Álgebra lineal

Bibliografía:

- Material clases teóricas, Álgebra Lineal, Universidad de Murcia (anónimo).

Capítulo 1

Espacios vectoriales

1.1. Cuerpos

Conjunto K con dos operaciones, **suma** $(+)$ y **producto** $(\cdot)$, tales que $\forall a, b, c \in K$:

1. **Propiedad asociativa de la suma:** $a + (b + c) = (a + b) + c$.
2. **Propiedad conmutativa de la suma:** $a + b = b + a$.
3. **Elemento neutro para la suma** o **nulo:** $\exists! 0 \in K : \forall a \in K, 0 + a = a$.
   Pongamos que existe otro 0 $(0')$, entonces $0 = 0 + 0' = 0'$.
4. **Inverso para la suma** u **opuesto:** $\forall a \in K, \exists! a' : a + a' = 0$. $-a := a'$.
   Pongamos que existe otro opuesto a'', entonces $a' = 0 + a' = (a'' + a) + a' = a'' + (a + a') = a'' + 0 = a''$.
5. **Propiedad asociativa del producto:** $a \cdot (b \cdot c) = (a \cdot b) \cdot c$.
6. **Propiedad conmutativa del producto:** $a \cdot b = b \cdot a$.
7. **Elemento neutro para el producto** o **unidad:** $\exists! 1 \in K : \forall a \in K, 1 \cdot a = a$.
8. **Inverso para el producto:** $\forall a \in K \backslash \{0\}, \exists! a'' : a \cdot a'' = 1$; $a^{-1} := \frac{1}{a} := a''$.
9. **Propiedad distributiva:** $a \cdot (b + c) = a \cdot b + a \cdot c$.

La congruencia $\mathbb{Z}_2 = \{0, 1\}$ con operaciones $0 + 0 = 1 + 1 = 0$, $0 + 1 = 1 + 0 = 1$, $0 \cdot 0 = 0 \cdot 1 = 1 \cdot 0 = 0$ y $1 \cdot 1 = 1$ es un cuerpo. Siempre existe un cuerpo F_{p^n}, formado por p^n elementos, donde $n \in \mathbb{N}$ y p es primo. Algunas propiedades: $\forall a, b, c \in K$:

1. $a + b = a + c \implies b = c$.
2. $0a = 0$.
3. $(-a)b = a(-b) = -(ab)$; $(-1)a = -a$; $(-a)(-b) = ab$.
4. $ab = 0 \implies a = 0 \vee b = 0$.
5. $ab = ac \implies a = 0 \vee b = c$.

1.1.1. El cuerpo de los números complejos

Si consideramos $\mathbb{R}\times\mathbb{R}=\{(a,b)|a,b\in\mathbb{R}\}$, con $(a,b)+(c,d)=(a+c,b+d)$ y $(a,b)\cdot(c,d)=(ac-bd,ad+bc)$, obtenemos el cuerpo de los números complejos ($\mathbb{C}$). El conjunto de elementos de $\mathbb{C}$ con forma $(a,0)$ es una copia del cuerpo $\mathbb{R}$. Llamamos **unidad imaginaria** a $i=(0,1)$, de forma que $i^2=i\cdot i=(-1,0)=-1$. Dado que $(b,0)i=(0,b)$, y como $(b,0)$ es el número real b, tenemos $(a,b)=a+bi$, lo que denominamos la **forma binomial.** a es la **componente real,** y b la **componente imaginaria.** Si $z=a+bi$, llamamos **conjugado** de z a $\overline{z}=a-bi$, de forma que $z=\overline{z}\iff z\in\mathbb{R}$.

Podemos representar un número complejo $z=a+bi$ como un punto del plano, con coordenadas (a,b). La distancia del punto al origen de coordenadas, llamada **módulo**, es $r=|z|=\sqrt{a^2+b^2}=\sqrt{z\overline{z}}$. El ángulo con el eje OX, llamado **argumento**, cumple que $a+bi=r(\cos(\alpha)+i\sin(\alpha))$. Esta es la **forma polar** o **módulo argumental** del complejo. La multiplicación en forma polar es: $[r(\cos(\alpha)+i\sin(\alpha))][r'(\cos(\alpha')+i\sin(\alpha'))]=rr'(\cos(\alpha+\alpha')+i\sin(\alpha+\alpha'))$.

El **Teorema Fundamental del Álgebra** nos dice que todo polinomio $P(x)=a_0+a_1X+a_2X^2+\cdots+a_nX^n$, con $n\geq 1$, $a_i\in\mathbb{C}$ y $a_n\neq 0$, tiene raíz compleja.

1.1.2. Característica de un cuerpo.

$$\forall a\in K, n\in\mathbb{N}, na=\underset{n\text{ veces}}{a+a+\cdots+a}$$

En particular,

$$n1=\underset{n\text{ veces}}{1+1+\cdots+1}$$

Por tanto, $na=(n1)a$, $n1+m1=(n+m)1$ y $(n1)(m1)=(nm)1$ para cualquier cuerpo K.

Un cuerpo tiene **característica cero** si $\forall n>0, n1\neq 0$. De lo contrario, se dice que tiene **característica** n, siendo n el menor natural tal que $n1=0$. Así, $na=(n1)a\implies na=0$. Dado que $ab=0\iff a=0\vee b=0$, tenemos que $\exists p,q<n: n=pq\implies 0=n1=(p1)(q1)\implies p1=0\vee q1=0$, por lo que la característica de un cuerpo es cero o un nº primo.

1.2. Espacios vectoriales

Un **espacio vectorial** sobre K, o K-espacio vectorial, es una terna $(V,+,\cdot)$ donde $V\neq\emptyset$, $+$ es una operación $V\times V\longrightarrow V$ y $\cdot$ es una operación $K\times V\longrightarrow V$, tales que $\forall u,v,w\in V,\alpha,\beta\in K$:

1. **Suma asociativa:** $u+(v+w)=(u+v)+w$.
2. **Suma conmutativa:** $u+v=v+u$.
3. **Vector cero** o **nulo:** $\exists 0_V:\forall u\in V, 0_V+u=u$.
4. **Opuesto para la suma:** $\forall u\in V,\exists u': u+u'=0_V;\ u':=-u$.
5. $\alpha\cdot(u+v)=\alpha\cdot u+\alpha\cdot v$.
6. $(\alpha+\beta)\cdot u=\alpha\cdot u+\beta\cdot u$.

7. $(\alpha \cdot \beta) \cdot u = \alpha \cdot (\beta \cdot u)$.

8. $1_K \cdot u = u$.

Llamamos **vectores** a los elementos de V y **escalares** a los de K. Todo cuerpo es espacio vectorial sobre sí mismo. $\mathbb{R}$ es espacio vectorial sobre $\mathbb{Q}$, y $\mathbb{C}$ sobre $\mathbb{R}$ y $\mathbb{Q}$.

El plano real $\mathbb{R}^2 = \{(x, y)|x, y \in \mathbb{R}\}$, con las operaciones $(x, y) + (x', y') = (x + x', y + y')$ y $\alpha(x, y) = (\alpha x, \alpha y)$, es un $\mathbb{R}$-espacio vectorial. El conjunto K^n de las n-uplas de elementos de K, $K^n = \{(x_1, x_2, \ldots, x_n|x_1, x_2, \ldots, x_n \in K\}$ es un K-espacio vectorial con las operaciones $(x_1, \ldots, x_n) + (y_1, \ldots, y_n) = (x_1 + y_1, \ldots, x_n + y_n)$ y $\alpha(x_1, \ldots x_n) = (\alpha x_1, \ldots, \alpha x_n)$. También, si $V_1, \ldots, V_n$ son K-espacios vectoriales, el conjunto $V_1 \times \cdots \times V_n = \{(v_1, v_2, \ldots, v_n)|v_1 \in V_1, v_2 \in V_2, \ldots, v_n \in V_n\}$, con operaciones similares, es un K-espacio vectorial llamado **espacio vectorial producto** de los V_i.

Una **matriz** A de tamaño $m \times n$ (con $m, n \in \mathbb{N}$) sobre K es una disposición en m filas y n columnas de $m \cdot n$ elementos de K. La matriz es **cuadrada** si $m = n$, **fila** si $m = 1$ y **columna** si $n = 1$. Llamamos $M_{m,n}(K)$ al conjunto de las matrices $m \times n$ sobre K, y $M_{n,n}(K) = M_n(K)$. Se representan de la siguiente forma:

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

$\forall A = (a_{ij}), B = (b_{ij}) \in M_{m,n}(K), A + B = (c_{ij}) \in M_{m,n}(K)$, donde $c_{ij} = a_{ij} + b_{ij}$ para cada elemento de la matriz. También, $\forall \alpha \in K, A \in M_{m,n}(K), \alpha A = (c_{ij}) \in M_{m,n}(K)$, donde $c_{ij} = \alpha a_{ij}$. Con estas operaciones, $M_{m,n}(K)$ es un K-espacio vectorial.

Un **polinomio** en una **indeterminada** X con **coeficientes** en K es una expresión de la forma

$$a_0 + a_1 X + a_2 X^2 + \cdots + a_n X^n$$

Donde n es un entero no negativo, $a_i \in K \forall i = 0, 1, \ldots n$. El conjunto de polinomios sobre K se llama $K(X)$ y es un espacio vectorial con las operaciones:

$$\begin{aligned} (a_0 + \cdots + a_n X^n) + (b_0 + \cdots + b_n X^n) &= (a_0 + b_0) + \cdots + (a_n + b_n) X^n \\ \alpha(a_0 + a_1 X + \cdots + a_n X^n) &= \alpha a_0 + \alpha a_1 X + \cdots + \alpha a_n X^n \end{aligned}$$

Si $\mathcal{S} \neq \emptyset$, el conjunto $\mathcal{F}(\mathcal{S}, K) = \{f \mid \mathcal{S} \to K\}$, formado por todas las aplicaciones de $\mathcal{S}$ en K, con operaciones $(f + g)(s) = f(s) + g(s)$ y $(\alpha f)(s) = \alpha f(s)$, es un K-espacio vectorial. Con estas mismas operaciones, el conjunto $\mathcal{C}([a, b], \mathbb{R}) = \{f \mid [a, b] \to \mathbb{R}|f \text{ continua}\}$ es un $\mathbb{R}$-espacio vectorial.

Propiedades de los espacios vectoriales: $\forall u, v, w \in V, \alpha, \beta \in K$:

1. $u + v = u + w \implies v = w$

   $$u+v = u+w \implies (-u)+(u+v) = (-u)+(u+w) = ((-u)+u)+v = ((-u)+u)+w = v = w$$

2. $0u = 0_V$

   $$0u + 0u = (0 + 0)u = 0u = 0u + 0_V \implies 0u = 0_V$$

3. $\alpha 0_V = 0_V$

$$\alpha 0_V + 0_V = \alpha 0_V = \alpha(0_V + 0_V) = \alpha 0_V + \alpha 0_V \implies \alpha 0_V = 0_V$$

4. $u \in V, \alpha u = 0_V \implies \alpha = 0 \vee u = 0_V$; $\alpha u = \alpha v \implies \alpha = 0 \vee u = v$; $\alpha u = \beta u \implies \alpha = \beta \vee u = 0_V$

$$\begin{matrix} \alpha u = 0_V \\ \alpha \neq 0 \end{matrix} \implies u = (\alpha^{-1} \cdot \alpha)u = \alpha^{-1}(\alpha u) = \alpha^{-1} \cdot 0_V = 0_V$$

$$\begin{matrix} \alpha u = \alpha v \\ \alpha \neq 0 \end{matrix} \implies \alpha^{-1}(\alpha u) = \alpha^{-1}(\alpha v) = 1 \cdot u = 1 \cdot v = u = v$$

Para la última demostración, usamos (5):

$$\begin{matrix} \alpha u = \beta u \\ u \neq 0_V \end{matrix} \implies 0_V = \alpha u + (-(\beta u)) \overset{(5)}{=} \alpha u + (-\beta)u = (\alpha + (-\beta))u \implies$$
$$\implies \alpha + (-\beta) = 0 \implies \alpha = \beta$$

5. $u \in V, (-\alpha)u = \alpha(-u) = -(\alpha u)$

$$(-\alpha)u + \alpha u = (-\alpha + \alpha)u = 0u = 0_V \implies (-\alpha)u = -(\alpha u)$$

$$\alpha(-u) + \alpha u = \alpha(-u + u) = \alpha \cdot 0_V = 0_V \implies \alpha(-u) = -(\alpha u)$$

También podemos llamar 0 a 0_V si esto no conlleva ambigüedad.

1.3. Subespacios vectoriales

$U \subseteq V$ $(U \neq \emptyset)$ es un **subespacio vectorial** de V $(U \leq V)$ si $\forall u, v \in U, u + v \in U$ y $\forall u \in U, \alpha \in K, \alpha u \in U$. De esta forma U es también un K-espacio vectorial.

Los subconjuntos $\{0\}$ y V son subespacios vectoriales de V, y $\{(x_1, \ldots, x_n) \in K^n | x_1 + \cdots + x_n = 0\}$ es un subespacio vectorial de K^n. El conjunto $\mathcal{P}_n$ de polinomios con coeficientes reales con grado menor o igual a n, junto con el 0, es un subespacio vectorial de $\mathbb{R}[X]$. También lo es $U_{a,b} = \{f \in \mathcal{C}([a, b], \mathbb{R}) \mid f(a) = f(b)\}$ respecto de $\mathcal{C}([a, b], \mathbb{R})$.

1.4. Combinaciones lineales. Sistemas de generadores

$u \in V$ es **combinación lineal** de $v_1, \ldots, v_n \in V$ si $\exists \alpha_1, \ldots, \alpha_n \in K : u = \alpha_1 v_1 + \cdots + \alpha_n v_n$. Se dice que es combinación lineal de vectores de $\mathcal{S}$ (con $\mathcal{S} \subseteq V$) si $\exists n \in \mathbb{N}, v_1, \ldots, v_n \in \mathcal{S} : u = \alpha_1 v_1 + \cdots + \alpha_n v_n$. Los escalares α_i se llaman **coeficientes** de la combinación. Así, un subconjunto no vacío $U \subseteq V$ es un subespacio vectorial de V si cualquier combinación lineal de vectores de U también está en U.

Si $\mathcal{S}$ es un subconjunto no vacío de V, el conjunto $< \mathcal{S} >$ de todas las combinaciones lineales de vectores en $\mathcal{S}$ es el menor subespacio vectorial tal que $\mathcal{S} \subseteq < \mathcal{S} >$. **Demostración:** $u \in \mathcal{S} \implies 1 \cdot u \in < \mathcal{S} >$. Si $u, v \in < \mathcal{S} >$, entonces existirán $u_1, \ldots, u_k, v_1, \ldots, v_l \in \mathcal{S}$ y $\alpha_1, \ldots \alpha_k, \beta_1, \ldots, \beta_l \in K$ tales que $u = \alpha_1 u_1 + \cdots + \alpha_k u_k$ y $v = \beta_1 v_1 + \cdots + \beta_l + v_l$, por lo que

$u+v$ también es combinación lineal de vectores en $\mathcal{S}$ y por tanto está en $< S >$. Si $\alpha \in K$, tenemos que $u = \alpha\alpha_1 u_1 + \cdots + \alpha\alpha_k u_k \in < \mathcal{S} >$. Finalmente, si U es un subespacio vectorial de V que contiene a $\mathcal{S}$, como toda combinación de vectores de U está en U, entonces $< \mathcal{S} > \subseteq U$.

Un subconjunto $\mathcal{S} \subseteq V$ es un **sistema de generadores** de V si $< \mathcal{S} >= V$. V es **de dimensión finita** o **finitamente generado** si tiene un sistema de generadores finito. Estas definiciones también son válidas para subespacios vectoriales. U es el subespacio **generado** por $\mathcal{S}$ si $U =< \mathcal{S} >$.

1.5. Dependencia e independencia lineal

Un conjunto $\mathcal{S} \subseteq V$ es **linealmente independiente** si la única forma de obtener el vector nulo como combinación lineal de vectores de $\mathcal{S}$ es tomando todos los coeficientes nulos. De lo contrario es **linealmente dependiente**. Así, $\{v\}$ es linealmente independiente si y sólo si $v \neq 0$, con lo que cualquier conjunto $\{0\} \subseteq \mathcal{S}$ es linealmente dependiente. En $\mathbb{C}_{\mathbb{R}}$, $\{1, i\}$ ($\mathbb{C}$ con escalares reales) es linealmente independiente, mientras que en $\mathbb{C}_{\mathbb{C}}$ es linealmente dependiente porque $1 + (i)i = 0$. Un subconjunto de un conjunto linealmente independiente también es linealmente dependiente.

Un conjunto $\{v_1, \ldots, v_m\}$ con al menos dos vectores es linealmente dependiente si y sólo si alguno de ellos es combinación lineal del resto.

$\Longrightarrow$] Se tiene que existen $\alpha_1, \ldots, \alpha_m$ no todos nulos con $\sum_{i=1}^m \alpha_i v_i = \alpha_1 v_1 + \cdots + \alpha_m v_m = 0$. Suponemos $\alpha_j \neq 0$. Entonces $\alpha_j v_j = -\sum_{i=1, i\neq j}^m \alpha_i v_i = -\alpha_1 v_1 - \cdots - \alpha_{j-1} v_{j-1} - \alpha_{j+1} v_{j+1} - \cdots - \alpha_m v_m$, por lo que $v_j = -\sum_{i=1, i\neq j}^m (\alpha_j^{-1} \alpha_i v_i)$.

$\Longleftarrow$] Si v_j es combinación lineal de $\{v_i\}_{1\leq i \leq m}$, existen escalares $\alpha_1, \ldots, \alpha_{j-1}, \alpha_{j+1}, \ldots, \alpha_m$ tales que $v_j = \alpha_1 v_1 \ldots, \alpha_{j-1} v_{j-1} + \alpha_{j+1} v_{j+1} + \cdots + \alpha_m v_m = \sum_{i=1, i\neq j}^m \alpha_i v_i$, de modo que $0 = \alpha_1 v_1 + \cdots + \alpha_{j-1} v_{j-1} + (-1) v_j + \alpha_{j+1} v_{j+1} + \cdots + \alpha_m v_m$.

1.6. Bases. Dimensión

Una **base** de un espacio vectorial es un sistema de generadores linealmente independiente. Así, $\{1\}$ es base de K_K, $\{(1, 0, \ldots, 0), (0, 1, 0, \ldots, 0), \ldots, (0, 0, \ldots, 0, 1)\}$ es **base canónica** de K^n y $\{1, X, \ldots, X^n, \ldots\}$ lo es de $\mathbb{R}[X]$. Si llamamos $A_{ij} \in M_{m,n}(K)$ a la matriz con un 1 en el lugar ij y 0 en el resto, entonces $\{A_{ij} \mid 1 \leq i \leq m, 1 \leq j \leq n\}$ es base de $M_{m,n}(K)$.

$\mathcal{B}$ es base de V si y sólo si todo $v \in V$ se expresa de modo único como combinación lineal de $\mathcal{B}$.

$\Longrightarrow$] Como $\mathcal{B}$ es base, es sistema de generadores, por lo que todo vector de V es combinación lineal de vectores de $\mathcal{B}$. Ahora, sea $v = \sum_{i=1}^n \alpha_i v_i = \sum_{i=1}^n \beta_i v_i$. Entonces $0 = (\alpha_1 v_1 + \cdots + \alpha_n v_n) - (\beta_1 v_1 + \cdots + \beta_n v_n)$. Entonces $0 = (\alpha_1 - \beta_1) v_1 + \cdots + (\alpha_n - \beta_n) v_n$, y como $\mathcal{B}$ es linealmente independiente, se tiene que $\alpha_1 - \beta_1 = \cdots = \alpha_n - \beta_n = 0$ y por tanto v se expresa de modo único.

$\Longleftarrow$] $\mathcal{B}$ es entonces sistema de generadores y queda demostrar que es linealmente dependiente. Sean $v_1, \ldots, v_m \in \mathcal{B}$, como el 0 se representa de modo único, se tiene que si $\alpha_1 v_1 + \cdots + \alpha_m v_m = 0 = 0 v_1 + \cdots + 0 v_m$ entonces $\alpha_1 = \cdots = \alpha_m = 0$.

Si $S = \{v_1, v_2, \dots, v_m\}$ es un conjunto linealmente independiente y $u \notin < S >$, entonces $S \cup \{u\}$ es linealmente independiente. **Demostración:** Supongamos $\alpha_1, \dots, \alpha_m, \beta$ tales que $\alpha_1 v_1 + \cdots + \alpha_m v_m + \beta u = 0$. Si $\beta \neq 0$ entonces $\beta u = -(\alpha v_1 + \cdots + \alpha_m v_m)$, de donde $u = -\beta^{-1}\alpha_1 v_1 - \cdots - \beta^{-1}\alpha_m v_m$ y por tanto $u \in < S >\#$. Por tanto $\beta = 0$, pero entonces $\alpha_1 v_1 + \cdots + \alpha_m v_m = 0$, de donde $\alpha_1 = \cdots = \alpha_m = 0$, por lo que $S \cup \{u\}$ es linealmente independiente.

Teorema: Todo espacio vectorial $V \neq \{0\}$ tiene una base. **Demostración** para espacios finitamente generados. Sea $V =< u_1, \dots, u_m >$ y $\mathcal{A}$ el conjunto de subconjuntos $\mathcal{L} \subseteq \{u_1, \dots, u_m\}$ linealmente independientes. Sabemos que $\mathcal{A} \neq \emptyset$ porque como $V \neq \{0\}$ existe un $u_{i_0} \neq 0$, de modo que $\{u_{i_0}\}$ es linealmente independiente. Sea $\mathcal{B} = \{v_1, \dots, v_n\} \in \mathcal{A}$ un elemento de $\mathcal{A}$ con un máximo de vectores, entonces $\mathcal{B}$ es base de V si es sistema de generadores de V. Sea u_i un elemento del conjunto de generadores de V. Si $u_i \in \mathcal{B}$, entonces $u_i \in < \mathcal{B} >$. Si no, entonces $\mathcal{B} \cup \{u_i\} \subseteq \{u_1, \dots, u_m\}$ tiene un elemento más que $\mathcal{B}$, que es un elemento de $\mathcal{A}$ con el número máximo de vectores, por lo que $\mathcal{B} \cup \{u_i\} \notin \mathcal{A}$ y será linealmente dependiente, pero entonces $u_i \in < \mathcal{B} >\#$. Acabamos de probar que $\{u_1, \dots, u_m\} \subseteq < \mathcal{B} >$, por lo que $V =< u_1, \dots, u_m > \subseteq < \mathcal{B} > \subseteq \mathcal{V}$, por lo que $< \mathcal{B} >= \mathcal{V}$ y ya hemos demostrado que $\mathcal{B}$ es base de V.

Teorema de Steinitz: Si $\{u_1, \dots, u_n\}$ es base de V y $\{v_1, \dots, v_m\}$ es un conjunto linealmente independiente, entonces se pueden sustituir m vectores de $\{u_1, \dots, u_n\}$ por los vectores $\{v_1, \dots, v_m\}$ y obtener una nueva base de V. Entonces $m \leq n$. **Demostración:** Vemos que, como $\{v_1, \dots, v_m\}$ es linealmente independiente, entonces $v_1 \neq 0$, y tenemos que $v_1 = \alpha_1 u_1 + \cdots + a_n u_n$ con algún $\alpha_i \neq 0$. Podemos suponer que $\alpha_1 \neq 0$, y queremos probar que $\{v_1, u_2, \dots, u_n\}$ es base. Primero probamos que es sistema de generadores: $\alpha_1 u_1 = v_1 - \alpha_2 u_2 - \cdots - \alpha_n u_n$, por lo que $u_1 = \alpha_1^{-1} v_1 - \alpha_1^{-1}\alpha_2 u_2 - \cdots - \alpha_1^{-1}\alpha_n u_n$, de modo que u_1 es combinación lineal de $\{v_1, u_2, \dots, u_n\}$ y $< v_1, u_2, \dots, u_n >$ contiene un sistema de generadores, por lo que $\{v_1, u_2, \dots, u_n\}$ también lo es. Ahora bien, sean $\beta_1, \dots, \beta_n$ tales que $\beta_1 v_1 + \cdots + \beta_n u_n = 0$. Entonces

$$\begin{aligned} 0 &= \beta_1 v_1 + \cdots + \beta_n u_n \\ &= \beta_1(\alpha_1 u_1 + \cdots + \alpha_n u_n) + \beta_2 u_2 + \cdots + \beta_n u_n \\ &= \beta_1 \alpha_1 u_1 + (\beta_1 \alpha_1 + \beta_2) u_2 + \cdots + (\beta_1 \alpha_n + \beta_n) u_n \end{aligned}$$

Por tanto, como $\alpha_1 \neq 0$, entonces $\beta_1 = 0$ y por tanto $\beta_2 = \cdots = \beta_n = 0$ y el nuevo conjunto es también linealmente independiente. De aquí además podemos concluir que todo conjunto de vectores linealmente independiente de un espacio vectorial puede completarse a una base (añadiendo vectores fuera del subespacio generado por este conjunto).

Teorema: Todas las bases de un espacio vectorial no nulo tienen el mismo número de elementos. **Demostración:** Sean $\{u_1, \dots, u_n\}$ y $\{v_1, \dots, v_m\}$ bases de V. Como $\{v_1, \dots, v_m\}$ es linealmente independiente, por el teorema de Steinitz tenemos que $m \leq n$. Análogamente, como $\{u_1, \dots, u_n\}$ también lo es, entonces $n \leq m$, por lo que $m = n$.

Definimos la **dimensión** de V ($\dim_K(V)$ o $\dim(V)$) como el número de elementos de cualquier base de V. Si $V = \{0\}$, entonces $\dim(V) = 0$. Si V no es finitamente generado, entonces $\dim_K(V) = \infty$. Por ejemplo, $\dim_K(K) = 1$, $\dim_K(K^n) = n$, $\dim(M_{m,n}(K)) = mn$ y $\dim(\mathbb{R}[X]) = \infty$.

Teorema: Si $\dim(V) = n$ entonces:

1. Todo conjunto linealmente independiente de n vectores es una base.

Consecuencia del teorema de Steinitz.

2. Todo conjunto de generadores de n vectores es una base.
Siempre va a haber una base contenida en él y que va a tener n vectores.

3. Si $U \neq \emptyset$ es un subespacio vectorial de V entonces $\dim(U) \leq n$ y además $\dim(U) = n \iff U = V$.
Si $\mathcal{B}'$ es base de U entonces es un conjunto de vectores de V linealmente independiente y tiene como máximo n elementos, por lo que $\dim(U) \leq \dim(V)$. Además, si $\dim(U) = \dim(V)$ entonces $\mathcal{B}'$ tiene n elementos y, por la primera propiedad, también es base de V, de modo que $U =< \mathcal{B}' >= \mathcal{V}$.

Definimos el **rango** de un conjunto de vectores como

$$\text{rang}(\{u_1, \dots, u_m\}) = \dim(< u_1, \dots, u_m >)$$

Así, si $\{v_1, \dots, v_m\} \subseteq V$, es fácil comprobar que:

1. $\text{rang}(\{v_1, \dots, v_i, \dots, v_j, \dots, v_m\}) = \text{rang}(\{v_1, \dots, v_j, \dots, v_i, \dots, v_m\})$.

2. $\alpha \neq 0 \implies \text{rang}(\{v_1, \dots v_i, \dots, v_m\}) = \text{rang}(\{v_1, \dots, \alpha v_i, \dots, v_m\})$.

3. $\text{rang}(\{v_1, \dots, v_i, \dots, v_j, \dots, v_m\}) = \text{rang}(\{v_1, \dots, v_i + \alpha v_j, \dots, v_j, \dots, v_m\})$.

Una forma de determinar el rango de un conjunto de vectores es ir haciendo estas operaciones y eliminando posibles vectores nulos hasta encontrar un conjunto linealmente independiente.

1.6.1. Operaciones elementales. Matrices escalonadas. Método Gauss-Jordan

En una matriz $A \in M_{m,n}(K)$, a intercambiar dos filas, multiplicar una fila por un $0 \neq \alpha \in K$ o añadir una fila a otra multiplicada por un $\alpha \in K$ se les llama **operaciones elementales por filas**. Las **operaciones elementales por columnas** se definen de forma análoga. Si B es la matriz resultante de aplicar una serie de operaciones elementales por filas a A, entonces el subespacio de K^n generado por las filas de A es el mismo que el generado por las filas de B.

Una matriz $A = (a_{ij}) \in M_{m,n}(K)$ está en forma **escalonada por filas** si las filas nulas, de haberlas, son las últimas, el primer elemento no nulo de cada fila no nula es un 1 (llamado **pivote**) y el pivote de cada fila no nula está en una columna posterior a la de cada uno de los pivotes anteriores. En las matrices escalonadas por filas, las filas no nulas son linealmente independientes.

Si en cada columna que contenga un pivote el resto de elementos son nulos, la matriz está en forma **escalonada reducida por filas**. Cambiando filas por columnas tendríamos una matriz en forma **escalonada por columnas** o **escalonada reducida por columnas**.

Método de eliminación Gauss-Jordan: Toda matriz se puede llevar a forma escalonada (también escalonada reducida) mediante operaciones elementales por filas. Algoritmo:

1. Encontrar el primer elemento a no nulo de la primera columna no nula e intercambiar su fila con la primera.

2. Multiplicarla por a^{-1} para obtener un pivote.

3. Hacer operaciones $Fila_i - cFila_1$, donde c es el elemento de cada fila debajo del pivote.

4. Repetir el proceso con la matriz resultado de eliminar la primera fila hasta terminar la matriz.

5. Para obtener la escalonada reducida, hacer operaciones $Fila_i - cFila_k$, donde $i < k$ y c es el elemento encima del pivote de la fila i.

Para obtener la base de un subespacio K^n generado por m vectores, escalonamos la matriz $m \times n$ cuyas filas son los vectores generadores del subespacio, y los vectores correspondientes a filas no nulas forman una base. Los vectores fila de cada nueva matriz son combinaciones lineales de los iniciales. Para obtener los coeficientes de estas combinaciones, anexamos la matriz identidad a la derecha de la original, separada por una línea, y le aplicamos también las operaciones, sin considerar estos coeficientes como parte de la matriz a escalonar.

1.7. Coordenadas. Cambio de base

Las **coordenadas** de un vector $v \in V$ respecto a la base $\mathcal{B} = \{u_1, \dots, u_n\}$ son la única n-upla $[v]_\mathcal{B} = (x_1, \dots, x_n) \in K^n$ con $v = x_1u_1 + \cdots + x_nu_n$, de forma que dos vectores de V son iguales si y solo si tienen las mismas coordenadas respecto a una base $\mathcal{B}$, y operar con vectores en V es equivalente a operar en K^n con las n-uplas de sus coordenadas, pues $[v + v']_\mathcal{B} = (x_1 + x'_1, \dots, x_n + x'_n) = [v]_\mathcal{B} + [v']_\mathcal{B}$ y $[rv]_\mathcal{B} = (rx_1, \dots, rx_n) = r[v]_\mathcal{B}$.

1.7.1. Cambio de coordenadas

Sean $\mathcal{B} = \{u_1, \dots, u_n\}$ y $\mathcal{B}' = \{u'_1, \dots, u'_n\}$ bases de V, y llamamos $[u'_j]_\mathcal{B} = (p_{1j}, \dots, p_{nj})$, de forma que $u'_j = \sum_{i=1}^n p_{ij}u_i$. Si $[v]_{\mathcal{B}'} = (x'_1, \dots, x'_n)$, entonces

$$v = x'_1u'_1 + \cdots + x'_nu'_n = \sum_{j=1}^n x'_ju'_j = \sum_{j=1}^n x'_j \left(\sum_{i=1}^n p_{ij}u_i\right) = \sum_{j=1}^n \sum_{i=1}^n x'_jp_{ij}u_i = \sum_{i=1}^n \left(\sum_{j=1}^n x'_jp_{ij}\right) u_i$$

De forma que si $[v]_\mathcal{B} = (x_1, \dots, x_n)$, entonces $x_i = \sum_{j=1}^n x'_jp_{ij}$. A las ecuaciones

$$\left.\begin{array}{ccccccc} x_1 & = & p_{11}x'_1 & + & \dots & + & p_{1n}x'_n \\ & \vdots & & & & & \\ x_n & = & p_{n1}x'_1 & + & \dots & + & p_{nn}x'_n \end{array}\right\}$$

las llamamos **ecuaciones del cambio de base de $\mathcal{B}'$ a $\mathcal{B}$.**

1.8. Producto de matrices

Sean $A = (a_{ij}) \in M_{m,n}(K)$ y $B = (b_{ij}) \in M_{n,p}(K)$, definimos $AB = (c_{ij}) \in M_{m,p}(K)$ tal que $c_{ij} = \sum_{k=1}^n a_{ik}b_{kj}$. En general no es conmutativo, aun si ambos productos se pueden efectuar y fuesen matrices del mismo tamaño. Propiedades:

1. **Asociativa:** $(AB)C = A(BC)$.

$$((AB)C)_{ij} = \sum_{k=1}^{p}(AB)_{ik}C_{kj} = \sum_{k=1}^{p}\sum_{l=1}^{n} A_{il}B_{lk}C_{kj} = \sum_{l=1}^{n}\sum_{k=1}^{p} A_{il}B_{lk}C_{kj} =$$
$$= \sum_{l=1}^{n} A_{il}\left(\sum_{k=1}^{p} B_{lk}C_{kj}\right) = \sum_{l=1}^{n} A_{il}(BC)_{lj} = (A(BC))_{ij}$$

2. **Distributiva respecto de la suma:** $A(B+C) = AB + AC$.

3. La **matriz identidad**, $I_n = (\delta_{ij})$ con $\delta_{ii} = 1$ y $\delta_{ij} = 0$ si $i \neq j$, satisface $AI_n = A$ y $I_nB = B$ para cada $A \in M_{m,n}(K)$ y $B \in M_{n,m}(K)$.

$$C_{ij} = \sum_{k=1}^{n} A_{ik}\delta_{kj} \underset{(\text{el resto son } = 0)}{=} A_{ij}\delta_{jj} = A_{ij}$$

   La demostración de que $I_nB = B$ es análoga.

4. $\alpha(AB) = (\alpha A)B = A(\alpha B)$.

$A \in M_n(K)$ es **invertible** si existe $B \in M_n(K)$ tal que $AB = BA = I_n$. Entonces B es la **matriz inversa** de A y se representa A^{-1}. Supongamos que B y C son inversas de A. Entonces $C = CI_n = C(AB) = (CA)B = I_nB = B$, por lo que la inversa es única.

Si $A, B \in M_n(K)$ son invertibles, entonces AB es también invertible, y $(AB)^{-1} = B^{-1}A^{-1}$:

$$(B^{-1}A^{-1})(AB) = B^{-1}(A^{-1}A)B = B^{-1}I_nB = B^{-1}B = I_n$$

Las ecuaciones de cambio de base se pueden expresar como

$$\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} p_{11} & p_{12} & \cdots & p_{1n} \\ p_{21} & p_{22} & \cdots & p_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ p_{n1} & p_{n2} & \cdots & p_{nn} \end{pmatrix} \begin{pmatrix} x'_1 \\ x'_2 \\ \vdots \\ x'_n \end{pmatrix}$$

donde las columnas de (p_{ij}) son los vectores de $\mathcal{B}'$ respecto a $\mathcal{B}$. Abreviadamente, $X_{\mathcal{B}} = M_{\mathcal{B}\mathcal{B}'}X'_{\mathcal{B}'}$, donde a $M_{\mathcal{B}\mathcal{B}'}$ la llamamos **matriz de cambio de base** de $\mathcal{B}'$ a $\mathcal{B}$. Podemos deducir que $M_{\mathcal{B}\mathcal{B}} = I_n$, $M_{\mathcal{B}''\mathcal{B}} = M_{\mathcal{B}''\mathcal{B}'}M_{\mathcal{B}'\mathcal{B}}$ y por tanto $M_{\mathcal{B}'\mathcal{B}}M_{\mathcal{B}\mathcal{B}'} = M_{\mathcal{B}\mathcal{B}'}M_{\mathcal{B}'\mathcal{B}} = I_n$. Así, toda matriz de cambio de base es invertible, y viceversa.

1.9. Operaciones con subespacios

Si $\{U_i\}_{i\in I} \neq \emptyset$ es un conjunto de subespacios vectoriales de V, entonces $\bigcap_{i\in I} U_i$ también es subespacio vectorial de V, pero en general $\bigcup_{i\in I} U_i$ no es subespacio vectorial. Llamamos **suma** de U_1 y U_2 al subespacio

$$U_1 + U_2 = \{u_1 + u_2 | u_1 \in U_1, u_2 \in U_2\}$$

Este es el menor subespacio vectorial V que contiene a U_1 y U_2. **Demostración:** Si $u \in U_1$, como $0 \in U_2$, se tiene que $u = u + 0 \in U_1 + U_2$, y viceversa, de modo que $U_1 \cup U_2 \subseteq U_1 + U_2$

y además $U_1 + U_2 \neq \emptyset$. Demostraremos ahora que es un espacio vectorial. Si $v, v' \in U_1 + U_2$, existirán $u_1, u_1' \in U$ y $u_2, u_2' \in U_2$ con $v = u_1 + u_2$ y $v' = u_1' + u_2'$, y si $\alpha, \alpha' \in K$, se tiene que

$$\alpha v + \alpha' v' = \alpha u_1 + \alpha u_2 + \alpha' u_1' + \alpha' u_2' = (\alpha u_1 + \alpha' u_1') + (\alpha u_2 + \alpha' u_2') \in U_1 + U_2$$

La **fórmula de Grassmann** o **teorema de Grassman** nos dice que si U_1 y U_2 son subespacios de V, y V tiene dimensión finita, entonces

$$\dim(U_1) + \dim(U_2) = \dim(U_1 \cap U_2) + \dim(U_1 + U_2)$$

Demostración. Sea $\{u_1, \dots, u_t\}$ base de $U_1 \cap U_2$, que completamos por un lado a la base $\{u_1, \dots, u_t, u_{t+1}, \dots, u_r\}$ de U_1 y por otro a la base $\{u_1, \dots, u_t, v_{t+1}, \dots, v_s\}$ de U_2. Entonces $\{u_1, \dots, u_t, u_{t+1}, \dots, u_r, v_{t+1}, \dots, v_s\}$ es sistema de generadores de $U_1 + U_2$. Queda ver que es además linealmente independiente.

Supongamos $\alpha_1 u_1 + \cdots + \alpha_t u_t + \alpha_{t+1} u_{t+1} + \cdots + \alpha_r u_r + \beta_{t+1} v_{t+1} + \cdots + \beta_s v_s = 0$. Entonces $\alpha_1 u_1 + \cdots + \alpha_r u_r = -\beta_{t+1} v_{t+1} - \cdots - \beta_s v_s \in U_1 \cap U_2$. Pero como $\{u_1, \dots, u_t\}$ es base de $U_1 \cap U_2$, se tiene que $-\beta_{t+1} v_{t+1} - \cdots - \beta_s v_s = \gamma_1 u_1 + \cdots + \gamma_t u_t$, por lo que $\gamma_1 u_1 + \cdots + \gamma_t u_t + \beta_{t+1} v_{t+1} + \cdots + \beta_s v_s = 0$. Al ser $\{u_1, \dots, u_t, v_{t+1}, \dots, v_s\}$ base de U_2, se tiene que $\beta_{t+1} = \cdots = \beta_s = 0$, por lo que $\alpha_1 u_1 + \cdots + \alpha_r u_r = 0$. Pero como $\{u_1, \dots, u_r\}$ es base de U_1 se tiene que $\alpha_1 = \cdots = \alpha_r = 0$.

Así, se tiene que $\{u_1, \dots, u_t, u_{t+1}, \dots, u_r, v_{t+1}, \dots, v_s\}$ es una familia linealmente independiente y por tanto una base del subespacio $U_1 + U_2$, que tendrá dimensión $\dim(U_1 + U_2) = r + s - t = \dim(U_1) + \dim(U_2) - \dim(U_1 \cap U_2)$.

U_1 y U_2 subespacios de V están en **suma directa** si $U_1 \cap U_2 = \{0\}$. Se dice entonces que $U_1 + U_2$ es una suma directa y se representa $U_1 \oplus U_2$. Por tanto $\dim(U_1 \oplus U_2) = \dim(U_1) + \dim(U_2)$.

Una suma de subespacios $U_1 + U_2$ es directa si y sólo si $\forall v \in U_1 + U_2, \exists! u_1 \in U_1, u_2 \in U_2 : v = u_1 + u_2$. Llamamos a u_1 y u_2 las **componentes** de v.

$\Longrightarrow$] Si $u_1 + u_2 = u_1' + u_2'$ con $u_1, u_1' \in U_1$ y $u_2, u_2' \in U_2$, entonces $u_1 - u_1' = u_2' - u_2 \in U_1 \cap U_2$, por lo que $u_1 - u_1' = u_2' - u_2 = 0$ y $u_1 = u_1'$ y $u_2 = u_2'$.

$\Longleftarrow$] Si $u \in U_1 \cap U_2$, entonces $u = u + 0 = 0 + u$, pero como u se expresa de modo único como suma de un vector de U_1 y otro de U_2, se tiene que $u = 0$, y por tanto $U_1 \cap U_2 = \{0\}$.

Dado U subespacio vectorial de V existe U', llamado **complementario** de U en V, tal que $V = U \oplus U'$, pues si expandimos la base $\{u_1, \dots, u_r\}$ de U a una base $\{u_1, \dots, u_r, u_{r+1}, \dots, u_n\}$ de V entonces $U' = < u_{r+1}, \dots, u_n >$ satisface la condición. El complementario no tiene por qué ser único.

Una suma $U_1 + \cdots + U_k$ es suma directa si todo vector de la suma se expresa de modo único como suma de un vector de cada U_i, lo que ocurre si y sólo si $\forall i \in \{1, \dots, k\}, U_i \cap (\sum_{1 \leq j \leq k, j \neq i} U_j) = \{0\}$.

Si $u_1, \dots, u_k \in V$, entonces $< u_1, \dots, u_k > = < u_1 > + \cdots + < u_k >$. Si además son linealmente independientes, entonces $< u_1, \dots, u_k > = < u_1 > \oplus \cdots \oplus < u_n >$. Así, $\{u_1, \dots, u_n\}$ es base de V si y sólo si $V = < u_1 > \oplus \cdots \oplus < u_n >$.

Capítulo 2

Aplicaciones lineales

$f : U \to V$ es una **aplicación lineal** u **homomorfismo de espacios vectoriales** si $f(u+u') = f(u) + f(u') \forall u, u' \in U$ y $f(\alpha u) = \alpha f(u) \forall \alpha \in K, u \in U$, es decir, si $f(\sum \alpha_i u_i) = \sum \alpha_i f(u_i)$. Ejemplos:

- La **aplicación identidad:** $Id_V : V \to V$ con $Id_V(v) = v$.
- La **aplicación inclusión:** $i : U \subseteq V \to V$ con $i(u) = u$.
- La **aplicación lineal nula:** $0 : U \to V$ con $0(u) = 0$.
- La **homotecia de razón α:** $h_\alpha : V \to V$ con $h_\alpha(v) = \alpha v$.
- Las **proyecciones de V sobre U y W,** con $V = U \oplus W$: $p_U : V \to U$ y $p_W : V \to W$, tales que si $v = u + w$ con $v \in V$, $u \in U$ y $w \in W$, entonces $p_U(v) = u$ y $p_W(v) = w$.
- La aplicación $f_A : K^n \to K^m$ con $A \in M_{m,n}(K)$, dada por

$$f_A(v) = A \begin{pmatrix} | \\ v \\ | \end{pmatrix}$$

Tenemos que $f(0_U) = 0_V$, y que $f(-u) = -f(u)$. Además, si $f : U \to V$ y $g : V \to W$ son aplicaciones lineales, $g \circ f : U \to W$ también lo es.

2.1. Aplicaciones lineales y subespacios. Núcleo e Imagen

El **núcleo** de f se define como $\mathrm{Nuc}(f) = \ker(f) = f^{-1}(\{0\})$, y la **imagen** de f como $\mathrm{Im}(f) = \{f(u)\}_{u \in U}$. Si $U' \leq U$ entonces $f(U') \leq V$, y si $V' \leq V$, entonces $\mathrm{Nuc}(f) \leq f^{-1}(V') \leq U$. En particular, $\mathrm{Nuc}(f)$ e $\mathrm{Im}(f)$ son espacios vectoriales, y si $U' =< u_1, \dots, u_r >\leq U$, entonces $f(U') =< f(u_1), \dots, f(u_r) >$. **Demostración:** Sean $u_1, u_2 \in U, \alpha_1, \alpha_2 \in K$, y sean $v_1 = f(u_1), v_2 = f(u_2) \in f(U')$. Entonces $\alpha_1 v_1 + \alpha_2 v_2 = \alpha_1 f(u_1) + \alpha_2 f(u_2) = f(\alpha_1 u_1 + \alpha_2 u_2) \in f(U')$, por lo que $f(U')$ es un espacio vectorial. Ahora bien, si V' es un subespacio de V entonces $\{0\} \subseteq V'$, por lo que $f^{-1}(\{0\}) = \mathrm{Nuc}(f) \subseteq f^{-1}(V')$. Entonces si $u_1, u_2 \in f^{-1}(V')$ y $\alpha_1, \alpha_2 \in K$, entonces $f(\alpha_1 u_1 + \alpha_2 u_2) = \alpha_1 f(u_1) + \alpha_2 f(u_2) \in V'$, y por lo tanto $\alpha_1 u_1 + \alpha_2 u_2 \in f^{-1}(V')$ y $f^{-1}(V')$ es un espacio vectorial.

Teorema: Para $f : U \to V$ con $\dim(U)$ finita, entonces $\dim(U) = \dim(\mathrm{Nuc}(f)) + \dim(\mathrm{Im}(f))$.
Demostración: Sea $\{v_1, \dots, v_n\}$ base de U. Entonces $\mathrm{Im}(f) =< f(v_1), \dots f(v_n) >$ es de dimensión finita. Ahora sea $\{v_1, \dots, v_k\}$ base de $\mathrm{Nuc}(f) \leq U$, con $k \leq n$. Entonces $f(v_1) = \cdots = f(v_k) = 0$, por lo que $\mathrm{Im}(f) =< f(v_1), \dots, f(v_k), f(v_{k+1}), \dots, f(v_n) >=< f(v_{k+1}), \dots, f(v_n) >$, por lo que $\{f(v_{k+1}), \dots, f(v_n)\}$ es sistema de generadores de $\mathrm{Im}(f)$. A continuación mostramos que es linealmente independiente. Sea $0 = \alpha_{k+1} f(v_{k+1}) + \cdots + \alpha_n f(v_n) = f(\alpha_{k+1} v_{k+1} + \cdots + \alpha_n v_n)$. Entonces $\alpha_{k+1} v_{k+1} + \cdots + \alpha_n v_n \in \mathrm{Nuc}(f)$, por lo que existen $\beta_1, \dots, \beta_k$ tales que $\alpha_{k+1} v_{k+1} + \cdots + \alpha_n v_n = \beta_1 v_1 + \cdots + \beta_k v_k$. Pero entonces $\beta_1 v_1 + \cdots + \beta_k v_k - \alpha_{k+1} v_{k+1} - \cdots - \alpha_n v_n = 0$, y como $\{v_1, \dots, v_n\}$ es base de U, se tiene que $\beta_1 = \cdots = \beta_k = \alpha_{k+1} = \cdots = \alpha_n = 0$, de modo que $\{v_{k+1}, \dots, v_n\}$ es linealmente independiente y por ello $\{f(v_{k+1}), \dots, f(v_n)\}$ también, por lo que es base de $\mathrm{Im}(f)$.

Llamamos **rango** de f a la dimensión de la imagen: $\mathrm{rang}(f) = \dim(\mathrm{Im}(f))$. Así, dada $f : U \to V$ y $\{u_1, \dots, u_n\}$ base de U, entonces $f(U) =< f(u_1), \dots, f(u_n) >$ y por tanto

$$\mathrm{rang}(f) = \dim(\mathrm{Im}(f)) = \dim(< f(u_1), \dots, f(u_n) >) = \mathrm{rang}(\{f(u_1), \dots, f(u_n)\})$$

Si $f : U \to V$ es una aplicación lineal y $\dim(U) = \dim(V) < \infty$, entonces

$$f \text{ inyectiva} \iff f \text{ suprayectiva} \iff f \text{ biyectiva} \iff \mathrm{rang}(f) = \dim(U) \iff \mathrm{Nuc}(f) = \{0\}$$

$\Longrightarrow$ 3] Equivalen al hecho de que, para $f : A \to B$ con A y B conjuntos finitos, es lo mismo decir que f sea inyectiva, suprayectiva o biyectiva.

$\Longrightarrow$ 4] $\mathrm{rang}(f) = \dim(\mathrm{Im}(U)) \overset{\text{(supray.)}}{=} \dim(V)$. Si f no fuera suprayectiva, entonces $\dim(\mathrm{Im}(U)) < \dim(V)$.

$\Longrightarrow$ 5] $u \in \mathrm{Nuc}(f) \implies f(u) = 0_V = f(0_U) \implies u = 0_U$

$\Longrightarrow$ 1] $\mathrm{Nuc}(f) = \{0\} \implies (f(u) = f(u') \implies 0 = f(u - u') \implies u - u' \in \mathrm{Nuc}(f) \implies u = u')$

El homomorfismo $f : U \to V$ es un **isomorfismo de espacios vectoriales** si es biyectivo, un **endomorfismo** de U si $U = V$ y un **automorfismo** es un endomorfismo biyectivo. Ahora, dado el isomorfismo $f : U \to V$, $f^{-1} : V \to U$ es una aplicación lineal y por tanto un isomorfismo. **Demostración:** Consideramos $u = f^{-1}(v)$ y $u' = f^{-1}(v')$. Entonces $f(u + u') = f(u) + f(u') = v + v'$ y por tanto $f^{-1}(v + v') = u + u' = f^{-1}(v) + f^{-1}(v')$. Del mismo modo, $f(\alpha u) = \alpha f(u) = \alpha v$, por lo que $f^{-1}(\alpha v) = \alpha u = \alpha f^{-1}(v)$.

U y V son **isomorfos** ($U \cong V$) si existe un isomorfismo $f : U \to V$. Podemos comprobar que la relación es de equivalencia, y si U y V son K-espacios vectoriales, entonces $U \cong V \iff \dim(U) = \dim(V)$.

$\Longrightarrow$] $U \cong V \implies \dim(U) = \dim(\mathrm{Nuc}(f)) + \dim(\mathrm{Im}(f)) = 0 + \dim(V)$

$\Longleftarrow$] Sean $f : U \to K^n$ y $g : V \to K^n$ isomorfismos con $f(u) = [u]_{\mathcal{B}}$ y $g(v) = [v]_{\mathcal{B}'}$; entonces $g^{-1} \circ f : U \to V$ también es un isomorfismo.

2.2. Determinación de una aplicación lineal

Si U y V son K-espacios vectoriales con $\mathcal{B} = \{u_1, \dots, u_n\}$ base de U y $v_1, \dots, v_n$ vectores cualesquiera de V, existe una única $f : U \to V$ con $f(u_i) = v_i \forall i$, pues es aquella dada por $f(\alpha_1 u_1 + \cdots + \alpha_n u_n) = \alpha_1 v_1 + \cdots + \alpha_n v_n$. Esto también se cumple para espacios de dimensión infinita.

También, si $\mathcal{B} = \{u_i\}_{i \in I}$ es base de U (la cual puede ser infinita) y $f : U \to V$ lineal entonces:

1. f es inyectiva si y sólo si $\{f(u_i)\}_{i\in I}$ es linealmente independiente.

   $\Longrightarrow]$ $0 = \alpha_1 f(u_1) + \cdots + \alpha_k f(u_k) = f(\alpha_1 u_1 + \cdots + \alpha_k u_k) \implies \alpha_1 u_1 + \cdots + \alpha_k u_k \in \mathrm{Nuc}(f) = \{0\} \implies \alpha_1 u_1 + \cdots + \alpha_k u_k = 0 \implies \alpha_1, \dots, \alpha_k = 0$.

   $\Longleftarrow]$ Partimos de que $\{f(u_i)\}_{i\in I}$ es linealmente independiente. Sea $u \in \mathrm{Nuc}(f)$, si $u = \alpha_1 u_1 + \cdots + \alpha_k u_k$ entonces $0 = f(u) = \alpha_1 f(u_1) + \cdots + \alpha_k f(u_k) \implies \alpha_i = 0 \forall i \implies u = 0$, por lo que entonces $\mathrm{Nuc}(f) = \{0\}$.

2. f es suprayectiva si y sólo si $\{f(u_i)\}_{i\in I}$ es una familia de generadores de V.

$$f \text{ suprayectiva} \iff f(U) = V \iff <\{f(u_i)\}_{i\in I}>= V$$

3. f es biyectiva, y por tanto isomorfismo, si y sólo si $\{f(u_i)\}_{i\in I}$ es base de V.

2.3. Representación matricial de una aplicación lineal. Rango de una matriz. Matrices equivalentes

Si U y V son K-espacios vectoriales de dimensión finita, $\mathcal{B} = \{u_1, \dots, u_n\}$ es base de U y $\mathcal{B}' = \{v_1, \dots, v_m\}$ base de V, y $f : U \to V$ es un homomorfismo, entonces para cada j existirán a_{ij} tales que

$$f(u_j) = \sum_{i=1}^{m} a_{ij} v_i$$

Así, si $[u]_{\mathcal{B}} = (x_1, \dots, x_n)$, entonces $u = \sum_{j=1}^{n} x_j u_j \in U$ y

$$f(u) = f\left(\sum_{j=1}^{n} x_j u_j\right) = \sum_{j=1}^{n} x_j f(u_j) = \sum_{j=1}^{n} x_j \left(\sum_{i=1}^{m} a_{ij} v_i\right) = \sum_{j=1}^{n}\sum_{i=1}^{m} (x_j a_{ij}) v_i = \sum_{i=1}^{m}\left(\sum_{j=1}^{n} a_{ij} x_j\right)$$

por lo que $[f(u)]_{\mathcal{B}'} = (\sum_{j=1}^{n} a_{1j} x_j, \dots, \sum_{j=1}^{n} a_{mj} x_j)$, de modo que, si $[f(u)]_{\mathcal{B}'} = (y_1, \dots, y_m)$, entonces

$$\begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix} = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

Siendo las columnas de (a_{ij}) los $[f(u_j)]_{\mathcal{B}'}$, es decir, las imágenes de los elementos de $\mathcal{B}$ respecto de $\mathcal{B}'$. Entonces $[f(u)]_{\mathcal{B}'} = A[u]_{\mathcal{B}}$, lo que se conoce como **representación matricial**

de f respecto a las bases $\mathcal{B}$ y $\mathcal{B}'$. A la matriz (a_{ij}) se le llama **matriz de f** o **matriz asociada a f respecto a las bases $\mathcal{B}$ y $\mathcal{B}'$**, y se denomina $M_{\mathcal{B}',\mathcal{B}}(f)$. Así,

$$[f(u)]_{\mathcal{B}'} = M_{\mathcal{B}',\mathcal{B}}(f)[u]_{\mathcal{B}}$$

Tenemos que $M_{\mathcal{B}',\mathcal{B}}(f)$ está completamente determinada por f, y de igual modo, f está unívocamente determinada por $M_{\mathcal{B}',\mathcal{B}}(f)$. Además, si $f : U \to V$ y $g : V \to W$ son aplicaciones lineales y $\mathcal{B}_1$, $\mathcal{B}_2$ y $\mathcal{B}_3$ son bases respectivas de U, V y W, entonces $M_{\mathcal{B}_3,\mathcal{B}_1}(g \circ f) = M_{\mathcal{B}_3,\mathcal{B}_2}(g)M_{\mathcal{B}_2,\mathcal{B}_1}(f)$. **Demostración:** Para cada $u \in U, v \in V$, tenemos que $M_{\mathcal{B}_2,\mathcal{B}_1}(f)[u]_{\mathcal{B}_1} = [f(u)]_{\mathcal{B}_2}$ y $M_{\mathcal{B}_3,\mathcal{B}_2}(g)[v]_{\mathcal{B}_2} = [g(v)]_{\mathcal{B}_3}$, por lo que

$$M_{\mathcal{B}_3,\mathcal{B}_2}(g)M_{\mathcal{B}_2,\mathcal{B}_1}(f)[u]_{\mathcal{B}_1} = M_{\mathcal{B}_3,\mathcal{B}_2}(g)[f(u)]_{\mathcal{B}_2} = [g(f(u))]_{\mathcal{B}_3} = [(g \circ f)(u)]_{\mathcal{B}_3}$$

y por la unicidad de la matriz de una aplicación lineal respecto a dos bases, se tiene que $M_{\mathcal{B}_3,\mathcal{B}_1}(g \circ f) = M_{\mathcal{B}_3,\mathcal{B}_2}(g)M_{\mathcal{B}_2,\mathcal{B}_1}(f)$.

Sean U y V K-espacios vectoriales con bases respectivas $\mathcal{B}$ y $\mathcal{B}'$, $f : U \to V$ una aplicación lineal y $A = M_{\mathcal{B}',\mathcal{B}}(f)$. Entonces f es un isomorfismo si y sólo si A es invertible, y entonces $A^{-1} = M_{\mathcal{B},\mathcal{B}'}(f^{-1})$.

$\Longrightarrow$] Sea $B = M_{\mathcal{B},\mathcal{B}'}(f^{-1})$:

$$AB = M_{\mathcal{B}',\mathcal{B}}(f)M_{\mathcal{B},\mathcal{B}'}(f^{-1}) = M_{\mathcal{B}',\mathcal{B}'}(f \circ f^{-1}) = M_{\mathcal{B}',\mathcal{B}'}(Id_V) = I_n$$

$$BA = M_{\mathcal{B},\mathcal{B}'}(f^{-1})M_{\mathcal{B}',\mathcal{B}}(f) = M_{\mathcal{B},\mathcal{B}}(f^{-1} \circ f) = M_{\mathcal{B},\mathcal{B}}(Id_U) = I_n$$

$\Longleftarrow$] Al ser invertible es cuadrada, por lo que $\dim(U) = \dim(V) = n$, y si consideramos $g : V \to U$ tal que $M_{\mathcal{B},\mathcal{B}'}(g) = A^{-1}$, se tiene que

$$M_{\mathcal{B},\mathcal{B}}(g \circ f) = M_{\mathcal{B},\mathcal{B}'}(g)M_{\mathcal{B}',\mathcal{B}}(f) = A^{-1}A = I_n = M_{\mathcal{B},\mathcal{B}}(Id_U) \implies g \circ f = Id_U$$

$$M_{\mathcal{B}',\mathcal{B}'}(f \circ g) = M_{\mathcal{B}',\mathcal{B}}(f)M_{\mathcal{B},\mathcal{B}'}(g) = AA^{-1} = I_n = M_{\mathcal{B}',\mathcal{B}'}(Id_V) \implies f \circ g = Id_V$$

Así, si $\mathcal{B}$ y $\mathcal{B}'$ son bases de V, como $M_{\mathcal{B},\mathcal{B}'} = M_{\mathcal{B},\mathcal{B}'}(Id_V)$, se tiene que

$$M_{\mathcal{B}',\mathcal{B}}^{-1} = (M_{\mathcal{B}',\mathcal{B}}(Id_V))^{-1} = M_{\mathcal{B},\mathcal{B}'}(Id_V^{-1}) = M_{\mathcal{B},\mathcal{B}'}(Id_V) = M_{\mathcal{B},\mathcal{B}'}$$

También se tiene que si $\mathcal{B}_1$ y $\mathcal{B}_2$ son bases de U y $\mathcal{B}_1'$ y $\mathcal{B}_2'$ son bases de V, entonces

$$M_{\mathcal{B}_2',\mathcal{B}_2}(f) = M_{\mathcal{B}_2',\mathcal{B}_2}(Id_V \circ f \circ Id_U) = M_{\mathcal{B}_2',\mathcal{B}_1'} \cdot M_{\mathcal{B}_1'\mathcal{B}_1}(f) \cdot M_{\mathcal{B}_1,\mathcal{B}_2}$$

Para $A, B \in M_{m,n}(K)$, A y B son **equivalentes** si existen matrices invertibles $P \in M_m(K)$ y $Q \in M_n(K)$ tales que $B = PAQ$. Esta es una relación de equivalencia.

Se llama **rango** de $A \in M_{m,n}(K)$ al máximo de columnas linealmente independientes consideradas como vectores de K^m, es decir, $\text{rang}(A) = \dim(< C_1, \dots, C_n >)$, y por tanto $\text{rang}(A) \leq m, n$. Dado que las columnas de $M_{\mathcal{B}',\mathcal{B}}(f)$ son las imágenes en f de los elementos de $\mathcal{B}$ sobre $\mathcal{B}'$, se tiene que $\text{rang}(M_{\mathcal{B}',\mathcal{B}}(f)) = \text{rang}(f)$, y como las matrices invertibles corresponden a isomorfismos, se tiene que $A \in M_n(K)$ es invertible si y sólo si $\text{rang}(A) = n$, para lo que basta con considerar el homomorfismo $f : K^n \to K^n$ con $M_{\mathcal{C},\mathcal{C}}(f) = A$.

Dados K-espacios vectoriales U y V con dimensiones respectivas n y m, y el homomorfismo $f : U \to V$, existen bases $\mathcal{B}$ de U y $\mathcal{B}'$ de V tales que

$$M_{\mathcal{B}',\mathcal{B}}(f) = \left(\begin{array}{c|c} I_r & 0 \\ \hline 0 & 0 \end{array}\right) \in M_{m,n}(K)$$

con $r = \text{rang}(f)$. **Demostración:** Si $r = \text{rang}(f)$ entonces $\dim(\text{Nuc}(f)) = n - r$. Además, si $\{u_{r+1}, \dots, u_n\}$ es base de $\text{Nuc}(f)$ que se extiende a la base $\mathcal{B} = \{u_1, \dots, u_r, u_{r+1}, \dots, u_n\}$ de U, entonces $f(u_{r+1}) = \cdots = f(u_n) = 0$, y $f(u_1), \dots, f(u_r)$ son linealmente dependientes, dado que si $\alpha_1 f(u_1) + \cdots + \alpha_r f(u_r) = 0$ entonces $\alpha_1 u_1 + \cdots + \alpha_r u_r \in \text{Nuc}(f) =< u_{r+1}, \dots, u_n >$ y como $< u_1, \dots, u_r > \cap < u_{r+1}, \dots, u_n >= \{0\}$, entonces $\alpha_1 u_1 + \cdots + \alpha_r u_r = 0$ y por tanto $\alpha_1 = \cdots = \alpha_r = 0$. Si extendemos este conjunto a la base $\mathcal{B}' = \{f(u_1), \dots, f(u_r), v_{r+1}, \dots, v_m\}$, se tiene la $M_{\mathcal{B}',\mathcal{B}}(f)$ buscada.

Toda $A \in M_{m,n}(K)$ es equivalente a una de esta forma, con $r = \text{rang}(A)$. **Demostración:** Sea $f : K^n \to K^m$ tal que $M_{\mathcal{C}',\mathcal{C}}(f) = A$. Entonces $M_{\mathcal{B}',\mathcal{B}}(f) = M_{\mathcal{B}',\mathcal{C}'} \cdot A \cdot M_{\mathcal{C},\mathcal{B}}$.

Llamamos **matriz traspuesta** de $A = (a_{ij}) \in M_{m,n}(K)$ a la matriz $A^t = (b_{ij}) \in M_{n,m}(K)$ con $b_{ij} = a_{ji}$. Se verifica que $(A^t)^t = A$, $(\alpha A)^t = \alpha A^t$, $(A + B)^t = A^t + B^t$, $(AC)^t = C^t A^t$, y si A es invertible entonces A^t también lo es y $(A^t)^{-1} = (A^{-1})^t$. Así, si

$$B = \left(\begin{array}{c|c} I_r & 0 \\ \hline 0 & 0 \end{array}\right) \in M_{m,n}(K)$$

con $r = \text{rang}(A)$ y $A = M_{m,n}(K)$, existirán $P \in M_m(K)$ y $Q \in M_n(K)$ tales que $B = PAQ$, por lo que $Q^t A^t P^t = (PAQ)^t = B^t$ con $\text{rang}(B^t) = \text{rang}(B)$, y por tanto $\text{rang}(A) = \text{rang}(A^t)$.

2.4. Matrices elementales. Aplicaciones

Llamamos **matriz elemental** de tamaño n a toda matriz obtenida al efectuar una operación elemental (por filas o columnas) en I_n.

- $E_n(i, j)$ es la resultante de intercambiar las filas i y j, o las columnas i y j, en I_n. $E_n(i, j)^{-1} = E_n(i, j)$.
- $E_n(r[i])$ es la resultante de multiplicar por r la fila i, o la columna i, en I_n. $E_n(r[i])^{-1} = E_n(r^{-1}[i])$.
- $E_n([i] + r[j])$ es la resultante de añadir a la fila i la fila j multiplicada por r, o a la columna j la columna i multiplicada por r, en I_n. $E_n([i] + r[j])^{-1} = E_n([i] - r[j])$.

Si B se obtiene al realizar una operación elemental por filas en A y E al realizar la misma en I_m, entonces $B = EA$. Del mismo modo, si B se obtiene de aplicar una operación elemental por columnas en A y E al aplicarla a I_n, entonces $B = AE$. Así, realizar una serie de estas operaciones en una matriz equivale a multiplicarla por uno o ambos lados por un producto de matrices elementales, el cual es invertible. Dada una matriz A, para obtener las matrices P y Q tales que

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

podemos partir de

$$\left(\begin{array}{c|c} A & I_m \\ \hline I_n & \end{array}\right)$$

y realizar operaciones elementales hasta llegar a una matriz de la forma

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

$A \in M_n(K)$ es invertible cuando tiene rango precisamente n, por lo que al hacer operaciones elementales por filas para obtener una matriz escalonada reducida, esta será precisamente I_n, de forma que $(E_k \cdots E_1)A = I_n$ y por tanto $A^{-1} = E_k \cdots E_1$, de forma que A^{-1} es la matriz resultante de efectuar en I_n las mismas operaciones elementales fila que se hacen en A. A efectos prácticos, formamos la matriz $\left(\begin{array}{c|c} A & I_n \end{array}\right)$ y hacemos operaciones elementales por filas hasta llegar a $\left(\begin{array}{c|c} I_n & A^{-1} \end{array}\right)$. Por otro lado, si $A^{-1} = E_k \cdots E_1$, entonces $A = (A^{-1})^{-1} = (E_k \cdots E_1)^{-1} = E_1^{-1} \cdots E_k^{-1}$, de forma que toda matriz invertible es producto de matrices elementales.

Capítulo 3

Sistemas de ecuaciones lineales

Una **ecuación lineal** en las n **incógnitas** $x_1, \ldots, x_n$ sobre el cuerpo K es una expresión de la forma $a_1x_1 + \cdots + a_nx_n = b$, con los $a_i \in K$ (**coeficientes**) y $b \in K$ (**término independiente**). Un **sistema de** m **ecuaciones lineales** con n incógnitas sobre el cuerpo K tiene la forma

$$\left.\begin{array}{ccccccc} a_{11}x_1 & + & \ldots & + & a_{1n}x_n & = & b_1 \\ & & & & & \vdots & \\ a_{m1}x_1 & + & \ldots & + & a_{mn}x_n & = & b_m \end{array}\right\}$$

Puede expresarse matricialmente de la forma $AX = B$, donde $A = (a_{ij}) \in M_{m,n}(K)$ es la **matriz de los coeficientes**, $B = (b_i) \in M_{m,1}(K)$ es la **matriz de los términos independientes** y $X = (x_i) \in M_{n,1}(K)$ es la matriz de incógnitas. A la matriz $(A|B) \in M_{m,n+1}(K)$ se le llama **matriz ampliada** del sistema. Un sistema es **homogéneo** si $B = 0$, y a cada sistema $AX = B$ se le puede asociar el sistema homogéneo $AX = 0$.

Se llama **solución** a toda n-upla $(r_1, \ldots, r_n) \in K^n$ tal que si $R = (r_i) \in M_{n,1}(K)$ entonces $AR = B$. Un sistema es **compatible** si tiene alguna solución, **determinado** si tiene solo una e **indeterminado** si tiene más; o **incompatible** si no tiene ninguna. **Discutir** un sistema es determinar su compatibilidad, y **resolverlo** es encontrar las soluciones.

Teorema: Si un sistema $AX = B$ tiene una solución P, todas las soluciones son de la forma $P + M$, donde M es solución de $AX = 0$.

3.1. Teorema de Rouché-Frobenius

Un sistema $AX = B$ es compatible si y sólo si $\text{rang}(A) = \text{rang}(A|B)$, en cuyo caso es determinado si $\text{rang}(A) = n$. En concreto, si $k = n - \text{rang}(A) > 0$, existen $u_1, \ldots, u_k$ soluciones linealmente independientes de $AX = 0$ tales que cualquier solución del sistema es de la forma $x_0 + \lambda_1 u_1 + \cdots + \lambda_k u_k$. Decimos del sistema que **depende de** k **parámetros** $\lambda_1, \ldots, \lambda_k$ o que tiene k **grados de libertad**.

Demostración: Si tenemos la aplicación $f_A : K^n \to K^m$ tal que $A = M_{\mathcal{C}',\mathcal{C}}(f)$, entonces si $K^n =< u_1, \ldots, u_n >$ y v es el vector definido por B, el conjunto de soluciones es $f^{-1}(v) =$

$\{x \in K^n | f(x) = v\}$. Entonces,

$$\exists x \in U : f(x) = v \iff v \in \mathrm{Im}(f) \iff < f(u_1), \ldots, f(u_n), v > = < f(u_1), \ldots, f(u_n) > \iff$$
$$\iff \dim(< f(u_1), \ldots, f(u_n), v >) = \dim(< f(u_1), \ldots, f(u_n) >) = \dim(\mathrm{Im}(f)) = \mathrm{rang}(f)$$

Por tanto $AX = B$ es compatible si y sólo si $\mathrm{rang}(A) = \mathrm{rang}(A|B)$. Por otro lado, si $f(x_0) = v$, las soluciones serán $f^{-1}(v) = \{u \in U | f(u) = v\} = x_0 + \mathrm{Nuc}(f)$. Como además $\dim(K^n) = \mathrm{rang}(f) + \dim(\mathrm{Nuc}(f))$, entonces $k := \dim(\mathrm{Nuc}(f)) = n - \mathrm{rang}(f)$, por lo que existen k soluciones linealmente independientes de $AX = 0$ (una base de $\mathrm{Nuc}(f)$). Por tanto las soluciones del sistema homogéneo serán combinaciones lineales de dicha base.

3.2. Resolución de sistemas de ecuaciones lineales. Método de Gauss

Dos sistemas de m ecuaciones lineales con n incógnitas sobre un mismo cuerpo son **equivalentes** si tienen las mismas soluciones. Si $P \in M_m(K)$, $(PA)X = PB$ es equivalente a $AX = B$, y en particular, si $E \in M_m(K)$ es una matriz elemental, $(EA)X = EB$ también lo es, por lo que al hacer operaciones elementales por filas sobre $(A|B)$ se obtiene un sistema con las mismas soluciones.

El **método de Gauss** comienza por convertir la matriz ampliada a una escalonada reducida por filas $(A'|B')$. Si obtenemos que $\mathrm{rang}(A|B) > \mathrm{rang}(A)$, el sistema es incompatible. Si $r = \mathrm{rang}(A) = \mathrm{rang}(A|B)$, las filas nulas de A' lo son de B'. Reordenamos las incógnitas, lo que equivale a reordenar las columnas de A', para conseguir un sistema de la forma

$$\left(\begin{array}{c|c} I_r & C \\ \hline 0 & 0 \end{array}\right) \begin{pmatrix} y_1 \\ \vdots \\ y_r \\ \hline y_{r+1} \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} b'_1 \\ \vdots \\ b'_r \\ \hline 0 \\ \vdots \\ 0 \end{pmatrix}$$

Donde $y_1, \ldots, y_n$ son los $x_1, \ldots, x_n$ reordenados de la misma forma que las columnas. Esto equivale a

$$I_r \begin{pmatrix} y_1 \\ \vdots \\ y_r \end{pmatrix} + C \begin{pmatrix} y_{r+1} \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} b'_1 \\ \vdots \\ b'_r \end{pmatrix} \Longrightarrow \begin{pmatrix} y_1 \\ \vdots \\ y_r \end{pmatrix} = \begin{pmatrix} b'_1 \\ \vdots \\ b'_r \end{pmatrix} - C \begin{pmatrix} y_{r+1} \\ \vdots \\ y_n \end{pmatrix}$$

De modo que al asignar valores arbitrarios a $y_{r+1}, \ldots, y_n$, que llamamos **incógnitas libres**, obtenemos valores de $y_1, \ldots, y_r$, que llamamos **incógnitas principales**.

Capítulo 4

Determinantes

4.1. Determinante de una matriz. Propiedades

Una aplicación $f : U_1 \times \cdots \times U_n \to V$ es una **aplicación multilineal** si es lineal en cada una de las n variables, es decir, si

$$f(u_1, \ldots, \alpha u_i + \beta u_i', \ldots, u_n) = \alpha f(u_1, \ldots, u_i, \ldots, u_n) + \beta f(u_1, \ldots, u_i', \ldots, u_n)$$

Una aplicación multilineal $f : U^n \to V$ se llama **aplicación n-lineal**. Si además $V = K$ es una **forma n-lineal**. Una forma n-lineal $f : U^n \to K$ es **alternada** si se anula en cada n-upla con dos componentes iguales, es decir, tal que $f(u_1, \ldots, u_k, \ldots, u_l, \ldots, u_n) = 0$ cuando $u_k = u_l$ (con $k \neq l$).

Una **aplicación determinante** $\det : M_n(K) \to K$ es una forma n-lineal alternada que a cada matriz cuadrada A le asigna un escalar, llamado **determinante** de A, que denotamos $\det(A)$, $|A|$ o $\det(A_1, \ldots, A_n)$ (donde A_i son las columnas de A), tal que $|I_n| = 1$. Algunas aplicaciones determinantes son:

1. La aplicación $|| : M_2(K) \to K$ dada por

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

2. La **regla de Sarrus**, aplicación $|| : M_3(K) \to K$ dada por

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

Las aplicaciones determinantes verifican que:

1.

$$\begin{vmatrix} a_1 & 0 & \cdots & 0 \\ 0 & a_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_n \end{vmatrix} = a_1 a_2 \cdots a_n$$

Si $\{e_1, \dots, e_n\}$ es la base canónica de K^n,

$$\begin{vmatrix} a_1 & 0 & \cdots & 0 \\ 0 & a_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_n \end{vmatrix} = \det(a_1 e_1, \dots, a_n e_n) = a_1 \cdots a_n \det(e_1, \dots, e_n) = a_1 \cdots a_n$$

2. Si A tiene una columna nula entonces $\det(A) = 0$.
   Si $A_i = 0$, entonces

$$\begin{gathered} \det(A_1, \dots, A_{i-1}, 0, A_{i+1}, \dots, A_n) = \det(A_1, \dots, A_{i-1}, 0+0, A_{i+1}, \dots, A_n) = \\ = \det(A_1, \dots, A_{i-1}, 0, A_{i+1}, \dots, A_n) + \det(A_1, \dots, A_{i-1}, 0, A_{i+1}, \dots, A_n) \end{gathered}$$

   luego $\det A = 0$.

3. Al intercambiar dos columnas, el determinante cambia de signo.

$$\begin{gathered} 0 = \det(A_1, \dots, A_i + A_j, \dots, A_i + A_j, \dots, A_n) = \\ = \det(A_1, \dots, A_i, \dots, A_i, \dots, A_n) + \det(A_1, \dots, A_i, \dots, A_j, \dots, A_n) + \\ + \det(A_1, \dots, A_j, \dots, A_i, \dots, A_n) + \det(A_1, \dots, A_j, \dots, A_j, \dots, A_n) = \\ = \det(A_1, \dots, A_i, \dots, A_j, \dots, A_n) + \det(A_1, \dots, A_j, \dots, A_i, \dots, A_n) \end{gathered}$$

4. Si a una columna se le añade otra multiplicada por un escalar, el determinante no varía.

$$\begin{gathered} \det(A_1, \dots, A_i + \alpha A_j, \dots, A_j, \dots, A_n) = \\ = \det(A_1, \dots, A_i, \dots, A_j, \dots, A_n) + \alpha \det(A_1, \dots, A_j, \dots, A_j, \dots, A_n) = \\ = \det(A_1, \dots, A_i, \dots, A_j, \dots, A_n) \end{gathered}$$

5. Si las columnas de una matriz cuadrada son linealmente dependientes, su determinante es 0. Por tanto una matriz no invertible tiene determinante 0.
   Habrá una columna que será combinación lineal del resto: $A_k = \sum_{j \neq k} \alpha_j A_j$. Así,

$$\begin{gathered} \det(A_1, \dots, A_k, \dots, A_n) = \det(A_1, \dots, \textstyle\sum_{j \neq k} \alpha_j A_j, \dots, A_n) = \\ = \textstyle\sum_{j \neq k} \alpha_j \det(A_1, \dots, A_j, \dots, A_n) = 0 \end{gathered}$$

   Ya que cada matriz del último sumatorio tiene dos columnas iguales.

De aquí podemos deducir que $|E_n(i,j)| = -1$, $|E_n(\alpha[i])| = \alpha$ y $|E_n([i] + \alpha[j])| = 1$, y que si $A, E \in M_n(K)$, siendo E una matriz elemental, entonces $|AE| = |A||E|$. Se deducen los siguientes teoremas:

1. Una matriz cuadrada A es invertible si y sólo si $|A| \neq 0$.

$\Longrightarrow$] Toda matriz invertible es producto de matrices elementales, y por lo anterior, $|A| = |I_n E_1 \cdots E_k| = |I_n||E_1| \cdots |E_k| = |E_1| \cdots |E_k|$. Como ninguno de los factores es nulo, se tiene que $|A| \neq 0$.

$\Longleftarrow$] Inmediato de la última propiedad.

2. Si $A, B \in M_n(K)$, entonces $|AB| = |A||B|$.
   Si alguna de las dos no es invertible, su producto tampoco (pues si lo fuera, A y B serían invertibles). En tal caso, $|AB| = 0 = |A||B|$. Si son ambas invertibles, existen matrices elementales $E_1, \dots, E_k$ con $B = E_1 \cdots E_k$, por lo que $|AB| = |AE_1 \cdots E_k| = |A||E_1| \cdots |E_k| = |A||B|$.

De aquí tenemos que $|A^{-1}| = |A|^{-1}$, pues $1 = |I_n| = |AA^{-1}| = |A||A^{-1}|$. Tenemos también que la aplicación determinante es única, pues $\det(A) = 0$ para matrices no invertibles y $\det(A) = |E_1| \cdots |E_k|$ para aquellas que sí lo son, y podemos entonces comprobar que esta operación está bien definida.

Teorema: $|A^t| = |A|$. **Demostración:** Si A no es invertible, A^t tampoco, por lo que $|A^t| = 0 = |A|$. Si lo es, existen $E_1, \dots, E_k$ con $A = E_1 \cdots E_k$, por lo que $|A^t| = |(E_1 \cdots E_k)^t| = |E_k^t \cdots E_1^t| = |E_k^t| \cdots |E_1^t| = |E_1| \cdots |E_k| = |E_1 \cdots E_k| = |A|$. Esto significa que todo lo relativo a determinantes que se diga para columnas también es válido para filas.

Si $A = (a_{ij}) \in M_n(K)$ e $i, j \in \{1, \dots, n\}$, llamamos **menor complementario** del elemento a_{ij} al determinante $|A_{ij}|$ de la matriz $A_{ij} \in M_{n-1}(K)$ resultado de eliminar la fila i y la columna j de A. Llamamos **adjunto** de a_{ij} en A al escalar $\Delta_{ij} \coloneqq (-1)^{i+j}|A_{ij}|$.

Teorema: Las aplicaciones $|| : M_n(K) \to K$ definidas para $n = 1$ como $|(a)| = a$ y para $n > 1$ como $|(a_{ij})| = a_{11}\Delta_{11} + \dots + a_{1n}\Delta_{1n}$ son aplicaciones determinante. **Demostración:** Para $n = 1$ es trivial. Ahora supongamos que la aplicación determinante está definida para $n-1$ y probamos que se cumplen las condiciones para $n-1$.

1. Multilineal: Sea $A = (a_{ij}) = (A_1, \dots, A_n) \in M_n(K)$ y $A' = (a'_{ij}) = (A_1, \dots, \alpha A_k, \dots, A_n)$. Entonces $a'_{ik} = \alpha a_{ik}$ y para $j \neq k$, $a'_{ij} = a_{ij}$. Si llamamos Δ_{ij} y Δ'_{ij} a los correspondientes adjuntos, $\Delta'_{ik} = \Delta_{ik}$ y para $j \neq k$, $\Delta'_{ij} = \alpha\Delta_{ij}$. Así,

$$\begin{gathered}|A'| = a'_{11}\Delta'_{11} + \dots + a'_{1n}\Delta'_{1n} = \\ = a_{11}\alpha\Delta_{11} + \dots + a_{1(k-1)}\alpha\Delta_{1(k-1)} + \alpha a_{1k}\Delta_{ik} + a_{1(k+1)}\alpha\Delta_{i(k+1)} + \dots + a_{1n}\alpha\Delta_{in} = \\ = \alpha(a_{11}\Delta_{11} + \dots + a_{1n}\Delta_{1n}) = \alpha|A|\end{gathered}$$

   Del mismo modo, sea $A = (a_{ij}) = (A_1, \dots, A'_k + A''_k, \dots, A_n)$ y sean $A' = (a'_{ij}) = (A_1, \dots, A'_k, \dots, A_n)$ y $A'' = (a''_{ij}) = (A_1, \dots, A''_k, \dots, A_n)$. Entonces $a_{ik} = a'_{ik} + a''_{ik}$ y si $j \neq k$, $a_{ij} = a'_{ij} = a''_{ij}$. Del mismo modo, $\Delta_{ik} = \Delta'_{ik} = \Delta''_{ik}$ y si $j \neq k$, $\Delta_{ij} = \Delta'_{ij} + \Delta''_{ij}$, por lo que

$$\begin{gathered}|A| = a_{11}\Delta_{11} + \dots + a_{1n}\Delta_{1n} = \\ = a_{11}(\Delta'_{11} + \Delta''_{11}) + \dots + a_{1(k-1)}(\Delta'_{1(k-1)} + \Delta''_{1(k-1)}) + (a'_{1k} + a''_{1k})\Delta_{1k} + \\ + a_{1(k+1)}(\Delta'_{1(k+1)} + \Delta''_{1(k+1)}) + \dots + a_{1n}(\Delta'_{1n} + \Delta''_{1n}) = \\ = a'_{11}\Delta'_{11} + \dots + a_{1n}\Delta'_{1n} + a''_{11}\Delta''_{11} + \dots + a_{1n}\Delta''_{1n} = |A'| + |A''|\end{gathered}$$

2. Alternada: Sea $A = (a_{ij}) = (A_1, \dots, A_n) \in M_n(K)$. Si para $r < s$ se tiene que $A_r = A_s$, entonces $a_r = a_s$ y para $j \neq r, s$, se tiene que $\Delta_{1j} = 0$, pues el menor complementario posee dos columnas iguales, por lo que $|A| = a_{11}\Delta_{11} + \dots + a_{1n}\Delta_{1n} = a_{1r}\Delta_{1r} + a_{1s}\Delta_{1s}$. Por otro lado, si llamamos A'_j al elemento de A_j resultado de eliminar a_{1j}, entonces

$$\begin{gathered}|A_{1s}| = |(A'_1, \dots, A'_{r-1}, A'_r, A'_{r+1}, \dots, A'_{s-1}, A'_{s+1}, \dots, A'_n)| = \\ = -|(A'_1, \dots, A'_{r-1}, A'_{r+1}, A'_r, \dots, A'_{s-1}, A'_{s+1}, \dots, A'_n)| = \dots = \\ = (-1)^{s-r-1}|(A'_1, \dots, A'_{r-1}, A'_{r+1}, \dots, A'_{s-1}, A'_r, A'_{s+1}, \dots, A_n)| = (-1)^{s-r-1}|A_{1r}|\end{gathered}$$

pues $A'_r = A'_s$. Por tanto

$$|A| = a_{1r}(-1)^{1+r}|A_{1r}| + a_{1s}(-1)^{1+s}(-1)^{s-r-1}|A_{1r}| =$$
$$= a_{1r}|A_{1r}|((-1)^{1+r} + (-1)^{1+2s-r-1}) = a_{1r}|A_{1r}|((-1)^{1+r} + (-1)^{-r}) = 0$$

3. $|I_n| = \delta_{11}\Delta_{11} + \cdots + \delta_{1n}\Delta_{1n} = \Delta_{11} = (-1)^{1+1}|I_{n-1}| = 1.$

Se puede probar que, para $1 \leq i \leq n$, cada aplicación dada por $|A| = a_{i1}\Delta_{i1} + \cdots + a_{in}\Delta_{in}$, que llamamos **desarrollo del determinante** de la matriz A por la i-ésima fila, también cumple las condiciones. Por otro lado, como $|A| = |A^t|$, también se puede desarrollar por filas. En la práctica se pueden hacer operaciones elementales para obtener ceros en una fila o columna y luego desarrollar por ella.

Determinantes de Vandermonde: Restando a cada fila la anterior por x_1, desarrollando, dividiendo lo resultante por cada elemento de la primera fila y repitiendo el proceso, se tiene que:

$$\begin{vmatrix} 1 & 1 & \cdots & 1 \\ x_1 & x_2 & \cdots & x_n \\ \vdots & \vdots & & \vdots \\ x_1^{n-1} & x_2^{n-1} & \cdots & x_n^{n-1} \end{vmatrix} = \begin{vmatrix} 1 & 1 & \cdots & 1 \\ 0 & x_2 - x_1 & \cdots & x_n - x_1 \\ \vdots & \vdots & & \vdots \\ 0 & x_2^{n-1} - x_1 x_2^{n-2} & \cdots & x_n^{n-1} - x_1 x_n^{n-2} \end{vmatrix} =$$
$$= (x_2 - x_1)\cdots(x_n - x_1)\begin{vmatrix} 1 & \cdots & 1 \\ x_2 & \cdots & x_n \\ \vdots & & \vdots \\ x_2^{n-2} & \cdots & x_n^{n-2} \end{vmatrix} = \cdots = \prod_{1\leq j<i\leq n}(x_i - x_j)$$

4.2. Desarrollo de un determinante por menores, desarrollo de Laplace

Se llama **submatriz** de A a la obtenida al eliminar determinadas filas y columnas de A. Toda matriz es submatriz de sí misma. Un **menor de orden** n es el determinante de una submatriz de tamaño $n \times n$. Si A es cuadrada, el **menor complementario** M' del menor M de A es el determinante de la matriz formada por las filas y columnas restantes. Un menor es de **clase par** o de **clase impar** según lo sea la suma de los índices de sus filas $(i_1, \dots, i_p)$ y columnas $(j_1, \dots, j_p)$. La **signatura** de un menor M es $\varepsilon(M) = (-1)^{(i_1+\cdots+i_p)+(j_1+\cdots+j_p)}$. Como $(1+\cdots+n)+(1+\cdots+n)$ es par, todo menor tiene la misma signatura que su complementario.

Llamamos χ_r al conjunto de combinaciones de r filas o columnas:

$$\chi_r = \{(i_1, \dots, i_r) \mid 1 \leq i_1 < \ldots < i_r \leq n\}$$

Si $I, J \in \chi_r$, llamamos A_{IJ} al menor determinado por las filas I y las columnas J de A. Entonces:

- Dado $I \in \chi_r$, $|A| = \sum_{J\in\chi_r} \varepsilon(A_{IJ})A_{IJ}A'_{IJ}$.
- Dado $J \in \chi_r$, $|A| = \sum_{I\in\chi_r} \varepsilon(A_{IJ})A_{IJ}A'_{IJ}$.

Esto es útil cuando A está formada por bloques de tamaño adecuado alguno de los cuales es nulo. De aquí se tiene que $\left|\left(\begin{array}{c|c} P & Q \\ \hline 0 & R \end{array}\right)\right| = |P||R|$.

4.3. Determinante de un endomorfismo

Si $\mathcal{B}$ y $\mathcal{B}'$ son bases de V y $f \in \mathrm{End}(V)$, entonces

$$M_{\mathcal{B}}(f) = M_{\mathcal{B}\mathcal{B}'} M_{\mathcal{B}'}(f) M_{\mathcal{B}'\mathcal{B}} = P^{-1} M_{\mathcal{B}'}(f) P$$

por lo que $|M_{\mathcal{B}}(f)| = |P|^{-1}|M_{\mathcal{B}'}(f)||P| = |M_{\mathcal{B}'}(f)|$. Así, llamamos **determinante del endomorfismo** f al de la matriz asociada a f respecto de cualquier base de V.

4.4. Matriz adjunta. Aplicación al cálculo de la inversa

Llamamos **matriz adjunta** de A a la matriz $\hat{A} = (\Delta_{ij}) \in M_n(K)$. **Teorema:** Si $A \in M_n(K)$, entonces $A \cdot \hat{A}^t = \hat{A}^t \cdot A = |A|I_n$. **Demostración:** Si $A = (a_{ij})$, $\hat{A} = (\Delta_{ij})$ y $\hat{A}^t = (b_{ij})$, entonces $b_{ij} = \Delta_{ji}$. Sea entonces $C = (c_{ij}) = A \cdot \hat{A}^t$, entonces $c_{ij} = \sum_{k=1}^n a_{ik}b_{kj} = \sum_{k=1}^n a_{ik}\Delta_{jk}$. Para $i \neq j$, esto corresponde al desarrollo por la fila j-ésima del determinante de la matriz que se diferencia de A en que tiene la fila i-ésima copiada en la j-ésima, por lo que entonces $c_{ij} = 0$. Si $i \neq j$, este es el desarrollo por la fila j-ésima de A, por lo que $c_{ii} = |A|$ y $C = |A|I_n$.

Como consecuencia, se tiene el **teorema** de que A es invertible si y sólo si $|A| \neq 0$ y entonces

$$A^{-1} = \frac{1}{|A|}\hat{A}^t$$

4.5. Cálculo del rango de una matriz por determinantes

El rango de A es el mayor de los órdenes de los menores no nulos de A. **Demostración:** Sean $A = (a_{ij}) \in M_{m,n}(K)$ con $A \neq 0$, $r = \mathrm{rang}(A)$ y p el mayor de los tamaños de los menores no nulos, que existe si $A \neq 0$. Si A' es una submatriz cuadrada de A de tamaño $p \times p$ con $|A'| \neq 0$, entonces la submatriz B formada por las filas de A' pero con todas las columnas de A tiene p columnas linealmente independientes (las de A') y por tanto también tiene p filas linealmente independientes, pero entonces A tiene al menos p filas linealmente independientes y $r \geq p$. Por otro lado, si $A_{i_1}, \dots, A_{i_r}$ son filas linealmente independientes de A y tomamos la submatriz $B \in M_{r,n}(K)$ formada por estas filas y todas las columnas, B tendrá rango r, luego tendrá r columnas $j_1, \dots, j_r$ linealmente independientes. Si tomamos la submatriz $A' \in M_r(K)$ formada por estas columnas, al ser linealmente independientes, $|A'| \neq 0$, luego $p \geq r$. Por tanto $p = r$.

Dados $A_i = (a_{1i}, \dots, a_{ni}) \in K^n$ con $A_1, \dots, A_r$ linealmente independientes y

$$\begin{vmatrix} a_{i_1 1} & \cdots & a_{i_1 r} \\ \vdots & \ddots & \vdots \\ a_{i_r 1} & \cdots & a_{i_r r} \end{vmatrix} \neq 0$$

$B = (b_1, \dots, b_n)$ es combinación lineal de $A_1, \dots, A_r$ si y sólo si para todo j,

$$\begin{vmatrix} a_{i_1 1} & \dots & a_{i_1 r} & b_{i_1} \\ \vdots & \ddots & \vdots & \vdots \\ a_{i_r 1} & \cdots & a_{i_r r} & b_{i_r} \\ a_{j1} & \cdots & a_{jr} & b_j \end{vmatrix} = 0$$

$\Longrightarrow$] Si son linealmente dependientes, los determinantes son nulos.

$\Longleftarrow$] Si todos son nulos, desarrollando por la última fila, se obtiene, para cada j, que

$$a_{j1}\begin{vmatrix} a_{i_1 2} & \cdots & a_{i_1 r} & b_{i_1} \\ \vdots & \ddots & \vdots & \vdots \\ a_{i_r 2} & \cdots & a_{i_r r} & b_{i_r} \end{vmatrix} \pm \cdots \pm b_j \begin{vmatrix} a_{i_1 1} & \cdots & a_{i_1 r} \\ \vdots & \ddots & \vdots \\ a_{i_r 1} & \cdots & a_{i_r r} \end{vmatrix} = 0$$

Por lo que

$$b_j = \frac{1}{\begin{vmatrix} a_{i_1 1} & \cdots & a_{i_1 r} \\ \vdots & \ddots & \vdots \\ a_{i_r 1} & \cdots & a_{i_r r} \end{vmatrix}} \left(\pm a_{j1} \begin{vmatrix} a_{i_1 2} & \cdots & b_{i_1} \\ \vdots & \ddots & \vdots \\ a_{i_r 2} & \cdots & b_{i_r} \end{vmatrix} \pm \cdots \pm a_{j_r} \begin{vmatrix} a_{i_1 1} & \cdots & b_{i1} \\ \vdots & \ddots & \vdots \\ a_{i_r 1} & \cdots & b_{i_r} \end{vmatrix} \right)$$

A efectos prácticos, esto significa que, una vez encontrado un menor no nulo de orden k en una matriz A, podemos *orlarlo* (obtener otro añadiendo una fila y una columna a la submatriz) de todas las formas posibles y, si todos los menores resultantes son nulos, entonces $\operatorname{rang}(A) = k$.

4.6. Regla de Cramer

Un sistema de ecuaciones lineales $AX = B$ es un **sistema de Cramer** si A es invertible. En tal caso tiene solución única $X = A^{-1}B$. **Regla de Cramer:** si las columnas de A son $(A_1, \ldots, A_n)$, entonces

$$x_i = \frac{\det(A_1, \ldots, A_{i-1}, B, A_{i+1}, \ldots, A_n)}{\det(A)}$$

Demostración: $A^{-1} = \frac{1}{|A|}\hat{A}^t$, y si $X = (x_i)$, $A = (a_{ij})$ y $\hat{A} = (\Delta_{ij})$, entonces $x_i = \sum_{j=1}^n \frac{1}{|A|}\Delta_{ji} b_j = \frac{1}{|A|}\sum_{j=1}^n \Delta_{ji} b_j$.

Si $A \in M_{m,n}(K)$ con $\operatorname{rang}(A) = r$, habrá un menor $M \neq 0$ de orden r, por lo que las $n - r$ últimas filas serán combinaciones lineales de las r primeras, y moviendo al lado derecho los $m - r$ coeficientes que no están en la submatriz de M, nos queda el sistema

$$\left.\begin{array}{rcl} a_{11}x_1 + \cdots + a_{1r}x_r & = & b_1 - (a_{1r+1}x_{r+1} + \cdots + a_{1n}x_n) \\ & \vdots & \\ a_{r1}x_1 + \cdots + a_{rr}x_r & = & b_r - (a_{rr+1}x_{r+1} + \cdots + a_{rn}x_n) \end{array}\right\}$$

que podemos resolver por Cramer.

Capítulo 5

Diagonalización de endomorfismos

5.1. Semejanza de matrices

$A, B \in M_n(K)$ son **semejantes** si $\exists P \in M_n(K) : B = P^{-1}AP$. Esta relación es de equivalencia, y si dos matrices son semejantes también son equivalentes y por tanto tienen el mismo rango, si bien el recíproco no se cumple.

Sea $A \in M_n(K)$ una matriz formada por **bloques** cuadrados en la diagonal y ceros en el resto:

$$A = \begin{pmatrix} \boxed{A_1} & & 0 \\ & \ddots & \\ 0 & & \boxed{A_t} \end{pmatrix}$$

con $A_i \in M_{n_i}(K)$ y $n_1 + \cdots + n_t = n$. Por el desarrollo de Laplace, su determinante es $|A| = |A_1| \cdots |A_t|$; su rango es la suma de los rangos de los A_i, y su potencia k-ésima es

$$A^k = \begin{pmatrix} \boxed{A_1^k} & & 0 \\ & \ddots & \\ 0 & & \boxed{A_t^k} \end{pmatrix}$$

Entonces, si B es semejante a A, $|B| = |P^{-1}AP| = |P|^{-1}|A||P| = |A|$, su rango es el de A y

$$B^k = \underbrace{(P^{-1}AP) \cdots (P^{-1}AP)}_{k \text{ veces}} = P^{-1}A^kP$$

5.2. Subespacios invariantes

Dado $f \in \mathrm{End}_K(V)$ y $W \leq V$, W es **invariante** por f si $f(W) \subseteq W$. Entonces la restricción de f a W es $f|_W \in \mathrm{End}_K(W)$. También se tiene que $\{0\}$ y V son invariantes de cada $f \in \mathrm{End}_K(V)$, y la suma e intersección de subespacios invariantes por f también son subespacios invariantes por f. Ahora supongamos que $V = W_1 \oplus \cdots \oplus W_t$, donde cada W_i es un subespacio invariante no nulo por $f \in \mathrm{End}(V)$. Si tomamos $\mathcal{B} = \mathcal{B}_1 \cup \cdots \cup \mathcal{B}_t$, siendo cada $\mathcal{B}_i$ una base de W_i, entonces

$$M_{\mathcal{B}}(f) = \begin{pmatrix} \boxed{A_1} & & 0 \\ & \ddots & \\ 0 & & \boxed{A_t} \end{pmatrix}$$

donde $A_i = M_{\mathcal{B}_i}(f|_W) \in M_{n_i}(K)$ para $n_i = \dim(W_i)$. Recíprocamente, si $M_{\mathcal{B}}(f)$ tiene dicha forma y

$$\mathcal{B} = \{v_1, \ldots, v_{n_1}, v_{n_1+1}, \ldots, v_{n_1+n_2}, \ldots, v_{n_1+\cdots+n_{t-1}+1}, \ldots, v_{n_1+\cdots+n_t}\}$$

entonces $W_1 =< v_1, \ldots, v_{n_1} >$, $W_2 =< v_{n_1+1}, \ldots, v_{n_1+n_2} >$, etc. son subespacios vectoriales invariantes por f.

5.3. Endomorfismos y matrices diagonalizables

$f \in \mathrm{End}_K(V)$ es **diagonalizable** si existe una base $\mathcal{B}$ de V tal que $M_{\mathcal{B}}(f)$ es diagonal, y una matriz cuadrada es diagonalizable si lo es el endomorfismo de K^n que cuya matriz respecto a la base canónica es A. Equivalentemente, una matriz cuadrada es diagonalizable si y sólo si es semejante a una matriz diagonal, y un endomorfismo es diagonalizable si su matriz asociada respecto a cualquier base lo es. Denotamos las matrices diagonales como

$$\begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix} = \mathrm{Diag}(\lambda_1, \ldots, \lambda_n)$$

5.4. Vectores y valores propios

Si $\mathcal{B} = \{u_1, \ldots, u_n\}$ y $M_{\mathcal{B}}(f) = \mathrm{Diag}(\lambda_1, \ldots, \lambda_n)$, se tiene que $f(u_i) = \lambda_1 u_1$. Así:

- Un **vector propio**, **autovector** o **vector característico** de f es un vector $v \neq 0$ para el que existe un $\lambda \in K$ con $f(v) = \lambda v$.

- Un **valor propio**, **autovalor** o **valor característico** de f es un escalar $\lambda \in K$ para el que existe un $v \in V \setminus \{0\}$ tal que $f(v) = \lambda v$. Decimos que λ es ***el* valor propio asociado al vector propio** v, o que v es ***un* vector propio asociado al valor propio** λ.

Así, $f \in \mathrm{End}(V)$ es diagonalizable si y sólo si existe una base de V formada por vectores propios de f.

5.5. Subespacios propios. Polinomio característico

Los vectores propios de f asociados a λ son todos los vectores no nulos de $\mathrm{Nuc}(f - \lambda Id)$. Así, $V_\lambda = \mathrm{Nuc}(f - \lambda Id) = \{v \in V \mid (f - \lambda Id)(v) = 0\} = \{v \in V \mid f(v) = \lambda v\}$ es el **subespacio propio** o **característico** correspondiente al valor propio λ. Así, $\lambda \in K$ es un valor propio de f

si y sólo si $\det(f-\lambda Id)=0$. **Demostración:** $\lambda \in K$ es valor propio si y sólo si existe $0 \neq v \in V_\lambda$, es decir, si $\mathrm{Nuc}(f-\lambda Id) \neq \{0\}$, pero entonces $0 < \dim(\mathrm{Nuc}(\lambda Id - f)) = \dim(V) - \mathrm{rang}(\lambda Id - f)$, es decir, $\mathrm{rang}(\lambda Id - f) < \dim(V)$ o, equivalentemente, $\det(\lambda Id - f) = 0$.

$P_f(x) \coloneqq \det(xId - f)$ es el **polinomio característico** de f, y $P_A(x) \coloneqq \det(xI_n - A)$ es el polinomio característico de A. Podemos comprobar que

$$P_A(x) = x^n - \mathrm{tr}(A)x^{n-1} + \cdots + (-1)^n \det(A)$$

donde $\mathrm{tr}(A)$ es la **traza** de A, la suma de los elementos de su diagonal. Obtenemos como resultado que los valores propios de $f \in \mathrm{End}(V)$ son las raíces de $P_f(x)$, y que f tiene a lo sumo $\dim(V)$ valores propios distintos.

5.6. Independencia de los subespacios propios

Si $\lambda_1, \ldots, \lambda_s$ son valores propios de f distintos dos a dos, entonces $\mathrm{Nuc}(\lambda_1 Id - f) + \cdots + \mathrm{Nuc}(\lambda_s Id - f)$ es suma directa, y en particular, vectores propios correspondientes a valores propios distintos dos a dos son linealmente independientes. **Demostración:** Para $s=2$, sean $v_1 \in \mathrm{Nuc}(\lambda_1 Id - f)$ y $v_2 \in \mathrm{Nuc}(\lambda_2 Id - f)$ con $0 = v_1 + v_2$. Entonces $0 = f(0) = f(v_1 + v_2) = f(v_1) + f(v_2) = \lambda_1 v_1 + \lambda_2 v_2$, pero también $0 = \lambda_2(v_1+v_2) = \lambda_2 v_1 + \lambda_2 v_2$. Restando, $0 = (\lambda_2 - \lambda_1)v_1$, pero como $\lambda_2 \neq \lambda_1$, entonces $v_1 = 0$ y por tanto v_2 también. Ahora sea $s > 2$ y supongamos el resultado cierto para $s-1$. Sean ahora $v_1 \in \mathrm{Nuc}(\lambda_1 Id - f), \ldots, v_s \in \mathrm{Nuc}(\lambda_s Id - f)$ con $0 = v_1 + \cdots + v_s$, entonces $0 = f(0) = f(v_1 + \cdots + v_s) = f(v_1) + \cdots + f(v_s) = \lambda_1 v_1 + \cdots + \lambda_s v_s$, pero también $0 = \lambda_s v_1 + \cdots + \lambda_s v_{k-1} + \lambda_s v_s$. Restando, $0 = (\lambda_s - \lambda_1)v_1 + \cdots + (\lambda_s - \lambda_{s-1})v_{s-1}$. Aplicando la hipótesis de inducción, queda que $v_1 = \cdots = v_{s-1} = 0$, luego $v_s = 0$.

También, si $f \in \mathrm{End}(V)$ tiene $\dim(V)$ autovalores, entonces es diagonalizable. **Demostración:** Si $\lambda_1, \ldots, \lambda_n$ son valores propios de f distintos dos a dos y $v_1, \ldots, v_n$ son vectores propios asociados a cada uno, entonces $v_1, \ldots, v_n$ son n vectores linealmente independientes en un espacio de dimensión n, por lo que constituyen una base formada por vectores propios de f.

5.7. Caracterización de los endomorfismos diagonalizables

Si $P(x)$ es un polinomio con coeficientes en K y $\lambda \in K$ es una raíz de $P(x)$, entonces λ tiene **multiplicidad** m en $P(x)$ si $(x-\lambda)^m | P(x)$ pero $\neg((x-\lambda)^{m+1} | P(x))$. Si una raíz tiene multiplicidad 1, es una raíz **simple**. De lo contrario es una raíz **múltiple**.

Dado $f \in \mathrm{End}_K(V)$ y λ un valor propio de f, si $d = \dim(\mathrm{Nuc}(\lambda Id - f))$ y m es la multiplicidad de λ en $P_f(x)$, entonces $d \leq m$. En particular, si el valor propio es una raíz simple, entonces $d = m = 1$. **Demostración:** Sea $\{v_1, \ldots, v_d\}$ una base de $\mathrm{Nuc}(\lambda Id - f)$ y

$\mathcal{B} = \{v_1, \dots, v_d, v_{d+1}, \dots, v_n\}$ una base de V. Entonces $M_{\mathcal{B}}(f)$ tiene forma

$$\left(\begin{array}{ccc|c} \lambda & & 0 & \\ & \ddots & & C \\ 0 & & \lambda & \\ \hline & 0 & & D \end{array}\right)$$

por lo que $P_f(x) = (x-\lambda)^d \det(xI_{n-d} - D) = (x-\lambda)^d Q(x)$, por lo que $d \leq m$.

Teorema de diagonalización: f es diagonalizable si y sólo si

$$P_f(x) = (x-\lambda_1)^{d_1} \cdots (x-\lambda_r)^{d_r}$$

con $\lambda_1, \dots, \lambda_r \in K$ distintos dos a dos, y $d_i = \dim(\text{Nuc}(\lambda_i Id - f))$.

$\Longrightarrow$] Sean $\lambda_1, \dots, \lambda_r$ los valores propios de f, existirá una base $\mathcal{B}$ de vectores propios en los que cada vector tendrá asociado un valor propio y pertenecerá por tanto al subespacio propio correspondiente. Agrupando, $M_{\mathcal{B}}(f)$ tendrá forma

$$\begin{pmatrix} \lambda_1 & & & & & & \\ & \ddots & & & 0 & & \\ & & \lambda_1 & & & & \\ & & & \ddots & & & \\ & & & & \lambda_r & & \\ & 0 & & & & \ddots & \\ & & & & & & \lambda_r \end{pmatrix}$$

donde cada λ_i se repetirá m_i veces, el número de vectores propios de la base del subespacio. Por tanto, $P_f(x) = (x-\lambda_1)^{m_1} \cdots (x-\lambda_r)^{m_r}$ tiene todas sus raíces en K. Además, si $d_i = \dim(\text{Nuc}(\lambda_i Id - f))$, se tiene que $\sum d_i = \dim(\text{Nuc}(\lambda_1 Id - f) \oplus \cdots \oplus \text{Nuc}(\lambda_r Id - f)) \leq \dim(V) = n$, y como en la base hay m_i vectores linealmente independientes de $\text{Nuc}(\lambda_i Id - f)$, entonces $m_i \leq d_i$, luego $n = \text{gr}(P_f(x)) = \sum m_i \leq \sum d_i \leq n$ y por tanto $m_i = d_i$.

$\Longleftarrow$] Si $P_f(x) = (x-\lambda_1)^{d_1} \cdots (x-\lambda_r)^{d_r}$ con $d_i = \dim(\text{Nuc}(f - \lambda Id))$, entonces $\dim(\text{Nuc}(\lambda_1 Id - f) \oplus \cdots \oplus \text{Nuc}(\lambda_r Id - f)) = d_1 + \cdots + d_r = \text{gr}(P_f(x)) = \dim(V)$, luego $V = \text{Nuc}(f - \lambda_1 Id) \oplus \cdots \oplus \text{Nuc}(f - \lambda_r Id)$ y la unión de las bases de cada subespacio será una base de V formada por vectores propios.

Así, para diagonalizar una matriz $A \in M_n(K)$ en matrices $A = M_{\mathcal{CB}} D M_{\mathcal{BC}}$, con D diagonal, obtenemos su polinomio característico, hallamos sus raíces, que serán los autovalores de A. Si la suma de sus multiplicidades da n, resolvemos cada ecuación $(\lambda Id - f)X = 0$ para obtener las bases de los subespacios propios, cuya dimensión debería coincidir con la multiplicidad del autovalor si A es diagonalizable. Entonces añadimos cada raíz en D tantas veces como sea su multiplicidad y razonamos que los vectores correspondientes de la base $\mathcal{B}$, y por tanto las correspondientes columnas de $M_{\mathcal{CB}}$, son los de la base de dicho subespacio propio.

5.8. Aplicaciones

$(x_n)_n \subseteq K$ verifica una **ecuación en diferencias lineales con coeficientes constantes** (homogénea) si para todo n satisface que $x_{n+r} + a_1 x_{n+r-1} + \cdots + a_r x_n = 0$. Llamamos a r el **orden** de la ecuación. Podemos definir entonces una sucesión auxiliar $(Y_n)_n \subseteq M_{r,1}(K)$ con $(Y_n)_i = x_{n+r-i}$. Se tiene entonces que $x_{n+r} = -a_1 x_{n+r-1} - \cdots - a_r x_n$, luego

$$Y_{n+1} = \begin{pmatrix} x_{n+r} \\ \vdots \\ x_{n+1} \end{pmatrix} = \begin{pmatrix} -a_1 x_{n+r-1} - \cdots - a_r x_n \\ x_{n+r-1} \\ \vdots \\ x_{n+1} \end{pmatrix} =$$

$$= \begin{pmatrix} -a_1 & -a_2 & \cdots & -a_r \\ 1 & & 0 & 0 \\ & \ddots & & \vdots \\ 0 & & 1 & 0 \end{pmatrix} \begin{pmatrix} x_{n+r-1} \\ \vdots \\ x_n \end{pmatrix} = AY_n$$

Un **sistema de ecuaciones en diferencias lineales de primer orden con coeficientes constantes** (homogéneo) es una relación entre los términos de unas sucesiones y sus términos inmediatamente anteriores:

$$\left.\begin{array}{rcl} x_{n+1} & = & a_{11}x_n + a_{12}y_n + a_{13}z_n \\ y_{n+1} & = & a_{21}x_n + a_{22}y_n + a_{23}z_n \\ z_{n+1} & = & a_{31}x_n + a_{32}y_n + a_{33}z_n \end{array}\right\}$$

Estos pueden expresarme matricialmente de la forma $Y_{n+1} = AY_n$ con $A = (a_{ij})$. Por recurrencia, en ambos casos se tiene que $Y_n = A^{n-1}Y_1 = A^n Y_0$. Entonces es útil diagonalizar A, si es posible, para poder calcular las potencias rápidamente.