ÁLGEBRA LINEAL I

15 de noviembre de 2012

UD 1

Sistemas de ecuaciones.

El método de Gauss-Jordan.

Teorema de Rouché- Frobenius.

El determinante.

Aplicaciones del determinante.

SISTEMAS DE ECUACIONES LINEALES.

EL MÉTODO DE GAUSS-JORDAN

SISTEMAS DE ECUACIONES LINEALES

Definición
Una ecuación lineal es una expresión de la forma:

$$a_1x_1 + a_2x_2 + \text{.................} + a_nx_n = b$$

donde los coeficientes a_i y el término independiente b son escalares:

$$a_1, a_2, ..., a_n, b \in K = \mathbb{R} \text{ ó } \mathbb{C}$$

y $x_1, x_2, ..., x_n$ son las incógnitas o variables.

La solución de una ecuación lineal es una serie de números $\alpha_1, \alpha_2, ..., \alpha_n$ que verifican:

$$a_1\alpha_1 + a_2\alpha_2 + ... + a_n\alpha_n = b$$

entonces la ecuación se denomina compatible.

- Una ecuación del tipo $0x_1 + 0x_2 + \cdots + 0x_n = b \Rightarrow 0 = b$ se denomina:
  - Ecuación trivial si $b = 0$
  - Ecuación incompatible si $b \neq 0$

Definición
Un *sistema lineal* es una lista finita de m ecuaciones con n incógnitas de la forma:

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

donde cada fila es una ecuación lineal diferente, donde tenemos n incógnitas $(x_1, x_2, \ldots, x_n)$ en todas la ecuaciones.

Otro tipo de notación sería:

$\sum_{j=1}^{n} a_{ij}x_j = b_i \ (i = 1, \ldots, m)$ en donde las incógnitas son x_j, los coeficientes $a_{ij} \in \mathbb{R}$ y los términos independientes $b_i \in \mathbb{R}$.

SISTEMA DE ECUACIONES HOMOGÉNEO

Definición
Un sistema lineal se denomina *homogéneo* si todos sus términos independientes son nulos:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = 0 \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n = 0 \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n = 0 \end{cases} \equiv \sum_{j=1}^{n} a_{ij}x_j = 0 \ (i = 1, \ldots, m)$$

Definición
Diremos que un sistema lineal es:

- *Compatible*: si admite alguna solución del sistema.
    - *Determinado*: si existe una única solución.
    - *Indeterminado*: si existen infinitas soluciones.
- *Incompatible*: si no existe ninguna solución al sistema.

Nota:
Un sistema homogéneo siempre admite la *solución trivial*:

$$x_j = 0 \qquad \forall\, j = 1, \ldots, n$$

COMBINACIONES LINEALES DE ECUACIONES

Definición
Dos sistemas de ecuaciones lineales son *equivalentes* si toda solución de uno lo es del otro.

Nuestra estrategia para resolver un sistema lineal consistirá en obtener sucesivamente sistemas equivalentes cada vez más sencillos.

Para obtener sistemas de ecuaciones lineales equivalentes podemos realizar las siguientes operaciones:

1. Intercambiar dos ecuaciones
2. Multiplicar una ecuación por un escalar no nulo
3. Añadir a una ecuación un múltiplo no nulo de otra

Si sabemos que una ecuación dada de nuestro sistema es combinación de otras, nos encontramos ante una ecuación superflua que podemos eliminar, simplificando así el sistema.

Nota: Las operaciones elementales conservan la independencia en las ecuaciones.

MÉTODO DE ELIMINACIÓN DE GAUSS

Definición
Un sistema lineal se denomina escalonado reducido cuando la primera incógnita de cada ecuación tiene coeficiente 1 y no aparece en el resto de ecuaciones:

$$\begin{cases} x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = b_1 \\ \qquad\qquad x_2 + \ldots + a_{2n}x_n = b_2 \\ \qquad\qquad\qquad\qquad\qquad \vdots \\ \qquad\qquad\qquad\qquad\quad x_n = b_m \end{cases}$$

Nota: Un sistema escalonado reducido no puede contener ni ecuaciones incompatibles ni triviales.

Método escalonamiento de Gauss-Jordan

Mediante este método se transforma cualquier sistema de ecuaciones en un sistema escalonado reducido, aplicando sucesivamente operaciones elementales.

Algoritmo de Gauss-Jordan

- Paso 0: hacemos $k = 1$
- Paso 1: examinamos las ecuaciones:
  - Si hay alguna incompatible, hemos terminado: el sistema es incompatible.
  - Si hay ecuaciones triviales se suprimen.
- Paso 2: se elige la ecuación cuya primera incógnita tenga índice mínimo.
- Paso 3: se divide la ecuación elegida por el coeficiente de su primera incógnita.
- Paso 4: a cada ecuación le restamos la ecuación obtenida anteriormente multiplicada por el coeficiente que en ella tiene su incógnita.
- Paso 5: repetimos con la siguiente incógnita hasta obtener un sistema reducido escalonado.

Otra manera de ver el Algoritmo consiste en aplicar las operaciones anteriores para obtener ecuaciones lineales equivalentes con el fin de ir obteniendo un sistema de ecuaciones escalonado:

1. Si es necesario intercambiar la primera ecuación con otra para que x_1 aparezca en la primera de las ecuaciones.
2. Eliminar x_1 de todas las ecuaciones excepto de la primera.
3. Repetir el proceso con $x_2, x_3, \ldots, x_n$ hasta obtener un sistema escalonado:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = b_1 \\ \qquad\qquad a_{22}x_2 + \ldots + a_{2n}x_n = b_2 \\ \qquad\qquad\qquad\qquad\qquad\quad \vdots \\ \qquad\qquad\qquad\qquad a_{rn}x_n = b_r \end{cases}$$

Nota:

- Las incógnitas de un sistema escalonado reducido se denominan:
  - *principales:* la primera variable de cada ecuación
  - *secundarias* o parámetros: el resto de variables
- Las ecuaciones de un sistema escalonado reducido son independientes.

DISCUSIÓN Y RESOLUCIÓN DE UN SISTEMA ESCALONADO REDUCIDO

Sea un sistema escalonado reducido con las incógnitas principales $x_1, \ldots, x_r$ y con las incógnitas secundarias $x_{r+1}, ..., x_n$ del modo:

$$\begin{cases} x_1 \qquad\qquad a_{1,r+1}x_{r+1} + \ldots + a_{1n}x_n = b_1 \\ \quad x_2 \qquad\quad a_{2,r+1}x_{r+1} + \ldots + a_{2n}x_n = b_2 \\ \qquad \ddots \qquad\qquad\qquad\qquad \vdots \\ \qquad\quad x_r \ + a_{r,r+1}x_{r+1} + \ldots + a_{rn}x_n = b_r \end{cases}$$

Para resolver este tipo de sistemas basta con despejar las incógnitas principales:

$$\begin{cases} x_1 = b_1 - a_{1,r+1}x_{r+1} - \ldots - a_{1n}x_n \\ x_2 = b_2 - a_{2,r+1}x_{r+1} - \ldots - a_{2n}x_n \\ \vdots \\ x_r = b_r - a_{r,r+1}x_{r+1} - \ldots - a_{rn}x_n \end{cases}$$

De este modo, denominamos $\lambda_j = x_j \ \forall \, j = r+1, ..., n$ para obtener las denominadas ecuaciones paramétricas:

$$\begin{cases} x_1 = b_1 - a_{1,r+1}\lambda_{r+1} - \ldots - a_{1n}\lambda \\ x_2 = b_2 - a_{2,r+1}\lambda_{r+1} - \ldots - a_{2n}\lambda_n \\ \vdots \\ x_r = b_r - a_{r,r+1}\lambda_{r+1} - \ldots - a_{rn}\lambda_n \\ x_{r+1} = \qquad \lambda_{r+1} \\ \vdots \\ x_n = \qquad\qquad\qquad \lambda_n \end{cases}$$

Una vez tenemos los denominados parámetros $\lambda_{r+1}, ..., \lambda_r \in K$, entonces si:

- $r = n$: el sistema no tiene incógnitas secundarias, por tanto es SCD con solución única.
- $r < n$: el sistema tiene al menos una incógnita secundaria, por tanto es SCI con una solución diferente para cada valor de $(\lambda_{r+1}, ..., \lambda_n) \in K^{n-r}$.

Proposición
Dos sistemas escalonados reducidos compatibles equivalentes son idénticos.
De esto se deduce que:

1. Un sistema compatible es equivalente a un único sistema reducido escalonado.
2. Dos sistemas compatibles equivalentes se pueden transformar uno en otro mediante operaciones elementales.

Demostración:
Un sistema y su homogéneo tienen las mismas incógnitas: $x_1, ..., x_n$
¿tienen las mismas incógnitas principales?
Suponemos que no:
$x_1, ..., x_{i-1}$ son las principales en los dos sistemas
x_i es principal sólo en uno de los sistemas
Si hacemos $x_{i+1} = ... = x_n = 0$ obtenemos otros dos sistemas equivalentes con las incógnitas $x_1, ..., x_i$ pero:

- el sistema con las incógnitas principales $x_1, ..., x_i$ no tiene incógnitas secundarias, por lo que es un SCD
- el sistema con las incógnitas principales $x_1, ..., x_{i-1}$tiene una incógnita secundaria x_i, por lo que es un SCI
  Esto es una contradicción, entonces los dos sistemas deben tener las mismas incógnitas principales.

¿tienen los mismos términos independientes?

Hacemos nulas las variables secundarias $x_{r+1}, ..., x_n = 0$, entonces:

$$x_1 = b_1, ..., x_r = b_r$$

$$x_1 = b'_1, ..., x_r = b'_r \quad \Longrightarrow \quad b_i = b'_i$$

Así pues, dos sistemas escalonados reducidos compatibles equivalentes son idénticos.

Veamos ahora la demostración de las dos deducciones:

1. *¿un sistema compatible es equivalente a un único sistema escalonado reducido?*

   Si un sistema compatible es equivalente a dos sistemas escalonados reducidos, éstos son compatibles equivalentes y por la proposición tenemos que son iguales, por tanto es equivalente a un único sistema escalonado reducido.

2. *¿dos sistemas compatibles equivalentes se pueden transformar uno en otro mediante operaciones elementales?*

   Si tenemos dos sistemas compatibles equivalentes, aplicamos el método de Gauss-Jordan para obtener dos sistemas escalonados reducidos que serán compatibles equivalentes, y por el enunciado de la proposición, éstos son idénticos, llegando a ellos mediante operaciones elementales.

TEOREMA DE ROUCHÉ-FROBENIUS

MATRICES

Definición

Una matriz $m \times n$ es una tabla de números del cuerpo K formada por m filas y n columas, tal que:

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} = (a_{ij}) \quad 1 \leq i \leq m, \quad 1 \leq j \leq m, \quad a_{ij} \in K$$

Denominamos *coeficiente* de una matriz A al elemento a_{ij} situado en la fila i y la columna j.

Tipos de matrices:

- Si $m = n$ la matriz A es una matriz cuadrada.
- Si $m = n, \quad a_{ij} = 0 \; \forall \, i \neq j, \quad a_{ij} = 1 \; \forall \, i = j$, A es la matriz identidad de orden n y la denotaremos por I_n. Para definir la matriz identidad se utiliza la denominada delta de Kronecker:

  $\delta_{ij} = \begin{cases} 1 & para \; i = j \\ 0 & para \; i \neq j \end{cases}, \quad I_n = (\delta_{ij})$

Denotaremos $\mathcal{M}_{m \times n}(K)$ al conjunto de todas las matrices $m \times n$ con coeficientes en K.

Operaciones elementales por filas

Estas operaciones son las mismas que para las ecuaciones lineales:

1. Sustituir una fila por el resultado de multiplicarla por un escalar no nulo.
2. Sustituir una fila por el resultado de sumarle otra multiplicada por un escalar no nulo.
3. Intercambiar dos filas.

Definición. Matrices equivalentes
Diremos que dos matrices A y B son equivalentes por filas si podemos transformar una en otra mediante operaciones elementales.

MATRIZ ESCALONADA

Definición. Matriz escalonada reducida por filas
Denominamos matriz escalonada reducida por filas a la matriz cuyas filas no nulas preceden a las nulas y el primer coeficiente no nulo de cada fila precede al primer coeficiente no nulo de la siguiente y es 1 (el único no nulo de su columna)

$$A = \begin{bmatrix} 1 & a_{12} & a_{13} & \cdots & \cdots & \cdots & a_{1n} \\ 0 & 1 & a_{23} & \cdots & \cdots & \cdots & a_{2n} \\ 0 & 0 & 1 & \cdots & \cdots & \cdots & a_{3n} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & \cdots & a_{1k} \\ 0 & 0 & 0 & \cdots & 0 & \cdots & 0 \end{bmatrix}$$

Nota:

$$A = \begin{bmatrix} 1 & 3 & 0 & 5 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2 \end{bmatrix} \text{ no es una matriz escalonada reducida por filas}$$

$$B = \begin{bmatrix} 1 & 3 & 0 & 5 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix} \text{ es una matriz escalonada reducida por filas}$$

Proposición. Escalonamiento de matrices
Toda matriz no nula es equivalente por filas a una matriz escalonada reducida.

Demostración:
Nos basamos en un algoritmo similar al de Gauss-Jordan para la demostración.

1. Se hace $k = 1$
2. Si todas las filas a partir de la k-ésima inclusive son nulas, hemos terminado.
3. Si no, buscamos un coeficiente no nulo $a_{i_k j_k}$con j_k mínimo.
4. Se divide la fila i_k-ésima por su primer coeficiente no nulo, $a_{i_k j_k}$, y se intercambia con la k-ésima fila.
5. A cada fila distinta de la nueva k-ésima le restamos ésta multiplicada por el coeficiente j_k-ésimo de aquélla.
6. Se hace *k=k+1* y volvemos al paso 1.

Ejemplo:

$$\text{Sea la matriz } A = \begin{bmatrix} 1 & 2 & 1 \\ 3 & 7 & 5 \\ 2 & 6 & 8 \end{bmatrix}$$

Tenemos $a_{11} = 1 \neq 0$

$$\begin{pmatrix} 1 & 2 & 1 \\ 3 & 7 & 5 \\ 2 & 6 & 8 \end{pmatrix} \begin{matrix} \\ f_2 - 3f_1 \\ f_3 - 2f_1 \end{matrix} \equiv \begin{pmatrix} 1 & 2 & 1 \\ 0 & 1 & 2 \\ 0 & 2 & 6 \end{pmatrix} \begin{matrix} \\ a_{22} = 1 \neq 0 \\ f_3 - 2f_2 \end{matrix} \equiv \begin{pmatrix} 1 & 2 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 2 \end{pmatrix} \begin{matrix} \\ f_2 - f_3 \\ f_3 : 2 \end{matrix} \equiv \begin{pmatrix} 1 & 2 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{matrix} f_1 - 2f_2 \\ \\ \\ \end{matrix} \equiv$$

$$\begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{matrix} f_1 - f_3 \\ \\ \\ \end{matrix} \equiv \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Proposición
Dos matrices A_1y $A_2 \in \mathcal{M}_{m\times n}(K)$ son equivalentes por filas si, y solo si, los sistemas homogéneos $A_1x^t = 0$ y $A_2x^t = 0$ tienen las mismas soluciones.

Demostración:

$(\Longrightarrow)$

Si A_1y A_2 son equivalentes por filas, las operaciones elementales que transforman A_1 en A_2 se traducen en operaciones elementales que transforman el sistema $A_1x^t = 0$ en el sistema $A_2x^t = 0$ y si los sistemas $A_1x^t = 0$ y $A_2x^t = 0$ son equivalentes entonces tienen las mismas soluciones.

$(\Longleftarrow)$

$A_1x^t = 0$ y $A_2x^t = 0$ tienen las mismas soluciones $\Longrightarrow$ $A_1x^t = 0$ y $A_2x^t = 0$ son equivalentes $\Longrightarrow A_1x^t = 0$ se transforma en $A_2x^t = 0$ por medio de operaciones elementales $\Longrightarrow A_1$ se transforma en A_2 por medio de operaciones elementales.

Proposición
Si dos matrices escalonadas reducidas $R_1,\ R_2 \in \mathcal{M}_{m\times n}(K)$ son equivalentes por filas, entonces son iguales.

Demostración:
Por la proposición anterior tenemos que:
R_1 *y* R_2 son equivalentes por filas $\Longleftrightarrow$ $R_1x^t = 0$ y $R_2x^t = 0$ tienen las mismas soluciones.
Como los sistemas escalonados reducidos $R_1x^t = 0$ y $R_2x^t = 0$ tienen las mismas soluciones, entonces los dos sistemas son iguales $\Longrightarrow$ $R_1 = R_2$

Corolario
Una matriz es equivalente por filas a una única matriz escalonada reducida.

Definición. Rango
Denominamos rango por filas de una matriz A, *rg(A)*, al número de filas no nulas de la única matriz escalonada reducida equivalente por filas a A, es decir, al máximo número de filas independientes de A.

Proposición
Sea $A \in \mathcal{M}_{m\times n}(K)$, entonces:

1. $rg(A) \le m,\ rg(A) \le n$
2. $rg(A)$ es máximo si $rg(A) = r, \quad r = min\{m,n\}$

*Demostración***:**
Suponemos que A tiene r filas independientes.
Denotamos A*'* a la matriz resultante de hacer cero las restantes s filas que dependen de las anteriores r filas:
$\begin{cases} Ax^t = 0 \\ A'x^t = 0 \end{cases}$ son equivalentes $\Longrightarrow$ $A,\ A'$ son equivalentes por filas $\Longrightarrow$ $rg(A) = rg(A') = r$ que por el escalonamiento de la matriz, el número de filas no nulas no puede exceder al número de columnas ni de filas.

Proposición
Si varias filas son independientes y les añadimos una que no depende de ellas, obtenemos de nuevo filas independientes.
En particular, dada una fila a no nula, entonces:

$$rg\begin{pmatrix} a & b \\ 0 & A \end{pmatrix} = 1 + rg(A)$$

Demostración:
Suponemos que las filas $a_1, a_2, \ldots, a_r$ son independientes y otra fila c no depende de ellas, para que esto ocurra, las filas $c, a_1, a_2, ..., a_r$ deben ser independientes, en caso contrario la fila c dependería de las anteriores, cosa que nos llevaría a una contradicción.

Definición. Matriz ampliada
Sea un sistema lineal:

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

cuya escritura matricial es:

$$Ax^t = b^t$$

con $A = (a_{ij})$ la matriz de coeficientes de las incógnitas del sistema.
Definimos la matriz ampliada como:

$$\widetilde{A} = \left(\begin{array}{ccc|c} a_{11} & \cdots & a_{1n} & b_1 \\ \vdots & & \vdots & \vdots \\ a_{m1} & \cdots & a_{mn} & b_m \end{array} \right) = (A|b^t)$$

Teorema de Rouché-Frobenius

Un sistema lineal $Ax^t = b^t$ con n incógnitas es compatible si, y sólo si, $rg(A) = rg(\widetilde{A})$.
En tal caso, el sistema compatible determinado si, y sólo si, $rg(A) = rg(\widetilde{A}) = n$.

Demostración:
Aplicando el algoritmo de Gauss-Jordan al sistema, a la matriz A y a su ampliada $\widetilde{A}$, obtenemos la matriz escalonada reducida R y $\widetilde{R}$ que sólo difieren en la última columna, entonces:

$$rg(R) = rg(\widetilde{R}) = rg(A) = rg(\widetilde{A})$$

Un sistema es compatible determinado si, y sólo si, todas sus variables son principales, es decir, si $rg(A) = rg(R) = n$.

Proposición
El rango por filas de una matriz coincide con su rango por columnas, es decir, el máximo número de filas independientes es igual al máximo número de columnas independientes.

Definición
La traspuesta de una matriz $A = (a_{ij})$ es la matriz $A^t = (b_{ij})$ definida por $b_{ij} = a_{ji}$. Las filas de la matriz traspuesta son las columnas de la matriz dada.

Nota: $\quad rg(A) = rg(A^t)$

OPERACIONES CON MATRICES

Una forma de representar matrices es por cajas, matrices menores que forman una matriz mayor:

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix}$$

donde el número de filas de A y B coincide, como también coinciden el número de filas de C y D, el número de columnas de A y C, y el número de columnas de B y D.

Operaciones lineales

Suma
La suma de dos matrices $A = (a_{ij}),\ B = (b_{ij}) \in \mathcal{M}_{m\times n}(K)$ es la matriz suma:

$$A + B = (a_{ij} + b_{ij}) \in \mathcal{M}_{m\times n}(K)$$

Propiedades:

- Asociativa: $A + (B + C) = (A + B) + C$
- Conmutativa: $A + B = B + A$
- Tiene elemento neutro, la matriz nula formada por ceros: $A + 0 = A$
- Toda matriz tiene su matriz opuesta:

$$\forall\, A = (a_{ij}), \quad \exists\, -A = (-a_{ij}) \,/\, A + (-A) = 0$$

- $(A + B)^t = A^t + B^t$

Producto por escalares
El producto de una matriz $A = (a_{ij}) \in \mathcal{M}_{m\times n}(K)$ y un número $\lambda \in K$ es la matriz:

$$\lambda A = (\lambda a_{ij})$$

Propiedades:

- Distributiva del producto respecto a la suma:
  $\lambda(A + B) = \lambda A + \lambda B$
  $(\lambda + \mu)A = \lambda A + \mu A$
- Asociativa: $(\lambda\mu)A = \lambda(\mu A)$
- El producto por el escalar $\lambda = 1$ no altera la matriz.
- $(\lambda A)^t = \lambda A^t$

Nota: El rango de una matriz no varia al multiplicarla por un escalar no nulo.

Proposición
Sean $A, B \in \mathcal{M}_{m\times n}(K)$ se cumple que:

$$|rg(A) - rg(B)| \leq rg(A + B) \leq rg(A) + rg(B)$$

Demostración:
¿$rg(A+B) \leq rg(A) + rg(B)$?

$$rg(A+B) = rg\begin{pmatrix} A+B \\ 0 \end{pmatrix} \leq rg\begin{pmatrix} A+B \\ B \end{pmatrix} = rg\begin{pmatrix} A \\ B \end{pmatrix} \leq rg(A) + rg(B)$$

¿$|rg(A) - rg(B)| \leq rg(A+B)$? $\iff$ ¿$rg(A) \leq rg(A+B) + rg(B)$?

$$rg(A) = rg(A+B-B) = rg((A+B)+(-B)) \leq$$

$$\leq rg(A+B) + rg(-B) = rg(A+B) + rg(B)$$

Producto de matrices
El producto de dos matrices $A = (a_{ij}) \in \mathcal{M}_{m\times n}(K),\ B = (b_{kj}) \in \mathcal{M}_{n\times p}(K)$ es la matriz:

$$A \cdot B = (c_{ij}) \in \mathcal{M}_{m\times p}(K)$$

cuyos coeficientes son: $c_{ij} = \sum\limits_{k=1}^{n} a_{ik}b_{kj}$

Propiedades:

- Asociativa:
  - $(A\cdot B)\cdot C = A\cdot(B\cdot C)$
  - $\lambda(A\cdot B) = (\lambda A)\cdot B = A\cdot(\lambda B) \quad \forall\,\lambda \in K$
- Distributiva del producto respecto a la suma:
  - $A\cdot(B+C) = A\cdot B + A\cdot C$
  - $(A+B)\cdot C = A\cdot C + B\cdot C$
- Sean I_m, I_n las matrices identidad de orden m y n:

$$I_m \cdot A = A = A \cdot I_n \quad \forall\, A \in \mathcal{M}_{m\times n}(K)$$

- $(A\cdot B)^t = B^t \cdot A^t$

Nota: El producto de matrices no es conmutativo.

Proposición.
Combinaciones de filas y columnas en el producto de matrices
Sea el producto de matrices $C = AB$ se cumple:
Una combinación de filas (respectivamente columnas) en C se obtiene haciendo la misma combinación en A y después multiplicando por B (respectivamente en B y después multiplicando por A).

Nota: Lo mismo vale para las operaciones elementales incluido el intercambio de filas o columnas.

Demostración:
La fila i-ésima de C se obtiene de multiplicar la fila i-ésima de A por las columnas de B, de ahí que cualquier intercambio de filas en A se produce igual en C.
Veamos ahora las combinaciones de filas:
Sean $\alpha, \beta \in K$, elegimos los índices i, j. Sea A' la matriz que se obtiene al cambiar la fila i-ésima a_i de A por la combinación $\alpha a_i + \beta a_j$ y C' la que resulta de hacer lo mismo en la matriz C.
¿C' y $D = A'B$ son iguales? $\iff$ son iguales las filas i-ésimas de ambas matrices

$$d_{il} = \sum_k a'_{ik}b_{kl} = \sum_k(\alpha a_{ik} + \beta a_{jk})b_{kl} = \alpha\sum_k a_{ik}b_{kl} + \beta\sum_k a_{jk}b_{kl} =$$

$$= \alpha c_{il} + \beta c_{jl} = c'_{il} \;\Rightarrow\; C' = A'B$$

Proposición
Sean las matrices $A \in \mathcal{M}_{m\times n}(K)$ y $B \in \mathcal{M}_{n\times p}(K)$ se cumple que:

$$rg(AB) \leq min\{rg(A), rg(B)\}$$

Demostración:
Sea $C = AB \in \mathcal{M}_{m\times p}(K)$. Si la fila i-ésima de A depende de sus anteriores, lo mismo le pasa a la matriz C y como el rango es el máximo de filas independientes:

$$C = AB, \quad rg(C) \leq rg(A)$$

Si la columna j-ésima de B depende de sus anteriores, lo mismo le pasa a la matriz C:

$$C = AB, \quad rg(C) \leq rg(B)$$

Otra manera de verlo es:

$$rg(C) = rg(C^t) = rg(B^t A^t) \leq rg(B^t) = rg(B)$$

Definición. Matrices cuadradas
Las *matrices cuadradas* son las que tienen el mismo número de filas que de columnas.
El conjunto de todas las matrices cuadradas de orden n se denota como $\mathcal{M}_n(K)$.
La diagonal principal de una matriz cuadrada $A = (a_{ij})$ está formada por los coeficientes a_{ii}.
Una *matriz triangular* es una matriz cuadrada cuyos coeficientes no nulos están en la diagonal principal o en un sólo lado respecto a la diagonal.
Una *matriz diagonal* es una matriz cuadrada cuyos coeficientes no nulos están en la diagonal principal.

Ejemplo: la matriz identidad I_n es una matriz diagonal.

Matrices invertibles

La multiplicación de matrices es una operación bien definida en $\mathcal{M}_n(K)$ por lo que pueden existir matrices inversas.

Definición
Una matriz cuadrada A de orden n se llama *invertible* o *regular* si existe otra matriz $C \in \mathcal{M}_n(K)$ tal que:

$$AC = CA = I_n$$

La matriz C es única y se denomina inversa de A:

$$C = A^{-1}$$

Demostración de la unicidad de la inversa:
Suponemos que existen C, C' matrices inversas de A, entonces:

$$AC = CA = I_n$$

$$AC' = C'A = I_n$$

$$C' = C'I_n = C'(CA) = C'(AC) = (C'A)C = I_nC = C \quad \Longrightarrow \quad C' = C$$

Proposición
Una matriz cuadrada es invertible si, y sólo si, tiene rango máximo.

Proposición
El producto de dos matrices es invertible si, y sólo si, son invertibles ambas matrices; ya que el rango del producto es máximo si, y sólo si, lo es el de cada factor.

Demostración:
($\Longrightarrow$)
Si AB es invertible entonces:

$$\exists C \;/\; C(AB) = I_n = (AB)C$$

Por la propiedad asociativa:

$$C(AB) = (CA)B = I_n = (AB)C = A(BC)$$

entonces: CA es la inversa de B y BC es la inversa de A.

($\Longleftarrow$)
Si A, B son invertibles entonces $B^{-1}A^{-1}$ es la inversa de AB

Nota: El conjunto de las matrices invertibles de orden n se denota $GL(n, K)$ o $GL(n)$ denominado *Grupo lineal* (es un grupo con la multiplicaión de matrices).

Propiedad
Si A es una matriz cuadrada invertible de orden m y B es una matriz $m \times n$, entonces:

$$rg(AB) = rg(B)$$

Demostración:
Sabemos que $rg(AB) \leq rg(B)$
Por otro lado:

$$rg(B) = rg((A^{-1}A)B) = rg(A^{-1}(AB)) \leq rg(AB)$$

Si hemos obtenido:

$$\begin{array}{l} rg(AB) \leq rg(B) \\ rg(B) \leq rg(AB) \end{array} \Longrightarrow rg(AB) = rg(B)$$

Cálculo de la inversa

La matriz inversa se obtiene en realidad resolviendo sistemas lineales.
Supongamos que A es una matriz cuadrada de orden n, consideremos los sistemas:

$$Ax^t = (1, 0, \ldots, 0)^t, \ldots, Ax^t = (0, \ldots, \overset{(j)}{1}, \ldots, 0)^t, \ldots, Ax^t = (0, \ldots, 0, 1)^t$$

cuyas soluciones, si existen, son las columnas de A^{-1}.
Podemos resolverlo aplicando operaciones elementales a la matriz $(A|I_n)$ y obtendremos una matriz escalonada reducida de la forma $(I_n|C)$ cuando A sea invertible, siendo $C = A^{-1}$.

$$(A|I_n) \rightsquigarrow (I_n|A^{-1})$$

Potenciación
Las potencias de una matriz cuadrada $A \in \mathcal{M}_n(K)$ se definen como:

$$A^n = \begin{cases} A^0 = 1 \\ A^1 = A \\ A^{k+1} = A^k \cdot A \qquad k \geq 1 \end{cases}$$

Propiedades:
Por la propiedad asociativa se cumple que:

$$A^p A^q = A^{p+q}$$

Como el producto de matrices no es conmutativo:

$$(AB)^p \neq A^p B^p$$

sólo sería cierta en el caso de que A y B conmutaran.
Si A y B conmutan, tiene sentido la fórmula de Newton:

$$(A+B)^p = \sum_{k=0}^{p} \binom{p}{k} A^k B^{p-k}$$

En algunas ocasiones nos valdremos de la fórmula de Newton para calcular potencias de una matriz, descomponiéndola en la suma de dos más simples (buscando, a ser posible, que una de ellas sea la matriz identidad y así simplificar los cálculos).

EL DETERMINANTE

Para una escritura más cómoda representaremos cada matriz de orden n como una upla de longitud n^2:

$$\mathcal{M}_n(K) \rightsquigarrow K^n \times \overset{(n)}{\cdots} \times K^n : A = (a_{ij}) \rightsquigarrow (a_1, \ldots, a_n),\ a_i = (a_{i1}, \ldots, a_{in}) \in K^n$$

Definición y proposición
Existe una única aplicación, llamada determinante:

$$det:\ \mathcal{M}_n(K) \to K$$

que cumple las siguientes propiedades:

1. $det(I_n) = 1$. Esta propiedad garantiza la unicidad del determinante.
2. Para cualquiera $\alpha, \beta \in K$, con:

$$a_i', a_i'', a_1, a_2, \ldots, a_i = \alpha a_i' + \beta a_i'', \ldots, a_n \in K^n$$

se cumple:

$$det(a_1, \ldots, \alpha a_i' + \beta a_i'', \ldots, a_n) =$$

$$= \alpha \cdot det(a_1, \ldots, a_i', \ldots, a_n) + \beta \cdot det(a_1, \ldots, a_i'', \ldots, a_n)$$

Las combinaciones de filas son compatibles con el determinante.

Nota: Si una fila es nula, el determinante es cero.

1. Si la matriz A tiene dos filas iguales, entonces: $det(A) = 0$.
   Cuando el rango no es máximo, el determinante se anula.

Propiedades
Las propiedades del determinante son en realidad reglas de cálculo.

1. Si se multiplica una fila por un escalar, el determinante queda multiplicado por dicho escalar.

$$det(a_1, \ldots, \lambda a_i, \ldots, a_n) = \lambda \cdot det(a_1, \ldots, a_i, \ldots, a_n)$$

   Esto se demuestra utilizando la segunda propiedad de la definición de determinante con $\alpha = \lambda, \beta = 0$.

2. Si a una fila le sumamos un múltiplo de otra, el determinante no varía.

$$det(a_1, \ldots, a_i + \mu a_j, \ldots, a_j, \ldots, a_n) =$$

$$= det(a_1, \ldots, a_i, \ldots, a_j, \ldots, a_n) + \mu det(a_1, \ldots, a_j, \ldots, a_j, \ldots, a_n) =$$

$$= det(a_1, \ldots, a_i, \ldots, a_j, \ldots, a_n) + \mu \cdot 0$$

3. Si intercambiamos dos filas, el determinanate cambia de signo.

   Suponemos que queremos cambiar las filas a_i, a_j:

$$0 = det(a_1, \dots, a_i + a_j, \dots, a_i + a_j, \dots, a_n) =$$

$$= det(a_1, \dots, a_i, \dots, a_i + a_j, \dots, a_n) + det(a_1, \dots, a_j, \dots, a_i + a_j, \dots, a_n) =$$

$$= det(a_1, \dots, a_i, \dots, a_i, \dots, a_n) + det(a_1, \dots, a_i, \dots, a_j, \dots, a_n) +$$

$$+ det(a_1, \dots, a_j, \dots, a_i, \dots, a_n) + det(a_1, \dots, a_j, \dots, a_j, \dots, a_n) =$$

$$= 0 + det(a_1, \dots, a_i, \dots, a_j, \dots, a_n) + det(a_1, \dots, a_j, \dots, a_i, \dots, a_n) + 0$$

Por tanto:

$$0 = det(a_1, \dots, a_i, \dots, a_j, \dots, a_n) + det(a_1, \dots, a_j, \dots, a_i, \dots, a_n)$$

$$det(a_1, \dots, a_i, \dots, a_j, \dots, a_n) = -det(a_1, \dots, a_j, \dots, a_i, \dots, a_n)$$

Nota: Si una matriz B se obtiene a partir de otra A mediante operaciones elementales por filas, entonces $det(B)$ se obtiene multiplicando $det(A)$ por los factores asociados a las operaciones elementales realizadas.

Propiedad. *Escalonamiento*
Podemos calcular el determinante de una matriz A de orden n haciendo operaciones elementales para convertir A en una matriz escalonada reducida R.

- Si $rg(A) = n \longrightarrow R = I_n \longrightarrow det(R) = 1$
- Si $rg(A) < n \longrightarrow R$ tiene alguna fila nula $\longrightarrow det(R) = 0$

$$\begin{array}{ccccc} A & \rightsquigarrow \cdots \rightsquigarrow & B \rightsquigarrow B' & \rightsquigarrow \cdots \rightsquigarrow & R \\ det(A) & \rightsquigarrow \cdots \rightsquigarrow & det(B) \overset{\cdot\mu}{\rightsquigarrow} det(B') & \rightsquigarrow \cdots \rightsquigarrow & det(R) = \begin{cases} 0 \\ 1 \end{cases} \end{array}$$

donde B' se obtiene realizando operaciones elementales con el factor $\mu \neq 0$ sobre B, entonces $det(B') = \mu det(B)$. Así, leyendo de derecha a izquierda podemos calcular $det(A)$.

Nota: El determinante de una matriz cuadrada es no nulo si, y sólo si, su rango es máximo.

Proposición
El determinante de una matriz cuadrada existe y es único.

Demostración:

Unicidad del determinante:

Para $n = 1$:

$$det(a) = det(a \cdot I_n) = a \cdot det(I_n) = a \cdot 1 = a$$

por lo que la función det está bien determinada.

Suponemos $n \geq 2$, vamos a obtener la fórmula denominada *Regla de Laplace*:

$$det(A) = \sum_{j=1}^{n} (-1)^{1+j} a_{1j} det(A_{1j})$$

donde A_{1j} es la matriz de orden $n-1$ que se obtiene al suprimir en A la primera fila y la columna j-ésima.

$$det \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = a_{11}a_{22} - a_{12}a_{21}$$

La regla de Laplace nos dice como calcular el determinante, luego prueba su unicidad por el método de inducción. Para empezar, escribimos la matriz dada por filas:

$$A = (a_1, \ldots, a_n)$$

y la primera de las filas como combinación lineal:

$$a_1 = \sum_{j=1}^{n} a_{1j} e_j, \quad e_j = (0, \ldots, 0, \overset{(j)}{1}, 0, \ldots, 0)$$

entonces:

$$det(A) = det(\sum_{j=1}^{n} a_{1j} e_j, a_2, \ldots, a_n) = \sum_{j=1}^{n} a_{1j} \cdot det(e_j, a_2, \ldots, a_n)$$

así afirmamos, lo que prueba la regla de Laplace, que se cumple:

$$det(e_j, a_2, \ldots, a_n) = (-1)^{1+j} det(A_{1j})$$

Hacemos operaciones elementales en A_{1j} para transformarla en su matriz escalonada reducida, que será I_{n-1} o una matriz con una fila nula.

Existencia del determinante:

Construimos la aplicación det por inducción sobre el orden n de las matrices.
Para $n = 1$:

$$det(a) = a$$

cumple las propiedades de la definición de determinante.

Suponemos construido el determinante de orden $n-1 \geq 1$ y definimos el de orden n mediante la regla de Laplace, entonces vamos a verificar que cumple las tres propiedades de la definición:

1. Es inmediata ya que al aplicar la regla de Laplace a $A = I_n$, el único coeficiente no nulo de la primera fila es el primero, que vale 1, y la matriz correspondiente es $A_{11} = I_{n-1}$ que por inducción tiene determinante 1.

2. ¿$det(\ldots, \alpha a'_i + \beta a''_i, \ldots) = \alpha det(\ldots, a'_i, \ldots) + \beta det(\ldots, a''_i, \ldots)$?

   Suponemos primero $i = 1$, entonces tenemos:

$$det(\alpha a'_1 + \beta a''_1, a_2, \ldots, a_n) = \sum_{j=1}^{n} (\alpha a'_{1j} + \beta a''_{1j})(-1)^{1+j} det(A_{1j}) =$$

$$= \alpha \sum_{j=1}^{n} a'_{1j}(-1)^{1+j} det(A_{1j}) + \beta \sum_{j=1}^{n} a''_{1j}(-1)^{1+j} det(A_{1j}) =$$

$$= \alpha det(a'_1, a_2, \ldots, a_n) + \beta det(a''_1, a_2, \ldots, a_n)$$

   En el caso $i = 1$ la operación de combinar filas queda verificada.

   Si la fila es $i > 1$, la operación de combinar filas se reproduce en todos los determinantes del sumatorio y por inducción podemos aplicar la propiedad 2 para obtener:

$$det(a_1, \ldots, \alpha a'_i + \beta a''_i, \ldots, a_n) = \sum_{j=1}^{n} a_{1j}(-1)^{1+j}(\alpha det(A'_{1j}) + \beta det(A''_{1j})) =$$

$$= \alpha \sum_{j=1}^{n} a_{1j}(-1)^{1+j} det(A'_{1j}) + \beta \sum_{j=1}^{n} a_{1j}(-1)^{1+j} det(A''_{1j}) =$$

$$= \alpha det(a_1, \ldots, a'_i, \ldots a_n) + \beta det(a_1, \ldots, a''_i, \ldots a_n)$$

   Por tanto queda verficada la segunda propiedad.

3. ¿El determinante vale 0 si hay dos filas iguales?

   Para $n = 2$:

$$det(A) = det \begin{pmatrix} a_{11} & a_{12} \\ a_{11} & a_{12} \end{pmatrix} = a_{11}a_{12} - a_{12}a_{11} = 0$$

   Suponemos $n \geq 3$ y que A tiene dos filas iguales. Si ninguna de las filas iguales es la primera, entonces todas las matrices A_{1j} de la regla de Laplace tienen una par de filas repetidas y por inducción tienen el determinante nulo, por tanto el sumatorio es nulo; y en consecuencia $det(A) = 0$.

   Ahora bien, suponemos que las filas iguales son las dos primeras (es lo mismo que si fuera la primera y otra fila i-ésima con $i > 2$, ya que intercambiando filas el determinante varía de signo, pero como será nulo el signo no nos importa).

$$det(A) = \sum_{j=1}^{n} (-1)^{1+j} a_{1j} det(A_{1j})$$

   aplicamos la regla de Laplace a cada determinante $det(A_{1j})$.

   La primera fila de cada matriz A_{1j} es:

$$\left(\overset{(1)}{a_{21}}, \ldots, \overset{(j-1)}{a_{2j-1}}, \overset{(j)}{a_{2j+1}}, \ldots, \overset{(n-1)}{a_{2n}} \right)$$

   luego:

$$det(A_{1j}) = \cdots + (-1)^{\sigma_k} a_{2k} det(B_{jk}) + \cdots$$

donde:

- B_{jk} es la matriz obtenida al suprimir en A_{1j} la fila y la columna donde se encuentra el término a_{2k}, esto es, al suprimir en A las filas primera y segunda, y las columnas j-ésima y k-ésima.
- $\sigma_k = \begin{cases} 1+k, & k<j \\ 1+(k-1)=k, & k>j \end{cases}$

Así:

$$det(A_{1j}) = \sum_{k=1}^{j-1}(-1)^{1+k}a_{2k}det(B_{jk}) + \sum_{k=j+1}^{n}(-1)^{k}a_{2k}det(B_{jk}) =$$

$$= \sum_{k=1}^{j-1}(-1)^{1+k}a_{2k}det(B_{jk}) - \sum_{k=j+1}^{n}(-1)^{1+k}a_{2k}det(B_{jk})$$

Llevando ésto al cálculo de $det(A)$ obtenemos:

$$det(A) = \sum_{j=1}^{n}(-1)^{1+j}a_{1j}(\sum_{k=1}^{j-1}(-1)^{1+k}a_{2k}det(B_{jk}) - \sum_{k=j+1}^{n}(-1)^{1+k}a_{2k}det(B_{jk})) =$$

$$= \sum_{1\leq k<j\leq n}(-1)^{j+k}a_{1j}a_{2k}det(B_{jk}) - \sum_{1\leq j<k\leq n}^{n}(-1)^{j+k}a_{1j}a_{2k}det(B_{jk})$$

Hacemos un cambio de índices en la sumación:

$$= \sum_{1\leq p<q\leq n}(-1)^{p+q}(a_{1q}a_{2p}det(B_{qp}) - a_{1p}a_{2q}det(B_{pq})) =$$

observamos que por construcción $B_{qp} = B_{pq}$:

$$= \sum_{1\leq p<q\leq n}(-1)^{p+q}(a_{1q}a_{2p} - a_{1p}a_{2q})det(B_{pq})$$

como nuestra matriz A tiene las dos primeras filas iguales:

$$a_{1q}a_{2p} - a_{1p}a_{2q} = 0 \implies det(A) = 0$$

Regla de Sarrus
Cálculo del determinante de orden 3 por la regla de Laplace, su resultado se conoce como Regla de Sarrus.

$$det\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} =$$

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

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

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

Nota: El determinante de una matriz diagonal es el producto de los coeficientes de su diagonal principal.

Proposición
El determinante de un producto de matrices cuadradas es el producto de sus determinantes.

Demostración:

Sea AB el producto de dos matrices cuadradas. El rango del producto es máximo si, y sólo si, lo es el de los factores. Así sólo nos ocuparemos del caso en que A y B tengan rango máximo, es decir, tengan determinante no nulo.

Si A tiene rango máximo, mediante operaciones elementales por filas, A es equivalente a una matriz escalonada reducida R con todas las filas no nulas, que es por tanto la identidad.

Si aplicamos esas mismas operaciones elementales a AB, obtenemos $RB = B$.

Invirtiendo las operaciones vemos que las mismas que transforman R en A, sirven para transformar $B = RB$ en AB.

Así, los factores que se aplican para obtener $det(A)$ a partir de $det(R) = 1$, son los mismos que se aplican para obtener $det(AB)$ a partir de $det(RB) = det(B)$.

Entonces:

$$det(AB) = det(A) \cdot det(B)$$

Proposición
Los determinantes por filas y por columnas coinciden.

Demostración:

¿$det_c(A) = det_f(A)$*?*

Si A no es invertible, tanto su rango por columnas como por filas son menores que su orden, entonces:

$$det_c(A) = 0 = det_f(A)$$

Suponemos que A es invertible. Para calcular el determinante por filas de A se hacen las operaciones elementales por filas hasta obtener la matriz identidad, así el $det_f(A)$ es el inverso del producto de los factores asociados a esas operaciones:

$$\begin{array}{ccccc} A & \rightsquigarrow \cdots \rightsquigarrow & B \rightsquigarrow B' & \rightsquigarrow \cdots \rightsquigarrow & R = I_n \\ det_f(A) & \rightsquigarrow \cdots \rightsquigarrow & det_f(B) \overset{\cdot\mu}{\rightsquigarrow} det_f(B') & \rightsquigarrow \cdots \rightsquigarrow & det_f(R) = 1 \end{array}$$

Podemos utilizar esta sucesión de operaciones elementales para calcular el determinante por columnas:

$$det_c(B) \overset{\cdot\mu}{\rightsquigarrow} det_c(B') \quad \Longrightarrow \quad det_c(B') = \mu det_c(B)$$

Empecemos por considerar la matriz I' que se obtiene a partir de la identidad I mediante la misma operación elemental por filas que B' a partir de B.

Por la forma en que se comporta el producto respecto a estas operaciones; como $B = IB$ entonces tenemos que $B' = I'B$.

Así tenemos:

$$det_c(B') = det_c(I'B) = det_c(I') \cdot det_c(B)$$

¿$det_c(I') = \mu$*?*

Es cierto porque las operaciones elementales por filas con operaciones por columnas del mismo tipo, por tanto son las mismas que para la matriz identidad I.

En conclusión:

$$det_f(A) = det_c(A)$$

Nota: Como el determinanate por filas coincide con el determinante por columnas, entonces:

$$det(A) = det(A^t)$$

Definición. *Traza*
Definimos la *traza* como la suma de los elementos de la diagonal principal de una matriz cuadrada:

$$tr(A) = \sum_i a_{ii}$$

La traza es lineal, por lo que conserva la suma y el producto por escalares.

Definición. *Determinante de Vandermonde*
Sean $a_1, \ldots, a_n \in K$, contruimos la matriz:

$$A = \begin{pmatrix} 1 & 1 & \cdots & 1 \\ a_1 & a_2 & \cdots & a_n \\ a_1^2 & a_2^2 & \cdots & a_n^2 \\ \vdots & \vdots & & \vdots \\ a_1^{n-1} & a_2^{n-1} & \cdots & a_n^{n-1} \end{pmatrix}$$

denominamos *determinante de Vandermonde*:

$$\triangle(a_1, \ldots, a_n) = \prod_{1 \leq i < j \leq n} (a_j - a_i)$$

Demostración: por inducción.

Para $n = 2$:

$$\triangle(a_1, a_2) = det \begin{pmatrix} 1 & 1 \\ a_1 & a_2 \end{pmatrix} = a_1 - a_2$$

Para $n = 3$:

$$\triangle(a_1, a_2, a_3) = det \begin{pmatrix} 1 & 1 & 1 \\ a_1 & a_2 & a_3 \\ a_1^2 & a_2^2 & a_3^2 \end{pmatrix} = det \begin{pmatrix} 1 & 1 & 1 \\ 0 & a_2 - a_1 & a_3 - a_1 \\ 0 & a_2(a_2 - a_1) & a_3(a_3 - a_1) \end{pmatrix} =$$

$$= (a_2 - a_1)(a_3 - a_1) det \begin{pmatrix} 1 & 1 \\ a_2 & a_3 \end{pmatrix} = (a_2 - a_1)(a_3 - a_1)(a_3 - a_2)$$

Para n:

$$\triangle(a_1, \ldots, a_n) = det \begin{pmatrix} 1 & 1 & \cdots & 1 \\ a_1 & a_2 & \cdots & a_n \\ a_1^2 & a_2^2 & \cdots & a_n^2 \\ \vdots & \vdots & & \cdots \\ a_1^{n-1} & a_2^{n-1} & \cdots & a_n^{n-1} \end{pmatrix} =$$

restamos a cada fila (excepto a la primera) el producto de a_1 por la anterior:

$$= det \begin{pmatrix} 1 & 1 & \cdots & 1 \\ 0 & a_2 - a_1 & \cdots & a_n - a_1 \\ 0 & a_2(a_2 - a_1) & \cdots & a_n(a_n - a_1) \\ \vdots & \vdots & & \cdots \\ 0 & a_2^{n-2}(a_2 - a_1) & \cdots & a_n^{n-2}(a_n - a_1) \end{pmatrix} =$$

desarrollamos el determinante por la primera columna y extraemos de cada columna los factores $(a_2 - a_1), \ldots, (a_n - a_1)$:

$$= (a_2 - a_1) \cdots (a_n - a_1) det \begin{pmatrix} 1 & 1 & \cdots & 1 \\ a_2 & a_3 & \cdots & a_n \\ a_2^2 & a_3^2 & \cdots & a_n^2 \\ \vdots & \vdots & & \cdots \\ a_2^{n-2} & a_3^{n-2} & \cdots & a_n^{n-2} \end{pmatrix} =$$

$$= (a_2 - a_1) \cdots (a_n - a_1) \triangle(a_2, \ldots, a_n)$$

Por la hipótesis de inducción:

$$\triangle(a_2, \ldots, a_n) = \prod_{2 \leq i < j \leq n} (a_j - a_i)$$

Entonces:

$$\triangle(a_1, \ldots, a_n) = (a_2 - a_1) \cdots (a_n - a_1) \triangle(a_2, \ldots, a_n) =$$

$$= (a_2 - a_1) \cdots (a_n - a_1) \prod_{2 \leq i < j \leq n} (a_j - a_i) = \prod_{1 \leq i < j \leq n} (a_j - a_i)$$

Nota: El determinante de Vandermonde es no nulo si, y sólo si, todos los escalares a_i son distintos.

APLICACIONES DEL DETERMINANTE

Definición
Un menor de orden s de una matriz $A \in \mathcal{M}_{m\times n}(K)$ es el determinante de una matriz cuadrada de orden s que se obtiene a partir de A eliminando $m-s$ de sus filas y $n-s$ columnas.

Proposición
Sea A una matriz no necesariamente cuadrada, y sea D un menor no nulo de A. Entonces el rango de A es el mayor de los órdenes de los menores no nulos que contienen a D.

Demostración:

Supongamos que $\overline{D}$ es un menor de orden s no nulo, cuya matriz denotamos M, por tener determinante no nulo, el rango de M es máximo, luego sus s filas son independientes. Cada una de estas filas es parte de alguna fila de A por lo que A tiene al menos s filas independientes:

$$rg(A) \geq s$$

Si $rg(A) > s$, podemos añadir a M una fila y una columna para obtener otro menor no nulo de orden $s+1$. A tiene que tener alguna fila que no dependa de esas s filas. Con ella y las s que ya tenemos formamos una submatriz A' de A que tiene $s+1$ filas independientes y tantas columnas como A. Además, A' contiene a nuestra submatriz M.

Si las demás columnas de A' dependieran de aquellas s, el $rg(A) = s$ y no es el caso, por lo que hay alguna columna de A' que no depende de las s en cuestión.

Sea M' la submatriz de A' cuyas columnas son las s columnas independientes de A' que contienen a las s columnas de M y a la nueva columna que no depende de ellas. Las $s+1$ columnas de M' son independientes, M' es cuadrada de orden $s+1$ con determinante no nulo y M' contiene a M, por lo que podemos concluir con que el $rg(A)$ es el mayor de los órdenes de los menores no nulos.

Matriz adjunta y matriz inversa
Sea $A = (a_{ij})$ una matriz cuadrada de orden n.

El ***adjunto*** de un elemento a_{ij} es:

$$\alpha_{ij} = (-1)^{i+j} det(A_{ij})$$

donde A_{ij} es la matriz que resulta al eliminar la fila i-ésima y la columna j-ésima.

Se cumple que:

$$\sum_{i=1}^{n} \alpha_{ij} a_{ik} = \sum_{i=1}^{n} (-1)^{i+j} a_{ik} det(A_{ij}) = \begin{cases} det(A) & si\ k = j \\ 0 & si\ k \neq j \end{cases}$$

Si $k = j$, la matriz es A; si $k \neq j$, la matriz es la matriz A' que resulta al sustituir la columna j-ésima de A por una copia de la k-ésima, y como A' tiene dos columnas iguales su determinante es 0.

Definimos la ***Matriz Adjunta*** de A como:

$$Adj(A) = (\alpha_{ij})^t$$

Si A es una matriz invertible:

$$A^{-1} = \frac{Adj(A)}{det(A)}$$

Nota: Una matriz A es invertible si, y sólo si, $det(A) \neq 0$.

Regla de Cramer
Consideremos un sistema $Ax^t = b^t$ con n incógnitas, y la matriz $A = (a_{ij})$ de rango n. Entonces:

$$x^t = (A^{-1}A)x^t = A^{-1}(Ax^t) = A^{-1}b^t = \frac{Adj(A) \cdot b^t}{det(A)}$$

Para calcular el valor de la incógnita j-ésima:

$$x_j = \frac{(\alpha_{1j}, \ldots, \alpha_{nj})\,(b_1, \ldots .b_n)^t}{det(A)} = \frac{\alpha_{1j}b_1 + \cdots + \alpha_{nj}b_n}{det(A)}$$

Para calcular la incógnita j-ésima se siguen los pasos siguentes:

1. Se calcula el determinante de la matriz del sistema.
2. Se calcula el determinante de la matriz después de reemplazar la columna j-ésima por la de términos independientes.
3. Se divide el resultado del punto 2 por el del punto 1.

Nota: La regla de Cramer sirve para resolver cualquier sistema compatible.

UD 2

Espacios vectoriales

Espacios vectoriales de tipo finito

ESPACIOS VECTORIALES

Definición
Un espacio vectorial sobre $\mathbb{K}$ es un conjunto E no vacío cuyos elementos se denominan vectores, dotado de dos operaciones:

- Suma: $+: E \times E \to E$
- Producto por escalares: $\cdot: E \times E \to E$

que cumplen:

1. Suma
   - Conmutativa:
   $$u + v = v + u \quad \forall\, u, v \in E$$
   - Asociativa:
   $$u + (v + w) = (u + v) + w \quad \forall\, u, v, w \in E$$
   - Existencia de elemento neutro:
   $$\exists 0 \in E \;\; / \;\; u + 0 = 0 + u = u \quad \forall\, u \in E$$
   - Existencia de elemento opuesto:
   $$\forall\, u \in E \quad \exists -u \in E \;\; / \;\; u + (-u) = 0$$
2. Producto por escalares:
   - Asociativa:
   $$\lambda\,(\mu u) = (\lambda\mu)\,u \quad \forall\, u \in E \;\; \forall\, \lambda, \mu \in \mathbb{K}$$
   - Existencia de elemento unicidad:
   $$\exists 1 \in \mathbb{K} \;\; / \;\; 1 \cdot u = u \quad \forall\, u \in E$$
   - Distributiva respecto a la suma de escalares:
   $$(\lambda + \mu)\,u = \lambda u + \mu u \quad \forall\, u \in E \;\; \forall\, \lambda, \mu \in \mathbb{K}$$
3. Distributiva respecto a la suma de vectores:
   $$\lambda\,(u + v) = \lambda u + \lambda v \quad \forall\, u, v \in E \;\; \forall\, \lambda \in \mathbb{K}$$

Proposición

Sea E un espacio vectorial sobre $\mathbb{K}$. Sean $u, v \in E$ y $\lambda, \mu \in \mathbb{K}$
Se cumplen:

1. El elemento neutro 0 es único
2. $\lambda\mu = 0 \iff \lambda = 0$ ó $\mu = 0$
3. $-u = (-1) \cdot u$: unicidad del opuesto
4. Si $\lambda u = \mu u,\; u \neq 0 \implies \lambda = \mu$
5. Si $\lambda u = \lambda v,\; \lambda \neq 0 \implies u = v$

Demostración:

1. El elemento 0 es único.
   Suponemos que $\exists\, 0'$ otro elemento neutro. Entonces:
   Como $0 + 0' = 0'$ por ser 0 neutro y $0 + 0' = 0$ por ser $0'$ neutro, entonces tenemos que $0 = 0'$
2. $\lambda\mu = 0 \iff \lambda = 0$ ó $\mu = 0$

   $(\Longleftarrow)$

   Si $\lambda = 0$ tenemos $v = \lambda u$

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

   si restamos v en ambos miembros:

   $$v - v = v + v - v \implies 0 = v + 0 = v$$

   Si $u = 0$ tenemos $v = \lambda u$

   $$v = \lambda \cdot 0 = \lambda\,(0+0) = \lambda \cdot 0 + \lambda \cdot 0 = v + v$$

   si sumamos $-v$ de la misma manera que antes:

   $$0 = v$$

   $(\Longrightarrow)$

   Si $\lambda u = 0$ y $\lambda \neq 0$, podemos dividir por λ :

   $$u = 1 \cdot u = \left(\frac{1}{\lambda} \cdot \lambda\right) u = \frac{1}{\lambda} \cdot (\lambda u) = \frac{1}{\lambda} \cdot 0 = 0$$

   Si $u \neq 0$:

   $$\lambda = 1 \cdot \lambda = \left(\frac{1}{u} \cdot u\right) \lambda = \frac{1}{u} \cdot (\lambda u) = \frac{1}{u} \cdot 0 = 0$$
3. $-u = (-1) \cdot u$: unicidad del opuesto

   $$0 = (1 + (-1)) \cdot u = 1 \cdot u + (-1) \cdot u = u + (-1) \cdot u$$

   si sumamos $-u$ en ambos lados:

   $$0 - u = u + (-1) \cdot u - u \implies -u = (-1) \cdot u$$
4. Si $\lambda u = \mu u,\ u \neq 0 \implies \lambda = \mu$

   $$0 = \mu u + (-\mu u) = \lambda u + (-\mu u) = (\lambda - \mu) \cdot u$$

   como $u \neq 0 \implies \lambda - \mu = 0 \implies \lambda = \mu$
5. Si $\lambda u = \lambda v,\ \lambda \neq 0 \implies u = v$

   $$0 = \lambda u + (-\lambda u) = \lambda v + (-\lambda u) = \lambda \cdot (v - u)$$

   como $\lambda \neq 0 \implies v - u = 0 \implies u = v$

Definición
El producto cartesiano de dos espacios vectoriales define una estructura de espacio vectorial sobre $\mathbb{K}$:

$$\mathbb{K}^n = \{ u = (x_1, \ldots, x_n) : \; x_i \in \mathbb{K} \;\; \forall\, i = 1, \ldots, n \}$$

Definición. Subespacio vectorial
Sea E un espacio vectorial.

Las operaciones de E restringidas a $V \subset E$ definen un nuevo espacio vectorial, en el que se cumple:

$$\lambda u + \mu v \in V \;\;\; \forall\, \lambda, \mu \in \mathbb{K} \;\;\; \forall\, u, v \in V$$

Definición
Un vector $u \in E$ depende linealmente de otros vectores $u_1, \ldots, u_r \in E$ si

$$\exists\, \lambda_1, \ldots, \lambda_r \in \mathbb{K} \;\; / \;\; u = \lambda_1 u_1 + \cdots + \lambda_r u_r$$

Diremos que los vectores $u_1, \ldots, u_r \in E$ son linealmente independientes si ninguno depende de los demás.

Proposición
Los vectores $u_1, \ldots, u_r \in E$ son linealmente independientes si, y sólo si, de cualquier combinación lineal

$$\lambda_1 u_1 + \cdots + \lambda_r u_r = 0$$

obtenemos la solución trivial: $\lambda_1 = \cdots = \lambda_r = 0$

ESPACIOS VECTORIALES DE TIPO FINITO

Definición
Un espacio vectorial E se denomina de tipo finito si:

$$\exists\, u_1, \dots, u_q \in E \ / \ \forall\, u \in E, \ \ u = \lambda_1 u_1 + \cdots + \lambda_q u_q$$

Todo vector es combinación lineal de una colección finita de vectores.

Diremos que $E = L\,[u_1, \dots, u_q]$ está generado por los u_j, es decir, $u_1, \dots, u_q$ forman un sistema de generadores.

$\mathbb{K}^n$ es un espacio de tipo finito:

$$(x_1, \dots, x_n) = \sum_{i=1}^{n} x_i e_i, \quad e_i = \left(0, \dots, 0, \overset{(j)}{1}, 0, \dots, 0\right)$$

de donde tenemos que $\mathbb{K}^n = L\,[e_1, \dots, e_n]$

Sistemas generadores de $\mathbb{K}^n$

Sabemos que $\mathbb{K}^n$ es de tipo finito, pero como reconocer una colección de vectores que genera el espacio:

$$\mathbb{K}^n = L\,[u_1, \dots, u_q]$$

Supongamos $u_j = \left(a_{1_j}, \dots, a_{n_j}\right) \in \mathbb{K}^n \ \ 1 \le j \le q$
Veamos si $u = (b_1, \dots, b_n)$ es combinación lineal de los vectores u_j:

$$u = \lambda_1 u_1 + \cdots + \lambda_q u_q$$

Si calculamos componente a componente:

$$\begin{cases} b_1 = \lambda_1 a_{11} + \cdots + \lambda_q a_{1q} \\ \vdots \\ b_n = \lambda_1 a_{n1} + \cdots + \lambda_q a_{nq} \end{cases}$$

de donde tenemos que u es combinación lineal de u_j si, y sólo si, el sistema tiene solución.

$$\mathbb{K}^n = L\,[u_1, \dots, u_q] \iff rg(A) = rg(A|b)$$

Como A tiene n filas podemos concluir:

1. Los u_j generan $\mathbb{K}^n \iff rg(A) = n$
2. El número máximo de generadores de $\mathbb{K}^n$ es n
3. n generadores d e$\mathbb{K}^n$ son siempre independientes, y n vectores independientes necesariamente generan $\mathbb{K}^n$

En conclusión, n vectores son independientes si, y sólo si, el rango de la matriz correspondiente es n, si, y sólo si, los vectores generan $\mathbb{K}^n$.

Definición
Denominamos base de un espacio vectorial a un sistema ordenado de generadores formado por vectores independientes.

Los vectores $e_1, \dots, e_n$ son una base en $\mathbb{K}^n$ denominada base estándar $\mathcal{E}_n$ o $\mathcal{E}$.

Definición
Supongamos que $\mathcal{B} = \{u_1, \ldots, u_n\}$ es una base de E.
Denominamos coordenadas de u respecto a $\mathcal{B}$ a los coeficientes $(x_1, \ldots, x_n)$ tales que:

$$u = x_1 u_1 + \cdots + x_n u_n, \quad x_1, \ldots, x_n \in \mathbb{K}$$

Proposición
Todas las bases de un espacio vectorial E de tipo finito tienen el mismo número de vectores.
Denominaremos a este número dimensión de E.

Nota:

$$dim(E) = 1 \;\; E \text{ es una recta}$$
$$dim(E) = 2 \;\; E \text{ es un plano}$$

Matriz de cambio de base

Sea E un espacio vectorial de tipo finito. Sean $\mathcal{B} = \{u_1, \ldots, u_n\}$, $\mathcal{B}' = \{v_1, \ldots, v_n\}$ dos bases de E.
Dado un vector $u \in E$ denotamos sus coordenadas:

$$(x_1, \ldots, x_n) \text{ respecto a } \mathcal{B}$$

$$(y_1, \ldots, y_n) \text{ respecto a } \mathcal{B}'$$

Entonces:

1. Existe una única matriz C tal que

   $$\forall\, u \in E \implies y^t = Cx^t$$

   ¿ C es única ?

   Veamos que las coordenadas respecto a la base $\mathcal{B}$ del vector u_j son $e_j = (0, \ldots, 0, \overset{(j)}{1}, 0, \ldots, 0)$; por tanto sus coordenadas respecto a la base $\mathcal{B}'$ se obtienen calculando el producto:

   $$C \cdot e_j^t$$

   que es la columna j-ésima de C, en consecuencia tenemos que C está completamente determinada por la condición deseada.

2. Si $\mathcal{B} = \mathcal{B}'$ entonces $C = I_n$

   Si $\mathcal{B}''$ es otra base tal que $C' = C(\mathcal{B}', \mathcal{B}'')$ entonces $C'C = C(\mathcal{B}, \mathcal{B}'')$.

   Si denotamos z las coordenadas respecto a $\mathcal{B}''$:

   $$z^t = C'y^t = C'\left(Cx^t\right) = (C'C)\, x^t \implies C'C = C(\mathcal{B}, \mathcal{B}'')$$

3. Si $\mathcal{B}'' = \mathcal{B} \implies C'C = I_n$, es decir, la matriz de cambio de base es invertible y su inversa es la matriz de cambio de base en orden contrario.

Proposición. Teorema de prolongación
Sea $E = L[u_1, \ldots, u_q]$ un espacio vectorial de tipo finito con $dim(E) = n$.
Si los vectores $v_1, \ldots, v_r$ son independientes, entonces $\exists\, u_{i_1}, \ldots, u_{i_{n-r}}$ de forma que:

$$v_1, \ldots, v_r, u_{i_1}, \ldots, u_{i_{n-r}} \text{ son una base}$$

Proposición
Todo subespacio V de un espacio finito E es de tipo finito.
Además,

$$dim(V) \leq dim(E)$$

Denominamos codimensión a la diferencia de dimensiones:

$$codim(V) = dim(E) - dim(V)$$

Un subespacio de codimensión 1 se denomina hiperplano.

Obtención de las ecuaciones implícitas y parámetricas
Sea $\mathcal{B} = \{u_1, u_2, u_3, u_4, u_5\}$ una base, consideremos los vectores:

$$\begin{aligned} v_1 &= u_1 + u_3 + 2u_4 \\ v_2 &= 2u_1 + u_2 + 3u_3 + 4u_4 \\ v_3 &= u_1 - u_2 + 2u_4 \end{aligned}$$

Extraemos la matriz $A = \begin{pmatrix} 1 & 2 & 1 \\ 0 & 1 & -1 \\ 1 & 3 & 0 \\ 2 & 4 & 2 \end{pmatrix}$ y estudiamos su rango:

$$\begin{pmatrix} 1 & 2 & 1 \\ 0 & 1 & -1 \\ 1 & 3 & 0 \\ 2 & 4 & 2 \end{pmatrix} \begin{matrix} \\ \\ f_3 - f_1 \\ f_4 - 2f_1 \end{matrix} \Longrightarrow \begin{pmatrix} 1 & 2 & 1 \\ 0 & 1 & -1 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{pmatrix} \; f_3 + f_2 \Longrightarrow \begin{pmatrix} 1 & 2 & 1 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \Longrightarrow rg(A) = 2$$

Las dos primeras columnas son independientes, de manera que:

$$rg\begin{pmatrix} 1 & 2 & x_1 \\ 0 & 1 & x_2 \\ 1 & 3 & x_3 \\ 2 & 4 & x_4 \end{pmatrix} = 2 \Longrightarrow det\begin{pmatrix} 1 & 2 & x_1 \\ 0 & 1 & x_2 \\ 1 & 3 & x_3 \end{pmatrix} = 0 \text{ y } det\begin{pmatrix} 1 & 2 & x_1 \\ 0 & 1 & x_2 \\ 2 & 4 & x_4 \end{pmatrix} = 0$$

Si desarrollamos los determinantes:

$$\begin{cases} x_3 + 2x_2 - x_1 - 3x_2 = 0 \\ x_4 + 4x_2 - 2x_1 - 4x_2 = 0 \end{cases} \Longrightarrow \begin{cases} -x_1 - x_2 + x_3 = 0 \\ -2x_1 + x_4 = 0 \end{cases} \Longrightarrow \begin{cases} x_1 + x_2 - x_3 = 0 \\ 2x_1 - x_4 = 0 \end{cases}$$

serían las ecuaciones implícitas

Para las paramétricas:

Denominamos: $\begin{cases} x_4 = \lambda \\ x_3 = \mu \end{cases} \Longrightarrow \begin{cases} x_1 + x_2 - \mu = 0 \\ 2x_1 - \lambda = 0 \end{cases} \Longrightarrow 2x_1 = \lambda \Longrightarrow x_1 = \frac{\lambda}{2}$

$\frac{\lambda}{2} + x_2 - \mu = 0 \Longrightarrow x_2 = \mu - \frac{\lambda}{2}$

Por tanto, las ecuaciones paramétricas son:

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

y observando los coeficientes de los parámetros, tenemos que:
$\left\{\left(\frac{1}{2}, \frac{-1}{2}, 0, 1\right), (0, 1, 1, 0)\right\}$ forman una base.

UD 3

Operaciones con subespacios.

Aplicaciones lineales.

OPERACIONES CON SUBESPACIOS

Intersecciones de subespacios
Sea $\{V_i\}_i$ una colección arbitraria de subespacios de E. Entonces su intersección

$$V = \bigcap_i V_i$$

es un subespacio vectorial de E.

Nota: Para calcular la intersección de dos subespacios formamos un sistema con las ecuaciones implícitas de los subespacios.

Generación de subespacios
Sea $X \subset E$. Consideremos la colección $\{V_i\}$ de todos los subespacios que contienen a X. Entonces

$$V = \bigcap_i V_i$$

es el menor subespacio que contiene a X y coincide con $L[X]$.

Suma de subespacios
Sean $V_1, \ldots, V_p$ subespacios vectoriales de E. El subespaico generado por su unión, $X = V_1 \cup \cdots \cup V_p$, se denomina suma y se denota:

$$L[X] = V_1 + \cdots + V_p = \sum_{i=1}^{p} V_p$$

Cociente módulo subespacio
Sea L un subespacio vectorial de E. Definimos en E la relación de equivalencia:

$$v\mathcal{R}u \iff v - w = u \in L$$

La clase de equivalencia de un vector es:

$$[v] = v + L = \{\, v + u : \; u \in L \,\}$$

Denominamos cociente módulo L al espacio vectorial E/L sobre $\mathbb{K}$.

Proposición
Cualquier subespacio de E es intersección de hiperplanos.
El número mínimo necesario es la codimensión del subespacio.

Proposición. Fórmula de Grassmann
Sean V, W dos subespacios de E. Entonces:

$$dim(V+W) = dim(V) + dim(W) - dim(V \cap W)$$

Demostración:
Suponemos que $V, W \neq \{0\}$ y que $V \cap W \neq \{0\}$, ya que en este caso la fórmula se cumple trivialmente.

Sea $\{u_1, \dots, u_d\}$ una base de $V \cap W$. Si le añadimos los vectores $v_1, \dots, v_p \in V$ obtenemos una base de V; y añadiendo los vectores $w_1, \dots, w_q \in W$ obtenemos una base de W. Así tenemos que:

$$\begin{aligned} dim(V \cap W) &= d \\ dim(V) &= d + p \\ dim(W) &= d + q \end{aligned}$$

Por tanto debemos probar que: $dim(V+W) = d + p + q$

Veamos si $\{u_1, \dots, u_d, v_1, \dots, v_p, w_1, \dots, w_q\}$ es una base de $V + W$.

Está claro que esta colección de vectores genera todo el espacio $V + W$ como consecuencia de que la suma consiste en la suma de vectores. Así sólo nos queda comprobar que sean independientes:

$$\sum_j \alpha_j u_j + \sum_k \lambda_k v_k + \sum_l \mu_l w_l = 0$$

$$\sum_j \alpha_j u_j + \sum_k \lambda_k v_k = -\sum_l \mu_l w_l \in V \cap W$$

$$\sum_j \alpha_j u_j + \sum_k \lambda_k v_k = \sum_j \beta_j u_j$$

Los u_j, v_k son independientes, por tanto $\lambda_k = 0$

$$\sum_j \alpha_j u_j + \sum_l \mu_l w_l = 0$$

Los u_j, w_l son independientes, por tanto $\alpha_j = \mu_l = 0$.

Así pues los vectores son linealmente independientes y generan el espacio $V + W$, por lo que tenemos que son una base y:

$$dim(V+W) = dim(V) + dim(W) - dim(V \cap W)$$

Proposición
Sea L un subespacio de E. Entonces E/L es de tipo finito y se cumple que:

$$dim(E/L) = dim(E) - dim(L) = codim(L)$$

APLICACIONES LINEALES

Definición
Sean E, F dos espacios vectoriales sobre un mismo cuerpo $\mathbb{K}$. Una aplicación $f: E \to F$ se denomina lineal cuando conserva las combinaciones lineales:

$$f(\lambda_1 u_1 + \cdots + \lambda_r u_r) = \lambda_1 f(u_1) + \cdots + \lambda_r f(u_r), \quad \lambda_i \in \mathbb{K},\ u_i \in E$$

Proposición
Sea $f: E \to F$ una aplicación lineal. Se cumple que:

1. $f(0) = 0$
2. $f(-u) = -f(u) \ \ \forall u \in E$

Ejemplo: Son aplicaciones lineales:

- Aplicación identidad: $Id_E: E \to E$
- Proyección: $\begin{array}{rl} \pi = \pi_L : & E \to E/L \\ & v \mapsto [v] \end{array}$
- Homotecia vectorial de razón λ: $\begin{array}{rl} h: & E \to E \\ & v \mapsto \lambda v \end{array}$

Definición
El conjunto de todas las aplicaciones lineales de un espacio vectorial E en otro F se denota por:

$$\mathcal{L}(E, F)$$

Nota: Las aplicaciones lineales respetan los subespacios vectoriales.

Supongamos una aplicación lineal $f: \mathbb{K}^3 \to \mathbb{K}^3$ que cumple:

$$\begin{array}{l} f(2,1,1) = (3,0,1) \\ f(1,0,-1) = (1,1,-1) \\ f(0,0,1) = (-1,2,0) \end{array}$$

Los tres vectores generan todo el subespacio, por tanto cualquier vector es combinación lineal de ellos:

$$(x, y, z) = \alpha(2,1,1) + \beta(1,0,-1) + \gamma(0,0,1)$$

$$\begin{cases} x = 2\alpha + \beta \\ y = \alpha \\ z = \alpha - \beta + \gamma \end{cases} \Longrightarrow \begin{cases} \alpha = y \\ \beta = x - 2\alpha = x - 2y \\ \gamma = z - \alpha + \beta = z - y + x - 2y = x - 3y + z \end{cases}$$

Por tanto:
$$\begin{aligned} f(x,y,z) &= \alpha(3,0,1) + \beta(1,1,-1) + \gamma(-1,2,0) = \\ &= y(3,0,1) + (x-2y)(1,1,-1) + (x-3y+z)(-1,2,0) = \\ &= (3y + x - 2y - x + 3y - z, x - 2y + 2x - 6y + 2z, y - x + 2y) = \\ &= (4y - z, 3x - 8y + 2z, -x + 3y) \end{aligned}$$

Proposición
Sea $f: E \to F$ una aplicaicón lineal. Sean $V \subset E$, $W \subset F$ subespacios vectoriales.
Entonces,

1. La imagen $f(V)$ es un subespacio vectorial de F:
$$f(V) = \{\, f(v) \in F: \ v \in V \,\}$$
2. La imagen inversa $f^{-1}(W)$ es un subespacio vectorial de E:
$$f^{-1}(W) = \{\, u \in E: \ f(u) \in W\}$$

Definición. Imagen y núcleo
Sea $f: E \to F$ una aplicación lineal.
Diremos que su imagen es:
$$im(f) = f(E) = \{\, f(u) \in F: \ u \in E \,\}$$
y su núcleo:
$$ker(f) = f^{-1}(0) = \{\, u \in E: \ f(u) = 0 \,\}$$

Proposición
Sea $f: E \to F$ una aplicación lineal.
Las siguientes afirmaciones son equivalentes:

(i) f conserva la independencia lineal: si $u_1, \ldots, u_p$ son independientes en E, entonces sus imégenes $f(u_1), \ldots, f(u_p)$ son independientes en F.

(ii) El núcleo de f es trivial: $ker(f) = \{0\}$

(iii) f es inyectiva

Demostración:

$(i) \Longrightarrow (ii)$
Para un sólo vector, ser independiente es ser no nulo.

$(ii) \Longrightarrow (iii)$
Si $f(u) = f(v) \implies f(u-v) = f(u) - f(v) = 0$
de donde $u - v \in ker(f) \implies u - v = 0 \implies u = v$ y por tanto f es inyectiva.

$(iii) \Longrightarrow (i)$
Sean $u_1, \ldots, u_p \in E$ vectores independientes.
Sea la combinación lineal: $\lambda_1 f(u_1) + \cdots + \lambda_p f(u_p) = 0$
Por linealidad:
$$\lambda_1 f(u_1) + \cdots + \lambda_p f(u_p) = f(\lambda_1 u_1 + \cdots + \lambda_p u_p)$$
Como $f(0) = 0$ y f es inyectiva:
$$\lambda_1 u_1 + \cdots + \lambda_p u_p = 0$$
Los u_j son independientes, luego todos los $\lambda_j = 0$, por tanto las imágenes son independientes.

Proposición
Sea $f : E \to F$ una aplicación lineal. Si f es biyectiva, entonces su inversa f^{-1} es también lineal.

En este caso diremos que f es un isomorfismo lineal y que E y F son isomorfos.

Nota: La composición de aplicaciones lineales es lineal.

Proposición
Sea $f : E \to F$ una aplicación lineal. Consideremos la proyección $\pi : E \to E/ker(f)$ y la inclusión $j : im(f) \hookrightarrow F$.
Entonces,

$$\overline{f} : E/ker(f) \to im(f) : [v] \mapsto f(v)$$

es la única aplicación lineal tal que:

$$f = j \circ \overline{f} \circ \pi$$

f tiene la factorización canónica:

$$\begin{array}{ccccccc} E & \xrightarrow{f} & F & & v & \xrightarrow{f} & f(v) \\ \downarrow \pi & & \uparrow j & : & \downarrow \pi & & \uparrow j \\ E/ker(f) & \xrightarrow{\overline{f}} & im(f) & & [v] & \xrightarrow{\overline{f}} & \overline{f}([v]) = f(v) \end{array}$$

donde π es suprayectiva, $\overline{f}$ es biyectiva y j es inyectiva.

De ahí obtenemos el primer teorema de isomorfía:

$$E/ker(f) \text{ e } im(f) \text{ son isomorfos}$$

Definición. Proyecciones y simetrías
Sea E un espacio vectorial. Sean V, W dos subespacios suplementarios: $E = V \oplus W$.
Cada vector $u \in E$ se escribe de una única manera como suma

$$u = v + w, \;\; v \in V, \;\; w \in W$$

De esta manera podemos definir las aplicaciones:

(i) La proyección $p : E \to E : u \mapsto v$, denominada de base V y dirección W.

(ii) La simetría $s : E \to E : u \mapsto v - w$, denominada de base V y dirección W.

Si $p : E \to E$ es una aplicación lineal, entonces:

$$im(p) = V \quad ker(p) = W$$

Si $s : E \to E$ es una aplicación lineal, entonces:

$$V = ker\,(Id_E - s) \quad W = ker\,(Id_E + s)$$

de donde:

$$im(p) = ker\,(Id_E - s) \quad ker(p) = ker\,(Id_E + s)$$

Proposición
Sea $f : E \to F$ una aplicación lineal, donde E, F son de tipo finito.
Entonces,

$$dim\left(E\right) = dim\left(ker\left(f\right)\right) + dim\left(im\left(f\right)\right)$$

Propiedades

1. Si f es inyectiva $\Longrightarrow$ $dim\left(E\right) \leq dim\left(F\right)$
2. Si f es suprayectiva $\Longrightarrow$ $dim\left(E\right) \geq dim\left(F\right)$
3. Si $dim\left(E\right) = dim\left(F\right) \iff f$ es biyectiva

Proposición
Sea E un espacio vectorial de tipo finito. Consideramos una base de E, $\mathcal{B} = \{u_1, \ldots, u_n\}$.
Sea F un espacio vectorial no necesariamente finito tal que $v_1, \ldots, v_n \in F$.
Entonces existe una única aplicación lineal $f : E \to F$ tal que:

$$f\left(u_1\right) = v_1, \ldots . f\left(u_n\right) = v_n$$

Demostración:
Si existe es única está claro ya que:

$$E = L\left[u_1, \ldots, u_n\right]$$

Construyamos f:
Cualquier vector $u \in E$ es de la forma:

$$u = x_1 u_1 + \cdots + x_n u_n$$

entonces definimos:

$$f(u) = x_1 f\left(u_1\right) + \cdots + x_n f\left(u_n\right) = x_1 v_1 + \cdots + x_n v_n$$

Ésto define bien la aplicación $f : E \to F$ ya que los x_j están detrerminados por u.
Veamos si es aplicación lineal:
Sean $u, w \in E$ con coordenadas $\left(x_j\right), \left(y_j\right)$ respectivamente, y con $\lambda, \mu \in \mathbb{K}$. Las coordenadas de $\lambda u + \mu w$ son $\left(\lambda x_j + \mu y_j\right)$, entonces:

$$f\left(\lambda u + \mu w\right) = \sum_j \left(\lambda x_j + \mu y_j\right) v_j = \sum_j \lambda x_j v_j + \sum_j \mu y_j v_j = \lambda f\left(u\right) + \mu f\left(w\right)$$

Por tanto, f es una aplicación lineal.

Definición. Ecuaciones y matriz de una aplicación lineal
Sea $f : E \to F$ una aplicación lineal entre espacios vectoriales de dimensión finita.
Sean las bases $\mathcal{B}_E = \{u_1, \ldots, u_n\}, \mathcal{B}_F = \{v_1, \ldots, v_m\}$.
Ddado un vector $u \in E$ denotamos sus coordenadas por $(x_1, \ldots, x_n)$ respecto de $\mathcal{B}_E$ y las coordenadas de $f(u)$ por $(y_1, \ldots, y_m)$ respecto de $\mathcal{B}_F$.
Entonces existe una única matriz M tal que $\forall\, u \in E$:

$$y^t = M x^t$$

M recibe el nombre de matriz de$f$ respecto a las bases $\mathcal{B}_E, \mathcal{B}_F$ y se denota por:

$$M = M_f\left(\mathcal{B}_E, \mathcal{B}_F\right)$$

Denominamos ecuaciones de f respecto a las bases dadas, a las igualdades:

$$y_i = \sum_{j=1}^{n} a_{ij} x_j$$

Si variamos las bases:

$$M_f\left(\mathcal{B}'_E, \mathcal{B}'_F\right) = C\left(\mathcal{B}_F, \mathcal{B}'_F\right) M_f\left(\mathcal{B}_E, \mathcal{B}_F\right) C\left(\mathcal{B}'_E, \mathcal{B}_E\right)$$

Sean las aplicaciones lineales $f : E \to F$, $g : F \to G$, y las bases $\mathcal{B}_E$, $\mathcal{B}_F$, $\mathcal{B}_G$. La matriz de la composición es:

$$M_{g \circ f}\left(\mathcal{B}_E, \mathcal{B}_G\right) = M_g\left(\mathcal{B}_F, \mathcal{B}_G\right) M_f\left(\mathcal{B}_E, \mathcal{B}_F\right)$$

Ejemplo: $f : \mathbb{K}^3 \to \mathbb{K}^3$

$$\begin{cases} f(2,1,1) = (3,0,1) \\ f(1,0,-1) = (1,1,-1) \\ f(0,0,1) = (-1,2,0) \end{cases}$$

Los vectores $\{(2,1,1), (1,0,-1), (0,0,1)\}$ forman una base $\mathcal{B}$:

$$M_f(\mathcal{B}, \mathcal{E}) = \begin{pmatrix} 3 & 1 & -1 \\ 0 & 1 & 2 \\ 1 & -1 & 0 \end{pmatrix}$$

Para hallar las ecuaciones de f respecto a la base estándar:

$$M_f(\mathcal{E}, \mathcal{E}) = C(\mathcal{E}, \mathcal{E})\, M_f(\mathcal{B}, \mathcal{E})\, C(\mathcal{E}, \mathcal{B})$$

donde $C(\mathcal{E}, \mathcal{B}) = C(\mathcal{B}, \mathcal{E})^{-1} = \begin{pmatrix} 2 & 1 & 0 \\ 1 & 0 & 0 \\ 1 & -1 & 1 \end{pmatrix}^{-1}$

Por tanto:

$$M_f(\mathcal{E}, \mathcal{E}) = M_f(\mathcal{B}, \mathcal{E})\, C(\mathcal{E}, \mathcal{B}) = M_f(\mathcal{B}, \mathcal{E})\, C(\mathcal{B}, \mathcal{E})^{-1} = \begin{pmatrix} 3 & 1 & -1 \\ 0 & 1 & 2 \\ 1 & -1 & 0 \end{pmatrix} \begin{pmatrix} 2 & 1 & 0 \\ 1 & 0 & 0 \\ 1 & -1 & 1 \end{pmatrix}^{-1} =$$

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

$$= \begin{pmatrix} 3 & 1 & -1 \\ 0 & 1 & 2 \\ 1 & -1 & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 & 0 \\ 1 & -2 & 0 \\ 1 & -3 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 4 & -1 \\ 3 & -8 & 2 \\ -1 & 3 & 0 \end{pmatrix}$$

de donde las ecuaciones de f respecto a la base estándar son:

$$\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} 0 & 4 & -1 \\ 3 & -8 & 2 \\ -1 & 3 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

Rango de una aplicación lineal

Sea $f : E \rightarrow F$ una aplicación lineal entre espacios de tipo finito.
Sean las bases $\mathcal{B}_E$, $\mathcal{B}_F$ cuyos vectores tienen de coordenadas x, y respectivamente.

Consideremos la matriz $M = M_f\left(\mathcal{B}_E, \mathcal{B}_F\right)$

Sabemos que la imagen está generada por las imágenes de los vectores de $\mathcal{B}_E$, como las columnas de M son las coordenadas de esas imágenes respecto a $\mathcal{B}_F$:

$$rg\left(f\right) = rg\left(M\right) = dim\left(im\left(f\right)\right)$$

$$rg\left(M\right) = codim\left(ker\left(f\right)\right) = dim\left(E\right) - dim\left(ker\left(f\right)\right)$$

Según el rango podemos saber que tipo de aplicación lineal es:

(i) f es inyectiva $\iff rg\left(f\right) = dim\left(E\right)$

(ii) f es suprayectiva $\iff rg\left(f\right) = dim\left(F\right)$

(iii) f es biyectiva $\iff rg\left(f\right) = dim\left(E\right) = dim\left(F\right)$