

Demostración. 1. Para demostrar esta propiedad, sean $\varphi_1, \varphi_2 \in \mathrm{an}(S)$ y sean $a, b \in K$. Consideremos $a\varphi_1 + b\varphi_2$ y veamos que pertenece a $\mathrm{an}(S)$. Si tomamos $v \in S$, $(a\varphi_1 + b\varphi_2)(v) = a\varphi_1(v) + b\varphi_2(v) = a \cdot 0 + b \cdot 0 = 0$, pues $\varphi_1(v) = \varphi_2(v) = 0$, ya que $\varphi_1, \varphi_2 \in \mathrm{an}(S)$.

2. Para demostrar esta propiedad, como $S \subseteq \mathcal{L}(S)$, si φ anula a todos los vectores de $\mathcal{L}(S)$, en particular tiene que anular a todos los de S, luego $\mathrm{an}(\mathcal{L}(S)) \subseteq \mathrm{an}(S)$. En el otro sentido, sea $\varphi \in \mathrm{an}(S)$. Cada vector de $\mathcal{L}(S)$ será de la forma $a_1v_1 + \cdots + a_kv_k$ con $v_1, \ldots, v_k \in S$. Entonces $\varphi(a_1v_1 + \cdots + a_kv_k) = a_1\varphi(v_1) + \cdots + a_k\varphi(v_k) = a_1 \cdot 0 + \cdots + a_k \cdot 0 = 0$. Por tanto, $\mathrm{an}(S) \subseteq \mathrm{an}(\mathcal{L}(S))$ y tendremos la igualdad.

$\square$

Si U es un subespacio vectorial de V, la 2ª propiedad nos permite calcular el $\mathrm{an}(U)$ calculando simplemente el anulador de una base de U.

Ejemplo 3.
Consideremos en $\mathbb{R}^4$ el subespacio vectorial $U = \{(x_1, x_2, x_3, x_3) \in \mathbb{R}^4 \,/\, x_1 + x_3 = 0, x_2 + 2x_4 = 0\}$. Todos los vectores de $\mathbb{R}^4$ los referiremos a $\mathbb{B}^4_U$ y todas las formas lineales de $(\mathbb{R}^4)^*$ a la base $(\mathcal{B}^4_U)^* = \{\varphi_1, \varphi_2, \varphi_3, \varphi_4\}$. (Recordemos que $\forall\, i \in \{1, 2, 3, 4\}$, $\varphi_i(x_1, x_2, x_3, x_4) = x_i$).
Calculemos, en primer lugar, una base de U. Unas ecuaciones paramétricas de U serían

$$\begin{cases} x_1 = & \lambda \\ x_3 = & -\lambda \\ x_4 = & \qquad +\mu \\ x_2 = & \qquad -2\mu \end{cases}$$

Así,

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \lambda \begin{pmatrix} 1 \\ 0 \\ -1 \\ 0 \end{pmatrix} + \mu \begin{pmatrix} 0 \\ -2 \\ 0 \\ 1 \end{pmatrix}$$

Lo que nos da que $\{(1,0,-1,0),(0,-2,0,1)\}$ es una base de U. Una forma arbitraria $\psi \in V^*$ será $\psi = a_1\varphi_1 + a_2\varphi_2 + a_3\varphi_3 + a_4\varphi_4$. Para que $\psi \in \mathrm{an}(U) = \mathrm{an}(\mathcal{L}(\{(1,0,-1,0),(0,-2,0,1)\}))$ tendremos que $\psi(1,0,-1,0) = \psi(0,-2,0,1) = 0$.
Entonces $(a_1\varphi_1+a_2\varphi_2+a_3\varphi_3+a_4\varphi_4)(1,0,-1,0) = a_1 - a_3 = 0$ y $(a_1\varphi_1+a_2\varphi_2+a_3\varphi_3+a_4\varphi_4)(0,-2,0,1) = -2a_2 + a_4 = 0$. Luego unas ecuaciones implícitas de $\mathrm{an}(U)$ respecto de $(\mathcal{B}_U^4)^*$ son

$$\begin{cases} a_1 - a_3 = 0 \\ -2a_2 + a_4 = 0 \end{cases}$$

Una base de $\mathrm{an}(U)$ vendrá dada por las formas $\psi_1 = \varphi_1 + \varphi_3$, $\psi_2 = \varphi_2 + 2\varphi_4$. Esto es, $\mathrm{an}(U) = \mathcal{L}(\{\varphi_1+\varphi_3, \varphi_2+2\varphi_4\})$. Donde $(\varphi_1+\varphi_3)(x_1,x_2,x_3,x_4) = x_1 + x_3$ y $(\varphi_2+2\varphi_4)(x_1,x_2,x_3,x_4) = x_2 + 2x_4$.

En general, si suponemos que $\dim_K(V) = n$ y U es un subespacio vectorial de V con dimensión r, siendo $\{u_1, u_2, \ldots, u_r\}$ una base de V, procediendo como en el ejemplo, obtendremos r ecuaciones implícitas para $\mathrm{an}(U)$. Esto nos dice que

$$\dim_K(\mathrm{an}(U)) = \dim_K(V) - \dim_K(U).$$

5. Aplicación lineal traspuesta.

Sean $V(K)$ y $V'(K)$ dos espacios vectoriales y $f : V \longrightarrow V'$ una aplicación lineal. Sean $\mathcal{B}$ una base de V y $\mathcal{B}'$ una base de V' de modo que $\mathcal{M}(f; \mathcal{B}' \leftarrow \mathcal{B}) = A$. Consideremos los espacios duales V^*, con base $\mathcal{B}^*$ y $(V')^*$ con base $(\mathcal{B}')^*$. Veamos cómo podemos definir una aplicación lineal de $(V')^*$ en V^*. Para ello, si $\varphi' \in (V')^*$, aparece el siguiente diagrama:

$$V \xrightarrow{f} V' \xrightarrow{\varphi'} K$$

de manera que $\varphi' \circ f$ es una aplicación lineal de V en K, es decir, $\forall\, \varphi' \in (V')^*$, $\varphi' \circ f \in V^*$.
Definimos entonces $f^t : (V')^* \longrightarrow V^*$ mediante $f^t(\varphi') = \varphi` \circ f$, $\forall\, \varphi' \in (V')^*$.
Tenemos entonces las siguientes propiedades:

1. f^t es una aplicación lineal.
2. $\mathcal{M}(f^t; \mathcal{B}^* \leftarrow (\mathcal{B}')^*) = A^t$.

Demostración. 1. Vamos a demostrar la primera propiedad: sean $\varphi_1', \varphi_2' \in (V')^*$, $a, b \in K$. Entonces $f^t(a\varphi_1' + b\varphi_2') = (a\varphi_1' + b\varphi_2') \circ f = a(\varphi_1' \circ f) + b(\varphi_2' \circ f) = af^t(\varphi_1') + bf^t(\varphi_2')$, ya que $\forall\, v \in V$, $((a\varphi_1' + b\varphi_2') \circ f)(v) = (a\varphi_1' + b\varphi_2')(f(v)) = a\varphi_1'(f(v)) + b\varphi_2'(f(v)) = a(\varphi_1'(f(v))) + b(\varphi_2'(f(v))) = a(\varphi_1' \circ f)(v) + b(\varphi_2' \circ f)(v)$.

2. Para demostrar 2, supongamos que $(\mathcal{B}')^* = \{\varphi_1, \varphi_2, \ldots, \varphi_n\}$. Sea φ_i, $i \in \{1, \ldots, n\}$, $\mathcal{M}(\varphi_i; \mathcal{B}') = \begin{pmatrix} \varphi_i(v_1') & \ldots & \varphi_i(v_n') \end{pmatrix}$, suponiendo que $\mathcal{B}' = \{v_1', \ldots, v_n'\}$. Luego

$$\mathcal{M}(\varphi_i; \mathcal{B}') = \begin{pmatrix} 0 & \ldots & 1 & \ldots & 0 \end{pmatrix}.$$

Entonces

$$\mathcal{M}(\varphi_i \circ f;\mathcal{B}) = \mathcal{M}(\varphi_i;\mathcal{B}')\mathcal{M}(f;\mathcal{B}' \leftarrow \mathcal{B}) = \begin{array}{c} \begin{array}{ccccc} & & i) & & \end{array} \\ \begin{pmatrix} 0 & \dots & 1 & \dots & 0 \end{pmatrix} \end{array} \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{i1} & a_{i2} & \dots & a_{in} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix} =$$

$$= \begin{pmatrix} a_{i1} & a_{i2} & \dots & a_{im} \end{pmatrix}.$$

Esta será la i-ésima columna de $\mathcal{M}(f^t;\mathcal{B}^* \leftarrow (\mathcal{B}')^*)$ con lo cual, esta matriz es A^t.

$\square$

La aplicación f^t se llama la aplicación lineal traspuesta de f.