Algebra Lineal

Miguel González
mgonzalez.contacto@gmail.com
miguelgg.com

Mayo 2018

$$\begin{tikzcd} G \arrow[r, "f"] \arrow[d, "\pi"'] & H \\ G/\ker(f) \arrow[ur, dashed, "\tilde{f}"'] & \end{tikzcd}$$

Revisado en 2020
Apuntes de la asignatura impartida por Orlando Villamayor
en la Universidad Autónoma de Madrid en Mayo 2018.

Acerca de este documento

Estos apuntes son una versión revisada de los de la asignatura Algebra Lineal del grado en matemáticas, tomados en Mayo 2018 por Miguel González. La asignatura fue impartida por Orlando Villamayor. A los apuntes originales se les ha añadido esta página, una imagen de portada, y breves párrafos explicativos en las zonas menos completas. Asimismo se han revisado las erratas y completado los contenidos faltantes.

Este documento es:

- Una recopilación ordenada y directa de las definiciones y resultados más importantes del tema en cuestión, al nivel de los estudios de grado.
- Una colección de demostraciones completas de dichos resultados (salvo en los casos más básicos).
- Una *guía* para revisar de manera rápida las ideas que se han adquirido previamente, o para consultar enunciados puntuales que puedan no haberse comprendido en su totalidad.

Este documento NO es:

- Un libro de texto de la asignatura.
- Una colección de ejercicios para practicar los conceptos adquiridos.
- Un listado de ejemplos para ilustrar las ideas tratadas. A pesar de ello, en ocasiones se incluyen ejemplos puntuales que puedan ser de especial interés o curiosidad, pero se intentan reducir al mínimo en virtud del primer punto de la lista anterior.

Sobre Algebra Lineal

El álgebra lineal se centra en el estudio de los espacios vectoriales y las transformaciones entre ellos, motivado en un inicio por la resolución de ecuaciones lineales homogéneas. En esta asignatura se introduce por primera vez de manera rigurosa una estructura algebraica, el espacio vectorial, y se estudian los morfismos entre distintas ocurrencias de esta estructura. Se presentan conceptos importantísimos en el álgebra, como el concepto de isomorfía o el de núcleo.

Requisitos previos

1. Conocimientos básicos de notación matemática.
2. Nociones básicas de teoría de conjuntos: inyectividad, sobreyectividad, biyectividad y conjunto cociente, entre otros.

Índice

1. Sistemas de ecuaciones lineales

Para motivar la introducción del espacio vectorial como estructura algebraica, vamos a discutir los sistemas de ecuaciones lineales y algunos resultados sobre su resolución, que quizá ya sean conocidos al lector.

Definición 1 (Ecuación lineal)**.** Una **ecuación lineal** en un cuerpo $\mathbb{K}$ (como por ejemplo $\mathbb{R}$) es una expresión de la forma:

$$a_1x_1 + a_2x_2 + ... + a_nx_n = b$$

con cada $a_i \in \mathbb{K}$ y $b \in \mathbb{K}$. Una *n-tupla* $(c_1, c_2, ..., c_n)$ de elementos de $\mathbb{K}$ es **solución** de la ecuación si la igualdad se verifica asignando el valor de cada c_i a la incógnita x_i correspondiente.

En $\mathbb{R}^n$, existen tres posibles casos con respecto al conjunto de soluciones de una ecuación. Si $a_1 = a_2 = ... = a_n = 0$ y $b = 0$, entonces dicho conjunto es todo $\mathbb{R}^n$. Si, en el caso anterior, por el contrario, $b \neq 0$, ningún valor satisface la ecuación y por lo tanto $\emptyset$ es el conjunto de soluciones. Si alguno de los coeficientes no es 0, el conjunto de soluciones es un subconjunto de $\mathbb{R}^n$.

Definición 2 (Sistema de ecuaciones lineales)**.** Un sistema de m ecuaciones lineales de n incógnitas en un cuerpo $\mathbb{K}$ es un conjunto de m ecuaciones lineales:

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

Una *n-tupla* $(c_1, c_2, ..., c_n)$ de n elementos de $\mathbb{K}$ es solución del sistema si es solución de cada una de las ecuaciones simultáneamente.

La **matriz de coeficientes** del sistema es la matriz que recoge los escalares que acompañan a cada incógnita:

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

La **matriz asociada** es la matriz A con la columna de términos independientes:

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

Esta última aporta toda la información del sistema y es equivalente trabajar con ella.

Un sistema es **homogéneo** si $b_1 = b_2 = ... = b_m = 0$, en cuyo caso la asignación $x_1 = x_2 = ... = x_n = 0$ es siempre solución. Para todo sistema, una propiedad interesante que podremos abordar con las herramientas del álgebra lineal, es que se verifica que si se reemplazan sus términos independientes por 0 (**sistema homogéneo equivalente**), y se suma a cada una de las soluciones de este nuevo sistema una solución fija del primero, se obtienen todas las soluciones del mismo.

Para obtener las soluciones de un sistema de ecuaciones, se realizan una serie de operaciones elementales que permiten alterar el sistema (o la matriz asociada) sin variar el conjunto de soluciones, para obtener expresiones que permitan *verlas* de manera explícita.

Proposición 1 (Operaciones elementales)**.** *Las siguientes operaciones elementales no alteran el conjunto de soluciones:*

1. *Intercambiar 2 ecuaciones de posición.*
2. *Multiplicar una ecuación por $\lambda \in \mathbb{K}$, $\lambda \neq 0$.*
3. *Sumar a una ecuación otra multiplicada por un escalar $\lambda \in \mathbb{K}$.*

Demostración. 1 es trivial. Para 2, veamos que si la ecuación en cuestión es la *i-ésima*, $a_{i1}c_1 + a_{i2}c_2 + ... + a_{in}c_n = b_i \iff \lambda(a_{i1}c_1 + a_{i2}c_2 + ... + a_{in}c_n) = \lambda b_i$, multiplicando por λ o λ^{-1} según convenga. El resto de ecuaciones se mantienen, luego las soluciones son las mismas. Para 3, veamos que si a la fila *i-ésima* se suma la *j-ésima* escalada, como $a_{i1}c_1 + a_{i2}c_2 + ... + a_{in}c_n = b_i$ y $a_{j1}c_1 + a_{j2}c_2 + ... + a_{jn}c_n = b_j \implies \lambda(a_{j1}c_1 + a_{j2}c_2 + ... + a_{jn}c_n) = \lambda b_j$, sigue el resultado sumando ambas expresiones (sumar elementos iguales da elementos iguales). A la inversa, si se tiene que $\sum_{k=1}^{n}(a_ik + \lambda a_jk)x_k = b_i$ y $a_{j1}c_1 + a_{j2}c_2 + ... + a_{jn}c_n = b_j$, basta con multiplicar la segunda expresión por $-\lambda$ y sumar ambas. □

Gracias a esto se puede llevar a cabo el **método de Gauss**, que consiste en aplicar estas transformaciones a la matriz $A|B$ de un sistema para dejar inalterado el conjunto de soluciones, hasta obtener una matriz que permita ver las soluciones:

Definición 3 (Matriz escalonada reducida)**.** Se dice que una matriz $A \in \mathcal{M}_{mn}(\mathbb{K})$ es **escalonada reducida** por filas si se verifica:

1. Todas las filas nulas se encuentran en la parte inferior de la matriz.
2. El primer elemento no nulo de cada fila no nula es el 1.
3. El primer elemento no nulo de cada fila aparece abajo y a la derecha del primer elemento no nulo de la fila anterior.
4. Cada columna que contenga al primer elemento no nulo de una fila contiene al 0 en el resto de posiciones.

Por ejemplo, la siguiente matriz es escalonada reducida:

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

El primer elemento no nulo de cada fila, o **pivote**, es de gran importancia a la hora de resolver sistemas, puesto que la variable con que se asocia se puede despejar en función del resto, que además no son pivotes de ninguna otra fila. Cada una de estas variables asociadas con el pivote se llama **variable principal** o **fija**, y el resto son **variables libres**. Por ejemplo, en el sistema de la matriz anterior, sigue que $x_1 = 2 - 2x_3$, $x_2 = 7 - 3x_3$ y $x_4 = 0$. Así, las soluciones son las tuplas de la forma $(2 - 2\lambda, 7 - 3\lambda, \lambda, 0)$ con $\lambda \in \mathbb{R}$.

Podemos resumir esta idea en:

Proposición 2 (Método de Gauss)**.** *Para resolver sistemas de ecuaciones lineales:*

1. *Se transforma la matriz asociada en escalonada reducida a través de operaciones elementales para no alterar el conjunto de soluciones. Para ello, primero se obtiene una matriz escalonada (que cumple las propiedades 1-3) y luego se reduce con los pivotes.*

2. *Se despejan las variables principales (asociadas a pivotes) en función de las libres (no asociadas a pivotes). Esto proporciona las soluciones permitiendo cualquier valor real en las variables libres.*

Cabe destacar que en un sistema de m ecuaciones con n incógnitas, al llevar a cabo este método, podemos encontrar distintos casos:

Observación 1. Sea p el número de pivotes de la matriz de coeficientes (equivalentemente, el número de filas no nulas). Entonces $p \leq n$, dado que cada pivote está abajo y a la derecha del anterior.

Observación 2. Sea $\hat{p}$ el número de pivotes de la matriz asociada ampliada. Entonces $p \leq \hat{p}$, dado que contiene los mismos pivotes, o bien uno más en la columna añadida.

Teorema 1 (Rouché-Frobenius)**.** *Sea $A \in \mathcal{M}_{mn}(\mathbb{K})$ la matriz de coeficientes de un sistema de n incógnitas y $A|B$ la matriz asociada ampliada con los términos independientes, tal que $A|B$ está en forma escalonada reducida por filas. Además, sea p el número de pivotes de A y $\hat{p}$ el número de pivotes de $A|B$. Entonces:*

- $p < \hat{p} \iff$ *el sistema es incompatible (El conjunto de soluciones es $\emptyset$)*
- $p = \hat{p} = n \iff$ *el sistema es compatible determinado (Solo hay una solución)*
- $p = \hat{p} < n \iff$ *el sistema es compatible indeterminado (Hay infinitas soluciones)*

Demostración. Si $p < \hat{p}$ es porque la columna B alberga un pivote, de modo que el resto de la fila es nula, así que se tiene la ecuación $0 = 1$. Si $p = \hat{p} = n$, todas las columnas de A tienen pivotes y por tanto todas las variables son principales, de modo que no hay variables libres y la solución es única. Si $p = \hat{p} < n$, hay variables no principales ya que el número de pivotes es menor al número de variables. $\square$

2. El espacio vectorial $\mathbb{R}^n$

En esta sección, antes de pasar a la introducción del espacio vectorial como estructura algebraica general, vamos a *jugar* con algunas de las propiedades que se cumplen en $\mathbb{R}^n$, que es el espacio más sencillo de imaginar (para $n = 2$ o $n = 3$ uno puede incluso visualizarlo como un plano o el espacio). El objetivo es que después no nos encontremos perdidos al dar el salto al concepto abstracto.

2.1. Los vectores de $\mathbb{R}^n$

Los elementos de $\mathbb{R}^n$ se denominarán vectores. A continuación se verán las principales operaciones con vectores. En primer lugar, las que dotan a $\mathbb{R}^n$ de las propiedades de un espacio vectorial sobre $\mathbb{R}$:

Definición 4 (Suma y producto por escalar)**.** Sean $\vec{x}, \vec{y} \in \mathbb{R}^n$. Se define el vector suma por $\vec{x} + \vec{y} = (x_1 + y_1, x_2 + y_2, ..., x_n + y_n)$, donde x_i y y_i son las coordenadas de $\vec{x}$ y $\vec{y}$, respectivamente.

Sea $\vec{x} \in \mathbb{R}^n$ y $\lambda \in \mathbb{R}$. Se define el producto como $\lambda\vec{x} = (\lambda x_1, \lambda x_2, ..., \lambda x_n)$.

Definición 5 (Combinación lineal)**.** Se dice que $w \in \mathbb{R}^n$ es combinación lineal de los vectores $\{v_1, v_2, ..., v_r\}$ si $w = \lambda_1 v_1 + \lambda_2 v_2 + ... + \lambda_r v_r$ para algunos escalares $\lambda_i \in \mathbb{R}$.

Si fijamos r vectores en $\mathbb{R}^n$ y definimos la matriz $A \in \mathcal{M}_{nr}(\mathbb{R})$ como la matriz que tiene por columnas las coordenadas de los vectores, y tenemos $w \in \mathbb{R}^n$, son equivalentes:

1. w es combinación lineal de $\{v_1, v_2, ..., v_r\}$
2. El sistema en $\mathbb{R}^r$ asociado a la matriz $A|W$, donde W es la columna de coordenadas de w, tiene solución.

En efecto, se trata del sistema que surge de imponer que $w = \lambda_1 v_1 + \lambda_2 v_2 + ... + \lambda_r v_r$, con las r incógnitas los escalares.

Definición 6 (Independencia lineal)**.** Se dice que un conjunto de vectores $\{v_1, v_2, ..., v_r\} \subset \mathbb{R}^n$ es **linealmente independiente** (o que los vectores son **linealmente independientes**) si y solo si la única combinación lineal del vector $0 := (0, 0, ..., 0)$ con ellos, es decir, $0 = \lambda_1 v_1 + \lambda_2 v_2 + ... + \lambda_r v_r$, es aquella en la que $\lambda_1 = \lambda_2 = ... = \lambda_r = 0$.

Es decir, son linealmente independientes si $0 = \lambda_1 v_1 + \lambda_2 v_2 + ... + \lambda_r v_r \iff \lambda_1 = \lambda_2 = ... = \lambda_r = 0$. (Nótese que la implicación $\Longleftarrow$ siempre se verifica, la interesante en este caso es $\Longrightarrow$).

En caso contrario, son **linealmente dependientes**.

Es fácil observar que son linealmente dependientes si y solo si algún vector se puede expresar como combinación lineal de los demás. (Basta con tomar aquel cuyo escalar no es cero, despejarlo y multiplicar por el inverso del escalar).

Observación 3. $0 \in \mathbb{R}^n$ es combinación lineal de cualquier conjunto de vectores.

En efecto, basta con tomar todos los escalares 0, y como $0v = 0 \ \forall v \in \mathbb{R}^n$ y $0+0 = 0$, sigue el resultado.

Observación 4. Determinar si un conjunto de r vectores en $\mathbb{R}^n$ es linealmente independiente se reduce a resolver un sistema homogéneo de r variables y n ecuaciones.

Se trata del siguiente sistema:

$$\lambda_1 \begin{pmatrix} v_{11} \\ v_{12} \\ \vdots \\ v_{1n} \end{pmatrix} + \lambda_2 \begin{pmatrix} v_{21} \\ v_{22} \\ \vdots \\ v_{2n} \end{pmatrix} + ... + \lambda_r \begin{pmatrix} v_{r1} \\ v_{r2} \\ \vdots \\ v_{rn} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

Si la única solución es la trivial, los vectores son linealmente independientes. Si hay más, serán linealmente independientes. También, por el teorema de Rouché-Frobenius, si $r > n$**, serán linealmente dependientes**, al solo poder haber un máximo de n pivotes.

Está idea la podemos expresar en términos de matrices:

Proposición 3. *Fijados r vectores $\{v_i\}_{1\leq i\leq r} \subset \mathbb{R}^n$, si definimos la siguiente matriz $A \in \mathcal{M}_{nr}(\mathbb{R})$:*

$$\begin{pmatrix} v_{11} & v_{21} & \dots & v_{r1} \\ v_{12} & v_{22} & \dots & v_{r2} \\ \vdots & \vdots & \ddots & \vdots \\ v_{1n} & v_{2n} & \dots & v_{rn} \end{pmatrix}$$

donde v_{ij} indica el i-ésimo vector, j-ésima coordenada, sigue que son equivalentes:

1. *Los vectores son linealmente independientes.*
2. *El sistema homogéneo asociado a A solo presenta la solución trivial.*
3. *La matriz escalonada reducida de A tiene r pivotes.*

Demostración. Sigue de la observación anterior y el teorema de Rouché-Frobenius. □

2.2. Rango de una familia de vectores y de una matriz

Proposición 4. *Sea $F = \{v_i\}_{1\leq i\leq r} \subset \mathbb{R}^n$ una familia de vectores. Sea $F' = \{v_i\}_{1\leq i\leq s} \subset \mathbb{R}^n, s > r$ la misma familia con $s - r$ vectores adicionales. Si F' es linealmente independiente, también lo es F. Contrapositivamente, si F es linealmente dependiente, lo es F'.*

Demostración. Supongamos que F' es linealmente independiente pero F es linealmente dependiente. Entonces, $\exists \lambda_1, \lambda_2, ...\lambda_r$, no nulos todos, tales que $\sum_{k=1}^{r} \lambda_k v_k = 0$. Entonces, sigue que $\lambda_1 v_1 + \lambda_2 v_2 + ... + \lambda_r v_r + 0v_{r+1} + ... + 0v_s = 0$ es una combinación lineal de 0 donde no todos los escalares son nulos, contradiciendo la independencia lineal de F'. □

Observación 5. Sea $\mathcal{F} = \{v_1, v_2, ..., v_r\} \subset \mathbb{R}^n$. $\exists p \leq r$, $p \in \mathbb{N}$ tal que $\mathcal{F}$ tiene una subfamilia $\mathcal{F}' = \{v_{i_1}, ..., v_{i_p}\} \subset \mathcal{F}$ tal que $\mathcal{F}'$ es linealmente independiente y toda subfamilia de $p+1$ vectores es linealmente dependiente.

Esto es así porque si vamos eliminando aquellos vectores que sean combinación lineal de los demás, llegará un momento en el que tengamos una familia linealmente independiente, o bien cero vectores. El tamaño de esa familia es el rango.

Definición 7. El **rango** de una familia de vectores es el tamaño del máximo subconjunto posible linealmente independiente, es decir, el p de la observación previa.

Proposición 5 (Hallar el rango de una matriz)**.** *El rango de una matriz A, o lo que es lo mismo, el de la familia de vectores asociada, se corresponde con el número de pivotes de la matriz en su forma escalonada, es decir* ***el número de filas no nulas*** *de la misma.*

Demostración. Esto es así puesto que las columnas que contienen un pivote corresponden a los vectores de $\mathcal{F}'$. Veámoslo:

1. **Los p vectores de las columnas con pivote son linealmente independientes.** En efecto, si consideramos el sistema:

$$A' \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_p \end{pmatrix} = 0$$

donde A' solo contiene en sus columnas a los p vectores, entonces **los mismos** pasos que transforman a A en escalonada, transforman a A' en escalonada (se trata de A ignorando los vectores que no se corresponden a pivotes), y quedan exactamente p pivotes (hemos dejado en A' todos los vectores con pivote), luego el sistema homogéneo solo admite la solución trivial. □

2. **Cualquier subfamilia de más de p vectores es linealmente dependiente.** Se procede igual que antes. Si consideramos el sistema:

$$A'' \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_k \end{pmatrix} = 0$$

donde A'' contiene k vectores de la familia, $k > p$, entonces en su forma escalonada esta matriz será igual a A excepto los vectores eliminados, es decir, hay a lo sumo p pivotes, pero $p < k$ luego el sistema admite soluciones no triviales. □

Sigue además por esta razón que $rg(A) \leq \text{mín}\{m, n\}$ si la matriz es de $m \times n$. Ahora vamos a ver un resultado en torno a $\mathcal{F}'$:

Proposición 6. *$w \in \mathbb{R}^n$ es combinación lineal de los vectores de $\mathcal{F}$ **si y solo si** lo es de los de $\mathcal{F}'$, es decir, del máximo subconjunto linealmente independiente.*

Demostración. Si denotamos una vez más por A a la matriz asociada a la familia $\mathcal{F}$ por columnas, entonces que w sea combinación lineal es lo mismo que decir que $A \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_r \end{pmatrix} = w$ tiene solución, es decir, después de escalonar $A|w$, únicamente hace falta que los términos independientes de las filas nulas sean iguales a 0 (para asegurar la existencia de soluciones). No obstante, si A solo contuviese las columnas con pivotes, es decir, los vectores de $\mathcal{F}'$, las filas no nulas serían las mismas y por tanto los términos independientes que deben ser nulos, los mismos. Es decir, las condiciones para que w sea combinación lineal de ambas familias son las mismas. □

Es decir, la dependencia lineal *no aporta* información extra: añadir a la familia vectores dependientes es irrelevante.

El teorema de Rouché-Frobenius se puede reescribir en términos del rango, con todo lo visto hasta ahora:

Teorema 2 (Rouché-Frobenius)**.** *Si A es la matriz de coeficientes de un sistema y $A|B$ la matriz ampliada, el sistema es compatible si y solo si $rg(A) = rg(A|B)$. Si además $rg(A) = n$, con n el número de columnas de A, entonces es compatible determinado.*

Finalmente, vamos a destacar dos aspectos del rango que serán de utilidad en la siguiente sección:

Proposición 7. *Sea* $A \in \mathcal{M}_{m\times n}(\mathbb{R})$. *Entonces:*

1. *Si* $rg(A) = n$, *el sistema homogéneo* $AX = 0$ *solo tiene la solución trivial. Si* $rg(A) < n$ *tiene infinitas soluciones.*

2. *Si* $rg(A) = m$, *dada cualquier columna de términos independientes* v, *el sistema* $AX = v$ *tiene solución. Sin embargo, si* $rg(A) < m$, *existe algún vector* v' *para el que* $AX = v'$ *no tiene solución.*

Demostración. *1* sigue del teorema de Rouché-Frobenius. Para 2, veamos que $A|v$ en forma escalonada reducida tiene que tener como mínimo rango m, puesto que es A con una columna más. Además, como tiene m filas, como mucho tiene rango m, luego tiene rango m (El mismo que A). Así, el sistema tiene solución. Sin embargo, si $rg(A) < m$, podemos convertirla a escalonada reducida, A', y agregar el vector columna $v'' = (0, 0, ..., 0, 1)$ al resultado. Como $rg(A) < m$, sigue que la fila inferior es nula y se tiene la expresión $0 = 1$, incompatible. Ahora, transformando hacia atrás $A'|v''$ se llega a $A|v'$ que sigue siendo incompatible. □

3. Aplicaciones lineales en $\mathbb{R}^n$

En toda estructura algebraica no solo es importante el estudio de la estructura en sí, sino también las *transformaciones* entre estructuras de la misma índole. Este concepto, en el contexto de los espacios vectoriales, se da en las conocidas como *aplicaciones lineales*. De nuevo, vamos a ejemplificarlas en $\mathbb{R}^n$ para que en secciones posteriores el salto a espacios abstractos sea inmediato.

Definición 8. Una aplicación $f : \mathbb{R}^s \mapsto \mathbb{R}^b$ es **lineal** si conserva las operaciones suma y producto, es decir, si:

$$f(a + b) = f(a) + f(b) \quad \forall a, b \in \mathbb{R}^s \qquad (1)$$

$$f(\lambda a) = \lambda f(a) \quad \forall \lambda \in \mathbb{R}, a \in \mathbb{R}^s \qquad (2)$$

Las transformaciones lineales obligatoriamente dejan fijo el origen de coordenadas:

Proposición 8. *Si* $f : \mathbb{R}^s \mapsto \mathbb{R}^b$ *es lineal se verifica que* $f(0) = 0$.

Demostración. Puesto que en $\mathbb{R}^s$ se tiene que $0 + 0 = 0$, entonces $f(0 + 0) = f(0)$, y entonces $f(0) + f(0) = f(0)$ de donde se sigue, restando, que $f(0) = 0$. □

Definición 9. Se define el **núcleo** de una aplicación lineal $f : \mathbb{R}^s \mapsto \mathbb{R}^b$ como el conjunto $Nu(f) = \{v \in \mathbb{R}^s : f(v) = 0\}$.

(La notación más típica internacionalmente es $\ker(f)$, del inglés *kernel*, y en español $Nuc(f)$, aunque en el contexto de estos apuntes se usará frecuentemente $Nu(f)$.)

Sigue de lo anterior que $0 \in Nu(f)$ para cualquier f lineal.

Proposición 9 (Caracterización de la inyectividad a través del núcleo)**.** *Dada una aplicación lineal* $f : \mathbb{R}^s \mapsto \mathbb{R}^b$, *se tiene que es inyectiva* $\iff Nu(f) = \{0\}$.

Demostración. Para $\implies$, queda claro que como $f(0) = 0$ y f es inyectiva, ese es el único valor de imagen 0, luego $Nu(f) = \{0\}$. Para $\impliedby$, véase que si $v, w \in \mathbb{R}^s$ cumplen que $f(v) = f(w)$, entonces $f(v) - f(w) = 0 \implies f(v - w) = 0$ por la linealidad, y si $Nu(f) = \{0\}$, entonces debe ser que $v - w = 0 \implies v = w$ luego es inyectiva. □

Las aplicaciones lineales guardan relación con las matrices:

Teorema 3. *Sea $A \in \mathcal{M}_{b\times s}(\mathbb{R})$. Definimos $f_A : \mathbb{R}^s \mapsto \mathbb{R}^r$ por $f_A(v) = Av$ (donde v y $f_A(v)$ se tratan como vectores columna). Entonces f_A es lineal.*

Demostración. $f_A(v+w) = Av + Aw = A(v+w) = f_A(v+w)$. Además $f_A(\lambda v) = A\lambda v = \lambda Av = \lambda f_A(v)$. □

Observación 6. $v \in Nu(f_A) \iff Av = 0$, es decir, si v es solución del homogéneo $AX = 0$.

Proposición 10 (Caracterización de inyectividad y sobreyectividad en funciones matriciales)**.** *Sea A una matriz $b \times s$ y $f_A : \mathbb{R}^s \mapsto \mathbb{R}^b$ la función definida anteriormente. Sigue que:*

1. *f_A es inyectiva $\iff rg(A) = s$.*

2. *f_A es sobreyectiva $\iff rg(A) = b$.*

Demostración. Será inyectiva si y solo si $Nu(f_A) = \{0\}$, esto es, $AX = 0$ solo presenta la solución trivial. Según una proposición previa, esto ocurre si y solo si $rg(A) = s$. Será sobreyectiva si y solo si para cualquier $w \in \mathbb{R}^b$, $\exists v \in \mathbb{R}^s$ con $Av = w$. Ahora bien, esto ocurre si y solo si $rg(A) = b$. □

Teorema 4. *Toda transformación lineal $f : \mathbb{R}^n \mapsto \mathbb{R}^m$ se puede expresar mediante una matriz asociada $A \in \mathcal{M}_{m\times n}(\mathbb{R})$ de forma que $f(v) = Av$.*

Demostración. Sea $f : \mathbb{R}^n \mapsto \mathbb{R}^m$ lineal. Denotamos los vectores de la base canónica por $e_1 = (1,0,0,...,0), e_2 = (0,1,0,...,0),..., e_n = (0,0,...,1)$. Ahora, un vector arbitrario de $\mathbb{R}^n$ se puede expresar así: $(x_1, x_2, ..., x_n) = x_1e_1 + x_2e_2 + ... + x_ne_n$. Por tanto, gracias a la linealidad de f, podemos hallar su imagen únicamente sabiendo las de los vectores canónicos. Es decir: **la imagen de los vectores canónicos restringe las demás**. Veamos: $f(x_1, x_2, ..., x_n) = f(x_1e_1 + x_2e_2 + ... + x_ne_n) = x_1f(e_1) + x_2f(e_2) + ... + x_nf(e_n)$. Ahora, si definimos:

$$A = \begin{pmatrix} f(e_1)_1 & f(e_2)_1 & \dots & f(e_n)_1 \\ f(e_1)_2 & f(e_2)_2 & \dots & f(e_n)_2 \\ \vdots & \vdots & \ddots & \vdots \\ f(e_1)_m & f(e_2)_m & \dots & f(e_n)_m \end{pmatrix}$$

que es la matriz que tiene por columnas las transformaciones de los e_i, y por tanto es de $m \times n$, se comprueba fácilmente que se verifica $f((x_1, x_2, ..., x_n)) = x_1f(e_1) + x_2f(e_2) + ... + x_nf(e_n) = A(x_1, x_2, ..., x_n)$ (este último vector como columna para el cálculo). □

Sigue como corolario que si dos funciones lineales f y f' entre los mismos espacios transforman los vectores canónicos a los mismos vectores, son iguales ($f = f'$). Es decir:

Proposición 11. *Sean $f, f' : \mathbb{R}^n \mapsto \mathbb{R}^m$ lineales. Entonces $f = f' \iff f(e_i) = f'(e_i)\ \forall i = 1, 2, ..., n$.*

Y por tanto si la matriz asociada a dos funciones es la misma, las funciones son iguales. Esto da lugar al importante resultado:

Teorema 5. *Existe una biyección entre $\{g : \mathbb{R}^n \mapsto \mathbb{R}^m \setminus g \text{ es lineal}\}$ y $\mathcal{M}_{m\times n}$.*

Hemos visto que cada matriz da lugar a una y solo una aplicación lineal y que cada aplicación lineal se escribe de forma única como matriz. De ahí sigue el resultado.

Proposición 12. *Sean $f, g : \mathbb{R}^n \mapsto \mathbb{R}^m$ aplicaciones lineales y $\lambda \in \mathbb{R}$. Entonces, también son lineales las aplicaciones $(f + g)(x) = f(x) + g(x)$ y $(\lambda f)(x) = \lambda f(x)$.*

Demostración. Sean $v, w \in \mathbb{R}^n$ y $\beta \in \mathbb{R}$. Veamos que $(f+g)(v+w) = f(v+w) + g(v+w) = f(v)+f(w)+g(v)+g(w) = (f+g)(v)+(f+g)(w)$. Además $(f+g)(\beta v) = f(\beta v)+g(\beta v) = \beta f(v)+\beta g(v) = \beta(f(v)+g(v)) = \beta(f+g)(v)$.

Asimismo, $(\lambda f)(v+w) = \lambda f(v+w) = \lambda f(v) + \lambda f(w) = (\lambda f)(v) + (\lambda f)(w)$ y $(\lambda f)(\beta v) = \lambda f(\beta w) = \lambda\beta f(v) = \beta(\lambda f)(v)$. □

De hecho, si se define la suma y producto por escalar de matrices como es habitual, entonces la biyección expuesta en el teorema 5 conserva la suma y el producto por escalar, ya que evidentemente si f y g son lineales $(f+g)(v) = (A+B)v = Av + Bv = f(v) + g(v)$, donde A y B son las matrices asociadas a f y g, y análogamente para el producto.

Proposición 13. *Sean $f : \mathbb{R}^n \mapsto \mathbb{R}^m$ y $g : \mathbb{R}^m \mapsto \mathbb{R}^s$ lineales. Entonces la composición $g \circ f(x) = g(f(x))$ es lineal.*

Demostración. Sean $v, w \in \mathbb{R}^n$ y $\lambda \in \mathbb{R}$. Sigue que $g(f(v+w)) = g(f(v)+f(w)) = g(f(v)) + g(f(w))$ y $g(f(\lambda v)) = g(\lambda f(v)) = \lambda g(f(v))$. □

Otra forma de verlo es que si $f(v) = Av$ y $g(w) = Bw$, entonces $g(f(v)) = B(Av) = (BA)v$ con BA una matriz de $s \times n$ luego es lineal también.

3.1. Operaciones de matrices y aplicaciones

Definición 10. Se definen las aplicaciones suma y producto de matrices así:

$+ : \mathcal{M}_{m\times n} \times \mathcal{M}_{m\times n} \mapsto \mathcal{M}_{m\times n}$ como $\{a_{ij}\}_{i,j} + \{b_{ij}\}_{i,j} = \{a_{ij} + b_{ij}\}_{i,j}$

$\cdot : \mathcal{M}_{m\times n} \times \mathcal{M}_{n\times s} \mapsto \mathcal{M}_{m\times s}$ como $\{a_{ij}\}_{i,j} + \{b_{ij}\}_{i,j} = \{\sum_{k=1}^{n} a_{ik}b_{kj}\}_{i,j}$

Observación 7. Como vimos antes, la suma de aplicaciones lineales es lineal. Si f y g son lineales y de matrices asociadas A y B, entonces $(f+g)$ tiene asociada la matriz $A+B$.

Esto es así porque $(f+g)(v) = Av + Bv = (A+B)v$, propiedad que se cumple de la definición.

Observación 8. Si A define la función lineal $f : \mathbb{R}^n \mapsto \mathbb{R}^m$, y tenemos $B \in \mathcal{M}_{n\times s}$, entonces la matriz AB es la que tiene por columnas las transformaciones de las columnas de B.

Veamos que dada una columna de B, su transformación es $A\begin{pmatrix} b_{1j} \\ b_{2j} \\ \vdots \\ b_{nj} \end{pmatrix} = \begin{pmatrix} c_{1j} \\ c_{2j} \\ \vdots \\ c_{nj} \end{pmatrix}$, que es la columna j de $C = AB$.

Observación 9. Si $f : \mathbb{R}^n \mapsto \mathbb{R}^m$ y $g : \mathbb{R}^m \mapsto \mathbb{R}^s$ son lineales con matrices asociadas A y B, entonces la composición $g \circ f$ es lineal y la matriz BA, que es de $s \times n$, es su matriz asociada.

Esto es así según se vio anteriormente. Así, aplicar sucesivas transformaciones a un vector consiste en irlo multiplicando por la matriz correspondiente de derecha a izquierda.

Observación 10. Si $A, B \in \mathcal{M}_{m\times m}$ son cuadradas, $AB = BA$ no se cumple en general.

El orden en la composición de funciones altera el producto final.

3.2. Inversa, identidad y matrices elementales

En esta sección estudiamos el caso en el que ambas dimensiones de la matriz son las mismas. Es decir, transformamos $\mathbb{R}^n$ en sí mismo.

Observación 11. Sea $A \in \mathcal{M}_{n\times n}(\mathbb{R})$ con $Rg(A) = n$. Entonces la aplicación que define es sobreyectiva e inyectiva, es decir, dado $w \in \mathbb{R}^n$, $\exists! v \in \mathbb{R}^n$ tal que $Av = w$.

Esto sigue de que el rango es igual al número de filas y al de columnas. Además, la matriz escalonada reducida asociada se trata de la identidad.

Proposición 14. *Sea $A \in \mathcal{M}_{n\times n}$ con $Rg(A) = n$, y $B \in \mathcal{M}_{n\times m}$. Entonces, existe* ***una única*** *matriz $m \times n$, X, tal que $AX = B$.*

Demostración. Si $m = 1$ se trata de la observación anterior. Si $m > 1$, sigue de la observación 8 que la columna j de C se obtiene de multiplicar A por la columna j de X, esto es, $AX_j = C_j$, pero de lo anterior, hay un único vector columna X_j que verifica esto. Como cada columna está determinada únicamente, hay una y solo una matriz X. □

Veamos como obtenerla. En definitiva, se busca resolver simultáneamente los sistemas $AX_j = B_j$ para cada columna j de 1 a m. Como la matriz de coeficientes es la misma, podemos colocar todos los sistemas en una matriz y transformarla a escalonada reducida para hallar las columnas de X:

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} & b_{11} & b_{12} & \dots & b_{1m} \\ a_{21} & a_{22} & \dots & a_{2n} & b_{21} & b_{22} & \dots & b_{2m} \\ \vdots & & & & & & & \\ a_{n1} & a_{n2} & \dots & a_{nn} & b_{n1} & b_{n2} & \dots & b_{nm} \end{pmatrix} \to \begin{pmatrix} 1 & 0 & \dots & 0 & x_{11} & x_{12} & \dots & x_{1m} \\ 0 & 1 & \dots & 0 & x_{21} & x_{22} & \dots & x_{2m} \\ \vdots & & & & & & & \\ 0 & 0 & \dots & 1 & x_{n1} & x_{n2} & \dots & x_{nm} \end{pmatrix}$$

Por lo tanto, para obtener la matriz X tal que $AX = B$ con A cuadrada y de rango máximo, se transforma elementalmente la matriz $(A|B)$ en $(I|X)$. Si queremos hallar una matriz tal que $XA = B$, basta con ver que $A^tX^t = B^t$ y aplicar el mismo método.

Definición 11. La **matriz inversa** de una matriz cuadrada de orden n dada, A, existe si $rg(A) = n$ y es aquella matriz X que verifica $AX = I$.

Está claro que existe en esas condiciones según lo que se ha visto antes, y para hallarla basta con particularizar el método anterior, transformando $(A|I)$ en $(I|A^{-1})$.

Proposición 15. *Una matriz A cuadrada de orden n y rango n es regular a derecha, es decir $AP = 0 \implies P = 0$.*

Demostración. Hemos visto que solo puede haber una matriz tal que $AX = 0$, y además $A0 = 0$, luego se verifica obligatoriamente que $X = 0$. Vemos que también es regular a izquierda, puesto que si $PA = 0$, entonces $A^tP^t = 0^t \implies P^t = 0$. □

Proposición 16. *Si X es la inversa de A, es decir $AX = I$, entonces $XA = I$ también.*

Demostración. Veamos que $A(XA - I) = AXA - AI = IA - AI = A - A = 0$, de modo que por la condición de regularidad, $XA - I = 0 \implies XA = I$.

La inversa de A se denota A^{-1}.

Definición 12. Las **matrices elementales** son las matrices cuadradas $E \in \mathcal{M}_{n\times n}$ que se corresponden a la matriz identidad tras aplicar una operación elemental (por ejemplo, sumar un múltiplo de una fila). Se verifica que si $A \in \mathcal{M}_{n\times m}$, entonces EA es la matriz A tras realizar la misma operación que corresponde a E.

Observación 12. Las matrices elementales son invertibles dado que la operación realizada no altera el rango y $rg(I_n) = n$. Además es fácil ver que cada operación elemental tiene otra que la deshace, luego dicha inversa es elemental, y multiplicar esas matrices elementales da la identidad de vuelta.

Proposición 17. *Si A es cuadrada de $n \times n$ y rango n, es decir, invertible, es producto de matrices elementales.*

Demostración. Se vio antes que la matriz se puede llevar por operaciones elementales a la identidad, es decir, $E_r E_{r-1} ... E_2 E_1 A = I$, luego, multiplicando por las inversas, $A = E_1^{-1} ... E_{r-1}^{-1} E_r^{-1}$ que son elementales. Además esas operaciones que llevan a la identidad son la inversa, ya que sigue que $E_r E_{r-1} ... E_2 E_1 A = A^{-1}$.

4. Espacios vectoriales

A continuación introduciremos finalmente el concepto de espacio vectorial. Todo conjunto con las operaciones y propiedades indicadas a continuación tiene estructura de espacio vectorial, y las propiedades que discutamos en un futuro para espacios vectoriales aplicarán en dicho conjunto.

Definición 13. Un **espacio vectorial** sobre un cuerpo $\mathbb{K}$ (en los ejemplos será por lo general $\mathbb{R}$) es un conjunto dotado de dos operaciones:

1. $+ : V \times V \mapsto V$, llamada suma, que asigna a dos vectores v, w otro $v + w$ (notación).
2. $\cdot : \mathbb{K} \times V \mapsto V$, llamada producto por escalar, que asigna a un elemento λ del cuerpo y a un vector v, otro $\lambda \cdot v$ (notación).

Que deben verificar las siguientes 8 propiedades:

1. $\forall v, w \in V,\ v + w = w + v$ (Conmutatividad de la suma)
2. $\forall v, w, z \in V,\ (v + w) + z = v + (w + z)$ (Asociatividad de la suma)
3. $\exists 0 \in V$ t.q. $\forall v \in V,\ v + 0 = v$ (Existencia de elemento neutro de la suma)
4. $\forall v \in V,\ \exists - v \in V$ t.q. $v + (-v) = 0$ (Existencia de inverso de la suma)
5. $\forall v \in V,\ 1 \cdot v = v$ $(1 \in \mathbb{K})$ (Identidad del producto escalar)
6. $\forall \lambda, \beta \in \mathbb{K},\ v \in V,\ (\lambda\beta) \cdot v = \lambda \cdot (\beta \cdot v)$ (Pseudoasociatividad del producto escalar)
7. $\forall \lambda, \beta \in \mathbb{K},\ v \in V,\ (\lambda + \beta) \cdot v = \lambda \cdot v + \beta \cdot v$ (Distributividad del vector)
8. $\forall \lambda \in \mathbb{K},\ v, w \in V,\ \lambda \cdot (v + w) = \lambda \cdot v + \lambda \cdot w$ (Distributividad del escalar)

Es decir, $(V, +)$ debe ser un grupo conmutativo (esta terminología hace referencia a otra de las estructuras algebraicas más importantes, si el lector no está familiarizado puede obviarla) y además se deben verificar las 4 últimas propiedades. Si el cuerpo es $\mathbb{R}$ se denomina **espacio vectorial real**.

Proposición 18. *Las siguientes son propiedades básicas que se infieren de la definición de espacio vectorial:*

1. $\forall v \in V,\ v + 0_1 = v \implies 0_1 = 0$ *(Unicidad del neutro)*
2. $\forall v \in V,\ v + v' = 0 \implies v' = -v$ *(Unicidad del opuesto)*
3. $v + v = v \iff v = 0$
4. $0 \cdot v = 0\ \forall v \in V$
5. $\forall \lambda \in \mathbb{K},\ \lambda \cdot 0 = 0\ (0 \in V)$
6. $(-1) \cdot v = -v\ \forall v \in V.$

Demostración.

1. Se tiene que $(-v) + (v + 0_1) = (-v) + (v) \implies 0 + 0_1 = 0 \implies 0_1 = 0.$
2. Se tiene que $(-v) + (v + v') = (-v) + 0 \implies 0 + v' = -v \implies v' = -v.$

3. Si $v = 0$, al ser neutro, $0 + 0 = 0$. Si $v + v = v$, entonces $v + v + (-v) = v + (-v) \implies v + 0 = 0 \implies v = 0$.

4. Se tiene que $0v = (0 + 0)v = 0v + 0v$, luego, por *3*, $0v = 0$.

5. Se tiene que $\lambda 0 = \lambda(0 + 0) = \lambda 0 + \lambda 0$ luego, por *3*, $\lambda 0 = 0$.

6. Se tiene que $v + (-1)v = 1v + (-1)v = (1 + (-1))v = 0v = 0$, luego, por *2*, sigue que $(-1)v = -v$.

□

Las definiciones de combinación lineal e independencia lineal se mantienen. Si $(V, +, \cdot)$ es espacio vectorial, entonces:

Definición 14 (Combinación lineal)**.** Se dice que $w \in V$ es combinación lineal de los vectores $\{v_1, v_2, ..., v_r\}$ si $w = \lambda_1 v_1 + \lambda_2 v_2 + ... + \lambda_r v_r$ para algunos escalares $\lambda_i \in \mathbb{K}$.

Definición 15 (Independencia lineal)**.** Se dice que un conjunto de vectores $\{v_1, v_2, ..., v_r\} \subset V$ es **linealmente independiente** (o que los vectores son **linealmente independientes**) si y solo si la única combinación lineal de 0 con ellos, es decir, $0 = \lambda_1 v_1 + \lambda_2 v_2 + ... + \lambda_r v_r$, es aquella en la que $\lambda_1 = \lambda_2 = ... = \lambda_r = 0$.

Es decir, son linealmente independientes si $0 = \lambda_1 v_1 + \lambda_2 v_2 + ... + \lambda_r v_r \iff \lambda_1 = \lambda_2 = ... = \lambda_r = 0$.

En caso contrario, son **linealmente dependientes**.

Es fácil observar que son linealmente dependientes si y solo si algún vector se puede expresar como combinación lineal de los demás. (Basta con tomar aquel cuyo escalar no es cero, despejarlo y multiplicar por el inverso del escalar).

Observación 13. $0 \in V$ es combinación lineal de cualquier conjunto de vectores. Por tanto, el conjunto $\{0, v_1, v_2, ..., v_r\} \subset V$ es linealmente dependiente.

En efecto, basta con tomar todos los escalares 0, y como $0v = 0\ \forall v \in V$ y $0 + 0 = 0$, sigue el resultado.

4.1. Subespacios vectoriales

Dentro de un espacio vectorial, uno puede restringir su atención a solo algunos de los elementos del mismo (un subconjunto). Si lo hace, lo más probable es que deje de tener ante sus ojos un espacio vectorial, porque quizás al sumar dos de esos elementos del subconjunto en cuestión, se obtenga un resultado fuera del subconjunto. No obstante, si el subconjunto es cerrado a las operaciones del espacio vectorial, puesto que el resto de propiedades las seguirá verificando, ese subconjunto *hereda* la estructura de espacio vectorial en sí mismo.

Definición 16. Dado un espacio vectorial V sobre $\mathbb{K}$, se dice que $S \subset V$ es **subespacio vectorial** si se verifican:

1. $0 \in S$ (Equivalentemente, S es no vacío)

2. $\forall v, w \in S$, $v + w \in S$

3. $\forall v \in S$, $\lambda \in \mathbb{K}$, $\lambda v \in S$

Observemos que si S es no vacío ya implica que contiene al 0, porque tomamos $v \in S$ y por las otras propiedades $v + (-1)v = 0 \in S$. Es evidente que S es en sí un espacio vectorial, ya que hereda la mayor parte de propiedades de V, y con las condiciones impuestas se garantiza que $0 \in S$, $\forall v \in S$, $-v \in S$ y que sea cerrado respecto de las operaciones.

Definición 17 (Subespacio generado)**.** Sea V un espacio sobre $\mathbb{K}$ y sea $\mathcal{F} = \{v_1, ..., v_r\} \subset V$ una familia de vectores no vacía. Se define el **subespacio generado por** $\mathcal{F}$ como el conjunto $\langle \mathcal{F} \rangle = \{\lambda_1 v_1 + \lambda_2 v_2 + ... + \lambda_r v_r \in V : \lambda_i \in \mathbb{K}\}$, es decir, el conjunto de combinaciones lineales de vectores de $\mathcal{F}$. Un subespacio es **finitamente generado** si $\mathcal{F}$ es finito.

Proposición 19. *Sea V espacio sobre $\mathbb{K}$ y $\mathcal{F} \subset V$ una familia de vectores no vacía.*

1. *$\langle \mathcal{F} \rangle$ es subespacio vectorial.*
2. *$\mathcal{F} \subset \langle \mathcal{F} \rangle$*
3. *Si S es un subespacio vectorial que también verifica $\mathcal{F} \subset S$, entonces necesariamente $\langle \mathcal{F} \rangle \subset S$. Es decir, el generado es el menor subespacio que contiene a $\mathcal{F}$.*

Demostración. Para 2, veamos que basta con poner $\lambda_i = 1$ y $\lambda_{j \neq i} = 0$ en una combinación lineal de $\mathcal{F}$ para obtener v_i. Para 1, veamos que es no vacío por lo anterior. Además, si $v, w \in \langle \mathcal{F} \rangle$, entonces $v = \sum \lambda_i v_i$ y $w = \sum \lambda'_i v_i$, de modo que $v + w = \sum (\lambda_i + \lambda'_i) v_i$ así que $v + w \in \langle \mathcal{F} \rangle$. Análogamente con el producto por escalar. Para 3, veamos que si S es subespacio y contiene a $\mathcal{F}$, por las propiedades de un subespacio, contiene a $\lambda_i v_i$ $\forall \lambda_i \in \mathbb{K}$, $v_i \in \mathcal{F}$. Entonces, por las mismas propiedades, contiene a $\sum \lambda_i v_i$ $\forall \lambda_i \in \mathbb{K}$ de modo que $\langle \mathcal{F} \rangle \subset S$. □

Observación 14. Si S, S' son subespacios de V, entonces $S \cap S'$ también lo es.

Razón. Sabemos que $0 \in S, S'$ luego $0 \in S \cap S'$. Además, si $v, w \in S \cap S'$, entonces $v, w \in S$ y $v, w \in S'$, de modo que $v + w \in S, S'$ luego $v + w \in S \cap S'$. Análogamente para el producto por escalar.

Proposición 20. *Si en $\mathcal{F} = \{v_1, ..., v_r\}$ se tiene que v_j es combinación lineal del resto de vectores de $\mathcal{F}$, y definimos $\mathcal{F}' = \mathcal{F} \backslash \{v_j\}$, sigue que $\langle \mathcal{F} \rangle = \langle \mathcal{F}' \rangle$.*

Demostración. Trivialmente $\langle \mathcal{F}' \rangle \subset \langle \mathcal{F} \rangle$, ya que un vector combinación lineal de $\mathcal{F}'$ se puede expresar como combinación lineal de $\mathcal{F}$ con $\lambda_j = 0$. Ahora, si $v \in \langle \mathcal{F} \rangle$, entonces $v = \sum \lambda_i v_i = (\sum_{i \neq j} \lambda_i v_i) + \lambda_j v_j$, pero como $v_j = \sum_{i \neq j} \lambda'_i v_i$ al ser combinación lineal, sigue que $v = \sum_{i \neq j} (\lambda_i + \lambda'_i) v_i$, es decir, combinación lineal de $\mathcal{F}'$. □

Teorema 6. *Sea $\mathcal{F} \subset V$ una familia finita de vectores no todos nulos. Entonces, $\exists \mathcal{F}' \subset \mathcal{F}$ tal que F' es linealmente independiente y además $\langle \mathcal{F}' \rangle = \langle \mathcal{F} \rangle$.*

Demostración. Si $\mathcal{F} = \{v_i\}_i$ es linealmente independiente hemos acabado. Si no, $\exists \lambda_1, \lambda_2, ..., \lambda_r \in \mathbb{K}$ con $\lambda_j \neq 0$ tales que $\sum \lambda_i v_i = 0$. Despejando, sigue que $v_j = \sum_{i \neq j} -\lambda_j^{-1} \lambda_i v_i$, es decir, es combinación lineal del resto. Sabemos que el conjunto $\mathcal{G} = \mathcal{F} \backslash \{v_j\}$ genera el mismo espacio que el de partida. Ahora, si $\mathcal{G}$ es linealmente independiente hemos acabado y si no repetimos el proceso. Como no todos los vectores son nulos, como mucho acabaremos con un conjunto de un vector no nulo, que es linealmente independiente, nunca con el vacío. □

Definición 18. Con anterioridad hemos hablado de familias no vacías. Para extender todos estos teoremas a familias en general, definimos $\langle \emptyset \rangle = \{0\}$, y $\emptyset$ es linealmente independiente. Es fácil comprobar que esto extiende los resultados anteriores añadiendo el vacío.

Proposición 21 (Caracterización de familias linealmente independientes)**.** *Sea $\mathcal{F} = \{v_1, ..., v_r\}$ una familia no vacía de vectores no nulos.*

Entonces $\mathcal{F}$ es linealmente independiente $\iff$ $\nexists i$ t.q. $v_i = \sum_{k<i} \lambda_k v_k$, es decir, no hay vector que sea combinación lineal de los anteriores en el conjunto.

Demostración. $\Longrightarrow$. Supongamos que algún vector v_i verifica $v_i = \sum_{k<i} \lambda_k v_k$. Entonces, no todos los λ_k son nulos dado que de otro modo v_i sería nulo. Pongamos $\lambda_j \neq 0$, con $1 \leq j < i$. Entonces, sigue que $0 = -v_i + \sum_{k<i} \lambda_k v_k + \sum_{r \geq k > i} 0 v_k$, que se trata de una combinación lineal de 0 en la que $\lambda_j \neq 0$, contradiciendo la independencia lineal.

$\Longleftarrow$. Supongamos que es linealmente dependiente. Entonces, hay escalares λ_i tales que $0 = \sum \lambda_i v_i$ con $\lambda_j \neq 0$, donde j es el máximo índice que verifica esto. Entonces $0 = \sum_{k \leq j} \lambda_k v_k$, pero, despejando $v_j = \sum_{k<j} (-\lambda_j^{-1}) \lambda_k v_k$ contradiciendo el supuesto. □

Este resultado refuerza la misma idea usada en la demostración del teorema 6.

Definición 19 (Base de un subespacio)**.** Se dice que un conjunto de vectores $W = \{w_1, w_2, ..., w_r\} \subset V$ es **base** de un subespacio $S \subset V$ si:

1. W genera S, esto es, $\langle W \rangle\ = S$.

2. W es linealmente independiente.

Por el teorema 6, todo subespacio finitamente generado tiene base.

Proposición 22. *Sea V espacio vectorial. Sean $\mathcal{F}_1 = \{v_1, ..., v_r\} \subset V$ y $\mathcal{F}_2 = \{w_1, ..., w_s\} \subset V$. Si $\mathcal{F}_1$ es linealmente independiente, y $\mathcal{F}_1 \subset\ \langle \mathcal{F}_2 \rangle$, sigue que $r \leq s$.*

Demostración. Se tiene $\forall j = 1, ..., r$ que $v_j = \sum_i a_{ji} w_i$ al ser generadores los w_i. Entonces, una combinación lineal de los vectores v_j igualada a 0 se puede escribir como:

$$\begin{pmatrix} a_{11} & a_{21} & \dots & a_{r1} \\ a_{12} & a_{22} & \dots & a_{r2} \\ \vdots & \vdots & \ddots & \vdots \\ a_{1s} & a_{2s} & \dots & a_{rs} \end{pmatrix} \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_r \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

ya que en cada columna se hallan los v_j en función de los escalares a_{ji} correspondientes, entonces el resultado es multiplicar a cada vector por un escalar λ_j correspondiente e igualarlo a 0. Al ser linealmente independientes los v_j, este sistema solo admite la solución trivial, de modo que $rg(A) = r$, donde A es la matriz de coeficientes. Ahora bien, $rg(A) \leq \text{mín}\{r, s\}$, de modo que $r \leq s$ □

Teorema 7 (Dimensión de un subespacio)**.** *Dado $S \subset V$ un subespacio finitamente generado, sigue que si $\mathcal{F}_1$ y $\mathcal{F}_2$ son ambos bases de S, entonces tienen los mismos elementos: $|\mathcal{F}_1| = |\mathcal{F}_2|$. Este número único se denominará **dimensión del subespacio**.*

Demostración. Usando el resultado previo, por ser bases, $\mathcal{F}_1$ es linealmente independiente y $\mathcal{F}_2$ genera S, luego $|\mathcal{F}_1| \leq |\mathcal{F}_2|$. Análogamente, $|\mathcal{F}_1| \geq |\mathcal{F}_2|$. □

Proposición 23. *Sea $S \subset V$ un subespacio de dimensión r y sea $S' \subset S$ otro subespacio. Entonces S' es finitamente generado con* $\dim(S) \leq r$.

Demostración. Si $S' = \{0\}$ es evidente. Si no, sea $v_1 \in S'$ no nulo. Por ser subespacio, $\langle v_1 \rangle \subset S'$. Si además son iguales, hemos acabado. Si no, sea $v_2 \in S'$, $v_2 \notin \langle v_1 \rangle$. Sigue ahora que $\{v_1, v_2\}$ es linealmente independiente al no ser v_2 combinación lineal de los anteriores. Ahora iteramos de esta forma: si se ha llegado a $\langle v_1, v_2, ..., v_s \rangle$ linealmente independientes, y aún no generan S', tomamos $v_{s+1} \in S'$ con $v_{s+1} \notin \langle v_1, v_2, ..., v_s \rangle$, para formar un conjunto linealmente independiente de un vector más, al no ser combinación de los anteriores. A lo sumo, tendremos un conjunto de r vectores linealmente independiente que genere S', dado que si fuesen más, habría más de r vectores linealmente independientes en S', y por tanto en S, contradiciendo que sea de dimensión r. □

Proposición 24 (Extensión de base)**.** *Sea $S \subset V$ un subespacio de dimensión r y $\mathcal{G} = \{v_1, ..., v_s\} \subset S$ un conjunto linealmente independiente. Entonces, existen vectores tales que $\{v_1, ..., v_s, v_{s+1}, ..., v_r\}$ es base.*

Demostración. En primer lugar cabe notar que $s \leq r$ por uno de los resultados previos. Ahora, $\langle G\rangle \subset S$, y si es igual hemos acabado. Si no, tomamos $v_{s+1} \notin \langle G\rangle$, $v_{s+1} \in S$, y lo añadimos obteniendo un conjunto linealmente independiente al no ser v_{s+1} combinación lineal de los anteriores. Continuamos hasta obtener r vectores, lo que formará base, ya que si no podríamos agregar otro y tener $r+1$ linealmente independientes, lo cual no es posible. □

Proposición 25. *Sea $S \subset V$ un subespacio y sea $\mathcal{F} = \{v_1, ..., v_r\}$ una base. Todo vector $v \in S$ se escribe **de manera única** como combinación lineal de la base.*

Demostración. Está claro que $v = \sum a_i v_i$ dado que $\mathcal{F}$ genera S. Ahora, si tenemos que $v = \sum b_i v_i$, sigue que $\sum a_i v_i - \sum b_i v_i = 0$, al ser ambos v, es decir $\sum (a_i - b_i) v_i = 0 \implies a_i - b_i = 0 \ \forall i = 1, 2, ..., r$ al ser linealmente independientes, de modo que $a_i = b_i \ \forall i = 1, 2, ..., r$ y la forma de escribirlo es única. □

Definición 20. Definimos a los únicos escalares $(a_1, ..., a_r)$ que cumplen $v = \sum a_i v_i$, con los v_i de la base ordenada $\mathcal{F}$, como las **coordenadas** de v respecto de la base $\mathcal{F}$.

Observación 15. Sea $S \subset V$ de dimensión r y $\mathcal{F} = \{v_1, ..., v_r\} \subset S$ un conjunto linealmente independiente. Entonces es base de S. Es decir, todos los conjuntos linealmente independientes de r vectores son base.

Demostración. Si no lo fuese, habría $v \in S$ con $v \notin \langle \mathcal{F}\rangle$, de modo que $\mathcal{F} \cup \{v\}$ sería linealmente independiente de $r + 1$ vectores, contradictorio. □

Observación 16. Sea $S \subset V$ de dimensión r y $\mathcal{F} = \{v_1, ..., v_r\} \subset S$ un conjunto con $\langle \mathcal{F}\rangle = S$. Entonces es base de S. Es decir, todos los conjuntos de generadores de r vectores son base.

Demostración. Supongamos que no lo es. Entonces, no es linealmente independiente, luego hay un vector v que es combinación lineal de los anteriores, y entonces $\mathcal{F} \backslash \{v\}$ sería un sistema de generadores de $r - 1$ vectores, contradictorio. □

Observación 17. Sea $S' \subset S$ con ambos dimensión d. Entonces $S' = S$.

Demostración. Sigue de las observaciones anteriores: una base de S' cualquiera tendrá d vectores, luego lo será de S de modo que son iguales.

Como corolario, sigue que $\dim S = 0 \iff S = \{0\}$, puesto que $\{0\}$ está contenido en cualquier espacio, luego dado S con $\dim S = 0$, se tiene el resultado.

4.2. Operaciones con subespacios

Las dos operaciones básicas son:

Definición 21 (Intersección)**.** Sean S_1 y S_2 subespacios de V. Su intersección $S_1 \cap S_2$ es subespacio de V.

Definición 22 (Suma)**.** Sean S_1 y S_2 subespacios. Su suma, $S_1 + S_2 = \{s + s' : s \in S_1, s' \in S_2\}$ es un subespacio.

La suma se define puesto que la simple unión de subespacios no basta para dar un subespacio. La demostración de la intersección de subespacios se dio en la observación 14. Por otra parte, la suma es subespacio dado que si $v, w \in S_1 + S_2$ entonces $v = s_v + s'_v$ y $w = s_w + s'_w$ con los $s \in S_1$ y los $s' \in S_2$, de manera que $v + w = s_v + s_w + s'_v + s'_w$, y como S_1 y S_2 son subespacios, $s_v + s_w \in S_1$ y $s'_v + s'_w \in S_2$ luego $v + w \in S_1 + S_2$. Análogamente con el producto por escalar: si $v \in S_1 + S_2$, $\lambda \in \mathbb{K}$, entonces $v = s_v + s'_v$ con $s_v \in S_1$ y $s'_v \in S_2$, de manera que $\lambda v = \lambda s_v + \lambda s'_v$, y como S_1 y S_2 son subespacios, $\lambda s_v \in S_1$ y $\lambda s'_v \in S_2$ luego $v + w \in S_1 + S_2$.

Proposición 26 (Obtención de generadores para la suma)**.** *Sea $\mathcal{F}_1$ tal que $< \mathcal{F}_1 > = S_1$ y $\mathcal{F}_2$ tal que $\langle \mathcal{F}_2\rangle = S_2$. Entonces, $\langle \mathcal{F}_1, \mathcal{F}_2\rangle = S_1 + S_2$. Es decir, uniendo sistemas de generadores de cada espacio se obtiene un sistema de generadores para la suma (que luego puede reducirse a base si se desea).*

Demostración. Sea $v \in S_1 + S_2$. Entonces, $v = s_1 + s_2$ con $s_i \in S_i$. Por lo tanto, $v = \sum \lambda_i v_i + \sum \beta_i w_i$, con v_i los vectores de $\mathcal{F}_1$ y w_i los de $\mathcal{F}_2$, al ser generadores. Pero entonces, v es combinación lineal de $\mathcal{F}_1 \cup \mathcal{F}_2$. Asimismo, si es combinación lineal de esas dos familias, se puede separar en dos vectores, uno de S_1 y otro de S_2, al ser generadoras.

Teorema 8 (Grassmann)**.** *Sean $S_1, S_2 \subset V$ subespacios vectoriales. Entonces, se verifica que:* $\dim(S_1 + S_2) = \dim(S_1) + \dim(S_2) - \dim(S_1 \cap S_2)$

Demostración. Sea $\mathcal{F}_3 = \{v_1, v_2, ..., v_r\}$ una base de $S_1 \cap S_2$. Como este espacio es subconjunto de S_1, podemos poner $\mathcal{F}_1 = \{v_1, v_2, ..., v_r, w_1, ..., w_s\}$ una base de S_1 (dado que $\mathcal{F}_3$ también es linealmente independiente en S_1), y por el mismo razonamiento, puede ser $\mathcal{F}_2 = \{v_1, v_2, ..., v_r, z_1, ..., z_p\}$ una base de S_2. Por una de las proposiciones previas, $\mathcal{F}_4 = \{v_1, ..., v_r, z_1, ..., z_p, w_1, ..., w_s\}$ genera $S_1 + S_2$.

Ahora probaremos que $\mathcal{F}_4$ es linealmente independiente en $S_1 + S_2$ y por tanto base. Consideremos $\sum \alpha_i v_i + \sum \beta_i w_i + \sum \gamma_i z_i = 0$. Ahora sea $v \in S_2 = -\sum \gamma_i z_i$, lo que podemos hacer porque $\mathcal{F}_2$ genera S_2. Entonces, $\sum \alpha_i v_i + \sum \beta_i w_i = v$, luego v es combinación lineal de $\mathcal{F}_1$ y por tanto $v \in S_1$, de donde sigue que $v \in S_1 \cap S_2$ por lo que $v = \sum \lambda_i v_i$. Igualando, queda que $\sum \lambda_i v_i = -\sum \gamma_i z_i \implies \sum \lambda_i v_i + \sum \gamma_i z_i = 0 \implies \gamma_i = 0 \; \forall i = 1, 2, ..., p$ al ser $\mathcal{F}_2$ independiente.

Finalmente, esto quiere decir que la combinación lineal inicial queda en $\sum \alpha_i v_i + \sum \beta_i w_i = 0 \implies \alpha_i = \beta_i = 0$ al ser $\mathcal{F}_1$ independiente. Hemos probado que todos los coeficientes son 0 de manera que $\mathcal{F}_4$ es base. Conocidas las dimensiones, sigue que $\dim(S_1 + S_2) = \dim(S_1) + \dim(S_2) - \dim(S_1 \cap S_2) \iff r + s + p = (r + s) + (r + p) - r$ lo cual es evidentemente cierto. □

Definición 23 (Suma directa)**.** Si $U, W \in V$ son subespacios, se dice que su suma es **directa**, en ocasiones denotado $U \oplus W$, si se cumple $U \cap W = \{0\}$. Naturalmente, por la fórmula anterior y el corolario de la observación 17, la suma es directa **si y solo si** $\dim U + W = \dim U + \dim W$.

Proposición 27 (Definicion alternativa de suma directa)**.** *La suma $U + W$ es directa $\iff \forall v \in U + W$, la expresión $v = u + w$ con $u \in U$, $w \in W$ es única.*

Demostración. $\implies$ Supongamos que $v = u + w = u' + w'$, $u, u' \in U$, $w, w' \in W$. Entonces, $u - u' = w' - w$ luego se trata de un vector de $U \cap W$ al estar el lado izquierdo de la igualdad en U y el derecho en W. Por lo tanto, $u - u' = 0 = w - w'$ luego $u = u'$, $w = w'$. Para $\impliedby$, supongamos que la suma no es directa, es decir, $\exists v \neq 0 \in U \cap W$. Entonces, $v = v + 0 = 0 + v$ considerando en el primer caso $v \in U$ y en el segundo $v \in W$, luego no es única su expresión. □

Proposición 28 (Formas de representar un subespacio)**.** *Un subespacio vectorial finitamente generado del espacio vectorial n-dimensional V se puede expresar con un sistema de generadores o con una base, como el espacio generado por estos. Asimismo, se puede expresar de forma paramétrica, donde unos parámatros en $\mathbb{K}$ determinan cada coordenada del espacio en una base de V, o, por último, como el espacio de soluciones de un sistema homogéneo. Es posible pasar de cualquier forma a las demás.*

Veamos que de sistema de generadores a base se pasa eliminando los vectores que dependan de otros hasta obtener la base, linealmente independiente.

Si se tiene la base, $\{u_1, u_2, ..., u_r\}$, podemos escribirla en la base B de V que usaremos para las paramétricas: $\{(u_1^1, u_1^2, ..., u_1^n)_B, (u_2^1, u_2^2, ..., u_2^n)_B, ..., (u_r^1, u_r^2, ..., u_r^n)_B\}$, de modo que un vector $(a_1, a_2, ..., a_n)_B$ de V escrito en B estará en el subespacio si $a_1 = \lambda_1 u_1^1 + \lambda_2 u_2^1 + ... \lambda_r u_r^1$, $a_1 = \lambda_1 u_1^2 + \lambda_2 u_2^2 + ... \lambda_r u_r^2$ y así (se trata de una combinación lineal de los u_i), obteniendo unas paramétricas con r parámetros. A su vez, estas definen las soluciones a un sistema de ecuaciones que, por verificarlo el 0, debe ser homogéneo. Se puede hallar eliminando parámetros o de otros métodos que permitan obtener un homogéneo a partir de sus soluciones.

Observación 18. El sistema homogéneo cuyas soluciones son el subespacio $S \subset V$, también llamado **ecuaciones cartesianas**, verifica que $\dim V - \dim S =$ número de ecuaciones cartesianas independientes.

Esto es así porque, según hemos visto anteriormente, hay $\dim S$ parámetros en las paramétricas, luego $\dim S$ variables libres en el sistema que tiene $\dim V$ variables, luego el rango de la matriz del sistema es $\dim V - \dim S$, y estas son las ecuaciones independientes del mismo.

Definición 24 (Subespacio complementario)**.** Sea $S \subset V$ un subespacio vectorial. Se dice que $S' \subset V$ es **complementario** de S si $S \oplus S' = V$, esto es, si $S + S' = V$ y además $S \cap S' = \{0\}$.

Proposición 29 (Existencia y obtención del complementario)**.** *Sea $S \subset V$ un subespacio tal que $\dim S = d$ y $\dim V = n$. Entonces, existe un subespacio complementario (de dimensión $n - d$).*

Demostración/Método. Sea $\{v_1, v_2, ..., v_d\}$ una base de S. Es linealmente independiente en V. Por la proposición de la extensión de base, sigue que podemos obtener $\{v_1, v_2, ..., v_d, v_{d+1}, ..., v_n\}$ una base de V. Consideremos $S' = < v_{d+1}, v_{d+2}, ..., v_n >$. Veamos que $\{v_{d+1}, v_{d+2}, ..., v_n\}$ lo genera, y además es linealmente independiente al ser subconjunto de la base de V que también lo es. Por tanto, es base de S' y $\dim S' = n-d$. Además, como el sistema de generadores de $S+S'$ obtenido concatenando los de ambos, es la base de V, hemos inferido que $S+S' = V$, luego $\dim(S+S') = \dim V = n = n-d+d = \dim S + \dim S'$, luego, por la fórmula de Grassmann, $\dim S \cap S' = 0 \implies S \cap S' = \{0\}$, así que la suma es directa. $\square$

Por tanto, el método para obtener un complementario es extender su base hasta una base del total. Los vectores añadidos formarán la base del complementario.

4.3. Espacio cociente

Otra operación crucial para generar espacios nuevos a partir de otros dados, es la del *cociente*. En general, las estructuras algebraicas con subestructuras (como el espacio vectorial y los subespacios) permiten generar nuevas estructuras tomando el conjunto cociente si se hace de manera adecuada para que las operaciones estén bien definidas.

Definición 25. Sea V un $\mathbb{K}$ espacio vectorial de dimensión n y sea $S \subset V$ un subespacio del mismo. Definimos la siguiente relación en V: $\forall v, w \in V$, $v \sim w \iff v - w \in S$. Es de equivalencia.

En efecto, veamos que $v \sim v \iff v - v = 0 \in S$, lo cual es cierto; además, $v \sim w \implies v - w \in S \implies -(v-w) \in S \implies w - v \in S \implies w \sim v$, y por último, $v \sim w \wedge w \sim z \implies v - w \in S \wedge w - z \in S \implies (v-w) + (w-z) \in S \implies v - z \in S \implies v \sim z$.

Definición 26. El conjunto de las clases de equivalencia de la relación $\sim$ es el **espacio cociente** y se denota en este caso V/S.

Observación 19. $v \in S \iff [v] = [0]$.

Es decir, estamos *colapsando* todo el subespacio S al punto 0, lo que da lugar a la nueva estructura.

Definición 27 (Operaciones en el conjunto cociente)**.** Definiremos las siguientes operaciones en el espacio cociente:
$+ : V/S \times V/S \mapsto V/S$, dada por $+([v], [w]) = [v + w]$, y
$\cdot : \mathbb{K} \times V/S \mapsto V/S$, dada por $\cdot(\lambda, [v]) = [\lambda v]$.
Están bien definidas y además convierten a V/S en otro $\mathbb{K}$ espacio vectorial.

Veamos que están bien definidas. Para ello, debemos ver que si $[v] = [v']$, entonces $[v]+[w] = [v']+[w]$, y además $\lambda[v] = \lambda[v']$. Para el primer caso, según hemos definido la aplicación, basta ver si $[v+w] = [v'+w]$, es decir, si $v + w \sim v' + w \implies (v + w) - (v' + w) \in S$. Ahora bien, esto es igual a $v - v' \in S$, que es cierto por hipótesis. Para el segundo caso, hay que ver si $\lambda v - \lambda v' \in S$, pero como $v - v' \in S$ esto sigue de la definición de subespacio.

Además, proporcionan estructura de espacio vectorial, ya que todas las propiedades se heredan de las de $+, \cdot$ en V dado que están definidas en base a estas. Por ejemplo, es fácil ver que $[0] = S$ es el neutro en este espacio.

Proposición 30. *El **epimorfismo natural** de paso al cociente, $\pi : V \mapsto V/S$ con $\pi(v) = [v]$ es una aplicación lineal sobreyectiva en cuanto a las operaciones de V y V/S.*

Demostración. La sobreyectividad es evidente dado que $\forall [v] \in V/S$, se tiene que $\pi(v) = [v]$, es decir, todos los representantes de la clase están en la preimagen, y la clase es no vacía. Veamos que es lineal: $\pi(v + w) = [v + w] = [v] + [w] = \pi(v) + \pi(w)$, y $\pi(\lambda v) = [\lambda v] = \lambda[v] = \lambda\pi(v)$. □

Proposición 31. *Sea $\{v_1, v_2, ..., v_n\}$ una base de V. Entonces, $\{[v_1], [v_2], ..., [v_n]\}$ genera V/S.*

Demostración. Sea $[z] \in V/S$. Tomamos un representante arbitrario, por ejemplo, z. Entonces $z = \sum_i \lambda_i v_i$ al ser base. Tomando clases, $[z] = [\sum_i \lambda_i v_i]$, y por la linealidad vista antes, $[z] = \sum_i \lambda_i [v_i]$. □

Para ver como encontrar una base de V/S y cuál es su dimensión, ir al corolario del Teorema 9 (Sección 5).

5. Aplicaciones lineales

Ahora vamos a generalizar lo que sabíamos de transformaciones de $\mathbb{R}^n$ a espacios vectoriales cualesquiera:

Definición 28. Sean V, W dos $\mathbb{K}$-espacios vectoriales. Se dice que $f : V \mapsto W$ es **lineal** si conserva las operaciones suma y producto, es decir, si:

$$f(a + b) = f(a) + f(b) \quad \forall a, b \in V \qquad (1)$$

$$f(\lambda a) = \lambda f(a) \quad \forall \lambda \in \mathbb{K}, a \in V \qquad (2)$$

Como en el caso de $\mathbb{R}^n$, las transformaciones lineales dejan fijo el origen de coordenadas, y preservan los opuestos:

Proposición 32. *Si $f : V \mapsto W$ es lineal se verifica que $f(0) = 0$.*
Si $f : V \mapsto W$ es lineal se verifica que $f(-v) = -f(v)$.

Demostración. Puesto que en V se tiene que $0 + 0 = 0$, entonces $f(0 + 0) = f(0) \mapsto f(0) + f(0) = f(0) \implies f(0) = 0$. Además, $f(-v) = f((-1)v) = (-1)f(v) = -f(v)$ □

Definición 29. Se define el **núcleo** de una aplicación lineal $f : V \mapsto W$ como el conjunto $Nu(f) = \{v \in V : f(v) = 0\}$. También se suele denotar $\ker(f)$.
Se define la **imagen** de una aplicación lineal como $Im(f) = \{f(w) : w \in W\}$.

Sigue de lo anterior que $0 \in Nu(f)$ para cualquier f lineal.

Proposición 33 (Caracterización de la inyectividad y la sobreyectividad)**.** *$f : V \mapsto W$ es inyectiva $\iff$ $Nu(f) = \{0\}$. $f : V \mapsto W$ es sobreyectiva $\iff$ $Im(f) = W$.*

Demostración. Inyectividad: Para $\implies$, queda claro que como $f(0) = 0$ y f es inyectiva, ese es el único valor de imagen 0, luego $Nu(f) = \{0\}$. Para $\impliedby$, véase que si $v, w \in V$ cumplen que $f(v) = f(w)$, entonces $f(v) - f(w) = 0 \implies f(v - w) = 0$ por la linealidad, y si $Nu(f) = \{0\}$, entonces debe ser que $v - w = 0 \implies v = w$ luego es inyectiva. La sobreyectividad es la definición. □

El núcleo y la imagen no se definieron de manera casual, sino que los vectores que conforman estos subconjuntos están organizados en forma de subespacio.

Proposición 34. *Si $f : V \mapsto W$ es lineal, $Nu(f)$ es subespacio vectorial de V e $Im(f)$ lo es de W.*

Demostración. Para $Nu(f)$, ya vimos que $0 \in Nu(f)$. Si $v, w \in Nu(f)$, $f(v) = 0 \wedge f(w) = 0 \implies f(v) + f(w) = 0 \implies f(v+w) = 0 \implies v + w \in Nu(f)$, y $\lambda f(v) = 0 \implies f(\lambda v) = 0 \implies \lambda v \in Nu(f)$.

Para $Im(f)$, también está claro que $0 \in Im(f)$. Si $v, w \in Im(f)$, se tiene que $f(r) = v \wedge f(s) = w$ con $r, s \in V$, luego $f(r+s) = v + w$ por tanto $v + w \in Im(f)$, y análogo para el producto. □

Proposición 35. *Sean $f : V \mapsto W$ y $g : W \mapsto Z$ lineales, con V, W, Z $\mathbb{K}$-espacios vectoriales. Entonces la composición $g \circ f(x) = g(f(x))$ es lineal.*

Demostración. Sean $v, w \in V$ y $\lambda \in \mathbb{K}$. Sigue que $g(f(v+w)) = g(f(v) + f(w)) = g(f(v)) + g(f(w))$ y $g(f(\lambda v)) = g(\lambda f(v)) = \lambda g(f(v))$. □

Definición 30. Una función lineal $f : V \mapsto W$ es **monomorfismo** si es inyectiva, **epimorfismo** si es sobreyectiva o **isomorfismo** si es biyectiva.

Observación 20. f es isomorfismo $\iff$ es epimorfismo y monomorfismo $\iff \exists g : W \mapsto V$ lineal t.q. $f(g(w)) = w = Id_W(w)$ y $g(f(v)) = v = Id_V(v)$.

En realidad, es este último criterio el que sirve como definición de isomorfismo en contextos algebraicos más generales, es decir, la existencia de inversa que además es compatible con las operaciones. Sin embargo, es fácil comprobar que *la inversa de una función lineal biyectiva también es lineal*, de tal modo que en el caso de espacios vectoriales es equivalente.

Observación 21. La composición de monomorfismos es monomorfismo y la de epimorfismos es epimorfismo.

Si h es epimorfismo, $Im(g \circ h) = Im(g)$. En cualquier caso $Im(g \circ h) \subset Im(g)$.

Observación 22. Si V es finitamente generado por $\{v_1, v_2, ..., v_n\}$ y $f : V \mapsto W$, sigue que $\{f(v_1), ..., f(v_n)\}$ genera $Im(f)$.

Razón. $w \in Im(f) \iff f(v) = w,\ v \in V \iff f(\sum_i \lambda_i v_i) = w \iff \sum_i \lambda_i f(v_i) = w$.

Teorema 9 (Teorema rango-nulidad). *Sea $f : V \mapsto W$ una transformación lineal entre espacios finitamente generados. Sigue que* $\dim V = \dim Nuc(f) + \dim Im(f)$.

Demostración. Sea $\dim V = n$, $\dim Nuc(f) = d$. Sea $\{w_1, ..., w_d\}$ una base de $Nuc(f)$. Al ser subespacio de V, podemos completar a base de V: $\{w_1, ..., w_d, z_1, ..., z_{n-d}\}$ será dicha base. Vamos a probar que $\{f(z_1), ..., f(z_{n-d})\}$ es base de $Im(f)$, finalizando la prueba.

En primer lugar, genera $Im(f)$, dado que por la observación anterior $\{f(w_1), f(w_2), ...f(w_d), f(z_1), ...f(z_{n-d})\}$ genera dicha imagen, pero esto es igual a $\{0, 0, ..., 0, f(z_1), ...f(z_{n-d})\}$, luego $\{f(z_1), ...f(z_{n-d})\}$ genera también la imagen.

En segundo lugar, es linealmente independiente: $\sum_i \lambda_i f(z_i) = 0 \iff f(\sum_i \lambda_i z_i) = 0 \iff \sum_i \lambda_i z_i \in Nuc(f) \iff \sum_i \lambda_i z_i = \sum_i \beta_i w_i \iff \sum_i \lambda_i z_i - \sum_i \beta_i w_i = 0 \implies \lambda_i = 0\ \forall i$, al ser esta última una combinación lineal de la base de V. □

Corolario. Si V es un espacio y S un subespacio, $\dim V/S = \dim V - \dim S$ y además, si $\{s_1, ..., s_d\}$ es base de S y $\{s_1, ..., s_d, v_1, ..., v_n\}$ es base de V, se tiene que $\{[v_1], ..., [v_n]\}$ es base de V/S.

Demostración. Basta con aplicar los resultados del teorema rango-nulidad al epimorfismo natural $\pi : V \mapsto V/S$, teniendo en cuenta que $Nuc(\pi) = S$.

Corolario 2. $\dim Nuc(f) = 0 \iff \dim Im(f) = \dim V \iff f$ es monomorfismo

Definición 31. Se dice que dos espacios V, W son *isomorfos* si $\exists f : V \mapsto W$, f isomorfismo. Se denota por $V \simeq W$.

A todos los efectos, ser *isomorfo* es como ser el mismo espacio (salvo quizás los nombres/la construcción de cada uno de ellos), puesto que todo lo que se realice en uno de los dos espacios, puede traducirse al otro a través del isomorfismo.

Proposición 36. *Sea V un $\mathbb{K}$-espacio vectorial de base $B = \{v_1, ..., v_n\}$. Entonces, la función $T : V \mapsto \mathbb{K}^n$ con $T(v) = (\lambda_1, \lambda_2, ..., \lambda_n)$, donde se asignan a cada vector sus coordenadas, es un isomorfismo. Por tanto, todo espacio de dimensión n es isomorfo a $\mathbb{K}^n$.*

Demostración. En primer lugar, es biyectiva puesto que vimos que a cada vector de V corresponde una y solo una expresión en B. Es lineal también, dado que $v + w = \sum_i \lambda_i v_i + \sum_i \beta_i w_i = \sum_i(\lambda_i + \beta_i)v_i$, luego la expresión en coordenadas de la suma es la suma de expresiones de coordenadas, e igual para el producto. □

Este isomorfismo, como se puede notar, depende de la base B escogida.

Corolario. Si V, W son $\mathbb{K}$-espacios de dimensión n, son isomorfos.

Razón. Existen f, g isomorfismos entre V y $\mathbb{K}^n$, y entre W y $\mathbb{K}^n$, respectivamente, luego su composición $g^{-1} \circ f$ es isomorfismo y va entre V y W.

Teorema 10. *Sean V, W dos $\mathbb{K}$-espacios y sea $B = \{v_1, ..., v_n\}$ base de V. Dados vectores $\{w_1, ..., w_n\} \subset W$, existe una y solo una transformación lineal h tal que $h(v_i) = w_i$, $i = 1, 2, ..., n$.*

Es decir, ***las transformaciones de los vectores de la base de V determinan biunívocamente la aplicación lineal****.*

Demostración. Existencia. Basta con definir $h : V \mapsto W$ tal que $h(\sum_i \lambda_i v_i) = \sum_i \lambda_i w_i$. Se comprueba fácilmente que las transformaciones de los vectores base son las pedidas, así como que es lineal: $h(v+w) = h(\sum_i \lambda_i v_i + \sum \beta_i v_i) = \sum_i(\lambda_i + \beta_i)w_i = \sum_i \lambda_i w_i + \sum_i \beta_i w_i = h(v) + h(w)$, e igual para el producto.

Unicidad. Sea h lineal con esas características. Entonces, $h(v) = h(\sum_i \lambda_i v_i)$, y por linealidad, $h(v) = \sum_i \lambda_i h(v_i) = \sum_i \lambda_i w_i$, y v era arbitrario, luego se trata en efecto de la misma función que la dada anteriormente. □

Por ejemplo, el isomorfismo de la proposición 36 viene dado por $T(v_1) = (1, 0, ..., 0)$, $T(v_2) = (0, 1, 0, ..., 0)$... Toda aplicación queda perfectamente caracterizada por la transformación de una base del espacio de partida.

Corolario. f es monomorfismo $\iff$ dada una base $\{v_1, ..., v_n\}$ de partida, el conjunto $\{f(v_1), ..., f(v_n)\}$ es linealmente independiente en el conjunto de llegada.

f es epimorfismo $\iff$ dada una base $\{v_1, ..., v_n\}$ de partida, el conjunto $\{f(v_1), ..., f(v_n)\}$ genera el conjunto de llegada.

Razón. En el caso del monomorfismo, es suficiente y necesario que $dim(Im(f)) = dimV$, y si el conjunto en cuestión es linealmente independiente, lo será en la imagen luego tendrá dimensión n al ser n vectores. Para el epimorfismo, recordemos que $\{f(v_1), ..., f(v_n)\}$ genera la imagen.

5.1. Matrices de aplicaciones lineales. Cambios de base.

La proposición 36 junto con el teorema 10 resulta de gran utilidad para dar una forma de calcular las transformaciones de cualquier vector. Teniendo en cuenta que las coordenadas de un vector preservan las operaciones (se puede operar con sus coordenadas, proposición 36) y que los vectores base definen la aplicación, tenemos el siguiente importante resultado:

Teorema 11 (Matriz de aplicación lineal). *Sea $f : V \mapsto W$ lineal y sean $B = \{v_1, ..., v_n\}$ y $B' = \{w_1, ..., w_m\}$ bases de V y W, respectivamente. Entonces, si obtenemos las coordenadas de los vectores $f(v_1), f(v_2), ..., f(v_n)$ en la base B', podemos definir la matriz:*

$$A = \begin{pmatrix} f(v_1)_1 & f(v_2)_1 & \dots & f(v_n)_1 \\ f(v_1)_2 & f(v_2)_2 & \dots & f(v_n)_2 \\ \vdots & \vdots & \ddots & \vdots \\ f(v_1)_m & f(v_2)_m & \dots & f(v_n)_m \end{pmatrix}$$

Donde $f(v_i)_j$ es la $j-$ésima coordenada en base B' del vector $f(v_i)$. En este caso, se verifica que $\forall v \in V$, $v = (\lambda_1, ..., \lambda_n)_B$, se tiene que $A(\lambda_1, ..., \lambda_n)^t = (\delta_1, ..., \delta_m)^t$, donde $(\delta_1, ..., \delta_m)_{B'} = f(v)$.

Es decir, la matriz transforma los vectores de V escritos en B a su imagen escrita en B'. Se suele denotar $M(f)_{BB'}$ o bien $|f|_{BB'}$.

Demostración. Si $v = \sum_i \lambda_i v_i$, con $(\lambda_1, ..., \lambda_n)_B$ sus coordenadas, entonces, por la linealidad de f, sigue que $f(v) = \sum_i \lambda_i f(v_i)$, lo que se trata de una combinación lineal de vectores en W. Supongamos que las coordenadas de cada $f(v_i)$ son $(\alpha_i^1, \alpha_i^2, ..., \alpha_i^m)_{B'}$. En ese caso, según hemos visto en la proposición 36, las coordenadas de $f(v)$ serán $(\sum_i^n \lambda_i \alpha_i^1, \sum_i^n \lambda_i \alpha_i^2, ..., \sum_i^n \lambda_i \alpha_i^m)_{B'}$, que se corresponde justamente con el producto de A por las coordenadas de v en B. □

Proposición 37 (Composición de funciones lineales)**.** *Si $f : V \mapsto W$ y $g : W \mapsto Z$ son lineales y B, B', B'' son bases de V, W y Z respectivamente, entonces se tiene que $M(g \circ f)_{BB''} = M(g)_{B'B''} M(f)_{BB'}$, es decir, el producto de derecha a izquierda de las matrices por separado.*

Demostración. Si tenemos v escrito en B, sigue que $g(f(v))_{B''} = M(g)_{B'B''} f(v)_{B'} = M(g)_{B'B''} M(f)_{BB'} v_B$. El razonamiento también se puede hacer obteniendo las imágenes de cada vector de B en B'' y dándose cuenta de que coincide con el producto de ambas matrices en ese orden. □

Esta proposición acerca de la composición explica por qué se realiza el producto de matrices de manera tan extravagante: no es más que para obtener la matriz de la composición de las funciones asociadas.

Proposición 38 (Matriz de cambio de base)**.** *Dadas dos bases B y B_1 del espacio V, la **matriz de cambio de base**, A, de B a B_1 permite pasar las coordenadas de cualquier vector de B a B_1. Para ello, basta con colocar en las columnas de la matriz las coordenadas de los vectores de B en la base B_1. Entonces se tiene que $Av_B^t = v_{B_1}^t$*

Demostración. Se puede interpretar como la matriz de la función $Id : V \mapsto V$ para la base B en el espacio de partida y B' en el espacio de llegada. Dado que para cualquier $v_i \in B$, $Id(v_i) = v_i$, efectivamente basta con poner las coordenadas de los mismos vectores de B en B_1. Se trata de un caso particular del teorema 11, con $A = M(Id)_{BB_1}$. □

Proposición 39 (Conversión de matrices de funciones)**.** *Sea $f : V \mapsto W$ lineal. Si $M(f)_{BB'}$ es la matriz de la aplicación en esas bases, y se tiene B_1 una base de V, sigue que $M(f)_{B_1B'} = M(f)_{BB'} M(Id)_{B_1B}$. Si B_1' es base de W, sigue que $M(f)_{BB_1'} = M(Id)_{B'B_1'} M(f)_{BB'}$.*

Demostración. Se va a probar el primer caso, siendo el segundo análogo. Podemos aprovechar que $f \circ Id = f$ para observar que si tomamos $Id : V \mapsto V$ con la base B_1 en partida y B en llegada, la composición queda con B_1 en partida y B' en llegada y basta con aplicar una de las la proposiciones previas. También se puede ver como un cambio, en primer lugar, de B_1 a B y luego una transformación a B'. □

Observemos que la matriz de cambio de base en sentido opuesto a una dada es su inversa, dado que si $M(Id)_{BB'} v_B = v'_B$ entonces $v_B = M(Id)_{BB'}^{-1} v'_B$, y en efecto convierte el vector de B' al mismo en B.

5.2. Endomorfismo proyección

Definición 32. Sea V un espacio vectorial. Se dice que $f : V \mapsto V$ es un **endomorfismo** si es lineal. (Es una aplicación que va de un espacio en sí mismo)

Definición 33. Sean S, T subespacios de V tales que $S \oplus T = V$. Se define la **proyección** sobre S como $p : V \mapsto V$ tal que si $v = s + t$, $s \in S$, $t \in T$ es la expresión única de v como suma de vectores de S y T, entonces $p(v) = s$.

Proposición 40. *La proyección $p : V \mapsto V$ es endomorfismo.*

Demostración. Sean $v = s + t$ y $v' = s' + t'$ con $s, s' \in S$, $t, t' \in T$ vectores de V. Observemos que entonces $v + v' = (s + s') + (t + t')$ y $\lambda v = \lambda s + \lambda t$ son las expresiones únicas de la suma directa de esos vectores. Así, $p(v + v') = s + s' = p(v) + p(v')$ y $p(\lambda v) = \lambda s = \lambda p(v)$. □

Proposición 41. *Estas son las propiedades más importantes de este endomorfismo:*

1. *Si $v \in S$, $p(v) = v$.*
2. *En particular, $p(p(v)) = p(v)$. Así, $p \circ p = p$.*
3. *$Im(p) = S$*
4. *$Nuc(p) = T$*
5. *Si p' es el endomorfismo proyección sobre T, entonces $p + p' = Id_V$, es decir $p' = Id_V - p$ o bien $p = Id_V - p'$.*
6. *$p' \circ p' = p'$.*

Demostración:

1. La expresión única de v será $v = v + 0$.
2. Como $p(v) \in S$ por definición, sigue por *1*.
3. Por definición, $Im(p) \subset S$. Además si $v \in S$, entonces $v = p(v) \in Im(p)$ luego $S \subset Im(p)$.
4. $v \in Nuc(p) \iff v = 0 + v_2$, $v_2 \in T \iff v \in T$.
5. Si la expresión de v es $v = s + t$, entonces $(p + p')(v) = s + t = v = Id_V(v)$.
6. Sigue de que p' es proyección. Otra prueba: $(Id_V - p) \circ (Id_V - p) = Id_V - p - p + p \circ p = Id_V - p$, donde se ha usado la linealidad para distribuir la composición.

□

Resulta que la segunda de esas propiedades es equivalente a ser una proyección:

Proposición 42. *Sea V subespacio vectorial y $p : V \mapsto V$ un endomorfismo tal que $p \circ p = p$. Entonces $V = Im(p) \oplus Nuc(p)$, y p es una proyección de V sobre su imagen.*

Demostración. Sabemos que $\dim V = \dim Im(p) + \dim Nuc(p)$. Por la fórmula de Grassmann, si $\dim(Im(p) \cap Nuc(p)) = 0$, la suma será directa, dado que tendremos que $\dim(Im + Nuc) = \dim(Im) + \dim(Nuc) = \dim(V)$. Para demostrar que esto se cumple, sea $v \in Im(p) \cap Nuc(p)$. Entonces, $v = p(w)$, $w \in V$. Ahora bien, $p(w) = p(p(w)) = p(v)$, y por estar en el núcleo, $p(v) = 0$ de donde $v = p(w) = 0$.

Es fácil ver que p es proyección ya que si $v = p(w) + n$ es la descomposición en un vector de la imagen y otro del núcleo, entonces $p(v) = p(p(w)) + p(n) = p(w) + 0 = p(w)$. □

5.3. Imágenes y preimágenes de morfismos

Proposición 43. *Sea $f : V \mapsto W$ lineal, y $S \subset V$, $T \subset W$ respectivos subespacios. Sigue que $f(S)$ es subespacio en W y $f^{-1}(T)$ lo es en V.*

Demostración. Ya sabemos que $f(S)$ es el espacio generado por las imágenes de una base de S. (Se trata de una mera restricción de f a S, que también es espacio). Si $v, w \in f^{-1}(T)$, entonces, $f(v), f(w) \in T$, y por ser subespacio, $f(v) + f(w) = f(v + w) \in T$, de manera que $v + w \in f^{-1}(T)$. Análogo para el producto. □

Observación 23. Si $f : V \mapsto W$ es epimorfismo y $T \subset W$ es subespacio, $f(f^{-1}(T)) = T$.

Razón. Siempre se tiene que $f(f^{-1}(T)) \subset T$. Para probar la otra inclusión, sea $v \in T$. Entonces, por ser epimorfismo, $\exists w \in f^{-1}(T)$ t.q. $f(w) = v$. Y como $f(w) \in f(f^{-1}(T))$, sigue que $v \in f(f^{-1}(T))$.

Proposición 44. *Si T es subespacio de W y $f : V \mapsto W$ es lineal, se tiene que $Nuc(f) \subset f^{-1}(T)$.*

Demostración. Como $0 \subset T$, sigue que $f^{-1}(0) \subset f^{-1}(T)$, pero $f^{-1}(0) = Nuc(f)$. □

Esto quizás justifica la nomenclatura de *núcleo*, dado que habita en todas los subespacios preimagen posibles.

Si es epimorfismo, este hecho es mucho más fuerte: **todo subespacio de W corresponde biunívocamente a un subespacio de V que contenga al núcleo**:

Proposición 45. *Sea $V \mapsto W$ epimorfismo. Entonces, la aplicación $\sigma : \{T \subset W$ subespacios$\} \mapsto \{S \subset V$ t.q. $Nuc(f) \subset S$, y subespacios$\}$ dada por $\sigma(T) = f^{-1}(T)$ es una biyección entre los subespacios de llegada y los subespacios de partida que contienen al núcleo.*

Demostración. Está bien definida por la proposición 41. Es inyectiva: si $f^{-1}(T) = f^{-1}(T')$, entonces $f(f^{-1}(T)) = f(f^{-1}(T')) \implies T = T'$ al ser epimorfismo. Es sobreyectiva: Sea $S \subset V$ subespacio tal que $Nuc(f) \subset S$. Entonces $f(S)$ verifica que $f^{-1}(f(S)) = S$, puesto que la inclusión a la derecha se cumple para toda función y todo conjunto, y la inclusión a izquierda sigue de que contenga al núcleo: $v \in f^{-1}(f(S)) \implies f(v) \in f(S) \implies \exists s \in S$ t.q $f(v) = f(s) \implies v - s \in Nuc(f) \implies v - s \in S$, y como $s \in S$, sigue que $(v - s) + s = v \in S$ □

Puede parecer artificial, pero la relación que se establece en esta proposición es crucial en el estudio de las estructuras algebraicas, no solo espacios vectoriales.

Como ejemplo, en el epimorfismo natural $\pi : V \mapsto V/S$ se tiene que hay una biyección entre los subespacios de V/S y aquellos de V que contengan a S. Esto facilita la comprensión de la estructura de V/S si uno comprende la de V.

5.4. Teoremas de isomorfía

Esta sección es una de las más abstractas de este documento, pero también una de las más enriquecedoras. En ella se presentan los teoremas de isomorfía, que son unos potentes resultados acerca de morfismos que se aplican no solo a los espacios vectoriales si no a multitud de distintas estructuras algebraicas.

El siguiente lema dará pie a los tres importantes teoremas de isomorfía:

Lema 1 (Lema de isomorfía)**.** *Sea $f : V \mapsto W$ lineal y $S \subset Nuc(f)$ un subconjunto del núcleo. $\exists \bar{f} : V/S \mapsto W$ tal que:*

1. *Si $\pi : V \mapsto V/S$ es el epimorfismo natural de paso al cociente, entonces $f = \bar{f} \circ \pi$. ($V \mapsto V/S \mapsto W$)*
2. *$Nuc(\bar{f}) = Nuc(f)/S$*

3. $\bar{f}$ es lineal.

Demostración. Tal $\bar{f}$ es la definida por $\bar{f}([v]) = f(v)$. En primer lugar, está bien definida, dado que si $[v] = [v']$, entonces $v - v' \in S \implies v - v' \in Nuc(f) \implies f(v - v') = 0 \implies f(v) = f(v')$, esto es, $\bar{f}([v]) = \bar{f}([v'])$. La composición se verifica, dado que $\bar{f}(\pi(v)) = \bar{f}([v]) = f(v)$. Finalmente, $[v] \in Nuc(\bar{f}) \iff f(v) = 0 \iff v \in Nuc(f) \iff [v] \in Nuc(f)/S$.

Para ver que es lineal, $\bar{f}([v] + [w]) = \bar{f}([v + w]) = f(v + w) = f(v) + f(w) = \bar{f}([v]) + \bar{f}([w])$ □

Teorema 12 (Primer teorema de isomorfía)**.** *Sea $f : V \mapsto W$ lineal. Entonces, $V/Nuc(f) \simeq Im(f)$.*

Demostración. Aplicamos el lema con $S = Nuc(f)$. En ese caso, $\bar{f}$ es epimorfismo sobre $Im(f)$, dado que como π es epimorfismo, $Im(\bar{f}) = Im(f)$ (dado que $Im(f) = f(V) = \bar{f}(\pi(V)) = \bar{f}(V/S) = Im(\bar{f})$). Además, $Nuc(\bar{f}) = Nuc(f)/Nuc(f) = [0]$, de modo que es monomorfismo. Por tanto, $\bar{f}$ es el isomorfismo buscado. □

Teorema 13 (Segundo teorema de isomorfía)**.** *Sea V espacio vectorial y sean S, T subespacios tales que $S \subset T \subset V$. Tiene sentido entonces considerar V/T, V/S pero también T/S. Además $T/S \subset V/S$ debido a que son las imágenes por el epimorfismo natural sobre S de T y V, luego tiene sentido considerar $(V/S)/(T/S)$. Se verifica que $(V/S)/(T/S) \simeq V/T$.*

Demostración. Consideramos $\pi : V \mapsto V/T$ el epimorfismo natural, y aplicamos a este morfismo el lema 4 con el conjunto S, que es subconjunto del núcleo T de ese morfismo por hipótesis. Entonces $\bar{\pi} : V/S \mapsto V/T$ verifica que $Nuc(\bar{\pi}) = T/S$, y por el primer teorema aplicado a $\bar{\pi}$, $(V/S)/(T/S) \simeq Im(\bar{\pi})$, pero por la misma razón que en la demostración anterior, $Im(\bar{\pi}) = Im(\pi) = V/T$ al ser epimorfismo, luego $(V/S)/(T/S) \simeq V/T$. □

Teorema 14 (Tercer teorema de isomorfía)**.** *Sean S, T subespacios del espacio V. Entonces $T/(S \cap T) \simeq (S + T)/(S)$.*

Demostración. Consideramos el monomorfismo inclusión $j : T \mapsto S + T$, y el epimorfismo natural $\pi : S + T \mapsto (S + T)/S$. Sea $\delta = \pi \circ j$ ($T \mapsto (S + T) \mapsto (S + T)/S$). Se verifica que es epimorfismo: Si $[v] \in (S + T)/S$, entonces $v = s + t$ con $s \in S$, $t \in T$, y se tiene que $\delta(t) = [t] = [t] + [0] = [t] + [s] = [t + s] = [v]$. Además, $v \in Nuc(\delta) \iff [v] = [0]$, es decir, si y solo si $v \in S$ además de $v \in T$, luego si y solo si $v \in S \cap T$. Luego, por el primer teorema aplicado a δ, $T/(S \cap T) \simeq Im(\delta)/(S) = (S + T)/S$ □

Observación 24. Dada $f : V \mapsto W$ lineal y $S \subset V$ subespacio, podemos considerar $f|_S : S \mapsto V$ la restriccion sobre S. Se cumple que $Nuc(f|_S)) = Nuc(f) \cap S$.

Razón. Son los vectores que van a parar al 0 a través de f, pero además deben estar en S.

Proposición 46. *Sea T complementario de $Nuc(f)$ en V, es decir, $T \oplus Nuc(f) = V$. Entonces $f|_T$ es un monomorfismo que verifica $Im(f|_T) = Im(f)$. Es decir, esta restricción hace inyectiva a la función sin alterar su imagen.*

Demostración. Como $Nuc(f|_T) = Nuc(f) \cap T = \{0\}$, es monomorfismo. Ahora, si $w \in Im(f)$, $\exists v \in V$ t.q. $f(v) = w$, y $v = t + n$, $t \in T$, $n \in Nuc(f)$, por la suma directa. Observemos que $f(t) = f(t) + 0 = f(t) + f(n) = f(t + n) = f(v) = w$, luego $w \in Im(f|_T)$. □

Observemos que si f ya era epimorfismo, la restricción la convierte en isomorfismo. Esto permite dar una prueba alternativa del corolario del teorema 9: considerando el epimorfismo natural $\pi : V \mapsto V/S$, si $\{s_1, ..., s_d\}$ es base de $S = Nuc(\pi)$, completamos hasta $\{s_1, ..., s_d, v_1, ..., v_n\}$ base de V, con lo que el subespacio de base $\{v_1, ..., v_n\}$ es complementario, y restringir π a el nos da un isomorfismo de misma imagen (V/S al ser sobreyectiva), luego $\pi(\{v_1, ..., v_n\}) = \{[v_1], ..., [v_n]\}$ es base de dicha imagen.

6. Espacio dual

Quizá el lector haya notado al discutir las aplicaciones lineales, que la *suma de aplicaciones lineales* y el *producto de un escalar por una aplicación lineal* vuelve a dar como resultado nuevas aplicaciones lineales, de una manera que sugiere que **las propias aplicaciones lineales sean un espacio vectorial**. En esta sección estudiaremos el *espacio dual*, que es precisamente un nuevo espacio vectorial que se induce a partir de otro, considerando algunas aplicaciones lineales sobre este.

Conceptos previos:

Definición 34. Sean V, W dos $\mathbb{K}$-espacios vectoriales. Se define el espacio $L(V,W) = \{f : V \mapsto W : \text{f es lineal}\}$, con las operaciones de suma: $(f_1 + f_2)(v) = f_1(v) + f_2(v)\ \forall v \in V$ y de producto por escalar: $(\lambda f)(v) = \lambda f(v)\ \forall v \in V$.

Vimos con anterioridad las distintas propiedades que llevan a afirmar que se trata en efecto de un espacio vectorial y que el resultado de aplicar esos operadores a funciones lineales proporciona otra lineal.

Observación 25. Si $\dim V = n$ y $\dim W = m$, $L(V,W)$ es isomorfo a $\mathcal{M}_{mn}(\mathbb{K})$. Así, $\dim L(V,W) = \dim V \cdot \dim W$.

Razón. Ya sabemos que existe una biyección entre aplicaciones lineales y matrices ya que cada función se caracteriza por donde manda los vectores de base de V. Fijadas $B \in V$ y $B' \in W$ bases, dicha biyección vimos además que se trataba de una función lineal teniendo en cuenta la suma/producto por escalar de $\mathcal{M}_{mn}(\mathbb{K})$.

Proposición 47. *Sea V' espacio vectorial, y sea $g : V' \mapsto V$ lineal. Esto permite definir la aplicación lineal:*

$\alpha_g : L(V,W) \mapsto L(V',W)$ *dada por* $\alpha_g(f) = f \circ g$.

Demostración. Ya sabemos que la composición será lineal. Vamos a probar que $\alpha_g(f+h) = (f+h)\circ g = f \circ g + h \circ g = \alpha_g(f) + \alpha_g(h)$. Para ello, sea $v \in V'$, entonces $(f+h)(g(v')) = f(g(v')) + h(g(v'))$ por definición, cumpliendo la afirmación. Análogo para el producto. □

Definición 35. Sea V un $\mathbb{K}$-espacio vectorial. Se define su **espacio dual** por $V^* = L(V, \mathbb{K})$.

Observación 26. $\dim V = \dim V^*$ por la observación 25.

Una transformación lineal $g : V' \mapsto V$ permite definir otra $\alpha_g : V^* \mapsto V'^*$ como se vio en la proposición anterior.

Definición 36. Sea $B = \{v_1, ..., v_n\} \subset V$ una base. Se define la **base dual** de B por el conjunto $B^* = \{\gamma_1, \gamma_2, ..., \gamma_n\} \subset V^*$ donde la forma lineal γ_i es la definida por $\gamma_i(v_i) = 1$ y $\gamma_i(v_j) = 0$, con $j \neq i$. (Recordemos que cada aplicación se caracteriza por las imágenes de la base)

Proposición 48. *Fijados V espacio vectorial y B base del mismo, se tiene que B^* es base de V^*.*

Demostración. Se trata de $\dim V$ vectores así que basta probar que son linealmente independientes. Supongamos que $\{\lambda_k\}$ son escalares tales que $\sum \lambda_k \gamma_k = 0$, donde 0 es la forma lineal nula (elemento neutro de V^*). En particular, evaluando en $v_i \in B$: $(\sum \lambda_k \gamma_k)(v_i) = \lambda_i = 0$, donde hemos usado la definición de los γ_k. Por tanto, $\lambda_i = 0\ \forall i = 1, 2, ..., \dim V$. □

Proposición 49. *Sea $B = \{v_1, ..., v_n\}$ base de V y $B^* = \{\gamma_1, ..., \gamma_n\}$ su base dual en V^*. Entonces si $\gamma \in V^*$, sigue que sus coordenadas son: $\gamma = (a_1, ..., a_n)_{B^*}$, donde $a_i = \gamma(v_i)$, es decir, vienen dadas por las imágenes de B. En otras palabras, coinciden con las entradas de la matriz $|\gamma|_{B,\{1\}} = (\gamma(v_1), ..., \gamma(v_n)) = (a_1, ..., a_n)$.*

Esto es, $\gamma = \sum_{k=1}^{n} \gamma(v_k)\gamma_k$.

Demostración. Para ver que en efecto se cumple la igualdad $\gamma = \sum_{k=1}^{n} \gamma(v_k)\gamma_k$, basta con que ambas funciones coincidan en los vectores de B, como sabemos. Sea $v_i \in B$, entonces $(\sum_{k=1}^{n} \gamma(v_k)\gamma_k)(v_i) = \gamma(v_i) \cdot \gamma_i(v_i) = \gamma(v_i)$, luego coinciden. $\square$

Teorema 15. *Sean V_1, V_2 $\mathbb{K}$-espacios vectoriales, con bases $B_1 = \{v_1, ..., v_n\}$ y $B_2 = \{w_1, ..., w_m\}$. Denotamos $B_1^* = \{\gamma_1, ..., \gamma_n\}$ y $B_2^* = \{\delta_1, ..., \delta_m\}$ sus bases duales. Sea $g : V_1 \mapsto V_2$ lineal. Definimos $g^* : V_2^* \mapsto V_1^*$ como se vio en la proposición previa, con $g^*(\delta) = \delta \circ g$.*

Entonces, se verifica que $|g|_{B_1 B_2} = (|g^*|_{B_2^* B_1^*})^t$, es decir, la matriz de g es la transpuesta de la de su función dual.

Demostración. Sea

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix} = |g|_{B_1 B_2}$$

la matriz de g en esas bases. Como la matriz de g^* se construye a partir de las imágenes de B_2^* escritas en B_1^*, para probar que es la traspuesta de A es suficiente con ver que $g^*(\delta_i) \equiv (a_{i1}, a_{i2}, ..., a_{in})_{B_1^*}$, $\forall i = 1, 2, ..., m$.

Ahora bien, si $g^*(\delta_i) \equiv (b_1, ..., b_n)_{B_1^*}$, entonces por la proposición 46 se tiene que $b_j = g^*(\delta_i)(v_j) = \delta_i(g(v_j)) = \delta_i(\sum_{k=1}^{m} a_{kj} w_k) = \sum_{k=1}^{m} a_{kj} \delta_i(w_k) = a_{ij}$, como queríamos probar. Hemos usado la linealidad de cada δ_i además de su definición. $\square$

Proposición 50. *Si tenemos aplicaciones lineales $g : V_1 \mapsto V_2$ y $h : V_2 \mapsto V_3$, sigue que $(h \circ g)^* = g^* \circ h^*$. (Obsérvese que ambas son de V_3^* en V_1^*).*

Demostración. Sea $\delta \in V_3^*$. Entonces, $(h \circ g)^*(\delta) = \delta \circ (h \circ g) = (\delta \circ h) \circ g = h^*(\delta) \circ g = g^*(h^*(\delta)) = g^* \circ h^*(\delta)$. Como δ era arbitrario sigue la igualdad. $\square$

Proposición 51. *Si $id_V : V \mapsto V$ es la función identidad, entonces $(id_V)^* : V^* \mapsto V^*$ verifica $(id_V)^* = id_{V^*}$.*

Demostración. Sea $\delta \in V^*$. Entonces $(id_V)^*(\delta) = \delta \circ id_V = \delta$ $\square$

Observación 27. Dada $\delta \in V^*$, con $\dim V = n$, se tiene $\dim Nuc(\delta) = n - 1$ salvo si $\delta = 0$, en cuyo caso $\dim Nuc(\delta) = n$, ya que la imagen es o bien $\mathbb{K}$ o bien $\{0\}$.

Definición 37 (Anulador de un subespacio)**.** Sea $S \subset V$ un subespacio vectorial de V. Se define su **anulador** $ann(S) = \{\delta \in V^* : S \subset Nuc(\delta)\} \subset V^*$, es decir, las formas lineales de V^* tales que $\delta(s) = 0$, $\forall s \in S$. Se denota también S^0.

Observación 28. Se tiene que $ann(\{0\}) = V^*$ y que $ann(V) = \{0\} \subset V^*$.

Razón. Ya se demostró que toda aplicación lineal lleva el 0 al 0. Además, la forma lineal que lleva todo V a 0 es única y es la nula.

Observación 29. Si $S_1 \subset S_2 \subset V$ son subespacios, entonces $S_2^0 \subset S_1^0$. Además, $(S_1 + S_2)^0 = S_1^0 \cap S_2^0$, y $(S_1 \cap S_2)^0 = S_1^0 + S_2^0$.

Razón. Si $\delta \in S_2^0$, entonces $S_2 \subset Nuc(\delta)$ por lo que $S_1 \subset Nuc(\delta)$ luego $\delta \in S_1^0$.

Si $\delta \in (S_1 + S_2)^0$, entonces $\forall s_1 \in S_1$, $s_2 \in S_2$, al estar en la suma, $\delta(s_1) = \delta(s_2) = 0$, por lo que $\delta \in S_1^0$ y $\delta \in S_2^0$. Además si $\delta \in S_1^0 \cap S_2^0$, entonces, si $s \in (S_1 + S_2)$, se tiene $s = s_1 + s_2$ para algunos s_1, s_2, no necesariamente únicos, y entonces $\delta(s) = \delta(s_1 + s_2) = \delta(s_1) + \delta(s_2) = 0 + 0 = 0$, luego $\delta \in (S_1 + S_2)^0$. Un razonamiento similar para la otra afirmación. $\square$

Teorema 16. *Sea $S \subset V$ subespacio. Entonces S^0 es subespacio de V^*, y además si* $\dim S = d$ *y* $\dim V = n$*, sigue que* $\dim S^0 = n - d$.

Demostración. Sea $\{v_1, ..., v_d\} \subset S$ base. Extendemos a $B = \{v_1, ..., v_d, v_{d+1}, ..., v_n\}$ base de V. Consideramos $B^* = \{\delta_1, ..., \delta_d, \delta_{d+1}, ..., \delta_n\}$ la dual. Vamos a probar que $S^0 =< \{\delta_{d+1}, ..., \delta_n\} >$, con lo que habremos probado que es subespacio, y que es de dimensión $n - d$. (Esa familia es l.independiente al ser subconjunto de una base). Sea $\delta \in S^0$. Entonces $\delta = \sum_1^n \lambda_i \delta_i$. Supongamos que algún λ_{i_0} con $1 \leq i_0 \leq d$ es no nulo. Entonces $\delta(v_{i_0}) = \lambda_{i_0} \neq 0$ por definición y δ no anula S, luego $\delta = \sum_{i=d+1}^n \lambda_i \delta_i$. Por otro lado, si $\delta = \sum_{i=d+1}^n \lambda_i \delta_i$, y $s' \in S$, entonces $\delta(s') = \delta(\sum_1^d \beta_j v_j) = \sum_1^d \beta_j \delta(v_j) = 0$ dado que en v_j todas las δ_i, $d + 1 \leq i \leq n$ valen 0 por definición. □

6.1. Aplicaciones bilineales y el espacio bidual

Uno podría preguntarse entonces qué ocurre si toma el espacio dual del espacio dual, es decir $(V^*)^*$. Lo que veremos en esta sección es una manera satisfactoria de ver que existe un fuerte isomorfismo entre ese y el V original. Si bien es cierto que V, V^* y $(V^*)^*$ son isomorfos por tener la misma dimensión, la relación que hay entre V y $(V^*)^*$ es mucho más natural y fuerte que el isomorfismo de cambio de base, dado que este isomorfismo que discutiremos ni siquiera requiere de elegir una base.

Definición 38. Sea $\mathbb{K}$ un cuerpo y V, W, Z tres $\mathbb{K}$-espacios vectoriales. Sea $\psi : V \times Z \mapsto W$ una aplicación. Se dice que es **bilineal** si $\psi_v : Z \mapsto W$ dada por $\psi_v(z) = \psi(v, z)$ es lineal $\forall v \in V$ y $\psi'_z : V \mapsto W$ dada por $\psi'_z(v) = \psi(v, z)$ es lineal $\forall z \in Z$. Es decir, si es lineal para cada variable.

Proposición 52. *La función evaluación, $\psi : V \times V^* \mapsto \mathbb{K}$, dada por $\psi(v, \gamma) \mapsto \gamma(v)$, que devuelve el resultado de evaluar el vector en la forma, es bilineal.*

Demostración. Sea $v \in V$. Entonces, $\psi(v, \gamma + \delta) = (\gamma + \delta)(v) = \gamma(v) + \delta(v) = \psi(v, \gamma) + \psi(v, \delta)$. Asimismo, $\psi(v, \lambda\gamma) = (\lambda\gamma)(v) = \lambda\gamma(v) = \lambda\psi(v, \gamma)$, donde $\gamma, \delta \in V^*$ y $\lambda \in \mathbb{K}$.

Asimismo, sea $\gamma \in V^*$. Entonces, $\psi(v + w, \gamma) = \gamma(v + w) = \gamma(v) + \gamma(w) = \psi(v, \gamma) + \psi(w, \gamma)$ al ser lineal, y $\psi(\lambda v, \gamma) = \gamma(\lambda v) = \lambda\gamma(v) = \lambda\psi(v, \gamma)$, donde $v, w \in V$ y $\lambda \in \mathbb{K}$. □

Teorema 17 (Bidual)**.** *Sea V un $\mathbb{K}$-espacio vectorial, V^* su dual, y $V^{**} = \{\gamma : V^* \mapsto \mathbb{K}$ lineales $\}$ el dual de este. Se tiene que $V \simeq V^{**}$ a través del isomorfismo natural $\alpha : V \mapsto V^{**}$ dado por $\alpha(v) = \psi_v \in V^{**}$, donde $\psi_v : V^* \mapsto \mathbb{K}$ es la evaluación $\psi_v(\delta) = \delta(v)$.*

Este isomorfismo no depende, ni esta asociado a ninguna base de V ni V^{**}, a diferencia de otros isomorfismos entre espacios de igual dimensión.

Demostración. Veamos que es lineal. $\alpha(v + w) = \psi_{v+w} = \psi_v + \psi_w = \alpha(v) + \alpha(w)$, y $\alpha(\lambda v) = \psi_{\lambda v} = \lambda\psi_v = \lambda\alpha(v)$, al ser bilineal la evaluación por la proposición anterior.

Ahora, como $\dim V = \dim V^* = \dim V^{**}$, basta con probar que es monomorfismo y el teorema rango-nulidad asegurará que es isomorfismo. Para ello probaremos que $Nuc(\alpha) = \{0\}$. Al ser lineal basta probar que $Nuc(\alpha) \subset \{0\}$. Sea $v \in Nuc(\alpha)$. Entonces, $\alpha(v) = 0_{V^{**}}$, es decir, $\alpha(v)(\delta) = 0 \ \forall \delta \in V^*$, o lo que es lo mismo, $\delta(v) = 0 \ \forall \delta \in V^*$. Ahora, sea una base cualquiera $B = \{v_1, ..., v_n\} \subset V$ y su dual $\{\delta_1, ..., \delta_n\}$. Si $v = \sum \lambda_i v_i$, como $\delta_i(v) = \lambda_i = 0 \ \forall i = 1, ..., n$, al tratarse de formas de V^*, sigue que $v = 0$. □

Proposición 53. *Dado $S \subset V$ subespacio, sigue que $S = ann(ann(S))$ con el isomorfismo natural α anterior.*

Demostración. Observemos primero que si $\dim S = d$ y $\dim V = n$, entonces $\dim ann(ann(S)) = n - (n - d) = d$ luego son de igual dimensión. Así, basta probar una inclusión a través del isomorfismo. Sea $v \in S$, entonces $\psi_v(\delta) = 0 \iff \delta(v) = 0 \iff \delta \in ann(S)$, es decir, $\alpha(v) = \psi_v$ anula las formas de $ann(S)$, luego $\psi_v \in ann(ann(S))$ y así $\alpha(S) = ann(ann(S))$.

Observación 30. Se puede probar que $(S \cap T)^0 = S^0 + T^0$ y $(S + T)^0 = S^0 \cap T^0$ con la teoría anterior.

Veremos la segunda. Tengamos en cuenta que $S, T \subset S + T$ de modo que $(S + T)^0 \subset S^0, T^0$ y por tanto $(S + T)^0 \subset (S^0 \cap T^0)$. Para la otra inclusión, basta con tomar anuladores a través del isomorfismo: $(S^0 \cap T^0)^0 \subset S + T$, pero además $(S^0)^0, (T^0)^0 \subset (S^0 \cap T^0)^0 \implies (S^0)^0 + (T^0)^0 \subset (S^0 \cap T^0)^0$, con lo que queda $S + T \subset (S^0 \cap T^0)^0 \subset (S + T)$ luego sus anuladores son iguales, así que coinciden en dimensión y por tanto la inclusión es una igualdad. $\square$

7. Formas multilineales y determinantes

En esta sección vamos a profundizar en las funciones multilineales, que son aquellas que aceptan varios vectores en su entrada, y exploraremos el *determinante*, que es una función multilineal con interesantes propiedades y numerosas aplicaciones en distintos campos de las matemáticas.

Proposición 54. *El conjunto de funciones bilineales $\psi : V \times Z \mapsto W$ adquiere estructura de espacio vectorial con las operaciones $(\psi)+(\varphi) = (\psi+\varphi)$ definida por $(\psi+\varphi)(v, z) = \psi(v, z)+\varphi(v, z)$ y $\lambda\psi = (\lambda\psi)$ definida por $(\lambda\psi)(v, z) = \lambda\psi(v, z)$.*

Demostración. Es fácil probar que la resultante de cada operación es multilineal también. Veamos que $(\psi + \varphi)(\lambda v + \beta v', z) = \psi(\lambda v + \beta v', z) + \varphi(\lambda v + \beta v', z) = \lambda\psi(v, z) + \beta\psi(v', z) + \lambda\varphi(v, z) + \lambda\varphi(v', z) = \lambda((\psi + \varphi)(v, z)) + \beta((\psi + \varphi)(v', z))$ y análogo para las demás pruebas. $\square$

Definición 39. Sean $V_1, ..., V_n, Z$ unos $\mathbb{K}$-espacios vectoriales. La función $\psi : V_1 \times V_2 \times ... \times V_n \mapsto Z$ es **n-lineal** o **multilineal** si lo es con respecto a cada variable, fijadas las demás. Esto es, si $\forall i = 1, ..., n$ se tiene que $\psi_{v_1 v_2 ... v_{i-1} v_{i+1} ... v_n} : V_i \mapsto Z$ dada por $\psi_{v_1 v_2 ... v_{i-1} v_{i+1} ... v_n}(v) = \psi(v_1, v_2, ..., v_{i-1}, v, v_{i+1}, ..., v_n)$ es lineal $\forall v_1 \in V_1, v_2 \in V_2, ...v_n \in V_n$.

Proposición 55. *El conjunto de aplicaciones multilineales de $V_1 \times ... \times V_n$ en Z adquiere estructura de espacio vectorial con las operaciones naturales de la proposición 51.*

7.1. Conceptos acerca de biyecciones, reordenamientos y trasposiciones

Definición 40. Se define $S_n = \{f : \{1, 2, ..., n\} \mapsto \{1, 2, ..., n\} \text{ biyectivas}\}$ el conjunto de biyecciones entre los primeros n naturales.

Algunas propiedades evidentes son que es cerrado a la composición y a la inversión, y $(f \circ g)^{-1} = g^{-1} \circ f^{-1}$.

Definición 41. Una **trasposición** $\tau_{ij} \in S_n$ es la biyección dada por $\tau_{ij}(i) = j$, $\tau_{ij}(j) = i$ y $\tau_{ij}(k) = k$ $\forall k \neq i$, $k \neq j$. Se requiere además que $i \neq j$. Se verifica que $\tau_{ij}^{-1} = \tau_{ij}$.

Un importante resultado de la *teoría de grupos* acerca del conjunto S_n, que no es más que las permutaciones de n elementos, indica que toda tal permutación puede interpretarse como la realización de una serie de trasposiciones de forma sucesiva:

Teorema 18. *Una biyección $\sigma \in S_n$ se puede escribir $\sigma = \tau_1 \circ \tau_2 \circ ... \circ \tau_n$ donde cada τ_i es una trasposición. Además, la paridad de n es única para cada biyección, aunque haya varias representaciones.*

Puesto que la paridad de n siempre es la misma independientemente de la descomposición escogida para una $\sigma \in S_n$, podemos definir:

Definición 42. Se define el **signo** de una biyección como $sig(\sigma) = 1$ si σ es composición de un número par de trasposiciones, o $sig(\sigma) = -1$ si lo es de un número impar. Se tiene como corolario que $sig(\sigma) = sig(\sigma^{-1})$

Definición 43. Una **reordenación** $(i_1, ..., i_n)$ es una $n-upla$ ordenada tal que $\{i_1, ..., i_n\} = \{1, 2, ..., n\}$. Cada reordenación da lugar a la biyección $\sigma \in S_n$ tal que $\sigma(k) = i_k$, y alternativamente, cada $\sigma \in S_n$ da lugar a la reordenación $(\sigma(1), \sigma(2), ..., \sigma(n))$.

Definición 44. La forma n-lineal sobre el $\mathbb{K}$-espacio V vectorial, $\psi : V \times V \times ... \times V \mapsto \mathbb{K}$, se denomina **alternada** si $\forall v_1, ..., v_n \in V$, se tiene que $\psi(v_{\sigma(1)}, ..., v_{\sigma(n)}) = sig(\sigma)\psi(v_1, ..., v_n)$, para cualquier reordenamiento σ. Consecuentemente, intercambiar dos vectores de sitio en la entrada solo varía el signo de la salida.

El espacio $\bigwedge_n(V) = \{\psi : V^n \mapsto \mathbb{K}$, multilineales y alternadas$\}$ es vectorial con las operaciones usuales.

(Nota para el lector entendido: en realidad, las funciones que cumplen la definición previa se suelen denominar antisimétricas, y la definición estándar de alternada es que la forma valga 0 *cuando dos de sus entradas son la misma. Estas dos definiciones son equivalentes si suponemos que la característica de* $\mathbb{K}$ *es distinta de* 2*, pero no así en el caso de característica* 2*. Para estos apuntes, vamos a tomar como definición la dada arriba.)*

7.2. El determinante

Definición 45. Sea $A = (a_{ij}) \in \mathcal{M}_n$. Se define su **determinante** por

$$det(A) = \sum_{\sigma \in S_n} sig(\sigma) \cdot a_{\sigma(1),1} \cdot a_{\sigma(2),2} \cdot ... \cdot a_{\sigma(n)n} = \sum_{\sigma \in S_n} sig(\sigma) \prod_{r=1}^{n} a_{\sigma(r)r}$$

es decir, la suma a través de todas las reordenaciones de las filas de la traza de cada reordenación, con su signo de paridad. Se denota también $|A|$.

Proposición 56. $det(A) = det(A^t)$

Demostración. Si $A = (a_{ij})$, entonces sigue que $det(A^t) = \sum_{\sigma \in S_n} sig(\sigma) \prod_{r=1}^{n} a_{r\sigma(r)} =$ $\sum_{\sigma \in S_n} sig(\sigma^{-1}) \prod_{r=1}^{n} a_{\sigma^{-1}(r)r} = \sum_{\sigma \in S_n} sig(\sigma) \prod_{r=1}^{n} a_{\sigma(r)r} = det(A)$, donde hemos usado que $sig(\sigma) = sig(\sigma^{-1})$ y que el conjunto de todas las reordenaciones es igual al de las inversas de todas las reordenaciones, así como que $\{(1, \sigma(1)), ...(n, \sigma(n))\} = \{(\sigma^{-1}(1), 1), ..., (\sigma^{-1}(n), n)\}$ ☐

Teorema 19. *El determinante es una forma multilineal alternada. Esto es,* $\phi : \mathbb{K}^n \times ... \times \mathbb{K}^n = (\mathbb{K}^n)^n \mapsto \mathbb{K}$ *dada por* $\phi(v_1, v_2, ..., v_n) = det(v_1|v_2|...|v_n)$*, es decir, el determinante de la matriz cuyas columnas son los vectores, es multilineal alternada.*

Demostración. Primero veremos que es lineal. Sea $k \in \{1, 2, ..., n\}$. Sean $v_1, ..., v_{k-1}, v_{k+1}, ..., v_n \in \mathbb{K}^n$ vectores cualesquiera. Vamos a probar que $det(v_1|...|v_{k-1}|v+v'|v_{k+1}|...|v_n) = det(v_1|...|v_{k-1}|v|v_{k+1}|...|v_n)+$ $det(v_1|...|v_{k-1}|v'|v_{k+1}|...|v_n)$. Sean las matrices anteriores $C = (c_{ij})$, $A = (a_{ij})$ y $B = (b_{ij})$ en ese orden. Observemos que $c_{ij} = a_{ij} = b_{ij}$ si $j \neq k$ y $c_{ik} = a_{ik} + b_{ik}$. Ahora,

$$det(C) = \sum_{\sigma \in S_n} sig(\sigma) \prod_{r=1}^{n} c_{\sigma(r)r} = \sum_{\sigma \in S_n} sig(\sigma)(a_{\sigma(k),k} + b_{\sigma(k),k}) \prod_{r=1, r\neq k}^{n} a_{\sigma(r)r} =$$

$$\sum_{\sigma \in S_n} sig(\sigma) \prod_{r=1}^{n} a_{\sigma(r)r} + \sum_{\sigma \in S_n} sig(\sigma) b_{\sigma(k),k} \prod_{r=1, r\neq k}^{n} a_{\sigma(r)r} =$$

$$\sum_{\sigma \in S_n} sig(\sigma) \prod_{r=1}^{n} a_{\sigma(r)r} + \sum_{\sigma \in S_n} sig(\sigma) \prod_{r=1}^{n} b_{\sigma(r)r} = det(A) + det(B)$$

Ahora probaremos que $det(v_1|...|v_{k-1}|\lambda v|v_{k+1}|...|v_n) = \lambda det(v_1|...|v_{k-1}|v|v_{k+1}|...|v_n)$, con $\lambda \in \mathbb{K}$. En este caso denotamos B y A las matrices anteriores en ese orden. Se tiene:

$$det(B) = \sum_{\sigma \in S_n} sig(\sigma) \prod_{r=1}^{n} b_{\sigma(r)r} = \sum_{\sigma \in S_n} sig(\sigma)(\lambda a_{\sigma(k),k}) \prod_{r=1, r\neq k}^{n} a_{\sigma(r)r} = \lambda \sum_{\sigma \in S_n} sig(\sigma) \prod_{r=1}^{n} a_{\sigma(r)r} = \lambda det(A)$$

Para ver que es alternada nos valdremos de que $S_n = \{\sigma \circ \tau : \sigma \in S_n\}$ donde τ es una trasposición cualquiera. En efecto, dada $\sigma \circ \tau \in S_n$ sin importar σ, y además, dada $\sigma \in S_n$ se tiene que $\sigma \circ \tau^{-1} \in S_n$ luego $(\sigma \circ \tau^{-1}) \circ \tau = \sigma$ y se prueba que la reordenación es de esa forma.

Ahora basta con probar que $det(v_1|...|v_l|...|v_k|...|v_n) = -det(v_1|...|v_k|...|v_l|...|v_n)$, es decir, que aplicando una trasposición cambia el signo. Esto es así porque después cualquier reordenación se puede hacer encadenando trasposiciones, y al ir variando el signo según la paridad verifica la definición de alternada. Si A y B son las matrices anteriores, y observando que $b_{ik} = a_{il}$ y $b_{il} = a_{ik}$, se tiene que

$$det(A) = \sum_{\sigma \in S_n} sig(\sigma) \prod_{r=1}^{n} a_{\sigma(r)r} = \sum_{\sigma \in S_n} sig(\sigma) b_{\sigma(l)k} b_{\sigma(k)l} \prod_{r=1, r\neq k, r\neq l}^{n} b_{\sigma(r)r} =$$

$$\sum_{\sigma \in S_n} sig(\sigma \circ \tau_{lk}) b_{\sigma\tau_{lk}(l)k} b_{\sigma\tau_{lk}(k)l} \prod_{r=1, r\neq k, r\neq l}^{n} b_{\sigma\tau_{lk}(r)r} = \sum_{\sigma \in S_n} -sig(\sigma) b_{\sigma(k)k} b_{\sigma(l)l} \prod_{r=1, r\neq k, r\neq l}^{n} b_{\sigma(r)r} = -det(B)$$

El paso final ha sido precisamente escribir cada reordenación como otra compuesta con la trasposición entre l y k, teniendo en cuenta que $sig(\sigma\tau) = -sig(\sigma)$ ya que cambia la paridad. □

Ahora veremos propiedades de las formas multilineales alternadas, que evidentemente se aplican al determinante:

Proposición 57. *Sea $\psi : E^n \mapsto \mathbb{K}$ multilineal. Entonces $\psi(v_1, ..., 0, ..., v_n) = 0$*

Demostración. ψ es lineal en la coordenada donde está el cero luego evaluada en 0 ahí debe dar cero. □

Proposición 58. *Sea $\psi : E^n \mapsto \mathbb{K}$ mutilineal y alternada, y sean $v_1, v_2, ..., v_n \in \mathbb{K}$ tales que $v_i = v_j$ para unos i, j. Entonces $\psi(v_1, ..., v_n) = 0$.*

Demostración. Por ser alternada, $\psi(v_1, ..., v_i, ..., v_j, ..., v_n) = -\psi(v_1, ..., v_j, ..., v_i, ..., v_n)$, pero por ser iguales, $\psi(v_1, ..., v_i, ..., v_j, ..., v_n) = \psi(v_1, ..., v_j, ..., v_i, ..., v_n)$, luego debe ser cero. □

(Segunda nota para el lector entendido: Esto es una continuación de la nota que aparecía en la definición de alternada. Obsérvese que esta demostración solo funciona si la característica de $\mathbb{K}$ es distinta de 2, dado que en caso contrario, el 1 también coincide con su opuesto y no se puede ejecutar este argumento. Es por esto que se suele tomar esta proposición como definición de alternada. Si se busca explícitamente trabajar en característica 2, tómese esta proposición como definición de alternada, y demuéstrese la "otra definición" a partir de esta.)

Proposición 59. *Sea $\psi : E^n \mapsto \mathbb{K}$ multilineal y alternada, y sean $v_1, ..., v_n \in \mathbb{K}$ n vectores linealmente dependientes. Entonces $\psi(v_1, ..., v_n) = 0$.*

Demostración. Se tiene que $\sum \lambda_i v_i = 0$ con $\lambda_{i_0} \neq 0$, luego, despejando, $v_{i_0} = \sum_{i \neq i_0} \beta_i v_i$ para algunos β. Entonces, $\psi(v_1, ..., v_{i_0}, ..., v_n) = \psi(v_1, ..., \sum_{i\neq i_0} \beta_i v_i, ..., v_n) = \sum_{i \neq i_0} \beta_i \psi(v_1, ..., v_i, ..., v_n) = \sum_{i \neq i_0} \beta_i 0 = 0$ dado que en cada evaluación de ψ en el sumatorio hay dos vectores repetidos. □

Observación 31. Si ψ es n-lineal alternada sobre E y $\dim E < n$, entonces $\psi = 0$. Esto es, $\bigwedge_n(E) = \{0\}$ en ese caso.

Razón. Toda n-upla de más vectores que la dimensión se trata de un conjunto linealmente dependiente.

Teorema 20. *Sea E un $\mathbb{K}$-espacio vectorial, de dimensión n. Entonces $\bigwedge_n(E)$ es de dimensión 1. Esto es, $\exists\psi : E^n \mapsto \mathbb{K}$ multilineal alternada tal que $\psi' = \lambda\psi$, $\forall\psi' : E^n \mapsto \mathbb{K}$ multilineal alternada.*

Demostración. Fijamos $B = \{e_1, ..., e_n\} \subset E$ base. Vamos a ver que toda $\psi \in \bigwedge_n(E)$ es múltiplo escalar de la aplicación det, el determinante de la matriz que tiene por columnas las coordenadas de los vectores de entrada escritas en B. Sean $v_1, ..., v_n \in E$. Tenemos que $v_i = \sum_{j_i=1}^n a_{j_i i}e_{j_i}$ $\forall i \in \{1, ..., n\}$. Sea $\psi \in \bigwedge_n(E)$. Entonces, $\psi(v_1, ..., v_n) = \psi(\sum_{j_1=1}^n a_{j_1 1}e_{j_1}, \sum_{j_2=1}^n a_{j_2 2}e_{j_2}, ..., \sum_{j_n=1}^n a_{j_n n}e_{j_n}) = \sum_{j_1,j_2,...,j_n} a_{j_1 1}a_{j_2 2}...a_{j_n n}\psi(e_{j_1}, ..., e_{j_n})$ al ser multilineal. Teniendo en cuenta que todas las combinaciones en que $j_k = j_l$ para algunos k, l se anulan al ser multilineal alternada, quedan solo las que no se repiten, es decir, las reordenaciones de naturales: $\psi(v_1, ..., v_n) = \sum_{\sigma\in S_n}(\prod_{r=1}^n a_{\sigma(r)r})\psi(e_{\sigma(1)}, ..., e_{\sigma(n)}) = \sum_{\sigma\in S_n}(\prod_{r=1}^n a_{\sigma(r)r})sig(\sigma)\psi(e_1, ..., e_n) = det(v_1, ..., v_n)\psi(e_1, ..., e_n)$, donde hemos usado al final que es alternada.

Así:

$$\psi = det_B \cdot \psi(e_1, ..., e_n)$$

Donde det_B indica el determinante para las coordenadas escritas en B. □

Observación 32. $det(Id) = det((1, 0, ..., 0), (0, 1, ..., 0), ...(0, 0, ..., 1)) = 1$

Razón. La única reordenación que no anula el sumando en el determinante es $\sigma = Id$, dando lugar al valor 1.

Observación 33. Si $\psi(e_1, ..., e_n)$ es no nulo para $\psi \in \bigwedge_n(V)$ con V de dimensión n, y ϕ' es otra función de $\bigwedge_n(V)$, despejando del resultado anterior, sigue que $\psi' = \frac{\psi'(e_1,...,e_n)}{\psi(e_1,...,e_n)}\psi$. Esto es consistente con que $det(e_1, ..., e_n) = 1$, para que $\psi' = \psi'(e_1, ..., e_n)det$.

A continuación desarrollaremos que el determinante preserva la multiplicación.

Definición 46. Sea V un $\mathbb{K}$-espacio y $f : E \mapsto E$ un endomorfismo. Sea $\psi : E^n \mapsto \mathbb{K}$ multilineal alternada. Entonces, $\psi_f : E^n \mapsto \mathbb{K}$ dada por $\psi_f(v_1, ..., v_n) = \psi(f(v_1), f(v_2), ..., f(v_n))$ también es multilineal alternada.

Razón. Fijadas $n - 1$ coordenadas, es evidente que la restante es lineal al ser composición de lineales. También, $\psi_f(v_1, ..., v_n) = \psi(f(v_1), ..., f(v_n)) = -\psi(f(v_{\tau(1)}), ..., f(v_{\tau_n})) = -\psi_f(v_{\tau(1)}, ..., v_{\tau_n})$ luego es alternada.

Teorema 21. *Sean $A, B \in \mathcal{M}_n(\mathbb{K})$ y sea $det : \mathbb{K}^n \times \mathbb{K}^n \times ... \times \mathbb{K}^n \mapsto \mathbb{K}$ el determinante n-lineal. Entonces, $det(BA) = det(B)det(A)$. (interpretando como determinante de una matriz el de sus columnas)*

Demostración. Trabajamos en base canónica de $\mathbb{K}^n$. Sea $f : \mathbb{K}^n \mapsto \mathbb{K}^n$ el endomorfismo de matriz B. Sea det_f la aplicación construida como en la definición anterior a partir del determinante y de f. Veamos que $det_f(A) = det(f(v_1), f(v_2), ..., f(v_n)) = det(Bv_1, Bv_2, ..., Bv_n) = det(BA)$, si $v_1, ..., v_n$ son las columnas de A. Por otra parte, $det_f = det_f(e_1, ..., e_n)det = det(Be_1, Be_2, ..., Be_n)det = det(B)det$. Por ende, $det_f(A) = det(B)det(A)$, así que son iguales. □

Proposición 60. *Si $S \in \mathcal{M}_n(\mathbb{K})$, entonces $rg(S) = n \iff S$ es invertible $\iff det(S) \neq 0$.*

Demostración. Sabemos que si $rg(S) = n$, S es invertible. Ahora, si S es invertible, $det(SS^{-1}) = 1 \implies det(S)det(S^{-1}) = 1$, luego $det(S) \neq 0$. Finalmente, probaremos que $det(S) \neq 0 \implies rg(S) = n$. Esto es así porque si $rg(S) < n$, entonces $det(S) = 0$ al ser las columnas linealmente dependientes. □

Proposición 61. *Sea $f : V \mapsto V$ endomorfismo con $\dim V = n$. Sean B y B' bases de V cualesquiera. Se tiene que $det(|f|_B) = det(|f|_{B'})$, es decir, el determinante es independiente de la elección de base.*

Demostración. Esto sigue de que $|f|_{B'} = |id|_{BB'}|f|_B|id|_{B'B} = |id|^{-1}_{B'B}|f|_B|id|_{B'B}$. Así, $det(|f|_{B'}) = \frac{det(|id|_{B'B})}{det|id|_{B'B}} det(|f|_B)$ □

Definición 47. Dada $A = (a_{ij})$ se define la **adjunta** $adj(A) \in \mathcal{M}_n\mathbb{K}$ como aquella tal que $adj(A)_{ij} = (-1)^{i+j}A_{ij}$, donde A_{ij} es el determinante de A tras eliminar la i-ésima fila y la j-ésima columna.

Teorema 22. *Se tiene que* $det(A) = \sum_{j=1}^n a_{ij} \cdot adj(A)_{ij}$ *sin importar* $i = \{1, 2, ..., n\}$.

De este método alternativo para calcular el determinante, se extrae también que $A \cdot (adj(A))^t$ es una matriz que vale $det(A)$ en la diagonal principal, ya que se estaría llevando a cabo una suma como la anterior. Es más, cualquier otro elemento de esta matriz es cero: si B es esta matriz, $b_{ij} = \sum_{k=1}^n a_{ik} \cdot adj(A)_{jk}$, y si $i \neq j$, esta suma también describe $det(A')$, donde A' es la matriz que tiene la i-ésima y la j-ésima filas iguales (en la j-ésima se copia la i-ésima), desarrollado por la j-ésima fila. Por tanto, al ser dos filas iguales, ese valor es cero. De aquí sigue directamente que:

Proposición 62. $A \cdot adj(A)^t = det(A) \cdot Id_n$

Esto, por ejemplo, da una manera de calcular la inversa de A, despejando de arriba. Además, estas matrices conmutan:

Observación 34. $adj(A)^t \cdot A = det(A) \cdot Id_n$

Razón. En este caso el elemento b_{ii} de la diagonal principal del producto es $b_{ii} = \sum_{k=1}^n adj(A)_{ki}a_{ki} = det(A)$ al ser el desarollo por la i-ésima columna. Asimismo, si $i \neq j$, se tiene $b_{ij} = \sum_{k=1}^n adj(A)_{ki}a_{kj} = 0$ al ser el desarrollo del determinante de A'', matriz con una copia de la j-ésima columna en la i-ésima, desarrollado por la i-ésima.

Teorema 23 (Cramer)**.** *Sea* $A = (a_{ij}) \in \mathcal{M}_n(\mathbb{K})$ *y sean* $b_1, ..., b_n \in \mathbb{K}$ *vectores dados. Consideramos el sistema* $A \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}$. *Se verifica entonces que:* $x_i = \frac{det(A_i)}{det(A)}$, *donde* A_i *es* A *con la i-ésima columna sustituida por la de términos independientes,* $\forall i \in \{1, ..., n\}$.

Demostración. Veamos que si $AX = B$, entonces, $adj(A)^tAX = adj(A)^tB$, entonces $det(A)Id_nX = adj(A)^tB$, de modo que:

$$\begin{pmatrix} det(A)x_1 \\ det(A)x_2 \\ \vdots \\ det(A)x_n \end{pmatrix} = adj(A)^tB$$

.

De aquí sigue que $det(A)x_i = [adj(A)^tB]_i = \sum_{k=1}^n adj(A)_{ki}b_k$, y este desarrollo coincide con el $det(A_i)$ descrito en el enunciado. □

Observación 35. Una matriz A con coeficientes enteros es invertible en $\mathcal{M}_n(\mathbb{Z})$ si y solo si $det(A) \in \{1, -1\} = \mathcal{U}(\mathbb{Z})$

En general, restringiendo la definición de determinante a cualquier anillo, puede obtenerse este resultado (invertibilidad si y solo si el determinante es unidad). Razón. Si es invertible, $det(A)det(A^{-1}) = 1$, para lo que $det(A) \in \{1, -1\}$, puesto que cualquier otro no tiene inverso multiplicativo en $\mathbb{Z}$. Por otra parte, si $det(A) = 1$, entonces $A \cdot adj(A)^t = I$ y si vale -1, $A \cdot (-adj(A)^t) = 1$, y todas son de coeficientes enteros dado que restringiendo la definición de determinante a una matriz de $\mathcal{M}_n(\mathbb{Z})$ se obtiene un entero. □

8. Estructura de los endomorfismos

Dado un $\mathbb{K}$-espacio vectorial, E, existen multitud de endomorfismos, como las homotecias: $f_\lambda : E \mapsto E$ tal que $f(v) = \lambda v$ fijado $\lambda \in \mathbb{K}$. Es fácil ver una característica de estas: su matriz es $\lambda \cdot I$ sin importar la base fijada para E. Vamos a estudiar cuándo y qué otros endomorfismos presentan aunque sea una base en la que la matriz resultante es diagonal y por tanto mucho más simple (no solo permite operar con mayor sencillez, por ejemplo, elevar a potencias altas, sino que también permite entender mejor cómo se comporta el endomorfismo).

Definición 48. Sea E un $\mathbb{K}$-espacio vectorial y $f : E \mapsto E$ un endomorfismo. Se dice que $v \in E$ es un **autovector** de f si:

1. $v \neq 0$
2. $\exists \lambda \in \mathbb{K}$ t.q. $f(v) = \lambda v$

Se dice en ese caso que λ es un **autovalor** de f, y que v es el autovector correspondiente a λ.

El siguiente es un criterio para hallar los autovalores:

Proposición 63. *Sea $f : E \mapsto E$. Entonces, λ es autovalor de $f \iff f - \lambda Id_E$ no es isomorfismo ($\equiv$ no es monomorfismo).*

Demostración. λ es autovalor $\iff \exists v \in E, v \neq 0$ t.q. $f(v) = \lambda v \iff f(v) - \lambda v = 0 \iff (f - \lambda Id_E)(v) = 0 \iff \exists v \neq 0, v \in Nuc(f - \lambda Id_E) \iff f - \lambda Id_E$ no es monomorfismo. □

En otras palabras:

Proposición 64. *Sea $f : E \mapsto E$. Entonces, λ es autovalor de f de matriz asociada $A \iff rg(A - \lambda I) < dimE \iff det(A - \lambda I) = 0$.*

Observación 36. 0 es autovalor de $f \iff f$ no es isomorfismo.

Proposición 65. *Sean $\{\lambda_i\}_1^s$ tales que $\lambda_i \neq \lambda_j$ cuando $i \neq j$, un conjunto de autovalores (distintos) de $f : E \mapsto E$. Entonces, los autovectores $\{v_1, ..., v_s\}$ correspondientes (cada v_i a su λ_i) son un conjunto linealmente independiente.*

Demostración. Razonamos por inducción. Está claro que $\{v_1\}$ es linealmente independiente al ser no nulo (es un autovector). Ahora, si se tiene que $s-1$ autovectores de f con las características anteriores son linealmente independientes, veremos que lo son s. Supongamos que $\sum_1^s a_i v_i = 0$. Tomando imágenes, $\sum_1^s a_i \lambda_i v_i = 0$. Multiplicando la expresión inicial por λ_1, $\sum_1^s a_i \lambda_1 v_i = 0$. Restando ambas, $\sum_2^s a_i(\lambda_i - \lambda_1) v_i = 0$. Ahora bien, estos son $s-1$ autovectores de valores distintos y, por hipótesis, $a_i(\lambda_i - \lambda_1) = 0$ $\forall i = 2, ..., s$. No obstante, como $\lambda_i \neq \lambda_1$ en ningún caso, sigue que $a_i = 0$ $\forall i = 2, ..., s$, y entonces queda que $a_1 v_1 = 0$, por lo que $a_1 = 0$ al ser v_1 no nulo. □

Corolario. Si $f : E \mapsto E$ es un endomorfismo en un espacio tal que $\dim E = n$, tiene a lo sumo n autovalores distintos.

Demostración. Si no, podríamos generar conjuntos de más de n vectores linealmente independientes, contradiciendo que sea de dimensión n. □

Observación 37. Si en un espacio n-dimensional se tienen n autovalores distintos del endomorfismo f, inducen una base $\{v_1, ..., v_n\}$ donde por definición la matriz de f es diagonal.

Definición 49. Un endomorfismo $f : V \mapsto V$ es **diagonalizable** si $\exists B \subset V$ base tal que $|f|_B$ es diagonal.

Observación 38. f es diagonalizable $\iff \exists \dim V$ autovectores distintos que forman base.

Razón. Si f es diagonalizable, la base en la cual la matriz es diagonal es de autovectores, puesto que si v_i está en dicha base, y a_{ii} es el i-ésimo elemento de la diagonal de la matriz, entonces $f(v_i) = a_{ii}v_i$ al ser la matriz de f y por tanto v_i es autovalor. Por la misma razón, si existe tal base, la matriz en dicha base es diagonal.

Definición 50. Sea $\{\lambda_1, ...\lambda_n\}$ el conjunto de los autovalores de $f : E \mapsto E$. Se define el **conjunto de autovectores** de λ_i por $S_{\lambda_i} = Nuc(f - \lambda_i Id_E) = \{v \in E : f(v) = \lambda_i v\}$. Se trata del vector 0 junto con los autovectores de λ_i. Evidentemente, al ser un núcleo, se trata de un subespacio.

Proposición 66. *S_{λ_i} definido como anteriormente es invariante por f. ($f(S_{\lambda_i}) \subset S_{\lambda_i}$)*

Demostración. Si $v \in S_{\lambda_i}$, entonces $f(v) = \lambda_i v \in S_{\lambda_i}$ al ser subespacio. □

Definición 51 (Suma directa de más de dos subespacios)**.** Se dice que los subespacios $S_1, ..., S_n$ están en suma directa si $\dim(S_1 + S_2 + ... + S_n) = \dim S_1 + \dim S_2 + ... + \dim S_n$. No es suficiente que estén dos a dos en suma directa.

Proposición 67. *Dado $f : E \mapsto E$ endomorfismo con E un $\mathbb{K}$-espacio vectorial, si $\{\lambda_1, ..., \lambda_s\}$ son todos los autovalores de f, entonces $S_{\lambda_1} + S_{\lambda_2} + ... + S_{\lambda_s}$ es suma directa.*

Demostración. Por inducción. Veamos que $S_{\lambda_1} + S_{\lambda_2}$ es directa: si $v \in S_{\lambda_1} \cap S_{\lambda_2}$, $f(v) = \lambda_1 v = \lambda_2 v$ al estar en ambos, pero como $\lambda_1 \neq \lambda_2$, sigue que $v = 0$ luego la suma es directa. Ahora probaremos que si $S_{\lambda_1} + ... + S_{\lambda_{s-1}}$ están en suma directa, entonces $(S_{\lambda_1} + ... + S_{\lambda_{s-1}}) + S_{\lambda_s}$ es directa, es decir, que $(S_{\lambda_1} + ... + S_{\lambda_{s-1}}) \cap S_{\lambda_s} = \{0\}$, dado que así inductivamente su dimensión será la suma de las dimensiones. Sea v no nulo en la intersección. Entonces $v = v_1 + ... + v_{s-1}$ de forma única, y no todos los vectores nulos, y además $v \in S_{\lambda_s}$. Por ese motivo, tomando funciones, $\sum_1^{s-1} \lambda_i v_i = \lambda_s v$, no todos los λ nulos al haber al menos 2 y ser distintos. Reordenando, queda una combinación lineal no trivial de dichos vectores igualada a cero (menos los nulos, que no son todos), contradiciendo que sean autovectores de valores distintos. □

Corolario. $\sum \dim(S_{\lambda_i}) \leq \dim E$, al estar la suma contenida en E y coincidir en dimensión. De hecho, $\sum \dim(S_{\lambda_i}) = \dim E \iff \bigoplus_1^s S_{\lambda_i} = E$, en cuyo caso las bases de cada S_{λ_i} concatenadas dan lugar a base de E, y al ser autovectores, es diagonalizable:

Proposición 68. $\sum \dim(S_{\lambda_i}) = \dim E \iff \bigoplus_1^s S_{\lambda_i} = E \iff f$ *es diagonalizable.*

Demostración. Con todo lo dicho anteriormente, basta probar que si f es diagonalizable la suma es directa. Pero esto sigue de que si la inclusión de la suma directa fuese estricta ($\bigoplus_1^s S_{\lambda_i} \subsetneq E$), habría $v \in E$ que no se podría expresar como combinación de autovectores al no estar en la suma directa, contradiciendo que f sea diagonalizable. □

Definición 52. La **multiplicidad geométrica** del autovalor λ_i es la dimensión de S_{λ_i}. Coincide con el número de veces que figura en la diagonzalización, si es diagonalizable, según lo visto anteriormente. Es decir, $mult(\lambda_i) = \dim(Nuc(f - \lambda_i Id))$.

De lo anterior se extrae que $\sum_1^s Mult(s) \leq \dim E$, con la igualdad si y solo si es diagonalizable.

Definición 53 (Autovalores/vectores de una matriz)**.** Dada $M \in \mathcal{M}_n(\mathbb{K})$, define un endomorfismo $f_M : \mathbb{K}^n \mapsto \mathbb{K}^n$ dado por $f_M(v) = Mv$. Los autovectores y autovalores de la matriz M son los del morfismo f_M. Por ello, λ es autovector $\iff \det(A - \lambda I) = 0$.

Definición 54. La matriz $A \in \mathcal{M}_n(\mathbb{K})$ es diagonalizable si y solo si lo es f_A, es decir, si y solo si $\exists C \in \mathcal{M}_n(\mathbb{K})$ tal que CAC^{-1} es diagonal.

8.1. Polinomio característico de un endomorfismo

Con el fin de facilitar la teoría de autovectores y diagonalización, se introduce un polinomio que encierra la información relacionada a estos aspectos.

Definición 55. Se define el **polinomio característico** de un endomorfismo/matriz como el polinomio dado por $P_f(X) = det(A - XI)$ donde A es la matriz del morfismo. Como veremos después, está bien definido al no depender de la base en la que se construye la matriz A.

Observación 39. $\lambda \in \mathbb{K}$ es autovalor de $f \iff P_f(\lambda) = 0 \iff (X - \lambda)|P_f(X)$.

Proposición 69. *Sea f un endomorfismo de matriz A y sea $C \in \mathcal{M}_n(\mathbb{K})$ una matriz de cambio de base, entonces la matriz CAC^{-1} es la de f en dicha otra base. El polinomio característico es el mismo para ambas.*

Demostración. $det(CAC^{-1} - XI) = det(C(A - XI)C^{-1}) = det(C)det(C^{-1})det(A - XI) = det(A - XI)$. □

Observación 40. Si $f : V \mapsto V$ es un endomorfismo con $\dim V = n$, entonces $\deg(P_f) = n$. Por tanto, esta es otra prueba de que a lo sumo hay n autovalores.

Esto sigue de la definición de determinante: con la reordenación identidad obtenemos la traza de la matriz (polinomio de grado n) y con cualquier otra algunos de los elementos no tendrán a la indeterminada X, y su grado será menor de n. La suma entonces será como mucho de grado n.

Definición 56. La matriz A es **triangulable** si es semejante a una triangular superior, esto es $\exists C$ de cambio de base tal que CAC^{-1} es triangular superior.

Proposición 70. *Si $A \in \mathcal{M}_n(\mathbb{K})$ tiene polinomio característico con menos de n raíces (contando varias veces las multiplicidades algebraicas), entonces no es triangulable, y por tanto tampoco diagonalizable.*

Demostración. Si fuese triangulable con el cambio C, el polinomio característico de CAC^{-1} sería únicamente la traza de $CAC^{-1} - XI = (b_{ij})_{i,j}$, que tendría n raíces al ser directamente $\prod_1^n (b_{ii} - X)$. Pero entonces el característico de A sería el mismo, contradiciendo la hipótesis. □

Teorema 24. *Sea $f : E \mapsto E$ un endomorfismo de polinomio $p_f(X)$ con los ceros en $\{\lambda_1, ..., \lambda_s\}$, los cuales son los autovalores de f. Entonces, $(X - \lambda_i)^{r_i}|p_f(X)$, donde $r_i = mult(\lambda_i) = \dim Nuc(f - \lambda_i Id)$, $\forall i \in \{1, ..., s\}$.*

En otras palabras, si R_i es la multiplicidad algebraica de λ_i como cero de $p_f(X)$, entonces $mult(\lambda_i) \leq R_i$.

Demostración. Fijado el i, sea $\{v_1, ..., v_{r_i}\}$ una base de S_{λ_i}. Extendemos a base de E: $B = \{v_1, ..., v_{r_i}, ..., v_n\}$. Ahora, como hasta v_{r_i} son autovectores de λ_i, sigue que la matriz del morfismo es:

$$|f|_B = \left(\begin{array}{c|c} \lambda_i I_{r_i} & A \\ \hline 0 & B \end{array} \right)$$

Donde A, B contienen las coordenadas cualesquiera de las imágenes de los vectores añadidos al extender. Por tanto, $P(X)_f = det(\lambda_i I_{r_i} - X)det(B) = (\lambda_i - X)^{r_i} det(B)$ donde $det(B)$ es otro polinomio. □

8.2. Polinomio mínimo

Definición 57. Dado un $\mathbb{K}$-espacio, E, y $f \in End(E)$, para cada polinomio $P(X) = \sum_{i=0}^n a_i X^i \in \mathbb{K}[X]$ podemos considerar el endomorfismo $P(f) = \sum_{i=0}^n a_i f^i$, donde por suma y producto por escalar de morfismos entendemos los usuales, y por producto de morfismos entendemos la composición (siendo $f^0 = Id$). Con esto, es evidente que $P(f) \in End(f)$.

Observación 41. Sea el que sea $f \in End(E)$, se tiene que los endomorfismos $P(f), Q(f)$ conmutan: $P(f) \circ Q(f) = Q(f) \circ P(f)$. Esto sigue de que $P(X)Q(X) = Q(X)P(X)$, con lo que los morfismos asociados deben ser los mismos.

Análogamente:

Definición 58. Dada $A \in \mathcal{M}_n(\mathbb{K})$, para cada polinomio $P(X) = \sum_{i=0}^n a_i X^i \in \mathbb{K}[X]$ podemos considerar la matriz $P(A) = \sum_{i=0}^n a_i A^i$, donde las operaciones entre matrices son las usuales, y $A^0 = I$).

Observación 42. Fijada $A \in \mathcal{M}_n(\mathbb{K})$, las matrices $P(A)$ y $Q(A)$ deben conmutar, al ser $P(X)Q(X) = Q(X)P(X)$.

Teorema 25 (Cayley-Hamilton)**.** *Sea E un $\mathbb{K}$-espacio vectorial, con $f \in End(E)$ y $P_f(X) \in \mathbb{K}[X]$ su polinomio característico. Se verifica que $P_f(f) = 0$, es decir, el endomorfismo dado por el polinomio característico en f es el nulo.*

(Equivalentemente) Sea $A \in \mathcal{M}_n(\mathbb{K})$ y $P_A(X) \in \mathbb{K}[X]$ su polinomio característico. Se verifica que $P_A(A) = 0$, es decir, la matriz dada dado por el polinomio característico en A es la nula.

Definición 59. Denotamos $I_f \subset \mathbb{K}[X]$ al subconjunto $I_f = \{q(X) \in \mathbb{K}[X] : q(f) = 0\}$. El teorema de Cayley-Hamilton indica que $P_f(X) \in I_f$.

Proposición 71. *$\exists m_f(X) \in \mathbb{K}[X]$ tal que $m_f(X) \in I_f$, y además: $I_f = \{m_f(X)h(X) : h(X) \in \mathbb{K}[X]\} = m_f(X)\mathbb{K}[X]$.*

Demostración. Sea $m_f(X) \in I_f$ el polinomio de menor grado no nulo de I_f, el cual existe porque por lo menos $P_f(X) \in I_f$. Observemos que entonces si $h(X) \in \mathbb{K}[X]$, el endomorfismo en f de $m_f(X)h(X)$ es: $m_f(f)h(f) = 0 \circ h(f) = 0$. Además, todos son de esa forma: sea $g(X) \in I_f$, entonces $g(X) = q(X)m_f(X) + r(X)$ donde $r(X)$ es nulo o bien $gr(r) < gr(m_f)$. Pero sabemos que $0 = g(f) = q(f) \circ m_f(f) + r(f) = q(f) \circ 0 + r(f) = r(f)$, por tanto $r(X) = 0$, ya que de otro modo estaría en I_f y con grado menor que m_f. □

Esto da pie a la definición:

Definición 60. Tal $m_f(X)$, mónico, se denomina **polinomio mínimo** de f. Se trata del polinomio mónico de menor grado tal que $m_f(f) = 0$, y todos los demás polinomios p con $p(f) = 0$ son múltiplos suyos.

El mismo razonamiento sirve para matrices, y entonces:

Definición 61. El **polinomio mínimo**, $m_A(X)$, de $A \in \mathcal{M}_n(\mathbb{K})$ se trata del polinomio mónico de menor grado tal que $m_A(A) = 0$, y todos los demás polinomios p con $p(A) = 0$ son múltiplos suyos.

Observación 43. $m_f(X) | p_f(X)$.

Teorema 26. *Todo factor irreducible de $p_f(X)$ divide a $m_f(X)$. Así, los autovalores de f anulan también a $m_f(X)$. Con matrices ocurre lo mismo.*

Teorema 27. *Si A, A' son matrices semejantes ($A' = PAP^{-1}$) entonces $m_A(X) = m_{A'}(X)$.*

8.3. Subespacios invariantes

En adelante todos los resultados se aplican para endomorfismos y matrices por igual (recordemos la identificación que hay entre ambos).

Definición 62. Un **subespacio invariante** por $f \in End(E)$ es $S \subset E$ un subespacio tal que $f(S) \subset S$. Además, tomando $B = \{v_1, .., v_s\}$ base de S y extendiendo a base de V, $\tilde{B} = \{v_1, ..., v_n\}$, es fácil ver que se tiene una cómoda expresión para la matriz de f:

$$|f|_{\tilde{B}} = \left(\begin{array}{c|c} |f|_S|_B & A \\ \hline 0 & B \end{array}\right)$$

con A y B las entradas restantes.

Observación 44. Sea $f \in End(E)$. Si $P_f(X)$ tiene un factor de grado 1, f tiene subespacio invariante de dimensión 1, dado que tendrá un autovalor λ_1 y el autovector correspondiente, v_1, genera dicho invariante como ya vimos. Si f tiene dos autovalores distintos λ_1, λ_2, también es fácil comprobar que $\{v_1, v_2\}$ dos autovectores correspondientes a estos generan un invariante de dimensión 2. Así sucesivamente.

Observación 45. Por tanto, si se considera E en los complejos, toda f da lugar a un subespacio invariante de dimensión 1.

Proposición 72. *Sea E un $\mathbb{R}$-espacio vectorial. Entonces todo $f \in End(E)$ admite un subespacio invariante de dimensión 2 o menor.*

Demostración. Se tiene que $P_f(X) = \lambda \prod_1^n Q_i(X)^{r_i}$, con cada $Q_i(X)$ irreducible y mónico, por tanto de grado 1 o 2. Si hay algún Q_i de grado 1 entonces $P_f(X)$ tiene una raíz real y hemos acabado. Si no, podemos considerar $P_f(X) \in \mathbb{R}[X]$ como polinomio de $\mathbb{C}[X]$, y E como un $\mathbb{C}$-espacio (permitimos escalares complejos). Veamos que $Q_1(X) = (X - \lambda)(X - \bar{\lambda})$ para algún $\lambda \in \mathbb{C} \backslash \mathbb{R}$, con lo cual, si ponemos $A \in \mathcal{M}_n(\mathbb{R})$ la matriz de f en alguna base, sigue que $Az = \lambda z$ para algún $z \in \mathbb{C}^n$ correspondiente a las coordenadas de un autovector en E considerado como complejo. En adelante trabajaremos con dichas coordenadas por comodidad.

Ahora bien, en ese caso $\bar{z}$ es el autovector de $\bar{\lambda}$, dado que $\bar{A}\bar{z} = \bar{\lambda}\bar{z}$, pero $\bar{A} = A$ al ser de coeficientes reales. Por tanto, el espacio $S = \langle z, \bar{z} \rangle$ es invariante de dimensión 2 bajo la matriz A, considerando el endomorfismo sobre E como $\mathbb{C}$-espacio. No obstante, $S = \langle z + \bar{z}, z - \bar{z} \rangle$, donde se ha realizado un cambio de base, y entonces $S =< Re(z), Im(z) >$, con $Re(z)$ el vector de partes reales de z e $Im(z)$ el de imaginarias (se han simplificado los escalares).

Ahora volvemos a considerar $S \subset E$ como $\mathbb{R}$-espacio, generado por dichos $Re(z) = \mathbf{x}, Im(z) = \mathbf{y}$. Supongamos que $\lambda = a + bi$. Veamos que $Az = A(\mathbf{x} + i\mathbf{y}) = (a + bi)(\mathbf{x} + i\mathbf{y})$, por tanto, igualando partes reales e imaginarias en cada coordenada de los dos vectores, sigue que: $A\mathbf{x} = a\mathbf{x} - b\mathbf{y}$ y $A\mathbf{y} = b\mathbf{x} + a\mathbf{y}$. Es decir, dicho $S = \langle \mathbf{x}, \mathbf{y} \rangle$ es invariante por f, dado que las imágenes de la base siguen siendo generadas por dicha base. Sigue además que la matriz de f en la base $\{\mathbf{x}, \mathbf{y}, v_3, ..., v_n\}$ que se obtiene de extender a base de E debe ser:

$$\tilde{A} = \left(\begin{array}{cc|c} a & b & X \\ -b & a & Y \\ \hline 0 & 0 & Z \end{array}\right)$$

Siendo el autovalor complejo $\lambda = a + bi$ □

Observación 46. Si generamos invariantes de orden 1 y 2 con cada factor del polinomio característico de $\mathbb{R}[X]$, tendrán intersección nula.

Esto es así porque cada invariante tiene los autovalores (reales, o complejos si ha sido con el procedimiento anterior) correspondientes a las raíces del factor con que han sido generados. De modo que de haber un vector no nulo en dos de ellos, tendría dos autovalores correspondientes.

Recordemos que si conseguimos encontrar una base del espacio E, podremos convertir la matriz del endomorfismo de $End(E)$ en una de gran simplicidad, compuesta por bloques en la diagonal (las submatrices del endomorfismo en el invariante) y ceros en el resto. El siguiente resultado será de importancia en las próximas discusiones:

Proposición 73. *Sean $A \in \mathcal{M}_n(\mathbb{K})$ y sea $B \in Conm(A) = \{B \in \mathcal{M}_n(\mathbb{K}) : AB = BA\}$. Sigue que $Nuc(B)$ es invariante por A.*

Demostración. Sea $v \in Nuc(B)$. Entonces $Bv = 0$, con lo que $ABv = A0 = 0$, pero $ABv = BAv$, luego $B(Av) = 0$ y entonces $Av \in Nuc(B)$. □

Ya vimos una fácil forma de encontrar matrices que conmuten:

Proposición 74. *Sea $P(X) \in \mathbb{K}[X]$ y $A \in \mathcal{M}_n$. Entonces $P(A) \in Conm(A)$, y por tanto $Nuc(P(A))$ es invariante por A.*

Demostración. Vimos con anterioridad que dados dos polinomios, las matrices resultantes de evaluar A en ellos conmutan. Consideramos X y $P(X)$, como $XP(X) = P(X)X$ entonces $AP(A) = P(A)A$. □

Proposición 75. *Sean $P(X), Q(X) \in \mathbb{K}[X]$ dos polinomios que no tengan factores irreducibles comunes en su factorización. Dada $A \in \mathcal{M}_n$, entonces $Nuc(P(A)) \oplus Nuc(Q(A))$, es decir, se hallan en suma directa.*

De hecho, dados $P_1, P_2, ..., P_n \in \mathbb{K}[X]$ coprimos dos a dos, entonces los invariantes por A, $P_1(A), P_2(A), ...P_n(A)$ están en suma directa, es decir, $\sum_1^n \dim(Nuc(P_i(A))) = \dim \sum_1^n Nuc(P_i(A))$.

Demostración. Probamos el primer enunciado, siendo el otro completamente análogo. Sea $v \in Nuc(P_1(A)) \cap Nuc(P_2(A))$. Al ser coprimos, la identidad de Bézout asegura que $\exists f, g \in \mathbb{K}[X]$ tales que $f(X)P_1(X) + g(X)P_2(X) = 1$, por lo tanto se tiene la igualdad de matrices: $f(A)P_1(A) + g(A)P_2(A) = Id$. Ahora aplicamos la matriz a v: $f(A)P_1(A)v + g(A)P_2(A)v = 0 + 0 = 0 = Id \cdot v$, pero la identidad es biyectiva, luego $v = 0$. □

Siempre que los espacios invariantes obtenidos así sean no nulos (cosa que en general no sucederá), esto permite una forma de hallar invariantes en suma directa (permitiendo que sus vectores sean completamente independientes) para simplificar la matriz. El siguiente teorema además nos da una forma de hallar los invariantes para que den lugar a base del espacio ambiente:

Teorema 28. *Sea $A \in \mathcal{M}_n(\mathbb{K})$ y $P_A(X) \in \mathbb{K}[X]$ el característico. Si $P_A(X) = \lambda \prod_{i=1}^s (X - \lambda_i)^{r_i}$ es producto de factores lineales, entonces $\mathbb{K}^n = \bigoplus_{i=1}^s Nuc((A - \lambda_i Id)^{r_i})$.*

Es decir:

Teorema 29. *Sea $f \in End(V)$ y $P_f(X) \in \mathbb{K}[X]$ el característico. Si $P_f(X) = \lambda \prod_{i=1}^s (X - \lambda_i)^{r_i}$ es producto de factores lineales, entonces $V = \bigoplus_{i=1}^s Nuc((f - \lambda_i Id)^{r_i})$.*

Esto permite lo que se conoce como **simplificación fuerte de la matriz** o del endomorfismo: si tomamos como base B la unión de las bases de cada $S_i = Nuc((f - \lambda_i Id)^{r_i})$, (es base por la suma directa), queda una matriz con bloques diagonales y ceros en el resto, gracias a ser invariantes. A continuación estudiamos una característica fundamental de estos bloques, pero antes:

Observación 47. Si denotamos $S_i = Nuc((f - \lambda_i Id)^{r_i})$ en el supuesto anterior, entonces $v \in S_i \iff (f - \lambda Id)^{r_i}(v) = 0$, es decir, $(f|_{S_i} - \lambda_i Id_{S_i})^{r_i} = 0$. Restringir a S_i ese morfismo es el nulo. Por tanto, $(X - \lambda_i)^{r_i}$ es un polinomio anulador de $f|_{S_i}$.

Proposición 76. *Manteniendo la notación anterior, el polinomio característico de $f|_{S_i}$ es $(X - \lambda_i)^{r_i}$, es decir, el mismo factor lineal irreducible elevado a la misma potencia que el que generó S_i.*

Demostración. Como $(X - \lambda_i)^{r_i}$ está en el anulador polinómico de $f|_{S_i}$, entonces $m_{f|_{S_i}}(X) | (X - \lambda_i)^{r_i}$, de modo que $m_{f|_{S_i}}(X) = (X - \lambda_i)^{t_i}$, con $t_i \leq r_i$. Ahora bien, como el mínimo tiene todos los factores irreducibles del característico, entonces $P_{f|_{S_i}} = (X - \lambda_i)^{p_i}$ para algún $p_i \geq t_i$. Además, si observamos la matriz de f en la base B de V descrita anteriormente como unión de las bases de invariantes, es inmediato que $P_f(X) = \prod_1^s P_{f|_{S_i}}(X)$, luego no queda otra que $P_{f|_{S_i}} = (X - \lambda_i)^{r_i}$. □

Es decir, el polinomio de cada "bloque", que es la matriz $|f|_{S_i}|$ en la base del invariante, es un factor lineal a una potencia (la potencia con la que aparecía). Ahora nos centramos en simplificar estos:

Definición 63. Un **bloque o matriz de Jordan elemental** es una matriz de $n \times n$ de la forma: $\begin{pmatrix} \lambda & 1 & 0 & \dots & 0 \\ 0 & \lambda & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & \lambda \end{pmatrix}$ con $\lambda \in \mathbb{K}$. Veamos que tiene característico $(X-\lambda)^n$. Se trata de una matriz casi diagonal, con 1 en la diagonal superior. Además, es muy fácil hallar las potencias de $A - \lambda I$ donde A es el bloque. Cada vez que aumenta el orden de la potencia, la diagonal de 1 se desplaza una posición hacia la derecha, hasta que eventualmente la matriz se convierte en la nula.

Teorema 30. *Dada $B \in \mathcal{M}_n(\mathbb{K})$ que tenga característico $P_B(X) = (X-\lambda)^n$ para un $\lambda \in \mathbb{K}$, $\exists C \in \mathcal{M}_r(\mathbb{K})$ no singular tal que $C^{-1}BC = A$, con A una matriz que tiene la diagonal de bloques de Jordan y el resto ceros.*

Es decir, toda matriz con el característico de esa forma (como los sub-bloques que quedaban al simplificar la matriz con invariantes) es semejante a una matriz con bloques de Jordan en la diagonal.

A continuación estudiaremos distintos conceptos que permitirán profundizar en la teoría de los bloques de Jordan. La ventaja de simplificar una matriz en bloques diagonales es que operar con ellos es mucho más fácil. Por ejemplo, elevar la matriz a una potencia se reduce a elevar los bloques. También se pueden multiplicar matrices a bloques. Para ver cómo demostrar que toda matriz de característico $(X-\lambda_i)^{r_i}$ se reduce a bloques de Jordan, comenzaremos viendo la idea de nilpotencia:

Definición 64. Sea $f \in End(V)$ $(B \in \mathcal{M}_n(\mathbb{K}))$. Se dice que es **nilpotente** si $\exists k \in \mathbb{N}$ t.q. $f^k = 0$ $(B^k = 0)$. Recordemos que f^k denota $f \circ f \circ ... \circ f$, k veces.

En adelante hablaremos equivalentemente de endomorfismos y matrices (aunque no se indique explícitamente para ambos). Un resultado sorprendente es que si una matriz no se anula tras elevarla un número de veces igual a su orden, no lo hace nunca:

Teorema 31. *Sea $B \in \mathcal{M}_n(\mathbb{K})$. Son equivalentes:*

1. *B es nilpotente.*

2. *$P_B(X) = X^n$*

3. *$B^n = 0$.*

Demostración. Si B es nilpotente, $\exists k$ t.q. $X^k \in I_B$, donde $I_B = \{P(X) \in \mathbb{K}[X] : P(B) = 0\}$ es el ideal anulador de B. Pero entonces, $m_B(X) = X^t$ para $t \leq k$, y como el mínimo y el característico comparten factores irreducibles, no queda otra que $P_B(X) = X^q$ para algún $q \in \mathbb{N}$. No obstante, el grado del característico debe ser n, luego $P_B(X) = X^n$. Por el teorema de Cayley-Hamilton, $B^n = 0$. Finalmente, si $B^n = 0$, entonces B es nilpotente. □

Por ejemplo, si B tiene característico $P_B(X) = \lambda \prod_1^s (X-\lambda_i)^{r_i}$, entonces la matriz $B_i - \lambda_i Id$, donde B_i es la matriz de la restricción del morfismo asociado a $Nuc(B - \lambda_i Id)^{r^i}$, es nilpotente, dado que ya vimos que B_i tenía característico $P_{B_i}(X) = (X-\lambda_i)^{r_i}$.

El siguiente teorema permitirá demostrar la descomposición en bloques de Jordan:

Teorema 32. *Sea $B \in \mathcal{M}_n(\mathbb{K})$. B es nilpotente $\iff$ $B \sim A$, donde A es una matriz que se compone únicamente de bloques $A_i = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 0 \end{pmatrix}$ en la diagonal y ceros en el resto, llamados **bloques nilpotentes elementales**.*

Observación 48. Ahora, dada una matriz A de característico $(X-\lambda)^r$, en virtud del teorema de Cayley-Hamilton, se tiene que $A-\lambda Id$ es nilpotente y por tanto $\exists C$ tal que, como $A-\lambda Id=(A-\lambda Id)+\lambda Id$, entonces $C^{-1}AC = C^{-1}(A-\lambda Id)C + C^{-1}\lambda IdC = A' + \lambda Id$, donde A' está compuesta de bloques nilpotentes, y por tanto $C^{-1}AC$ está compuesta de bloques de Jordan.

Para demostrar el teorema, en primer lugar, daremos la observación:

Proposición 77. *Dado $\mathbb{K}^n$ y $S\subset T$ dos subespacios suyos, si $\{[v_1],...,[v_l]\}\subset T/S$ y $\{w_1,...,w_m\}\subset S$, entonces $\{v_1,...,v_l,w_1,...,w_m\}$ es base de T.*

Demostración. Al ser el número adecuado de vectores, basta ver que generan T. Sea $z\in T$. Como $[z]=\sum\lambda[v_i]$, entonces $[z-\sum\lambda_i v_i]=[0]$, dada la linealidad de las clases. Entonces, $z-\sum\lambda_i v_i\in S$, con que $z-\sum\lambda_i v_i=\sum\beta_k w_k$, luego $z=\sum\lambda_i v_i+\sum\beta_k w_k$. □

Ahora, para ver qué base es la que proporciona esa matriz compuesta de bloques nilpotentes, deberemos tratar las cadenas de subespacios:

Proposición 78. *Sea $g\in End(E)$. Entonces, $Nuc(g)\subset Nuc(g^2)\subset ...\subset Nuc(g^r)\subset ...$*

Demostración. Sea $i\in\mathbb{N}$ y $v\in Nuc(g^i)$. Entonces, $g^{i+1}(v)=g(g^i(v))=g(0)=0$. □

Observación 49. $v\in Nuc(g^i)\iff g(v)\in Nuc(g^{i-1})$.

Razón. $g^i(v)=0\iff g^{i-1}g(v)=0$.

Proposición 79. *Por la observación anterior, $g: Nuc(g^i)\mapsto Nuc(g^{i-1})$ está bien definida. Consideremos el paso al cociente $\pi: Nuc(g^{i-1})\mapsto Nuc(g^{i-1})/Nuc(g^{i-2})$. Sea $\alpha_i=\pi\circ g$, $\alpha_i: Nuc(g^i)\mapsto Nuc(g^{i-1})/Nuc(g^{i-2})$.*

Entonces, $Nuc(\alpha_i)=Nuc(g^{i-1})$, y por el primer teorema de isomorfía se induce un monomorfismo entre $\frac{Nuc(g)}{Nuc(g^{i-1})}$ y $\frac{Nuc(g^{i-1})}{Nuc(g^{i-2})}$.

Demostración. Puesto que $Nuc(\pi)=Nuc(g^{i-2})$, entonces $v\in Nuc(\alpha)\iff g(v)\in Nuc(g^{i-2})\iff v\in Nuc(g^{i-1})$, en virtud de la observación anterior. □

Visto esto, podemos continuar profundizando en las matrices nilpotentes:

Observación 50. Sea $B\in\mathcal{M}_n(\mathbb{K})$ una matriz nilpotente. Entonces, consideramos los espacios $\{0\}\subset N(B)\subset N(B^2)\subset ...\subset N(B^n)=\mathbb{K}^n$.

Observación 51. Dados subespacios $S_1\subset ...\subset S_n$, entonces: $dim(S_n)=dim(S_n/S_{n-1})+dim(S_{n-1}/S_{n-2})+...+dim(S_2/S_1)+dim(S_1)$.

Definición 65. Asociamos a la cadena $N(B)\subset N(B^2)\subset ...\subset N(B^n)=\mathbb{K}^n$, con B nilpotente, un **diagrama** consistente en filas de cajas, donde la inferior contiene $dim(N(B))$ cajas, la siguiente $dim(N(B^2)/N(B_1))$, y así hasta $dim(N(B^n)/N(B^{n-1}))$. El número de cajas desde abajo hasta la i-ésima fila es, en virtud de la observación anterior, $dim(N(B^i))$. Se usará en las discusiones posteriores.

Observación 52. A causa de la proposición 74, se tiene $dim(N(B^{i+1})/N(B^i))\leq dim(N(B^i)/N(B^{i-1}))$, por tanto, el número de cajas en cada fila del diagrama decrece según se asciende.

Proposición 80. *Si $N(B^i)=N(B^{i+1})$, entonces $N(B^i)=N(B^k)$ $\forall k\geq i$, es decir, la cadena se estabiliza, cesando el número de cajas en el diagrama a partir de ese punto. Como sabemos, además, en cualquier caso cesarán a partir de la n-ésima fila.*

Demostración. Se tendrá entonces que $dim(N(B^{i+1})/N(B^i))=0$, luego, por la proposición 74, $dim(N(B^{i+1+k})/N(B^{i+k}))=0$, es decir, $N(B^{i+1+k})=N(B^{i+k})$ $\forall k\geq 0$. □

8.4. Obtención de la base de Jordan

Ahora veremos un **ejemplo de obtención de la base** en la que la matriz nilpotente B se descompone en bloques nilpotentes elementales. Esto, como sabemos, lleva a la base que convierte el bloque de característico $(X-\lambda)^r$ en matriz de Jordan.

Supongamos que B tiene el siguiente diagrama:

$$\begin{array}{||c|c|c||c} X & & & (N(B^3)/N(B^2)) \\ \hline X & X & & (N(B^2)/N(B)) \\ \hline X & X & X & (N(B)) \end{array}$$

Con la correspondiente cadena $N(B) \subset N(B^2) \subset N(B^3) = \mathbb{K}^6$. Al ser nilpotente aseguramos que la cadena cesa. Vamos a rellenarlo de tal forma que cada fila contenga los representantes de las clases de una base de cada cociente. Para ello, tomamos $v \in N(B^3)$, $v \notin N(B^2)$, cuya clase, $[v]$ formará base de $N(B^3)/N(B^2)$ al ser de dimensión 1. Lo agregamos al diagrama:

$$\begin{array}{||c|c|c||c} v & & & (N(B^3)/N(B^2)) \\ \hline X & X & & (N(B^2)/N(B)) \\ \hline X & X & X & (N(B)) \end{array}$$

Ahora, observemos que da igual qué familia F agregásemos a la fila superior, $[B(F)]$ es linealmente independiente en $N(B^2)/N(B)$. Esto sigue de la proposición 74: la aplicación $\tilde{\alpha}_B$ es monomorfismo, y $\tilde{\alpha}_B([v]) = \alpha_B(v) = [B(v)]$, es decir, la imagen de todo vector a través de este monomorfismo es la clase de la imagen por B, y serán linealmente independientes. Por tanto, podemos anotarlos en el diagrama:

$$\begin{array}{||c|c|c||c} v & & & (N(B^3)/N(B^2)) \\ \hline Bv & X & & (N(B^2)/N(B)) \\ \hline X & X & X & (N(B)) \end{array}$$

Ahora, extendemos los vectores linealmente dependientes que hemos obtenido a base: $\{[B_v],[w]\}$ base de $N(B^2)/N(B)$. Esto quiere decir que $w \in N(B^2)$, $w \notin N(B)$ y además son linealmente independientes. Agregamos los nuevos representantes al diagrama:

$$\begin{array}{||c|c|c||c} v & & & (N(B^3)/N(B^2)) \\ \hline Bv & w & & (N(B^2)/N(B)) \\ \hline X & X & X & (N(B)) \end{array}$$

Siguiendo el razonamiento anterior, podemos bajar esta nueva fila:

$$\begin{array}{||c|c|c||c} v & & & (N(B^3)/N(B^2)) \\ \hline Bv & w & & (N(B^2)/N(B)) \\ \hline B^2v & Bw & X & (N(B)) \end{array}$$

Y extendemos a base con un nuevo vector (z):

$$\begin{array}{||c|c|c||c} v & & & (N(B^3)/N(B^2)) \\ \hline Bv & w & & (N(B^2)/N(B)) \\ \hline B^2v & Bw & z & (N(B)) \end{array}$$

Ahora que tenemos los representantes de una base de cada cociente, hacemos uso de la observación 72: todos los vectores del diagrama son base de $N(B^3) = \mathbb{K}^6$, dado que los de las dos últimas filas lo son de $N(B^2)$ (observación 72), y esos con la superior de $N(B^3)$ (una vez más).

Ahora, ordenamos la base por columnas convenientemente: $F = \{B^2v, Bv, v, Bw, w, z\}$. No es difícil comprobar que, en efecto:

$$|B|_F = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

En bloques nilpotentes. Se observa que cada columna del diagrama es un bloque, de orden las filas correspondientes. □

Puesto que esta última sección ha sido de una densidad teórica importante, para justificar todos los pasos que garantizan la existencia de la forma canónica de Jordan, acabamos el documento con un breve resumen:

Observación 53 (Hallar forma de Jordan). En una matriz cuyo característico factoriza completamente en lineales:

1. Partir la matriz en bloques diagonales, cada uno correspondiente al invariante $Nuc(A - \lambda_i Id)^{r_i}$, concatenando bases de todos ellos (están en suma directa dando el total). Cada bloque tiene característico $(X - \lambda_i I)^{r_i}$.

2. Por cada bloque, $A_i - \lambda_i Id$ es entonces nilpotente de orden r_i. Hallar su forma canónica nilpotente y sumar $\lambda_i Id$ para obtener su forma de Jordan. Si se quiere obtener su base, es posible con el procedimiento anterior.

3. Dados todos los bloques en forma de Jordan, ya se tiene la matriz final en forma de Jordan. La matriz de cambio de base, si se realizó el procedimiento anterior, se obtiene con facilidad yuxtaponiendo cada una de las submatrices de paso.

Observemos también que el **número de cajas** del diagrama es el grado del característico y la **altura del diagrama** es el grado del mínimo.