

MARIANO SUÁREZ-ÁLVAREZ

ÁLGEBRA LINEAL

7 DE AGOSTO DE 2020



This work is licensed under a Creative Commons
“Attribution-NonCommercial-ShareAlike 4.0 International” license.



Índice

1 Espacios vectoriales	1
§1. Cuerpos	1
§2. Espacios vectoriales	6
§3. Subespacios	10
§4. Combinaciones lineales	13
§5. Dependencia e independencia lineal	18
§6. Bases	23
§7. Dimensión	27
§8. Sumas de subespacios	34
§9. Sumas directas de subespacios	37
§10. Complementos y codimensión	42
§11. Coordenadas y cambio de base	46
§12. Una digresión: el principio de inclusión–exclusión para subespacios	55
2 Funciones lineales	75
§1. Funciones lineales	75
§2. Imagen, preimagen y núcleo	81
§3. Monomorfismos, epimorfismos e isomorfismos	82
§4. El teorema de la dimensión	86
§5. El espacio de homomorfismos entre dos espacios vectoriales	90
§6. La matriz asociada a una función lineal	95
§7. Proyectores	106
§8. Cocientes	112
§9. Los teoremas de isomorfismo	119

§10. Exactitud	124
§11. Subespacios invariantes	134
3 Dualidad	137
§1. El espacio dual	137
§2. Anuladores	143
§3. Funciones transpuestas	153
§4. El rango de una matriz	158
4 Determinantes	163
§1. Funciones multilíneales alternantes	163
§2. Funciones multilíneales alternantes de grado máximo	169
§3. El determinante de una matriz	177
§4. Permutaciones y sus signos	181
§5. La fórmula de Leibniz	188
§6. La fórmula de Laplace y la regla de Cramer	190
§7. El rango de una matriz	195
§8. Tres determinantes	196
§9. Espacios de funciones multilíneales alternantes	202
5 Autovectores y autovalores	215
§1. Autovectores y autovalores	215
§2. Ejemplos	219
§3. Diagonalizabilidad	224
§4. El polinomio característico	228
§5. Homomorfismos de álgebras e ideales	243
§6. El polinomio minimal	247
§7. Descomposición primaria y diagonalizabilidad	261
§8. Endomorfismos triangularizables y semisimples	268
6 Formas normales	279
§1. Equivalencia de funciones lineales y de matrices	280
§2. Equivalencia a derecha de funciones lineales y de matrices	288
§3. Conjugación de endomorfismos y de matrices cuadradas	293
§4. Endomorfismos descomponibles e indescomponibles	294
§5. Endomorfismos nilpotentes	297
§6. La forma normal de Jordan	318

§7. Algunas ejemplos y aplicaciones	323
7 Espacios con producto interno	331
§1. Espacios con producto interno	331
§2. Normas y métricas	335
§3. Ortogonalidad	340
§4. Complementos ortogonales	347
§5. Proyectores ortogonales	350
§6. El teorema de representación de Riesz	353
§7. Funciones adjuntas	355
§8. Funciones lineales autoadjuntas	360
§9. Funciones lineales normales	362
§10. Funciones unitarias y ortogonales	366
§11. Matrices ortonormal y unitarias	367
§12. Una familia importante de ejemplos	369
Referencias	377

Capítulo 1

Espacios vectoriales

§1. Cuerpos

1.1.1. Un *cuerpo* es un conjunto \mathbb{k} dotado de dos operaciones

$$+ : \mathbb{k} \times \mathbb{k} \rightarrow \mathbb{k}, \quad \cdot : \mathbb{k} \times \mathbb{k} \rightarrow \mathbb{k},$$

a las que llamamos la *suma* y el *producto* de \mathbb{k} , respectivamente, que satisfacen las siguientes condiciones:

- (K₁) la suma es asociativa: $(a + b) + c = a + (b + c)$ para cada $a, b, c \in \mathbb{k}$;
- (K₂) la suma es commutativa: $a + b = b + a$ para cada $a, b \in \mathbb{k}$;
- (K₃) hay un elemento neutro para la suma: existe un elemento $0 \in \mathbb{k}$ tal que $a + 0 = a$ y $0 + a = a$ para todo $a \in \mathbb{k}$;
- (K₄) todo elemento de \mathbb{k} posee un opuesto aditivo: para todo $a \in \mathbb{k}$ existe un elemento $a' \in \mathbb{k}$ tal que $a + a' = 0$ y $a' + a = 0$;
- (K₅) el producto es asociativo: $(a \cdot b) \cdot c = a \cdot (b \cdot c)$ para cada $a, b, c \in \mathbb{k}$;
- (K₆) el producto es commutativo: $a \cdot b = b \cdot a$ para cada $a, b \in \mathbb{k}$;
- (K₇) hay un elemento neutro para el producto: existe un elemento $1 \in \mathbb{k}$ tal que $a \cdot 1 = a$ y $1 \cdot a = a$ para todo $a \in \mathbb{k}$;
- (K₈) todo elemento de \mathbb{k} distinto de 0 posee un inverso multiplicativo: para cada $a \in \mathbb{k}$ distinto de 0, existe $a' \in \mathbb{k}$ tal que $a \cdot a' = 1$ y $a' \cdot a = 1$;
- (K₉) el producto se distribuye sobre la suma: $a \cdot (b + c) = a \cdot b + a \cdot c$ y $(b + c) \cdot a = b \cdot a + c \cdot a$ para cada $a, b, c \in \mathbb{k}$;
- (K₁₀) los elementos neutros de la suma y del producto son distintos: $0 \neq 1$.

La condición (K₃) afirma que existe un elemento neutro para la suma y, de hecho, implica que

existe *exactamente* uno: en efecto, si 0 y $0'$ son dos elementos de \mathbb{k} que son neutros para la suma, entonces

$$0 = 0 + 0' = 0'.$$

La primera de estas igualdades es consecuencia de que $0'$ es un elemento neutro para la suma, y la segunda de que 0 lo es. De manera similar, si a es un elemento de \mathbb{k} , la condición (K_4) nos dice que existe un elemento opuesto a a , esto es, un elemento $a' \in \mathbb{k}$ tal que $a + a' = 0$ y $a' + a = 0$, y este elemento está, de hecho, únicamente determinado por a : si a'' es otro elemento opuesto a a , de manera que $a + a'' = 0$ y $a'' + a = 0$, entonces tenemos que

$$\begin{aligned} a' &= a' + 0 && \text{por } (K_3) \\ &= a' + (a + a'') && \text{porque } a'' \text{ es un opuesto a } a \\ &= (a' + a) + a'' && \text{por } (K_1) \\ &= 0 + a'' && \text{porque } a' \text{ es un opuesto a } a \\ &= a''. && \text{por } (K_3), \text{ otra vez.} \end{aligned}$$

Como no hay entonces ambigüedad posible, escribiremos desde ahora en adelante $-a$ al elemento opuesto de a . Observemos que lo que acabamos de hacer es probar que

si $a \in \mathbb{k}$, entonces $-a$ es único elemento que sumado con a da 0

Esto tiene una consecuencia evidente: siempre que tengamos un elemento x de \mathbb{k} y queramos mostrar que $x = -a$ bastará que probemos que $x + a = 0$. Un ejemplo muy sencillo de esto aparece en la prueba de la siguiente observación:

1.1.2. Lema. *Sea \mathbb{k} un cuerpo. Si $a \in \mathbb{k}$, entonces $-(-a) = a$.*

Demostración. Queremos mostrar que a es igual a $-(-a)$, es decir, al opuesto de $-a$ y, de acuerdo a lo que dijimos recién, para hacer esto es suficiente con mostrar que la suma de a y $-a$ es 0: pero esto es claro, precisamente porque $-a$ es el opuesto de a . \square

Procediendo de la misma forma pero ahora con el producto de nuestro cuerpo, es fácil ver que el elemento 1 que es neutro para el producto cuya existencia afirma la condición (K_7) es uno sólo y que el inverso de todo elemento distinto de 0 que nos da la condición (K_8) está bien determinado. Gracias a esto último, podemos escribir sin ninguna ambigüedad a^{-1} al inverso de un elemento $a \in \mathbb{k} \setminus \{0\}$.

1.1.3. Ejemplos.

- (a) Los conjuntos \mathbb{Q} , \mathbb{R} y \mathbb{C} de los números racionales, reales y complejos, respectivamente, dotados de sus operaciones usuales de suma y producto, son cuerpos.
- (b) Sea $\mathbb{Q}(\sqrt{2})$ el conjunto de todos los números reales de la forma $a + b\sqrt{2}$ con $a, b \in \mathbb{Q}$. Si x e y son dos elementos de $\mathbb{Q}(\sqrt{2})$, entonces tanto la suma $x + y$ como el producto xy también

están en $\mathbb{Q}(\sqrt{2})$: en efecto, si $x = a + b\sqrt{2}$ e $y = c + d\sqrt{2}$, con $a, b, c, d \in \mathbb{Q}$, entonces

$$x + y = (a + c) + (b + d)\sqrt{2}, \quad xy = (ac + 2bd) + (ad + bc)\sqrt{2}$$

y es claro que los números $a + c$, $b + d$, $ac + 2bd$ y $ad + bc$ pertenecen a \mathbb{Q} . Esto nos dice que podemos restringir las operaciones de \mathbb{R} a $\mathbb{Q}(\sqrt{2})$, obteniendo de esta forma operaciones

$$+ : \mathbb{Q}(\sqrt{2}) \times \mathbb{Q}(\sqrt{2}) \rightarrow \mathbb{Q}(\sqrt{2}), \quad \cdot : \mathbb{Q}(\sqrt{2}) \times \mathbb{Q}(\sqrt{2}) \rightarrow \mathbb{Q}(\sqrt{2}).$$

Estas operaciones hacen de $\mathbb{Q}(\sqrt{2})$ un cuerpo. La verificación de todas las condiciones de la definición de 1.1.1 es inmediata salvo la de (K₈), que es consecuencia de la siguiente observación: si x es un elemento no nulo de $\mathbb{Q}(\sqrt{2})$, entonces existen $a, b \in \mathbb{Q}$ no simultáneamente nulos tales que $x = a + b\sqrt{2}$ y es

$$x^{-1} = \frac{a}{a^2 + 2b^2} - \frac{b}{a^2 + 2b^2}\sqrt{2} \in \mathbb{Q}(\sqrt{2})$$

Procediendo de manera similar, para cada número racional $r \in \mathbb{Q}$ podemos construir un cuerpo $\mathbb{Q}(\sqrt{r})$ cuyos elementos son todos los números reales de la forma $a + b\sqrt{r}$ con $a, b \in \mathbb{Q}$.

- (c) El conjunto $\mathbb{F}_2 = \{0, 1\}$ es un cuerpo si lo dotamos de las operaciones de suma y producto dadas por

$+$	0	1
0	0	1
1	1	0

\cdot	0	1
0	0	0
1	0	1

- (d) El conjunto $\mathbb{F}_3 = \{0, 1, 2\}$ es un cuerpo si lo dotamos de las operaciones de suma y producto dadas por

$+$	0	1	2
0	0	1	2
1	1	2	0
2	2	0	1

\cdot	0	1	2
0	0	0	0
1	0	1	2
2	0	2	1

- (e) Si p es un número primo, entonces tenemos la relación de equivalencia \equiv sobre el conjunto de los enteros \mathbb{Z} dada por la congruencia módulo p y, por lo tanto, podemos considerar el conjunto cociente $\mathbb{Z}_p := \mathbb{Z}/\equiv$, cuyos elementos son las clases de equivalencia de la relación \equiv en \mathbb{Z} . Si a es un entero, escribamos $[a]$ a la clase de equivalencia de a en \mathbb{Z} con respecto a \equiv .

Es fácil verificar que hay operaciones $+$, $\cdot : \mathbb{Z}_p \times \mathbb{Z}_p \rightarrow \mathbb{Z}_p$ en el conjunto \mathbb{Z}_p tales que

$$[a] + [b] = [a + b], \quad [a] \cdot [b] = [ab]$$

para cualquier elección de a y b en \mathbb{Z} . Más aún, el conjunto \mathbb{Z}_p dotado de estas operaciones $+$ y \cdot de suma y producto es un cuerpo que claramente tiene exactamente p elementos. Una notación común para este cuerpo es \mathbb{F}_p . Los cuerpos de los ejemplos 1.1.3(c) y 1.1.3(d) que

dimos arriba son esencialmente los que se obtienen tomando $p = 2$ y $p = 3$, respectivamente, en esta construcción.

Más generalmente, se puede mostrar que si p es un número primo y $n \in \mathbb{N}$ un entero positivo, entonces existe un cuerpo \mathbb{F}_{p^n} con p^n elementos, que éste es esencialmente único — llamamos a estos cuerpos **cuerpos de Galois**, por Evariste Galois (1811–1832) — y que los cuerpos que se obtienen de esta forma son los únicos cuerpos con finitos elementos.

- (f) Si X es una variable y \mathbb{k} es un cuerpo, el conjunto $\mathbb{k}[X]$ de los polinomios con coeficientes en \mathbb{k} en la variable X no es un cuerpo cuando lo dotamos de sus operaciones usuales de suma y producto: por ejemplo, el polinomio X no es el polinomio nulo y sin embargo no posee en $\mathbb{k}[X]$ un inverso multiplicativo. En cambio, en conjunto $\mathbb{k}(X)$ de las **funciones racionales**, esto es, el de las fracciones $p(X)/Q(X)$ con p y q elementos de $\mathbb{k}[X]$ y q no nulo, es un cuerpo con respecto a sus operaciones usuales.
- (g) Los conjuntos \mathbb{N} , \mathbb{N}_0 y \mathbb{Z} de los enteros positivos, de los enteros no negativos y de los enteros, respectivamente, no son cuerpos cuando los dotamos de sus operaciones usuales: en \mathbb{N} no hay un elemento neutro para la suma, en \mathbb{N}_0 no todo elemento posee un opuesto, en \mathbb{Z} no todo elemento no nulo posee un inverso. \diamond

La característica de un cuerpo

1.1.4. Si \mathbb{k} es un cuerpo, podemos definir una sucesión $(u_n)_{n \geq 0}$ de elementos de \mathbb{k} de la siguiente manera: ponemos $u_0 := 0$ y, para cada $n \in \mathbb{N}$, $u_n := u_{n-1} + 1$.

Lema. Sea \mathbb{k} un cuerpo y sea $(u_n)_{n \geq 0}$ la sucesión de elementos de \mathbb{k} que acabamos de definir.

- (i) Para cada elección de n y m en \mathbb{N}_0 se tiene que $u_n + u_m = u_{n+m}$.
- (ii) Para cada elección de n y m en \mathbb{N}_0 se tiene que $u_n \cdot u_m = u_{nm}$.

Demostración. Para probar (i) fijamos $n \in \mathbb{N}_0$ y procedemos haciendo inducción con respecto a m .

- Si $m = 0$, entonces $u_n + u_m = u_n + u_0 = u_n + 0 = u_n = u_{n+0}$ simplemente por definición.
- Por otro lado, si $m \geq 1$ y $u_n + u_{m-1} = u_{n+m-1}$, entonces

$$u_n + u_m = u_n + u_{(m-1)+1} = u_n + u_{m-1} + 1 = u_{n+m-1} + 1 = u_{n+m}.$$

Para probar (ii) hacemos lo mismo: fijamos $n \in \mathbb{N}_0$ y hacemos inducción con respecto a m .

- Si $m = 0$, entonces $u_n \cdot u_m = u_n \cdot u_0 = u_n \cdot 0 = 0 = u_0 = u_{n0}$.
- Por otro lado, si $m \geq 1$ y $u_n \cdot u_{m-1} = u_{n(m-1)}$, entonces

$$\begin{aligned} u_n \cdot u_m &= u_n \cdot (u_{m-1} + u_1) = u_n \cdot u_{m-1} + u_n \cdot u_1 = u_{n(m-1)} + u_n = u_{n(m-1)+n} \\ &= u_{nm} \end{aligned}$$

gracias a lo que ya probamos. \square

1.1.5. Si \mathbb{k} es un cuerpo y $(u_n)_{n \geq 0}$ es la sucesión que definimos arriba, hay dos posibilidades. O bien $u_n \neq 0$ para todo $n \in \mathbb{N}$, y entonces decimos que el cuerpo \mathbb{k} tiene **características 0**, o bien no es ese el caso, y entonces llamamos **características** de \mathbb{k} al numero

$$\chi(\mathbb{k}) := \min\{n \in \mathbb{N} : u_n = 0\},$$

que es positivo y, como $1 \neq 0$ en \mathbb{k} , mayor o igual que 2.

1.1.6. La observación más importante que podemos hacer sobre la característica de un cuerpo es la siguiente:

Proposición. *La característica de un cuerpo es 0 cero o un número primo.*

Demostración. Usando esto, probemos ahora la proposición. Supongamos que el cuerpo \mathbb{k} tiene característica positiva, sea $p = \chi(\mathbb{k})$ su característica, de manera que $p \geq 2$, y supongamos que p es un número compuesto. Hay entonces dos enteros n y m mayores que 1 tales que $p = nm$ y, por lo tanto, $u_n u_m = u_{nm} = u_p = 0$. Como \mathbb{k} es un cuerpo, esto implica que alguno de u_n o u_m es nulo: esto es absurdo, ya que claramente $n < p$ y $m < p$ y elegimos a p como el menor entero positivo tal que $u_p = 0$. \square

1.1.7. Ejemplos.

- (a) Si \mathbb{k} es $\mathbb{Q}, \mathbb{R}, \mathbb{C}$ o, de hecho, cualquier subcuerpo de \mathbb{C} , entonces \mathbb{k} tiene característica 0. En efecto, es inmediato verificar que para todo $n \in \mathbb{N}$ se tiene que $u_n = n \neq 0$.
- (b) Los cuerpos \mathbb{F}_2 y \mathbb{F}_3 de los Ejemplos 1.1.3(c) y 1.1.3(d) tienen característica 2 y 3, respectivamente. Más generalmente, si p es un número primo, el cuerpo \mathbb{F}_p y, para cada $r \in \mathbb{N}$, el cuerpo \mathbb{F}_{p^r} del Ejemplo 1.1.3(e) tienen característica p . \diamond

1.1.8. Proposición. *Un cuerpo finito tiene característica positiva.*

Demostración. Sea \mathbb{k} un cuerpo finito y supongamos, para llegar a una contradicción, que la característica de \mathbb{k} es nula. Si $(u_n)_{n \geq 0}$ es la sucesión construida arriba, tenemos entonces que $u_n \neq 0$ para todo $n \in \mathbb{N}$. Ahora bien, como K es finito la función $n \in \mathbb{N}_0 \mapsto u_n \in K$ no puede ser inyectiva y existe entonces dos enteros i y j tales que $0 \leq i < j$ y $u_i = u_j$. Tenemos entonces que $u_{j-i} + u_i = u_j$, así que $u_{j-i} = u_j - u_i = 0$, y esto es absurdo, ya que $j - i \in \mathbb{N}$. \square

1.1.9. La característica de un cuerpo es la misma que la de cualquier subcuerpo o supercuerpo:

Proposición. *Sea \mathbb{K} un cuerpo y \mathbb{k} un subcuerpo de \mathbb{K} . Las características de \mathbb{K} y de \mathbb{k} coinciden.*

Demostración. Sean $(u_n)_{n \geq 0}$ y $(U_n)_{n \geq 0}$ las sucesión definidas como en 1.1.4 en \mathbb{k} y en \mathbb{K} , respectivamente. Como \mathbb{k} es un subcuerpo de \mathbb{K} , tenemos que $1_{\mathbb{k}} = 1_{\mathbb{K}}$ y una inducción evidente muestra que $u_n = U_n$ para todo $n \in \mathbb{N}$. La proposición sigue inmediatamente de esto. \square

1.1.10. Una última observación que podemos hacer es la siguiente:

Proposición. Si \mathbb{k} es un cuerpo de característica positiva p , entonces $P = \{u_n : n \in \mathbb{N}_0\}$ es un subcuerpo de \mathbb{k} de cardinal p .

Llamamos a P el *subcuerpo primo* de \mathbb{k} .

Demostración. Sea \mathbb{k} un cuerpo de característica positiva p y sea P el conjunto descripto en el enunciado de la proposición. Para ver que P es un subcuerpo de \mathbb{k} tenemos que mostrar que contiene al 0 y al 1 de \mathbb{k} , que es cerrado por sumas y productos, y que contiene al inverso de todos sus elementos no nulos. Es claro que $1 = u_1$ y que $0 = u_0$ están en P , y que P es cerrado por sumas y productos es consecuencia inmediata del Lema 1.1.4.

Si $n \in \mathbb{N}$, el algoritmo de la división nos da enteros no negativos q y r tales que $n = qp + r$ y $0 \leq r < p$, y es

$$u_n = u_{qp+r} = u_q \cdot u_p + u_r$$

porque $u_p = 0$. Vemos así que

$$P = \{u_0, \dots, u_{p-1}\}.$$

Más aún, los p elementos listados aquí son distintos dos a dos, así que P tiene exactamente p elementos. En efecto, si no fuese ese el caso, habría enteros i y j tales que $0 \leq i < j < p$ y $u_i = u_j$, y entonces tendríamos que $u_{j-i} + u_i = u_j$, así que $u_{j-i} = u_j - u_i = 0$: esto es absurdo, porque $j - i < p$ y p es el menor entero no positivo tal que $u_p = 0$.

Para terminar, mostremos que P contiene el inverso de cualquiera de sus elementos no nulos. Sea x un elemento no nulo de P . Como vimos, existe $i \in \llbracket 0, p-1 \rrbracket$ tal que $x = u_i$ y, como $x \neq 0 = u_0$, tenemos que de hecho $0 < i < p$. En particular, los números i y p son coprimos, ya que p es primo, y existen enteros x e y tales que $xi + yp = 1$ e $y < 0$. Esto implica que

$$u_x \cdot u_i = u_{xi} = u_{1+(-y)p} = u_1 + u_{-y} \cdot u_p = 1,$$

ya que $u_p = 0$ y, por lo tanto, que $u_i^{-1} = u_x \in P$. La proposición queda así probada. \square

§2. Espacios vectoriales

1.2.1. Fijemos un cuerpo \mathbb{k} . Un *espacio vectorial sobre \mathbb{k}* es un conjunto V dotado de dos operaciones

$$+ : V \times V \rightarrow V, \quad \cdot : \mathbb{k} \times V \rightarrow V$$

que llamamos la *suma* de V y la *multiplicación escalar*, que satisfacen las siguientes condiciones:

- (V₁) la suma es asociativa: $(x + y) + z = x + (y + z)$ para cada $x, y, z \in V$;
- (V₂) la suma es commutativa: $x + y = y + x$ para cada $x, y \in V$;
- (V₃) hay un elemento neutro para la suma: existe un elemento $0_V \in V$ tal que $x + 0_V = x$ y $0_V + x = x$ para todo $x \in V$;
- (V₄) todo elemento de V posee un opuesto: para todo $x \in V$ existe un elemento $x' \in V$ tal que $x + x' = 0_V$ y $x' + x = 0_V$;
- (V₅) la multiplicación escalar es asociativa: $a \cdot (b \cdot x) = (a \cdot b) \cdot x$ para cada $a, b \in \mathbb{k}$ y cada $x \in V$;
- (V₆) la multiplicación escalar es unitaria: $1 \cdot x = x$ para todo $x \in V$;
- (V₇) la multiplicación escalar se distribuye sobre la suma de \mathbb{k} y sobre la suma de V :

$$\begin{aligned} a \cdot (x + y) &= a \cdot x + a \cdot y \text{ para todo } a \in \mathbb{k} \text{ y todo } x, y \in V; \\ (a + b) \cdot x &= a \cdot x + b \cdot x \text{ para todo } a, b \in \mathbb{k} \text{ y todo } x \in V. \end{aligned}$$

Estas condiciones implican que hay en V un único elemento neutro para la suma: si 0_V y $0'_V$ son dos elementos neutros, se tiene que $0_V = 0_V + 0'_V = 0'_V$. Decimos que es el **cero** de V . De forma similar, si x es un elemento de V , entonces hay un único elemento opuesto a x : si x' y x'' son dos elementos opuestos a x , entonces

$$x' = x' + 0_V = x' + (x + x'') = (x' + x) + x'' = 0_V + x'' = x'',$$

y esto nos permite escribir sin ambigüedad $-x$ para denotar al elemento opuesto de x .

1.2.2. Si \mathbb{k} es un cuerpo y V es un espacio vectorial sobre \mathbb{k} , llamamos a los elementos de \mathbb{k} **escalares** y a los elementos de V **vectores**. Si $\mathbb{k} = \mathbb{R}$, el cuerpo de los números reales, decimos generalmente que V es un **espacio vectorial real** en lugar de que es un espacio vectorial sobre \mathbb{R} y si $\mathbb{k} = \mathbb{C}$, el cuerpo de los números complejos, que es un **espacio vectorial complejo**.

1.2.3. Ejemplos.

- (a) Sea $n \in \mathbb{N}$, sea \mathbb{k} un cuerpo y sea \mathbb{k}^n el conjunto de las n -uplas de elementos de \mathbb{k} , que consideraremos *siempre* como columnas. Si dotamos a \mathbb{k}^n de las operaciones $+ : \mathbb{k}^n \times \mathbb{k}^n \rightarrow \mathbb{k}^n$ y $\cdot : \mathbb{k} \times \mathbb{k}^n \rightarrow \mathbb{k}^n$ definidas de manera que

$$\begin{aligned} (x_1, \dots, x_n)^t + (y_1, \dots, y_n)^t &= (x_1 + y_1, \dots, x_n + y_n)^t, \\ a \cdot (x_1, \dots, x_n)^t &= (a \cdot x_1, \dots, a \cdot x_n)^t \end{aligned}$$

para cada par de elementos $(x_1, \dots, x_n)^t$ e $(y_1, \dots, y_n)^t$ de \mathbb{k}^n y cada $a \in \mathbb{k}$, entonces \mathbb{k}^n es un espacio vectorial sobre \mathbb{k} .

- (b) Más generalmente, sean $m, n \in \mathbb{N}$, sea \mathbb{k} un cuerpo y sea $M_{n,m}(\mathbb{k})$ el conjunto de las matrices de n filas y m columnas con entradas en \mathbb{k} . El conjunto $M_{n,m}(\mathbb{k})$ es un espacio vectorial sobre \mathbb{k} con las operaciones de suma $+ : M_{n,m}(\mathbb{k}) \times M_{n,m}(\mathbb{k}) \rightarrow M_{n,m}(\mathbb{k})$ y de multiplicación por escalares $\cdot : \mathbb{k} \times M_{n,m}(\mathbb{k}) \rightarrow M_{n,m}(\mathbb{k})$ usuales:

- si $a = (a_{i,j})$ y $b = (b_{i,j})$ son dos elementos de $M_{n,m}(\mathbb{k})$, entonces la suma $a + b$ es la matriz $(c_{i,j})$ con $c_{i,j} = a_{i,j} + b_{i,j}$ para cada $i \in \llbracket n \rrbracket$ y $j \in \llbracket m \rrbracket$; y
- si $\lambda \in \mathbb{k}$ es un escalar y $a = (a_{i,j}) \in M_{n,m}(\mathbb{k})$, entonces $\lambda \cdot a$ es la matriz $c = (c_{i,j})$ con $c_{i,j} = \lambda a_{i,j}$ para cada $i \in \llbracket n \rrbracket$ y $j \in \llbracket m \rrbracket$.

Notemos que si $m = 1$, el espacio $M_{n,1}(\mathbb{k})$ coincide con el espacio \mathbb{k}^n que describimos en el ejemplo anterior. \diamond

1.2.4. Ejemplos.

- (a) Sea X un conjunto no vacío y sea \mathbb{k} un cuerpo. El conjunto \mathbb{k}^X de todas las funciones $X \rightarrow \mathbb{k}$ es un espacio vectorial sobre \mathbb{k} con respecto a las operaciones

$$+ : \mathbb{k}^X \times \mathbb{k}^X \rightarrow \mathbb{k}^X, \quad \cdot : \mathbb{k} \times \mathbb{k}^X \rightarrow \mathbb{k}^X$$

tales que

$$(f + g)(x) = f(x) + g(x), \quad (a \cdot f)(x) = a \cdot f(x)$$

para cada $f, g \in \mathbb{k}^X$, cada $a \in \mathbb{k}$ y cada $x \in X$.

- (b) Más generalmente, si X es un conjunto no vacío y V es un espacio vectorial sobre un cuerpo \mathbb{k} , el conjunto $W = V^X$ de todas las funciones $X \rightarrow V$ es un espacio vectorial sobre \mathbb{k} con respecto a las operaciones

$$+ : V^X \times V^X \rightarrow V^X, \quad \cdot : \mathbb{k} \times V^X \rightarrow V^X$$

tales que

$$(f + g)(x) = f(x) + g(x), \quad (a \cdot f)(x) = a \cdot f(x)$$

para cada $f, g \in V^X$, cada $a \in \mathbb{k}$ y cada $x \in X$. \diamond

1.2.5. Ejemplos.

- (a) Sea \mathbb{k} un cuerpo y sea $V = \mathbb{k}[X]$ el conjunto de los polinomios en la variable X con coeficientes en \mathbb{k} . Si dotamos a V de las operaciones usuales de suma de polinomios y de multiplicación por escalares, entonces V es un espacio vectorial.
- (b) Sea V un conjunto con un único elemento, que denotamos $*$. Hay exactamente una función $+ : V \times V \rightarrow V$ y hay exactamente una función $\cdot : \mathbb{k} \times V \rightarrow V$, y es fácil verificar que V es un espacio vectorial con respecto a esas operaciones. El cero de V es $*$ y entonces podemos escribir $V = \{0_V\}$. Decimos que un espacio vectorial con un único es un **espacio vectorial nulo** y lo escribimos, cuando esto no dé lugar a confusiones, simplemente 0.
- (c) Si K es un cuerpo y $\mathbb{k} \subseteq K$ es un subcuerpo de K , entonces K es un espacio vectorial sobre \mathbb{k} con respecto a la operación de suma $+ : K \times K \rightarrow K$ de K y a la multiplicación $\cdot : \mathbb{k} \times K \rightarrow K$ por elementos de \mathbb{k} que se obtiene restringiendo la multiplicación $\cdot : K \times K \rightarrow K$ de K a $\mathbb{k} \times K$. De esta forma podemos ver al cuerpo \mathbb{C} como un espacio vectorial sobre \mathbb{R} o sobre \mathbb{Q} , por ejemplo, y a $\mathbb{Q}(\sqrt{2})$ como un espacio vectorial sobre \mathbb{Q} . \diamond

1.2.6. Proposición. Sea \mathbb{k} un cuerpo y sea V un espacio vectorial sobre \mathbb{k} .

- (i) Para todo $x \in V$ es $0 \cdot x = 0_V$.
- (ii) Para todo $a \in \mathbb{k}$ es $a \cdot 0_V = 0_V$.
- (iii) Para todo $x \in V$ es $(-1) \cdot x = -x$.

Demostración. Observemos primero que

$$\text{si } x \in V \text{ es tal que } x + x = x, \text{ entonces } x = 0_V. \quad (1)$$

En efecto, si x es un elemento de V que satisface esa condición, tenemos que

$$x = x + 0_V = x + (x + (-x)) = (x + x) + (-x) = x + (-x) = 0_V.$$

Si $x \in V$, entonces

$$0 \cdot x + 0 \cdot x = (0 + 0) \cdot x = 0 \cdot x$$

y nuestra observación (1) nos dice que $0 \cdot x = 0_V$: esto prueba la afirmación (i) de la proposición.
Si ahora $a \in \mathbb{k}$, entonces

$$a \cdot 0_V + a \cdot 0_V = a \cdot (0_V + 0_V) = a \cdot 0_V$$

y, usando otra vez la observación (1), concluimos que $a \cdot 0_V = 0_V$, como afirma la parte (ii) de la proposición. Para ver que vale la parte (iii), consideramos un elemento $x \in V$ y calculamos que

$$x + (-1) \cdot x = 1 \cdot x + (-1) \cdot x = (1 + (-1)) \cdot x = 0 \cdot x = 0_V$$

y, de manera similar, que $(-1) \cdot x + x = 0_V$. Estas dos igualdades nos dicen que $(-1) \cdot x$ es un elemento de V opuesto a x y, como sabemos que hay uno sólo, que $(-1) \cdot x = -x$. \square

CONVENCIÓN

A partir de este momento fijaremos un cuerpo \mathbb{k} arbitrario en todo lo que sigue y todos los espacios vectoriales que consideremos serán espacios vectoriales sobre \mathbb{k} . Por otro lado, escribiremos simplemente 0 para denotar al cero del cuerpo \mathbb{k} , al cero de todos los espacios vectoriales con los que estemos trabajando, y a un espacio vectorial nulo.

§3. Subespacios

1.3.1. Sea V un espacio vectorial. Un subconjunto U de V es un *subespacio* de V si

- (S₁) $0 \in U$;
- (S₂) si $x, y \in U$, entonces $x + y \in U$; y
- (S₃) si $a \in \mathbb{k}$ y $x \in U$, entonces $a \cdot x \in U$.

Cuando ése es el caso, escribimos $U \leq V$.

1.3.2. Un subespacio de un espacio vectorial es, de manera natural, él mismo un espacio vectorial: esto es lo que afirma el siguiente resultado.

Proposición. *Sea V un espacio vectorial. Si U es un subespacio de V , entonces U es un espacio vectorial con respecto a las operaciones*

$$+ : (x, y) \in U \times U \longmapsto x + y \in U \quad \cdot : (a, x) \in \mathbb{k} \times U \longmapsto a \cdot x \in U$$

obtenidas de las de V por restricción.

Notemos que es precisamente porque U satisface las condiciones (S₂) y (S₃) que las operaciones $+ : V \times V \rightarrow V$ y $\cdot : \mathbb{k} \times V \rightarrow V$ de V se restringen a operaciones $+ : U \times U \rightarrow U$ y $\cdot : \mathbb{k} \times U \rightarrow U$. Desde ahora en adelante, cada vez que consideremos un subespacio de un espacio vectorial, lo dotaremos implícitamente de la estructura de espacio vectorial que nos da esta proposición.

Demostración. Que las condiciones (V₁), (V₂), (V₅), (V₆) y (V₇) se cumplen en U es consecuencia inmediata de que se cumplen en V . La condición (S₁) implica que (V₃) se cumple en U , ya que 0 , que está en U por hipótesis, es un elemento neutro para la suma de U . Resta ver, entonces, que la condición (V₄) se satisface en U . Pero si $x \in U$, entonces sabemos que $x' = (-1) \cdot x \in U$ por (S₃) y, según la Proposición 1.2.6(iii), que se tiene que $x + x' = 0$ y $x' + x = 0$ en V : en vista de la forma en que están definidas las operaciones de U , esto nos dice que x' es el opuesto de x en U . \square

1.3.3. Si V es un espacio vectorial, entonces los subconjuntos $\{0\}$ y V de V son subespacios de V —la verificación de las condiciones de la definición 1.3.1 es inmediata en los dos casos. Estos son los *subespacios triviales* de V . El subespacio $\{0\}$ es un espacio vectorial nulo y casi siempre escribiremos simplemente 0 en lugar de $\{0\}$: lo llamamos el *subespacio nulo* de V . Por otro lado, llamamos al subespacio V de V el *subespacio propio* de V . Los subespacio de V distintos de $\{0\}$ y de V son sus *subespacios no triviales*.

1.3.4. Ejemplos.

(a) Sea $n \in \mathbb{N}$ y sea $V = \mathbb{k}^n$. Es fácil ver que los siguientes subconjuntos de V son subespacios

de V :

$$\begin{aligned} U_1 &= \{(x_1, \dots, x_n)^t \in V : x_1 = 0\}, \\ U_2 &= \{(x_1, \dots, x_n)^t \in V : x_1 + \dots + x_n = 0\}, \\ U_3 &= \{(x_1, \dots, x_n)^t \in V : x_1 = x_2, x_3 = x_4\}. \end{aligned}$$

- (b) Sea $X = (0, 1)$ el intervalo unidad abierto de \mathbb{R} y sea $V = \mathbb{R}^X$ el espacio vectorial real de todas las funciones $X \rightarrow \mathbb{R}$, como en el Ejemplo 1.2.4(a). Los siguientes subconjuntos de V son subespacios de V :

$$\begin{aligned} U_1 &= \{f \in V : f \text{ es continua}\}, \\ U_2 &= \{f \in V : f \text{ tiene derivada en todo punto de } X\}, \\ U_3 &= \{f \in V : f \text{ tiene derivada en todo punto de } X \text{ y } f' \text{ es continua en } X\}, \\ U_4 &= \{f \in V : f \text{ es continua y } f(\frac{1}{2}) = 0\}, \\ U_5 &= \{f \in V : f \text{ es continua y } \int_{1/3}^{2/3} f(x) dx = 0\}. \end{aligned}$$

◇

1.3.5. Proposición. Sea V un espacio vectorial.

- (i) Si U_1 y U_2 son subespacios de V , entonces la intersección $U_1 \cap U_2$ también lo es.
- (ii) Más generalmente, si $\{U_i : i \in I\}$ es una familia arbitraria de subespacios de V , entonces la intersección $\bigcap_{i \in I} U_i$ es un subespacio de V .

Demostración. Basta que probemos la segunda parte, ya que contiene a la primera como un caso particular. Sea entonces $\{U_i : i \in I\}$ una familia de subespacios de V y pongamos $U = \bigcap_{i \in I} U_i$. Para ver que U es un subespacio de V , verificamos una a una las condiciones de la definición 1.3.1:

- Si $i \in I$, entonces $0 \in U_i$, porque U_i es un subespacio de V , y, por lo tanto, $0 \in \bigcap_{i \in I} U_i = U$.
- Sean x e y dos elementos de U . Si $i \in I$, entonces x e y están en U_i , ya que $U \subseteq U_i$, y entonces $x + y \in U_i$, porque U_i satisface la condición (S₂). Esto nos dice que $x + y \in U$.
- Sean $a \in \mathbb{k}$ y $x \in U$. Si $i \in I$ entonces $x \in U_i$ y, como U_i es un subespacio de V , $a \cdot x \in U_i$. Vemos así que $a \cdot x \in \bigcap_{i \in I} U_i = U$.

Esto completa la prueba de la proposición. □

1.3.6. A diferencia de lo que ocurre con la intersección de subespacios de un espacio vectorial, la unión de subespacios no es en general un subespacio. De hecho, vale la siguiente observación:

Proposición. Sea V un espacio vectorial y sean U_1 y U_2 dos subespacios de V . La unión $U_1 \cup U_2$ es un subespacio de V si y solamente si coincide con U_1 o con U_2 .

Notemos que esto último ocurre exactamente cuando o bien $U_1 \subseteq U_2$ o bien $U_2 \subseteq U_1$.

Demostración. Si $U_1 \cup U_2$ coincide con U_1 o con U_2 , es claro que se trata de un subespacio. Veamos la implicación recíproca. Supongamos que $U_1 \cup U_2$ es un subespacio de V y que ni U_1 no está

contenido en U_2 ni U_1 está contenido en U_2 . Existen entonces vectores $x \in U_1 \setminus U_2$ e $y \in U_2 \setminus U_1$. Como ambos pertenecen a $U_1 \cup U_2$ y esta unión es un subespacio de V , se sigue de esto que $x + y \in U_1 \cup U_2$. Esto es absurdo: si $x + y \in U_1$, entonces $y = (x + y) - x \in U_1$, contradiciendo la elección de y , y si $x + y \in U_2$, entonces $x = (x + y) - y \in U_2$, contradiciendo la de x . \square

1.3.7. La Proposición 1.3.6 nos dice que la unión de dos subespacios es un subespacio solo en situaciones muy especiales, pero no es difícil encontrar ejemplos de uniones de *muchos* subespacios que sí son subespacios. Por ejemplo, si \mathbb{k} es un cuerpo finito con q elementos y $n \in \mathbb{N}$ es mayor que 2, entonces el espacio vectorial \mathbb{k}^n es finito, tiene q^n elementos, y cada vector $v \in \mathbb{k}^n$ está contenido en $S_v = \{\lambda v : \lambda \in \mathbb{k}\}$, que es un subespacio propio de \mathbb{k}^n porque tiene exactamente q elementos, menos que los que tiene \mathbb{k}^n : esto nos dice que

$$\mathbb{k}^n = \bigcup_{v \in \mathbb{k}^n} S_v,$$

una unión finita de subespacios propios. Esto solo puede ocurrir cuando el cuerpo es finito, salvo en casos triviales: es lo que afirma la primera parte de la siguiente proposición.

Proposición. *Sea V un espacio vectorial que tiene al menos 3 subespacios.*

- (i) *Si \mathbb{k} es un cuerpo infinito, entonces V no es unión de un número finito de subespacios propios.*
 - (ii) *Si \mathbb{k} es un cuerpo finito de q elementos, entonces V no es unión de menos que q subespacios propios.*
-

Si V tiene menos que 3 subespacios, entonces los únicos subespacios son el nulo 0 y V mismo, así que si V no es unión de subespacios propios. Por otro lado, veremos más adelante en la Proposición 1.11.8 que cuando el cuerpo \mathbb{k} tiene un número finito q de elementos V es unión de $q + 1$ subespacios propios y no de menos: esto mejora la parte (ii).

Demostración. Sea V un espacio vectorial con al menos 3 subespacios y supongamos que $r \in \mathbb{N}$ y que S_1, \dots, S_r son subespacios propios de V tales que $V = \bigcup_{i=1}^r S_i$. Más aún, supongamos que elegimos a r y a los subespacios de manera tal que r sea lo mínimo posible. Para probar la proposición bastará que probemos que el cuerpo \mathbb{k} tiene menos que r elementos.

Para llegar a un absurdo, supongamos que por el contrario \mathbb{k} tiene más que r elementos. Como los subespacios S_1, \dots, S_r de V son propios y $V = \bigcup_{i=1}^r S_i$, necesariamente $r > 1$. Por otro lado, el subespacio V_r no es nulo: si lo fuera, entonces tendríamos que $V = \bigcup_{i=1}^{r-1} S_i$, en contra de la minimalidad de r .

Como S_r no es el subespacio nulo, podemos elegir $x \in S_r \setminus 0$, y como S_r es un subespacio propio de V , podemos elegir $y \in V \setminus S_r$. La función $f : \lambda \in \mathbb{k} \setminus \{0\} \mapsto x + \lambda y \in V$ es inyectiva: si λ y μ son dos elementos de $\mathbb{k} \setminus \{0\}$ tales que $f(\lambda) = f(\mu)$, entonces $x + \lambda y = x + \mu y$ y, por lo tanto, $(\lambda - \mu)y = 0$: como $y \neq 0$ esto nos dice que $\lambda - \mu = 0$, esto es, que $\lambda = \mu$. Por otro lado, si $\lambda \in \mathbb{k} \setminus \{0\}$, entonces $f(\lambda) \notin S_r$: en efecto, de pertenecer $f(\lambda)$ a S_r tendríamos que $y = \lambda^{-1}(x - (x + \lambda y)) = \lambda^{-1}(x - f(\lambda)) \in S_r$, contradiciendo la forma en que elegimos a y .

Ahora bien, para cada $\lambda \in \mathbb{k} \setminus \{0\}$ es $f(\lambda) \in V = \bigcup_{i=1}^r S_i$ y $f(\lambda) \notin S_r$, así que existe $j \in [\![r-1]\!]$ con $f(\lambda) \in V_j$. Como el conjunto $\mathbb{k} \setminus \{0\}$ tiene más que $r-1$ elementos, el principio de Dirichlet implica que existen $k \in [\![r-1]\!]$ y λ y $\mu \in \mathbb{k} \setminus \{0\}$ tales que $f(\lambda), f(\mu) \in S_k$ y $\lambda \neq \mu$ y, por lo tanto, que

$$(\mu - \lambda)x = \mu(x + \lambda y) - \lambda(x + \mu y) = \mu f(\lambda) - \lambda f(\mu) \in S_k.$$

Como $\mu - \lambda \neq 0$, podemos concluir que $x \in S_k \subseteq \bigcup_{i=1}^{r-1} S_i$.

Hemos probado con todo esto que cada elemento no nulo de S_r pertenece a $\bigcup_{i=1}^{r-1} S_i$ y, por lo tanto, que $V = \bigcup_{i=1}^r S_i = \bigcup_{i=1}^{r-1} S_i$: esto es absurdo en vista de la forma en que elegimos a r . \square

§4. Combinaciones lineales

1.4.1. Sea V un espacio vectorial. Si $S \subseteq V$ es un subconjunto de V , decimos que un vector $v \in V$ es **combinación lineal de los elementos de S** si existen $n \in \mathbb{N}_0$, elementos x_1, \dots, x_n de S y escalares $a_1, \dots, a_n \in \mathbb{k}$ tales que

$$x = a_1x_1 + \dots + a_nx_n.$$

En esta definición permitimos que n sea igual a 0: como una suma de cero elementos de V tiene valor 0, esto significa que el elemento 0 de V es una combinación lineal de los elementos de S . Por otro lado, observemos que en esta definición no incluimos la condición de que los vectores x_1, \dots, x_n sean distintos dos a dos.

Escribimos $\langle S \rangle$ al conjunto de todos los elementos de V que son combinación lineal de elementos de S . Cuando $S = \{x_1, \dots, x_n\}$ es un conjunto finito dado por la enumeración de sus elementos, escribimos $\langle x_1, \dots, x_n \rangle$ en lugar de $\langle \{x_1, \dots, x_n\} \rangle$.

1.4.2. Una primera observación, casi evidente, que podemos hacer es que la construcción de $\langle S \rangle$ a partir de S es una operación *monótona* en el siguiente sentido:

Proposición. *Sea V un espacio vectorial. Si S y T son subconjuntos de V y $S \subseteq T$, entonces se tiene que $\langle S \rangle \subseteq \langle T \rangle$.*

Demostración. En efecto, si $S \subseteq T \subseteq V$, entonces todo vector de V que es combinación lineal de elementos de S es, claramente, combinación lineal de elementos de T . \square

1.4.3. Una segunda observación que hay que hacer es que cualquiera sea el subconjunto S con el que empecemos el conjunto $\langle S \rangle$ es un subespacio:

Proposición. *Sea V un espacio vectorial. Si S es un subconjunto de V , entonces el subconjunto $\langle S \rangle$ es un subespacio de V que contiene a S .*

Demostración. Sea $S \subseteq V$ un subconjunto. Como siempre, verificamos cada una de las condiciones de la definición 1.3.1:

- Que 0 es un elemento de $\langle S \rangle$ es consecuencia de la observación hecha en 1.4.1.
- Supongamos que x e y son dos elementos de $\langle S \rangle$, de manera que existen $n, m \in \mathbb{N}_0$, elementos x_1, \dots, x_n e y_1, \dots, y_m de S , y escalares $a_1, \dots, a_n, b_1, \dots, b_m$ de \mathbb{k} tales que $x = a_1x_1 + \dots + a_nx_n$ e $y = b_1y_1 + \dots + b_my_m$. Pongamos $z_i = x_i$ y $c_i = a_i$ para cada $i \in \llbracket n \rrbracket$ y $z_i = y_{i-n}$ y $c_i = b_{i-n}$ para cada $i \in \llbracket n+1, n+m \rrbracket$. Tenemos entonces que

$$x + y = (a_1x_1 + \dots + a_nx_n) + (b_1y_1 + \dots + b_my_m) = c_1z_1 + \dots + c_{n+m}z_{n+m},$$

y esto muestra que $x + y$ es una combinación lineal de elementos de S , esto es, que es un elemento de $\langle S \rangle$.

- Sean ahora $a \in \mathbb{k}$ y $x \in \langle S \rangle$, de manera que existen $n \in \mathbb{N}_0$, vectores $x_1, \dots, x_n \in S$ y escalares $a_1, \dots, a_n \in \mathbb{k}$ tales que $x = a_1x_1 + \dots + a_nx_n$. Como

$$a \cdot x = a \cdot (a_1x_1 + \dots + a_nx_n) = a(a_1)_1 + \dots + a(a_n)_n = (aa_1)x_1 + \dots + (aa_n)x_n,$$

vemos que $a \cdot x$ es una combinación lineal de elementos de S y, entonces, que está en $\langle S \rangle$.

Para terminar, nos queda probar que $S \subseteq \langle S \rangle$, pero esto es inmediato: si $x \in S$, entonces $x = 1 \cdot x$ y esto muestra que x es una combinación lineal de elementos de S . \square

1.4.4. En vista de la Proposición 1.4.3 tiene sentido, dado un subconjunto S de V , llamar a $\langle S \rangle$ el **subespacio generado por S** , ya que se trata efectivamente de un subespacio. Por otro lado, si U es un subespacio de V y S es un subconjunto de V tal que $U = \langle S \rangle$, decimos que el subespacio U está **generado** por S y que el conjunto S **genera** a U .

1.4.5. Ejemplos.

- (a) Sea $n \in \mathbb{N}$, sea $V = \mathbb{k}^n$, y para cada $i \in \llbracket n \rrbracket$ escribamos e_i al elemento

$$(0, \dots, 0, \underbrace{1}_i, 0, \dots, 0)$$

de \mathbb{k}^n cuyas componentes son todas nulas salvo la i -ésima, que es igual a 1. El conjunto $S = \{e_1, \dots, e_n\}$ genera a V . En efecto, si $x = (x_1, \dots, x_n)$ es un vector de V , entonces claramente vale que

$$x = x_1e_1 + \dots + x_ne_n.$$

- (b) Sea $\mathbb{k}[X]$ el espacio vectorial de los polinomios en la variable X con coeficientes en \mathbb{k} . El conjunto $S = \{X^i : i \in \mathbb{N}_0\}$ de los monomios genera a $\mathbb{k}[X]$ como espacio vectorial: en efecto, por la definición misma de lo que es un polinomio todo elemento de $\mathbb{k}[X]$ es una combinación lineal de elementos de S .
- (c) Sea X un conjunto no vacío y sea $V = \mathbb{k}^X$ el espacio vectorial de todas las funciones $X \rightarrow \mathbb{k}$.

Para cada $x \in X$ sea $\delta_x : X \rightarrow \mathbb{k}$ la función tal que

$$\delta_x(y) = \begin{cases} 1, & \text{si } y = x; \\ 0, & \text{en cualquier otro caso.} \end{cases}$$

Sea, finalmente, $S = \{\delta_x : x \in X\}$. Afirmamos que

el conjunto S genera a V si y solamente si el conjunto X es finito.

Para verlo, supongamos primero que X es finito, sea n el número de sus elementos y sean x_1, \dots, x_n sus elementos; observemos que $x_i \neq x_j$ si i y j son dos elementos distintos de $\llbracket n \rrbracket$. Si $f \in V$, entonces

$$f = f(x_1)\delta_{x_1} + \dots + f(x_n)\delta_{x_n}. \quad (2)$$

En efecto, si y es un elemento de X , entonces existe $i \in \llbracket n \rrbracket$ tal que $y = x_i$, y evaluando la función que aparece en esta igualdad a la derecha vemos que

$$(f(x_1)\delta_{x_1} + \dots + f(x_n)\delta_{x_n})(y) = f(x_1)\delta_{x_1}(y) + \dots + f(x_n)\delta_{x_n}(y) = f(x_i) = f(y),$$

ya que $\delta_{x_j}(y) = 0$ si $j \in \llbracket n \rrbracket \setminus \{i\}$ y $\delta_{x_i}(y) = 1$. La igualdad (2) nos dice que $f \in \langle S \rangle$, y como esto es así cualquiera sea $f \in V$, que el conjunto S genera a V en este caso.

Supongamos ahora que X es infinito y mostremos que S no genera a V . Sea $f : X \rightarrow \mathbb{k}$ el elemento de V tal que $f(x) = 1$ para cada $x \in X$ y, para llegar a un absurdo, supongamos que $f \in \langle S \rangle$, de manera que existen $m \in \mathbb{N}_0$, elementos $x_1, \dots, x_m \in X$ y escalares $a_1, \dots, a_m \in \mathbb{k}$ tales que

$$f = a_1\delta_{x_1} + \dots + a_m\delta_{x_m}. \quad (3)$$

Como estamos suponiendo que X es infinito, existe un elemento $x \in X$ distinto de todos los elementos x_1, \dots, x_m , y entonces evaluando ambos lados de la igualdad (3) en x vemos que

$$1 = f(x) = (a_1\delta_{x_1} + \dots + a_m\delta_{x_m})(x) = 0.$$

Esto, por supuesto, es absurdo, y podemos concluir entonces que $V \not\subseteq \langle S \rangle$.

- (d) Sean $m, n \in \mathbb{N}$ y consideremos el espacio vectorial $M_{m,n}(\mathbb{k})$ de las matrices de tamaño $m \times n$ con coeficientes en \mathbb{k} . Si $(k, l) \in \llbracket m \rrbracket \times \llbracket n \rrbracket$, hay una matriz $E^{k,l} = (e_{i,j}^{k,l})_{i,j} \in M_{m,n}(\mathbb{k})$ que para cada $(i, j) \in \llbracket m \rrbracket \times \llbracket n \rrbracket$ tiene coeficiente (i, j) -ésimo

$$E_{i,j}^{k,l} = \begin{cases} 1 & \text{si } k = l \text{ y } l = j; \\ 0 & \text{en cualquier otro caso.} \end{cases}$$

En otras palabras, la matriz $E^{k,l}$ tiene todos sus coeficientes nulos salvo el (k, l) -ésimo, que es igual a 1. Afirmamos que

el conjunto $S = \{E^{k,l} : k \in \llbracket m \rrbracket, l \in \llbracket n \rrbracket\}$ genera a $M_{m,n}(\mathbb{k})$.

En efecto, si $A = (a_{i,j})$ es un elemento de $M_{m,n}(\mathbb{k})$, es inmediato verificar que

$$A = \sum_{k=1}^m \sum_{l=1}^n a_{k,l} E^{k,l} \in \langle S \rangle.$$

Observemos que S tiene exactamente mn elementos.

- (e) Si V es un espacio vectorial nulo, entonces $V = \langle \emptyset \rangle$. En efecto, sabemos que $\langle \emptyset \rangle$ es un subespacio de V , así que necesariamente contiene al elemento cero de V y, como ése es de hecho el único elemento de V , vale la igualdad que afirmamos. \diamondsuit

1.4.6. Proposición. *Sea V un espacio vectorial, sea U un subespacio de V y sea S un subconjunto de V . El subespacio $\langle S \rangle$ está contenido en U si y solamente si $S \subseteq U$.*

Demostración. Si $\langle S \rangle \subseteq U$, entonces $S \subseteq U$, ya que sabemos que $S \subseteq \langle S \rangle$: esto nos dice que la condición del enunciado es necesaria. Veamos que es suficiente.

Supongamos que $S \subseteq U$, sea $x \in \langle S \rangle$ y mostremos que $x \in U$. Como x es combinación lineal de elementos de S , existen $n \in \mathbb{N}_0$, $x_1, \dots, x_n \in S$ y $a_1, \dots, a_n \in \mathbb{k}$ tales que $x = a_1x_1 + \dots + a_nx_n$. Para cada $i \in [n]$ tenemos que $a_i x_i \in U$, ya que $x_i \in S \subseteq U$ y U es un subespacio, y entonces, juntando todo, vemos que $x = a_1x_1 + \dots + a_nx_n \in U$, como queríamos. \square

1.4.7. Un corolario sencillo y útil de esta proposición es el siguiente:

Corolario. *Sea V un espacio vectorial. Un subconjunto S de V es un subespacio de V si y solamente si $S = \langle S \rangle$.*

Demostración. Sea S un subconjunto de V . Si $S = \langle S \rangle$, entonces claramente S es un subespacio de V , ya que $\langle S \rangle$ lo es: esto prueba la implicación directa. Supongamos, para probar la recíproca, que S es un subespacio de V . Sabemos que $S \subseteq \langle S \rangle$, así que para ver que $S = \langle S \rangle$ es suficiente que mostremos que $\langle S \rangle \subseteq S$. Ahora bien, como por supuesto tenemos que $S \subseteq S$ y S es un subespacio, la Proposición 1.4.6 nos dice que $\langle S \rangle \subseteq S$, como queremos. \square

1.4.8. El siguiente punto que tenemos que hacer es una caracterización importante del subespacio generado por un conjunto:

Proposición. *Sea V un espacio vectorial. Si S es un subconjunto de V , entonces el subespacio $\langle S \rangle$ es el menor subespacio de V que contiene a S y, de hecho, coincide con la intersección de todos los subespacios de V que contienen a S ,*

$$\langle S \rangle = \bigcap_{\substack{U \leq V \\ S \subseteq U}} U.$$

Demostración. Llamemos W a la intersección que aparece en el enunciado. Como es una intersección de subespacios de V , la Proposición 1.3.5(ii) nos dice que W es un subespacio, y como todos

los subespacios que aparecen en la intersección contienen a S , tenemos claramente que $S \subseteq W$.

Afirmamos que W es el menor subespacio de V que contiene a S : en efecto, si U es otro subespacio de V que contiene a S , entonces U es uno de los espacios que aparecen en la intersección que define a W y, en consecuencia, es $W \subseteq U$. En particular, como $\langle S \rangle$ es un subespacio de V que contiene a S , esto nos dice que $W \subseteq \langle S \rangle$. Por otro lado, como W contiene a S y es un subespacio, la Proposición 1.4.6 nos dice que $\langle S \rangle \subseteq W$. Vemos así que $\langle S \rangle = W$ y esto completa la prueba de la proposición. \square

1.4.9. Si S es un subconjunto de V y x un vector, la definición 1.4.1 nos dice que x pertenece a $\langle S \rangle$ si existen $n \in \mathbb{N}_0$, vectores x_1, \dots, x_n y escalares $a_1, \dots, a_n \in \mathbb{k}$ tales que $x = a_1x_1 + \dots + a_nx_n$. Una observación sencilla y útil es que cuando ése es el caso, podemos elegir n , los vectores x_1, \dots, x_n y los escalares a_1, \dots, a_n de manera que esos n vectores sean distintos dos a dos y que los escalares correspondientes sean todos no nulos.

Lema. *Sea V un espacio vectorial y sea S un subconjunto de V . Un vector x de V pertenece a $\langle S \rangle$ si y solamente si existen $n \in \mathbb{N}_0$, vectores $x_1, \dots, x_n \in S$ distintos dos a dos y escalares $a_1, \dots, a_n \in \mathbb{k}$ todos no nulos tales que $x = a_1x_1 + \dots + a_nx_n$.*

Usaremos este resultado muchas veces de manera implícita, sin mencionarlo.

Demostración. Sea $x \in V$. Es claro que si se cumple la condición del enunciado se tiene que $x \in \langle S \rangle$, así que bastará que probemos que esa condición es necesaria. Supongamos entonces que $x \in \langle S \rangle$. Es posible entonces elegir $n \in \mathbb{N}_0$, $x_1, \dots, x_n \in S$ y $a_1, \dots, a_n \in \mathbb{k}$ de manera que se tenga $x = a_1x_1 + \dots + a_nx_n$. Más aún, es posible hacer esas elecciones de manera que el entero no negativo n sea lo menor posible: en otras palabras, de manera que x no sea combinación lineal de *menos* que n elementos de S .

Si hubiera índices i y j en $\llbracket n \rrbracket$ tales que $i < j$ y $x_i = x_j$, tendríamos que

$$x = a_1x_1 + \dots + a_ix_i + \dots + a_jx_j + \dots + a_nx_n = a_1x_1 + \dots + (a_i + a_j)x_i + \dots + \widehat{a_jx_j} + \dots + a_nx_n$$

y esto es imposible, porque el último miembro de esta cadena de igualdades es una combinación lineal de $n - 1$ elementos de S . Vemos así que los vectores x_1, \dots, x_n son distintos dos a dos. Por otro lado, si hubiera un índice $i \in \llbracket n \rrbracket$ tal que $a_i = 0$, entonces tendríamos que

$$x = a_1x_1 + \dots + a_ix_i + \dots + a_nx_n = a_1x_1 + \dots + \widehat{a_ix_i} + \dots + a_nx_n,$$

y esto es imposible por la misma razón que antes: los escalares a_1, \dots, a_n son, por lo tanto, todos no nulos. Esto prueba la necesidad de la condición del enunciado. \square

§5. Dependencia e independencia lineal

1.5.1. Sea V un espacio vectorial sobre \mathbb{k} . Si $n \in \mathbb{N}$ y x_1, \dots, x_n son vectores de V , una *relación de dependencia lineal* entre los vectores x_1, \dots, x_n es una igualdad de la forma

$$a_1x_1 + \dots + a_nx_n = 0,$$

con escalares $a_1, \dots, a_n \in \mathbb{k}$. Decimos que esa relación de dependencia lineal es *trivial* si los n escalares a_1, \dots, a_n son nulos y que es *no trivial* en caso contrario. Los vectores x_1, \dots, x_n son *linealmente dependientes* si existe una relación de dependencia lineal no trivial entre ellos, y *linealmente independientes* en caso contrario.

De manera similar, un subconjunto S de V es *linealmente dependiente* si existen un entero positivo $n \in \mathbb{N}$ y elementos x_1, \dots, x_n de S distintos dos a dos tales que hay una relación de dependencia lineal no trivial entre los vectores x_1, \dots, x_n , y es *linealmente independiente* en caso contrario. Se sigue inmediatamente de estas definiciones que

un subconjunto S de V es linealmente independiente si y solamente si todo subconjunto finito de S es linealmente independiente

y que

un subconjunto S de V es linealmente dependiente si y solamente si posee un subconjunto finito que es linealmente dependiente.

Por otro lado, las nociones de dependencia e independiente lineal de un subconjunto S de un espacio vectorial V son independientes de V en el siguiente sentido: si U es un subespacio de V que contiene a S , entonces

S es linealmente dependiente o independiente en tanto subconjunto de U si y solamente si lo es en tanto subconjunto de V .

Usaremos estas tres observaciones de manera implícita en lo que sigue.

1.5.2. Ejemplos.

- Sean $n \in \mathbb{N}$ y $x_1, \dots, x_n \in V$. Si alguno de estos n vectores es nulo o hay dos de ellos que son iguales, entonces los vectores x_1, \dots, x_n son linealmente dependientes.
- El subconjunto vacío \emptyset de V es linealmente independiente. Un subconjunto S de V con un solo elemento es linealmente independiente si y solamente si ese elemento no es nulo.
- Si $n \in \mathbb{N}$ y x_1, \dots, x_n son vectores de V , entonces las frases «los vectores x_1, \dots, x_n son linealmente independientes» y la frase «el conjunto $\{x_1, \dots, x_n\}$ es linealmente independiente» significan cosas distintas. Por ejemplo, si $n = 2$ y $x_1 = x_2 \neq 0$, entonces los vectores x_1, x_2 son linealmente dependientes, mientras que el conjunto $\{x_1, x_2\}$ es linealmente independiente, ya que coincide con el conjunto $\{x_1\}$. \diamond

1.5.3. Proposición. Sea V un espacio vectorial y sean S y T dos subconjuntos de V .

- (i) Si $S \subseteq T$ y S es linealmente dependiente, entonces T es linealmente dependiente.
- (ii) Si $S \subseteq T$ y T es linealmente independiente, entonces S es linealmente independiente.
- (iii) Si $0 \in S$, entonces S es linealmente dependiente.

Informalmente, la primera de estas afirmaciones nos dice que *agrandar* un conjunto linealmente dependiente no puede volverlo linealmente independiente, y la segunda que *achicar* un conjunto linealmente independiente no puede volverlo linealmente dependiente.

Demostración. (i) Supongamos que $S \subseteq T$ y que S es linealmente dependiente. Existen entonces $n \in \mathbb{N}$, vectores $x_1, \dots, x_n \in S$ distintos dos a dos, y escalares $a_1, \dots, a_n \in \mathbb{k}$ no todos nulos tales que $a_1x_1 + \dots + a_nx_n = 0$. Como $S \subseteq T$, cada uno de los vectores x_1, \dots, x_n está en T , y es claro que esto implica que T es linealmente dependiente.

(ii) Esta afirmación es la implicación contrarrecíproca de la de (i).

(iii) Si S contiene a 0, podemos tomar $n = 1$, $x_1 = 0$ y $a_1 = 1$: como $a_1x_1 = 0$, esto nos dice que el conjunto S es linealmente dependiente. \square

1.5.4. Un conjunto es linealmente independiente cuando no es linealmente dependiente: esa es la definición, pero es útil explicitar qué significa exactamente eso como en la siguiente proposición.

Proposición. Sea V un espacio vectorial. Un subconjunto S de V es linealmente independiente si y solo si satisface la siguiente condición:

cada vez que x_1, \dots, x_n son elementos de S distintos dos a dos y $a_1, \dots, a_n \in \mathbb{k}$ son escalares tales que $a_1x_1 + \dots + a_nx_n = 0$, se tiene que de hecho $a_1 = \dots = a_n = 0$.

Demostración. En efecto, la condición que aparece en el enunciado dice simplemente que el conjunto no es linealmente dependiente. \square

1.5.5. Una forma útil de pensar en la dependencia lineal es la que establece la siguiente proposición: un conjunto S es linealmente dependiente cuando, a los fines de generar al subespacio $\langle S \rangle$, es redundante, en el sentido de que podemos sacarle un elemento sin cambiar el espacio que genera.

Proposición. Sea V un espacio vectorial. Un subconjunto S de V es linealmente dependiente si y solo si existe $x \in S$ tal que $\langle S \setminus \{x\} \rangle = \langle S \rangle$.

Demostración. Sea S un subconjunto de V y supongamos primero que S es linealmente dependiente, de manera que existen un entero positivo $n \in \mathbb{N}$, vectores $x_1, \dots, x_n \in S$ distintos dos a dos, y escalares $a_1, \dots, a_n \in \mathbb{k}$ no todos nulos tales que

$$a_1x_1 + \dots + a_nx_n = 0. \tag{4}$$

Como los escalares no son todos nulos, existe $i \in \llbracket n \rrbracket$ tal que $a_i \neq 0$ y entonces de la igualdad (4) vemos que

$$x_i = (-a_i^{-1}a_1)x_1 + \cdots + \widehat{(-a_i^{-1}a_i)x_i} + \cdots + (-a_i^{-1}a_n)x_n.$$

Como los n vectores x_1, \dots, x_n son distintos dos a dos, los $n - 1$ vectores $x_1, \dots, \widehat{x_i}, \dots, x_n$ que aparecen aquí a la derecha están todos en $S \setminus \{x_i\}$, así que tenemos que $x_i \in \langle S \setminus \{x_i\} \rangle$. Como también $S \setminus \{x_i\} \subseteq \langle S \setminus \{x_i\} \rangle$, de esto deducimos que, de hecho,

$$S = \{x_i\} \cup (S \setminus \{x_i\}) \subseteq \langle S \setminus \{x_i\} \rangle$$

y, entonces, que $\langle S \rangle \subseteq \langle S \setminus \{x_i\} \rangle$. Se tiene claramente que $\langle S \setminus \{x_i\} \rangle \subseteq \langle S \rangle$, así que juntando todo podemos concluir que $\langle S \setminus \{x_i\} \rangle = \langle S \rangle$: la condición del enunciado es, por lo tanto, necesaria.

Veamos que también es suficiente. Supongamos que existe $x \in S$ tal que $\langle S \setminus \{x\} \rangle = \langle S \rangle$. En particular, como $x \in S \subseteq \langle S \rangle = \langle S \setminus \{x\} \rangle$, existen $n \in \mathbb{N}_0$, $x_1, \dots, x_n \in S \setminus \{x\}$ y $a_1, \dots, a_n \in \mathbb{k}$ tales que

$$x = a_1x_1 + \cdots + a_nx_n. \quad (5)$$

Más aún, de acuerdo al Lema 1.4.9, podemos suponer que los vectores x_1, \dots, x_n son distintos dos a dos. Como están todos en $S \setminus \{x\}$, son todos distintos de x y, por lo tanto, lo que tenemos es que los $n + 1$ vectores x, x_1, \dots, x_n son distintos dos a dos. Ahora bien, de la igualdad (5) se deduce que

$$(-1)x + a_1x_1 + \cdots + a_nx_n = 0$$

y, como los $n + 1$ vectores que aparecen aquí están en S y son distintos dos a dos, y claramente no todos los coeficientes son nulos, esto muestra que el conjunto S es linealmente dependiente. \square

1.5.6. De manera similar a la proposición anterior, podemos dar una caracterización frecuentemente cómoda del subespacio que genera un subconjunto en términos de la dependencia lineal:

Lema. Sea V un espacio vectorial y sea S un subconjunto de V que es linealmente independiente. Un vector x de V que no está en S pertenece a $\langle S \rangle$ si y solamente si el conjunto $S \cup \{x\}$ es linealmente dependiente.

Demostración. Sea x un vector de V tal que $x \notin S$. Supongamos primero que el conjunto $S \cup \{x\}$ es linealmente dependiente, de manera que existen $n \in \mathbb{N}$, vectores $x_1, \dots, x_n \in S \cup \{x\}$ distintos dos a dos y escalares $a_1, \dots, a_n \in \mathbb{k}$ no todos nulos tales que

$$a_1x_1 + \cdots + a_nx_n = 0. \quad (6)$$

No puede ser que todos los vectores x_1, \dots, x_n sean distintos de x : en ese caso estarían todos en S , y esto es imposible ya que S es linealmente independiente. Podemos suponer, entonces, y sin pérdida de generalidad, que $x_1 = x$. Observemos que esto implica, ya que los vectores x_1, \dots, x_n están en $S \cup \{x\}$ y son distintos dos a dos, que los vectores x_2, \dots, x_n están en S .

El escalar a_1 no puede ser nulo: si lo fuera, tendríamos que

$$a_2x_2 + \cdots + a_nx_n = a_1x_1 + \cdots + a_nx_n = 0,$$

con los vectores x_2, \dots, x_n distintos dos a dos y en S , y los escalares a_2, \dots, a_n no todos nulos, y esto es imposible, otra vez, porque el conjunto S es linealmente independiente. Podemos entonces deducir de la igualdad (6) que

$$x = x_1 = (-a_1^{-1}a_2)x_2 + \cdots + (-a_1^{-1}a_n)x_n \in \langle S \rangle.$$

Probemos ahora la implicación recíproca: supongamos que $x \in \langle S \rangle$, de manera que existen un entero $n \in \mathbb{N}_0$, vectores $x_1, \dots, x_n \in S$ distintos dos a dos, y escalares $a_1, \dots, a_n \in \mathbb{k}$ tales que $x = a_1x_1 + \cdots + a_nx_n$. Como x no pertenece a S por hipótesis, los $n+1$ vectores x, x_1, \dots, x_n de $S \cup \{x\}$ son distintos dos a dos y, como $(-1)x + a_1x_1 + \cdots + a_nx_n = 0$ y en la combinación lineal que aparece aquí a la izquierda no todos los coeficientes son nulos, el conjunto $S \cup \{x\}$ es linealmente dependiente. \square

1.5.7. Ejemplos.

- (a) Sea $n \in \mathbb{N}$. En el espacio \mathbb{k}^n los n vectores e_1, \dots, e_n son linealmente independientes. En efecto, si $\alpha_1, \dots, \alpha_n$ son escalares de \mathbb{k} , entonces es fácil ver que

$$\alpha_1e_1 + \cdots + \alpha_ne_n = (\alpha_1, \dots, \alpha_n)^t$$

y entonces si la combinación lineal que aparece a la izquierda en esta igualdad coincide con el elemento nulo de \mathbb{k}^n , es claro que $\alpha_1 = \cdots = \alpha_n = 0$.

- (b) Sea X un conjunto no vacío, sea \mathbb{k}^X el espacio del Ejemplo 1.2.4(a) y para cada $x \in S$ sea $\delta_x : X \rightarrow \mathbb{k}$ el elemento de \mathbb{k}^X descripto en el Ejemplo 1.4.5(c). Afirmamos que el conjunto $S = \{\delta_x : x \in X\}$ es linealmente independiente. Para verlo, sean $n \in \mathbb{N}$, sean x_1, \dots, x_n elementos de X distintos dos a dos, y sean $\alpha_1, \dots, \alpha_n$ escalares de \mathbb{k} , y supongamos que

$$\alpha_1\delta_{x_1} + \cdots + \alpha_n\delta_{x_n} = 0. \tag{7}$$

Si $i \in \llbracket n \rrbracket$, entonces $\delta_{x_j}(x_i) = 0$ si $j \in \llbracket n \rrbracket \setminus \{i\}$ y $\delta_{x_i}(x_i) = 1$, así que la igualdad de arriba implica que

$$0 = (\alpha_1\delta_{x_1} + \cdots + \alpha_n\delta_{x_n})(x_i) = \alpha_1\delta_{x_1}(x_i) + \cdots + \alpha_n\delta_{x_n}(x_i) = \alpha_i.$$

Como esto es así cualquiera sea el elemento i de $\llbracket n \rrbracket$, vemos que la relación de dependencia lineal (7) es trivial.

- (c) Sean $m, n \in \mathbb{N}$, sea $M_{m,n}(\mathbb{k})$ el espacio vectorial de las matrices de tamaño $m \times n$ con coeficientes en \mathbb{k} , y para cada $(k, l) \in \llbracket m \rrbracket \times \llbracket n \rrbracket$ sea $E^{k,l}$ la matriz descripta en el Ejemplo 1.4.5(d). El conjunto

$$S = \{E^{k,l} : k \in \llbracket m \rrbracket, l \in \llbracket n \rrbracket\}$$

es linealmente independiente. En efecto, si para cada $k \in \llbracket m \rrbracket$ y $l \in \llbracket n \rrbracket$ tenemos un escalar $a_{k,l} \in \mathbb{k}$ de manera que se tiene que

$$\sum_{k=1}^m \sum_{l=1}^n a_{k,l} E^{k,l} = 0,$$

es inmediato verificar que la matriz que aparece a la izquierda en esta igualdad tiene, para cada $(i, j) \in \llbracket m \rrbracket \times \llbracket n \rrbracket$, coeficiente (i, j) -ésimo igual a $a_{i,j}$, así que este escalar es nulo.

- (d) Sea $V = \mathbb{k}[X]$, el espacio vectorial de los polinomios sobre \mathbb{k} , y consideremos el conjunto

$$\mathcal{B}_1 = \{X^i : i \in \mathbb{N}_0\}$$

de las potencias de X . Mostremos que es linealmente independiente. Supongamos que $n \in \mathbb{N}$, que f_1, \dots, f_n son n elementos de \mathcal{B}_1 distintos dos a dos, y que $a_1, \dots, a_n \in \mathbb{k}$ son escalares tales que

$$a_1 f_1 + \dots + a_n f_n = 0. \quad (8)$$

En vista de la definición del conjunto \mathcal{B}_1 , para cada $k \in \llbracket n \rrbracket$ existe $i_k \in \mathbb{N}_0$ tal que $f_k = X^{i_k}$. Más aún, como los polinomios f_1, \dots, f_n son distintos dos a dos, los números i_1, \dots, i_n son distintos dos a dos. Ahora bien, en lado derecho de la igualdad (8) es igual al polinomio

$$a_1 X^{i_1} + a_2 X^{i_2} + \dots + a_n X^{i_n}$$

y aquella igualdad nos dice que este polinomio es el polinomio nulo: como las potencias de X que aparecen aquí son distintas dos a dos, vemos que, de hecho, cada uno de los coeficientes es nulo: esto es, que $a_1 = \dots = a_n = 0$. Podemos concluir entonces, como queríamos, que el conjunto \mathcal{B}_1 es linealmente independiente.

Consideremos ahora el conjunto

$$\mathcal{B}_2 = \{f_i : i \in \mathbb{N}_0\}$$

con

$$f_i = \sum_{k=0}^i X^k$$

para cada $i \in \mathbb{N}_0$, de manera que, por ejemplo, $f_0 = 1$ y $f_3 = 1 + X + X^2 + X^3$. Queremos mostrar que el conjunto \mathcal{B}_2 es linealmente independiente, y para ello probaremos que todo subconjunto finito de \mathcal{B}_2 es linealmente independiente haciendo inducción sobre su cardinal.

Sea S un subconjunto finito de \mathcal{B}_2 . Si el cardinal de S es 0, de manera que $S = \emptyset$, entonces S es linealmente independiente. Supongamos ahora que $n \in \mathbb{N}_0$ y que sabemos que todo subconjunto de \mathcal{B}_2 de cardinal n es linealmente independiente, y sea S un subconjunto de \mathcal{B}_2 de cardinal exactamente $n+1$, de manera que hay enteros $i_1, \dots, i_{n+1} \in \mathbb{N}_0$ distintos dos a dos tales que $S = \{f_{i_1}, \dots, f_{i_{n+1}}\}$, y sin pérdida de generalidad podemos suponer que están

ordenados de forma creciente, esto es, que $i_1 < \dots < i_{n+1}$. Supongamos que $\alpha_1, \dots, \alpha_{n+1} \in \mathbb{k}$ son escalares tales que

$$\alpha_1 f_{i_1} + \dots + \alpha_{n+1} f_{i_{n+1}} = 0. \quad (9)$$

Los polinomios f_{i_1}, \dots, f_{i_n} tiene todos grado menor que i_{n+1} , mientras que $f_{i_{n+1}}$ tiene grado $n+1$ y es mónico: el coeficiente de $X^{i_{n+1}}$ en el polinomio que aparece a la izquierda en la igualdad (9) es entonces α_{n+1} y, como el polinomio es nulo, vemos que $\alpha_{n+1} = 0$. Esto implica que $\alpha_1 f_{i_1} + \dots + \alpha_n f_{i_n} = 0$, que es una relación de dependencia lineal entre los elementos del subconjunto $\{f_{i_1}, \dots, f_{i_n}\}$ de \mathcal{B}_2 . La hipótesis de inducción nos dice entonces que todos los coeficientes que aparecen en ella son nulos. Así, tenemos que todos los coeficientes que aparecen en (9) son nulos, y, por lo tanto, que el conjunto S con el que empezamos es linealmente independiente. \diamond

§6. Bases

1.6.1. Sea V un espacio vectorial. Decimos que un subconjunto \mathcal{B} de V es una **base** de V si

- \mathcal{B} es un conjunto linealmente independiente, y
- \mathcal{B} genera a V , de manera que $V = \langle \mathcal{B} \rangle$.

Es importante observar que las dos condiciones de esta definición compiten entre sí: mientras más grande es un subconjunto del espacio V más fácil es que lo genere, a la vez que mientras más chico es más fácil es que sea linealmente independiente.

1.6.2. La siguiente proposición explicita dos caracterizaciones importantes de las bases de un espacio vectorial:

Proposición. *Sea V un espacio vectorial. Si \mathcal{B} es un subconjunto de V , entonces las siguientes afirmaciones son equivalentes:*

- (a) \mathcal{B} es una base de V .
- (b) \mathcal{B} es un conjunto linealmente independiente y es maximal con esta propiedad: todo subconjunto S de V tal que $\mathcal{B} \subsetneq S$ es linealmente dependiente.
- (c) \mathcal{B} genera a V y es minimal con esta propiedad: ningún subconjunto propio de \mathcal{B} genera a V .

Demostración. Sea \mathcal{B} un subconjunto de V .

(a) \Rightarrow (b) Supongamos que \mathcal{B} es una base de V . En vista de la definición de 1.6.1, el conjunto \mathcal{B} genera a V y es linealmente independiente. Tenemos que mostrar que es maximal con esta última propiedad. Sea $S \subseteq V$ un subconjunto tal que $S \supsetneq \mathcal{B}$, de manera que existe algún vector $x \in S \setminus \mathcal{B}$. Es claro que $\mathcal{B} \subseteq S \setminus \{x\}$ y, en consecuencia, que

$$V = \langle \mathcal{B} \rangle \subseteq \langle S \setminus \{x\} \rangle \subseteq \langle S \rangle \subseteq V,$$

de manera que $\langle S \setminus \{x\} \rangle = \langle S \rangle$. La Proposición 1.5.5 nos dice, entonces, que el conjunto S es linealmente dependiente. Vemos de esta forma que todo subconjunto que contiene propiamente a \mathcal{B} es linealmente dependiente, como queríamos.

(b) \Rightarrow (c) Supongamos ahora que el subconjunto \mathcal{B} es linealmente independiente y maximal con esa propiedad. Mostremos que \mathcal{B} genera a V y que es minimal con esta propiedad.

Empezamos por mostrar que $V = \langle \mathcal{B} \rangle$. Sea $x \in V$. Si $x \in \mathcal{B}$, es claro que $x \in \langle \mathcal{B} \rangle$, así que supongamos que no es ése el caso. El conjunto $\mathcal{B} \cup \{x\}$ es entonces un subconjunto de V que contiene propiamente a \mathcal{B} y la maximalidad de este último implica que $\mathcal{B} \cup \{x\}$ es un conjunto linealmente dependiente. Existen entonces $n \in \mathbb{N}$, vectores $x_1, \dots, x_n \in \mathcal{B} \cup \{x\}$ distintos dos a dos, y escalares $a_1, \dots, a_n \in \mathbb{k}$ no nulos tales que

$$a_1x_1 + \dots + a_nx_n = 0. \quad (10)$$

Observemos que no puede ser que todos los vectores x_1, \dots, x_n pertenezcan a \mathcal{B} , porque este conjunto es linealmente independiente: uno de ellos es entonces igual a x y, sin pérdida de generalidad, podemos suponer que, de hecho, es $x_1 = x$ y, por lo tanto, que los otros vectores x_2, \dots, x_n pertenecen a \mathcal{B} . En ese caso, la igualdad (10) implica que

$$x = x_1 = (-a_1^{-1}a_2)x_2 + \dots + (-a_1^{-1}a_n)x_n \in \langle \mathcal{B} \rangle.$$

Esto muestra que $\langle \mathcal{B} \rangle = V$, esto es, que \mathcal{B} genera a V , como queríamos.

Sea ahora T un subconjunto propio de \mathcal{B} y supongamos que T genera a V . Como $T \subsetneq \mathcal{B}$, existe un vector $x \in \mathcal{B} \setminus T$, y como $x \in V = \langle T \rangle$, existen $n \in \mathbb{N}_0$, vectores $x_1, \dots, x_n \in T$ distintos dos a dos, y escalares $a_1, \dots, a_n \in \mathbb{k}$ tales que $x = a_1x_1 + \dots + a_nx_n$. Esto nos dice que

$$a_1x_1 + \dots + a_nx_n + (-1)x = 0. \quad (11)$$

Como los vectores x_1, \dots, x_n son distintos dos a dos y son todos distintos de x , ya que están en T y x no, y los escalares que aparecen en la combinación lineal (11) no son todos nulos, vemos que el conjunto \mathcal{B} es linealmente dependiente. Esto es absurdo, ya que contradice nuestra hipótesis sobre \mathcal{B} . Esta contradicción provino de suponer que T , que es un subconjunto propio de \mathcal{B} , genera a V , y prueba, entonces, que \mathcal{B} es un subconjunto generador minimal de V , como queríamos.

(c) \Rightarrow (a) Supongamos finalmente que el subconjunto \mathcal{B} genera a V y que es minimal con esa propiedad. Para ver que se trata de una base, alcanza con que mostremos que es linealmente independiente. Supongamos, para llegar a una contradicción, que no lo es. La Proposición 1.5.5 nos dice que entonces existe $x \in \mathcal{B}$ tal que $\langle \mathcal{B} \setminus \{x\} \rangle = \langle \mathcal{B} \rangle$ y esto es absurdo, ya que en ese caso $\mathcal{B} \setminus \{x\}$ es un subconjunto propio de \mathcal{B} que genera a V . Esto completa la prueba de la proposición. \square

1.6.3. Decimos que un espacio vectorial V es *finitamente generado* si existe un subconjunto *finito* $S \subseteq V$ que genera a V , esto es, tal que $V = \langle S \rangle$.

1.6.4. Ejemplos.

- (a) Si $n \in \mathbb{N}$, el espacio $V = \mathbb{k}^n$ es finitamente generado. En efecto, mostramos en 1.4.5(a) que si para cada $i \in \llbracket n \rrbracket$ escribimos e_i al vector $(0, \dots, 1, \dots, 0)$, con el 1 en la posición i -ésima, entonces el conjunto $S = \{e_1, \dots, e_n\}$ genera a V . Como este conjunto S es finito, tenemos que V es finitamente generado.

Más generalmente, vimos en el Ejemplo 1.4.5(d) que para cada elección de m y n en \mathbb{N} , el espacio de matrices $M_{m,n}(\mathbb{k})$ puede ser generado por un conjunto de mn elementos, así que se trata de un espacio vectorial finitamente generado.

- (b) Sea $V = \mathbb{k}[X]$: queremos mostrar que V no es finitamente generado. Supongamos que, por el contrario, existe un subconjunto S de V que es finito y que genera a V . No puede ser que S sea vacío o que el único elemento de S sea el polinomio nulo, porque si fuera ese el caso tendríamos que $\langle S \rangle = 0$, el subespacio nulo de V , y V mismo no es nulo. Podemos entonces considerar el número

$$d = \max\{\deg f : f \in S, f \neq 0\},$$

ya que el conjunto de números naturales cuyo máximo estamos tomando es no vacío y finito. Ahora bien, como estamos suponiendo que el conjunto S genera a V y X^{d+1} es un elemento de V , existen $n \in \mathbb{N}_0$, $f_1, \dots, f_n \in S$ y $a_1, \dots, a_n \in \mathbb{k}$ tales que

$$X^{d+1} = a_1 f_1 + \dots + a_n f_n.$$

Esto es imposible: a la derecha del signo igual tenemos una combinación lineal de polinomios cada uno de los cuales es o nulo o de grado a lo sumo igual a d , así que el resultado es un polinomio o nulo o de grado a lo sumo igual a d , mientras que a la izquierda del signo igual tenemos un polinomio que tiene grado exactamente igual a $d+1$.

- (c) Sea X un conjunto no vacío, sea $V = \mathbb{k}^X$ es espacio vectorial de las funciones $X \rightarrow \mathbb{k}$, y para $x \in X$ sea $\delta_x : X \rightarrow \mathbb{k}$ la función que describimos en 1.4.5(c). Allí mostramos que cuando el conjunto X es finito el conjunto $S = \{\delta_x : x \in X\}$ genera a V , así que el espacio V es finitamente generado en ese caso. Por el contrario, si el conjunto X es infinito, el espacio V no es finitamente generado: probaremos esto más adelante en el Ejemplo 1.7.3. ◇

- 1.6.5.** La razón por la que nos interesa la noción de finita generación de un espacio vectorial es que vale la siguiente proposición:

Proposición. *Sea V un espacio vectorial. Si V es finitamente generado, entonces V posee una base finita. Más precisamente, si S es un subconjunto finito de V que lo genera, entonces existe una base de V contenida en S .*

Demostración. Sea $S = \{x_1, \dots, x_n\}$ un subconjunto finito de V tal que $\langle S \rangle = V$. Claramente S posee subconjuntos que generan a V —por ejemplo, S mismo es uno de ellos— y todos ellos tienen cardinal finito: hay entonces un subconjunto T de S de cardinal mínimo entre todos ellos,

esto es, existe un subconjunto T de S que genera a V y tal que todo subconjunto de S con menos elementos que T no genera a V . En particular, ningún subconjunto *propio* de T genera a V y, de acuerdo a la Proposición 1.6.2, esto significa que T es una base de V . \square

1.6.6. Esta proposición es importante en la práctica: si tenemos un espacio vectorial V y un subconjunto S de V que es finito y que genera a V , la proposición nos dice cómo encontrar explícitamente una base de V . Basta considerar los subconjuntos de V y encontrar entre ellos uno que sea una base — la proposición afirma que alguno es efectivamente una base, así que el procedimiento tiene garantizado el éxito. Si S tiene cardinal n , entonces S posee 2^n subconjuntos: para encontrar una base tenemos que considerar a lo sumo esa cantidad de conjuntos.

Podemos obtener una base de formas más eficientes. Supongamos que el subconjunto S de V es finito y genera a V , que n es el cardinal de S , y que x_1, \dots, x_n son los elementos de S . Si llevamos a cabo el siguiente procedimiento, al terminar el conjunto \mathcal{B} es una base de V contenida en S .

```

1    $\mathcal{B} \leftarrow \emptyset$ 
2   para cada  $i$  de 1 a  $n$ 
3       |   si  $x_i \notin \langle \mathcal{B} \rangle$  entonces
4           |       |    $\mathcal{B} \leftarrow \mathcal{B} \cup \{x_i\}$ 
5       |   fin
6   fin

```

Verifiquemos que esto realmente funciona. Pongamos $\mathcal{B}_0 = \emptyset$, para cada $i \in \llbracket n \rrbracket$ escribamos \mathcal{B}_i a valor que tiene la variable \mathcal{B} en el momento en que se termina de ejecutar por vez i -ésima el cuerpo del bucle que empieza en la línea 2, y para cada $i \in \llbracket 0, n \rrbracket$ sea $P(i)$ la afirmación

el conjunto \mathcal{B}_i es linealmente independiente, está contenido en $\{x_1, \dots, x_n\}$ y genera el mismo subespacio que $\{x_1, \dots, x_i\}$.

Observemos que la afirmación $P(0)$ es evidente y que si mostramos que $P(n)$ vale entonces tendremos que \mathcal{B}_n , que es el valor de la variable \mathcal{B} al terminar el procedimiento, es linealmente independiente, está contenido en $\{x_1, \dots, x_n\}$ y genera el mismo subespacio que este último conjunto: en otras palabras, que al terminar el procedimiento la variable \mathcal{B} contiene una base de V contenida en $\{x_1, \dots, x_n\}$, como queremos.

Para hacer esto, podemos proceder por inducción: sea $k \in \llbracket 0, n \rrbracket$ tal que $k < n$ y tal que la afirmación $P(k)$ vale y mostremos que también vale entonces $P(k+1)$. Ahora bien, en vista de la forma del procedimiento, hay dos posibilidades para el conjunto \mathcal{B}_{k+1} :

- Si $x_{k+1} \in \langle \mathcal{B}_k \rangle$, entonces $\mathcal{B}_{k+1} = \mathcal{B}_k$: es claro en este caso que \mathcal{B}_{k+1} es linealmente, está contenido en $\{x_1, \dots, x_{k+1}\}$ y que genera el mismo subespacio que este conjunto.
- Si, en cambio, se tiene que $x_{k+1} \notin \langle \mathcal{B}_k \rangle$, entonces $\mathcal{B}_{k+1} = \mathcal{B}_k \cup \{x_{k+1}\}$. Es fácil ver que como \mathcal{B}_k es linealmente independiente y $x_{k+1} \notin \mathcal{B}_k$, el conjunto \mathcal{B}_{k+1} es él mismo linealmente independiente. Por otro lado, como \mathcal{B}_k está contenido en $\{x_1, \dots, x_k\}$ y genera

el mismo subespacio que este conjunto, entonces $\mathcal{B}_{k+1} = \mathcal{B}_k \cup \{x_{k+1}\}$ está contenido en $\{x_1, \dots, x_{k+1}\}$ y genera el mismo subespacio que este conjunto.

En definitiva, en cualquiera de los dos casos vemos que vale la afirmación $P(k+1)$. Esto prueba, como dijimos, que el algoritmo funciona.

1.6.7. La Proposición 1.6.5 afirma que todo espacio vectorial finitamente generado posee una base. De hecho, esa hipótesis no es necesaria para obtener la conclusión, ya que vale el siguiente teorema:

Teorema. *Todo espacio vectorial posee una base. Más precisamente, si V es un espacio vectorial y S es un subconjunto linealmente independiente de V , entonces existe una base \mathcal{B} de V tal que $S \subseteq \mathcal{B}$.*

En estas notas no probaremos este resultado. La demostración del teorema se hace usualmente usando el llamado *Axioma de Elección* vía el *Lema de Zorn*, que es equivalente a él — más aún, Andreas Blass mostró en [Bla84] que el teorema es equivalente al Axioma de Elección. En un sentido que puede hacerse muy preciso, a pesar de que en vista del teorema todo espacio vectorial posee una base, no es en general posible construirla explícitamente.

§7. Dimensión

1.7.1. En la sección anterior mostramos que un espacio vectorial que es finitamente generado posee bases. En general, posee muchas: la observación fundamental que haremos en esta sección es que todas ellas tienen el mismo número de elementos. Para llegar a ese resultado el paso más importante es la siguiente proposición:

Proposición. *Sea V un espacio vectorial y sea $n \in \mathbb{N}_0$. Si V posee una base con n elementos, entonces todo subconjunto de V con más que n elementos es linealmente dependiente.*

Demostración. Supongamos que $\mathcal{B} = \{x_1, \dots, x_n\}$ es una base de V con n elementos, sea T un subconjunto de V que es linealmente independiente, y supongamos, para llegar a un absurdo, que hay en T más que n elementos. En particular, podemos elegir $n+1$ vectores y_1, \dots, y_{n+1} en T que son linealmente independientes y distintos dos a dos.

Mostraremos abajo que para cada $r \in \llbracket 0, n \rrbracket$ vale que

$$\begin{aligned} &\text{existen índices } i_1, \dots, i_{n-r} \in \llbracket n \rrbracket, \text{ distintos dos a dos, tales que el conjunto} \\ &\{y_1, \dots, y_r, x_{i_1}, \dots, x_{i_{n-r}}\} \text{ es una base de } V. \end{aligned} \tag{12}$$

En particular, mostraremos que esto vale cuando $r = n$, de manera que el conjunto $S = \{y_1, \dots, y_n\}$ es una base de V . Esto implica que S es linealmente independiente y que $y_{n+1} \notin \langle S \rangle$: de acuerdo al Lema 1.5.6, de estas dos cosas y de que $y_{n+1} \notin S$ se sigue que $S \cup \{y_{n+1}\}$ es linealmente dependiente, y esto es absurdo porque elegimos a los vectores y_1, \dots, y_{n+1} de manera que sean linealmente

independientes. Esta contradicción proviene de haber supuesto que el conjunto T tiene más que n elementos: la conclusión de la proposición es, por lo tanto, cierta.

Para probar que vale la afirmación (12) procedemos por inducción con respecto a r , que es un elemento del conjunto $\llbracket 0, n \rrbracket$. Notemos que cuando $r = 0$ no hay nada que probar, ya que basta tomar $i_j = j$ para cada $j \in \llbracket n \rrbracket$.

Sea entonces $k \in \llbracket 0, n - 1 \rrbracket$ y supongamos que la afirmación (12) vale cuando $r = k$, de manera que existen índices $i_1, \dots, i_{n-k} \in \llbracket n \rrbracket$, distintos dos a dos, tales que el conjunto

$$H = \{y_1, \dots, y_k, x_{i_1}, \dots, x_{i_{n-k}}\}$$

es una base de V . Como H genera a V , existen escalares $a_1, \dots, a_k, b_1, \dots, b_{n-k} \in \mathbb{k}$ tales que

$$y_{k+1} = a_1 y_1 + \dots + a_k y_k + b_1 x_{i_1} + \dots + b_{n-k} x_{i_{n-k}}. \quad (13)$$

No puede ser que todos los escalares b_1, \dots, b_{n-k} sean nulos: de ser ése el caso, a partir de la igualdad (13) obtendríamos una relación de dependencia lineal no trivial entre elementos del conjunto H , que es linealmente independiente. Así, alguno de esos escalares es no nulo: existe $j \in \llbracket n - k \rrbracket$ tal que $b_j \neq 0$. De la igualdad (13), entonces, podemos deducir que

$$\begin{aligned} x_{i_j} &= (-b_j^{-1} a_1) y_1 + \dots + (-b_j^{-1} a_k) y_k + b_j^{-1} y_{k+1} \\ &\quad + (-b_j^{-1} b_1) x_{i_1} + \dots + \widehat{(-b_j^{-1} b_j) x_{i_j}} + \dots + (-b_j^{-1} b_{n-k}) x_{i_{n-k}}. \end{aligned}$$

Sea $H' = \{y_1, \dots, y_{k+1}, x_{i_1}, \dots, \widehat{x_{i_j}}, \dots, x_{i_{n-k}}\}$. La expresión que encontramos para x_{i_j} nos dice que ese vector pertenece a $\langle H' \rangle$ y entonces que

$$\langle H' \rangle = \langle H' \cup \{x_{i_j}\} \rangle = \langle H \cup \{y_{k+1}\} \rangle \supseteq \langle H \rangle = V,$$

de manera que H' genera a V . Para ver que H' es una base nos resta probar que es linealmente independiente. Sean entonces $c_1, \dots, c_{k+1}, d_1, \dots, \widehat{d_j}, \dots, d_{n-k} \in \mathbb{k}$ escalares tales que

$$c_1 y_1 + \dots + c_{k+1} y_{k+1} + d_1 x_{i_1} + \dots + \widehat{d_j x_{i_j}} + \dots + d_{n-k} x_{i_{n-k}} = 0. \quad (14)$$

Ahora bien, si en esta igualdad usamos la expresión (13) para y_{k+1} , vemos que

$$\begin{aligned} c_1 y_1 + \dots + c_k y_k + c_{k+1} (a_1 y_1 + \dots + a_k y_k + b_1 x_{i_1} + \dots + b_{n-k} x_{i_{n-k}}) \\ + d_1 x_{i_1} + \dots + \widehat{d_j x_{i_j}} + \dots + d_{n-k} x_{i_{n-k}} = 0 \end{aligned}$$

y, agrupando términos, que

$$\begin{aligned} (c_1 + c_{k+1} a_1) y_1 + \dots + (c_k + c_{k+1} a_k) y_k \\ + (c_{k+1} b_1 + d_1) x_{i_1} + \dots + (c_{k+1} b_j + \widehat{d_j}) x_{i_j} + \dots + (c_{k+1} b_{n-k} + d_{n-k}) x_{i_{n-k}} = 0. \end{aligned}$$

Como el conjunto H es linealmente independiente, los coeficientes que aparecen en esta relación de dependencia lineal son todos nulos, esto es, valen las igualdades

$$\begin{aligned} c_q + c_{k+1}a_q &= 0 && \text{para cada } q \in \llbracket k \rrbracket, \\ c_{k+1}b_q + d_q &= 0 && \text{para cada } q \in \llbracket n-k \rrbracket \setminus \{j\}, \\ c_{k+1}b_j &= 0. \end{aligned}$$

Recordando que el escalar b_j no es nulo, la última de estas ecuaciones nos dice que $c_{k+1} = 0$ y, usando eso, todas las otras nos dicen que $c_1 = \dots = c_k = 0$ y que $d_1 = \dots = \widehat{d}_j = \dots = d_{n-k} = 0$. Así, hemos concluido que todos los coeficientes que aparecen en la relación de dependencia lineal (14) son todos nulos, como queríamos. Esto completa la inducción que prueba nuestra afirmación (12). \square

1.7.2. La afirmación (12) en la que basamos la prueba de esta proposición es conocida como el *Lema de Intercambio de Steinitz*, por Ernst Steinitz (1871–1928), que usó un argumento completamente similar en el contexto de la teoría de cuerpos.

1.7.3. Ejemplo. En el Ejemplo 1.6.4(c) vimos que cuando X es un conjunto no vacío y finito el espacio de funciones \mathbb{k}^X es finitamente generado. Podemos mostrar ahora que, como anunciamos ahí, si X es por el contrario infinito entonces el espacio \mathbb{k}^X no es finitamente generado.

En efecto, supongamos que X es un conjunto infinito pero que \mathbb{k}^X tiene dimensión finita, y sea $n = \dim \mathbb{k}^X$. Como X es infinito, podemos elegir $n+1$ elementos x_1, \dots, x_{n+1} en X distintos dos a dos y considerar el conjunto de funciones $S = \{\delta_{x_1}, \dots, \delta_{x_{n+1}}\}$. Ahora bien, como los $n+1$ puntos son distintos dos a dos, el conjunto S es linealmente independiente: esto contradice a la Proposición 1.7.1, ya que S tiene más que n elementos. \diamond

1.7.4. La consecuencia más importante de la Proposición 1.7.1 es el siguiente resultado, que es fundamental en todo lo que sigue:

Teorema. *Sea V un espacio vectorial. Si V posee alguna base finita, entonces todas sus bases son finitas y tienen el mismo número de elementos.*

Demostración. Supongamos que \mathcal{B} es una base finita de V con n elementos y que \mathcal{B}' es otra base cualquiera de V . Si \mathcal{B}' tiene más que n elementos, podemos elegir $n+1$ vectores y_1, \dots, y_{n+1} distintos dos a dos en \mathcal{B}' : el conjunto $\{y_1, \dots, y_{n+1}\}$ es entonces linealmente independiente y tiene $n+1$ elementos: esto es absurdo, ya que contradice a la Proposición 1.7.1. Vemos así que el conjunto \mathcal{B}' es necesariamente finito y que tiene a lo sumo n elementos. Esto prueba la primera afirmación del teorema e, intercambiando los roles de \mathcal{B} y de \mathcal{B}' en el razonamiento que acabamos de hacer, también la segunda. \square

1.7.5. Decimos que un espacio vectorial V tiene **dimensión finita** si posee una base finita y, en ese caso, que el cardinal de esa base, que es un elemento de \mathbb{N}_0 al que escribimos $\dim V$, es la **dimensión** de V ; en caso contrario decimos que V tiene **dimensión infinita**. El Teorema 1.7.4 nos

dice que todas las bases de un espacio vectorial de dimensión finita tienen la misma cantidad de elementos, así que la definición de la dimensión tiene sentido.

Observemos que un espacio tiene dimensión finita si y solamente si es finitamente generado. En efecto, si tiene una base finita entonces esa base es un conjunto finito que lo genera y, recíprocamente, si posee un subconjunto finito que lo genera, la Proposición 1.6.5 nos dice que posee una base contenida en ese conjunto y, por lo tanto, finita.

1.7.6. Ejemplos.

- (a) Si $n \in \mathbb{N}_0$, entonces $\dim \mathbb{k}^n = n$, ya que el conjunto $\{e_1, \dots, e_n\}$ es una de sus bases. Más generalmente, si $m, n \in \mathbb{N}$ entonces $\dim M_{m,n}(\mathbb{k}) = mn$, ya que el conjunto

$$S = \{E^{k,l} : k \in [\![k]\!], l \in [\![n]\!]\}$$

de las matrices descriptas en el Ejemplo 1.4.5(d) es una base de $M_{m,n}(\mathbb{k})$.

- (b) Si $n \in \mathbb{N}$ y $\mathbb{k}[X]_{\leq n}$ es el espacio vectorial de los polinomios de grado a lo sumo n , entonces $\dim \mathbb{k}[X]_{\leq n} = n + 1$, ya que el conjunto $\{1, X, \dots, X^n\}$ de $n + 1$ elementos es una de sus bases.
(c) Si X es un conjunto no vacío y finito, entonces $\dim \mathbb{k}^X = \#X$, ya que el conjunto $\{\delta_x : x \in X\}$ es una base de \mathbb{k}^X . \diamond

1.7.7. Por definición, que un espacio tenga dimensión infinita significa que no posee ninguna base finita. Es a veces útil tener una caracterización *positiva* de esa propiedad, como la siguiente:

Proposición. *Un espacio vectorial tiene dimensión infinita si y solamente si posee un subconjunto infinito y linealmente independiente, y cuando ese es el caso contiene subconjuntos finitos linealmente independientes de cualquier cardinal finito.*

Demostración. Sea V un espacio vectorial y supongamos primero que V tiene un subconjunto S que es infinito y linealmente independiente y que al mismo tiempo posee una base \mathcal{B} que es finita. Si n es el cardinal de \mathcal{B} , entonces S tiene más que n elementos y, por lo tanto, tiene un subconjunto T con $n + 1$ elementos. Este conjunto T es linealmente independiente y tiene más elementos que \mathcal{B} : esto es absurdo, ya que contradice a la Proposición 1.7.1. Esto muestra que la condición de la proposición es suficiente.

Supongamos que ahora, para probar la necesidad de esa condición, que V tiene dimensión infinita y mostremos inductivamente que

existe una sucesión $(S_i)_{i \geq 0}$ de subconjuntos de V linealmente independientes y tales que para cada $i \in \mathbb{N}_0$ el conjunto S_i tiene exactamente i elementos y $S_i \subsetneq S_{i+1}$. (15)

Empezamos poniendo $S_0 = \emptyset$, que es un subconjunto finito linealmente independiente de V con 0 elementos. Para continuar, supongamos que $n \in \mathbb{N}_0$ y que ya construimos el conjunto S_n . Como S_n es finito, no puede ser que $\langle S_n \rangle$ sea igual a V , porque en ese caso S_n contendría, de acuerdo a la Proposición 1.6.5, una base de V y ésta sería finita, contradiciendo nuestra hipótesis sobre V . Así,

existe un vector $x \in V \setminus \langle S_n \rangle$ y podemos considerar el conjunto $S_{n+1} = S_n \cup \{x\}$, que claramente contiene a S_n . Como $x \notin \langle S_n \rangle$, en particular $x \notin S_n$ y, en consecuencia, el conjunto S_{n+1} tiene $n + 1$ elementos. Veamos que S_{n+1} es linealmente independiente. Para ello, supongamos que, por el contrario, existen elementos x_1, \dots, x_k de S_{n+1} distintos dos a dos y escalares $a_1, \dots, a_k \in \mathbb{k}$ no nulos tales que

$$a_1x_1 + \dots + a_kx_k = 0. \quad (16)$$

Los vectores x_1, \dots, x_k no pueden ser todos distintos de x , porque en ese caso pertenecerían todos a S_n y la igualdad (16) sería una relación de dependencia lineal no trivial entre elementos de S_n , que es linealmente independiente. Existe entonces $i \in [k]$ tal que $x_i = x$ y, como consecuencia de esto, tenemos que $x_1, \dots, \widehat{x_i}, \dots, x_k \in S_n$. Se sigue entonces de (16) que

$$x = x_i = (-a_i^{-1}a_1)x_1 + \dots + \widehat{(-a_i^{-1}a_i)x_i} + \dots + (-a_i^{-1}a_k)x_k \in \langle S_n \rangle,$$

y esto es imposible en vista de la forma en que elegimos el vector x . Esta contradicción nos dice que no puede ser que S_{n+1} sea linealmente dependiente, así que es linealmente independiente.

Todo esto prueba la afirmación (15). Para probar la proposición, consideremos el conjunto $S = \bigcup_{n \geq 1} S_n$, que es infinito —ya que para cada $n \in \mathbb{N}$ contiene a S_n , que tiene exactamente n elementos— y mostremos que es linealmente independiente.

Si no lo fuese, existirían $k \in \mathbb{N}$, elementos y_1, \dots, y_k de S distintos dos a dos y escalares $b_1, \dots, b_k \in \mathbb{k}$ todos no nulos tales que

$$b_1y_1 + \dots + b_ky_k = 0. \quad (17)$$

Ahora bien, para cada $i \in [k]$ existe $m_i \in \mathbb{N}_0$ tal que $y_i \in S_{m_i}$, y podemos considerar el entero $m = \max\{m_1, \dots, m_k\}$. Se sigue de (15) que $y_1, \dots, y_k \in S_m$ y entonces la igualdad (17) implica que el conjunto S_m es linealmente dependiente. Esto es absurdo y esta contradicción termina la prueba de la proposición. Observemos incidentalmente que el subconjunto S de V que acabamos de construir no genera en general al espacio V . \square

1.7.8. Una aplicación inmediata de la proposición que acabamos de obtener es la prueba de que todo subespacio de un espacio de dimensión finita tiene dimensión finita.

Proposición. *Sea V un espacio vectorial de dimensión finita. Si U es un subespacio de V , entonces U tiene dimensión finita y vale que $\dim U \leq \dim V$.*

Demostración. Sea U un subespacio de V y sea $n = \dim V$. Si U tuviera dimensión infinita, la Proposición 1.7.7 nos diría que posee un subconjunto S linealmente independiente con $n + 1$ elementos: como S está contenido en V , esto es absurdo en vista de la Proposición 1.7.1.

Vemos así que U tiene que ser necesariamente de dimensión finita. Sea \mathcal{B} una base de U . Como \mathcal{B} es un subconjunto linealmente independiente de V , la Proposición 1.7.1 afirma que tiene a lo sumo n elementos, y esto significa precisamente que $\dim U \leq \dim V$. \square

1.7.9. Una base de un espacio vectorial es un subconjunto que es linealmente independiente y que lo genera, y en general ninguna de estas dos condiciones implica la otra. Cuando el espacio tiene dimensión finita, sin embargo, podemos decir más:

Proposición. *Sea V un espacio vectorial de dimensión finita n y sea S un subconjunto de V de exactamente n elementos. Las siguientes tres afirmaciones son equivalentes:*

- (a) *El conjunto S es una base de V .*
- (b) *El conjunto S es linealmente independiente.*
- (c) *El conjunto S genera a V .*

Demostración. La implicación $(a) \Rightarrow (b)$ es evidente. Supongamos, para probar la implicación $(b) \Rightarrow (c)$, que el conjunto S es linealmente independiente y que S no genera a V . Existe entonces un vector v en $V \setminus \langle S \rangle$: como $v \notin \langle S \rangle$ y S es linealmente independiente, el Lema 1.5.6 nos dice que el conjunto $S \cup \{v\}$ es linealmente independiente, y esto es imposible porque este conjunto tiene $n + 1$ elementos, un número mayor que la dimensión de V .

Veamos, finalmente, que la implicación $(c) \Rightarrow (a)$ vale. Supongamos que S genera a V y que no es una base, de manera que es linealmente dependiente. Según la Proposición 1.5.5 existe $x \in S$ tal que $\langle S \setminus \{x\} \rangle = \langle S \rangle = V$ y, por lo tanto, el conjunto $S \setminus \{x\}$ genera a V . Se sigue de esto que hay una base \mathcal{B} contenida en $S \setminus \{x\}$, y esta base tiene a lo sumo $n - 1$ elementos, ya que $S \setminus \{x\}$ tiene esa cantidad de elementos: esto es absurdo, ya que $n = \dim V$. \square

1.7.10. Otra aplicación útil de la proposición con la que empezamos esta sección es la siguiente, que nos dice que todo subconjunto linealmente independiente de un espacio vectorial de dimensión finita puede *completarse* a una base del espacio.

Proposición. *Sea V un espacio vectorial de dimensión finita. Si S es un subconjunto de V que es linealmente independiente, entonces existe una base \mathcal{B} de V tal que $S \subseteq \mathcal{B}$.*

Demostración. Sea n la dimensión de V y sea S un subconjunto de V que es linealmente independiente. Notemos que S es necesariamente finito y que tiene a lo sumo n elementos. Claramente existen subconjuntos de V que contienen a S y que son linealmente independientes – por ejemplo, S mismo — y todos ellos, de acuerdo a la Proposición 1.7.1, son finitos y de cardinal menor o igual a n . Podemos entonces elegir un subconjunto \mathcal{B} de V que contiene a S y es linealmente independiente, y que tiene el máximo cardinal posible entre todos los subconjuntos de V con esas dos propiedades. Mostremos que este conjunto \mathcal{B} es una base de V . Como contiene a S , esto probará la proposición.

Si T es un subconjunto de V que es linealmente independiente y tal que $T \not\supseteq \mathcal{B}$, entonces existe $x \in T \setminus \mathcal{B}$ y el conjunto $T \cup \{x\}$ contiene a S , es linealmente independiente, y tiene un elemento más que T : esto es absurdo, porque contradice la forma en que elegimos a T . Vemos así que \mathcal{B} es un subconjunto de V linealmente independiente y maximal con esa propiedad, así que es una base, como queremos. \square

1.7.11. Una forma útil de pensar en la dimensión de un espacio vectorial de dimensión finita es como una medida de su tamaño, que nos permite reducir comparaciones de espacios y subespacios a comparaciones de las dimensiones de esos espacios y subespacios. Un ejemplo de esto es el siguiente resultado:

Proposición. *Sea V un espacio vectorial de dimensión finita y sea U un subespacio de V .*

- (i) *Es $U = V$ si y solamente si $\dim U = \dim V$.*
- (ii) *Es $U = 0$ si y solamente si $\dim U = 0$.*

Demostración. (i) La condición es evidentemente necesaria. Para ver que es suficiente, supongamos que $\dim U = \dim V$ y sea \mathcal{B} una base de U . Se tiene entonces, por supuesto, que $\langle \mathcal{B} \rangle = U$; por otro lado, como \mathcal{B} es un subconjunto linealmente independiente de V con cantidad de elementos igual a $\dim V$, genera a V y, en consecuencia, $\langle \mathcal{B} \rangle = V$. Así, es $U = V$, como queremos.

(ii) Sabemos que la dimensión de un espacio nulo es 0. Recíprocamente, si $\dim U = 0$, entonces el conjunto vacío \emptyset es una base de U y esto es posible solamente si $U = 0$. \square

1.7.12. Otro ejemplo de la utilidad de la dimensión como medida del tamaño de los espacios es el siguiente resultado que nos será útil más adelante:

Proposición. *Sea V un espacio vectorial de dimensión finita. Si $(F_r)_{r \in \mathbb{N}}$ es una sucesión de subespacios de V tales que*

- o bien $F_r \subseteq F_{r+1}$ para todo $r \in \mathbb{N}$
- o bien $F_r \supseteq F_{r+1}$ para todo $r \in \mathbb{N}$,

entonces existe $r_0 \in \mathbb{N}$ tal que $F_r = F_{r_0}$ para todo $r \in \mathbb{N}$ con $r \geq r_0$.

En otras palabras, toda sucesión creciente o decreciente de subespacios de un espacio V de dimensión finita se estabiliza a partir de cierto punto: la prueba que daremos deduce la proposición de la afirmación bien conocida de que toda sucesión creciente y acotada o decreciente de elementos de \mathbb{N}_0 se estabiliza a partir de algún momento.

Demostración. Sea n la dimensión de V y para cada $r \in \mathbb{N}$ sea n_r la de F_r . Supongamos que vale la primera de las dos hipótesis del enunciado, de manera que $F_r \subseteq F_{r+1}$ para cada $r \in \mathbb{N}$. De acuerdo a la Proposición 1.7.8, tenemos entonces que $n_r = \dim F_r \leq \dim F_{r+1} = n_{r+1}$ para cada $r \in \mathbb{N}$ y, por lo tanto, que la sucesión de números $(n_r)_{r \geq 1}$ de \mathbb{N}_0 es creciente. Como además $n_r = \dim F_r \leq \dim V = n$ para todo $r \in \mathbb{N}$, esa sucesión de números es acotada: se sigue de esto que existe $r_0 \in \mathbb{N}$ tal que $n_r = n_{r_0}$ para todo $r \in \mathbb{N}$ con $r \geq r_0$. Pero entonces si $r \in \mathbb{N}$ es tal que $r \geq r_0$ tenemos que $F_r \subseteq F_{r_0}$ y que $\dim F_r = \dim F_{r_0}$, así que, de hecho, es $F_{r_0} = F_r$, como afirma la proposición.

Si hubiéramos supuesto que vale no la primera sino la segunda de las dos hipótesis del enunciado, procediendo exactamente de la misma forma obtendríamos que la sucesión $(n_r)_{r \geq 1}$ es ahora decreciente. Como es una sucesión de elementos de \mathbb{N}_0 , se deduce de ello inmediatamente que

existe $r_0 \in \mathbb{N}$ tal que $n_r = n_{r_0}$ para todo $r \in \mathbb{N}$ con $r \geq r_0$, y podemos continuar el razonamiento de la misma forma que antes. \square

§8. Sumas de subespacios

1.8.1. Sea V un espacio vectorial. Si $n \geq 0$ y U_1, \dots, U_n son subespacios de V , entonces denotamos $U_1 + \dots + U_n$ o $\sum_{i=1}^n U_i$ y llamamos **suma de los subespacios U_1, \dots, U_n** al subespacio de V generado por la unión $U_1 \cup \dots \cup U_n$, esto es, ponemos

$$U_1 + \dots + U_n = \langle U_1 \cup \dots \cup U_n \rangle.$$

Las convenciones usuales implican que si $n = 0$ la unión $U_1 \cup \dots \cup U_n$ denota al conjunto vacío, así que la suma $U_1 + \dots + U_n$ es en ese caso el subespacio nulo de V .

1.8.2. De acuerdo a la definición, la suma de una familia de subespacios es el conjunto de las combinaciones lineales de los elementos de la unión de éstos. La descripción alternativa de esta suma dada en la siguiente proposición es muchas veces útil:

Proposición. *Sea V un espacio vectorial. Si $n \geq 0$ y U_1, \dots, U_n son subespacios de V , entonces*

$$U_1 + \dots + U_n = \{x_1 + \dots + x_n : x_1 \in U_1, \dots, x_n \in U_n\}.$$

Demostración. Si $n = 0$ la afirmación es evidente, así que podemos suponer que $n > 0$. Escribamos T al conjunto que aparece a la derecha en la igualdad que tenemos que probar. Para ver que coincide con $U_1 + \dots + U_n$, es decir, con $\langle U_1 \cup \dots \cup U_n \rangle$, mostraremos que es un subespacio de V que contiene a $U_1 \cup \dots \cup U_n$ y que es el menor subespacio de V con esta propiedad: en vista de la Proposición 1.4.8, esto es suficiente.

Empecemos por mostrar que el subconjunto T es un subespacio de V :

- Como $0 \in U_i$ para cada $i \in \llbracket n \rrbracket$, es $0 = 0 + \dots + 0 \in T$.
- Sean x e y dos elementos de T , de manera que para cada $i \in \llbracket n \rrbracket$ existen x_i e y_i en U_i tales que $x = x_1 + \dots + x_n$ e $y = y_1 + \dots + y_n$. Entonces

$$x + y = (x_1 + \dots + x_n) + (y_1 + \dots + y_n) = (x_1 + y_1) + \dots + (x_n + y_n)$$

y esto nos dice que $x + y$ está en T , ya que $x_i + y_i \in U_i$ para cada $i \in \llbracket n \rrbracket$.

- Sean $a \in \mathbb{k}$ y $x \in T$. Existen entonces $x_1 \in U_1, \dots, x_n \in U_n$ tales que $x = x_1 + \dots + x_n$ y, en consecuencia,

$$a \cdot x = a \cdot (x_1 + \dots + x_n) = ax_1 + \dots + ax_n.$$

Como $ax_i \in U_i$ para cada $i \in \llbracket n \rrbracket$, esto muestra que $ax \in T$.

Para terminar, veamos que T es el menor subespacio de V que contiene a $U_1 \cup \dots \cup U_n$: supongamos que W es un subespacio de V que contiene a $U_1 \cup \dots \cup U_n$ y mostremos que necesariamente también contiene a T . Ahora bien, si $x \in T$, entonces existen $x_1 \in U_1, \dots, x_n \in U_n$ tales que $x = x_1 + \dots + x_n$ y, como $x_i \in U_i \subseteq W$ para cada $i \in \llbracket n \rrbracket$, esto implica que $x \in W$, ya que W es un subespacio. \square

1.8.3. Frecuentemente describimos un subespacio de un espacio vectorial dando un conjunto que lo genera. La siguiente proposición nos dice cómo describir la suma de una familia de subespacios vía un conjunto que lo genere cuando conocemos un conjunto generador para cada uno de los sumandos:

Proposición. *Sea V un espacio vectorial, sea $n \geq 0$ y consideremos subespacios U_1, \dots, U_n y subconjuntos S_1, \dots, S_n de V . Si para cada $i \in \llbracket n \rrbracket$ el conjunto S_i genera a U_i , entonces la unión $S_1 \cup \dots \cup S_n$ genera a $U_1 + \dots + U_n$.*

En otras palabras, vale la igualdad $\langle S_1 \rangle + \dots + \langle S_n \rangle = \langle S_1 \cup \dots \cup S_n \rangle$.

Demostración. Supongamos que para cada $i \in \llbracket n \rrbracket$ es $U_i = \langle S_i \rangle$ y sea $S = S_1 \cup \dots \cup S_n$. Como para cada $i \in \llbracket n \rrbracket$ es $S_i \subseteq S$, se tiene que $U_i = \langle S_i \rangle \subseteq \langle S \rangle$. Así, es $U_1 \cup \dots \cup U_n \subseteq \langle S \rangle$ y, por lo tanto, $U_1 + \dots + U_n = \langle U_1 \cup \dots \cup U_n \rangle \subseteq \langle S \rangle$. Por otro lado, como claramente $S \subseteq U_1 + \dots + U_n$ y $U_1 + \dots + U_n$ es un subespacio de V , es $\langle S \rangle \subseteq U_1 + \dots + U_n$. Concluimos de esta forma que $\langle S \rangle = U_1 + \dots + U_n$, como afirma la proposición. \square

1.8.4. Si A y B son dos conjuntos finitos, el llamado *Principio de Inclusión–Exclusión* nos permite calcular el cardinal de la unión $A \cup B$ en término de los de A , de B y de la intersección $A \cap B$: afirma que se tiene que

$$\#(A \cup B) + \#(A \cap B) = \#A + \#B.$$

El siguiente resultado es un análogo de este resultado para la dimensión de espacios vectoriales.

Proposición. *Sea V un espacio vectorial. Si U_1 y U_2 son subespacios de V de dimensión finita, entonces tanto $U_1 \cap U_2$ como $U_1 + U_2$ son subespacios de dimensión finita de V y vale que*

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

Más aún, existe una base \mathcal{B} de $U_1 + U_2$ tal que las intersecciones $\mathcal{B} \cap U_1$, $\mathcal{B} \cap U_2$ y $\mathcal{B} \cap U_1 \cap U_2$ son bases de U_1 , de U_2 y de $U_1 \cap U_2$, respectivamente.

Demostración. Sean U_1 y U_2 subespacios de dimensión finita de V . Como $U_1 \cap U_2$ es un subespacio de U_1 y este último tiene dimensión finita, la Proposición 1.7.8 nos dice que $U_1 \cap U_2$ tiene dimensión finita. Sea $n = \dim U_1 \cap U_2$ y sea $\mathcal{B}_0 = \{x_1, \dots, x_n\}$ una base de $U_1 \cap U_2$. Sean, por otro lado, $r = \dim U_1$ y $s = \dim U_2$.

Como \mathcal{B}_0 es linealmente independiente y está contenido en U_1 , las Proposiciones 1.7.8 y 1.7.10 nos dicen que $r \geq n$ y que hay vectores $y_1, \dots, y_{r-n} \in U_1$ tales que $\mathcal{B}_1 = \{x_1, \dots, x_n, y_1, \dots, y_{r-n}\}$

es una base de U_1 . De manera similar, es $s \geq n$ y existen vectores $z_1, \dots, z_{s-n} \in U_2$ tales que el conjunto $\mathcal{B}_2 = \{x_1, \dots, x_n, z_1, \dots, z_{s-n}\}$ es una base de U_2 . Mostraremos que

el conjunto $\mathcal{B} = \{x_1, \dots, x_n, y_1, \dots, y_{r-n}, z_1, \dots, z_{s-n}\}$ es una base de $U_1 + U_2$ con exactamente $r + s - n$ elementos.

Por un lado, esto implicará que $U_1 + U_2$ tiene dimensión finita y, por otro, que

$$\dim(U_1 + U_2) = r + s - n = \dim U_1 + \dim U_2 - \dim(U_1 \cap U_2),$$

como afirma la proposición. Que la base \mathcal{B} tiene además la propiedad descripta en la última oración del enunciado es claro.

Como $\mathcal{B} = \mathcal{B}_1 \cup \mathcal{B}_2$, la Proposición 1.8.3 nos dice que $\langle \mathcal{B} \rangle = \langle \mathcal{B}_1 \cup \mathcal{B}_2 \rangle = U_1 + U_2$ y, en consecuencia, que el conjunto \mathcal{B} genera a $U_1 + U_2$. Veamos que es linealmente independiente. Supongamos que $a_1, \dots, a_n, b_1, \dots, b_{r-n}, c_1, \dots, c_{s-n} \in \mathbb{k}$ son escalares tales que

$$a_1x_1 + \dots + a_nx_n + b_1y_1 + \dots + b_{r-n}y_{r-n} + c_1z_1 + \dots + c_{s-n}z_{s-n} = 0. \quad (19)$$

Si ponemos $x = a_1x_1 + \dots + a_nx_n$, $y = b_1y_1 + \dots + b_{r-n}y_{r-n}$ y $z = c_1z_1 + \dots + c_{s-n}z_{s-n}$, la igualdad (19) nos dice que $U_1 \ni -(x + y) = z \in U_2$ y, entonces, que $z \in U_1 \cap U_2$. Como \mathcal{B}_0 es una base de $U_1 \cap U_2$, existen escalares $d_1, \dots, d_n \in \mathbb{k}$ tales que $d_1x_1 + \dots + d_nx_n = z = c_1z_1 + \dots + c_{s-n}z_{s-n}$, es decir,

$$d_1x_1 + \dots + d_nx_n + (-c_1)z_1 + \dots + (-c_{s-n})z_{s-n} = 0.$$

Ahora bien: el conjunto \mathcal{B}_2 es linealmente independiente, así que de esta última igualdad podemos concluir que $c_1 = \dots = c_{s-n} = 0$. Usando esto y la ecuación (19) vemos que

$$a_1x_1 + \dots + a_nx_n + b_1y_1 + \dots + b_{r-n}y_{r-n} = 0$$

y, como el conjunto \mathcal{B}_1 es linealmente independiente, que $a_1 = \dots = a_n = 0$ y $b_1 = \dots = b_{r-n} = 0$. Así, todos los coeficientes que aparecen en (19) son nulos. Esto implica, primero, que los $r + s - n$ vectores $x_1, \dots, x_n, y_1, \dots, y_{r-n}, z_1, \dots, z_{s-n}$ son distintos dos a dos y, en segundo lugar, que el conjunto \mathcal{B} es linealmente independiente. Esto prueba (18), como queríamos. \square

1.8.5. Corolario. *Sea V un espacio vectorial. Si $n \in \mathbb{N}$ y U_1, \dots, U_n son subespacios de V de dimensión finita, entonces el subespacio $U_1 + \dots + U_n$ de V también tiene dimensión finita y*

$$\dim(U_1 + \dots + U_n) \leq \dim U_1 + \dots + \dim U_n.$$

Demostración. Procedamos haciendo inducción con respecto al entero n , notando que cuando $n = 1$ la afirmación es evidente. Supongamos entonces que $r \in \mathbb{N}$, que la afirmación del corolario es cierta cuando $n = r$, y que U_1, \dots, U_{r+1} son subespacios de V de dimensión finita. Es $U_1 + \dots + U_{r+1} = (U_1 + \dots + U_r) + U_{r+1}$ y sabemos gracias a la hipótesis inductiva que la suma $U_1 + \dots + U_r$ tiene dimensión finita. De acuerdo a la Proposición 1.8.4 la suma $U_1 + \dots + U_{r+1}$ tiene

entonces dimensión finita y

$$\begin{aligned}\dim(U_1 + \cdots + U_{r+1}) &= \dim((U_1 + \cdots + U_r) + U_{r+1}) \\ &= \dim((U_1 + \cdots + U_r)) + \dim U_{r+1} - \dim((U_1 + \cdots + U_r) \cap U_{r+1}).\end{aligned}$$

Como la tercera de estas dimensiones es un entero no negativo, esto es

$$\leq \dim((U_1 + \cdots + U_r)) + \dim U_{r+1}$$

y la hipótesis inductiva nos dice que esto es

$$\leq \dim U_1 + \cdots + \dim U_r + \dim U_{r+1}.$$

Esto completa la inducción. \square

§9. Sumas directas de subespacios

1.9.1. Sea V un espacio vectorial, sea $k \in \mathbb{N}_0$, y sean U_1, \dots, U_k subespacios de V . Decimos que los subespacios U_1, \dots, U_k son **independientes** si para cada $i \in [k]$ se tiene que

$$U_i \cap (U_1 + \cdots + \widehat{U_i} + \cdots + U_k) = 0.$$

Si además vale que $V = U_1 + \cdots + U_k$, entonces decimos que V es la **suma directa** de U_1, \dots, U_k y escribimos $V = U_1 \oplus \cdots \oplus U_k$.

Es importante observar que cuando U_1, \dots, U_k son subespacios independientes de un espacio V , la notación $U_1 + \cdots + U_k$ y la notación $U_1 \oplus \cdots \oplus U_k$ denotan el *mismo* subespacio de V : la única diferencia es que la segunda deja explícito el hecho de que los subespacios en cuestión son independientes y sólo la usamos en ese caso.

1.9.2. Ejemplo. La definición de independencia de subespacios generaliza a la de independencia lineal de vectores en el siguiente sentido. Si x_1, \dots, x_k son elementos no nulos de V , entonces las siguientes dos afirmaciones son equivalentes:

- (a) Los vectores x_1, \dots, x_k son linealmente independientes.
- (b) Los subespacios $\langle x_1 \rangle, \dots, \langle x_k \rangle$ son independientes.

Veamos que (a) \Rightarrow (b). Supongamos que los vectores x_1, \dots, x_k son linealmente independientes, sea $i \in [n]$, y sea $u \in \langle x_i \rangle \cap (\langle x_1 \rangle + \cdots + \widehat{\langle x_i \rangle} + \cdots + \langle x_k \rangle)$. Como $u \in \langle x_1 \rangle + \cdots + \widehat{\langle x_i \rangle} + \cdots + \langle x_k \rangle$, existen $u_1 \in \langle x_1 \rangle, \dots, u_i \in \langle x_i \rangle, \dots, u_k \in \langle x_k \rangle$ tales que $u = u_1 + \cdots + \widehat{u_i} + \cdots + u_k$, y entonces existen escalares $\lambda_1, \dots, \widehat{\lambda_i}, \dots, \lambda_k \in \mathbb{k}$ tales que $u_j = \lambda_j x_j$ para cada $j \in [k] \setminus \{i\}$. Por otro lado, como $u \in \langle x_i \rangle$, existe $\lambda_i \in \mathbb{k}$ tal que $u = -\lambda_i x_i$. Juntando todo, vemos que

$$\lambda_1 x_1 + \cdots + \widehat{\lambda_i x_i} + \cdots + \lambda_k x_k = u = -\lambda_i x_i$$

y, por lo tanto, que

$$\lambda_1x_1 + \cdots + \lambda_kx_k = 0.$$

Como los vectores x_1, \dots, x_k son linealmente independientes, esto implica que $\lambda_i = 0$ y, en consecuencia, que $u = \lambda_i x_i = 0$. Vemos así que la intersección $\langle x_i \rangle \cap (\langle x_1 \rangle + \cdots + \widehat{\langle x_i \rangle} + \cdots + \langle x_k \rangle)$ es nula.

Para probar la implicación $(b) \Rightarrow (a)$ supongamos que los subespacios $\langle x_1 \rangle, \dots, \langle x_k \rangle$ son independientes y $\lambda_1, \dots, \lambda_k \in \mathbb{k}$ escalares tales que $\lambda_1x_1 + \cdots + \lambda_kx_k = 0$. Si estos escalares no son todos nulos, de manera que existe $i \in [k]$ con $\lambda_i \neq 0$, entonces

$$\langle x_i \rangle \ni -\lambda_i x_i = \lambda_1x_1 + \cdots + \widehat{\lambda_i x_i} + \cdots + \lambda_kx_k \in \langle x_1 \rangle + \cdots + \widehat{\langle x_i \rangle} + \cdots + \langle x_k \rangle,$$

así que

$$\lambda_i x_i \in \langle x_i \rangle \cap (\langle x_1 \rangle + \cdots + \widehat{\langle x_i \rangle} + \cdots + \langle x_k \rangle) = 0.$$

Esto es absurdo, ya que $x_i \neq 0$ y $\lambda_i \neq 0$. Vemos así que $\lambda_i = 0$ para todo $i \in [k]$ y, por lo tanto que los vectores x_1, \dots, x_k son linealmente independientes.

Notemos que para la equivalencia $(a) \Leftrightarrow (b)$ necesitamos que los vectores x_1, \dots, x_n sean no nulos. Por ejemplo, si es $x_i = 0$ para cada $i \in [k]$ entonces ciertamente los vectores x_1, \dots, x_k no son linealmente independientes pero los subespacios $\langle x_1 \rangle, \dots, \langle x_k \rangle$ sí son independientes. \diamond

1.9.3. La siguiente observación bien sencilla nos permite hacer pruebas por inducción:

Lema. Sea V un espacio vectorial, sea $k \in \mathbb{N}_0$, y sean U_1, \dots, U_k subespacios de V . Si los subespacios U_1, \dots, U_k son independientes, entonces los subespacios U_1, \dots, U_{k-1} son independientes y, en consecuencia, el subespacio $U_1 + \cdots + U_{k-1}$ es la suma directa de U_1, \dots, U_{k-1} .

Demostración. En efecto, si los subespacios U_1, \dots, U_k de V son independientes e $i \in [k-1]$, entonces

$$U_i \cap (U_1 + \cdots + \widehat{U_i} + \cdots + U_{k-1}) \subseteq U_i \cap (U_1 + \cdots + \widehat{U_i} + \cdots + U_{k-1} + U_k)$$

y esta última intersección es el subespacio nulo de V . \square

1.9.4. La definición de independencia de subespacios está enunciada en términos de operaciones de conjuntos y subespacios. La siguiente proposición nos da una versión equivalente pero en términos de operaciones entre vectores.

Proposición. Sea V un espacio vectorial, sea $k \in \mathbb{N}_0$, y sean U_1, \dots, U_k subespacios de V . Las siguientes tres afirmaciones son equivalentes:

- (a) Es $V = U_1 \oplus \cdots \oplus U_k$.
- (b) Para cada $x \in V$ existen únicos $x_1 \in U_1, \dots, x_k \in U_k$ tales que $x = x_1 + \cdots + x_k$.
- (c) Es $V = U_1 + \cdots + U_k$ y si $x_1 \in U_1, \dots, x_k \in U_k$ son tales que $x_1 + \cdots + x_k = 0$, entonces

$$x_1 = \dots = x_k = 0.$$

Demostración. (a) \Rightarrow (b) Supongamos que $V = U_1 \oplus \dots \oplus U_k$ y sea $x \in V$. Como $V = U_1 + \dots + U_k$, la Proposición 1.8.2 nos dice que existen $x_1 \in U_1, \dots, x_k \in U_k$ tales que $x = x_1 + \dots + x_k$. Así, vale la afirmación de existencia de (b). Para ver la unicidad, supongamos que los vectores $x'_1 \in U_1, \dots, x'_k \in U_k$ también son tales que $x = x'_1 + \dots + x'_k$. Se tiene entonces que

$$(x_1 - x'_1) + \dots + (x_k - x'_k) = (x_1 + \dots + x_k) - (x'_1 + \dots + x'_k) = x - x = 0.$$

Si $i \in \llbracket k \rrbracket$, esto implica que

$$-(x_i - x'_i) = (x_1 - x'_1) + \dots + (\widehat{x_i - x'_i}) + \dots + (x_k - x'_k)$$

y como el lado izquierdo de esta igualdad está en U_i y el lado derecho en $U_1 + \dots + \widehat{U_i} + \dots + U_k$, vemos que, de hecho, es

$$x_i - x'_i \in U_i \cap (U_1 + \dots + \widehat{U_i} + \dots + U_k).$$

La hipótesis nos dice que esta intersección es el subespacio nulo de V , así que $x_i - x'_i = 0$, esto es, $x_i = x'_i$. Concluimos así que también vale la afirmación de unicidad de (b).

(b) \Rightarrow (c) Si vale (b), entonces que $V = U_1 + \dots + U_k$ es consecuencia inmediata de la Proposición 1.8.2 y la última afirmación de (c) es el caso particular de (b) en el que $x = 0$.

(c) \Rightarrow (a) Supongamos que vale la condición (c) y mostremos que entonces también vale la condición de la definición 1.9.1. Sea $i \in \llbracket k \rrbracket$ y sea x un elemento de $U_i \cap (U_1 + \dots + \widehat{U_i} + \dots + U_k)$. Esto significa, por un lado, que $x \in U_i$ y, por otro, que existen $x_1 \in U_1, \dots, \widehat{x_i} \in U_i, \dots, x_k \in U_k$ tales que $x = x_1 + \dots + \widehat{x_i} + \dots + x_k$. Como esto nos dice que

$$x_1 + \dots + x_{i-1} - x + x_{i+1} + \dots + x_k = 0,$$

la hipótesis (c) implica, entre otras cosas, que $x = 0$. Así, la intersección $U_i \cap (U_1 + \dots + \widehat{U_i} + \dots + U_k)$ sólo contiene al vector nulo y esto es lo que queríamos probar. \square

1.9.5. Proposición. *Sea V un espacio vectorial. Si $k \in \mathbb{N}_0$ y U_1, \dots, U_k son subespacios independientes de V , todos de dimensión finita, entonces el subespacio $U_1 \oplus \dots \oplus U_k$ de V también tiene dimensión finita y*

$$\dim(U_1 \oplus \dots \oplus U_k) = \dim U_1 + \dots + \dim U_k.$$

Más aún, si para cada $i \in \llbracket k \rrbracket$ el conjunto \mathcal{B}_i es una base de U_i , entonces la unión $\mathcal{B} = \mathcal{B}_1 \cup \dots \cup \mathcal{B}_k$ es disjunta y es una base de $U_1 \oplus \dots \oplus U_k$.

Demostración. Probemos la primera afirmación haciendo inducción sobre k .

- Si $k = 0$ o $k = 1$, no hay nada que probar.

- Sea ahora $n \geq 2$, supongamos que la proposición es cierta si $k = n - 1$ y sean U_1, \dots, U_n subespacios independientes de V . De acuerdo al Lema 1.9.3, los subespacios U_1, \dots, U_{n-1} son independientes, así que la hipótesis inductiva nos permite concluir que

$$\dim(U_1 \oplus \cdots \oplus U_{n-1}) = \dim U_1 + \cdots + \dim U_{n-1}.$$

Por otro lado, los subespacios $U_1 \oplus \cdots \oplus U_{n-1}$ y U_n de V son independientes, ya que la hipótesis nos dice que $(U_1 + \cdots + U_{n-1}) \cap U_n = 0$, y entonces la Proposición 1.8.4 nos dice que

$$\begin{aligned}\dim(U_1 \oplus \cdots \oplus U_n) &= \dim((U_1 \oplus \cdots \oplus U_{n-1}) \oplus U_n) \\ &= \dim(U_1 \oplus \cdots \oplus U_{n-1}) + \dim U_n \\ &= \dim U_1 + \cdots + \dim U_{n-1} + \dim U_n.\end{aligned}$$

Esto completa la inducción.

Probemos ahora la segunda afirmación del enunciado. Supongamos que $k \in \mathbb{N}_0$, que los subespacios U_1, \dots, U_k son independientes, y que para cada $i \in \llbracket k \rrbracket$ el conjunto \mathcal{B}_i es una base de U_i . Si i y j son elementos de $\llbracket k \rrbracket$ distintos, entonces

$$\mathcal{B}_i \cap \mathcal{B}_j \subseteq U_i \cap U_j \subseteq U_i \cap (U_1 + \cdots + \widehat{U_j} + \cdots + U_k) = 0,$$

así que, como 0 no pertenece a la intersección $\mathcal{B}_i \cap \mathcal{B}_j$, esta intersección es vacía. Vemos así que los conjuntos $\mathcal{B}_1, \dots, \mathcal{B}_k$ son disjuntos dos a dos y, por lo tanto, que la unión $\mathcal{B} = \mathcal{B}_1 \cup \cdots \cup \mathcal{B}_k$ es disjunta. En particular, tenemos que

$$\#\mathcal{B} = \#(\mathcal{B}_1 \cup \cdots \cup \mathcal{B}_k) = \#\mathcal{B}_1 + \cdots + \#\mathcal{B}_k = \dim U_1 + \cdots + \dim U_k = \dim(U_1 \oplus \cdots \oplus U_k).$$

Como además \mathcal{B} genera a $U_1 \oplus \cdots \oplus U_k$, ya que

$$\langle \mathcal{B} \rangle = \langle \mathcal{B}_1 \cup \cdots \cup \mathcal{B}_k \rangle = \langle \mathcal{B}_1 \rangle + \cdots + \langle \mathcal{B}_k \rangle = U_1 \oplus \cdots \oplus U_k,$$

la Proposición 1.7.9 nos dice que \mathcal{B} es una base de $U_1 \oplus \cdots \oplus U_k$. □

1.9.6. La primera parte de la Proposición 1.9.5 que acabamos de probar tiene una recíproca parcial que nos da un criterio numérico para decidir de una familia de subespacios es independiente o no.

Proposición. *Sea V un espacio vectorial y sean U_1, \dots, U_k subespacios de V de dimensión finita. Si $\dim(U_1 + \cdots + U_k) = \dim U_1 + \cdots + \dim U_k$, entonces los subespacios U_1, \dots, U_k son independientes y la suma $U_1 + \cdots + U_k$ es directa.*

Demostración. Supongamos que $\dim(U_1 + \cdots + U_k) = \dim U_1 + \cdots + \dim U_k$. Para probar la proposición es suficiente que mostremos que los subespacios U_1, \dots, U_k son independientes, porque la segunda afirmación sigue inmediatamente de eso. Supongamos, para llegar a una contradicción, que no son independientes, de manera que existe $i \in \llbracket k \rrbracket$ tal que la intersección

$U_i \cap (U_1 + \dots + \widehat{U}_i + \dots + U_k)$ no es el subespacio nulo de V y tiene, por lo tanto, dimensión positiva. Se tiene entonces que

$$\begin{aligned}\dim(U_1 + \dots + U_k) &= \dim(U_i + (U_1 + \dots + \widehat{U}_i + \dots + U_k)) \\ &= \dim U_i + \dim(U_1 + \dots + \widehat{U}_i + \dots + U_k) - \dim U_i \cap (U_1 + \dots + \widehat{U}_i + \dots + U_k) \\ &< \dim U_i + \dim(U_1 + \dots + \widehat{U}_i + \dots + U_k) \\ &\leq \dim U_i + (\dim U_1 + \dots + \widehat{\dim U_i} + \dots + \dim U_k) \\ &= \dim U_1 + \dots + \dim U_k = \dim(U_1 + \dots + U_k).\end{aligned}$$

Esto es, por supuesto, absurdo. \square

1.9.7. La segunda parte de la Proposición 1.9.5 también tiene una recíproca parcial:

Proposición. Sea V un espacio vectorial y sea \mathcal{B} una base de V . Si $\mathcal{B}_1, \dots, \mathcal{B}_n$ son subconjuntos de \mathcal{B} disjuntos dos a dos tales que $\mathcal{B} = \bigcup_{i=1}^n \mathcal{B}_i$, entonces $V = \langle \mathcal{B}_1 \rangle \oplus \dots \oplus \langle \mathcal{B}_n \rangle$.

Demostración. Sean $\mathcal{B}_1, \dots, \mathcal{B}_n$ subconjuntos de \mathcal{B} disjuntos dos a dos y tales que $\mathcal{B} = \bigcup_{i=1}^n \mathcal{B}_i$. Como claramente $V = \langle \mathcal{B} \rangle = \langle \mathcal{B}_1 \cup \dots \cup \mathcal{B}_n \rangle = \langle \mathcal{B}_1 \rangle + \dots + \langle \mathcal{B}_n \rangle$, solo tenemos que mostrar que los subespacios $\langle \mathcal{B}_1 \rangle, \dots, \langle \mathcal{B}_n \rangle$ son independientes. Sea $i \in [\![n]\!]$. Es

$$\begin{aligned}\langle \mathcal{B}_i \rangle \cap (\langle \mathcal{B}_1 \rangle + \dots + \widehat{\langle \mathcal{B}_i \rangle} + \dots + \langle \mathcal{B}_n \rangle) &= \langle \mathcal{B}_i \rangle \cap \langle \mathcal{B}_1 \cup \dots \cup \widehat{\mathcal{B}_i} \cup \dots \cup \mathcal{B}_n \rangle \\ &= \langle \mathcal{B}_i \rangle \cap \langle \mathcal{B} \setminus \mathcal{B}_i \rangle.\end{aligned}$$

Sea x un elemento de este subespacio. Como x está en $\langle \mathcal{B}_i \rangle$, existen $r \in \mathbb{N}_0$, elementos y_1, \dots, y_r de \mathcal{B}_i distintos dos a dos, y escalares $\alpha_1, \dots, \alpha_r \in \mathbb{k}$ tales que $x = \alpha_1 y_1 + \dots + \alpha_r y_r$. Por otro lado, como x está en $\langle \mathcal{B} \setminus \mathcal{B}_i \rangle$, existen $s \in \mathbb{N}_0$, elementos z_1, \dots, z_s de $\mathcal{B} \setminus \mathcal{B}_i$ distintos dos a dos, y escalares $\beta_1, \dots, \beta_s \in \mathbb{k}$ tales que $x = \beta_1 z_1 + \dots + \beta_s z_s$. Ahora bien, como \mathcal{B}_i y $\mathcal{B} \setminus \mathcal{B}_i$ son conjuntos disjuntos, tenemos que los $r+s$ elementos $y_1, \dots, y_r, z_1, \dots, z_s$ de \mathcal{B} son disjuntos dos a dos y

$$\alpha_1 y_1 + \dots + \alpha_r y_r + (-\beta_1) z_1 + \dots + (-\beta_s) z_s = 0$$

es una relación de dependencia lineal entre ellos: como el conjunto \mathcal{B} es linealmente independiente, todos los coeficientes que aparecen en ella tienen que ser nulos. En particular, tenemos que $\alpha_1 = \dots = \alpha_r = 0$ y, por lo tanto que $x = 0$. Vemos así que el subespacio intersección $\langle \mathcal{B}_i \rangle \cap (\langle \mathcal{B}_1 \rangle + \dots + \widehat{\langle \mathcal{B}_i \rangle} + \dots + \langle \mathcal{B}_n \rangle)$ solo contiene al vector nulo, que es lo que queríamos. \square

1.9.8. Otra situación en la que podemos dar una condición suficiente para la independencia de una familia de subespacios es la del próximo resultado:

Proposición. Sea V un espacio vectorial, sean V_1, \dots, V_r subespacios de V y supongamos que $V = V_1 \oplus \dots \oplus V_r$. Si para cada $i \in [\![r]\!]$ tenemos un subespacio U_i de V_i y ponemos $U = U_1 + \dots + U_r$, entonces esta suma es directa, de manera que $U = U_1 \oplus \dots \oplus U_r$.

Demostración. Es suficiente que probemos que los subespacios U_1, \dots, U_k son independientes: esto es inmediato, ya que para cada $i \in \llbracket k \rrbracket$ tenemos que

$$U_i \cap (U_1 + \cdots + \widehat{U_i} + \cdots + U_k) \subseteq V_i \cap (V_1 + \cdots + \widehat{V_i} + \cdots + V_k) = 0,$$

porque los subespacios V_1, \dots, V_k son independientes por la hipótesis. \square

§10. Complementos y codimensión

1.10.1. Sea V un espacio vectorial. Si S es un subespacio de V , llamamos *complemento de S en V* a todo subespacio T de V tal que $V = S \oplus T$.

1.10.2. Ejemplos.

- (a) Cualquiera sea V , el único complemento del subespacio nulo 0 en V es V mismo, mientras que el único complemento del subespacio impropio V en V es el subespacio nulo 0 .
- (b) Si V tiene dimensión 2 y $\{e_1, e_2\}$ es una base de V , para todo $\lambda \in \mathbb{k}$ el subespacio $\langle \lambda e_1 + e_2 \rangle$ es un complemento para el subespacio $\langle e_1 \rangle$ y, de hecho, es fácil ver que todos sus complementos son de esa forma. \diamond

1.10.3. En general, como muestra este último ejemplo, un subespacio de un espacio vectorial posee muchos complementos. La siguiente observación es útil cuando queremos compararlos:

Lema. *Sea V un espacio vectorial y sea S un subespacio de V . Si T y T' son dos complementos de S en V y $T \subseteq T'$, entonces, de hecho, $T = T'$.*

Demostración. Sea $x \in T'$. Como $V = S \oplus T$, existen $y \in S$ y $z \in T$ únicamente determinados tales que $x = y + z$. Por otro lado, como $V = S \oplus T'$, existen $y' \in S$ y $z \in T'$, también únicamente determinados, tales que $x = y' + z'$. Pero entonces

$$0 = x - x = (y + z) - (y' + z') = (y - y') + (z - z'),$$

así que $S \ni y - y' = -(z - z') \in T'$, ya que $T \subseteq T'$. Vemos de esta forma que $z - z' \in S \cap T' = 0$, esto es, que $z' = z \in T$. Esto prueba que $T' \subseteq T$. \square

1.10.4. Una forma de construir complementos de subespacios es usando bases adaptadas:

Proposición. *Sea V un espacio vectorial y sea W un subespacio de V . Si \mathcal{B} es una base de V y el conjunto $\mathcal{B}' = \mathcal{B} \cap W$ genera a W , entonces \mathcal{B}' es una base de W y el conjunto $\mathcal{B}'' = \mathcal{B} \setminus \mathcal{B}'$ es base de un complemento de W en V .*

Demostración. Sea \mathcal{B} una base de V tal que $\mathcal{B}' = \mathcal{B} \cap W$ genera a W y sea $\mathcal{B}'' = \mathcal{B} \setminus \mathcal{B}'$. Como \mathcal{B} es la unión disjunta de \mathcal{B}' y \mathcal{B}'' , la Proposición 1.9.7 nos dice que

$$V = \langle \mathcal{B}' \rangle \oplus \langle \mathcal{B}'' \rangle = W \oplus \langle \mathcal{B}'' \rangle,$$

esto es, que el subespacio $\langle \mathcal{B}'' \rangle$ es un complemento para W en V que tiene como base al conjunto \mathcal{B}'' , ya que este es linealmente independiente. \square

1.10.5. Corolario. Si V es un espacio vectorial, entonces todo subespacio de V posee un complemento en V .

Demostración. Sea V un espacio vectorial y sea W un subespacio de V . Sabemos que W posee una base \mathcal{B}' , y como este conjunto \mathcal{B}' es un subconjunto linealmente independiente de V , el Teorema 1.6.7 nos dice que existe una base \mathcal{B} de V tal que $\mathcal{B}' \subseteq \mathcal{B}$. Observemos que $\mathcal{B}' = \mathcal{B} \cap W$: que \mathcal{B}' está contenido en $\mathcal{B} \cap W$ es evidente, y si hubiera un elemento x en esta intersección que no está en \mathcal{B}' , entonces como sí esta en $W = \langle \mathcal{B}' \rangle$, tendríamos que el conjunto $\mathcal{B}' \cup \{x\}$ es linealmente dependiente, lo que es imposible ta está contenido en \mathcal{B} . Ahora, como $\mathcal{B} \cap W$ genera a W , la Proposición 1.10.4 nos dice que el subespacio $\langle \mathcal{B} \setminus \mathcal{B}' \rangle$ es un complemento para W en V . \square

1.10.6. De la misma forma que si un espacio vectorial tiene una base finita, todas las bases son finitas y tienen el mismo cardinal, el siguiente resultado nos dice que si un subespacio de un espacio vectorial tiene un complemento de dimensión finita, entonces todos sus complementos tienen dimensión finita y, de hecho, la misma.

Proposición. Sea V un espacio vectorial, sea S un subespacio de V y sean T_1 y T_2 dos complementos de S en V . Si T_1 tiene dimensión finita, entonces T_2 también tiene dimensión finita y, de hecho, $\dim T_1 = \dim T_2$.

Demostración. Supongamos que T_1 tiene dimensión finita, sea $n = \dim T_1$ y sea $\mathcal{B} = \{x_1, \dots, x_n\}$ una base de T_1 . Para cada $i \in [n]$ existen $s_i \in S$ e $y_i \in T_2$ tales que $x_i = s_i + y_i$, ya que $V = S \oplus T_2$. Sea $y \in T_2$. Como $V = S \oplus T_1$, existen $s \in S$ y $x \in T_1$ tales que $y = s + x$, y como \mathcal{B} es una base de T_1 , existen escalares $a_1, \dots, a_n \in \mathbb{k}$ tales que $x = a_1x_1 + \dots + a_nx_n$. Observemos que

$$\begin{aligned} y &= s + x \\ &= s + a_1x_1 + \dots + a_nx_n \\ &= s + a_1(s_1 + y_1) + \dots + a_n(s_n + y_n) \\ &= (s + a_1s_1 + \dots + a_ns_n) + (a_1y_1 + \dots + a_ny_n), \end{aligned}$$

de manera que

$$y - (a_1y_1 + \dots + a_ny_n) = (s + a_1s_1 + \dots + a_ns_n).$$

Como el lado izquierdo de esta igualdad está en T_2 y el derecho en S , de que $S \cap T_2 = 0$ podemos deducir que, de hecho, ambos lados son nulos y, en particular, que $y = a_1y_1 + \dots + a_ny_n$. Esto nos dice que el conjunto $\{y_1, \dots, y_n\}$ genera a T_2 : así, este subespacio de V tiene dimensión finita y $\dim T_2 \leq n = \dim T_1$.

Ahora que sabemos que T_2 también tiene dimensión finita podemos repetir el razonamiento que acabamos de hacer pero con los roles de T_1 y de T_2 intercambiados: concluiremos que $\dim T_1 \leq \dim T_2$ y, en definitiva, que T_1 y T_2 tienen la misma dimensión. \square

1.10.7. Si V es un espacio vectorial, decimos que un subespacio S de V tiene **codimensión finita en V** si posee un complemento en V de dimensión finita y en caso contrario que tiene **codimensión infinita en V** . Observemos que la Proposición 1.10.6 nos dice que si S tiene codimensión finita entonces todos sus complementos tienen dimensión finita y, de hecho, que todos tienen la misma dimensión: llamamos **codimensión de S en V** a la dimensión común de sus complementos.

1.10.8. Podemos dar un criterio útil para reconocer los subespacios de codimensión finita:

Proposición. *Sea V un espacio vectorial. Un subespacio S de V tiene codimensión finita en V si y solo si existe un subespacio T de V de dimensión finita tal que $S + T = V$.*

Demostración. Sea S un subespacio de V . Es evidente que si S tiene codimensión finita se satisface la condición del enunciado, así que esta es necesaria. Veamos que es suficiente. Supongamos que existe un subespacio T de V dimensión finita tal que $S + T = V$. La intersección $S \cap T$ es un subespacio del espacio de dimensión finita T , así que, de acuerdo a la Proposición 1.10.9, posee un complemento en T : existe entonces un subespacio U de T , necesariamente de dimensión finita, tal que $T = S \cap T \oplus U$. Observemos que

$$V = S + T = S + S \cap T + U = S + U,$$

ya que $S \cap T \subseteq S$. Por otro lado, como $(S \cap T) \cap U = 0$, tenemos que

$$S \cap U = (S \cap U) \cap U \subseteq (S \cap T) \cap U = 0.$$

Así, tenemos que $V = S \oplus U$ y, por lo tanto, que U es un complemento de dimensión finita de S en V , de manera que S tiene codimensión finita en V . \square

1.10.9. Si la dimensión de un subespacio es una medida de su tamaño y su codimensión una medida de cuánto le falta para ser el espacio total, entonces el siguiente resultado es bien natural:

Proposición. *Sea V un espacio vectorial. Si V tiene dimensión finita, entonces todo subespacio S de V tiene codimensión finita y*

$$\dim S + \text{codim}_V S = \dim V.$$

Demostración. Supongamos que V tiene dimensión finita y que S es un subespacio de V . De acuerdo al Corolario 1.10.5, el subespacio S tiene un complemento, esto es, hay un subespacio T

de V tal que $V = S \oplus T$. La Proposición 1.9.5 nos dice ahora que

$$\dim V = \dim S \oplus T = \dim S + \dim T = \dim S + \text{codim}_V S,$$

que es lo que afirma la proposición. \square

1.10.10. Proposición. *Sea V un espacio vectorial. Si S y S' son subespacios de V tales que $S \subseteq S'$, entonces las siguientes afirmaciones son equivalentes:*

- (a) *S tiene codimensión finita en V .*
- (b) *S tiene codimensión finita en S' y S' tiene codimensión finita en V .*

Cuando estas afirmaciones valen, entonces

$$\text{codim}_V S = \text{codim}_{S'} S + \text{codim}_V S'.$$

Demostración. Sean S y S' dos subespacios de V tales que $S \subseteq S'$.

Supongamos, para empezar, que S tiene codimensión finita en V y sea T un complemento de dimensión finita para S en V , de manera que $V = S \oplus T$. Como S y $S' \cap T$ son subespacios de S' , tenemos que $S + S' \cap T \subseteq S'$. Por otro lado, si $x \in S'$, entonces existe $y \in S$ y $t \in T$ tal es que $x = y + t$, ya que $V = S \oplus T$, y $t = x - y \in S'$ porque $x \in S'$ e $y \in S \subseteq S'$: vemos así que $x = y + t \in S + S' \cap T$ y, por lo tanto, que $S' \subseteq S + S' \cap T$. Hemos mostrado que $S' = S + S' \cap T$. Más aún, como $S \cap (S' \cap T) = S \cap T = 0$, tenemos que $S' = S \oplus S' \cap T$. En particular, como $S' \cap T$ tiene dimensión finita porque es un subespacio de T , vemos que S tiene codimensión finita en S' y que $\text{codim}_{S'} S = \dim S \cap T$.

Por otro lado, es $S' + T \supseteq S + T = V$ y T tiene dimensión finita, así que S' tiene codimensión finita en V . Esto prueba la implicación $(a) \Rightarrow (b)$. Además, si U es un complemento para $S' \cap T$ en T , que existe porque T tiene dimensión finita, tenemos entonces que $S' \cap T \oplus U = T$ y que

$$V = S \oplus T = S \oplus S' \cap T \oplus U = S' \oplus U,$$

así que $\text{codim}_V S' = \dim U$. Juntando todo, vemos que

$$\text{codim}_V S = \dim T = \dim S' \cap T \oplus U = \dim S' \cap T + \dim U = \text{codim}_{S'} S + \text{codim}_V S'.$$

Supongamos ahora que vale la condición (b) del enunciado, de manera que existe subespacios T de S' y U de V tales que $S' = S + T$ y $V = S' + U$. Entonces $T + U$ tiene dimensión finita y

$$S + (T + U) = (S + T) + U = S' + U = V,$$

así que S tiene codimensión finita en V , esto es, vale la condición (a) del enunciado. \square

§11. Coordenadas y cambio de base

1.11.1. Si V es un espacio vectorial de dimensión finita y $n = \dim V$, una *ordenada* de V es una n -upla $\mathcal{B} = (x_1, \dots, x_n)$ de n elementos de V tal que el conjunto $\{x_1, \dots, x_n\}$ es una base de V . Así, dar una base ordenada de V es lo mismo que dar una base de V y un orden (total) entre sus elementos. Por ejemplo, (e_1, e_2) y (e_2, e_1) son dos bases ordenadas distintas de \mathbb{k}^2 , que corresponden a los dos órdenes distintos que se puede dar a la base estándar $\{e_1, e_2\}$ de ese espacio vectorial. De manera similar, la base estándar $\{e_1, e_2, e_3\}$ de \mathbb{k}^3 puede ordenarse de seis maneras distintas, dando lugar a seis bases ordenadas distintas de este espacio.

1.11.2. Sea V un espacio vectorial de dimensión finita, sea $n = \dim V$, y sea $\mathcal{B} = (x_1, \dots, x_n)$ una base ordenada de V . Para cada $x \in V$ existen escalares $a_1, \dots, a_n \in \mathbb{k}$ tales que $x = a_1x_1 + \dots + a_nx_n$, ya que los elementos de \mathcal{B} generan a V , y estos escalares están únicamente determinados por x , ya que los elementos de \mathcal{B} son linealmente independientes: llamamos al vector $(a_1, \dots, a_n)^t \in \mathbb{k}^n$ el *vector de coordenadas* de x con respecto a la base \mathcal{B} y lo denotamos $[x]_{\mathcal{B}}$.

La observación más importante a hacer sobre esta construcción es la siguiente:

Proposición. *Sea V un espacio vectorial de dimensión finita, sea $n = \dim V$, y sea \mathcal{B} una base ordenada de V . La función $c_{\mathcal{B}} : x \in V \mapsto [x]_{\mathcal{B}} \in \mathbb{k}^n$ es una biyección.*

Demostración. Consideremos la función $d : \mathbb{k}^n \rightarrow V$ que en cada $a = (a_1, \dots, a_n) \in \mathbb{k}^n$ toma el valor $d(a) = a_1x_1 + \dots + a_nx_n$. Para probar la proposición es suficiente que mostremos que $c_{\mathcal{B}}$ y d son funciones mutuamente inversas.

- Si $x \in V$ y $c_{\mathcal{B}}(x) = (a_1, \dots, a_n)$, entonces de la definición misma de la función $c_{\mathcal{B}}$ se sigue que $d(c_{\mathcal{B}}(x)) = a_1x_1 + \dots + a_nx_n = x$.
- Por otro lado, si $a = (a_1, \dots, a_n) \in \mathbb{k}^n$ y x es el vector $d(a) = a_1x_1 + \dots + a_nx_n$ de V , entonces claramente $c_{\mathcal{B}}(d(a)) = c_{\mathcal{B}}(x) = a$.

Vemos así que $d \circ c_{\mathcal{B}} = \text{id}_V$ y que $c_{\mathcal{B}} \circ d = \text{id}_{\mathbb{k}^n}$, como queríamos. \square

1.11.3. Si \mathcal{B} es una base ordenada de un espacio vectorial V de dimensión finita, podemos ver al vector de coordenadas $[x]_{\mathcal{B}}$ de un elemento x de V con respecto a la base \mathcal{B} como un *nombre* para x que lo identifica completamente. Claro, un espacio vectorial tiene en general muchas bases ordenadas y, por lo tanto, a cada uno de sus elementos podemos asignarle muchos nombres, uno por cada una de esas bases ordenadas. Es importante describir qué relación hay entre todos estos nombres: ese es nuestro objetivo inmediato.

Sea n la dimensión del espacio V y sean $\mathcal{B} = (x_1, \dots, x_n)$ y $\mathcal{B}' = (y_1, \dots, y_n)$ dos bases ordenadas de V . Para cada $i \in \llbracket n \rrbracket$ sabemos que existen escalares $c_{1,i}, \dots, c_{n,i} \in \mathbb{k}$, únicamente determinados, tales que

$$x_i = c_{1,i}y_1 + \dots + c_{n,i}y_n,$$

y podemos entonces considerar la matriz

$$C(\mathcal{B}, \mathcal{B}') = \begin{pmatrix} c_{1,1} & \cdots & c_{1,n} \\ \vdots & \ddots & \vdots \\ c_{n,1} & \cdots & c_{n,n} \end{pmatrix} \in M_n(\mathbb{k}),$$

a la que llamamos **matriz de cambio de base de \mathcal{B} a \mathcal{B}'** . Las columnas de esta matriz son los vectores de coordenadas de los elementos de la base \mathcal{B} con respecto a la base \mathcal{B}' , en el orden dado.

La razón por la que nos interesamos en esta matriz es que permite expresar de manera sencilla las coordenadas de un vector cualquiera de un espacio vectorial con respecto de una base en términos de sus coordenadas con respecto a otra:

Proposición. *Sea V un espacio vectorial de dimensión finita y sean \mathcal{B} y \mathcal{B}' dos bases ordenadas de V . Para cada $x \in V$ se tiene que*

$$[x]_{\mathcal{B}'} = C(\mathcal{B}, \mathcal{B}') \cdot [x]_{\mathcal{B}}$$

y, de hecho, la matriz $C(\mathcal{B}, \mathcal{B}')$ es la única con esta propiedad.

Demostración. Sea $n = \dim V$ y supongamos que $\mathcal{B} = (x_1, \dots, x_n)$, que $\mathcal{B}' = (y_1, \dots, y_n)$ y que $C(\mathcal{B}, \mathcal{B}') = (c_{i,j})$, de manera que para cada $i \in \llbracket n \rrbracket$ se tiene que

$$x_i = c_{1,i}y_1 + \cdots + c_{n,i}y_n.$$

Sea $x \in V$ y supongamos que $[x]_{\mathcal{B}} = (a_1, \dots, a_n)$. Tenemos entonces que

$$\begin{aligned} x &= a_1x_1 + \cdots + a_nx_n \\ &= a_1(c_{1,1}y_1 + \cdots + c_{n,1}y_n) + \cdots + a_n(c_{1,n}y_1 + \cdots + c_{n,n}y_n) \\ &= (c_{1,1}a_1 + \cdots + c_{1,n}a_n)y_1 + \cdots + (c_{n,1}a_1 + \cdots + c_{n,n}a_n)y_n, \end{aligned}$$

y esto significa que para cada $i \in \llbracket n \rrbracket$ la coordenada i -ésima de x con respecto a la base \mathcal{B}' es precisamente la componente i -ésima del vector $C(\mathcal{B}, \mathcal{B}') \cdot [x]_{\mathcal{B}}$. Esto prueba la primera afirmación de la proposición.

Para ver la segunda, basta observar que si $A \in M_n(\mathbb{k})$ es una matriz tal que $[x]_{\mathcal{B}'} = A \cdot [x]_{\mathcal{B}}$ para todo $x \in V$, en particular se tiene que para cada $i \in \llbracket n \rrbracket$ es

$$A \cdot e_i = A \cdot [x_i]_{\mathcal{B}} = [x_i]_{\mathcal{B}'},$$

con e_i el i -ésimo vector de la base ordenada estándar de \mathbb{k}^n . Como $A \cdot e_i$ es la i -ésima columna de A , vemos de esta forma que cada columna de A es el vector de coordenadas del correspondiente vector de la base \mathcal{B} con respecto a la base \mathcal{B}' : así, la matriz A es la matriz $C(\mathcal{B}, \mathcal{B}')$. \square

1.11.4. Proposición. Sea V un espacio vectorial de dimensión finita y sea $n = \dim V$.

- (i) Si \mathcal{B} es una base ordenada de V , entonces $C(\mathcal{B}, \mathcal{B}) = I_n$, la matriz identidad de $M_n(\mathbb{k})$.
- (ii) Si $\mathcal{B}, \mathcal{B}'$ y \mathcal{B}'' son bases ordenadas de V , entonces

$$C(\mathcal{B}, \mathcal{B}'') = C(\mathcal{B}', \mathcal{B}'') \cdot C(\mathcal{B}, \mathcal{B}').$$

- (iii) Si \mathcal{B} y \mathcal{B}' son bases ordenadas de V , entonces la matriz $C(\mathcal{B}, \mathcal{B}')$ es inversible y

$$C(\mathcal{B}, \mathcal{B}')^{-1} = C(\mathcal{B}', \mathcal{B}).$$

Demostración. La primera afirmación es evidente. Veamos la segunda: sean $\mathcal{B} = (x_1, \dots, x_n)$, $\mathcal{B}' = (y_1, \dots, y_n)$ y $\mathcal{B}'' = (z_1, \dots, z_n)$ tres bases ordenadas de V , y sean $C(\mathcal{B}, \mathcal{B}') = (a_{i,j})$ y $C(\mathcal{B}', \mathcal{B}'') = (b_{i,j})$ las matrices de cambio de base de \mathcal{B} a \mathcal{B}' y de \mathcal{B}' a \mathcal{B}'' , respectivamente. Esto significa que para cada $i \in \llbracket n \rrbracket$ es

$$x_i = a_{1,i}y_1 + \cdots + a_{n,i}y_n$$

y

$$y_i = b_{1,i}z_1 + \cdots + b_{n,i}z_n.$$

Usando esta segunda igualdad en la primera vemos que para cada $i \in \llbracket n \rrbracket$ se tiene, de hecho, que

$$\begin{aligned} x_i &= a_{1,i}y_1 + \cdots + a_{n,i}y_n \\ &= a_{1,i}(b_{1,1}z_1 + \cdots + b_{n,1}z_n) + \cdots + a_{n,i}(b_{1,n}z_1 + \cdots + b_{n,n}z_n) \\ &= (b_{1,1}a_{1,i} + \cdots + b_{1,n}a_{n,i})z_1 + \cdots + (b_{n,1}a_{1,i} + \cdots + b_{n,n}a_{n,i})z_n. \end{aligned}$$

Así, el vector de coordenadas de x_i con respecto a la base \mathcal{B}'' es la i -ésima columna de la matriz $C(\mathcal{B}', \mathcal{B}'') \cdot C(\mathcal{B}, \mathcal{B}')$ y, por lo tanto, esta matriz coincide con $C(\mathcal{B}, \mathcal{B}'')$, como afirma la parte (ii) de la proposición.

Para ver la última parte, sean \mathcal{B} y \mathcal{B}' dos bases ordenadas de V . De acuerdo a lo que ya probamos, es

$$C(\mathcal{B}', \mathcal{B}) \cdot C(\mathcal{B}, \mathcal{B}') = C(\mathcal{B}, \mathcal{B}) = I_n$$

y

$$C(\mathcal{B}, \mathcal{B}') \cdot C(\mathcal{B}', \mathcal{B}) = C(\mathcal{B}', \mathcal{B}') = I_n,$$

y entonces es claro que las matrices $C(\mathcal{B}, \mathcal{B}')$ y $C(\mathcal{B}', \mathcal{B})$ son mutuamente inversas. En particular, la matriz $C(\mathcal{B}, \mathcal{B}')$ es inversible y tenemos que $C(\mathcal{B}, \mathcal{B}')^{-1} = C(\mathcal{B}', \mathcal{B})$. La proposición queda así probada. \square

1.11.5. De acuerdo a la Proposición 1.11.4(iii), la matriz de cambio de base entre dos bases ordenadas de un espacio vectorial de dimensión finita es una matriz inversible. Es útil muchas veces saber que, de hecho, toda matriz inversible aparece de esta forma, en el siguiente sentido:

Proposición. Sea V un espacio vectorial de dimensión finita y sea $n = \dim V$. Si \mathcal{B} es una base ordenada de V y $C \in GL_n(\mathbb{k})$ es una matriz inversible, entonces existe una y solo una base ordenada \mathcal{B}' de V tal que $C(\mathcal{B}, \mathcal{B}') = C$.

Demostración. Sea $\mathcal{B} = (x_1, \dots, x_n)$ una base ordenada de V y sea $C = (c_{i,j})$ una matriz inversible de tamaño $n \times n$. Escribamos $D = (d_{i,j})$ a la matriz inversa de C , de manera que, en particular, se tiene que $C \cdot D = I_n$. Esta igualdad matricial nos dice que para cada elección de i y j en $\llbracket n \rrbracket$ se tiene que

$$\sum_{k=1}^n c_{i,k} d_{k,j} = \delta_{i,j}. \quad (20)$$

Para cada $i \in \llbracket n \rrbracket$ consideremos el vector

$$y_i = d_{1,i}x_1 + \dots + d_{n,i}x_n. \quad (21)$$

Afirmamos que los n vectores y_1, \dots, y_n que así obtenemos son linealmente independientes. Para probarlo, supongamos que tenemos una relación de dependencia lineal

$$\alpha_1 y_1 + \dots + \alpha_n y_n = 0 \quad (22)$$

entre ellos, con $\alpha_1, \dots, \alpha_n \in \mathbb{k}$. En vista de la forma en que definimos a los vectores y_i , esto significa que

$$\begin{aligned} 0 &= \alpha_1 y_1 + \dots + \alpha_n y_n \\ &= \alpha_1(d_{1,1}x_1 + \dots + d_{n,1}x_n) + \dots + \alpha_n(d_{1,n}x_1 + \dots + d_{n,n}x_n) \\ &= (d_{1,1}\alpha_1 + \dots + d_{1,n}\alpha_n)x_1 + \dots + (d_{n,1}\alpha_1 + \dots + d_{n,n}\alpha_n)x_n. \end{aligned}$$

Como \mathcal{B} es una base, esta igualdad implica que cada uno de los coeficientes que aparecen en esta última expresión es nulo, esto es, que

$$d_{i,1}\alpha_1 + \dots + d_{i,n}\alpha_n = 0 \quad \text{para cada } i \in \llbracket n \rrbracket.$$

Usando esto vemos que si $k \in \llbracket n \rrbracket$ es

$$0 = \sum_{i=1}^n c_{k,i}(d_{i,1}\alpha_1 + \dots + d_{i,n}\alpha_n) = \left(\sum_{i=1}^n c_{k,i}d_{i,1} \right) \alpha_1 + \dots + \left(\sum_{i=1}^n c_{k,i}d_{i,n} \right) \alpha_n$$

y, recordando las igualdades de (20), esto es

$$= \delta_{k,1}\alpha_1 + \dots + \delta_{k,n}\alpha_n = \alpha_k.$$

Concluimos de esta forma que cada uno de los coeficientes que aparecen en (22) es nulo y, en definitiva, que los vectores y_1, \dots, y_n son linealmente independientes, como dijimos. Como son n y n es la dimensión de V , la Proposición 1.7.9 nos dice que $\mathcal{B}' = (y_1, \dots, y_n)$ es una base ordenada

de V . Más aún, y en vista de la igualdad (21) que usamos para definir los vectores de esta base, es claro que la matriz de cambio de base $C(\mathcal{B}', \mathcal{B})$ es precisamente la matriz D y, entonces, que

$$C(\mathcal{B}, \mathcal{B}') = C(\mathcal{B}', \mathcal{B})^{-1} = C.$$

Esto prueba la afirmación de existencia de la proposición. Veamos ahora la de unicidad.

Supongamos que $\mathcal{B}'' = (z_1, \dots, z_n)$ es otra base ordenada de V tal que $C(\mathcal{B}, \mathcal{B}'') = C$. Se tiene entonces que $C(\mathcal{B}'', \mathcal{B}) = C^{-1} = D$ y, por lo tanto, para cada $i \in \llbracket n \rrbracket$ el vector $[z_i]_{\mathcal{B}}$ es la i -ésima columna de la matriz D , así que coincide con el vector $[y_i]_{\mathcal{B}}$: como las coordenadas de un vector lo determinan completamente, esto implica que, de hecho, $z_i = y_i$. Vemos así que $\mathcal{B}'' = \mathcal{B}'$, que es lo que teníamos que probar. \square

Algunos resultados de conteo

1.11.6. Cuando nuestro cuerpo de base \mathbb{k} es finito tiene sentido e interés plantearse todo tipo de problemas de conteo. En la base de todas las respuestas a estos problemas está la siguiente consecuencia inmediata de la Proposición 1.11.2.

Corolario. *Supongamos que \mathbb{k} es un cuerpo finito de q elementos. Si V es un espacio vectorial de dimensión finita, entonces V tiene exactamente $q^{\dim V}$ elementos.*

Demostración. Si V es un espacio vectorial de dimensión finita y \mathcal{B} es una base ordenada, entonces la proposición nos da una biyección $c_{\mathcal{B}} : V \rightarrow \mathbb{k}^n$ con $n = \dim V$. Como el conjunto \mathbb{k}^n tiene q^n elementos, el conjunto V tiene también ese cardinal. \square

1.11.7. Este corolario tiene una aplicación sencilla e importante al estudio de los cuerpos finitos.

Proposición. *Si K es un cuerpo finito, entonces existen un número primo p y un entero positivo r tales que el cardinal de K es p^r .* \square

Esto prueba parte de nuestras afirmaciones del Ejemplo 1.1.3(e).

Demostración. Sea K un cuerpo finito. De acuerdo a la Proposición 1.1.8 sabemos que la característica de K es un número primo p y en vista de la Proposición 1.1.10 el cuerpo K posee un subcuerpo P con exactamente p elementos. Como vimos en el Ejemplo 1.2.5(c), podemos considerar al cuerpo K como un espacio vectorial sobre P . Más aún, K es un espacio vectorial finitamente generado: en efecto, está generado por él mismo, que es un conjunto finito. Se sigue de esto, entonces, que K es un espacio vectorial de dimensión finita sobre P . Más aún, como K tiene más que un elemento, su dimensión $r := \dim K$ es un número positivo. El Corolario 1.11.6 nos dice entonces, como el cuerpo P tiene p elementos, que K tiene p^r elementos. \square

1.11.8. Una segunda aplicación es dar una mejora de la segunda parte de la Proposición 1.3.7.

Proposición. Supongamos que \mathbb{k} es un cuerpo finito de q elementos. Un espacio vectorial de dimensión finita con al menos 3 subespacios es unión de $q + 1$ subespacios propios y no de menos.

Un espacio vectorial con menos que 3 subespacios tiene dimensión como mucho 1, así que no es unión de subespacios propios.

Demostración. Sea V un espacio vectorial de dimensión finita con al menos 3 subespacios y sea $n := \dim V$ su dimensión. La Proposición 1.3.7 nos dice que V no es unión de menos que q subespacios propios. Supongamos que S_1, \dots, S_q son subespacios propios de V . Como estos subespacios son propios, entonces $\dim S_i \leq n - 1$ para cada $i \in [q]$ y, por lo tanto, $\#S_i \leq q^{n-1}$. Usando esto, el hecho de que todos tienen a vector nulo en común y, finalmente, que $q \geq 2$, vemos que

$$\#\left(\bigcup_{i=1}^q S_i\right) \leq \sum_{i=1}^q \#S_i - 1 = q \cdot q^{n-1} - 1 < q^n = \#V$$

y, en particular, que la unión $\bigcup_{i=1}^q S_i$ está estrictamente contenida en V . Esto muestra que V no es unión de q subespacios propios.

Sea ahora $\mathcal{B} = (x_1, \dots, x_n)$ una base ordenada de V y pongamos

$$S_\infty := \langle x_2, x_3, \dots, x_n \rangle.$$

y, para cada $\lambda \in \mathbb{k}$,

$$S_\lambda := \langle x_1 + \lambda x_2, x_3, \dots, x_n \rangle.$$

Todos estos $q + 1$ subespacios de V son propios: en efecto, cada uno de ellos está generado por menos que n elementos, así que tienen dimensión estrictamente menor que la de V . Por otro lado, tenemos que

$$V = S_\infty \cup \bigcup_{\lambda \in \mathbb{k}} S_\lambda.$$

En efecto, si $x \in V$, entonces hay escalares $\alpha_1, \dots, \alpha_n \in \mathbb{k}$ tales que $x = \alpha_1 x_1 + \dots + \alpha_n x_n$, y o bien $\alpha_1 = 0$, y en ese caso $x \in S_\infty$, o bien $\alpha_1 \neq 0$, y en ese caso $x \in S_{\alpha_2/\alpha_1}$. Vemos así que V es unión de $q + 1$ subespacios propios. \square

1.11.9. Otra cosa que podemos contar cuando nuestro cuerpo de base es finito es conjuntos linealmente independientes:

Proposición. Supongamos que \mathbb{k} es un cuerpo finito de q elementos. Si V es un espacio vectorial de dimensión finita n y $k \in [0, n]$, entonces el número de k -uplas ordenadas (x_1, \dots, x_k) de vectores de V linealmente independientes es

$$q^{k(k-1)/2} \prod_{i=0}^{k-1} (q^{n-i} - 1). \tag{23}$$

En particular, el número de bases ordenadas de V es

$$q^{n(n-1)/2} \prod_{i=0}^{n-1} (q^{n-i} - 1). \quad (24)$$

Demostración. Es suficiente que probemos la primera afirmación, ya que la segunda se obtiene de ella eligiendo $k = n$. Para cada $k \in \mathbb{N}_0$ escribamos B_n al conjunto de todas las k -uplas ordenadas de vectores de V linealmente independientes y b_n a su cardinal. El conjunto B_0 contiene un único elemento, la 0-upla $(\)$, y, por lo tanto, $b_0 = 1$: esto coincide con el valor del producto (23) cuando $k = 0$.

Por otro lado, supongamos que $l \in \llbracket 0, n-1 \rrbracket$ y que el número b_l está dado por el producto (23) con $k = l$. Es claro que hay una función $\pi : B_{l+1} \rightarrow B_l$ que en cada $(x_1, \dots, x_{l+1}) \in B_{l+1}$ toma el valor

$$\pi((x_1, \dots, x_{l+1})) = (x_1, \dots, x_l).$$

Observemos que si $\xi = (x_1, \dots, x_l)$ es un elemento de B_l , entonces los elementos de B_{l+1} cuya imagen por π es ξ son precisamente las $(l+1)$ -uplas de la forma (x_1, \dots, x_l, y) con $y \in V \setminus \langle x_1, \dots, x_l \rangle$, y el número de estas es

$$\#(V \setminus \langle x_1, \dots, x_l \rangle) = \#V - \#\langle x_1, \dots, x_l \rangle = q^n - q^l = q^l(q^{n-l} - 1),$$

ya que $\dim V = n$ y $\dim \langle x_1, \dots, x_l \rangle = l$. Tenemos entonces que

$$b_{l+1} = \#B_{l+1} = \# \left(\bigcup_{\xi \in B_l} \pi^{-1}(\xi) \right) = \sum_{\xi \in B_l} \#(\pi^{-1}(\xi)),$$

ya que la unión es disjunta, y como todos los sumandos son iguales a $q^l(q^{n-l} - 1)$, esto es

$$= b_l \cdot (q^n - q^l) = q^{l(l-1)/2} \prod_{i=0}^{l-1} (q^{n-i} - 1) \cdot q^l(q^{n-l} - 1) = q^{(l+1)l/2} \prod_{i=0}^l (q^{n-i} - 1),$$

que es precisamente el producto (23) con $k = l + 1$. □

1.11.10. Si q es un entero distinto de 1, para cada $n \in \mathbb{N}_0$ escribimos

$$[n]_q := 1 + \cdots + q^{n-1} = \frac{q^n - 1}{q - 1}$$

y

$$[n]_q! := [1]_q \cdots [n]_q.$$

Con esta notación, si $n, k \in \mathbb{N}_0$ y $0 \leq k \leq n$, podemos escribir al número (23) de k -uplas ordenadas de vectores de un espacio vectorial de dimensión n sobre un cuerpo de q elementos en la forma

$$q^{k(k-1)/2} (q-1)^k \frac{[n]_q!}{[n-k]_q!}$$

y el número (24) de bases ordenadas de ese espacio en la forma

$$q^{n(n-1)/2}(q-1)^n[n]_q!$$

Combinando estas dos enumeraciones, obtenemos fácilmente el número de subespacios de una dimensión fija:

Corolario. *Supongamos que \mathbb{k} es un cuerpo finito de q elementos. Si V es un espacio vectorial de dimensión finita n y $k \in \mathbb{N}_0$, entonces V posee exactamente*

$$\binom{n}{k}_q := \frac{[n]_q!}{[k]_q![n-k]_q!}$$

subespacios de dimensión k .

Demostración. Sea V un espacio vectorial de dimensión n y sea $k \in \mathbb{N}_0$. Sabemos que hay

$$q^{k(k-1)/2}(q-1)^k \frac{[n]_q!}{[n-k]_q!} \quad (25)$$

k -uplas ordenadas de vectores de V linealmente independientes, y cada una de ellas genera un subespacio de dimensión k de V . Por otro lado, cada subespacio de dimensión k de V posee exactamente

$$q^{k(k-1)/2}(q-1)^k [k]_q! \quad (26)$$

bases ordenadas. Esto implica, claro, que el número de subespacios de V de dimensión k es el cociente entre el número (25) y el número (26), que es el que aparece en el enunciado del corolario. \square

1.11.11. Es evidente la de definición de estos números que si $n, k \in \mathbb{N}_0$ y $0 \leq k \leq n$, entonces

$$\binom{n}{k}_q = \binom{n}{n-k}_q, \quad \binom{n}{0}_q = \binom{n}{n}_q = 1. \quad (27)$$

Menos evidentes son las siguientes identidades:

Proposición. *Sea p un número primo, $r \in \mathbb{N}$ y $q := p^r$. Si $n, k \in \mathbb{N}$ y $0 < k < n$, entonces*

$$\binom{n}{k}_q = \binom{n-1}{k-1}_q + q^{n-k} \binom{n-1}{k}_q = q^k \binom{n-1}{k-1}_q + \binom{n-1}{k}_q.$$

Las dos igualdades de esta proposición pueden probarse sin ninguna dificultad usando directamente la definición numérica de los coeficientes q -binomiales. Dejamos eso como ejercicio para el lector, prefiriendo aquí dar una demostración vía biyecciones.

Demostración. Fijemos un cuerpo \mathbb{k} finito de q elementos. Sean $n, k \in \mathbb{N}$ tales que $0 < k < n$ y consideremos el espacio vectorial $V = \mathbb{k}^n$ y su subespacio $H = \{(x_1, \dots, x_n)^t \in V : x_n = 0\}$. Para

cada $l \in \mathbb{N}_0$ escribamos $\mathcal{S}_l(V)$ y $\mathcal{S}_l(H)$ el conjunto de todos los subespacios de dimensión l de V y de H , respectivamente. Si $S \in \mathcal{S}_k(V)$, entonces o bien $S \subseteq H$ o bien $S \not\subseteq H$, por supuesto, así que si escribimos

$$\mathcal{S}'_k(V) \coloneqq \{S \in \mathcal{S}_k(V) : S \subseteq H\}, \quad \mathcal{S}''_k(V) \coloneqq \{S \in \mathcal{S}_k(V) : S \not\subseteq H\},$$

tenemos que el conjunto $\mathcal{S}_k(V)$ se descompone como una unión disjunta,

$$\mathcal{S}_k(V) = \mathcal{S}'_k(V) \sqcup \mathcal{S}''_k(V). \quad (28)$$

Claramente $\mathcal{S}'_k(V) = \mathcal{S}_k(H)$, así que

$$\#(\mathcal{S}'_k(V)) = \binom{n-1}{k}_q. \quad (29)$$

Queremos ahora determinar el cardinal del conjunto $\mathcal{S}''_k(V)$. Sea

$$\mathcal{V} \coloneqq \{(x_1, \dots, x_n)^t \in V : x_n = 1\},$$

que claramente tiene q^{n-1} elementos. Si $v \in \mathcal{V}$ y $T \in \mathcal{S}_{k-1}(H)$, entonces $v \notin T$ y, por lo tanto, el subespacio $\langle v \rangle + T$ de V tiene dimensión k y no está contenido en H . Esto nos dice que hay una función

$$\Phi : (v, T) \in \mathcal{V} \times \mathcal{S}_{k-1}(H) \mapsto \langle v \rangle + T \in \mathcal{S}''_k(V).$$

Sea $S \in \mathcal{S}''_k(V)$. Mostremos que

$$\Phi^{-1}(S) = \{(u, S \cap H) : u \in S \cap \mathcal{V}\}$$

probando las dos inclusiones:

- Si $u \in S \cap \mathcal{V}$, entonces $u \notin S \cap H$, así que $\Phi(u, S \cap H) = \langle u \rangle + S \cap H = S$.
- Sean $u \in \mathcal{V}$ y $T' \in \mathcal{S}_{k-1}(H)$ tales que $\Phi(u, T') = \langle u \rangle + T' = S$. Claramente $u \in S \cap V$ y es $(\langle u \rangle + T') \cap H = S \cap H$ y la primera de estas intersecciones es igual a T' : en efecto, si $\lambda \in \mathbb{k}$ y $t \in T'$ son tales que $\lambda u + t \in H$, entonces la n -ésima coordenada del vector $\lambda u + t$, que es precisamente λ , es nula. Vemos así que $T' = S \cap H$.

Ahora, si v es un elemento cualquiera de $S \cap \mathcal{V}$ entonces es inmediato verificar que

$$S \cap \mathcal{V} = \{v + h : h \in S \cap H\}$$

así que, como $S \cap H$ es un espacio de dimensión $k-1$,

$$\#(\Phi^{-1}(S)) = \#(S \cap \mathcal{V}) = \#(S \cap H) = q^{k-1}.$$

Se sigue de esto que

$$q^{n-1} \cdot \binom{n-1}{k-1}_q = \#(\mathcal{V} \times \mathcal{S}_{k-1}(H)) = \sum_{S \in \mathcal{S}''_k(V)} \#(\Phi^{-1}(S)) = q^{k-1} \cdot \#(\mathcal{S}''_k(V)),$$

de manera que

$$\#(\mathcal{S}_k''(V)) = q^{n-k} \binom{n-1}{k-1}_q$$

y, por lo tanto, en vista de (28) y (29), que vale la primera igualdad del enunciado. La segunda se obtiene de ella usando la primera identidad de (27). \square

§12. Una digresión: el principio de inclusión–exclusión para subespacios

1.12.1. Arriba, al presentar la Proposición 1.8.4, comparamos su conclusión con el llamado *Principio de Inclusión-Exclusión*, que cuando A y B son dos conjuntos finitos nos dice que

$$\#(A \cup B) + \#(A \cap B) = \#A + \#B.$$

Esta igualdad tiene una extensión natural al caso en el que tenemos un número finito arbitrario de conjuntos finitos. Por ejemplo, cuando $n = 3$ o $n = 4$ tenemos que

$$\begin{aligned} \#(A_1 \cup A_2 \cup A_3) &= \#A_1 + \#A_2 + \#A_3 - \#(A_1 \cap A_2) - \#(A_1 \cap A_3) - \#(A_2 \cap A_3) \\ &\quad + \#(A_1 \cap A_2 \cap A_3) \end{aligned} \tag{30}$$

y que

$$\begin{aligned} \#(A_1 \cup A_2 \cup A_3 \cup A_4) &= \#A_1 + \#A_2 + \#A_3 + \#A_4 - \#(A_1 \cap A_2) - \#(A_1 \cap A_3) \\ &\quad - \#(A_1 \cap A_4) - \#(A_2 \cap A_3) - \#(A_2 \cap A_4) - \#(A_3 \cap A_4) \\ &\quad + \#(A_1 \cap A_2 \cap A_3) + \#(A_1 \cap A_2 \cap A_4) + \#(A_1 \cap A_3 \cap A_4) \\ &\quad + \#(A_2 \cap A_3 \cap A_4) - \#(A_1 \cap A_2 \cap A_3 \cap A_4) \end{aligned} \tag{31}$$

Uno puede preguntarse si fórmulas análogas valen para subespacios de un espacio vectorial y sus dimensiones: la respuesta es que, en general, no valen. Por ejemplo, si $V = \mathbb{k}^2$ y $U_1 = \langle e_1 \rangle$, $U_2 = \langle e_2 \rangle$ y $U_3 = \langle e_1 + e_2 \rangle$, entonces $U_1 + U_2 + U_3 = \mathbb{k}^2$, la dimensión de cualquiera de los tres subespacios es 1, y la intersección de dos o mas de los tres subespacios en el subespacio nulo de \mathbb{k}^2 , así que

$$\begin{aligned} \dim(U_1 + U_2 + U_3) &= 2 \neq 3 = \dim U_1 + \dim U_2 + \dim U_3 - \dim U_1 \cap U_2 - \dim U_1 \cap U_3 \\ &\quad - \dim U_2 \cap U_3 + \dim U_1 \cap U_2 \cap U_3. \end{aligned}$$

Si uno repasa la demostración de las igualdades (30) y (31) observa que depende del hecho de que la intersección de conjuntos se distribuye sobre uniones, esto es, de que si A_1, \dots, A_n y B son conjuntos, entonces

$$(A_1 \cup \dots \cup A_n) \cap B = (A_1 \cap B) \cup \dots \cup (A_n \cap B).$$

El análogo de esta igualdad para subespacios de un espacio vectorial es falso. Si U_1 , U_2 y U_3 son los subespacios de \mathbb{k}^2 que mencionamos arriba, entonces se tiene por ejemplo que

$$(U_1 + U_2) \cap U_3 = U_3 \neq 0 = U_1 \cap U_3 + U_2 \cap U_3.$$

En esta sección nos proponemos mostrar que esta posible falla de distributividad de la intersección de espacios vectoriales por sobre sumas es el único obstáculo para que valga un análogo de la Proposición 1.8.4 para *tres* subespacios.

CONVENCIÓN

En esta sección asumiremos que todos los espacios vectoriales con los que trabajamos tienen dimensión finita.

1.12.2. Empecemos observando una cosa que sí vale siempre: la llamada *Ley Modular*, originalmente identificada por **Richard Dedekind** (1831–1916), que es una forma débil de la distributividad para subespacios que enunciamos en la segunda parte de la siguiente proposición.

Proposición. *Sea V un espacio vectorial y sean S , T y U tres subespacios de V .*

- (i) *Se tiene que $(S + U) \cap T \supseteq S \cap T + U \cap T$.*
- (ii) *Si $S \subseteq T$, entonces $(S + U) \cap T = S \cap T + U \cap T$.*

Notemos que en la segunda parte podríamos haber escrito S en lugar $S \cap T$ a la derecha del signo $=$, ya que ahí es $S \subseteq T$.

Demostración. (i) Como $S + U$ contiene a S y a U , es claro que $(S + U) \cap T$ contiene a $S \cap T$ y a $U \cap T$, y, como es un subespacio, que contiene a la suma $S \cap T + U \cap T$, como afirma el enunciado.

(ii) Supongamos que $S \subseteq T$ y sea $x \in (S + U) \cap T$, de manera que $x \in T$ y que existen $s \in S$ y $u \in U$ tales que $x = s + u$. Como $T \ni x - s = u \in U$, tenemos que $x - s \in U \cap T$ y, por lo tanto, que

$$x = s + (x - s) \in S + U \cap T = S \cap T + U \cap T.$$

Junto con la parte (i), esto nos dice que $(S + U) \cap T = S \cap T + U \cap T$. \square

1.12.3. Si S , T y U son subespacios de dimensión finita de un espacio vectorial V , llamamos **defecto** de la terna (S, T, U) al entero

$$\mathcal{D}(S, T, U) := \dim(S + T) \cap U - \dim(S \cap U + T \cap U).$$

De acuerdo a la primera parte de la Proposición 1.12.2 este número es un entero no negativo. La siguiente proposición nos dice que es precisamente la diferencia entre la dimensión de la suma $S + T + U$ y la expresión que nos daría un principio de inclusión-exclusión para subespacios:

Proposición. Sea V un espacio vectorial. Si S, T y U son tres subespacios de V , entonces

$$\begin{aligned}\dim(S + T + U) &= \dim S + \dim T + \dim U - \dim S \cap T - \dim S \cap U - \dim T \cap U \\ &\quad + \dim S \cap T \cap U - \mathcal{D}(S, T, U).\end{aligned}$$

Demostración. Es

$$\dim(S \cap U + T \cap U) = \dim S \cap U + \dim T \cap U - \dim S \cap T \cap U. \quad (32)$$

y, como $S + T + U = (S + T) + U$, es

$$\begin{aligned}\dim(S + T + U) &= \dim(S + T) + \dim U - \dim(S + T) \cap U \\ &= \dim S + \dim T - \dim S \cap T + \dim U - \dim(S + T) \cap U, \\ &= \dim S + \dim T - \dim S \cap T + \dim U - \dim(S \cap U + T \cap U) \\ &\quad + (\dim(S \cap U + T \cap U) - \dim(S + T) \cap U)\end{aligned}$$

y esto, de acuerdo a la igualdad (32), es

$$\begin{aligned}&= \dim S + \dim T - \dim S \cap T + \dim U - \dim S \cap U - \dim T \cap U \\ &\quad + \dim S \cap T \cap U + (\dim(S \cap U + T \cap U) - \dim(S + T) \cap U),\end{aligned}$$

como afirma el enunciado. \square

1.12.4. Usando la proposición que acabamos de probar podemos resolver completamente el problema que nos planteamos:

Corolario. Sea V un espacio vectorial y sean S, T y U tres subespacios de V . El defecto $\mathcal{D}(S, T, U)$ es independiente del orden de los tres subespacios, y se anula si y solamente si vale cualquiera de las cuatro igualdades

$$\begin{aligned}(S + T) \cap U &= S \cap U + T \cap U, \\ (T + U) \cap S &= T \cap S + U \cap S, \\ (U + S) \cap T &= U \cap T + S \cap T,\end{aligned}$$

y

$$\begin{aligned}\dim(S + T + U) &= \dim S + \dim T + \dim U - \dim S \cap T - \dim S \cap U - \dim T \cap U \\ &\quad + \dim S \cap T \cap U.\end{aligned}$$

Demostración. La Proposición 1.12.3 nos dice que el defecto $\mathcal{D}(S, T, U)$ es igual al entero

$$\begin{aligned}\dim(S + T + U) - \dim S - \dim T - \dim U + \dim S \cap T + \dim S \cap U + \dim T \cap U \\ - \dim S \cap T \cap U.\end{aligned}$$

Como este último no cambia si permutamos S , T y U , esto implica que el valor de $\mathcal{D}(S, T, U)$ tampoco lo hace. Como $S \cap U + T \cap U \subseteq (S + T) \cap U$, que $\mathcal{D}(S, T, U)$ se anule es equivalente a que valga la primera de las tres igualdades que aparecen en el enunciado, y la independencia de $\mathcal{D}(S, T, U)$ del orden de los tres subespacio implica entonces que también es equivalente a la segunda y a la tercera. Finalmente, que la cuarta igualdad sea equivalente a la anulación del defecto es consecuencia de la Proposición 1.12.3. \square

1.12.5. Decimos que una terna (S, T, U) de subespacios de un espacio vectorial es **distributiva** si el defecto $\mathcal{D}(S, T, U)$ es nulo. Esto ocurre si vale cualquiera de las tres primeras igualdades del corolario anterior — esas condiciones no son totalmente simétricas: el siguiente lema nos da una formulación alternativa que tiene la ventaja de involucrar a los tres subespacios de la misma forma.

Lema. *Sea V un espacio vectorial y sean S , T y U tres subespacios de V . La terna (S, T, U) es distributiva si y solamente si*

$$S \cap T + T \cap U + U \cap S = (S + T) \cap (T + U) \cap (U + S).$$

Demostración. Supongamos primero que la terna (S, T, U) es distributiva. Es inmediato que

$$S \cap T + T \cap U + U \cap S \subseteq (S + T) \cap (T + U) \cap (U + S),$$

ya que cada uno de los tres sumandos que aparecen a la izquierda del símbolo \subseteq están contenidos en cada uno de los tres intersecandos que aparecen a la derecha. Supongamos, para probar la contención recíproca, que x es un elemento de $(S + T) \cap (T + U) \cap (U + S)$. Esto nos dice que existen $s_1, s_2 \in S$, $t_1, t_2 \in T$ y $u_1, u_2 \in U$ tales que

$$x = s_1 + t_1 = t_2 + u_1 = u_2 + s_2.$$

En particular, tenemos que $s_1 = t_2 - t_1 + u_1 \in T + U$, así que $s_1 \in (T + U) \cap S$ y, de manera similar, que $t_1 \in (U + S) \cap T$. Se sigue de esto que

$$x = s_1 + t_1 \in (T + U) \cap S + (U + S) \cap T = S \cap T + T \cap U + U \cap S,$$

lo que prueba la inclusión que necesitábamos.

Supongamos ahora que vale la igualdad del enunciado y probemos que la terna (S, T, U) es distributiva. Como U está contenido en $U + S$ y en $U + T$, tenemos que

$$(S + T) \cap U = (S + T) \cap (U + S) \cap (U + T) \cap U$$

y la hipótesis nos dice que esto es

$$= (S \cap T + \underbrace{T \cap U + S \cap U}_{\text{}}) \cap U.$$

La suma marcada aquí está contenida en U , así que usando la Ley Modular 1.12.2(ii) vemos que

$$\begin{aligned}(S + T) \cap U &= S \cap T \cap U + (T \cap U + S \cap U) \cap U \\ &= S \cap T \cap U + T \cap U + S \cap U\end{aligned}$$

y como el primero de estos tres sumandos está contenido en el segundo, esto es, de hecho,

$$= T \cap U + U \cap S.$$

Esto nos dice que la terna (S, T, U) es distributiva, como queríamos. \square

1.12.6. Otra forma de caracterizar las ternas distributivas es vía tres igualdades que son «duales» a las del Corolario 1.12.4, en el sentido que se obtienen de ellas intercambiando intersecciones y sumas:

Proposición. *Sea V un espacio vectorial y sean S, T y U tres subespacios de V . La terna (S, T, U) es distributiva si y solamente si vale cualquiera de las siguientes tres igualdades:*

$$\begin{aligned}S \cap T + U &= (S + U) \cap (T + U), \\ T \cap U + S &= (T + S) \cap (U + S), \\ U \cap S + T &= (U + T) \cap (S + T).\end{aligned}$$

Demostración. Supongamos que la terna (S, T, U) es distributiva. Como $S + U$ contiene a U , la Ley Modular 1.12.2(ii) nos dice que

$$(S + U) \cap (T + U) = (S + U) \cap T + U,$$

y la distributividad que esto es

$$\begin{aligned}&= S \cap T + U \cap T + U \\ &= S \cap T + U.\end{aligned}$$

Vemos así que vale la primera de las igualdades del enunciado. Que valen las otras dos es consecuencia de que la distributividad de (S, T, U) no depende del orden de los tres subespacios.

Para probar la recíproca, supongamos que vale la primera de las igualdades del enunciado. La Ley Modular nos dice que

$$S \cap U + T \cap U = (S \cap U + T) \cap U,$$

y la hipótesis que esto es

$$\begin{aligned}&= (S + T) \cap (U + T) \cap U \\ &= (S + T) \cap U,\end{aligned}$$

de manera que la terna (S, T, U) es distributiva. \square

Es interesante notar que las dos partes de la demostración que acabamos de hacer son duales: una se obtiene de la otra intercambiando sumas e intersecciones.

1.12.7. La Proposición 1.8.4 afirma que

si V es un espacio vectorial y S y T son dos subespacios de V , entonces existe una base \mathcal{B} de $S + T$ tal que las intersecciones $\mathcal{B} \cap S$, $\mathcal{B} \cap T$ y $\mathcal{B} \cap (S \cap T)$ son bases de S , de T y de $S \cap T$, respectivamente.

Más aún, en la prueba de esa proposición vimos que

si \mathcal{B}_0 es una base de $S \cap T$ y \mathcal{B}_1 y \mathcal{B}_2 son conjuntos disjuntos de \mathcal{B}_0 tales que $\mathcal{B}_0 \cup \mathcal{B}_1$ y $\mathcal{B}_0 \cup \mathcal{B}_2$ son bases de S y de T , respectivamente, entonces $\mathcal{B}_1 \cap \mathcal{B}_2 = \emptyset$ y el conjunto $\mathcal{B} = \mathcal{B}_0 \cup \mathcal{B}_1 \cup \mathcal{B}_2$ es una base de $S + T$ tal que las intersecciones $\mathcal{B} \cap S$, $\mathcal{B} \cap T$ y $\mathcal{B} \cap (S \cap T)$ son bases de S , de T y de $S \cap T$, respectivamente.

El siguiente lema es una recíproca parcial de esto:

Lema. *Sea V un espacio vectorial, sean S y T subespacios de V y sea \mathcal{B} una base de $S + T$. Si los conjuntos $\mathcal{B}_S = \mathcal{B} \cap S$ y $\mathcal{B}_T = \mathcal{B} \cap T$ son bases de S y de T , entonces $\mathcal{B}_S \cap \mathcal{B}_T$, que coincide con $\mathcal{B} \cap (S \cap T)$, y $\mathcal{B}_S \cup \mathcal{B}_T$, que coincide con $\mathcal{B} \cap (S + T)$, son bases de $S \cap T$ y de $S + T$, respectivamente.*

Demostración. Supongamos que $\mathcal{B}_S = \mathcal{B} \cap S$ y $\mathcal{B}_T = \mathcal{B} \cap T$ son bases de S y de T . El conjunto $\mathcal{B}_S \cap \mathcal{B}_T$ es linealmente independiente, ya que está contenido en \mathcal{B} , y está contenido en $S \cap T$, así que

$$\#(\mathcal{B}_S \cap \mathcal{B}_T) \leq \dim S \cap T. \quad (35)$$

De manera similar, el conjunto $\mathcal{B}_S \cup \mathcal{B}_T$ es linealmente independiente, ya que coincide con el subconjunto $\mathcal{B} \cap (S \cup T)$ de \mathcal{B} , y está contenido en $S + T$, así que

$$\#(\mathcal{B}_S \cup \mathcal{B}_T) \leq \dim(S + T). \quad (36)$$

La Proposición 1.8.4 junto con las dos desigualdades que obtuvimos y el Principio de Inclusión–Exclusión para conjuntos nos dicen que

$$\begin{aligned} \#\mathcal{B}_S + \#\mathcal{B}_T &= \dim S + \dim T \\ &= \dim(S + T) + \dim S \cap T \\ &\geq \#(\mathcal{B}_S \cup \mathcal{B}_T) + \#(\mathcal{B} \cap \mathcal{B}_T) \\ &= \#\mathcal{B}_S + \#\mathcal{B}_T. \end{aligned}$$

Esto implica, claro, que todas las desigualdades que aquí aparecen son igualdades y, ya que valen las dos desigualdades (35) y (36), que esas desigualdades son ellas también igualdades. Por supuesto, de eso podemos deducir inmediatamente que $\mathcal{B}_S \cap \mathcal{B}_T$ y $\mathcal{B}_S \cup \mathcal{B}_T$ son bases de $S \cap T$ y de $S + T$, como afirma el lema. \square

1.12.8. El resultado análogo a (33) para tres subespacios es falso: por ejemplo, si $V = \mathbb{k}^2$, $S = \langle e_1 \rangle$, $T = \langle e_2 \rangle$ y $U = \langle e_1 + e_2 \rangle$ es claro que no hay ninguna base de V que contenga bases de S , de T y de U . Otra vez, falta de distributividad es el único obstáculo:

Proposición. Sea V un espacio vectorial y sean S , T y U tres subespacios de V . La terna (S, T, U) es distributiva si y solamente si existe una base \mathcal{B} de $S + T + U$ tal que las intersecciones de \mathcal{B} con cada uno de los subespacios S , T y U son bases de esos subespacios.

Demostración. Supongamos primero que la terna (S, T, U) es distributiva. Sea \mathcal{B}_{STU} una base de $S \cap T \cap U$. Por supuesto,

$$\dim S \cap T \cap U = \#\mathcal{B}_{STU}.$$

Como $S \cap T \cap U$ está contenido en $S \cap T$, podemos completar \mathcal{B}_{STU} a una base de $S \cap T$: existe un conjunto \mathcal{B}_{ST} disjunto de \mathcal{B}_{STU} tal que $\mathcal{B}_{STU} \cup \mathcal{B}_{ST}$ es una base de $S \cap T$, y entonces

$$\dim S \cap T = \#\mathcal{B}_{STU} + \#\mathcal{B}_{ST}.$$

De manera similar, existen conjuntos \mathcal{B}_{SU} y \mathcal{B}_{TU} disjuntos de \mathcal{B}_{STU} tales que $\mathcal{B}_{STU} \cup \mathcal{B}_{SU}$ y $\mathcal{B}_{STU} \cup \mathcal{B}_{TU}$ son bases de $S \cap U$ y de $T \cap U$, respectivamente, y

$$\dim S \cap U = \#\mathcal{B}_{STU} + \#\mathcal{B}_{SU},$$

$$\dim T \cap U = \#\mathcal{B}_{STU} + \#\mathcal{B}_{TU}.$$

La forma en que elegimos los conjuntos \mathcal{B}_{STU} , \mathcal{B}_{ST} y \mathcal{B}_{SU} implica, de acuerdo a la observación (34), que son disjuntos dos a dos y que la unión $\mathcal{B}_{STU} \cup \mathcal{B}_{ST} \cup \mathcal{B}_{SU}$ es una base de $S \cap T + S \cap U$. Como este es un subespacio de S , existe entonces un conjunto \mathcal{B}_S disjunto de $\mathcal{B}_{STU} \cup \mathcal{B}_{ST} \cup \mathcal{B}_{SU}$ tal que $\mathcal{B}_{STU} \cup \mathcal{B}_{ST} \cup \mathcal{B}_{SU} \cup \mathcal{B}_S$ es una base de S y

$$\dim S = \#\mathcal{B}_{STU} + \#\mathcal{B}_{ST} + \#\mathcal{B}_{SU} + \#\mathcal{B}_S.$$

Por la misma razón, hay conjuntos \mathcal{B}_T y \mathcal{B}_U que son disjuntos de $\mathcal{B}_{STU} \cup \mathcal{B}_{ST} \cup \mathcal{B}_{TU}$ y de $\mathcal{B}_{STU} \cup \mathcal{B}_{SU} \cup \mathcal{B}_{TU}$ y tales que $\mathcal{B}_{STU} \cup \mathcal{B}_{ST} \cup \mathcal{B}_{TU} \cup \mathcal{B}_T$ y $\mathcal{B}_{STU} \cup \mathcal{B}_{SU} \cup \mathcal{B}_{TU} \cup \mathcal{B}_U$ son bases de S y de T , respectivamente, y

$$\dim T = \#\mathcal{B}_{STU} + \#\mathcal{B}_{ST} + \#\mathcal{B}_{TU} + \#\mathcal{B}_T,$$

$$\dim U = \#\mathcal{B}_{STU} + \#\mathcal{B}_{SU} + \#\mathcal{B}_{TU} + \#\mathcal{B}_U.$$

Observemos que el conjunto

$$\mathcal{B} := \mathcal{B}_{STU} \cup \mathcal{B}_{ST} \cup \mathcal{B}_{SU} \cup \mathcal{B}_{TU} \cup \mathcal{B}_S \cup \mathcal{B}_T \cup \mathcal{B}_U \tag{37}$$

genera a $S + T + U$, ya que contiene bases de los subespacios S , T y U . Por otro lado,

$$\begin{aligned}
& \dim(S + T + U) \\
& \leq \#(\mathcal{B}_{STU} \cup \mathcal{B}_{ST} \cup \mathcal{B}_{SU} \cup \mathcal{B}_{TU} \cup \mathcal{B}_S \cup \mathcal{B}_T \cup \mathcal{B}_U) \\
& \leq \#\mathcal{B}_{STU} + \#\mathcal{B}_{ST} + \#\mathcal{B}_{SU} + \#\mathcal{B}_{TU} + \#\mathcal{B}_S + \#\mathcal{B}_T + \#\mathcal{B}_U \\
& = (\#\mathcal{B}_{STU} + \#\mathcal{B}_{ST} + \#\mathcal{B}_{SU} + \#\mathcal{B}_S) + (\#\mathcal{B}_{STU} + \#\mathcal{B}_{ST} + \#\mathcal{B}_{TU} + \#\mathcal{B}_T) \\
& \quad + (\#\mathcal{B}_{STU} + \#\mathcal{B}_{SU} + \#\mathcal{B}_{TU} + \#\mathcal{B}_U) - (\#\mathcal{B}_{STU} + \#\mathcal{B}_{ST}) - (\#\mathcal{B}_{STU} + \#\mathcal{B}_{SU}) \\
& \quad - (\#\mathcal{B}_{STU} + \#\mathcal{B}_{TU}) + (\#\mathcal{B}_{STU}) \\
& = \dim S + \dim T + \dim U - \dim S \cap T - \dim S \cap U - \dim T \cap U + \dim S \cap T \cap U \\
& = \dim(S + T + U),
\end{aligned}$$

ya la terna (S, T, U) es distributiva. Esto implica que el conjunto \mathcal{B} tiene cardinal igual a la dimensión de $S + T + U$ y, como además genera a ese subespacio, la Proposición 1.7.9 nos dice que es, de hecho, una base de este. Observemos que esto implica que los 7 conjuntos que aparecen en la unión que define a \mathcal{B} en (37) son disjuntos dos a dos. Vemos así que la base \mathcal{B} de $S + T + U$ tiene la propiedad descripta en el enunciado. La Figura 1.1 de la página siguiente resume esta construcción.

Probemos ahora la implicación recíproca. Supongamos que hay una base \mathcal{B} de $S + T + U$ tal que las intersecciones $\mathcal{B}^S := \mathcal{B} \cap S$, $\mathcal{B}^T := \mathcal{B} \cap T$ y $\mathcal{B}^U := \mathcal{B} \cap U$ son bases de S , de T y de U , respectivamente. De acuerdo al Lema 1.12.7, los conjuntos $\mathcal{B}^S \cap \mathcal{B}^U = \mathcal{B} \cap (S \cap U)$ y $\mathcal{B}^T \cap \mathcal{B}^U = \mathcal{B} \cap (T \cap U)$ son bases de $S \cap U$ y de $T \cap U$ y, entonces,

$$\mathcal{B}^S \cap \mathcal{B}^U \cup \mathcal{B}^T \cap \mathcal{B}^U \text{ es una base de } S \cap U + T \cap U. \quad (38)$$

Por otro lado, el mismo lema nos dice que $\mathcal{B}^S \cup \mathcal{B}^T$ es una base de $S + T$ y que

$$(\mathcal{B}^S \cap \mathcal{B}^T) \cap \mathcal{B}^U \text{ es una base de } (S + T) \cap U. \quad (39)$$

Como $(\mathcal{B}^S \cup \mathcal{B}^T) \cap \mathcal{B}^U = \mathcal{B}^S \cap \mathcal{B}^U \cup \mathcal{B}^T \cap \mathcal{B}^U$, de (38) y (39) vemos que

$$S \cap U + T \cap U = (S + T) \cap U,$$

esto es, que la terna (S, T, U) es distributiva. □

1.12.9. Todas las construcciones de esta sección son del mismo tipo: empezando con tres subespacios de un espacio vectorial hacemos sumas e intersecciones repetidas veces. ¿Cuántos subespacios del espacio ambiente podemos construir haciendo sumas e intersecciones de esta forma? La respuesta a esta pregunta fue dada por Dedekind:

Proposición. Si (S, T, U) es una terna distributiva de subespacios de un espacio vectorial V , se pueden construir haciendo sumas e intersecciones sucesivas a partir de S , T y U a lo sumo 20

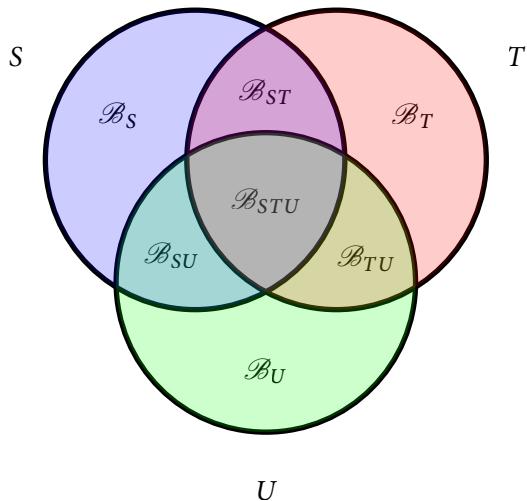


Figura 1.1. La construcción de la prueba de la Proposición 1.12.8.

subespacios de V y todos ellos aparecen en la siguiente lista:

- 0,
- $S \cap T \cap U,$
- $S \cap T,$ $S \cap U,$ $T \cap U,$
- $S \cap T + S \cap U,$ $S \cap T + T \cap U,$ $S \cap U + T \cap U,$
- $S,$ $T,$ $U,$
- $S \cap T + U \cap S + U \cap T,$
- $S + T \cap U,$ $T + S \cap U,$ $U + S \cap T,$
- $S + T,$ $S + U,$ $T + U$
- $S + T + U,$
- $V.$

Es posible elegir V, S, T y U de manera que estos 20 subespacios sean distintos dos a dos.

El número exacto de subespacios que podemos obtener a partir de una terna distributiva (S, T, U) bien puede ser *menor* que 20: por ejemplo, si $0 \not\subseteq S = T = U \not\subseteq V$, entonces las 20 expresiones listadas en la proposición producen solamente 3 subespacios. En la Figura 1.2 de la página siguiente están representados los 20 subespacios del enunciado y todas las relaciones de inclusión entre ellos.

Demostración. Salvo posiblemente por V , los subespacios listados en el enunciado se obtienen

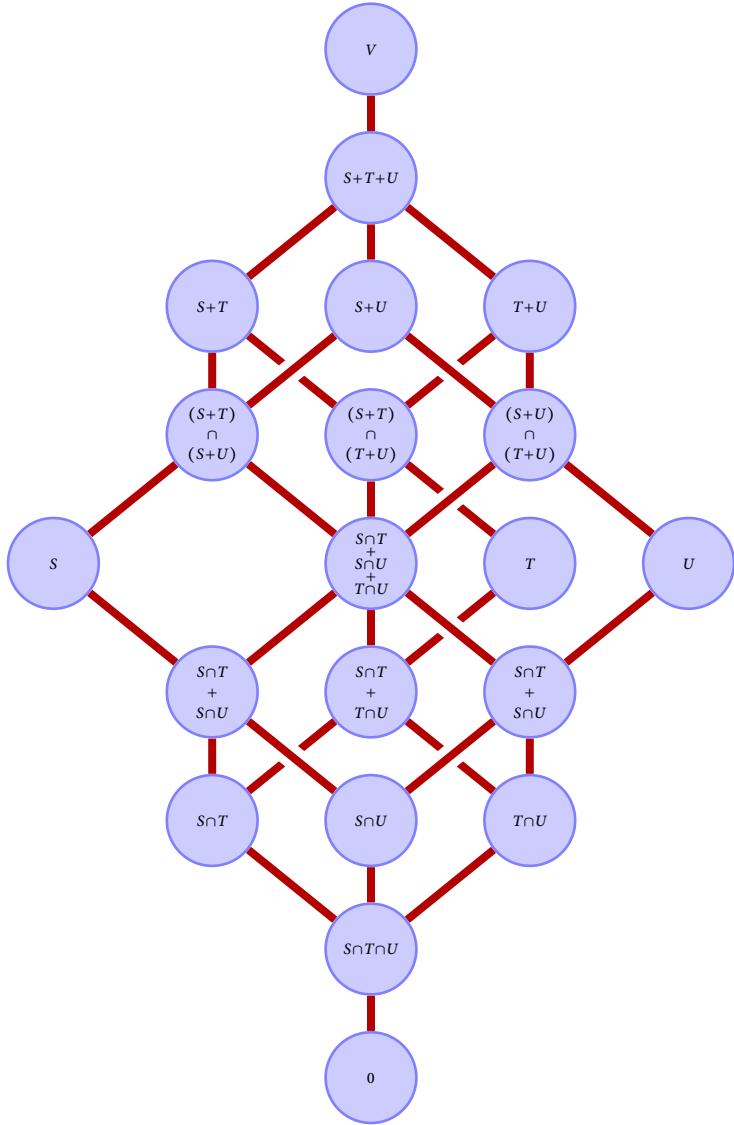


Figura 1.2. Los 20 subespacios que se pueden construir a partir de una terna distributiva (S, T, U) .

haciendo sumas de a lo sumo tres de los siguientes siete subespacios de V :

$$\begin{array}{lll} S \cap T \cap U, \\ S \cap T, & S \cap U, & T \cap U, \\ S, & T, & U. \end{array} \quad (40)$$

Observemos que

*si elegimos cuatro de estos siete subespacios, al menos dos de ellos son comparables*¹.

En efecto, elijamos cuatro de ellos. Si uno de los que elegimos es $S \cap T \cap U$, entonces ese está contenido en cada uno de los otros tres. Supongamos entonces que $S \cap T \cap U$ no es uno de los que elegimos. Alguno de los que sí elegimos tiene que ser uno de los de la segunda fila — digamos, por ejemplo, que elegimos a $S \cap T$: si además elegimos a S o a T , entonces elegimos dos que son comparables, y si no elegimos a ninguno de estos dos, entonces entre los que elegimos están necesariamente $S \cap U$ y U , que son comparables. Una consecuencia de esto es que la suma de cuatro o más de los subespacios listados en (40) es igual a una suma de tres o menos de ellos y, en consecuencia, que la suma de cualesquiera dos de los subespacios de la lista de la proposición aparece también en esa lista.

Usando la distributividad de la terna (S, T, U) y las Proposiciones 1.12.5 y 1.12.6 vemos inmediatamente que la lista del enunciado es la misma que la siguiente:

$$\begin{array}{lll} 0, \\ S \cap T \cap U, \\ S \cap T, & S \cap U, & T \cap U, \\ S \cap (T + U), & T \cap (S + U), & U \cap (S + T), \\ S, & T, & U, \quad (S + T) \cap (S + U) \cap (T + U), \\ (S + T) \cap (S + U), & (S + T) \cap (T + U), & (S + U) \cap (T + U), \\ S + T, & S + U, & T + U \\ S + T + U, \\ V. \end{array}$$

Es claro ahora que los subespacios, salvo posiblemente por 0, se obtienen haciendo intersecciones de a lo sumo tres de los de los siguientes siete:

$$\begin{array}{lll} S, & T, & U, \\ S + T, & S + U, & T + U, \\ S + T + U \end{array} \quad (41)$$

Como antes, cada elección de cuatro de estos siete subespacios contiene dos que son comparables, y esto implica que toda intersección de cuatro o más subespacios de los listados en (41) es igual

¹Decimos que dos subespacios son **comparables** si uno de ellos contiene al otro.

a una intersección de a lo sumo tres de ellos y, en definitiva, que la intersección de dos de los subespacios de la lista del enunciado de la proposición aparece en la misma lista.

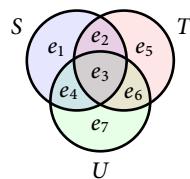
La conclusión de todo esto es que la lista del enunciado es cerrada por sumas y por intersecciones y, como cada uno de los subespacios que aparece en ella es se obtiene haciendo sumas e intersecciones a partir de S , de T y de U , que esa lista contiene *todos* los subespacios de V que pueden obtenerse de esa forma. Esto prueba la primera afirmación de la proposición.

Para ver la verdad de la segunda afirmación es suficiente considerar el siguiente ejemplo: elegimos $V = \mathbb{k}^8$ y los subespacios

$$S = \langle e_1, e_2, e_3, e_4 \rangle,$$

$$T = \langle e_2, e_3, e_5, e_6 \rangle,$$

$$U = \langle e_3, e_4, e_6, e_7 \rangle.$$



Un cálculo sencillo muestra que la terna (S, T, U) que así obtenemos es distributiva y que los 20 subespacios listados en el enunciado son distintos dos a dos. \square

1.12.10. En la prueba de la Proposición 1.12.9 la hipótesis de distributividad de la terna con la que empezamos fue usada varias veces. La siguiente proposición nos dice qué pasa si la terna no es distributiva.

Proposición. Sean S , T y U tres subespacios de un espacio vectorial V . A partir de S , T y U se pueden construir haciendo sumas e intersecciones sucesivas a lo sumo 30 subespacios de V y todos ellos aparecen en la siguiente lista:

0,

$S \cap T \cap U$,

$S \cap T$,

$S \cap U$,

$T \cap U$,

$S \cap T + S \cap U$,

$S \cap T + T \cap U$,

$S \cap U + T \cap U$,

$S \cap (T + U)$,

$U \cap (S + T)$,

$T \cap (S + U)$,

$S \cap T + U \cap S + U \cap T$,

S ,

T ,

U ,

$(S + T \cap U) \cap (T + U)$,

$(T + S \cap U) \cap (S + U)$,

$(U + S \cap T) \cap (S + T)$,

$S + T \cap U$,

$T + S \cap U$,

$U + S \cap T$,

$(S + T) \cap (S + U) \cap (T + U)$,

$(S + T) \cap (T + U)$,

$(S + U) \cap (T + U)$,

$(S + T) \cap (S + U)$,

$S + U$,

$T + U$

$S + T + U$,

V .

Es posible elegir los subespacios S , T y U de manera que estos 30 subespacios sean distintos dos a dos.

De la misma forma que con la Proposición 1.12.9 el número exacto de subespacios que podemos obtener a partir de los tres subespacios S , T y U puede ser *menor* que 30. En la Figura 1.3 de la página siguiente están representados los 30 subespacios del enunciado y todas las relaciones de inclusión entre ellos.

Demostración. Claramente las 30 expresiones que aparecen en el enunciado producen subespacios de V que se obtienen haciendo sumas y productos sucesivamente a partir de S , T y U , y una verificación sencilla pero larga muestra que la intersección y la suma de dos cualesquiera de ellos aparece en la lista. Esto prueba la primera afirmación del enunciado.

Para ver la segunda, consideremos el espacio $V = \mathbb{k}^8$ y sus tres subespacios

$$S = \langle e_1, e_5, e_6, e_7 \rangle, \quad T = \langle e_2, e_4, e_6, e_8 \rangle, \quad U = \langle e_3, e_4, e_5, e_7 + e_8 \rangle.$$

Otra vez, una verificación directa muestra que con esta elección los 30 subespacios del enunciado son distintos dos a dos. \square

1.12.11. El ejemplo que dimos para mostrar que los 30 subespacios de la Proposición 1.12.10 pueden ser distintos dos a dos tiene como espacio ambiente uno de dimensión 8. Esto no es casual: es posible mostrar que si V tiene dimensión *menor* que 8 entonces al menos dos de los 30 coinciden. Por otro lado, uno puede preguntarse cuántos subespacios uno puede obtener haciendo sumas e intersecciones si uno empieza no con tres sino con algún otro número de subespacios. Sorprendentemente, es fácil dar ejemplos para mostrar que, en general, el número es infinito si empezamos con más de tres subespacios.

1.12.12. Ejemplo. Consideremos el espacio vectorial racional $V = \mathbb{Q}^3$ y sus cuatro subespacios

$$R = \langle e_1 \rangle, \quad S = \langle e_2 \rangle, \quad T = \langle e_3 \rangle, \quad U = \langle e_1 + e_2 + e_3 \rangle.$$

Pongamos

$$Y = (S + T) \cap (R + U)$$

y construyamos recursivamente una sucesión $(X_n)_{n \geq 1}$ de subespacios de V poniendo

$$X_1 = (R + S) \cap (T + U)$$

y, para cada $n \in \mathbb{N}$,

$$X_{n+1} = ((R + T) \cap (Y + X_n) + U) \cap (R + S).$$

Calculando explícitamente, vemos que $Y = \langle (0, 1, 1)^t \rangle$ y que para cada $n \in \mathbb{N}$ es $X_n = \langle (n, 1, 0)^t \rangle$. En particular, es $X_n \neq X_m$ siempre que n y m son elementos distintos de \mathbb{N} : esto muestra que se pueden construir infinitos subespacios de V a partir de R , S , T y U haciendo sumas e intersecciones.

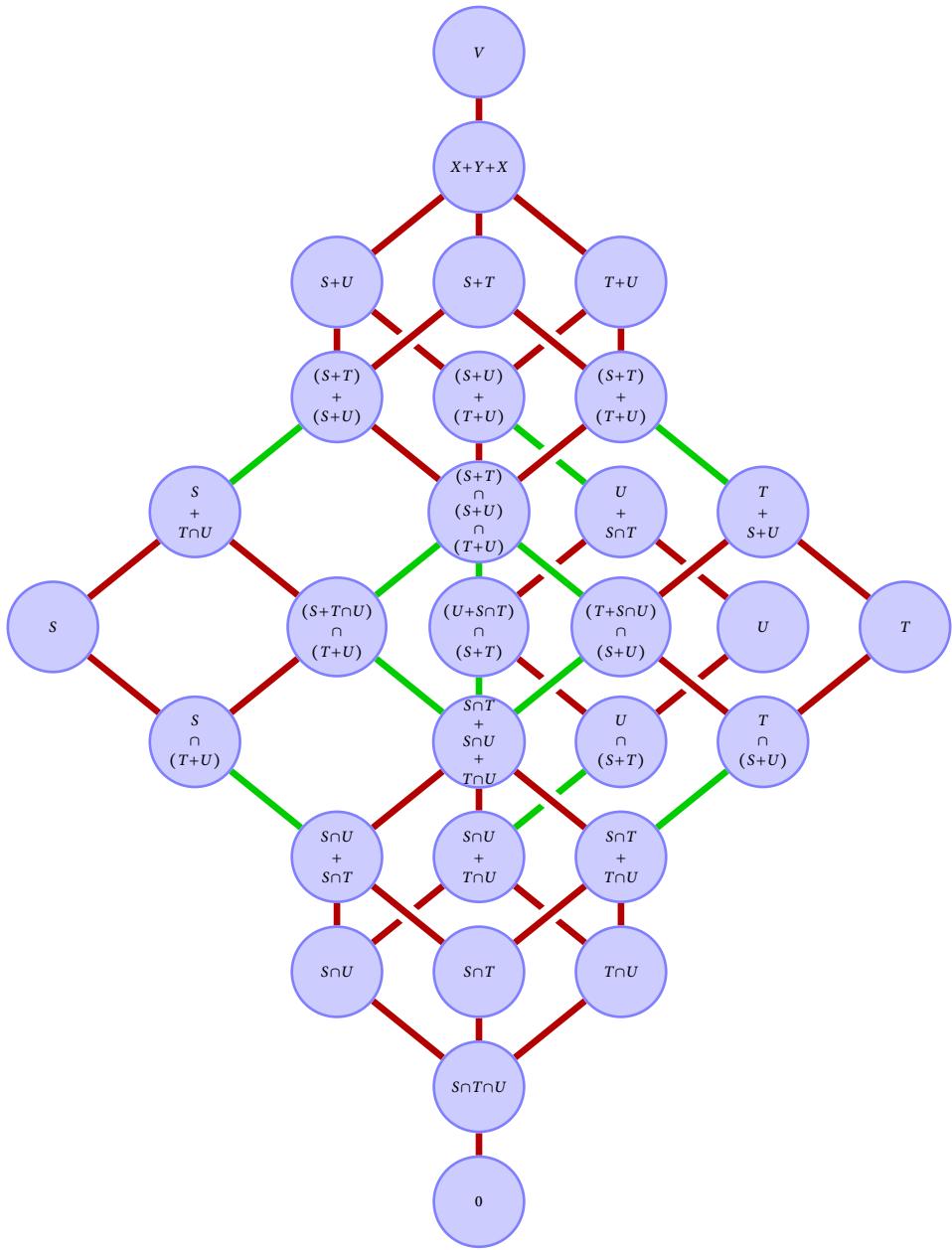


Figura 1.3. Los 30 subespacios que se pueden construir a partir de tres subespacios S , T y U . Las líneas verdes marcan inclusiones que cuando la terna (S, T, U) es distributiva resultan siempre igualdades.

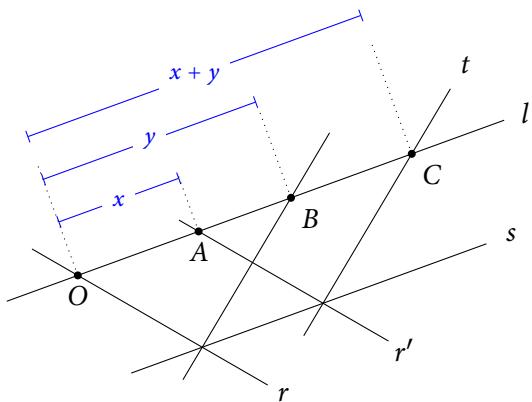


Figura 1.4. Una versión «afín» de la construcción del Ejemplo 1.12.12.

Esta construcción que a primera vista parece totalmente inmotivada puede entenderse de manera muy natural en términos de la geometría del plano proyectivo. Se trata, en efecto de la versión proyectiva de la siguiente construcción en el plano usual:

Dados tres puntos O, A y B sobre una recta l , trazamos por O una recta r distinta de l , la recta r' por A paralela a r , y una recta s distinta de l y paralela a ella. Si t es la paralela a la recta por $r \cap s$ y B que pasa por $r' \cap s$, y C es el punto en que t interseca a l , entonces el segmento OC es la suma algebraica de los segmentos OA y OB .

La Figura 1.4 da un esquema de esto. \diamond

Reticulados de subespacios

1.12.13. Si V es un espacio vectorial, escribamos $\mathcal{L}(V)$ al conjunto de todos los subespacios de V . Un **reticulado** en V es un subconjunto \mathcal{R} de $\mathcal{L}(V)$ que satisface las siguientes dos condiciones:

(**R₁**) $0 \in \mathcal{R}$ y $V \in \mathcal{R}$.

(**R₂**) Si S y T están en \mathcal{R} , entonces $S + T$ y $S \cap T$ están en \mathcal{R} .

1.12.14. El siguiente resultado es completamente similar a la Proposición 1.4.8:

Proposición. *Sea V un espacio vectorial. Si G es un subconjunto de $\mathcal{L}(V)$, entonces existe exactamente un reticulado $\mathcal{R}(G)$ en V que contiene a G y tal que*

todo reticulado \mathcal{S} en V que contiene a G contiene también a $\mathcal{R}(G)$.

Los elementos de $\mathcal{R}(G)$ son precisamente los subespacios de V que pueden obtenerse de los elementos de G haciendo sumas e intersecciones.

Llamamos al reticulado \mathcal{G} que nos da esta proposición el *reticulado en V generado por G* . Cuando el conjunto $G = \{S_1, \dots, S_n\}$ es finito y está dado por enumeración, escribimos normalmente $\mathcal{R}(S_1, \dots, S_n)$ en lugar de $\mathcal{R}(\{S_1, \dots, S_n\})$.

Demostración. Sea G un subconjunto de $\mathcal{L}(V)$ y sea R el conjunto de todos los reticulados en V que contienen a G . Este conjunto R no es vacío —por ejemplo, $\mathcal{L}(V)$ mismo es un elemento de R — así que podemos considerar la intersección

$$\mathcal{R}(G) = \bigcap_{L \in R} L.$$

Es inmediato verificar que $\mathcal{R}(G)$ es un reticulado en V que contiene a G . Más aún, si L es un reticulado en V que contiene a G , entonces $L \in R$ y, por lo tanto, $\mathcal{R}(G) \subseteq L$. Vemos así que la primera afirmación de la proposición es cierta.

Sea ahora $\mathcal{R}'(G)$ el subconjunto de $\mathcal{L}(V)$ de todos los subespacios de V que pueden obtenerse a partir los elementos de G haciendo sumas e intersecciones. Los subespacios 0 y V de V están en $\mathcal{R}(G)$, porque son la suma y la intersección de cero elementos de G , y es evidente, por la definición misma de $\mathcal{R}'(G)$, que es cerrado por sumas e intersecciones. Así, $\mathcal{R}'(G)$ es un reticulado en V que claramente contiene a G y, por lo que ya probamos, se tiene que $\mathcal{R}(G) \subseteq \mathcal{R}'(G)$.

Veamos ahora, para terminar, que vale la inclusión recíproca. Observemos primero que si S es un elemento de $\mathcal{R}'(G)$, de manera que puede obtenerse a partir de elementos de G haciendo sumas e intersecciones, entre todas las posibles formas en que pudo obtenerse de esa manera hay alguna que requiere el menor número posible de esas operaciones: escribamos n_S ese número. Probaremos que todo $S \in \mathcal{R}'(G)$ pertenece a $\mathcal{R}(G)$ haciendo inducción con respecto a n_S .

Si $n_S = 0$, entonces S es un elemento de G y vimos arriba que pertenece a $\mathcal{R}(G)$. Si en cambio es $n_S > 0$, entonces S es o la suma o la intersección de dos subespacios de V , digamos U y T , cada uno de los cuales puede obtenerse a partir de elementos de G haciendo sumas e intersecciones y, más todavía, un momento de reflexión es suficiente para convencernos de que $n_T < n_S$ y $n_U < n_S$. La hipótesis inductiva, entonces, nos dice que T y U están en $\mathcal{R}(G)$ y, como este conjunto es un reticulado, que S también está ahí. \square

1.12.15. Sea V un espacio vectorial.

- Si S es un subespacio de V , entonces

$$\mathcal{R}(S) = \{0, S, V\}.$$

En efecto, este conjunto es un reticulado en V y todos sus elementos pueden obtenerse a partir de elementos de G haciendo sumas e intersecciones. Vemos así que

$$|\mathcal{R}(S)| \leq 3,$$

y eligiendo V y S bien podemos hacer que se tenga una igualdad.

- Si S y T son dos subespacios de V , entonces

$$\mathcal{R}(S, T) = \{0, S \cap T, S, T, S + T, V\}$$

porque, otra vez, este conjunto es un reticulado en V y sus elementos se obtienen a partir de S y de T haciendo sumas e intersecciones. En particular, es

$$|\mathcal{R}(S, T)| \leq 6.$$

Es fácil ver que se pueden elegir V, S y T de manera que esta cota se alcance.

- Si S, T y U son tres subespacios de un espacio vectorial V , entonces la Proposición 1.12.10 nos da todos los elementos del reticulado $\mathcal{R}(S, T, U)$ y nos dice que

$$|\mathcal{R}(S, T, U)| \leq 30$$

y que esta cota puede alcanzarse.

- Por otro lado, el Ejemplo 1.12.12 muestra que si G es un subconjunto de $\mathcal{L}(V)$ que tiene al menos cuatro elementos, entonces $\mathcal{R}(G)$ puede ser infinito.

1.12.16. Decimos que un reticulado \mathcal{R} en un espacio vectorial V es **distributivo** si satisface la condición

(R₃) Si S, T y U son elementos de \mathcal{R} , entonces la terna (S, T, U) es distributiva .

1.12.17. Es fácil verificar que un reticulado en un espacio vectorial que está generado por un subespacio o por dos es distributivo. Cuando está generado por tres subespacios, tenemos el siguiente resultado:

Proposición. Sea V un espacio vectorial y sean S, T y U tres subespacios de V . El reticulado $\mathcal{R}(S, T, U)$ es distributivo si y solamente si la terna (S, T, U) es distributiva.

Demostración. Que la condición para que el reticulado $\mathcal{R}(S, T, U)$ sea distributivo es necesaria es evidente, ya que los subespacios S, T y U pertenecen a $\mathcal{R}(S, T, U)$. Veamos la suficiencia.

Supongamos entonces que la terna (S, T, U) es distributiva. Se deduce inmediatamente de la Proposición 1.12.8 que hay una base \mathcal{B} de V tal que las intersecciones $\mathcal{B} \cap S, \mathcal{B} \cap T$ y $\mathcal{B} \cap U$ son bases de S , de T y de U , respectivamente. A partir de esto y usando el Lema 1.12.7 es fácil ver que, de hecho, para cada $W \in \mathcal{R}(S, T, U)$ se tiene que $\mathcal{B} \cap W$ es una base de W .

Sean ahora A, B y C tres elementos de $\mathcal{R}(S, T, U)$. Ese mismo lema implica que $(A + B) \cap C$ y $A \cap C + B \cap C$ tienen a los conjuntos

$$((\mathcal{B} \cap A) \cup (\mathcal{B} \cap B)) \cap (\mathcal{B} \cap C)$$

y

$$((\mathcal{B} \cap A) \cap (\mathcal{B} \cap C)) \cup ((\mathcal{B} \cap A) \cap (\mathcal{B} \cap C))$$

como bases. Como estas dos bases son iguales, es claro que $(A + B) \cap C = A \cap C + B \cap C$. El reticulado $\mathcal{R}(S, T, U)$ es, por lo tanto, distributivo. \square

1.12.18. El siguiente resultado, originalmente obtenido por *Romolo Musti* y *Ettore Buttafuoco* en [MB56], es la generalización de la Proposición 1.12.17 a reticulados generados por un número finito arbitrario de subespacios:

Proposición. *Sea V un espacio vectorial, sea $n \in \mathbb{N}$ y sean S_1, \dots, S_n subespacios de V . El reticulado $\mathcal{R}(S_1, \dots, S_n)$ es distributivo si y solamente si siempre que $l \in [\![n]\!]$, $1 \leq i_1 < \dots < i_l \leq n$, y $k \in [\![2, l-1]\!]$ la terna*

$$(S_{i_1} + \dots + S_{i_{k-1}}, S_{i_k}, S_{i_{k+1}} \cap \dots \cap S_{i_l})$$

es distributiva. □

La prueba de esto puede encontrarse en el artículo original [MB56] o, por ejemplo, en la sección 1.6 del libro [PPo5] de *Alexander Polishchuk* y *Leonid Positselski*.

1.12.19. La Proposición 1.12.8 que probamos arriba se generaliza también al caso de un reticulado generador cualquier número finito de subespacios:

Proposición. *Sea V un espacio vectorial, sea $n \in \mathbb{N}$ y sean S_1, \dots, S_n subespacios de V . El reticulado $\mathcal{R}(S_1, \dots, R_n)$ es distributivo si y solamente si existe una base \mathcal{B} de V tal que para cada $i \in [\![n]\!]$ la intersección $\mathcal{B} \cap S_i$ es una base de S_i .* □

En [PPo5, Proposition 1.7.1] hay una demostración de esto. Usando nuestro Lema 1.12.7 podemos ver fácilmente que en la situación de la proposición si el reticulado $\mathcal{R}(S_1, \dots, S_n)$ es distributivo y \mathcal{B} es una base de V con la propiedad descripta en la proposición, entonces para todo $T \in \mathcal{R}(S_1, \dots, R_n)$ la intersección $\mathcal{B} \cap T$ es una base de T . Como V tiene dimensión finita, la base \mathcal{B} tiene un número finito de subconjuntos y podemos concluir de esto que $\mathcal{R}(S_1, \dots, S_n)$ es finito. Esto prueba el siguiente corolario.

Corolario. *Sea V un espacio vectorial, sea $n \in \mathbb{N}$ y sean S_1, \dots, S_n subespacios de V . Si el reticulado $\mathcal{R}(S_1, \dots, R_n)$ es distributivo, entonces es finito.* □

De hecho, el cardinal del reticulado $\mathcal{R}(S_1, \dots, S_n)$ puede acotarse por un número que depende solamente de n : es posible mostrar que para cada $n \in \mathbb{N}$ hay un número $D_n \in \mathbb{N}$ tal que

para toda elección de n subespacios S_1, \dots, S_n de un espacio vectorial de dimensión finita tal que el reticulado $\mathcal{R}(S_1, \dots, S_n)$ es distributivo se tiene que

$$|\mathcal{R}(S_1, \dots, S_n)| \leq D_n$$

y

existen un espacio vectorial de dimensión finita V y n subespacios de V tales que el reticulado $\mathcal{R}(S_1, \dots, S_n)$ es distributivo y tiene exactamente D_n elementos.

Llamamos a D_n en n -ésimo **número de Dedekind** y el problema de la determinación explícita de estos números es conocido como el **problema de Dedekind**, estudiado por *Richard Dedekind*

n	D_n
1	3
2	6
3	20
4	168
5	7 581
6	7 828 354
7	2 414 682 040 998
8	56 130 437 228 687 557 907 788

Cuadro 1.1. Los números de Dedekind conocidos.

en [Dedoo]. Los únicos para los que sabemos hacerlo son los primeros ocho, que listamos en el Cuadro 1.1 — el octavo fue calculado por *Doug Wiedemann* [Wie91] recién en 1991. No se conoce ninguna fórmula cerrada para D_n , pero *D. Kleitman* y *G. Markowsky* probaron en [KM75] que

$$\binom{n}{\lfloor n/2 \rfloor} \leq \log_2 D_n \leq \binom{n}{\lfloor n/2 \rfloor} \left(1 + O\left(\frac{\log n}{n}\right)\right).$$

La sucesión de los números de Dedekind es la sucesión A000372 de [OEIS]; allí puede encontrarse mucha más información sobre ella.

1.12.20. La hipótesis de que un conjunto finito de subespacios genere un reticulado distributivo es suficiente para que valga el Principio de Inclusión–Exclusión para la dimensión de la suma de todos ellos: esto generaliza el Corolario 1.12.4.

Proposición. *Sea V un espacio vectorial, sea $n \in \mathbb{N}$ y sean S_1, \dots, S_n subespacios de V . Si el reticulado $\mathcal{R}(S_1, \dots, S_n)$ es distributivo, entonces*

$$\dim(S_1 + \dots + S_n) = \sum_{k=1}^n (-1)^{k+1} \sum_{1 \leq j_1 < \dots < j_k \leq n} \dim(S_{j_1} \cap \dots \cap S_{j_k}).$$

Omitimos la demostración de esto — nos limitamos a observar que es exactamente igual a la del Principio de Inclusión–Exclusión para conjuntos. Por otro lado, notemos que cuando $n = 3$ sabemos que la validez de la igualdad de la proposición es *equivalente* a la distributividad del reticulado, pero que es solo una condición necesaria para la distributividad cuando $n > 4$.

Capítulo 2

Funciones lineales

§1. Funciones lineales

2.1.1. Sean V y W espacios vectoriales. Decimos que una función $f : V \rightarrow W$ es **lineal**, o que es un **homomorfismo** de espacios vectoriales, si

$$f(x + y) = f(x) + f(y), \quad f(\lambda \cdot x) = \lambda \cdot f(x)$$

para cada $x, y \in V$ y cada $\lambda \in \mathbb{k}$. Es inmediato verificar que esto ocurre si y solamente si cada vez que x y y son elementos de V y α y β son escalares de \mathbb{k} se tiene que

$$f(\alpha x + \beta y) = \alpha f(x) + \beta f(y).$$

Es en esta forma en la que generalmente verificaremos que una función es lineal.

Una función lineal $f : V \rightarrow V$ cuyos dominio y codominio son el mismo espacio V es un **endomorfismo** de ese espacio V .

2.1.2. Ejemplos.

- Si V es un espacio vectorial, entonces la función identidad $\text{id}_V : x \in V \mapsto x \in V$ es lineal. Más generalmente, si U es un subespacio de V , entonces la función de inclusión $x \in U \mapsto u \in V$ es lineal.
- Si V y W son espacios vectoriales, entonces la función constante $x \in V \mapsto 0 \in W$ con valor el elemento cero de W es lineal. Llamamos a esta función la **función nula** de V a W .
- Sean m y n enteros positivos. Si $A \in M_{m,n}(\mathbb{k})$ es una matriz de m filas y n columnas con entradas en \mathbb{k} , entonces la función $x \in \mathbb{k}^n \mapsto Ax \in \mathbb{k}^m$ es lineal. Mostraremos más adelante que, de hecho, todas las funciones lineales $\mathbb{k}^n \rightarrow \mathbb{k}^m$ son de esta forma.
- Sea X un conjunto y sea \mathbb{k}^X el espacio vectorial de todas las funciones $X \rightarrow \mathbb{k}$. Si $x \in X$, entonces la función $f \in \mathbb{k}^X \mapsto f(x) \in \mathbb{k}$ es lineal. Esta función es la **evaluación** en el punto x .
- Sea $C(\mathbb{R})$ el espacio vectorial real de todas las funciones continuas $\mathbb{R} \rightarrow \mathbb{R}$ y sea $C^1(\mathbb{R})$ el

subespacio de $C(\mathbb{R})$ de aquellas funciones que son derivables y tienen derivada continua. La función $f \in C^1(\mathbb{R}) \mapsto f' \in C(\mathbb{R})$ es lineal.

- (f) Sea V un espacio vectorial de dimensión finita n y sea $\mathcal{B} = (x_1, \dots, x_n)$ una base ordenada de V . La función $c_{\mathcal{B}} : x \in V \mapsto [x]_{\mathcal{B}} \in \mathbb{k}^n$ que manda cada vector de V a la n -upla ordenada de sus coordenadas con respecto a \mathcal{B} es una función lineal. Veamos por qué en detalle.

Sean x e y dos elementos de V , sean α y β dos escalares, y supongamos que los vectores de coordenadas de x y de y con respecto a \mathcal{B} son $[x]_{\mathcal{B}} = (a_1, \dots, a_n)^t$ e $[y]_{\mathcal{B}} = (b_1, \dots, b_n)^t$, de manera que

$$x = a_1x_1 + \dots + a_nx_n, \quad y = b_1x_1 + \dots + b_nx_n.$$

Claramente es

$$\begin{aligned} \alpha x + \beta y &= \alpha(a_1x_1 + \dots + a_nx_n) + \beta(b_1x_1 + \dots + b_nx_n) \\ &= (\alpha a_1 + \beta b_1)x_1 + \dots + (\alpha a_n + \beta b_n)x_n, \end{aligned}$$

así que el vector de coordenadas de $\alpha x + \beta y$ con respecto a la base ordenada \mathcal{B} es

$$[\alpha x + \beta y]_{\mathcal{B}} = (\alpha a_1 + \beta b_1, \dots, \alpha a_n + \beta b_n)^t.$$

En vista de la forma en que están definidas las operaciones en \mathbb{k}^n , entonces, tenemos que

$$c_B(\alpha x + \beta y) = [\alpha x + \beta y]_{\mathcal{B}} = \alpha[x]_{\mathcal{B}} + \beta[y]_{\mathcal{B}} = \alpha c_{\mathcal{B}}(x) + \beta c_{\mathcal{B}}(y).$$

Esto nos dice precisamente que la función $c_{\mathcal{B}}$ es lineal. \diamond

2.1.3. La clase de las funciones lineales es cerrada por composición:

Proposición. Sean U , V y W espacios vectoriales. Si $f : U \rightarrow V$ y $g : V \rightarrow W$ son funciones lineales, entonces la composición $g \circ f : U \rightarrow W$ también es una función lineal.

Demostración. Sean $f : U \rightarrow V$ y $g : V \rightarrow W$ funciones lineales. Para ver que la composición $g \circ f : U \rightarrow W$ es lineal, verificamos la condición de la definición. Si x e y son dos vectores de U y α y β dos escalares de \mathbb{k} , entonces

$$\begin{aligned} (g \circ f)(\alpha x + \beta y) &= g(f(\alpha x + \beta y)) \\ &= g(\alpha f(x) + \beta f(y)) \quad \text{porque la función } f \text{ es lineal} \\ &= \alpha g(f(x)) + \beta g(f(y)) \quad \text{porque la función } g \text{ es lineal} \\ &= \alpha(g \circ f)(x) + \beta(g \circ f)(y). \end{aligned}$$

Esto prueba la proposición. \square

2.1.4. Dos funciones son iguales si tienen dominios iguales, codominios iguales y toman valores iguales en cada punto de su dominio común. La siguiente proposición nos dice que cuando las

funciones son lineales alcanza con que tomen valores iguales en cada punto de un subconjunto de su dominio que lo genere.

Proposición. Sean V y W dos espacios vectoriales y sea S un subconjunto de V que lo genera. Si $f, g : V \rightarrow W$ son dos funciones lineales y $f(x) = g(x)$ para cada $x \in S$, entonces $f = g$.

Demostración. Supongamos que $f, g : V \rightarrow W$ son dos funciones lineales que satisfacen la condición del enunciado y sea $x \in V$. Como S genera a V , existen $n \in \mathbb{N}_0$, vectores $x_1, \dots, x_n \in S$ y escalares $\alpha_1, \dots, \alpha_n \in \mathbb{k}$ tales que $x = \alpha_1x_1 + \dots + \alpha_nx_n$, y entonces

$$\begin{aligned} f(x) &= f(\alpha_1x_1 + \dots + \alpha_nx_n) \\ &= \alpha_1f(x_1) + \dots + \alpha_nf(x_n) \\ &= \alpha_1g(x_1) + \dots + \alpha_ng(x_n) \quad \text{por la hipótesis hecha sobre } f \text{ y } g \\ &= g(\alpha_1x_1 + \dots + \alpha_nx_n) \\ &= g(x). \end{aligned}$$

Esto muestra que $f = g$, como queremos. \square

2.1.5. El siguiente resultado nos da una forma sistemática de construir funciones lineales sobre un espacio de dimensión finita.

Proposición. Sea V un espacio vectorial de dimensión finita y sea \mathcal{B} una base de V . Si W es un espacio vectorial y $\phi : \mathcal{B} \rightarrow W$ es una función, entonces existe una y sólo una función lineal $f : V \rightarrow W$ tal que $f(x) = \phi(x)$ para cada $x \in \mathcal{B}$.

Demostración. Sea $n = \dim V$ y supongamos que x_1, \dots, x_n son los elementos de la base \mathcal{B} de V . Sea W un espacio vectorial y sea $\phi : \mathcal{B} \rightarrow W$ una función. Construimos una función $f : V \rightarrow W$ de la siguiente manera: si $x \in V$, entonces hay escalares $a_1, \dots, a_n \in \mathbb{k}$ únicamente determinados por x tales que $x = a_1x_1 + \dots + a_nx_n$, y podemos poner

$$f(x) = a_1\phi(x_1) + \dots + a_n\phi(x_n).$$

Veamos que la función que obtenemos así es lineal y que satisface la condición del enunciado:

- Supongamos que x y x' son elementos de V y que α y β son escalares de \mathbb{k} , y sean $a_1, \dots, a_n, b_1, \dots, b_n \in \mathbb{k}$ los escalares tales que $x = a_1x_1 + \dots + a_nx_n$ y $x' = b_1x_1 + \dots + b_nx_n$. De acuerdo a la definición de f , es

$$f(x) = a_1\phi(x_1) + \dots + a_n\phi(x_n), \quad f(x') = b_1\phi(x_1) + \dots + b_n\phi(x_n).$$

Por otro lado, como

$$\begin{aligned} \alpha x + \beta x' &= \alpha(a_1x_1 + \dots + a_nx_n) + \beta(b_1x_1 + \dots + b_nx_n) \\ &= (\alpha a_1 + \beta b_1)x_1 + \dots + (\alpha a_n + \beta b_n)x_n, \end{aligned}$$

la definición de f nos dice que

$$f(\alpha x + \beta x') = (\alpha a_1 + \beta b_1)\phi(x_1) + \cdots + (\alpha a_n + \beta b_n)\phi(x_n)$$

y esto es

$$\begin{aligned} &= \alpha(a_1\phi(x_1) + \cdots + a_n\phi(x_n)) + \beta(b_1\phi(x_1) + \cdots + b_n\phi(x_n)) \\ &= \alpha f(x) + \beta f(x'). \end{aligned}$$

Vemos así que la función f es lineal.

- Si $i \in \llbracket n \rrbracket$, es claro que $x_i = 0x_1 + \cdots + 0x_{i-1} + 1x_i + 0x_{i+1} + \cdots + 0x_n$, así que la definición de f nos dice que

$$f(x_i) = 0\phi(x_1) + \cdots + 0\phi(x_{i-1}) + 1\phi(x_i) + 0\phi(x_{i+1}) + \cdots + 0\phi(x_n) = \phi(x_i).$$

Con esto, la afirmación de existencia de la proposición queda probada. Veamos la de unicidad. Supongamos que $g : V \rightarrow W$ es otra función lineal tal que $g(x) = \phi(x)$ para cada $x \in \mathcal{B}$: como f y g coinciden sobre cada elemento del conjunto \mathcal{B} , que genera a V , y son lineales, la Proposición 2.1.4 nos dice que $f = g$. \square

2.1.6. La proposición que acabamos de probar nos dice que si V es un espacio vectorial y \mathcal{B} es una base de V , entonces toda función $\phi : \mathcal{B} \rightarrow W$ con valores en un espacio vectorial W puede extenderse de manera única a una función lineal $f : V \rightarrow W$. Podemos representar esta situación con el siguiente diagrama:

$$\begin{array}{ccc} \mathcal{B} & \xrightarrow{\phi} & W \\ \downarrow & \nearrow f & \\ V & & \end{array}$$

Podemos explicitar esto de una forma bien concreta: si $n = \dim V$ y x_1, \dots, x_n son los elementos de \mathcal{B} , entonces para cada elección de n elementos y_1, \dots, y_n de W existe una y solo una función lineal $f : V \rightarrow W$ tal que $f(x_i) = y_i$ para cada $i \in \llbracket n \rrbracket$, que es la función lineal que nos da la proposición a partir de la función $\phi : \mathcal{B} \rightarrow W$ tal que $\phi(x_i) = y_i$ para cada $i \in \llbracket n \rrbracket$. Usaremos esto constantemente en todo lo que sigue.

2.1.7. Sabemos que la noción de independencia de subespacios generaliza a la de independencia lineal de vectores: en ese mismo sentido, el siguiente resultado sobre sumas directas generaliza a la Proposición 2.1.5:

Proposición. Sea V un espacio vectorial, sea $n \in \mathbb{N}$ y sean S_1, \dots, S_n subespacios de V tales que $V = S_1 \oplus \cdots \oplus S_n$. Si W es un espacio vectorial y $f_1 : S_1 \rightarrow W, \dots, f_n : S_n \rightarrow W$ son funciones lineales, entonces existe una y sola una función lineal $f : V \rightarrow W$ tal que para cada $i \in \llbracket n \rrbracket$ y cada $x \in S_i$ se tiene que $f(x) = f_i(x)$.

Demostración. Sean W un espacio vectorial y $f_1 : S_1 \rightarrow W, \dots, f_n : S_n \rightarrow W$ funciones lineales. Si $x \in V$, entonces existen $x_1 \in S_1, \dots, x_n \in S_n$, todos únicamente determinados por x , tales que $x = x_1 + \dots + x_n$. Podemos entonces considerar la función

$$f : x \in V \mapsto f_1(x_1) + \dots + f_n(x_n) \in W.$$

Esta función es lineal: en efecto, si x e y son dos elementos de V y a y b son dos escalares de \mathbb{k} , entonces claramente

$$ax + by = (ax_1 + by_1) + \dots + (ax_n + by_n),$$

así que $(ax + by)_i = ax_i + by_i$ para todo $i \in \llbracket n \rrbracket$ y, por lo tanto,

$$\begin{aligned} f(ax + by) &= f_1(ax_1 + by_1) + \dots + f_n(ax_n + by_n) \\ &= (af_1(x_1) + bf_1(y_1)) + \dots + (af_n(x_n) + bf_n(y_n)) \\ &= a(f_1(x_1) + \dots + f_n(x_n)) + \dots + a(f_1(y_1) + \dots + f_n(y_n)) \\ &= af(x) + bf(y). \end{aligned}$$

Por otro lado, si $i \in \llbracket n \rrbracket$ y $x \in S_i$, entonces los vectores $x_1 \in S_1, \dots, x_n \in S_n$ con $x = x_1 + \dots + x_n$ son tales que $x_i = x$ y $x_j = 0$ para cada $j \in \llbracket n \rrbracket \setminus \{i\}$: esto implica, claro, que

$$f(x) = f_1(x_1) + \dots + f_n(x_n) = f_i(x),$$

esto es, que f tiene la propiedad descripta en el enunciado de la propiedades. \square

2.1.8. La Proposición 2.1.5 nos permite construir funciones lineales sobre espacios vectoriales de dimensión finita, pero usándola podemos extender el resultado a espacios vectoriales arbitrarios.

Proposición. Sea V un espacio vectorial y sea \mathcal{B} una base de V . Si W es un espacio vectorial y $\phi : \mathcal{B} \rightarrow W$ es una función, entonces existe una y sólo una función lineal $f : V \rightarrow W$ tal que $f(x) = \phi(x)$ para cada $x \in \mathcal{B}$.

Demostración. Sea W un espacio vectorial y sea $\phi : \mathcal{B} \rightarrow W$ una función. Para cada subconjunto finito S de \mathcal{B} , el subespacio $\langle S \rangle$ de V tiene dimensión finita y tiene a S como base: la Proposición 2.1.5 nos dice entonces que existe exactamente una función lineal $f_S : \langle S \rangle \rightarrow W$ tal que $f_S(x) = \phi(x)$ para cada $x \in S$.

Sea x un elemento de V . Existe un subconjunto finito S de \mathcal{B} tal que $x \in \langle S \rangle$, y entonces podemos calcular $f_S(x)$. Este elemento de W depende solamente de x y no del subconjunto S elegido. Para verlo, supongamos que T es otro subconjunto finito de \mathcal{B} tal que $x \in \langle T \rangle$. Sean x_1, \dots, x_u los elementos de $S \cap T$, y_1, \dots, y_s los elementos de $S \setminus T$ y z_1, \dots, z_t los de $T \setminus S$. Como x está en el subespacio generado por S , existen escalares $a_1, \dots, a_u, b_1, \dots, b_s$ tales que

$$x = a_1x_1 + \dots + a_ux_u + b_1y_1 + \dots + b_sy_s. \tag{1}$$

De manera similar, como x está en el subespacio generado por T , existen escalares $a'_1, \dots, a'_u, c_1, \dots, c_s$ tales que

$$x = a'_1 x_1 + \dots + a'_u x_u + c_1 z_1 + \dots + c_s z_t. \quad (2)$$

De estas dos expresiones para x deducimos que

$$(a_1 - a'_1)x_1 + \dots + (a_u - a'_u)x_u + b_1 y_1 + \dots + b_s y_s + (-c_1)z_1 + \dots + (-c_t)z_t = 0. \quad (3)$$

Como los $u + s + t$ elementos $x_1, \dots, x_u, y_1, \dots, y_s, z_1, \dots, z_t$ de \mathcal{B} son distintos dos a dos, así que, como el conjunto \mathcal{B} es linealmente independiente, los coeficientes que aparecen en (3) son todos nulos: es $b_i = 0$ para cada $i \in [s]$, $c_j = 0$ para cada $j \in [t]$ y $a_k = a'_k$ para cada $k \in [u]$.

De (1) y la forma en que se define a la función f_S tenemos que

$$f_S(x) = a_1 \phi(x_1) + \dots + a_u \phi(x_u) + b_1 \phi(y_1) + \dots + b_s \phi(y_s) = a_1 \phi(x_1) + \dots + a_u \phi(x_u).$$

De la misma forma, de (2) y la definición de f_T , vemos que

$$f_T(x) = a'_1 \phi(x_1) + \dots + a'_u \phi(x_u) + c_1 \phi(y_1) + \dots + c_t \phi(y_t) = a_1 \phi(x_1) + \dots + a_u \phi(x_u).$$

Concluimos de esta forma que $f_S(x) = f_T(x)$, como queríamos.

Afirmamos ahora que hay una función $f : V \rightarrow W$ lineal y una sola tal que

$$\text{si } x \in V \text{ y } S \text{ es un subconjunto finito de } \mathcal{B} \text{ tal que } x \in \langle S \rangle, \text{ entonces } f(x) = f_S(x). \quad (4)$$

Para verlo, consideremos el subconjunto f de $V \times W$ de todas los pares ordenados (x, y) tales que existe un subconjunto finito S de \mathcal{B} tal que $x \in \langle S \rangle$ e $y = f_S(x)$. Claramente f es una relación de V a W : mostremos que, de hecho, se trata de una función $f : V \rightarrow W$.

- Si $x \in V$, sabemos que existe un subconjunto finito S de \mathcal{B} tal que $x \in \langle S \rangle$, y entonces el par ordenado $(x, f_S(x))$ está en f .
- Sean ahora (x, y) y (x, y') dos elementos de f que tienen la misma primera coordenada. De acuerdo a la definición de f , existen subconjuntos finitos S y T de \mathcal{B} tales que $x \in \langle S \rangle$, $x \in \langle T \rangle$, $y = f_S(x)$ e $y' = f_T(x)$. Lo que probamos al principio de esta prueba, entonces, nos dice que $y = y'$.

Observemos la forma en que construimos a f hace evidente que esta función tiene la propiedad (4). Veamos ahora que se trata de una función lineal.

- Sean x e y dos elementos de V y sean α y β dos escalares de \mathbb{k} . Sabemos que existen subconjuntos finitos S y T de \mathcal{B} tales que $x \in \langle S \rangle$ e $y \in \langle T \rangle$, y entonces claramente los tres vectores x , y y $\alpha x + \beta y$ pertenecen a $\langle S \cup T \rangle$. Como $S \cup T$ es un subconjunto finito de \mathcal{B} , de acuerdo a (4) tenemos que

$$f(\alpha x + \beta y) = f_{S \cup T}(\alpha x + \beta y) = \alpha f_{S \cup T}(x) + \beta f_{S \cup T}(y) = \alpha f(x) + \beta f(y),$$

porque la función $f_{S \cup T}$ es lineal.

La función f tiene la propiedad descripta en el enunciado. En efecto, si $x \in \mathcal{B}$, entonces claramente $x \in \langle \{x\} \rangle$ y, por lo tanto, $f(x) = f_{\{x\}}(x) = \phi(x)$, por la forma en que definimos a $f_{\{x\}}$. Finalmente, la función f que construimos es la única función lineal $V \rightarrow W$ que tiene esa propiedad: si $g : V \rightarrow W$ es otra función lineal tal que $g(x) = \phi(x)$ para todo $x \in \mathcal{B}$, entonces $f = g$ porque f y g coinciden en un subconjunto que genera a su dominio. \square

2.1.9. Una aplicación bien sencilla pero importante de las Proposiciones 2.1.5 y 2.1.8 es el siguiente resultado sobre extensión de funciones lineales:

Proposición. Sean V y W espacios vectoriales y sea S un subespacio de V . Si $f : S \rightarrow W$ es una función lineal, entonces existe otra función lineal $g : V \rightarrow W$ tal que $g(x) = f(x)$ para todo $x \in S$.

En otras palabras, toda función lineal definida sobre un subespacio de V puede extenderse a una función lineal definida sobre todo V .

Demostración. Sea $f : S \rightarrow W$ una función lineal definida sobre S y sea \mathcal{B}_S una base de S . Como \mathcal{B}_S es un subconjunto linealmente independiente de V , sabemos que existe una base \mathcal{B} de V tal que $\mathcal{B}_S \subseteq \mathcal{B}$. Por otro lado, sabemos que existe una función lineal $g : V \rightarrow W$ tal que $g(x) = f(x)$ para cada $x \in \mathcal{B}_S$ y $g(x) = 0$ para cada $x \in \mathcal{B} \setminus \mathcal{B}_S$. Claramente tenemos que $\mathcal{B}_S \subseteq \text{Nu}(f - g)$, así que $S = \langle \mathcal{B}_S \rangle \subseteq \text{Nu}(f - g)$ y, por lo tanto, es $f(x) = g(x)$ para todo $x \in S$. \square

§2. Imagen, preimagen y núcleo

2.2.1. Sean V y W espacios vectoriales y sea $f : V \rightarrow W$ una función lineal. Si U es un subespacio de V , la *imagen de U por f* es el subconjunto

$$f(U) = \{f(x) : x \in U\}$$

de W . Por otro lado, si U es un subespacio de W , la *preimagen de U por f* es el subconjunto

$$f^{-1}(U) = \{x \in V : f(x) \in U\}$$

de V . Llamamos *imagen de f* y escribimos $\text{Im}(f)$ a la imagen $f(V)$ de V por f , y llamamos *núcleo de f* y escribimos $\text{Nu}(f)$ a la preimagen $f^{-1}(0)$ del subespacio nulo de W .

2.2.2. Estas dos operaciones producen subespacios:

Proposición. Sean V y W espacios vectoriales y sea $f : V \rightarrow W$ una función lineal.

- (i) Si U es un subespacio de W , entonces la preimagen $f^{-1}(U)$ de U por f es un subespacio de V .
En particular, el núcleo $\text{Nu}(f)$ de f es un subespacio de V .
- (ii) Si U es un subespacio de V , entonces la imagen $f(U)$ de U por f es un subespacio de W .
En particular, la imagen $\text{Im}(f)$ de f es un subespacio de W .

Demostración. (i) Sea U un subespacio de W . Verifiquemos que $f^{-1}(U)$ es un subespacio de V :

- Como $f(0) = 0 \in U$, es $0 \in f^{-1}(U)$.
- Si $x, y \in f^{-1}(U)$, entonces $f(x) \in U$ y $f(y) \in U$. Como U es un subespacio de W , se sigue de esto que $f(x+y) = f(x) + f(y) \in U$, de manera que $x+y \in f^{-1}(U)$.
- Si $\lambda \in \mathbb{k}$ y $x \in f^{-1}(U)$, tenemos que $f(\lambda x) = \lambda f(x) \in U$, porque $f(x) \in U$, y entonces $\lambda x \in f^{-1}(U)$.

(ii) Sea ahora U un subespacio de V . Otra vez, verificamos una a una las condiciones para que $f(U)$ sea un subespacio de W .

- Como $0 \in U$, es $0 = f(0) \in f(U)$.
- Si $x, y \in f(U)$, existen $v, w \in U$ tales que $x = f(v)$ e $y = f(w)$ y entonces

$$x + y = f(v) + f(w) = f(v + w) \in f(U),$$

ya que $v + w \in U$.

- Si $\lambda \in \mathbb{k}$ y $x \in f(U)$, de manera que existe $v \in U$ tal que $x = f(v)$, es

$$\lambda x = \lambda f(v) = f(\lambda v) \in f(U)$$

porque $\lambda v \in U$.

Esto muestra que $f(U)$ es un subespacio de W . □

§3. Monomorfismos, epimorfismos e isomorfismos

2.3.1. Sean V y W espacios vectoriales y sea $f : V \rightarrow W$ una función lineal. Decimos que

- f es un **monomorfismo** si es inyectiva, y que
- f es un **epimorfismo** si es sobreyectiva.

2.3.2. La siguiente proposición nos da el criterio que usamos más frecuentemente para verificar que una función lineal es un monomorfismo: que su núcleo sea el subespacio nulo.

Proposición. Sean V y W espacios vectoriales y sea $f : V \rightarrow W$ una función lineal. Las siguientes afirmaciones son equivalentes:

- La función f es un monomorfismo.
- El núcleo $\text{Nu}(f)$ de f es el subespacio nulo de V .
- Si S es un subconjunto linealmente independiente de V , entonces la imagen $f(S)$ de S por f es un subconjunto linealmente independiente de W .

Demostración. (a) \Rightarrow (b) Supongamos que f es inyectiva. Si $x \in \text{Nu}(f)$, entonces $f(x) = 0 = f(0)$, así que la hipótesis implica que $x = 0$: vemos que el núcleo $\text{Nu}(f)$ sólo contiene al vector nulo de V y que, en consecuencia, es el subespacio nulo de V .

(b) \Rightarrow (c) Supongamos que $\text{Nu}(f) = 0$ y probemos la afirmación contrarrecíproca de la de (c). Sea S un subconjunto de V y supongamos que su imagen $f(S)$ por f es linealmente dependiente. Esto implica que existen elementos $y_1, \dots, y_n \in f(S)$ distintos dos a dos y escalares $a_1, \dots, a_n \in \mathbb{k}$ no todos nulos tales que $a_1y_1 + \dots + a_ny_n = 0$. Ahora bien, como los vectores y_1, \dots, y_n están en $f(S)$, existen vectores $x_1, \dots, x_n \in S$ tales que $y_i = f(x_i)$ para cada $i \in \llbracket n \rrbracket$ y, como aquéllos son distintos dos a dos, éstos también lo son. Observemos que

$$f(a_1x_1 + \dots + a_nx_n) = a_1f(x_1) + \dots + a_nf(x_n) = a_1y_1 + \dots + a_ny_n = 0,$$

así que $a_1x_1 + \dots + a_nx_n \in \text{Nu}(f)$. La hipótesis de que $\text{Nu}(f) = 0$ nos dice entonces que, de hecho, es $a_1x_1 + \dots + a_nx_n = 0$ y esto muestra que el conjunto S es linealmente dependiente.

(c) \Rightarrow (a) Si f no es un monomorfismo, entonces existen dos vectores distintos x e y en V tales que $f(x) = f(y)$ y entonces la imagen del conjunto $S = \{x - y\}$, que es linealmente independiente, es el conjunto $f(S) = \{0\}$, que no lo es. \square

2.3.3. La siguiente proposición es el análogo de la Proposición 2.3.2 para epimorfismos.

Proposición. Sean V y W espacios vectoriales y sea $f : V \rightarrow W$ una función lineal. Las siguientes afirmaciones son equivalentes:

- (a) La función f es un epimorfismo.
- (b) La imagen $\text{Im}(f)$ de f es W .
- (c) Si S es un subconjunto de V que genera a V , entonces la imagen $f(S)$ de S por f genera a W .

Demostración. La equivalencia entre las afirmaciones (a) y (b) es inmediata.

(a) \Rightarrow (c) Supongamos que f es un epimorfismo, sea S un subconjunto de V tal que $V = \langle S \rangle$ y mostremos que $W = \langle f(S) \rangle$. Sea $y \in W$. Como la función f es sobreyectiva, existe $x \in V$ tal que $f(x) = y$ y, como S genera a V , existen $x_1, \dots, x_n \in S$ y $a_1, \dots, a_n \in \mathbb{k}$ tales que $x = a_1x_1 + \dots + a_nx_n$. Entonces

$$y = f(x) = f(a_1x_1 + \dots + a_nx_n) = a_1f(x_1) + \dots + a_nf(x_n) \in \langle f(S) \rangle,$$

ya que, por supuesto, los vectores $f(x_1), \dots, f(x_n)$ están en $f(S)$.

(c) \Rightarrow (a) Probemos la implicación contrarrecíproca: supongamos que f no es sobreyectiva y mostremos que existe un subconjunto S de V que genera a V y tal que $f(S)$ no genera a W . De hecho, podemos poner simplemente $S = V$, que claramente genera a V . Sabemos que $f(S)$ es un subespacio de W y, como f no es sobreyectiva, que es un subespacio propio: esto nos dice que $\langle f(S) \rangle = f(S) \subsetneq W$, como queríamos. \square

2.3.4. Una función lineal $f : V \rightarrow W$ entre espacios vectoriales V y W es un **isomorfismo** si existe otra función lineal $g : W \rightarrow V$ tal que $g \circ f = \text{id}_V$ y $f \circ g = \text{id}_W$; observemos que en ese caso esta

función g también es un isomorfismo, al que llamamos el **isomorfismo inverso** de f .

Por otro lado, decimos que el espacio V es **isomorfo** a otro espacio W , y escribimos $V \cong W$, si existe algún isomorfismo $V \rightarrow W$.

2.3.5. Proposición. Sean V y W espacios vectoriales. Una función lineal $f : V \rightarrow W$ es un isomorfismo si y solamente si es biyectiva, y esto ocurre si y solamente si es a la vez un monomorfismo y un epimorfismo. Cuando es ése el caso, la función inversa $f^{-1} : W \rightarrow V$ es también un isomorfismo.

Demostración. Sea $f : V \rightarrow W$ un isomorfismo y sea $g : W \rightarrow V$ el isomorfismo inverso de f , de manera que $g \circ f = \text{id}_V$ y $f \circ g = \text{id}_W$. Si $x \in \text{Nu}(f)$, entonces $x = g(f(x)) = g(0) = 0$: esto nos dice que la función f es inyectiva. Por otro lado, si $y \in W$, entonces $y = f(g(y)) \in \text{Im}(f)$: la función f es por lo tanto sobreyectiva. Así, f es biyectiva y esto prueba la necesidad de la condición del enunciado.

Veamos ahora su suficiencia. Supongamos que $f : V \rightarrow W$ es una función lineal biyectiva y mostremos que se trata de un isomorfismo. Ahora bien, como f es biyectiva, posee una función inversa $f^{-1} : W \rightarrow V$ y es suficiente que probemos ésta es una función lineal, ya que por supuesto vale que $f^{-1} \circ f = \text{id}_V$ y $f \circ f^{-1} = \text{id}_W$.

- Sean x e y dos elementos de W . Como $f^{-1} \circ f = \text{id}_V$ y f es lineal,

$$f(f^{-1}(x + y)) = x + y = f(f^{-1}(x) + f(f^{-1}(y))) = f(f^{-1}(x) + f^{-1}(y)),$$

de manera que los vectores $f^{-1}(x + y)$ y $f^{-1}(x) + f^{-1}(y)$ tienen la misma imagen por f . Como f es inyectiva, esto implica que $f^{-1}(x + y) = f^{-1}(x) + f^{-1}(y)$.

- Sean $\lambda \in \mathbb{k}$ y $x \in W$. Usando otra vez que $f^{-1} \circ f = \text{id}_V$ y que f es lineal vemos que

$$f(f^{-1}(\lambda x)) = \lambda x = \lambda f(f^{-1}(x)) = f(\lambda f^{-1}(x))$$

y, como f es inyectiva, que $f^{-1}(\lambda x) = \lambda f^{-1}(x)$.

Esto completa la prueba de la proposición. □

2.3.6. Las Proposiciones 2.3.2 y 2.3.3 juntas nos dan la siguiente caracterización de los isomorfismos:

Proposición. Sean V y W espacios vectoriales y sea $f : V \rightarrow W$ una función lineal. Las siguientes afirmaciones son equivalentes:

- La función f es un isomorfismo.
- Es $\text{Nu}(f) = 0$ e $\text{Im}(f) = W$.
- La imagen $f(\mathcal{B})$ de toda base \mathcal{B} de V es una base de W .

Demostración. Esto sigue inmediatamente de las Proposiciones 2.3.2 y 2.3.3, junto con el resultado de la Proposición 2.3.5 que nos dice que una función lineal es un isomorfismo si y solamente si es a la vez un monomorfismo y un epimorfismo. □

2.3.7. Proposición. *La relación de isomorfismo entre espacios vectoriales es una relación de equivalencia, esto es:*

- *es reflexiva: cualquiera sea el espacio vectorial V , se tiene que $V \cong V$;*
- *es simétrica: si V y W son espacios vectoriales y $V \cong W$, entonces $W \cong V$; y*
- *es transitiva: si U , V y W son espacios vectoriales y se tiene que $U \cong V$ y $V \cong W$, entonces también es $U \cong W$.*

Demostración. Veamos las tres afirmaciones de a una:

- Si V es un espacio vectorial, sabemos que la función identidad $\text{id}_V : x \in V \mapsto x \in V$ es lineal y, como claramente es biyectiva, se trata de un isomorfismo: se sigue de esto que $V \cong V$.
- Supongamos que V y W son espacios vectoriales y supongamos que $V \cong W$, de manera que existe un isomorfismo $f : V \rightarrow W$. De acuerdo a la Proposición 2.3.5, sabemos que la función inversa $f^{-1} : W \rightarrow V$ también es un isomorfismo y entonces $W \cong V$.
- Sean U , V y W espacios vectoriales tales que $U \cong V$ y $V \cong W$, de manera que existen isomorfismos $f : U \rightarrow V$ y $g : V \rightarrow W$. La Proposición 2.1.3 nos dice que la composición $g \circ f : U \rightarrow W$ es una función lineal. Como es además biyectiva, se trata de un isomorfismo y, por lo tanto, $U \cong W$. \square

2.3.8. Dos espacios isomorfismos son para el álgebra lineal esencialmente el mismo. Un ejemplo de ello es el siguiente resultado:

Proposición. *Sean V y W espacios vectoriales.*

- (i) *Si $V \cong W$ y V tiene dimensión finita, entonces W tiene dimensión finita y $\dim V = \dim W$.*
- (ii) *Recíprocamente, si V y W tienen dimensión finita y $\dim V = \dim W$, entonces $V \cong W$.*

Observemos que se deduce de la primera parte de esta proposición que si $V \cong W$ y V tiene dimensión infinita, entonces W también tiene dimensión infinita.

Demostración. (i) Si $V \cong W$ y V tiene dimensión finita, entonces hay un isomorfismo $f : V \rightarrow W$ y una base finita \mathcal{B} de V . De acuerdo a la Proposición 2.3.6, el conjunto $f(\mathcal{B})$ es una base de W y, como es finita, W tiene dimensión finita. Más precisamente, como f es una biyección, \mathcal{B} y $f(\mathcal{B})$ tienen la misma cantidad de elementos y, por lo tanto, $\dim V = \dim W$.

(ii) Supongamos que V y W tienen dimensión finita y que $\dim V = \dim W$, y sea $n = \dim V$. Sean $\mathcal{B} = \{x_1, \dots, x_n\}$ y $\mathcal{B}' = \{y_1, \dots, y_n\}$ bases de V y de W , respectivamente. La Proposición 2.1.5 nos dice que existen funciones lineales $f : V \rightarrow W$ y $g : W \rightarrow V$ tales que $f(x_i) = y_i$ y $g(y_i) = x_i$ para cada $i \in \llbracket n \rrbracket$.

La función $g \circ f : V \rightarrow V$ es lineal y es tal que $(g \circ f)(x_i) = x_i$ para cada $i \in \llbracket n \rrbracket$: se trata entonces de una función lineal $V \rightarrow V$ que coincide sobre la base \mathcal{B} de V con la función identidad $\text{id}_V : V \rightarrow V$, y la Proposición 2.1.5 implica entonces que $g \circ f = \text{id}_V$. De la misma forma, podemos ver que $f \circ g = \text{id}_W$ y, en consecuencia, concluir que f es un isomorfismo. Así, V y W son espacios vectoriales isomorfos, como afirma la proposición. \square

2.3.9. Una consecuencia inmediata de la proposición que acabamos de probar es que la dimensión nos permite clasificar a menos de isomorfismo a los espacios vectoriales de dimensión finita:

Corolario. *Dos espacios vectoriales de dimensión finita son isomorfos si y solamente si tienen la misma dimensión.* □

Observemos que esto significa que todo espacio vectorial de dimensión finita es isomorfo a alguno de la forma \mathbb{k}^n con $n \in \mathbb{N}_0$ y, por otro lado, que dos espacios vectoriales de esa forma son isomorfismos si y solamente si son iguales. En otras palabras, el conjunto $\{\mathbb{k}^n : n \in \mathbb{N}_0\}$ es un sistema completo de representantes para las clases de isomorfismo de espacios vectoriales de dimensión finita.

2.3.10. Demos una aplicación importante de los resultados de esta sección. Si $m, n \in \mathbb{N}$, decimos que una matriz $A \in M_{m,n}(\mathbb{k})$ es **inversible** si existe otra matriz $B \in M_{n,m}(\mathbb{k})$ tal que $AB = I_m$ y $BA = I_n$. En principio esta definición no pide que los enteros n y m sean iguales, pero es cierto que las matrices inversibles son necesariamente cuadradas.

Lema. *Sean $m, n \in \mathbb{N}$ y sean $A \in M_{m,n}(\mathbb{k})$ y $B \in M_{n,m}(\mathbb{k})$. Si $AB = I_m$ y $BA = I_n$, entonces $m = n$.*

Demostración. Supongamos que $AB = I_m$ y $BA = I_n$, como en el enunciado, y consideremos las funciones $f : x \in \mathbb{k}^n \mapsto Ax \in \mathbb{k}^m$ y $g : y \in \mathbb{k}^m \mapsto By \in \mathbb{k}^n$, que son lineales. Si $x \in \mathbb{k}^n$ e $y \in \mathbb{k}^m$, entonces $g(f(x)) = BAx = I_n x = x$ y $f(g(y)) = BAY = I_m y = y$: esto nos dice que $g \circ f = \text{id}_{\mathbb{k}^n}$ y que $f \circ g = \text{id}_{\mathbb{k}^m}$, de manera que f y g son isomorfismos mutuamente inversos y, en particular, que $\mathbb{k}^m \cong \mathbb{k}^n$. Se tiene, entonces, que $m = \dim \mathbb{k}^m = \dim \mathbb{k}^n = n$. □

§4. El teorema de la dimensión

2.4.1. Estamos en posición de probar uno de los resultados más importantes del álgebra lineal:

Teorema. *Sean V y W dos espacios vectoriales y sea $f : V \rightarrow W$ una función lineal. Son equivalentes las siguientes afirmaciones:*

- (a) *El espacio vectorial V tiene dimensión finita.*
- (b) *Tanto el núcleo $\text{Nu}(f)$ como la imagen $\text{Im}(f)$ de f tienen dimensión finita.*

Cuando estas afirmaciones se cumplen, se tiene además que

$$\dim V = \dim \text{Nu}(f) + \dim \text{Im}(f). \tag{5}$$

Demostración. Supongamos primero que V tiene dimensión finita y sea $n = \dim V$. Como $\text{Nu}(f)$ es un subespacio de V , la Proposición 1.7.8 nos dice que tiene dimensión finita. Sea r su dimensión

y sea $\{x_1, \dots, x_r\}$ una de sus bases; observemos que los vectores x_1, \dots, x_r son distintos dos a dos. Como $\text{Nu}(f)$ está contenido en V , este conjunto $\{x_1, \dots, x_r\}$ está contenido en V y es allí linealmente independiente. La Proposición 1.7.10 nos dice que existen vectores $x_{r+1}, \dots, x_n \in V$, necesariamente distintos dos a dos, tales que el conjunto $\mathcal{B} = \{x_1, \dots, x_r, x_{r+1}, \dots, x_n\}$ es una base de V . Mostraremos que

$$\text{el subconjunto } \mathcal{B}' = \{f(x_{r+1}), \dots, f(x_n)\} \text{ de } W \text{ tiene } n - r \text{ elementos y es una base de } \text{Im}(f). \quad (6)$$

Se sigue de esto que $\text{Im}(f)$ tiene dimensión finita, que $\dim \text{Im}(f) = n - r$ y, en consecuencia, que

$$\dim V = n = r + (n - r) = \dim \text{Nu}(f) + \dim \text{Im}(f).$$

Esto probará que la afirmación (a) del teorema implica la afirmación (b) y que cuando vale aquella vale la igualdad (5) del enunciado.

Veamos entonces que vale (6). Sea primero $y \in \text{Im}(f)$, de manera que existe $x \in V$ tal que $y = f(x)$. Como \mathcal{B} es una base de V , existen escalares $a_1, \dots, a_n \in \mathbb{k}$ tales que $x = a_1x_1 + \dots + a_nx_n$ y entonces

$$\begin{aligned} y &= f(x) = f(a_1x_1 + \dots + a_nx_n) \\ &= b_1f(x_1) + \dots + a_nf(x_n) \\ &= a_{r+1}f(x_{r+1}) + \dots + a_nf(x_n) \in \langle \mathcal{B}' \rangle, \end{aligned}$$

porque $f(x_1) = \dots = f(x_r) = 0$. Vemos así que el conjunto \mathcal{B}' genera a $\text{Im}(f)$.

Por otro lado, supongamos que $b_1, \dots, b_{n-r} \in \mathbb{k}$ son escalares tales que

$$b_1f(x_{r+1}) + \dots + b_{n-r}f(x_n) = 0.$$

Como f es lineal, esto nos dice que $f(b_1x_{r+1} + \dots + b_{n-r}x_n) = 0$ y, en consecuencia, que el vector $b_1x_{r+1} + \dots + b_{n-r}x_n$ está en el núcleo $\text{Nu}(f)$ de f . Como el conjunto $\{x_1, \dots, x_r\}$ es una base de este subespacio, existen entonces escalares $c_1, \dots, c_r \in \mathbb{k}$ tales que

$$c_1x_1 + \dots + c_rx_r = b_1x_{r+1} + \dots + b_{n-r}x_n.$$

Ahora bien: como el conjunto \mathcal{B} es una base, es linealmente independiente y esta última igualdad implica que, entre otras cosas, $b_1 = \dots = b_{n-r} = 0$. Concluimos de esta forma que los $n - r$ vectores $f(x_{r+1}), \dots, f(x_n)$ son distintos dos a dos y que el conjunto \mathcal{B}' es linealmente independiente.

Para completar la prueba del teorema, tenemos que mostrar que vale la implicación $(b) \Rightarrow (a)$. Supongamos que los espacios $\text{Nu}(f)$ e $\text{Im}(f)$ tienen dimensión finita, sean $r = \dim \text{Nu}(f)$ y $s = \dim \text{Im}(f)$, y sean $\mathcal{B}' = \{x_1, \dots, x_r\}$ y $\mathcal{B}'' = \{y_1, \dots, y_s\}$ bases de $\text{Nu}(f)$ y de $\text{Im}(f)$, respectivamente. Si $i \in [s]$, el vector y_i está en $\text{Im}(f)$, así que existe $z_i \in V$ tal que $f(z_i) = y_i$. En vista de la Proposición 1.6.5, para ver que V tiene dimensión finita bastará que veamos que el conjunto $\mathcal{B} = \{x_1, \dots, x_r, z_1, \dots, z_s\}$ genera a V .

Sea entonces $x \in V$. Como $f(x) \in \text{Im}(f)$ y el conjunto \mathcal{B}'' es una base de $\text{Im}(f)$, existen escalares $b_1, \dots, b_s \in \mathbb{k}$ tales que $f(x) = b_1y_1 + \dots + b_sy_s$ y entonces

$$\begin{aligned} f(x - (b_1z_1 + \dots + b_sz_s)) &= f(x) - f(b_1z_1 + \dots + b_sz_s) \\ &= f(x) - (b_1f(z_1) + \dots + b_sf(z_s)) \\ &= f(x) - (b_1y_1 + \dots + b_sy_s) \\ &= f(x) - f(x) \\ &= 0. \end{aligned}$$

Así, el vector $x - (b_1z_1 + \dots + b_sz_s)$ está en $\text{Nu}(f)$ y, como \mathcal{B}' es una base de ese subespacio, existen escalares $a_1, \dots, a_r \in \mathbb{k}$ tales que

$$x - (b_1z_1 + \dots + b_sz_s) = a_1x_1 + \dots + a_rx_r.$$

Por supuesto, esto nos dice que

$$x = a_1x_1 + \dots + a_rx_r + b_1z_1 + \dots + b_sz_s \in \langle \mathcal{B} \rangle$$

y prueba, en definitiva, que \mathcal{B} genera a V , como queríamos. \square

2.4.2. Sea $f : V \rightarrow W$ es una función lineal. Si la imagen $\text{Im}(f)$ tiene dimensión finita, llamamos **rango** de f al número

$$\text{rg}(f) = \dim \text{Im}(f).$$

De manera similar, si el núcleo $\text{Nu}(f)$ tiene dimensión finita, llamamos **nulidad** de f al número

$$\text{nul}(f) = \dim \text{Nu}(f).$$

Usando este lenguaje, podemos enunciar parte del teorema que acabamos de probar en la siguiente forma:

Corolario. *Sea $f : V \rightarrow W$ una función lineal. Si V tiene dimensión finita, entonces*

$$\dim V = \text{rg}(f) + \text{nul}(f).$$

Es por esto que el Teorema 2.4.3 es conocido como el *Teorema de la nulidad y el rango*.

2.4.3. Es útil explicitar dos casos particulares del Teorema 2.4.1:

Proposición. *Sean V y W espacios vectoriales y sea $f : V \rightarrow W$ una función lineal.*

- (i) *Si f es sobreyectiva y V tiene dimensión finita, entonces W también tiene dimensión finita y $\dim W \leq \dim V$.*
- (ii) *Si f es inyectiva y W tiene dimensión finita, entonces V también tiene dimensión finita y $\dim V \leq \dim W$.*

Demostración. (i) Supongamos que f es sobreyectiva y que V tiene dimensión finita. Esto último, de acuerdo al Teorema 2.4.1, implica que $\text{Nu}(f)$ y $W = \text{Im}(f)$ tienen dimensión finita y que

$$\dim V = \dim \text{Nu}(f) + \dim \text{Im}(f) \geq \dim \text{Im}(f) = \dim W.$$

(ii) Supongamos ahora que f es inyectiva y que W tiene dimensión finita. La imagen $\text{Im}(f)$ de f , que es un subespacio de W , tiene entonces dimensión finita. Por otro lado, como la función f es inyectiva, su núcleo $\text{Nu}(f)$ es el subespacio nulo de V , que tiene dimensión finita —nula, de hecho— y el Teorema 2.4.1 nos dice que V tiene dimensión finita y que

$$\dim V = \dim \text{Nu}(f) + \dim \text{Im}(f) = \dim \text{Im}(f) \leq \dim W,$$

como afirma el enunciado. \square

2.4.4. Otro caso especial importante del Teorema 2.4.1 es el siguiente:

Proposición. Sean V y W espacios vectoriales y sea $f : V \rightarrow W$ una función lineal. Si V y W tienen dimensión finita y $\dim V = \dim W$, entonces las siguientes afirmaciones son equivalentes:

- (a) f es un isomorfismo.
- (b) f es un monomorfismo.
- (c) f es un epimorfismo.

Notemos que esto se aplica, en particular, cuando V y W son el mismo espacio vectorial de dimensión finita y, por lo tanto, la función f es un endomorfismo de V .

Demostración. Supongamos que los espacios V y W tienen dimensión finita y que $\dim V = \dim W$. Es claro que si f es un isomorfismo, entonces es un monomorfismo y un epimorfismo: esto nos dice que (a) implica a (b) y a (c). Veamos las implicaciones recíprocas. Del Teorema 2.4.1 sabemos que

$$\dim V = \dim \text{Nu}(f) + \dim \text{Im}(f). \quad (7)$$

Si f es un monomorfismo, entonces el núcleo de f es el subespacio nulo de V y la igualdad (7) nos dice que $\dim W = \dim V = \dim \text{Im}(f)$: como $\text{Im}(f)$ es un subespacio de W , esto implica que $\text{Im}(f) = W$, esto es, que f es sobreyectiva. Si, en cambio, f es un epimorfismo, entonces $\text{Im}(f) = W$, de manera que $\dim \text{Im}(f) = \dim W = \dim V$ y, de acuerdo a (7), $\dim \text{Nu}(f) = 0$. Esto significa que $\text{Nu}(f) = 0$ y, en consecuencia, que f es inyectiva. \square

2.4.5. Una aplicación sencilla del Teorema 2.4.1 es la siguiente:

Proposición. Sea $f : V \rightarrow W$ una función lineal y sea U un subespacio de W . Si tanto U como el núcleo $\text{Nu}(f)$ tienen dimensión finita, entonces $f^{-1}(U)$ tiene dimensión finita y

$$\dim f^{-1}(U) \leq \dim \text{Nu}(f) + \dim U.$$

Observemos que aquí V y W bien pueden tener dimensión infinita.

Demostración. Restringiendo el dominio y el codominio de f obtenemos una función lineal

$$g : x \in f^{-1}(U) \mapsto f(x) \in U \cap \text{Im}(f).$$

Esta función es sobreyectiva y su núcleo es precisamente $\text{Nu}(f)$. Si U y $\text{Nu}(f)$ tienen dimensión finita, entonces el núcleo y el codominio de g tienen dimensión finita, así que el Teorema 2.4.1 nos dice que el dominio $f^{-1}(U)$ de g tiene dimensión finita y que

$$\dim f^{-1}(U) = \dim \text{Nu}(g) + \dim \text{Im}(g) \leq \dim \text{Nu}(f) + \dim U \cap \text{Im}(f) \leq \dim \text{Nu}(d) + \dim U,$$

como afirma la proposición. \square

2.4.6. Generalmente usamos la proposición anterior vía el siguiente caso particular:

Corolario. Si $f : U \rightarrow V$ y $g : V \rightarrow W$ son funciones lineales que tienen núcleos de dimensión finita, entonces la composición $g \circ f : U \rightarrow W$ también tiene dimensión finita y, de hecho,

$$\dim \text{Nu}(g \circ f) \leq \dim \text{Nu}(f) + \dim \text{Nu}(g).$$

Como antes, los espacios U , V y W no necesariamente tienen aquí dimensión finita.

Demostración. El núcleo de $g \circ f$ es el subespacio $f^{-1}(\text{Nu}(g))$, así que el corolario es consecuencia inmediata de la Proposición 2.4.5. \square

§5. El espacio de homomorfismos entre dos espacios vectoriales

2.5.1. Si V y W son dos espacios vectoriales, escribimos $\text{hom}(V, W)$ al conjunto de todas las funciones lineales $V \rightarrow W$. Es un subconjunto del espacio vectorial W^V de *todas* las funciones $V \rightarrow W$ que construimos en el Ejemplo 1.2.4(b) y, de hecho, se trata de un subespacio. Probemos esto en detalle.

- El elemento nulo de W^V es la función constante nula $x \in V \mapsto 0 \in W$ y esta función es lineal, por lo que pertenece a $\text{hom}(V, W)$.
- Si $f, g \in \text{hom}(V, W)$, entonces $f + g \in \text{hom}(V, W)$: en efecto, cada vez que $x, y \in V$ y $\alpha, \beta \in \mathbb{k}$ tenemos que

$$\begin{aligned} (f + g)(\alpha x + \beta y) &= f(\alpha x + \beta y) + g(\alpha x + \beta y) \\ &= \alpha f(x) + \beta f(y) + \alpha g(x) + \beta g(y) \\ &= \alpha(f(x) + g(x)) + \beta(f(y) + g(y)) \\ &= \alpha(f + g)(x) + \beta(f + g)(y). \end{aligned}$$

- Si $\lambda \in \mathbb{k}$ y $f \in \text{hom}(V, W)$ entonces $\lambda f \in \text{hom}(V, W)$, ya que si $x, y \in V$ y $\alpha, \beta \in \mathbb{k}$ es

$$\begin{aligned}
(\lambda f)(\alpha x + \beta y) &= \lambda f(\alpha x + \beta y) \\
&= \lambda(\alpha f(x) + \beta f(y)) \\
&\stackrel{*}{=} \lambda \alpha f(x) + \lambda \beta f(y) \\
&= \alpha \lambda f(x) + \beta \lambda f(y) \\
&= \alpha(\lambda f)(x) + \beta(\lambda f)(y).
\end{aligned}$$

Es interesante observar que la igualdad marcada con un asterisco * es la primera vez en estas notas en la que necesitamos usar la comutatividad de la multiplicación del cuerpo \mathbb{k} .

Desde ahora en adelante, cada vez que veamos a $\text{hom}(V, W)$ lo consideraremos con la estructura de espacio vectorial que tiene por ser un subespacio de W^V .

2.5.2. Si V es un espacio vectorial, escribimos muchas veces $\text{End}(V)$ en lugar de $\text{hom}(V, V)$.

Funciones bilineales

2.5.3. Si U, V y W son espacios vectoriales, decimos que una función $\beta : V \times W \rightarrow U$ es **bilineal**, o también **\mathbb{k} -bilineal**, si

$$\begin{aligned}
\beta(x + x', y) &= \beta(x, y) + \beta(x', y), & \beta(\lambda x, y) &= \lambda \beta(x, y), \\
\beta(x, y + y') &= \beta(x, y) + \beta(x, y'), & \beta(x, \lambda y) &= \lambda \beta(x, y)
\end{aligned}$$

para cada $x, x' \in V, y, y' \in W, \lambda \in \mathbb{k}$. Observemos que las primeras dos condiciones nos dicen que para todo $y \in W$ la función $x \in V \mapsto \beta(x, y) \in U$ es lineal, mientras que las dos últimas condiciones que para todo $x \in V$ la función $y \in W \mapsto \beta(x, y) \in U$ es lineal: decimos por eso que una función bilineal es una que es lineal en cada una de sus dos variables o separadamente lineal.

2.5.4. Ejemplos.

- Si U, V y W son espacios vectoriales, la función constante $(x, y) \in U \times V \mapsto 0 \in W$ con valor el elemento nulo de W es bilineal.
- La multiplicación $(x, y) \in \mathbb{k} \times \mathbb{k} \rightarrow xy \in \mathbb{k}$ del cuerpo \mathbb{k} es una función bilineal.
- Si $n, m, k \in \mathbb{N}$, la función

$$(X, Y) \in M_{n,m}(\mathbb{k}) \times M_{m,k}(\mathbb{k}) \mapsto X \cdot Y \in M_{n,k}(\mathbb{k})$$

dada por la multiplicación matricial es bilineal. \diamond

2.5.5. La siguiente proposición nos da un ejemplo natural de función bilineal.

Proposición. Sean V y W dos espacios vectoriales. La función

$$\varepsilon : (f, x) \in \text{hom}(V, W) \times V \mapsto f(x) \in W$$

es bilineal.

Llamamos a esta función la **evaluación**.

Demostración. Sean $f, f' \in \text{hom}(V, W)$, $x, x' \in V$ y $\lambda \in \mathbb{k}$. Calculando, vemos que

$$\begin{aligned}\varepsilon(f + f', x) &= (f + f')(x) = f(x) + f'(x) = \varepsilon(f, x) + \varepsilon(f', x), \\ \varepsilon(\lambda f, x) &= (\lambda f)(x) = \lambda f(x) = \lambda \varepsilon(f, x), \\ \varepsilon(f, x + x') &= f(x + x') = f(x) + f(x') = \varepsilon(f, x) + \varepsilon(f, x')\end{aligned}$$

y

$$\varepsilon(f, \lambda x) = f(\lambda x) = \lambda f(x) = \lambda \varepsilon(f, x),$$

y estas cuatro igualdades nos dicen que la función ε es bilineal. \square

2.5.6. Otra función bilineal importante es la dada por la composición de funciones lineales:

Proposición. *Sean U, V y W espacios vectoriales. La función*

$$\beta : (f, g) \in \text{hom}(V, W) \times \text{hom}(U, V) \mapsto f \circ g \in \text{hom}(U, W)$$

es bilineal.

Demostración. Sean $f, f' \in \text{hom}(V, W)$, $g \in \text{hom}(U, V)$ y $\lambda \in \mathbb{k}$, y verifiquemos las dos primeras condiciones de la definición:

- Para cada $x \in U$ es

$$\begin{aligned}\beta(f + f', g)(x) &= ((f + f') \circ g)(x) \\ &= (f + f')(g(x)) \\ &= f(g(x)) + f'(g(x)) \\ &= (f \circ g)(x) + (f' \circ g)(x) \\ &= \beta(f, g)(x) + \beta(f', g)(x) \\ &= (\beta(f, g) + \beta(f', g))(x)\end{aligned}$$

así que $\beta(f + f', g) = \beta(f, g) + \beta(f', g)$,

- Para cada $x \in U$, calculamos que

$$\begin{aligned}\beta(\lambda f, g)(x) &= ((\lambda f) \circ g)(x) \\ &= (\lambda f)(g(x)) \\ &= \lambda f(g(x)) \\ &= \lambda(f \circ g)(x) \\ &= \lambda \beta(f, g)(x),\end{aligned}$$

así que $\beta(\lambda f, g) = \lambda \beta(f, g)$.

La verificación de las otras dos condiciones es enteramente similar. \square

Álgebras

2.5.7. Un *álgebra sobre \mathbb{k}* es un espacio vectorial A dotado de una operación de multiplicación

$$\cdot : A \times A \rightarrow A$$

que satisface las siguientes tres condiciones:

- (A₁) Es una función bilineal.
- (A₂) Es asociativa: para cada $a, b, c \in A$ se tiene que $(a \cdot b) \cdot c = a \cdot (b \cdot c)$.
- (A₃) Posee un elemento neutro: existe un elemento $1 \in A$ tal que para cada $a \in A$ se tiene que $a \cdot 1 = a$ y $1 \cdot a = a$.

Es importante observar que no exigimos que la multiplicación de un álgebra sea commutativa y, de hecho, en general no lo es.

2.5.8. Ejemplos.

- (a) Veamos al cuerpo \mathbb{k} con su estructura usual de espacio vectorial sobre \mathbb{k} . Si dotamos a \mathbb{k} de la multiplicación $\cdot : \mathbb{k} \times \mathbb{k} \rightarrow \mathbb{k}$ que tiene en tanto cuerpo, \mathbb{k} es un álgebra sobre \mathbb{k} .

Más generalmente, si K es un cuerpo y $\mathbb{k} \subseteq K$ es un subcuerpo de K , vimos en el Ejemplo 1.2.5(c) que podemos ver a K como un espacio vectorial sobre \mathbb{k} de manera natural y, de hecho, si lo dotamos de su multiplicación $K \times K \rightarrow K$, que es \mathbb{k} -bilineal, resulta ser un álgebra sobre \mathbb{k} .

- (b) Sea $\mathbb{k}[X]$ el espacio vectorial de los polinomios en la variable X con coeficientes en \mathbb{k} . La multiplicación de polinomios es una operación $\cdot : \mathbb{k}[X] \times \mathbb{k}[X] \rightarrow \mathbb{k}[X]$ que es bilineal y asociativa y que posee un elemento neutro, así que hace de $\mathbb{k}[X]$ un álgebra sobre \mathbb{k} .

- (c) Sea X un conjunto no vacío y sea \mathbb{k}^X el espacio vectorial de todas las funciones $X \rightarrow \mathbb{k}$, como en el Ejemplo 1.2.4(a). Si dotamos a \mathbb{k}^X de la operación usual $\cdot : \mathbb{k}^X \times \mathbb{k}^X \rightarrow \mathbb{k}^X$ de multiplicación de funciones, entonces es un álgebra sobre \mathbb{k} . ◇

2.5.9. Los dos ejemplos de álgebras más importantes para nosotros son los que da la siguiente proposición.

Proposición.

- (i) *El espacio vectorial $\text{End}(V)$ de los endomorfismos de un espacio vectorial V es un álgebra sobre \mathbb{k} con respecto a la operación de multiplicación*

$$\cdot : (f, g) \in \text{End}(V) \times \text{End}(V) \mapsto f \circ g \in \text{End}(V)$$

dada por la composición.

- (ii) *Sea $n \in \mathbb{N}$. El espacio vectorial $\mathbf{M}_n(\mathbb{k})$ de las matrices cuadradas de n filas y n columnas es un álgebra sobre \mathbb{k} con respecto a la operación de multiplicación matricial*

$$\cdot : (A, B) \in \mathbf{M}_n(\mathbb{k}) \times \mathbf{M}_n(\mathbb{k}) \mapsto AB \in \mathbf{M}_n(\mathbb{k}).$$

Demostración. (i) Sabemos de la Proposición 2.5.6 que esta operación es bilineal. Es asociativa porque la composición de funciones es una operación asociativa y claramente tiene a la función identidad $\text{id}_V \in \text{End}(V)$ como elemento neutro.

(ii) Esto es consecuencia de la distributividad y la asociatividad de la multiplicación matricial, y de que la matriz identidad I_n es un elemento neutro para la multiplicación. \square

2.5.10. Es importante tener en cuenta siempre que las álgebras $\text{End}(V)$ de endomorfismos de un espacio vectorial y $M_n(\mathbb{k})$ de matrices en general no son cuerpos: en general, la multiplicación no es conmutativa, no todo elemento no nulo tiene inverso, y el producto de dos elementos no nulos puede ser nulo. Por ejemplo, en el álgebra $M_2(\mathbb{k})$ tenemos que

$$\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix} \neq \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix},$$

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

y que la matriz $\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$, que no es nula, no es inversible.

Homomorfismos de álgebras e ideales

2.5.11. Si A y B son dos álgebras, una función lineal $f : A \rightarrow B$ es un *homomorfismo de álgebras* si $f(1_A) = 1_B$ y cada vez x e y son elementos de A se tiene que

$$f(x \cdot y) = f(x) \cdot f(y).$$

En ese caso es fácil ver que para cada $x \in A$ se tiene que $f(x^i) = f(x)^i$ para todo $i \in \mathbb{N}_0$.

2.5.12. Un homomorfismo de álgebras $f : A \rightarrow B$ es un *isomorfismo de álgebras* si es un isomorfismo de espacios vectoriales. En ese caso sabemos que la función f es biyectiva, que posee una función inversa $f^{-1} : B \rightarrow A$: más aún, esta función inversa también es un homomorfismo de álgebras. En efecto, sabemos que f^{-1} es lineal, y si x e y son dos elementos de B , entonces

$$f(f^{-1}(x \cdot y)) = x \cdot y = f(f^{-1}(x)) \cdot f(f^{-1}(y)) = f(f^{-1}(x) \cdot f^{-1}(y))$$

porque f es un homomorfismo de álgebras y, como f es inyectiva, esto nos dice que

$$f^{-1}(x \cdot y) = f^{-1}(x) \cdot f^{-1}(y).$$

2.5.13. Si A es un álgebra, un subespacio I de A es un *ideal* si cada vez que $x \in I$ e $y \in A$ se tiene que xy e yx están en I . Cada homomorfismo de álgebras nos da un ejemplo de un ideal:

Proposición. Sean A y B dos álgebras. Si $f : A \rightarrow B$ es un homomorfismo de álgebras, entonces el núcleo $\text{Nu}(f)$ es un ideal de A .

Demostración. Sea $f : A \rightarrow B$ un homomorfismo de álgebras. Si $x \in \text{Nu}(f)$ e $y \in A$, entonces

$$f(xy) = f(x)f(y) = 0f(y) = 0,$$

porque f es un homomorfismo y, de manera similar, $f(yx) = 0$. Como $\text{Nu}(f)$ es un subespacio de A , esto nos dice que se trata de un ideal. \square

2.5.14. En general, la multiplicación de un álgebra no es conmutativa. Vale sin embargo la siguiente observación importante, que usaremos muchas veces:

Lema. Sean A y B dos álgebras y sea $f : A \rightarrow B$ un homomorfismo de álgebras. Si A es conmutativa y b y b' son elementos de $f(A)$, entonces $bb' = b'b$.

Demostración. Como b y b' son elementos de $f(A)$, existen a y a' en A tales que $b = f(a)$ y $b' = f(a')$, y entonces

$$bb' = f(a)f(a') = f(aa') = f(a'a) = f(a')f(a) = b'b,$$

porque A es conmutativa. \square

§6. La matriz asociada a una función lineal

2.6.1. Sean V y W dos espacios vectoriales de dimensión finita, sean $n = \dim V$ y $m = \dim W$ sus dimensiones, y sean $\mathcal{B} = (x_1, \dots, x_n)$ y $\mathcal{B}' = (y_1, \dots, y_m)$ bases ordenadas de V y de W , respectivamente. Si $f : V \rightarrow W$ es una función lineal, para cada $i \in \llbracket n \rrbracket$ hay escalares $a_{1,i}, \dots, a_{m,i} \in \mathbb{k}$ bien determinados tales que

$$f(x_i) = a_{1,i}y_1 + \dots + a_{m,i}y_m,$$

las coordenadas de $f(x_i)$ con respecto a la base \mathcal{B}' . De esta manera obtenemos mn elementos de \mathbb{k} , m por cada uno de los elementos de \mathcal{B} , que podemos ordenar en forma de una matriz

$$\begin{pmatrix} a_{1,1} & \cdots & a_{1,n} \\ \vdots & \ddots & \vdots \\ a_{m,1} & \cdots & a_{m,n} \end{pmatrix} \in M_{m,n}(\mathbb{k})$$

de m filas y n columnas con entradas en \mathbb{k} , a la que llamamos **la matriz de f con respecto a las bases ordenadas \mathcal{B} y \mathcal{B}'** y escribimos $[f]_{\mathcal{B}}^{\mathcal{B}'}$. Explicitamente: para cada $i \in \llbracket n \rrbracket$, la i -ésima columna de esta matriz es el vector de coordenadas $[f(x_i)]_{\mathcal{B}'}$ de la imagen por f del i -ésimo vector de la base \mathcal{B} con respecto a la base \mathcal{B}' .

2.6.2. Ejemplos.

- (a) Sean $m, n \in \mathbb{N}$, sea $A = (a_{i,j}) \in M_{m,n}(\mathbb{k})$ una matriz y consideremos la función lineal $f_A : x \in \mathbb{k}^n \mapsto Ax \in \mathbb{k}^m$. Sean $\mathcal{B} = (x_1, \dots, x_n)$ y $\mathcal{B}' = (y_1, \dots, y_m)$ las bases ordenadas estándares de \mathbb{k}^n y de \mathbb{k}^m , respectivamente. Como el vector $f_A(x_i) = Ax_i$ es la columna i -ésima de la matriz A , esto es, $(a_{1,i}, \dots, a_{m,i})^t$, y este vector se escribe en la forma

$$f_A(x_i) = a_{1,i}y_1 + \cdots + a_{m,i}y_m$$

en la base \mathcal{B}' , vemos que la matriz de f_A con respecto a las bases ordenadas \mathcal{B} y \mathcal{B}' es precisamente la matriz A , esto es, que

$$[f_A]_{\mathcal{B}'}^{\mathcal{B}} = A.$$

- (b) Sea V un espacio vectorial de dimensión finita y sea $n = \dim V$. Si $\text{id}_V : V \rightarrow V$ es la función identidad de V , entonces para cada base ordenada \mathcal{B} de V se tiene que

$$[\text{id}_V]_{\mathcal{B}}^{\mathcal{B}} = I_n = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & 1 \end{pmatrix},$$

la matriz identidad de n filas y n columnas.

Observemos que es cierto porque estamos usando la misma base \mathcal{B} tanto para el dominio como para el codominio de id_V : si usásemos bases distintas, la matriz que obtendríamos sería otra. Por ejemplo, si $B = \mathbb{k}^2$, $\mathcal{B} = (x_1, x_2)$ es la base ordenada estándar de V y $\mathcal{B}' = (y_1, y_2)$ es la base ordenada de V que tiene $y_1 = x_1$ e $y_2 = x_1 + x_2$, entonces $\text{id}_V(x_1) = y_1$ e $\text{id}_V(x_2) = -y_1 + y_2$, así que

$$[\text{id}_V]_{\mathcal{B}'}^{\mathcal{B}} = \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix}.$$

- (c) Sea $n \in \mathbb{N}$ y consideremos el espacio vectorial $V = \mathbb{k}[X]_{\leq n}$ de los polinomios con coeficientes en \mathbb{k} de grado a lo sumo n . Sea $f : p \in V \mapsto p' \in V$ el endomorfismo de V que asigna a cada elemento de V su derivada formal y sea $\mathcal{B} = (1, X, \dots, X^n)$ la base ordenada usual de V . Sabemos que $f(1) = 0$ y que para cada $i \in \llbracket n \rrbracket$ es $f(X^i) = iX^{i-1}$. Esto implica inmediatamente que la matriz de f con respecto a las bases \mathcal{B} de su dominio y \mathcal{B} de su codominio es la matriz cuadrada de $n+1$ filas y columnas

$$[f]_{\mathcal{B}}^{\mathcal{B}} = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 2 & 0 & & \vdots \\ \ddots & \ddots & \ddots & & \\ & 0 & n-1 & 0 & \\ \vdots & & 0 & n & \\ 0 & \cdots & & & 0 \end{pmatrix}. \quad \diamond$$

2.6.3. La razón principal por la que estamos interesados en la matriz de una función lineal es la siguiente:

Proposición. Sean V y W dos espacios vectoriales de dimensión finita y sean \mathcal{B} y \mathcal{B}' bases ordenadas de V y de W , respectivamente. Si $f : V \rightarrow W$ es una función lineal y $x \in V$, entonces

$$[f(x)]_{\mathcal{B}'} = [f]_{\mathcal{B}}^{\mathcal{B}'} \cdot [x]_{\mathcal{B}}$$

y, de hecho, $[f]_{\mathcal{B}'}^{\mathcal{B}}$ es la única matriz con esta propiedad.

En palabras: el vector de coordenadas del vector $f(x)$ con respecto a la base ordenada \mathcal{B}' es el producto de la matriz $[f]_{\mathcal{B}}^{\mathcal{B}'}$ de f y el vector de coordenadas $[x]_{\mathcal{B}}$ de x con respecto a la base \mathcal{B} .

Demostración. Sean $n = \dim V$ y $m = \dim W$, y sean $\mathcal{B} = (x_1, \dots, x_n)$ y $\mathcal{B}' = (y_1, \dots, y_m)$. Sea $f : V \rightarrow W$ una función lineal y sea $x \in V$. Si $[x]_{\mathcal{B}} = (a_1, \dots, a_n)^t \in \mathbb{k}^n$ es el vector de coordenadas de x y $[f]_{\mathcal{B}'}^{\mathcal{B}} = (b_{i,j}) \in M_{m,n}(\mathbb{k})$ es la matriz de f , entonces

$$x = a_1x_1 + \dots + a_nx_n$$

y

$$f(x_j) = b_{1,j}y_1 + \dots + b_{m,j}y_m$$

para cada $j \in \llbracket n \rrbracket$, de manera que

$$\begin{aligned} f(x) &= f(a_1x_1 + \dots + a_nx_n) \\ &= a_1f(x_1) + \dots + a_nf(x_n) \\ &= a_1(b_{1,1}y_1 + \dots + b_{m,1}y_m) + \dots + a_n(b_{1,n}y_1 + \dots + b_{m,n}y_m) \\ &= (b_{1,1}a_1 + \dots + b_{1,n}a_n)y_1 + \dots + (b_{m,1}a_1 + \dots + b_{m,n}a_n)y_n. \end{aligned}$$

Esto significa que

$$[f(x)]_{\mathcal{B}'} = \begin{pmatrix} b_{1,1}a_1 + \dots + b_{1,n}a_n \\ \vdots \\ b_{m,1}a_1 + \dots + b_{m,n}a_n \end{pmatrix} = \begin{pmatrix} b_{1,1} & \cdots & b_{1,n} \\ \vdots & \ddots & \vdots \\ b_{m,1} & \cdots & b_{m,n} \end{pmatrix} \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} = [f]_{\mathcal{B}}^{\mathcal{B}'} \cdot [x]_{\mathcal{B}},$$

como afirma la proposición.

Para ver la unicidad, supongamos que $A \in M_{m,n}(\mathbb{k})$ es otra matriz tal que $[f(x)]_{\mathcal{B}'} = A \cdot [x]_{\mathcal{B}}$ para todo $x \in V$. Si $i \in \llbracket n \rrbracket$, entonces la columna i -ésima de A es, escribiendo e_i al vector i -ésimo de la base ordenada estándar de \mathbb{k}^n , el vector $A \cdot e_i$, y

$$A \cdot e_i = A \cdot [x_i]_{\mathcal{B}} = [f(x_i)]_{\mathcal{B}'},$$

que es precisamente la columna i -ésima de $[f]_{\mathcal{B}}^{\mathcal{B}'}$. Esto muestra que, de hecho, $A = [f]_{\mathcal{B}}^{\mathcal{B}'}$. \square

2.6.4. La matriz de una función lineal con respecto a bases ordenadas del dominio y del codominio depende, por supuesto, de la elección de esas bases. La siguiente proposición nos dice precisamente cómo es esta dependencia.

Proposición. *Sean V y W dos espacios vectoriales de dimensión finita y sean \mathcal{B}_1 y \mathcal{B}'_1 dos bases ordenadas de V y \mathcal{B}_2 y \mathcal{B}'_2 dos bases ordenadas de W . Si $f : V \rightarrow W$ es una función lineal, entonces*

$$[f]_{\mathcal{B}'_2}^{\mathcal{B}'_1} = C(\mathcal{B}_2, \mathcal{B}'_2) \cdot [f]_{\mathcal{B}_2}^{\mathcal{B}_1} \cdot C(\mathcal{B}'_1, \mathcal{B}_1).$$

Demostración. Si $x \in V$, se tiene que

$$\begin{aligned} C(\mathcal{B}_2, \mathcal{B}'_2) \cdot [f]_{\mathcal{B}_2}^{\mathcal{B}_1} \cdot C(\mathcal{B}'_1, \mathcal{B}_1) \cdot [x]_{\mathcal{B}'_1} &= C(\mathcal{B}_2, \mathcal{B}'_2) \cdot [f]_{\mathcal{B}_2}^{\mathcal{B}_1} \cdot [x]_{\mathcal{B}_1} \\ &= C(\mathcal{B}_2, \mathcal{B}'_2) \cdot [f(x)]_{\mathcal{B}_2} \\ &= [f(x)]_{\mathcal{B}'_2}. \end{aligned}$$

En vista de la afirmación de unicidad que afirma la Proposición 2.6.3, esto implica la igualdad de que aparece en el enunciado. \square

2.6.5. La aplicación más inmediata de la asignación de una matriz a cada función lineal es el siguiente resultado, que nos permite calcular la dimensión de los espacios de homomorfismos cuando tienen dimensión finita.

Proposición. *Sean V y W dos espacios vectoriales de dimensión finita, sean $n = \dim V$ y $m = \dim W$ sus dimensiones, y sean \mathcal{B} y \mathcal{B}' bases ordenadas de V y de W , respectivamente. La función*

$$\Phi : f \in \hom(V, W) \mapsto [f]_{\mathcal{B}'}^{\mathcal{B}} \in M_{m,n}(\mathbb{k})$$

es un isomorfismo de espacios vectoriales. En particular, el espacio vectorial $\hom(V, W)$ tiene dimensión finita y

$$\dim \hom(V, W) = \dim V \cdot \dim W.$$

Demostración. Supongamos que $\mathcal{B} = (x_1, \dots, x_n)$ y $\mathcal{B}' = (y_1, \dots, y_m)$. Mostremos primero que la función Φ es lineal:

- Sean $f, g \in \hom(V, W)$, y supongamos que $\Phi(f) = [f]_{\mathcal{B}'}^{\mathcal{B}} = (a_{i,j})$ y $\Phi(g) = [g]_{\mathcal{B}'}^{\mathcal{B}} = (b_{i,j})$. Esto significa que para cada $i \in \llbracket n \rrbracket$ es

$$f(x_i) = a_{1,i}y_1 + \cdots + a_{m,i}y_m$$

y

$$g(x_i) = b_{1,i}y_1 + \cdots + b_{m,i}y_m,$$

y entonces

$$\begin{aligned}(f + g)(x_i) &= f(x_i) + g(x_i) \\ &= (a_{1,i}y_1 + \dots + a_{m,i}y_m) + (b_{1,i}y_1 + \dots + b_{m,i}y_m) \\ &= (a_{1,i} + b_{1,i})y_1 + \dots + (a_{m,i} + b_{m,i})y_m.\end{aligned}$$

Vemos así que

$$\Phi(f + g) = [f + g]_{\mathcal{B}'}^{\mathcal{B}} = (a_{i,j} + b_{i,j}) = \Phi(f) + \Phi(g).$$

- Sean ahora $\lambda \in \mathbb{k}$ y $f \in \text{hom}(V, W)$. Sea $\Phi(f) = [f]_{\mathcal{B}'}^{\mathcal{B}} = (a_{i,j})$, de manera que para cada $i \in \llbracket n \rrbracket$ es $f(x_i) = a_{1,i}y_1 + \dots + a_{m,i}y_m$. Como entonces

$$\begin{aligned}(\lambda f)(x_i) &= \lambda f(x_i) \\ &= \lambda(a_{1,i}y_1 + \dots + a_{m,i}y_m) \\ &= (\lambda a_{1,i})y_1 + \dots + (\lambda a_{m,i})y_m,\end{aligned}$$

se tiene que

$$\Phi(\lambda f) = [\lambda f]_{\mathcal{B}'}^{\mathcal{B}} = (\lambda a_{i,j}) = \lambda(a_{i,j}) = \lambda\Phi(f).$$

Para ver que la función Φ es inyectiva, mostremos que tiene núcleo nulo. Sea $f \in \text{hom}(V, W)$ y supongamos que $\Phi(f)$ es el elemento nulo de $M_{m,n}(\mathbb{k})$, esto es, que la matriz $[f]_{\mathcal{B}'}^{\mathcal{B}} = (a_{i,j})$ tiene todas sus entradas nulas. Esto significa que $f(x_i) = 0$ para cada $i \in \llbracket n \rrbracket$ y, entonces, que f coincide con la función nula $0 : V \rightarrow W$ en cada elemento de la base \mathcal{B} . Sabemos, de la Proposición 2.1.5 que esto implica que $f = 0$. Así, $\text{Nu}(\Phi) = 0$, como queríamos.

Por otro lado, la función Φ es sobreyectiva: si $A = (a_{i,j})$ es un elemento de $M_{m,n}(\mathbb{k})$, la Proposición 2.1.5 nos dice que existe una función lineal $f : V \rightarrow W$ tal que $f(x_i) = a_{1,i}y_1 + \dots + a_{m,i}y_m$ para cada $i \in \llbracket n \rrbracket$ y es claro que esta función tiene matriz $[f]_{\mathcal{B}'}^{\mathcal{B}} = A$. Así, $\Phi(f) = A$ y A está en la imagen de Φ .

Hemos probado así que Φ es un isomorfismo de espacios vectoriales. Como el espacio de matrices $M_{m,n}(\mathbb{k})$ tiene dimensión finita y, de hecho, su dimensión es mn , la Proposición 2.3.8 nos dice que $\text{hom}(V, W)$ tiene dimensión finita y que su dimensión es mn . Esto completa la prueba de la proposición. \square

2.6.6. La Proposición 2.6.5 dice, entre otras cosas, que una condición suficiente para que el espacio $\text{hom}(V, W)$ de homomorfismos de un espacio vectorial V a otro W tenga dimensión finita es que tanto V como W tengan dimensión finita. De hecho, esa condición es también casi necesaria: si alguno de V o W tiene dimensión infinita y ninguno de los dos es nulo, entonces $\text{hom}(V, W)$ tiene dimensión infinita. Para probar esto es necesario tener a mano el Teorema 1.6.7.

2.6.7. La regla que asigna matrices a funciones lineales es buena en el sentido de que las matrices reflejan propiedades de las funciones que les dan lugar y nos permiten operar con ellas. El ejemplo más sencillo de esto es el siguiente:

Proposición. Sean U , V y W tres espacios vectoriales de dimensión finita y sean \mathcal{B} , \mathcal{B}' y \mathcal{B}'' bases ordenadas de U , de V y de W , respectivamente. Si $f : U \rightarrow V$ y $g : V \rightarrow W$ son funciones lineales, entonces la matriz de la composición $g \circ f : U \rightarrow W$ con respecto a las bases ordenadas \mathcal{B} y \mathcal{B}'' es

$$[g \circ f]_{\mathcal{B}''}^{\mathcal{B}} = [g]_{\mathcal{B}''}^{\mathcal{B}'} \cdot [f]_{\mathcal{B}'}^{\mathcal{B}}.$$

Demostración. Sea $x \in U$. Es

$$\begin{aligned} [(g \circ f)(x)]_{\mathcal{B}''} &= [g(f(x))]_{\mathcal{B}''} = [g]_{\mathcal{B}''}^{\mathcal{B}'} \cdot [f(x)]_{\mathcal{B}'} = [g]_{\mathcal{B}''}^{\mathcal{B}'} \cdot ([f]_{\mathcal{B}'}^{\mathcal{B}} \cdot [x]_{\mathcal{B}'}) \\ &= ([g]_{\mathcal{B}''}^{\mathcal{B}'} \cdot [f]_{\mathcal{B}'}^{\mathcal{B}}) \cdot [x]_{\mathcal{B}'}, \end{aligned}$$

así que la afirmación de unicidad de la Proposición 2.6.3 nos dice que $[g]_{\mathcal{B}''}^{\mathcal{B}'} \cdot [f]_{\mathcal{B}'}^{\mathcal{B}} = [g \circ f]_{\mathcal{B}''}^{\mathcal{B}}$, como afirma el enunciado. \square

2.6.8. El caso particular de la Proposición 2.6.7 en el que los tres espacios U , V y W coinciden es importante: lo destacamos en la primera parte de la siguiente proposición.

Proposición.

- (i) Sea V un espacio vectorial de dimensión finita, sea $n = \dim V$ y sea \mathcal{B} una base ordenada de V . La función

$$\Phi : f \in \text{End}(V) \mapsto [f]_{\mathcal{B}}^{\mathcal{B}} \in M_n(\mathbb{k})$$

es un isomorfismo de álgebras.

- (ii) Si $n \in \mathbb{N}$ y para cada matriz $A \in M_n(\mathbb{k})$ consideramos la función $f_A : x \in \mathbb{k}^n \mapsto Ax \in \mathbb{k}^n$, entonces

$$\rho : A \in M_n(\mathbb{k}) \mapsto f_A \in \text{End}(\mathbb{k}^n)$$

es un isomorfismo de álgebras.

Demostración. (i) La Proposición 2.6.5 nos dice que Φ es un isomorfismo de espacios vectoriales, la Proposición 2.6.7 que $\Phi(f \cdot g) = \Phi(f) \cdot \Phi(g)$ siempre que f y g están en $\text{End}(V)$, y el Ejemplo 2.6.2(b) que $\Phi(\text{id}_V) = I_n$. Juntando todo, vemos que la función Φ del enunciado es un isomorfismo de álgebras.

(ii) Pongamos $V = \mathbb{k}^n$ en (i) y elijamos como \mathcal{B} a la base ordenada estándar de V . Si $A \in M_n(\mathbb{k})$, sabemos del Ejemplo 2.6.2(a) que $\Phi(\rho(A)) = [f_A]_{\mathcal{B}}^{\mathcal{B}} = A$ y, por otro lado, para cada $g \in \text{End}(V)$ tenemos que

$$\rho(\Phi(g))(x) = f_{[g]_{\mathcal{B}}^{\mathcal{B}}}(x) = [g]_{\mathcal{B}}^{\mathcal{B}} \cdot x = [g]_{\mathcal{B}}^{\mathcal{B}} \cdot [x]_{\mathcal{B}} = [g(x)]_{\mathcal{B}} = g(x), \quad (8)$$

para todo $x \in \mathbb{k}^n$, así que $\rho(\Phi(g)) = g$: juntando estas dos cosas vemos que Φ y ρ son isomorfismos de espacios vectoriales mutuamente inversos y como el primero es un isomorfismo de álgebras,

que el segundo también lo es. Observemos que en (8) usamos dos veces el hecho de que para todo vector y de \mathbb{k}^n se tiene que $y = [y]_{\mathcal{B}}$. \square

2.6.9. Otro ejemplo de cómo podemos obtener información de una función lineal gracias a la matriz que le corresponde es el siguiente:

Proposición. Sean V y W dos espacios vectoriales de dimensión finita, sean $n = \dim V$ y $m = \dim W$ sus dimensiones, y sean \mathcal{B} y \mathcal{B}' bases ordenadas de V y de W , respectivamente.

(i) Si $f : V \rightarrow W$ es un isomorfismo, entonces $m = n$, la matriz $[f]_{\mathcal{B}'}^{\mathcal{B}}$ es cuadrada e inversible, y su inversa es la matriz de la función inversa f^{-1} con respecto a las bases ordenadas \mathcal{B}' y \mathcal{B} , esto es,

$$[f^{-1}]_{\mathcal{B}}^{\mathcal{B}'} = ([f]_{\mathcal{B}'}^{\mathcal{B}})^{-1}.$$

(ii) Recíprocamente, si $f : V \rightarrow W$ es una función lineal tal que la matriz $[f]_{\mathcal{B}'}^{\mathcal{B}}$ es inversible, entonces la función f es un isomorfismo.

Demostración. (i) Como f es un isomorfismo, sabemos de la Proposición 2.3.8 que tiene que ser $m = n$. De acuerdo a la Proposición 2.6.7 y el Ejemplo 2.6.2(b) sabemos que

$$[f^{-1}]_{\mathcal{B}}^{\mathcal{B}'} \cdot [f]_{\mathcal{B}'}^{\mathcal{B}} = [f^{-1} \circ f]_{\mathcal{B}}^{\mathcal{B}} = [\text{id}_V]_{\mathcal{B}}^{\mathcal{B}} = I_n$$

y que

$$[f]_{\mathcal{B}'}^{\mathcal{B}} \cdot [f^{-1}]_{\mathcal{B}}^{\mathcal{B}'} = [f \circ f^{-1}]_{\mathcal{B}}^{\mathcal{B}'} = [\text{id}_W]_{\mathcal{B}'}^{\mathcal{B}'} = I_n,$$

así que la matriz $[f]_{\mathcal{B}'}^{\mathcal{B}}$ es inversible y la matriz $[f^{-1}]_{\mathcal{B}}^{\mathcal{B}'}$ es su inversa.

(ii) Supongamos que $f : V \rightarrow W$ es una función lineal tal que la matriz $A = [f]_{\mathcal{B}'}^{\mathcal{B}}$ es inversible. El Lema 2.3.10 implica que A es cuadrada, esto es, que $n = m$. De acuerdo a la Proposición 2.6.5, existe una función lineal $g : W \rightarrow V$ tal que $[g]_{\mathcal{B}}^{\mathcal{B}'} = A^{-1}$. Es

$$[g \circ f]_{\mathcal{B}}^{\mathcal{B}} = [g]_{\mathcal{B}}^{\mathcal{B}'} \cdot [f]_{\mathcal{B}'}^{\mathcal{B}} = A^{-1} \cdot A = I_n = [\text{id}_V]_{\mathcal{B}}^{\mathcal{B}}$$

y, otra vez de acuerdo a la Proposición 2.6.5, esto nos dice que $g \circ f = \text{id}_V$. De manera similar podemos ver que $f \circ g = \text{id}_W$ y, en definitiva, que f y g son isomorfismos inversos. \square

La traza de una función lineal

2.6.10. Usando la asignación de matrices a funciones lineales podemos construir un invariante importante de los endomorfismos de un espacio vectorial. Empezamos recordando qué es la traza de una matriz cuadrada y sus propiedades más básicas.

2.6.11. Si $n \in \mathbb{N}$ y $A = (a_{i,j})$ es una matriz de tamaño $n \times n$ con entradas en el cuerpo \mathbb{k} , la **traza** de A es el escalar

$$\text{tr}(A) := \sum_{i=1}^n a_{i,i}.$$

Obtenemos de esta forma una función

$$\text{tr} : A \in M_n(\mathbb{k}) \mapsto \text{tr}(A) \in \mathbb{k}.$$

Proposición. (i) Para cada $n \in \mathbb{N}$ la función $\text{tr} : M_n(\mathbb{k}) \rightarrow \mathbb{k}$ es lineal y sobreyectiva.

(ii) Si $n, m \in \mathbb{N}$, $A \in M_{m,n}(\mathbb{k})$ y $B \in M_{n,m}(\mathbb{k})$, entonces

$$\text{tr}(AB) = \text{tr}(BA).$$

(iii) Si $n \in \mathbb{N}$, $A, C \in M_n(\mathbb{k})$ y la matriz C es inversible, entonces

$$\text{tr}(CAC^{-1}) = \text{tr}(A).$$

Demostración. (i) Fijemos $n \in \mathbb{N}$, sean $A = (a_{i,j})$ y $B = (b_{i,j})$ dos matrices de $M_n(\mathbb{k})$ y sean α y β dos escalares de \mathbb{k} . Es $\alpha A + \beta B = (\alpha a_{i,j} + \beta b_{i,j})$, así que

$$\text{tr}(\alpha A + \beta B) = \sum_{i=1}^n (\alpha a_{i,i} + \beta b_{i,i}) = \alpha \sum_{i=1}^n a_{i,i} + \beta \sum_{i=1}^n b_{i,i} = \alpha \text{tr}(A) + \beta \text{tr}(B).$$

Esto nos dice que la función $\text{tr} : M_n(\mathbb{k}) \rightarrow \mathbb{k}$ es lineal. Para ver que es sobreyectiva es suficiente con mostrar que no es idénticamente nula, ya que su codominio es un espacio de dimensión 1 y no posee, por lo tanto, subespacios propios no nulos. Sea $E = (e_{i,j})$ el elemento de $M_n(\mathbb{k})$ cuya única componente no nula es $e_{1,1}$, que es igual a 1: claramente tenemos que $\text{tr}(E) = 1$ y, por lo tanto, que $\text{tr} \neq 0$, como queremos.

(ii) Sean ahora $n, m \in \mathbb{N}$ y sean $A = (a_{i,j}) \in M_{m,n}(\mathbb{k})$ y $B = (b_{j,k}) \in M_{n,m}(\mathbb{k})$ dos matrices. Si escribimos $AB = (c_{i,j})$, entonces para cada $i, j \in [\![m]\!]$ es $c_{i,j} = \sum_{k=1}^n a_{i,k} b_{k,j}$, así que

$$\text{tr}(AB) = \sum_{i=1}^m c_{i,i} = \sum_{i=1}^m \sum_{k=1}^n a_{i,k} b_{k,i}.$$

De la misma forma, si $BA = (d_{i,j})$, entonces para cada $i, j \in [\![n]\!]$ es $d_{i,j} = \sum_{k=1}^m b_{i,k} a_{k,j}$, así que

$$\text{tr}(BA) = \sum_{i=1}^n d_{i,i} = \sum_{i=1}^n \sum_{k=1}^m b_{i,k} a_{k,i}.$$

Claramente, entonces, tenemos que $\text{tr}(AB) = \text{tr}(BA)$.

(iii) Sea finalmente $n \in \mathbb{N}$ y sean $A, C \in M_n(\mathbb{k})$ matrices tales que C es inversible. De acuerdo a la segunda parte de la proposición, tenemos que

$$\text{tr}(CAC^{-1}) = \text{tr}(C^{-1}CA) = \text{tr}(IA) = \text{tr}(A),$$

y esto es lo que afirma el enunciado. □

2.6.12. De acuerdo a la segunda parte de esta proposición, siempre que $n, m \in \mathbb{N}$, $A \in M_{m,n}(\mathbb{k})$ y $B \in M_{n,m}(\mathbb{k})$, tenemos que $\text{tr}(AB) = \text{tr}(BA)$. Decimos por ello que la traza es **cíclicamente invariant**. Se deduce inmediatamente de esto el siguiente resultado más general: si $k, n_0, \dots, n_k \in \mathbb{N}$

con $n_0 = n_k$ y tenemos matrices $A_1 \in M_{n_0, n_1}$, $A_2 \in M_{n_1, n_2}$, ..., $A_k \in M_{n_{k-1}, n_k}$, de manera que tienen sentido los productos $A_1 A_2 \cdots A_k$ y $A_2 \cdots A_k A_1$ y son matrices cuadradas de tamaños $n_0 \times n_0$ y $n_1 \times n_1$, respecto, entonces

$$\text{tr}(A_1 A_2 \cdots A_k) = \text{tr}(A_2 \cdots A_k A_1).$$

Describimos esto diciendo que la traza de un producto de matrices es invariantes por permutaciones cíclicas de los factores. Es importante notar que, a pesar de esto, no es cierto que la traza de un producto de matrices no dependa del orden de los factores: es fácil encontrar tres matrices A , B y C en $M_2(\mathbb{k})$ tales que $\text{tr}(ABC) \neq \text{tr}(ACB)$.

2.6.13. Usando estas propiedades de la función traza, podemos hacerla siguiente observación importante:

Proposición. *Sea V un espacio vectorial de dimensión finita, sea $n = \dim V$ y sea $f : V \rightarrow V$ un endomorfismo de V . Si \mathcal{B} y \mathcal{B}' son dos bases ordenadas de V , entonces*

$$\text{tr}([f]_{\mathcal{B}}^{\mathcal{B}}) = \text{tr}([f]_{\mathcal{B}'}^{\mathcal{B}'}).$$

Demostración. Sean \mathcal{B} y \mathcal{B}' dos bases ordenadas de V y sea $C(\mathcal{B}, \mathcal{B}') = (c_{i,j})$ la matriz de cambio de base de \mathcal{B} a \mathcal{B}' . La Proposición 1.11.4(iii) nos dice que esta matriz es inversible y la Proposición 2.6.4 que

$$[f]_{\mathcal{B}'}^{\mathcal{B}'} = C(\mathcal{B}, \mathcal{B}') \cdot [f]_{\mathcal{B}}^{\mathcal{B}} \cdot C(\mathcal{B}', \mathcal{B}) = C(\mathcal{B}, \mathcal{B}') \cdot [f]_{\mathcal{B}}^{\mathcal{B}} \cdot C(\mathcal{B}, \mathcal{B}')^{-1}.$$

Usando la Proposición 2.6.11(iii), entonces, vemos que

$$\text{tr}([f]_{\mathcal{B}'}^{\mathcal{B}'}) = \text{tr}\left(C(\mathcal{B}, \mathcal{B}') \cdot [f]_{\mathcal{B}}^{\mathcal{B}} \cdot C(\mathcal{B}, \mathcal{B}')^{-1}\right) = \text{tr}([f]_{\mathcal{B}}^{\mathcal{B}}),$$

como afirma la proposición. \square

2.6.14. Si V es un espacio vectorial de dimensión finita y $f : V \rightarrow V$ es un endomorfismo de V , la **traza** $\text{tr}(f)$ de f es el escalar que se obtiene eligiendo una base ordenada \mathcal{B} de V y poniendo

$$\text{tr}(f) := \text{tr}([f]_{\mathcal{B}}^{\mathcal{B}}).$$

La Proposición 2.6.13 nos dice exactamente que esto tiene sentido, porque el lado derecho de esta definición no depende de la base ordenada \mathcal{B} elegida sino solamente de f . Obtenemos de esta forma una nueva función

$$\text{tr} : \text{End}(V) \rightarrow \mathbb{k}.$$

2.6.15. Definimos a la traza de un endomorfismo usando la traza de su matriz: no es sorprendente, entonces, que la primera tenga propiedades similares a la segunda.

Proposición. Sean V y W espacios vectoriales de dimensión finita.

- (i) La función $\text{tr} : \text{End}(V) \rightarrow \mathbb{k}$ es lineal y sobreyectiva.
- (ii) Si $f : V \rightarrow W$ y $g : W \rightarrow V$ son funciones lineales, entonces

$$\text{tr}(f \circ g) = \text{tr}(g \circ f).$$

- (iii) Si $f : V \rightarrow V$ y $g : W \rightarrow V$ son funciones lineales y g es un isomorfismo, entonces

$$\text{tr}(g \circ f \circ g^{-1}) = \text{tr}(f).$$

Demostración. (i) Sean f y g elementos de $\text{End}(V)$ y sean α y β escalares de \mathbb{k} . Si $n = \dim V$ y $\mathcal{B} = (x_1, \dots, x_n)$ es una base ordenada de V , entonces

$$\begin{aligned} \text{tr}(\alpha f + \beta g) &= \text{tr}([\alpha f + \beta g]_{\mathcal{B}}^{\mathcal{B}}) = \text{tr}(\alpha[f]_{\mathcal{B}}^{\mathcal{B}} + \beta[g]_{\mathcal{B}}^{\mathcal{B}}) \\ &= \alpha \text{tr}([f]_{\mathcal{B}}^{\mathcal{B}}) + \beta \text{tr}([g]_{\mathcal{B}}^{\mathcal{B}}) = \alpha \text{tr}(f) + \beta \text{tr}(g). \end{aligned}$$

Esto prueba la linealidad. Por otro lado, sabemos que hay una función lineal $e : V \rightarrow E$ tal que $e(x_1) = x_1$ y $e(x_i) = 0$ para cada $i \in [2, n]$, y la matriz $[e]_{\mathcal{B}}^{\mathcal{B}}$ es precisamente la matriz que cuya traza calculamos en la prueba de la Proposición 2.6.11: tenemos entonces que $\text{tr}(e) = 1$.

(ii) Si $f : V \rightarrow W$ y $g : W \rightarrow V$ son funciones lineales y \mathcal{B} y \mathcal{B}' son bases ordenadas de V y de W , respectivamente, tenemos que

$$\begin{aligned} \text{tr}(f \circ g) &= \text{tr}([f \circ g]_{\mathcal{B}'}^{\mathcal{B}'}) = \text{tr}([f]_{\mathcal{B}'}^{\mathcal{B}'} \cdot [g]_{\mathcal{B}'}^{\mathcal{B}'}) = \text{tr}([g]_{\mathcal{B}'}^{\mathcal{B}'} \cdot [f]_{\mathcal{B}'}^{\mathcal{B}'}) = \text{tr}([g \circ f]_{\mathcal{B}}^{\mathcal{B}}) \\ &= \text{tr}(g \circ f). \end{aligned}$$

(iii) Finalmente, si $f : V \rightarrow V$ y $g : W \rightarrow V$ son funciones lineales y g es un isomorfismo, entonces usando lo que acabamos de probar vemos que

$$\text{tr}(g \circ f \circ g^{-1}) = \text{tr}(g^{-1} \circ g \circ f) = \text{tr}(\text{id}_V \circ f) = \text{tr}(f). \quad \square$$

2.6.16. Las propiedades de la traza descriptas en la Proposiciones 2.6.11 y 2.6.15 las determinan casi completamente: ese es el contenido de la segunda parte de la siguiente proposición.

Proposición. Sea $n \in \mathbb{N}$.

- (i) El núcleo de la función traza $\text{tr} : \mathbf{M}_n(\mathbb{k}) \rightarrow \mathbb{k}$ es el subespacio generado por el conjunto de las matrices de la forma

$$AB - BA, \quad \text{con } A, B \in \mathbf{M}_n(\mathbb{k}).$$

- (ii) Si $t : \mathbf{M}_n(\mathbb{k}) \rightarrow \mathbb{k}$ es una función lineal tal que $t(AB) = t(BA)$ para cada elección de A y B en $\mathbf{M}_n(\mathbb{k})$, entonces existe $\lambda \in \mathbb{k}$ tal que $t = \lambda \text{tr}$.

Demostración. (i) Como sabemos que $\text{tr} : M_n(\mathbb{k}) \rightarrow \mathbb{k}$ es lineal y sobreyectiva, el Teorema 2.4.1 nos dice que

$$\dim \text{Nu}(\text{tr}) = \dim M_n(\mathbb{k}) - \dim \mathbb{k} = n^2 - 1. \quad (9)$$

Sea $K := \langle AB - BA : A, B \in M_n(\mathbb{k}) \rangle$ el subespacio descripto en el enunciado. Para cada A y B de $M_n(\mathbb{k})$ es

$$\text{tr}(AB - BA) = \text{tr}(AB) - \text{tr}(BA) = 0,$$

así que $AB - BA \in \text{Nu}(\text{tr})$: esto implica que, de hecho, $K \subseteq \text{Nu}(\text{tr})$ y, en particular, que

$$\dim K \leq \dim \text{Nu}(\text{tr}) = n^2 - 1.$$

Para cada $i, j \in \llbracket n \rrbracket$ sea $E_{i,j}$ la matriz de $M_n(\mathbb{k})$ que tiene todas sus entradas nulas salvo la (i, j) -ésima, que es igual a 1. Es fácil calcular que cualesquiera sean $i, j, k, l \in \llbracket n \rrbracket$ se tiene que

$$E_{i,j}E_{k,l} - E_{k,l}E_{i,j} = \delta_{j,k}E_{i,j} - \delta_{l,i}E_{k,j}.$$

Usando esto vemos que si i, j y k en $\llbracket n \rrbracket$ son tales que $i \neq j = k$, entonces

$$E_{i,j} = E_{i,k}E_{k,j} - E_{k,j}E_{i,k} \in K,$$

y que si $i \in \llbracket 2, n \rrbracket$ entonces

$$E_{i,i} = E_{1,1} + \sum_{j=i}^{i-1} (E_{i+1,i+1} - E_{i,i}) = E_{1,1} + \sum_{j=i}^{i-1} (E_{i+1,i}E_{i,i+1} - E_{i,i+1}E_{i+1,i}) \in \langle E_{1,1} \rangle + K.$$

Esto nos dice que

$$M_n(\mathbb{k}) = \langle E_{i,j} : i, j \in \llbracket n \rrbracket \rangle \subseteq \langle E_{1,1} \rangle + K$$

y, por lo tanto, de acuerdo al Corolario 1.8.5 y la igualdad (9), que

$$n^2 = \dim M_n(\mathbb{k}) \leq \dim \langle E_{1,1} \rangle + \dim K \leq 1 + (n^2 - 1) = n^2.$$

Podemos concluir entonces que $\dim K = n^2 - 1$, en definitiva, como $K \subseteq \text{Nu}(f)$ y $\dim \text{Nu}(f) = n^2 - 1$, que $K = \text{Nu}(f)$. Esto es lo que queríamos probar.

(ii) Sea $t : M_n(\mathbb{k}) \rightarrow \mathbb{k}$ una función que satisface la condición del enunciado y supongamos, ya que en caso contrario lo que tenemos que probar es evidente que $t \neq 0$. Si A y B están en $M_n(\mathbb{k})$, entonces

$$t(AB - BA) = t(AB) - t(BA) = 0,$$

así que $AB - BA \in \text{Nu}(t)$. En vista de la afirmación (i) que ya probamos, esto implica que $\text{Nu}(\text{tr}) \subseteq \text{Nu}(t)$. Por otro lado, como la función t no es nula y su codominio tiene dimensión 1, es sobreyectiva y usando otra vez el Teorema 2.4.1 vemos que

$$\dim \text{Nu}(t) = \dim M_n(\mathbb{k}) - \dim \mathbb{k} = n^2 - 1 = \dim \text{Nu}(\text{tr}).$$

De todo esto se deduce que $\text{Nu}(t) = \text{Nu}(\text{tr})$. Si ahora $A \in M_n(\mathbb{k})$, entonces

$$\text{tr}(A - \text{tr}(A)E_{1,1}) = \text{tr}(A) - \text{tr}(A)\text{tr}(E_{1,1}) = 0,$$

ya que $\text{tr}(E_{1,1}) = 1$, así que $A - \text{tr}(A)E_{1,1} \in \text{Nu}(\text{tr}) = \text{Nu}(t)$ y, por lo tanto,

$$t(A - \text{tr}(A)E_{1,1}) = t(A) - \text{tr}(A)t(E_{1,1}),$$

de manera que

$$t(A) = t(E_{1,1})\text{tr}(A).$$

Vemos así que si ponemos $\lambda = t(E_{1,1})$, entonces $t = \lambda \text{tr}$, como queremos. \square

2.6.17. Si $n \in \mathbb{N}$ y A y B son matrices de $M_n(\mathbb{k})$, el *comutador* de A y B es la matriz

$$[A, B] := AB - BA.$$

Usando el hecho de que la multiplicación matricial es bilineal y asociativa, es inmediato verificar que la función

$$[-, -] : M_n(\mathbb{k}) \times M_n(\mathbb{k}) \rightarrow M_n(\mathbb{k})$$

es bilineal y que siempre que A, B y C están en $M_n(\mathbb{k})$ se tiene que

$$\begin{aligned} [A, B] &= -[B, A], \\ [A, [B, C]] + [B, [C, A]] + [C, [A, B]] &= 0. \end{aligned}$$

Esta última igualdad —que resulta ser extraordinariamente importante— es conocida como la *identidad de Jacobi*, recordando a *Carl Gustav Jacob Jacobi*.

La primera parte de la Proposición 2.6.16 que probamos arriba nos dice que el núcleo de la función traza $\text{tr} : M_n(\mathbb{k}) \rightarrow \mathbb{k}$ está generado como espacio vectorial por comutadores, esto es, que toda matriz de traza nula es combinación lineal de comutadores. De hecho, es cierto que toda matriz de traza nula es un comutador: esto fue probado en 1937 por *Kenjiro Shoda* en [Sho37] cuando el cuerpo \mathbb{k} tiene característica nula y en 1957 por *Abraham Albert* y *Benjamin Muckenhoupt* en general.

§7. Proyectores

2.7.1. Sea V un espacio vectorial. Un *proyector* es una función lineal $f : V \rightarrow V$ tal que $f^2 = f$. Decimos también que la función lineal f es *idempotente* en ese caso.

2.7.2. La propiedad más importante de los proyectores es la que explicita la siguiente proposición: un proyector nos da una descomposición en suma directa del espacio.

Proposición. Sea V un espacio vectorial. Si $f : V \rightarrow V$ es un proyector, entonces

$$V = \text{Nu}(f) \oplus \text{Im}(f).$$

Más aún, un vector $x \in V$ pertenece a $\text{Im}(f)$ si y solamente si $f(x) = x$.

Demostración. Sea $f : V \rightarrow V$ un proyector. Si $x \in V$, entonces

$$f(x - f(x)) = f(x) - f(f(x)) = f(x) - f(x) = 0,$$

así que $x - f(x) \in \text{Nu}(f)$ y, por lo tanto, $x = (x - f(x)) + f(x) \in \text{Nu}(f) + \text{Im}(f)$. Esto nos dice que $V = \text{Nu}(f) + \text{Im}(f)$. Por otro lado, si $x \in \text{Nu}(f) \cap \text{Im}(f)$, entonces $f(x) = 0$ porque $x \in \text{Nu}(f)$ y existe $y \in V$ tal que $x = f(y)$ como $x \in \text{Im}(f)$. Se sigue de esto que y de que $f \circ f = f$ que $x = f(y) = f(f(y)) = f(x) = 0$. Vemos entonces que $\text{Nu}(f) \cap \text{Im}(f) = 0$ y, por lo tanto, que $V = \text{Nu}(f) \oplus \text{Im}(f)$, como afirma la proposición.

Sea $x \in X$. Si $f(x) = x$, es claro que $x \in \text{Im}(f)$. Recíprocamente, si $x \in \text{Im}(f)$, entonces existe $y \in V$ tal que $x = f(y)$, y entonces $f(x) = f(f(y)) = f(y) = x$. \square

2.7.3. La primera parte del siguiente resultado nos da una especie de recíproca de la Proposición 2.7.2, mientras que la segunda nos da un criterio sencillo para verificar que dos proyectores son iguales.

Proposición. Sea V un espacio vectorial.

- (i) Si N e I son dos subespacios de V tales que $V = N \oplus I$, entonces existe un proyector $f : V \rightarrow V$ y uno solo tal que $N = \text{Nu}(f)$ e $I = \text{Im}(f)$.
- (ii) Dos proyectores $f, g : V \rightarrow V$ son iguales si y solamente si $\text{Nu}(f) = \text{Nu}(g)$ e $\text{Im}(f) = \text{Im}(g)$.

Demostración. (i) Sean N e I subespacios de V tales que $V = N \oplus I$. Si $x \in V$, entonces existen $y_x \in N$ y $z_x \in I$ tales que $x = y_x + z_x$, y están únicamente determinados por x : hay entonces una función $f : x \in V \mapsto z_x \in V$. Mostremos que f es una función lineal que es un proyector y que su núcleo y su imagen son N e I , respectivamente.

- Sean x y x' elementos de V . Como $x = y_x + z_x$ y $x' = y_{x'} + z_{x'}$, tenemos que

$$x + y = (y_x + y_{x'}) + (z_x + z_{x'}),$$

con $y_x + y_{x'} \in N$ y $z_x + z_{x'} \in I$: esto nos dice que

$$f(x + x') = z_x + z_{x'} = f(x) + f(x').$$

Por otro lado, si $\lambda \in \mathbb{k}$, entonces $\lambda x = \lambda y_x + \lambda z_x$, con $\lambda y_x \in N$ y $\lambda z_x \in I$, así que

$$f(\lambda x) = \lambda z_x = \lambda f(x).$$

Podemos concluir con todo esto que la función f es lineal.

- Sea $x \in V$. Como $z_x \in I$, es $y_{z_x} = 0$ y $z_{z_x} = z_x$, así como $f(x) = z_x$, es que

$$f(f(x)) = z_{z_x} = z_x = f(x).$$

Esto nos dice que la función f es un proyector.

- Si $x \in N$, entonces claramente $y_x = x$ y $z_x = 0$, así que $f(x) = 0$: esto significa que $N \subseteq \text{Nu}(f)$. Por otro lado, si $x \in \text{Nu}(f)$, entonces $f(x) = z_x = 0$ y, por lo tanto, $x = y_x + z_x = y_x \in N$. Vemos así que el núcleo de f es $\text{Nu}(f) = N$.
- Si $x \in I$, entonces $y_x = 0$ y $z_x = x$, así que $x = f(x) \in \text{Im}(f)$. Recíprocamente, si $x \in \text{Im}(f)$, entonces existe $x' \in V$ tal que $x = f(x')$ y, por lo tanto, $x = z_{x'} \in I$. De esto deducimos que la imagen de f es $\text{Im}(f) = I$.

Nos queda probar la afirmación de unicidad. Sea $g : V \rightarrow V$ un proyector que tiene núcleo e imagen N e I , respectivamente. Si $x \in V$, entonces tenemos que $x - g(x) \in \text{Nu}(g) = N$, $g(x) \in \text{Im}(g) = I$, y $x = (x - g(x)) + g(x)$, así que $y_x = x - g(x)$ y $z_x = g(x)$: se sigue de esto, claro, que $f(x) = z_x = g(x)$. Vemos así que las funciones f y g son la misma función.

(ii) Es evidente que la condición es necesaria, así que bastará que probemos su suficiencia. Sean $f, g : V \rightarrow V$ dos proyectores tales que $\text{Nu}(f) = \text{Nu}(g)$ y $\text{Im}(f) = \text{Im}(g)$, y sea $x \in V$. Existen $y \in \text{Nu}(f)$ y $z \in \text{Im}(f)$ tales que $x = y + z$, ya que $V = \text{Nu}(f) \oplus \text{Im}(f)$. Es $f(x) = f(y) + f(z) = f(z) = z$, porque $f(y) = 0$ y $f(z) = z$. De manera similar, y usando la hipótesis, tenemos que $g(x) = g(y) + g(z) = g(z) = z$, así que $f(x) = g(x)$. \square

2.7.4. Las Proposiciones 2.7.2 y 2.7.3 juntas tienen la siguiente consecuencia:

Corolario. Sea V un espacio vectorial y sean

- $\text{Proy}(V)$ el conjunto de todos los proyectores $V \rightarrow V$ y
- $\mathcal{D}_2(V)$ el conjunto de todos los pares ordenados (N, I) de subespacios de V tales que $V = N \oplus I$.

Hay una función

$$f \in \text{Proy}(V) \mapsto (\text{Nu}(f), \text{Im}(f)) \in \mathcal{D}_2(V)$$

y es una biyección.

Esto nos dice que es esencialmente lo mismo tener un proyector de V que una descomposición de V como suma directa de dos subespacios.

Demostración. La Proposición 2.7.2 nos dice que si $f \in \text{Proy}_2(V)$ entonces $(\text{Nu}(f), \text{Im}(f))$ es un elemento de $\mathcal{D}_2(V)$, de manera que, en particular, existe una función como la descripta en el enunciado. La primera parte de la Proposición 2.7.3 afirma precisamente que esta función es sobreyectiva y la segunda que es inyectiva. \square

2.7.5. La Proposición 2.7.3 nos dice que si tenemos una descomposición en suma directa $V = N \oplus I$, hay un proyector $f : V \rightarrow V$ que tiene núcleo e imagen N e I , respectivamente. Dado f , es fácil construir otro proyector $g : V \rightarrow V$ que tiene núcleo e imagen I y N : basta considerar el

homomorfismo $g = \text{id}_V - f$. En efecto, g es un proyector porque

$$\begin{aligned} g \circ g &= (\text{id}_V - f) \circ (\text{id}_V \circ f) \\ &= \text{id}_V \circ \text{id}_V - f \circ \text{id}_V - \text{id}_V \circ f + f \circ f \\ &= \text{id}_V - f - f + f \\ &= \text{id}_V - f \\ &= g. \end{aligned}$$

Por otro lado, si $x \in V$, es

$$x \in \text{Nu}(g) \iff g(x) = x - f(x) = 0 \iff f(x) = x \iff x \in \text{Im}(f),$$

de manera que $\text{Nu}(g) = \text{Im}(f) = I$. De manera similar, si $x \in \text{Im}(g)$, entonces existe $y \in V$ tal que $x = g(y) = y - f(y)$, así que $f(x) = f(y) - f(f(y)) = f(y) - f(y) = 0$ y $x \in \text{Nu}(f)$. Recíprocamente, si $x \in \text{Nu}(f)$, entonces $x = x - f(x) = g(x) \in \text{Im}(g)$. Esto nos dice que $\text{Im}(g) = \text{Nu}(f) = N$, como dijimos.

Calculando, podemos ver los morfismos f y g satisfacen las identidades

$$f \circ f = f, \quad g \circ g = g, \quad f + g = \text{id}_V, \quad f \circ g = g \circ f = 0.$$

La siguiente proposición generaliza esto.

2.7.6. Proposición. Sea V un espacio vectorial y sea $n \in \mathbb{N}$.

(i) Si U_1, \dots, U_n son subespacios de V tales que $V = U_1 \oplus \dots \oplus U_n$, entonces existen proyectores $f_1, \dots, f_n : V \rightarrow V$ únicamente determinados tales que

$$\begin{aligned} f_i \circ f_j &= 0 && \text{para cada } i, j \in [\![n]\!] \text{ tales que } i \neq j, \\ f_1 + \dots + f_n &= \text{id}_V, \\ \text{Im}(f_i) &= U_i && \text{para cada } i \in [\![n]\!]. \end{aligned}$$

(ii) Recíprocamente, si $f_1, \dots, f_n : V \rightarrow V$ son proyectores tales que

$$\begin{aligned} f_i \circ f_j &= 0 && \text{para cada } i, j \in [\![n]\!] \text{ tales que } i \neq j, \\ f_1 + \dots + f_n &= \text{id}_V, \end{aligned}$$

entonces $V = \text{Im}(f_1) \oplus \dots \oplus \text{Im}(f_n)$.

Demostración. (i) Sean U_1, \dots, U_n subespacios de V tales que $V = U_1 \oplus \dots \oplus U_n$. Para cada $x \in V$ existen $x_1 \in U_1, \dots, x_n \in U_n$, únicamente determinados por x , tales que $x = x_1 + \dots + x_n$. Hay entonces funciones $f_1, \dots, f_n : V \rightarrow V$ tales que para cada $x \in V$ y cada $i \in [\![n]\!]$ es $f_i(x) = x_i$. Verifiquemos que estas funciones satisfacen las condiciones del enunciado.

- Sea $i \in [\![n]\!]$. Si x e y son elementos de V , entonces tenemos claramente que

$$x + y = (x_1 + y_1) + \dots + (x_n + y_n),$$

con $x_j + y_j \in U_j$ para cada $j \in \llbracket n \rrbracket$, así que $f_i(x + y) = x_i + y_i = f_i(x) + f_i(y)$. De manera similar, si $x \in V$ y $\lambda \in \mathbb{k}$, entonces $\lambda x = \lambda x_1 + \dots + \lambda x_n$, con $\lambda x_j \in U_j$ para cada $j \in \llbracket n \rrbracket$, así que $f_i(\lambda x) = \lambda x_i = \lambda f_i(x)$. Esto muestra que la función f_i es lineal. Se trata, además, de un proyector. En efecto, si $x \in V$, entonces $x_i = 0 + \dots + x_i + \dots + 0$, esto es, $(x_i)_j = 0$ si $j \in \llbracket n \rrbracket$ es tal que $i \neq j$, y $(x_i)_i = x_i$, de manera que $f_i(f_i(x)) = x_i = f(x)$: vemos de esta forma que $f_i \circ f_i = f_i$.

- De la definición de la función f_i es claro que $\text{Im}(f_i) \subseteq U_i$. Por otro lado, si $x \in U_i$, entonces es $x_j = 0$ si $j \in \llbracket n \rrbracket$ es distinto de i y $x_i = x$, así que $x = f_i(x) \in \text{Im}(f_i)$. Vemos con esto que $\text{Im}(f_i) = U_i$.
- Si i y j son elementos distintos de $\llbracket n \rrbracket$ y $x \in V$, entonces $f_j(x) \in U_j$, así que $f_i(f_j(x)) = 0$. Esto nos dice que $f_i \circ f_j = 0$.
- Finalmente, si $x \in V$ y $x = x_1 + \dots + x_n$ con $x_1 \in U_1, \dots, x_n \in U_n$, entonces $f_i(x) = x_i$ para cada $i \in \llbracket n \rrbracket$, así que

$$(f_1 + \dots + f_n)(x) = f_1(x) + \dots + f_n(x) = x_1 + \dots + x_n = x.$$

Vemos así que $f_1 + \dots + f_n = \text{id}_V$.

Para ver la unicidad, supongamos que $f'_1, \dots, f'_n : V \rightarrow V$ es otra elección de proyectores tales que $f'_i \circ f'_j = 0$ si i y j son elementos distintos de $\llbracket n \rrbracket$ e $\text{Im}(f'_i) = U_i$ para cada $i \in \llbracket n \rrbracket$, y mostremos que se tiene que $f'_i = f_i$ para cada $i \in \llbracket n \rrbracket$. Para ello, como f_i y f'_i son proyectores, es suficiente que mostremos que sus núcleos coinciden y que sus imágenes coinciden. Como $\text{Im}(f_i) = U_i = \text{Im}(f'_i)$, bastará que nos ocupemos de los núcleos.

Sea $i \in \llbracket n \rrbracket$. Si j es un elemento de $\llbracket n \rrbracket$ distinto de i , entonces

$$f_i(U_j) = f_i(\text{Im}(f_j)) = \text{Im}(f_i \circ f_j) = 0.$$

Esto nos dice que el subespacio $U_1 + \dots + \widehat{U_i} + \dots + U_n$ está contenido en $\text{Nu}(f_i)$. Como ambos subespacios de V son complementos de $U_i = \text{Im}(f_i)$, se sigue de esto y del Lema 1.10.3 que, de hecho, esos subespacios son iguales. Exactamente el mismo argumento prueba que $\text{Nu}(f'_i)$ es igual a $U_1 + \dots + \widehat{U_i} + \dots + U_n$, así que $\text{Nu}(f_i) = \text{Nu}(f'_i)$, como queremos.

(ii) Supongamos ahora que $f_1, \dots, f_n : V \rightarrow V$ son proyectores tales que $f_i \circ f_j$ siempre que i y j son elementos distintos de $\llbracket n \rrbracket$ y $f_1 + \dots + f_n = \text{id}_V$. Esta última igualdad implica que si $x \in V$ se tiene que

$$x = \text{id}_V(x) = f_1(x) + \dots + f_n(x) \in \text{Im}(f_1) + \dots + \text{Im}(f_n)$$

y, por lo tanto, que $V = \text{Im}(f_1) + \dots + \text{Im}(f_n)$. Mostremos que esta suma es directa.

Sea $i \in \llbracket n \rrbracket$ y sea x un vector que está tanto en $\text{Im}(f_i)$ como en $\text{Im}(f_1) + \dots + \widehat{\text{Im}(f_i)} + \dots + \text{Im}(f_n)$. Esto significa que existen vectores $y_1, \dots, y_n \in V$ tales que $x = f_i(y_i)$ y

$$x = f_1(y_1) + \dots + \widehat{f_i(y_i)} + \dots + f_n(y_n). \tag{10}$$

Si $j \in \llbracket n \rrbracket$ es distinto de i , entonces

$$\begin{aligned} 0 &= f_j(f_i(y_i)) \\ &= f_j(x) \\ &= f_j(f_1(y_1) + \cdots + \widehat{f_i(y_i)} + \cdots + f_n(y_n)) \\ &= f_j(f_1(y_1)) + \cdots + \widehat{f_j(f_i(y_i))} + \cdots + f_j(f_n(y_n)) \\ &= f_j(f_j(y_j)) \\ &= f_j(y_j). \end{aligned}$$

Esto nos dice que todos los sumandos que aparecen a la derecha de la igualdad (10) son nulos, así que $x = 0$. Vemos así que $\text{Im}(f_i) \cap (\text{Im}(f_1) + \cdots + \widehat{\text{Im}(f_i)} + \cdots + \text{Im}(f_n)) = 0$. \square

2.7.7. Si V es un espacio vectorial y $n \in \mathbb{N}$, una n -upla (f_1, \dots, f_n) de funciones lineales $V \rightarrow V$ es un *sistema completo de proyectores ortogonales dos a dos* en V si

- los homomorfismos f_1, \dots, f_n son proyectores,
- si $i \neq j$ son elementos distintos de $\llbracket n \rrbracket$, es $f_i \circ f_j = 0$, y
- $f_1 + \cdots + f_n = \text{id}_V$.

Usando esta noción, podemos dar el siguiente corolario de la Proposición 2.7.6, que es, de hecho, equivalente a ella.

Corolario. Sea V un espacio vectorial, sea $n \in \mathbb{N}$ y consideremos

- el conjunto $\mathcal{S}_n(V)$ de los sistemas completos de proyectores ortogonales dos a dos (f_1, \dots, f_n) en V , y
- el conjunto $\mathcal{D}_n(V)$ de las n -uplas (U_1, \dots, U_n) de subespacios de V tales que $V = U_1 \oplus \cdots \oplus U_n$.

Si (f_1, \dots, f_n) es un elemento de $\mathcal{S}_n(V)$, entonces $(\text{Im}(f_1), \dots, \text{Im}(f_n))$ es un elemento de $\mathcal{D}_n(V)$, y la función

$$(f_1, \dots, f_n) \in \mathcal{S}_n(V) \mapsto (\text{Im}(f_1), \dots, \text{Im}(f_n)) \in \mathcal{D}_n(V)$$

es una biyección.

Demostración. La segunda parte de la Proposición 2.7.6 nos dice que la primera afirmación del corolario es cierta, mientras que la primera parte de esa proposición nos dice que la función es biyectiva. \square

§8. Cocientes

2.8.1. Sea V un espacio vectorial y sea W un subespacio de V . Decimos que dos vectores x e y de V son **congruentes módulo W** si $x - y \in W$ y en ese caso escribimos $x \equiv y$ o, si queremos poner de manifiesto al subespacio W , $x \equiv_W y$. Esto define una relación en el conjunto V y se trata, de hecho, de una relación de equivalencia:

- Es reflexiva: si $x \in V$, entonces $x - x = 0 \in W$, así que $x \equiv x$.
- Es simétrica: si $x, y \in V$ son tales que $x \equiv y$, de manera que $x - y \in W$, entonces tenemos que $y - x = -(x - y) \in W$ y, por lo tanto, que $y \equiv x$.
- Es transitiva: si $x, y, z \in V$ son tales que $x \equiv y$ e $y \equiv z$, entonces $x - y$ e $y - z$ son elementos de W y $x - z = (x - y) + (y - z)$ también, lo que implica que $x \equiv z$.

Podemos, por lo tanto, considerar el conjunto cociente de V por esta relación de congruencia: lo escribimos V/W . Si x es un vector de V , escribiremos $[x]$ a su clase de equivalencia: así, $[x] = \{y \in V : x \equiv y\}$; si necesitamos explicitar el subespacio W , escribimos $[x]_W$. Por otro lado, si u es un elemento de V/W y x es un vector de V que está en u , de manera que $u = [x]$, decimos que x es un **representante** de u en V .

Queremos hacer del conjunto V/W un espacio vectorial y para ello tenemos que construir operaciones de suma y de multiplicación por escalares.

- Si x, x', y, y' son elementos de V tales que $x \equiv x'$ e $y \equiv y'$, entonces

$$(x + y) - (x' + y') = (x - x') + (y - y') \in W,$$

así que $x + y \equiv x' + y'$. Se sigue de esto que hay una función

$$+ : V/W \times V/W \rightarrow V/W$$

tal que cada vez que x e y son elementos de W se tiene que $[x] + [y] = [x + y]$.

- Si x e y son elementos de V tales que $x \equiv y$ y λ es un escalar, entonces $\lambda x \equiv \lambda y$, ya que $\lambda x - \lambda y = \lambda(x - y) \in W$. Esto implica que hay una operación

$$\cdot : \mathbb{k} \times V/W \rightarrow V/W$$

tal que cada vez que $\lambda \in \mathbb{k}$ y $x \in V$ vale que $\lambda \cdot [x] = [\lambda x]$.

2.8.2. Proposición. *Sea V un espacio vectorial y sea W un subespacio de V . El conjunto V/W dotado con las operaciones de suma y producto por escalares de \mathbb{k} construidas arriba es un espacio vectorial. La función $\pi : x \in V \mapsto [x] \in V/W$ es una función lineal sobreductiva cuyo núcleo es precisamente el subespacio W .*

Siempre que en la situación de esta proposición consideremos el conjunto V/W será dotado de esta estructura de espacio vectorial. Lo llamamos el **espacio cociente de V por W** . La función π es la **proyección canónica** de V a V/W .

Demostración. Para ver que V/W es un espacio vectorial, tenemos que mostrar que se cumplen las condiciones de la definición 1.2.1.

- La suma es asociativa: si u, v y w son elementos de V/W , entonces existen $x, y, z \in V$ tales que $u = [x]$, $v = [y]$ y $w = [z]$, y

$$\begin{aligned}(u + v) + w &= ([x] + [y]) + [z] = [x + y] + [z] = [(x + y) + z] \\ &= [x + (y + z)] = [x] + [y + z] = [x] + ([y] + [z]) \\ &= u + (v + w).\end{aligned}$$

- La suma es conmutativa: si u y v son elementos de V/W , entonces existen $x, y \in V$ tales que $u = [x]$ y $v = [y]$, y se tiene que

$$u + v = [x] + [y] = [x + y] = [y + x] = [y] + [x] = v + u.$$

- La clase $[0]$ es un elemento neutro para la suma: si u es un elemento de V/W , existe $x \in V$ tal que $c = [x]$ y entonces

$$[0] + u = [0] + [x] = [0 + x] = [x] = u$$

y

$$u + x = [x] + [0] = [x + 0] = [x] = u.$$

- Todo elemento de V/W posee un opuesto: si u es un elemento de V/W , entonces existe $x \in V$ tal que $u = [x]$ y

$$u + [-x] = [x] + [-x] = [x + (-x)] = [0]$$

y

$$[-x] + u = [-x] + [x] = [(-x) + x] = [0],$$

así que $[-x]$ es un opuesto para u .

- La multiplicación por escalares es asociativa: si $a, b \in \mathbb{k}$ y $u \in V/W$, entonces existe $x \in V$ tal que $u = [x]$ y

$$\begin{aligned}a \cdot (b \cdot u) &= a \cdot (b \cdot [x]) = a \cdot [b \cdot x] = [a \cdot (b \cdot x)] = [(a \cdot b) \cdot x] \\ &= (a \cdot b) \cdot [x] = (a \cdot b) \cdot u.\end{aligned}$$

- La multiplicación por escalares es unitaria: si $u \in V/W$, entonces existe $x \in V$ tal que $u = [x]$ y

$$1 \cdot u = 1 \cdot [x] = [1 \cdot x] = [x] = u.$$

- La multiplicación por escalares se distribuye sobre la suma de V/W : si $a \in \mathbb{k}$ y $u, v \in V/W$, entonces existen $x, y \in V$ tales que $u = [x]$ y $v = [y]$, y

$$\begin{aligned}a \cdot (u + v) &= a \cdot ([x] + [y]) = a \cdot [x + y] = [a(x + y)] = [a \cdot x + a \cdot y] \\ &= [a \cdot x] + [a \cdot y] = a \cdot [x] + a \cdot [y] = a \cdot u + a \cdot v.\end{aligned}$$

- La multiplicación por escalares se distribuye sobre la suma de \mathbb{k} : si $a, b \in \mathbb{k}$ y $u \in V/W$, entonces existe $x \in V$ tal que $u = [x]$ y

$$\begin{aligned}(a+b) \cdot u &= (a+b) \cdot [x] = [(a+b) \cdot x] = [a \cdot x + b \cdot x] = [a \cdot x] + [b \cdot x] \\ &= a \cdot [x] + b \cdot [x] = a \cdot u + b \cdot u.\end{aligned}$$

Para ver que la función π del enunciado es una función lineal, observamos que si $\alpha, \beta \in \mathbb{k}$ son escalares y $x, y \in V$, entonces

$$\begin{aligned}\pi(\alpha \cdot x + \beta \cdot y) &= [\alpha \cdot x + \beta \cdot y] = [\alpha \cdot x] + [\beta \cdot y] = \alpha \cdot [x] + \beta \cdot [y] \\ &= \alpha \cdot \pi(x) + \beta \cdot \pi(y).\end{aligned}$$

Si $x \in V$ está en el núcleo de π , entonces su imagen $\pi(x) = [x]$ es el elemento neutro de V/W , es decir, $[x] = [0]$. Esto significa que $x \equiv 0$ y, por lo tanto, que $x = x - 0 \in W$: vemos así que $\text{Nu}(\pi) \subseteq W$. Recíprocamente, si $x \in W$, entonces $x - 0 \in W$, de manera que $x \equiv 0$ y, por lo tanto, $\pi(x) = [x] = [0]$, es decir, $x \in \text{Nu}(\pi)$. \square

2.8.3. Una consecuencia inmediata de lo que acabamos de probar es:

Corolario. *Sea V un espacio vectorial. Todo subespacio de V es el núcleo de una función lineal de dominio V .*

Demostración. Si W es un subespacio de V , entonces la proyección $\pi : V \rightarrow V/W$ es lineal y su núcleo es W , y esto prueba el corolario. \square

2.8.4. Hagamos una observación trivial pero útil:

Lema. *Sea V un espacio vectorial, sea W un subespacio de V y sea $\pi : V \rightarrow V/W$ la proyección canónica al cociente.*

- (i) *El cociente V/W es un espacio nulo si y solamente si $W = V$.*
- (ii) *La proyección $\pi : V \rightarrow V/W$ es un isomorfismo si y solamente si $W = 0$.*

Demostración. El espacio V/W es nulo si y solamente si para todo $v \in V$ se tiene que $[x] = [0]$, es decir, si y solamente si $x = x - 0 \in W$. Por otro lado, como la proyección π siempre es sobreyectiva, se trata de un isomorfismo si y solamente si es inyectiva, y esto ocurre si y solamente si su núcleo, que es precisamente W , es nulo. \square

2.8.5. Notemos que de acuerdo a este lema, si V es un espacio vectorial la proyección canónica $\pi : V \rightarrow V/0$ al cociente de V por su subespacio nulo es un isomorfismo. Muchas veces veremos a esta función como una identificación, esto es, consideraremos a V y a $V/0$ como el mismo espacio, identificando cada vector v de V con la clase $\pi(v) = [v]$ de $V/0$, que como conjunto es simplemente $\{v\}$.

2.8.6. El siguiente resultado es lo que nos va a permitir en casi todos los casos construir funciones lineales cuyo dominio es un cociente:

Proposición. *Sea V un espacio vectorial, sea W un subespacio de V y sea $\pi : V \rightarrow V/W$ la proyección canónica al cociente. Si U es un espacio vectorial y $f : V \rightarrow U$ es una función lineal tal que $W \subseteq \text{Nu}(f)$, entonces existe exactamente una función lineal $\tilde{f} : V/W \rightarrow U$ tal que $f = \tilde{f} \circ \pi$, de manera que para cada $x \in V$ es $\tilde{f}([x]) = f(x)$.*

Demostración. Sea U un espacio vectorial y sea $f : V \rightarrow U$ una función lineal tal que $W \subseteq \text{Nu}(f)$. Para cada $x, y \in V$ vale que

$$x \equiv y \implies f(x) = f(y).$$

En efecto, si $x \equiv y$ entonces $x - y \in W \subseteq \text{Nu}(f)$ y por lo tanto $f(x) - f(y) = f(x - y) = 0$. Consideremos ahora el subconjunto

$$\tilde{f} = \{([x], f(x)) \in V/W \times U : x \in V\}$$

de $V/W \times U$. Mostremos que se trata, de hecho, de una función $\tilde{f} : V/W \rightarrow U$.

- Si $u \in V/W$, entonces existe $x \in V$ tal que $u = [x]$, y entonces el par $([x], f(x))$, cuya primera componente es u , está en \tilde{f} .
- Supongamos ahora que (u, y) y (u, y') son dos elementos de \tilde{f} cuyas primeras componentes coinciden. Esto significa que existen vectores x y x' en V tales que $(u, y) = ([x], f(x))$ y $(u, y') = ([x'], f(x'))$. En particular, tenemos que $[x] = u = [x']$, así que $x \equiv x'$ y, por lo que observamos al principio, $f(x) = f(x')$, esto es, $y = y'$.

Notemos que para cada $x \in V$ se tiene que $\tilde{f}([x]) = f(x)$, y esto nos dice que $\tilde{f} \circ \pi = f$. Además de esto, la función \tilde{f} es lineal. En efecto, si $\alpha, \beta \in \mathbb{k}$ y $u, v \in V/W$, existen $x, y \in V$ tales que $u = [x]$ y $v = [y]$, y entonces

$$\begin{aligned} \tilde{f}(\alpha u + \beta v) &= \tilde{f}(\alpha[x] + \beta[y]) = \tilde{f}([\alpha x + \beta y]) = f(\alpha x + \beta y) = \alpha f(x) + \beta f(y) \\ &= \alpha \tilde{f}([x]) + \beta \tilde{f}([y]) = \alpha \tilde{f}(u) + \beta \tilde{f}(v). \end{aligned}$$

Finalmente, supongamos que $g : V/W \rightarrow U$ es otra función tal que $g \circ \pi = f$. Si $u \in V/W$, entonces existe $x \in V$ tal que $u = [x]$ y entonces

$$g(u) = g(\pi(x)) = f(x) = \tilde{f}(\pi(x)) = \tilde{f}(u).$$

Así, es $g = f$ y, en consecuencia, la función \tilde{f} está únicamente determinada. \square

2.8.7. La Proposición 2.8.6 admite varias variaciones útiles. Mencionemos una que nos será útil más tarde.

Proposición. Sean U , V y W tres espacios vectoriales y sean U' y V' subespacios de U y de V , respectivamente. Si $f : U \times V \rightarrow W$ es una función bilineal tal que cada vez que $u \in U$ y $v \in V$ se tiene que

$$f(u, v) = 0 \text{ si } u \in U' \text{ o } v \in V', \quad (11)$$

entonces existe exactamente una función bilineal $\tilde{f} : U/U' \times V/V' \rightarrow W$ tal que

$$\tilde{f}([u]_{U'}, [v]_{V'}) = f(u, v)$$

cualesquiera sean $u \in U$ y $v \in V$.

Demostración. Sea $f : U \times V \rightarrow W$ una función bilineal que satisface la condición del enunciado y consideremos el subconjunto

$$\tilde{f} = \{(([u]_{U'}, [v]_{V'}), f(u, v)) \in (U/U' \times V/V') \times W : u \in U, v \in V\}$$

del producto cartesiano $(U/U' \times V/V') \times W$. Se trata de una función $U/U' \times V/V' \rightarrow W$.

- Si $x \in U/U'$ e $y \in V/V'$, entonces existen $u \in U$ y $v \in V$ tales que $x = [u]_{U'}$ e $y = [v]_{V'}$ y por lo tanto $((x, y), f(u, v)) =(([u]_{U'}, [v]_{V'}), f(u, v)) \in \tilde{f}$.
- Sean ahora $x \in U/U'$ e $y \in V/V'$, por un lado, $y w_1, w_2 \in W$, por otro, tales que los pares ordenados $((x, y), w_1)$ y $((x, y), w_2)$ están en \tilde{f} . Esto significa que existen $u_1, u_2 \in U$ y $v_1, v_2 \in V$ tales que $x = [u_1]_{U'} = [u_2]_{U'}$, $y = [v_1]_{V'} = [v_2]_{V'}$, $w_1 = f(u_1, v_1)$ y $w_2 = f(u_2, v_1)$. Como $[u_1]_{U'} = [u_2]_{U'}$ y $[v_1]_{V'} = [v_2]_{V'}$, tenemos que $u_2 - u_1 \in U'$ y $v_2 - v_1 \in V'$, así que

$$\begin{aligned} w_2 &= f(u_2, v_2) \\ &= f(u_2 - u_1 + u_1, v_2 - v_1 + v_1) \\ &= f(u_2 - u_1, v_2 - v_1) + f(u_1, v_2 - v_1) + f(u_2 + u_1, v_2) + f(u_1, u_2) \\ &= f(u_1, v_1) = w_1, \end{aligned}$$

ya que f es bilineal y satisface la condición (11)

Tenemos entonces, como dijimos, una función $\tilde{f} : U/U' \times V/V' \rightarrow W$, y es claro, de acuerdo a su definición, que

$$\tilde{f}([u]_{U'}, [v]_{V'}) = f(u, v)$$

para todo $u \in U$ y todo $v \in V$. Mostremos que además es bilineal.

- Si $x_1, x_2 \in U/U'$, $y \in V/V'$ y $a, b \in \mathbb{k}$, entonces existen $u_1, u_2 \in U$ y $v \in V$ tales que

$x_1 = [u_1]_{U'}$, $x_2 = [u_2]_{U'}$ y $y = [v]_{V'}$, de manera que

$$\begin{aligned}\bar{f}(ax_1 + bx_2, y) &= \bar{f}(a[u_1] + b[u_2], [v]) \\ &= \bar{f}([au_1 + bu_2], [v]) \\ &= f(au_1 + bu_2, v) \\ &= af(u_1, v) + bf(u_2, v) \\ &= a\bar{f}([u_1], [v]) + b\bar{f}([u_2], [v]) \\ &= a\bar{f}(x_1, y) + b\bar{f}(x_2, y),\end{aligned}$$

así que \bar{f} es lineal con respecto a su primer argumento. Un razonamiento completamente análogo prueba que también es lineal con respecto al segundo.

Para terminar, tenemos que probar que la unicidad que se afirma en la proposición. Sea entonces $g : U/U' \times V/V' \rightarrow W$ una función bilineal tal que $g([u]_{U'}, [v]_{V'}) = f(u, v)$ siempre que $u \in U$ y $v \in V$. Si $x \in U/U'$ e $y \in V/V'$, entonces existen $u \in U$ y $v \in V$ tales que $x = [u]_{U'}$ e $y = [v]_{V'}$ y por lo tanto,

$$g(x, y) = g([u]_{U'}, [v]_{V'}) = f(u, v) = \bar{f}([u]_{U'}, [v]_{V'}) = \bar{f}(x, y).$$

Vemos con esto que $g = \bar{f}$, como queremos. \square

2.8.8. Hay una relación estrecha entre el cociente de un espacio por un subespacio y los complementos de este:

Proposición. *Sea V un espacio vectorial, sea W un subespacio de V y sea $\pi : V \rightarrow V/W$ la proyección canónica al cociente. Si C es un complemento de W en V , entonces la restricción $\pi|_C : C \rightarrow V/W$ de la proyección π a C es un isomorfismo.*

Demostración. Sea C un complemento de W en V , de manera que tenemos una descomposición $V = W \oplus C$, y sea $\pi|_C : C \rightarrow V/W$ la restricción de π a C . Si $u \in V/W$ y $x \in V$ es tal que $u = [x]$, entonces existen $w \in W$ y $c \in C$ con $x = w + c$ y, como $x - c = w \in W$, se tiene que $\pi(c) = [c] = [x] = u$. Esto nos dice que la restricción $\pi|_C$ es sobreyectiva. Por otro lado, si $c \in C$ es tal que $\pi|_C(c) = [c]$ es el cero $[0]$ de V/W , entonces $c - 0 \in W$, es decir, $c \in W$: como $V = W \oplus C$, vemos que $c \in W \cap C = 0$ y, en definitiva, que $\pi|_C$ es un isomorfismo. \square

2.8.9. Una aplicación importante de la proposición que acabamos de probar es la determinación de la dimensión de un cociente, cuando esta es finita.

Proposición. *Sea V un espacio vectorial, sea W un subespacio de V y sea $\pi : V \rightarrow V/W$ la proyección canónica al cociente.*

(i) *El subespacio W tiene codimensión finita en V si y solamente si el cociente V/W tiene dimensión finita y en ese caso se tiene que*

$$\dim V/W = \text{codim}_V W.$$

(ii) Si V tiene dimensión finita, entonces V/W también y

$$\dim V/W = \dim V - \dim W.$$

Demostración. (i) Supongamos primero que W tiene codimensión finita en V y fijemos un complemento C , necesariamente de dimensión finita, de W en V . De acuerdo a la Proposición 2.8.8, hay un isomorfismo $C \cong V/W$ así que por un lado V/W tiene dimensión finita, porque C tiene dimensión finita, y, por otro, tenemos que $\text{codim}_V W = \dim C = \dim V/W$.

Supongamos ahora que el cociente V/W tiene dimensión finita, sea n su dimensión y sea $\mathcal{B} = \{u_1, \dots, u_n\}$ una base de $V(W)$. Para cada $i \in \llbracket n \rrbracket$ sea $x_i \in V$ tal que $u_i = [x_i]$, consideremos el subespacio $C = \langle x_1, \dots, x_n \rangle$ de V y mostremos que $W + C = V$: esto probará que W tiene codimensión finita en V . Sea $x \in V$. Como \mathcal{B} es una base de V/W , existen escalares $a_1, \dots, a_n \in \mathbb{k}$ tales que $\pi(x) = a_1u_1 + \dots + a_nu_n$. De esto se sigue que

$$[x] = \pi(x) = a_1u_1 + \dots + a_nu_n = a_1[x_1] + \dots + a_n[x_n] = [a_1x_1 + \dots + a_nx_n],$$

de manera que el vector $w := x - (a_1x_1 + \dots + a_nx_n)$ pertenece a W y, por lo tanto,

$$x = w + (a_1x_1 + \dots + a_nx_n) \in W + C,$$

como queremos.

(ii) Si V tiene dimensión finita, entonces V/W tiene también, ya que a función $\pi : V \rightarrow V/W$ es sobreyectiva, y la Proposición 1.10.9 nos dice que el subespacio W tiene codimensión finita y que $\text{codim}_V W = \dim V - \dim W$: la igualdad del enunciado es entonces consecuencia inmediata de la parte (i). \square

El conúcleo de una función lineal

2.8.10. Si $f : V \rightarrow W$ es una función lineal, la imagen $\text{Im}(f)$ es un subespacio de W y podemos entonces considerar el espacio cociente $W/\text{Im}(f)$, al que llamamos el **conúcleo** de la función f y escribimos $\text{Con}(f)$. Como se trata de un cociente de W , tenemos disponible la proyección canónica $\pi : W \rightarrow \text{Con}(f)$.

2.8.11. Para construir funciones lineales con dominio en un conúcleo usamos la Proposición 2.8.6:

Proposición. Sean V y W espacios vectoriales, sea $f : V \rightarrow W$ una función lineal y sea $\pi : W \rightarrow \text{Con}(f)$ la proyección canónica al conúcleo de f . Si U es otro espacio vectorial y $g : W \rightarrow U$ es una función lineal tal que $\text{Im}(f) \subseteq \text{Nu}(g)$, entonces existe una y solo una fun-

ción lineal $\tilde{g} : \text{Con}(f) \rightarrow U$ tal que $\tilde{g} \circ \pi = g$.

$$\begin{array}{ccc} W & \xrightarrow{g} & Y \\ \pi \downarrow & \nearrow \tilde{g} & \\ \text{Con}(f) & & \end{array}$$

Demostración. Esto es simplemente un caso particular de la Proposición 2.8.6. \square

2.8.12. Vimos en la Proposición 2.3.2 que una función lineal es un monomorfismo si y solamente si su núcleo es nulo. El conúcleo tiene un rol similar para los epimorfismos.

Proposición. Una función lineal $f : V \rightarrow W$ es un epimorfismo si y solamente si su conúcleo $\text{Con}(f)$ es un espacio nulo.

Demostración. Si $f : V \rightarrow W$ es una función lineal, el conúcleo $\text{Con}(f)$ es el cociente $W/\text{Im}(f)$, y sabemos que este cociente es un espacio nulo si y solamente si $\text{Im}(f) = W$, es decir, si y solamente si la función f es sobreyectiva. \square

§9. Los teoremas de isomorfismo

2.9.1. Nuestro objetivo en esta sección es probar los llamados *teoremas de isomorfismo de Noether*, que son fundamentales en la manipulación de los espacios cociente que definimos en la sección anterior. Fueron enunciados y probados originalmente —en el contexto de los módulos y en casos especiales— por *Emmy Noether* en [Noe26] y, tres años más tarde —en el contexto de los grupos y con toda generalidad— por *Bartel Leendert van der Waerden* en su libro *Moderne Algebra* [vdW30], que es considerado como el tratado fundacional del álgebra moderna.

2.9.2. El *primer teorema de isomorfismo* nos da una descripción a menos de isomorfismo de la imagen de una función lineal en términos de un cociente de su dominio:

Proposición. Si $f : V \rightarrow W$ es una función lineal, entonces hay un isomorfismo

$$\phi : V/\text{Nu}(f) \rightarrow \text{Im}(f)$$

tal que para cada $x \in V$ es $\phi([x]) = f(x)$.

Demostración. Sea $f : V \rightarrow W$ una función lineal y consideremos la función

$$g : x \in V \mapsto f(x) \in \text{Im}(f),$$

que es lineal y sobreyectiva. Es inmediato que $\text{Nu}(g) = \text{Nu}(f)$, así que la Proposición 2.8.6 nos dice que existe una función lineal $\tilde{g} : V / \text{Nu}(f) \rightarrow \text{Im}(f)$ tal que $\tilde{g}([x]) = g(x) = f(x)$ para todo $x \in V$. Para probar la proposición, entonces, hay que mostrar que se trata de un isomorfismo, ya que si ése es el caso podremos poner $\phi = \tilde{g}$.

- Sea primero $u \in V / \text{Nu}(f)$ tal que $\tilde{g}(u) = 0$. Si $x \in V$ es un representante de u , de manera que $u = [x]$, entonces $0 = \tilde{g}(u) = \tilde{g}([x]) = f(x)$ y $x \in \text{Nu}(f)$: esto nos dice que $x \equiv 0$ y, por lo tanto, que $u = [x] = [0]$ es el elemento nulo del cociente $V / \text{Nu}(f)$. Así, la función \tilde{g} es inyectiva.
- Por otro lado, si $z \in \text{Im}(f)$, existe $x \in V$ tal que $f(x) = z$ y entonces $\tilde{g}([x]) = f(x) = z$: vemos así que \tilde{g} es sobreyectiva. \square

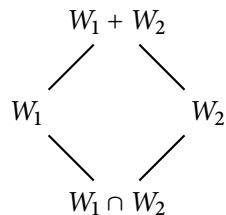
2.9.3. El segundo teorema de isomorfismo afirma que ciertos subcocientes —esto es, cocientes de subespacios— de un espacio vectorial son isomorfos:

Proposición. *Sea V un espacio vectorial. Si W_1 y W_2 son dos subespacios de V , entonces hay un isomorfismo*

$$\phi : \frac{W_1}{W_1 \cap W_2} \rightarrow \frac{W_1 + W_2}{W_2}$$

tal que para cada $x \in W_1$ se tiene que $\phi([x]_{W_1 \cap W_2}) = [x]_{W_2}$.

Es usual acompañar el enunciado de esta proposición con el siguiente diagrama



en el que las líneas representan las relaciones de inclusión entre los subespacios de V que aparecen en el enunciado: la conclusión de la proposición es que los cocientes correspondientes a lados opuestos de este rombo son isomorfos.

Demostración. Si $\pi : W_1 + W_2 \rightarrow (W_1 + W_2)/W_2$ es la proyección canónica e $\iota : W_1 \rightarrow W_1 + W_2$ es la inclusión, podemos considerar la composición

$$f = \pi \circ \iota : W_1 \rightarrow \frac{W_1 + W_2}{W_2}.$$

Ahora bien: si $x \in W_1 \cap W_2$, es

$$f(x) = \pi(\iota(x)) = \pi(x) = [x]_{W_2}$$

y, como $x \in W_2$, este último elemento del cociente $(W_1 + W_2)/W_2$ es nulo. De acuerdo a la Proposición 2.8.6, entonces, existe una función lineal $\tilde{f} : W_1/(W_1 \cap W_2) \rightarrow (W_1 + W_2)/W_2$ tal que para cada $x \in W_1$ es

$$\tilde{f}([x]_{W_1 \cap W_2}) = f(x) = [x]_{W_2}.$$

Mostremos que \tilde{f} es un isomorfismo:

- Supongamos primero que $u \in W_1/(W_1 \cap W_2)$ está en el núcleo de \tilde{f} . Si $x \in W_1$ es un representante de u , de manera que $u = [x]_{W_1 \cap W_2}$, entonces

$$\tilde{f}(u) = f(x) = [x]_{W_2} = 0$$

y esto significa que $x \in W_2$. Se sigue de esto que $x \in W_1 \cap W_2$ y, por lo tanto, que $u = [x]_{W_1 \cap W_2} = [0]_{W_1 \cap W_2}$, el elemento nulo de $W_1/(W_1 \cap W_2)$. Vemos así que la función \tilde{f} es inyectiva.

- Sea, por otro lado, v un elemento de $(W_1 + W_2)/W_2$, de manera que existe $y \in W_1 + W_2$ tal que $v = [y]_{W_2}$. Esto implica que existen $y_1 \in W_1$ e $y_2 \in W_2$ con $y = y_1 + y_2$ y, en particular, que $y - y_1 = y_2 \in W_2$, esto es, que $[y]_{W_2} = [y_1]_{W_2}$. Pero entonces $\tilde{f}([y_1]) = [y_1]_{W_2} = [y]_{W_2} = v$, con lo que v está en la imagen de \tilde{f} . La función \tilde{f} es, por lo tanto, sobreyectiva. \square

2.9.4. El *tercer teorema de isomorfismo* se ocupa de cocientes iterados, esto es, de describir cocientes de espacios cocientes. Para enunciarlo necesitamos hacer antes la siguiente observación.

Si V es un espacio vectorial y W_1 y W_2 son subespacios de V tales que $W_1 \subseteq W_2$, entonces el conjunto W_2/W_1 es un subconjunto de V/W_2 , esto es, toda clase de congruencia módulo W_2 en W_1 es una clase de congruencia módulo W_2 en V . Para verlo, basta observar que si $u \in W_1/W_2$ y $x \in u$, de manera que $u = [x]$, entonces

$$\{y \in W_1 : y \equiv_{W_2} x\} = \{y \in V : y \equiv_{W_2} x\}.$$

Una consecuencia inmediata de esto es que

(12)

un elemento de V/W_2 está en W_1/W_2 si y solamente si tiene un representante en V que está en W_1 .

Más aún, el subconjunto W_1/W_2 de V/W_2 es un subespacio: no es vacío, ya que contiene a $[0]_{W_2}$, y si $a, b \in \mathbb{k}$ y $u, v \in W_1/W_2$, de manera que existen $x, y \in W_1$ con $u = [x]$ y $v = [y]$, se tiene que

$$au + bv = a[x] + b[y] = [ax + by] \in W_1/W_2$$

ya que, por supuesto, $ax + by \in W_1$.

Proposición. *Sea V un espacio vectorial. Si W_1 y W_2 son subespacios de V tales que $W_1 \subseteq W_2$, entonces hay un isomorfismo*

$$\phi : \frac{V}{W_2} \rightarrow \frac{W_1/W_2}{W_2/W_1}$$

tal que para cada $x \in V$ vale que $\phi([x]_{W_2}) = [[x]_{W_1}]_{W_2/W_1}$.

Demostración. Sean $\pi_1 : V \rightarrow V/W_1$ y $\pi_2 : V/W_1 \rightarrow (V/W_1)/(W_2/W_1)$ las proyecciones canónicas y sea

$$f = \pi_2 \circ \pi_1 : V \rightarrow \frac{(V/W_1)}{(W_2/W_1)}$$

su composición, de manera que para cada $x \in V$ es $f(x) = [[x]_{W_1}]_{W_2/W_1}$. Si $x \in W_2$, entonces el elemento $[x]_{W_1}$ de V/W_1 pertenece a W_2/W_1 , así que $[[x]_{W_1}]_{W_2/W_1}$ es el cero de $(V/W_1)/(W_2/W_1)$. La Proposición 2.8.6 nos dice por lo tanto que existe una función lineal

$$\tilde{f} : \frac{V}{W_2} \rightarrow \frac{V/W_1}{W_2/W_1}$$

tal que $\tilde{f}([x]_{W_2}) = [[x]_{W_1}]_{W_2/W_1}$ para todo $x \in V$. Para probar la proposición mostraremos que esta función \tilde{f} es un isomorfismo.

- Supongamos que $u \in V/W_2$ es tal que $\tilde{f}(u) = 0$ y sea $x \in V$ un representante de u , de manera que $u = [x]_{W_2}$ y que $\tilde{f}(u) = [[x]_{W_1}]_{W_2/W_1}$ es el cero de $(V/W_1)/(W_2/W_1)$. Esto nos dice que $[x]_{W_1}$ es un elemento de W_2/W_1 y, de acuerdo a la observación (12) que hicimos arriba, que existe $x' \in W_2$ tal que $[x]_{W_1} = [x']_{W_1}$ y, por lo tanto, $x - x' \in W_1$. Se sigue de esto que $x = (x - x') + x' \in W_1 + W_2 = W_2$, ya que $W_1 \subseteq W_2$, y entonces que $u = [x]_{W_2}$ es el cero de V/W_2 . Esto prueba que la función \tilde{f} es inyectiva.
- Sea ahora v un elemento de $(V/W_1)/(W_2/W_1)$. Existe $w \in V/W_1$ tal que $v = [w]_{W_2/W_1} = w$, y a su vez, existe $x \in V$ tal que $w = [x]_{W_1}$. Se sigue de esto, entonces, que

$$\tilde{f}([x]_{W_2}) = [[x]_{W_1}]_{W_2/W_1} = [w]_{W_2/W_1} = v$$

y, por lo tanto, que la función \tilde{f} es sobreyectiva. \square

2.9.5. Junto con los tres teoremas de isomorfismo anteriores es importante el siguiente *teorema de correspondencia*, que describe los subespacios de un espacio cociente.

Proposición. Sea V un espacio vectorial y sea W un subespacio de V . Sea $\pi : V \rightarrow V/W$ es la proyección canónica y sean

- $\mathcal{L}(V)_W$ es el conjunto de todos los subespacios de V que contienen a W y
- $\mathcal{L}(V/W)$ es el conjunto de todos los subespacios de V/W .

La función

$$\phi : U \in \mathcal{L}(V)_W \mapsto \pi(U) \in \mathcal{L}(V/W)$$

es una biyección con cuya función inversa es

$$\psi : T \in \mathcal{L}(V/W) \mapsto \pi^{-1}(T) \in \mathcal{L}(V)_W.$$

Más aún, si U_1 y U_2 son dos elementos de $\mathcal{L}(V)_W$, se tiene que

$$U_1 \subseteq U_2 \iff \phi(U_1) \subseteq \phi(U_2),$$

y si U_1 y U_2 son dos elementos de $\mathcal{L}(V/W)$, se tiene que

$$T_1 \subseteq T_2 \iff \psi(T_1) \subseteq \psi(T_2).$$

Demostración. Si U es un subespacio de V , su imagen $\pi(U)$ por π es un subespacio de V/W , así que podemos considerar la función

$$\phi : U \in \mathcal{L}(V)_W \mapsto \pi(U) \in \mathcal{L}(V/W).$$

Por otro lado, si T es un subespacio de V/W sabemos que la preimagen $\pi^{-1}(T)$ es un subespacio de V y, ya que $0 \in T$, es $W = \pi^{-1}(0) \subseteq \pi^{-1}(T)$. Esto significa que hay una función

$$\psi : T \in \mathcal{L}(V/W) \mapsto \pi^{-1}(T) \in \mathcal{L}(V)_W.$$

Afirmamos que ϕ y ψ con biyecciones inversas:

- Si $U \in \mathcal{L}(V)_W$, entonces $\psi(\phi(U)) = \pi^{-1}(\pi(U))$. Esto claramente contiene a U ; por otro lado, si $x \in \pi^{-1}(\pi(U))$, entonces $\pi(x) \in \pi(U)$ y existe $u \in U$ tal que $\pi(x) = \pi(u)$: esto implica que $x - u \in \text{Nu}(\pi) = W \subseteq U$ y, por lo tanto, que $x \in U$. Así, es $\psi(\phi(U)) = U$.
- Si ahora $T \in \mathcal{L}(V/W)$, entonces $\phi(\psi(T)) = \pi(\pi^{-1}(T))$. Esto está evidentemente contenido en T y, a su vez, contiene a T porque la función π es sobreyectiva. Esto nos dice que $\phi(\psi(T)) = T$.

Esto prueba la primera afirmación de la proposición. La segunda es consecuencia de que

- si $U_1, U_2 \in \mathcal{L}(V)_W$, entonces

$$U_1 \subseteq U_2 \implies \phi(U_1) = \pi(U_1) \subseteq \pi(U_2) = \phi(U_2),$$

- si $T_1, T_2 \in \mathcal{L}(V/W)$, entonces

$$T_1 \subseteq T_2 \implies \psi(T_1) = \pi^{-1}(T_1) \subseteq \pi^{-1}(T_2) = \psi(T_2)$$

y de que ϕ y ψ son biyecciones inversas. □

§10. Exactitud

2.10.1. Si $n \geq 0$, un diagrama de la forma

$$U_0 \xrightarrow{f_1} U_1 \xrightarrow{f_2} U_2 \xrightarrow{f_3} \dots \xrightarrow{f_n} U_n \xrightarrow{f_{n+1}} U_{n+1} \quad (13)$$

en el que U_0, \dots, U_{n+1} son espacios vectoriales y $f_1 : U_0 \rightarrow U_1, \dots, f_{n+1} : U_n \rightarrow U_{n+1}$ son funciones lineales es

- un **complejo** si para cada $i \in \llbracket n \rrbracket$ se tiene que $\text{Im}(f_i) \subseteq \text{Nu}(f_{i+1})$, y
- una **sucesión exacta** si para cada $i \in \llbracket n \rrbracket$ se tiene que $\text{Im}(f_i) = \text{Nu}(f_{i+1})$.

y en ambos casos decimos que la longitud del diagrama es n . Cuando queremos referirnos al hecho de que en (13) se tiene que $\text{Im}(f_i) = \text{Nu}(f_{i+1})$ hablamos de la **exactitud del diagrama en U_i** . Por supuesto, el diagrama completo es exacto si y solamente si es exacto en cada uno de U_1, \dots, U_n .

Claramente una sucesión exacta es un complejo. En general, cuando alguno de los espacios vectoriales de (13) es un espacio nulo, las funciones que salen o que llegan a él en el diagrama son necesariamente nulas, así que no es necesario ponerles un nombre.

Una sucesión exacta de longitud 3 que empieza y termina con espacios nulos, de manera que es de la forma

$$0 \longrightarrow U_1 \xrightarrow{f_2} U_2 \xrightarrow{f_3} U_3 \longrightarrow 0$$

es tradicionalmente llamada una **sucesión exacta corta**.

2.10.2. Ejemplos.

(a) Un diagrama de la forma

$$0 \longrightarrow U \xrightarrow{f} V$$

es exacto si y solamente si el morfismo f es inyectivo, mientras que uno de la forma

$$U \xrightarrow{f} V \longrightarrow 0$$

es exacto si y solamente si el morfismo f es sobreinyectivo. De esto se sigue que un diagrama de la forma

$$0 \longrightarrow U \xrightarrow{f} V \longrightarrow 0$$

es exacto si y solamente si el morfismo f es un isomorfismo.

(b) Si $f : V \rightarrow W$ es una función lineal, $\iota : \text{Nu}(f) \rightarrow V$ a función de inclusión del núcleo de f en V y $\pi : W \rightarrow \text{Con}(f)$ la proyección canónica de W al conúcleo de f , entonces el diagrama

$$0 \longrightarrow \text{Nu}(f) \xrightarrow{\iota} V \xrightarrow{f} W \xrightarrow{\pi} \text{Con}(f) \longrightarrow 0$$

es exacto. Si además consideramos la función $\tilde{f} : v \in V \mapsto f(v) \in \text{Im}(f)$ que se obtiene correstringiendo f al subespacio $\text{Im}(f)$ de su codominio, tenemos una sucesión exacta corta

$$0 \longrightarrow \text{Nu}(f) \xrightarrow{\iota} V \xrightarrow{\tilde{f}} \text{Im}(f) \longrightarrow 0$$

Como casos particulares de esto, tenemos que si $f : V \rightarrow W$ es sobreyectiva hay una sucesión exacta corta

$$0 \longrightarrow \text{Nu}(f) \xrightarrow{\iota} V \xrightarrow{f} W \longrightarrow 0$$

y que si $f : V \rightarrow W$ es inyectiva hay una sucesión exacta corta

$$0 \longrightarrow V \xrightarrow{f} W \xrightarrow{\pi} \text{Con}(f) \longrightarrow 0 \quad \diamond$$

2.10.3. Empecemos por una observación bien sencilla:

Proposición. Si

$$0 \longrightarrow U \xrightarrow{f} V \xrightarrow{g} W \longrightarrow 0 \tag{14}$$

es una sucesión exacta corta de espacios vectoriales, entonces hay isomorfismos

$$\tilde{f} : U \rightarrow \text{Nu}(g), \quad \tilde{g} : \text{Con}(f) \rightarrow W$$

tales que $\tilde{f}(u) = f(u)$ para todo $u \in U$ y $\tilde{g}([v]) = g(v)$ para cada $w \in W$.

Demostración. Supongamos que tenemos una sucesión exacta corta como la del enunciado. La exactitud en V nos dice que $\text{Im}(f) = \text{Nu}(g)$, así que f toma valores en $\text{Nu}(g)$ y, por lo tanto, podemos correstringirla a ese subespacio de su codominio para obtener una función $\tilde{f} : U \rightarrow \text{Nu}(g)$ tal que $\tilde{f}(u) = f(u)$ para todo $u \in U$. La exactitud del diagrama (14) en U nos dice que la función f es un monomorfismo, así que claramente \tilde{f} también lo es. Por otro lado, como $\text{Im}(f) = \text{Nu}(g)$, es evidente que \tilde{f} es un epimorfismo. En definitiva, se trata de un isomorfismo.

Sea $\pi : V \rightarrow \text{Con}(f)$ la proyección canónica de W al conúcleo de f . Como $\text{Nu}(g) \subseteq \text{Im}(f)$, la Proposición 2.8.11 nos dice que existe una función $\tilde{g} : \text{Con}(f) \rightarrow W$ tal que $\tilde{g} \circ \pi = g$, es decir, tal que $\tilde{g}([v]) = g(v)$ para todo $v \in V$. Si $w \in W$, entonces existe $v \in V$ tal que $g(v) = w$, ya que la función g es sobreyectiva porque el diagrama (14) es exacto en W , y por lo tanto $\tilde{g}w = g(v) = \tilde{g}([v]) \in \text{Im}(\tilde{g})$: esto nos dice que la función \tilde{g} es sobreyectiva. Por otro lado, si $x \in \text{Con}(f)$ es un elemento del núcleo de \tilde{g} , entonces existe $v \in V$ tal que $x = [v]$ y $g(x) = \tilde{g}(x) = 0$: vemos así que $v \in \text{Nu}(g) = \text{Im}(f)$ y, en consecuencia, que $x = [v] = 0$ en $\text{Con}(f) \setminus \text{Im}(f)$. Así, \tilde{g} también es un monomorfismo y, juntando todo, un isomorfismo. \square

2.10.4. Cuando trabajamos con complejos y sucesiones exactas generalmente estamos interesados en relacionar las dimensiones de los espacios que intervienen en ellos. Un primer ejemplo de ello es el siguiente:

Proposición. Si el diagrama

$$U \xrightarrow{f} V \xrightarrow{g} W$$

de espacios vectoriales es exacto y U y W tienen dimensión finita, entonces V también tiene dimensión finita.

Demostración. Supongamos que tenemos un diagrama como el del enunciado en el que los espacios U y W tienen dimensión finita. De acuerdo al Teorema 2.4.1, para ver que V tiene dimensión finita es suficiente que probemos que $\text{Nu}(g)$ y $\text{Im}(g)$ tienen dimensión finita. El primero coincide con $\tilde{\mathcal{I}}f$, y tiene dimensión finita porque el dominio U de f tiene dimensión finita, mientras que el segundo $\text{Im}(g)$ tiene dimensión finita porque es un subespacio de W . \square

2.10.5. Si hacemos una hipótesis más fuerte, podemos concluir algo más preciso:

Proposición. Si

$$0 \longrightarrow U \xrightarrow{f} V \xrightarrow{g} W \longrightarrow 0$$

es una sucesión exacta corta de espacios vectoriales y o bien V tiene dimensión finita o bien U y W tienen dimensión finita, entonces de hecho los tres tienen dimensión finita y

$$\dim V = \dim U + \dim W.$$

Demostración. Consideremos un diagrama como el del enunciado. De acuerdo a la Proposición 2.10.3, hay un isomorfismo $U \cong \text{Nu}(g)$, y la exactitud del diagrama en W nos dice que la función g es sobreyectiva, de manera que $\text{Im}(g) = W$.

Si U y W tienen dimensión finita, la Proposición 2.10.4 nos dice que V tiene dimensión finita. Si en cambio V tiene dimensión finita, entonces el Teorema 2.4.1 nos dice que $\text{Nu}(f)$ e $\text{Im}(g)$ tienen dimensión finita y, por lo tanto, que U y W tienen dimensión finita. Esto prueba la primera afirmación del enunciado. Para ver que vale la segunda, notemos que si los tres espacios del diagrama tienen dimensión finita, entonces del teorema sabemos que

$$\dim V = \dim \text{Nu}(g) + \dim \text{Im}(g) = \dim U + \dim W,$$

como queremos. \square

2.10.6. A partir de la proposición que acabamos de probar obtenemos fácilmente el siguiente resultado importante:

Proposición. Si $n \in \mathbb{N}$ y

$$0 \xrightarrow{f_1} U_1 \xrightarrow{f_2} U_2 \xrightarrow{f_3} \dots \xrightarrow{f_{n-1}} U_{n-1} \xrightarrow{f_n} U_n \xrightarrow{f_{n+1}} 0$$

es una sucesión exacta de espacios vectoriales de dimensión finita de longitud n que empieza y termina con espacios nulos, entonces

$$\sum_{i=1}^n (-1)^n \dim U_i = 0.$$

Demostración. Para cada $i \in \llbracket n \rrbracket$ sabemos que

$$\dim U_i - \dim \text{Nu}(f_{i+1}) - \dim \text{Im}(f_{i+1}) = 0,$$

y entonces

$$\begin{aligned} 0 &= \sum_{i=1}^n (-1)^i (\dim U_i - \dim \text{Nu}(f_{i+1}) - \dim \text{Im}(f_{i+1})) \\ &= \sum_{i=1}^n (-1)^i \dim U_i - \sum_{i=1}^n (-1)^i (\dim \text{Nu}(f_{i+1}) + \dim \text{Im}(f_{i+1})). \end{aligned} \quad (15)$$

La exactitud del diagrama nos dice que para cada $i \in \llbracket n \rrbracket$ es $\dim \text{Nu}(f_{i+1}) = \dim \text{Im}(f_i)$. Usando eso y poniendo $d_i := \dim \text{Im}(f_i)$ para cada $i \in \llbracket n+1 \rrbracket$, tenemos que

$$\begin{aligned} \sum_{i=1}^n (-1)^i (\dim \text{Nu}(f_{i+1}) + \dim \text{Im}(f_{i+1})) &= \sum_{i=1}^n (-1)^i (d_i + d_{i+1}) \\ &= (d_1 + d_2) - (d_2 + d_3) + (d_3 + d_4) - (d_4 + d_5) + \cdots + (-1)^n (d_n + d_{n+1}) \\ &= d_1 + (-1)^{n-1} d_{n+1}, \end{aligned}$$

ya que todos los otros sumandos se cancelan, y esto es

$$= 0,$$

ya que los números d_1 y d_{n+1} son nulos: se trata de las dimensiones de las imágenes de las funciones $f_1 : 0 \rightarrow U_1$ y $f_{n+1} : U_n \rightarrow 0$. Volviendo a (15) con esta información, vemos que la afirmación de la proposición es cierta. \square

2.10.7. Decimos que un diagrama de la forma

$$\begin{array}{ccc} U & \xrightarrow{f} & V \\ g \downarrow & & \downarrow h \\ W & \xrightarrow{k} & T \end{array}$$

en el que U, V, W y T son espacios vectoriales y $f : U \rightarrow V, g : U \rightarrow W, h : V \rightarrow T$ y $k : W \rightarrow T$ son funciones lineales **commuta** si $k \circ g = h \circ f$.

2.10.8. Usaremos el siguiente lema para probar la Proposición 2.10.9 más abajo.

Lema. Sea

$$\begin{array}{ccccccc} U & \xrightarrow{f} & V & \xrightarrow{g} & W & \longrightarrow & 0 \\ \downarrow a & & \downarrow b & & \downarrow c & & \\ 0 & \longrightarrow & U' & \xrightarrow{f'} & V' & \xrightarrow{g'} & W' \end{array}$$

un diagrama de espacios vectoriales y funciones lineales en el que las dos filas son sucesiones exactas y los dos cuadrados comutan.

(i) Es $f(\text{Nu}(a)) \subseteq \text{Nu}(b)$ y $g(\text{Nu}(b)) \subseteq \text{Nu}(c)$, de manera que hay funciones lineales

$$\bar{f} : u \in \text{Nu}(a) \mapsto f(u) \in \text{Nu}(b), \quad \bar{g} : v \in \text{Nu}(b) \mapsto g(v) \in \text{Nu}(c).$$

(ii) Hay funciones lineales

$$\bar{f}' : \text{Con}(a) \rightarrow \text{Con}(b), \quad \bar{g}' : \text{Con}(b) \rightarrow \text{Con}(c)$$

tales que para cada $u' \in U'$ y para cada $v' \in V'$ se tiene que

$$\bar{f}'([u']) = [f'(u')], \quad \bar{g}'([v']) = [g'(v')].$$

Demostración. (i) Que el cuadrado de la izquierda del diagrama del enunciado commute significa que $b \circ f = f' \circ a$. Se sigue de esto que si $u \in \text{Nu}(a)$, de manera que $a(u) = 0$, entonces $b(f(u)) = f'(a(u)) = f'(0) = 0$, así que $f(u) \in \text{Nu}(b)$. Vemos así que $f(\text{Nu}(a)) \subseteq \text{Nu}(b)$ y, por lo tanto, hay una función $\bar{f}' : \text{Con}(a) \rightarrow \text{Con}(b)$. Exactamente de la misma forma, a partir de la conmutatividad del cuadrado de la derecha del diagrama podemos ver que $g(\text{Nu}(b)) \subseteq \text{Nu}(c)$ y que hay una función $\bar{g} : v \in \text{Nu}(b) \mapsto g(v) \in \text{Nu}(c)$.

(ii) Sean $q_a : U' \rightarrow \text{Con}(a)$, $q_b : V' \rightarrow \text{Con}(b)$ y $q_c : W' \rightarrow \text{Con}(c)$ las proyecciones canónicas de U , V y W a los conúcleos de las funciones a , b y c , respectivamente.

Consideremos la función $q_b \circ f' : U' \rightarrow \text{Con}(b)$. La conmutatividad del primer cuadrado del diagrama del enunciado nos dice que $b \circ f = f' \circ a$, así que si $u' \in \text{Im}(a)$ y $u \in U$ es tal que $u' = a(u)$, entonces

$$(q_b \circ f')(u') = q_b(f'(a(u))) = q_b(b(f(u))) = [b(f(u))] = 0$$

en $\text{Con}(b) = V'/\text{Im}(b)$, ya que $b(f(u))$ está, por supuesto, en $\text{Im}(b)$. Vemos así que

$$(q_b \circ f')(\text{Im}(a)) = 0$$

y aplicando la Proposición 2.8.11 a la función $q_b \circ f'$ podemos concluir que existe una función lineal $\bar{f}' : \text{Con}(a) \rightarrow \text{Con}(b)$ tal que $\bar{f}'([u']) = [f'(u')]$ para todo $u' \in U'$. El mismo argumento pero aplicado al segundo cuadrado del diagrama muestra que existe una función lineal $\bar{g}' : \text{Con}(b) \rightarrow \text{Con}(c)$ tal que $\bar{g}'([v']) = [g'(v')]$ para cada $v' \in V'$. \square

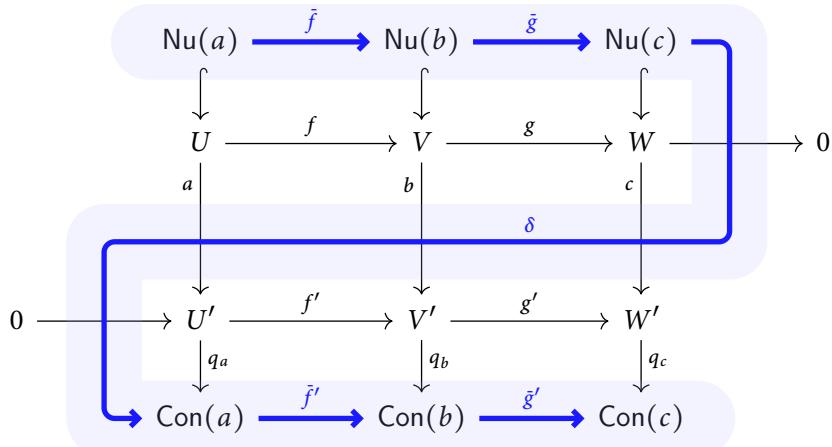


Figura 2.1. El diagrama completo de la Proposición 2.10.9.

2.10.9. El siguiente resultado, conocido como el *Lema de la serpiente*, es extremadamente importante.

Proposición. Sea

$$\begin{array}{ccccccc} U & \xrightarrow{f} & V & \xrightarrow{g} & W & \longrightarrow & 0 \\ \downarrow a & & \downarrow b & & \downarrow c & & \\ 0 & \longrightarrow & U' & \xrightarrow{f'} & V' & \xrightarrow{g'} & W' \end{array}$$

un diagrama de espacios vectoriales y funciones lineales en el que las dos filas son sucesiones exactas y los dos cuadrados comutan. Existe una función lineal $\delta : \text{Nu}(c) \rightarrow \text{Con}(a)$ tal que

$$\text{Nu}(a) \xrightarrow{\tilde{f}} \text{Nu}(b) \xrightarrow{\tilde{g}} \text{Nu}(c) \xrightarrow{\delta} \text{Con}(a) \xrightarrow{\tilde{f}'} \text{Con}(b) \xrightarrow{\tilde{g}'} \text{Con}(c)$$

es una sucesión exacta, con \tilde{f} , \tilde{g} , \tilde{f}' y \tilde{g}' las funciones del Lema 2.10.8.

La función δ es conocida como el **morfismo de conexión**. En la Figura 2.1 puede verse una versión completada del diagrama de esta proposición en la que puede apreciarse claramente por qué el resultado se llama el Lema de la serpiente.

Demostración. Sea w un elemento de $\text{Nu}(c)$. Como la función g es sobreyectiva, existe $v \in V$ tal que $g(v) = w$ y tenemos que

$$g'(b(v)) = c(g(v)) = c(w) = 0,$$

de manera que $b(v) \in \text{Nu}(g') = \text{Im}(f')$: existe entonces un vector u' en U' —y solo uno, ya que la función f' es inyectiva— tal que $f'(u') = b(v)$. Podemos entonces considerar la clase $[u']$ de u'

en $\text{Con}(a) = U' / \text{Im}(a)$. Afirmamos que

la clase $[u']$ en $\text{Con}(a)$ depende solamente del elemento w de $\text{Nu}(c)$ con el que empezamos y no de la elección del elemento v de V tal que $g(v) = w$ que usamos para construirla.

En efecto, supongamos que v_1 es otro elemento de V tal que $g(v_1) = w$. En ese caso es

$$g(v - v_1) = g(v) - g(v_1) = w - w = 0,$$

así que $v - v_1 \in \text{Nu}(g) = \text{Im}(f)$ y, por lo tanto, existe $u \in U$ tal que $v - v_1 = f(u)$. De esto se deduce que

$$b(v_1) = b(v) - b(f(u)) = f'(v') - f'(a(u)) = f'(v' - a(u))$$

así que el único elemento de U' cuya imagen por f' es $b(v_1)$ es $v' - a(u)$ y su clase en $\text{Con}(a)$ es $[v' - a(u)]$, que coincide con $[v']$, ya que $a(u) \in \text{Im}(a)$. Esto prueba lo que queremos.

En vista de lo que acabamos de probar, tenemos que existe una función $\delta : \text{Nu}(c) \rightarrow \text{Con}(a)$ que tiene la siguiente propiedad característica:

si $w \in \text{Nu}(c)$ y $x \in \text{Con}(a)$, entonces $\delta(w) = x$ si y solamente si existen $v \in V$ y $u' \in U'$ tales que $g(v) = w$, $f'(u') = b(v)$ y $x = [u']$.

$$\begin{array}{ccccc} & & v & \xrightarrow{g} & w \\ & & b \downarrow & & \downarrow c \\ u' & \xrightarrow{f'} & \bullet & \xrightarrow{g'} & 0 \\ q_a \downarrow & & & & \\ \delta(w) & & & & \end{array}$$

Nos queda por probar la exactitud de la sucesión que aparece en el enunciado de la proposición y haremos eso lugar a lugar.

- *Exactitud en $\text{Nu}(b)$.* Si $u \in \text{Nu}(a)$, entonces $\tilde{g}(\tilde{f}(u)) = g(f(u)) = 0$ porque $g \circ f = 0$: esto nos dice que $\text{Im}(\tilde{f}) \subseteq \text{Nu}(\tilde{g})$. Recíprocamente, si $v \in \text{Nu}(b)$ es tal que $\tilde{g}(v) = 0$, entonces $g(v) = \tilde{g}(v) = 0$ y $v \in \text{Nu}(g) = \text{Im}(f)$, así que existe $u \in U$ tal que $f(u) = v$. Es $f'(a(u)) = b(f(u)) = b(v) = 0$ y, como la función f' es inyectiva, $a(u) = 0$, esto es, $u \in \text{Nu}(a)$. Vemos así que $v = f(u) = \tilde{f}(u) \in \text{Im}(\tilde{f})$. Esto prueba que $\text{Nu}(\tilde{g}) \subseteq \text{Im}(\tilde{f})$.
- *Exactitud en $\text{Nu}(c)$.* Sea $v \in \text{Nu}(b)$ y escribamos $w := \tilde{g}(v) = g(v) \in \text{Nu}(c)$. Como $g(v) = w$ y $f'(0) = 0 = b(v)$, sabemos que $\delta(g(v)) = \delta(w) = [0]$, el cero de $\text{Con}(a)$. Hemos probado que $\text{Im}(\tilde{g}) \subseteq \text{Nu}(\delta)$.

Supongamos ahora que $w \in \text{Nu}(\delta)$, de manera que $w \in \text{Nu}(c)$ y existen entonces $v \in V$ y $u' \in U'$ tales que $g(v) = w$, $f'(u') = b(v)$ y en $\text{Con}(a)$ es $[u'] = 0$. Esto último significa que $u' \in \text{Im}(a)$, así que existe $u \in U$ tal que $a(u) = u'$, y entonces

$$b(v - f(u)) = b(v) - b(f(u)) = f'(u') - f'(a(u)) = f'(u') - f'(u') = 0.$$

Vemos así que $v - f(u) \in \text{Nu}(b)$ y, por lo tanto, que

$$w = g(v) - g(f(v)) = g(v - f(v)) = \tilde{g}(v - f(v)) \in \text{Im}(\tilde{g}).$$

La conclusión de esto es que $\text{Nu}(\delta) \subseteq \text{Im}(\tilde{g})$.

- *Exactitud en $\text{Con}(a)$.* Sea $w \in \text{Nu}(c)$. Sabemos que existen $v \in V$ y $u' \in U'$ tales que $g(v) = w$, $f'(u') = b(v)$ y $\delta(w) = [u']$ en $\text{Con}(a)$, así que

$$\bar{f}'(\delta(w)) = \bar{f}'([u']) = [f'(u')] = [b(v)] = 0$$

en $\text{Con}(b)$ ya que, por supuesto, $b(v) \in \text{Im}(b)$. Esto nos dice que $\text{Im}(\delta) \subseteq \text{Nu}(\bar{f}')$.

Sea ahora $x \in \text{Con}(a)$ tal que $\bar{f}'(x) = 0$, y sea $u' \in U'$ tal que $x = [u']$, de manera que $\bar{f}'(x) = \bar{f}'([u']) = [f'(u')]$: como esto es el elemento nulo de $\text{Con}(b)$, tenemos que $f'(u') \in \text{Im}(b)$, esto es, existe $v \in V$ tal que $f'(u') = b(v)$. Observemos que $c(g(v)) = g'(b(v)) = g'(f'(u')) = 0$, así que $w := g(v)$ es un elemento de $\text{Nu}(c)$. Más aún, como $g(v) = w$ y $f'(u') = b(v)$, es $\delta(w) = [u'] = x$. Hemos probado con esto que $\text{Nu}(\bar{f}') \subseteq \text{Im}(\delta)$.

- *Exactitud en $\text{Con}(b)$.* Si $x \in \text{Con}(a)$, entonces existe $u' \in U'$ tal que $x = [u']$ y, por lo tanto,

$$\bar{g}'(\bar{f}'(x)) = \bar{g}'(\bar{f}'([u'])) = \bar{g}'([f'(u')]) = [g'(f'(u'))] = 0,$$

ya que $g' \circ f' = 0$. Vemos así que $\text{Im}(\bar{f}') \subseteq \text{Nu}(\bar{g}')$.

Finalmente, si $x \in \text{Con}(b)$ es tal que $\bar{g}'(x) = 0$, entonces existe $v' \in V'$ tal que $x = [v']$ y, por lo tanto, $[g'(v')] = \bar{g}'([v']) = \bar{g}(x) = 0$ en $\text{Con}(c)$. Esto significa que $g'(v') \in \text{Im}(c)$, es decir, que existe $w \in W$ tal que $g'(v') = c(w)$ y, como la función g es sobreyectiva, que existe $v \in V$ tal que $w = g(v)$. Pero entonces

$$g'(v' - b(v)) = g'(v') - g'(b(v)) = c(w) - c(g(v)) = c(w) - v(w) = 0,$$

así que $v' - b(v) \in \text{Nu}(g') = \text{Im}(f')$ y existe $u' \in U'$ tal que $v' - b(v) = f'(u')$. Podemos considerar la clase $[u']$ de u' en $\text{Con}(a)$, y en $\text{Con}(b)$ es

$$x = [v'] = [v' - b(v)]$$

ya que $b(v) \in \text{Im}(b)$, y esto es

$$= [f'(u')] = \bar{f}'([u']) \in \text{Im}(\bar{f}').$$

Esto prueba que $\text{Nu}(\bar{g}') \subseteq \text{Im}(\bar{f}')$.

La proposición queda así completamente demostrada. □

2.10.10. Una consecuencia casi directa del Lema de la serpiente es:

Proposición. Si $f : U \rightarrow V$ y $g : V \rightarrow W$ son dos funciones lineales, entonces hay una sucesión exacta de la forma

$$0 \longrightarrow \text{Nu}(f) \longrightarrow \text{Nu}(g \circ f) \longrightarrow \text{Nu}(h) \longrightarrow \text{Con}(f) \longrightarrow \text{Con}(g \circ f) \longrightarrow \text{Con}(g) \longrightarrow 0$$

Observemos que en este enunciado no decimos cuáles son las funciones lineales que aparecen en la sucesión exacta: en la demostración quedarán completamente descriptos, pero en casi todas las aplicaciones de este resultado lo importante es, sobre todo, saber que existen.

Demostración. Consideremos el diagrama

$$\begin{array}{ccccccc} 0 & \longrightarrow & U & \xrightarrow{\text{id}_U} & U & \longrightarrow & 0 \\ \downarrow & & \downarrow f & & \downarrow g \circ f & & \\ 0 & \longrightarrow & \text{Nu}(g) & \xrightarrow{\iota_g} & V & \xrightarrow{g} & W \end{array}$$

en el que la función ι_g es la inclusión de $\text{Nu}(g)$ en V . Es inmediato verificar que las dos filas son exactas y que los dos cuadrados comutan, así que el Lema de la serpiente 2.10.9 nos da una sucesión exacta

$$0 \longrightarrow \text{Nu}(f) \xrightarrow{\bar{\text{id}}_U} \text{Nu}(g \circ f) \xrightarrow{\delta_1} \text{Nu}(g) \xrightarrow{\iota_g} \text{Con}(f) \xrightarrow{\tilde{g}} \text{Con}(g \circ f) \quad (16)$$

Aquí usamos el hecho de que el núcleo del morfismo $0 \rightarrow \text{Nu}(g)$ es nulo y su conúcleo se puede identificar con $\text{Nu}(g)$.

De manera similar, el diagrama

$$\begin{array}{ccccc} U & \xrightarrow{f} & V & \xrightarrow{q_f} & \text{Con}(f) \longrightarrow 0 \\ \downarrow g \circ f & & \downarrow g & & \downarrow \\ 0 & \longrightarrow & W & \xrightarrow{\text{id}_W} & W \longrightarrow 0 \end{array}$$

tiene filas exactas y cuadrados comutativos, así que el Lema de la serpiente 2.10.9 nos da una sucesión exacta

$$\text{Nu}(g \circ f) \xrightarrow{\tilde{f}} \text{Nu}(g) \xrightarrow{\tilde{q}_f} \text{Con}(f) \xrightarrow{\delta_2} \text{Con}(g \circ f) \xrightarrow{\bar{\text{id}}_W} \text{Con}(g) \longrightarrow 0 \quad (17)$$

donde estamos identificando al núcleo y al conúcleo de $\text{Con}(f) \rightarrow 0$ con $\text{Con}(f)$ y 0.

Afirmamos ahora que el morfismo \tilde{g} que aparece en la sucesión exacta (16) coincide con el morfismo de conexión $\tilde{\delta}_2$ de la sucesión exacta (17). En efecto, si $x \in \text{Con}(f)$, entonces existe $v \in V$ tal que $x = [v] = q_f(v)$, y como $\text{id}_W(g(v)) = g(v)$, la propiedad característica del morfismo de conexión δ_2 nos dice que $\delta_2(x) = [w(v)]$ en el conúcleo de $\text{Con}(f) \rightarrow 0$, y esto significa, de acuerdo a la identificación que hicimos, que $\delta_2(x) = \tilde{g}([v]) = \tilde{g}(x)$.

Sabiendo ahora que $\tilde{g} = \delta_2$, la exactitud de los diagramas (16) y (17) implica inmediatamente la exactitud del diagrama

$$0 \rightarrow \text{Nu}(f) \xrightarrow{\bar{\text{id}}_U} \text{Nu}(g \circ f) \xrightarrow{\delta_1} \text{Nu}(g) \xrightarrow{\iota_g} \text{Con}(f) \xrightarrow{\tilde{g}} \text{Con}(g \circ f) \xrightarrow{\bar{\text{id}}_W} \text{Con}(g) \rightarrow 0$$

La proposición queda así probada. \square

Funciones lineales de tipo Fredholm

2.10.11. Decimos que una función lineal $f : V \rightarrow W$ es *tipo Fredholm* si su núcleo $\text{Nu}(f)$ tiene dimensión finita y su imagen $\text{Im}(f)$ tiene codimensión finita en W , y en ese caso el *índice* de f es el entero

$$\text{ind}(f) = \dim \text{Nu}(f) - \text{codim}_W \text{Im}(f).$$

De acuerdo a la Proposición 2.8.9(i), la imagen $\text{Im}(f)$ tiene codimensión finita en W si y solamente si el cociente $W/\text{Im}(f)$ tiene dimensión finita, este cociente es precisamente el conúcleo $\text{Con}(f)$ de f y, más aún, en ese caso tenemos que $\text{codim}_W \text{Im}(f) = \dim \text{Con}(f)$.

El nombre dado a esta clase de funciones recuerda a Erik Ivar Fredholm (1866–1927, Suecia).

2.10.12. Ejemplos.

- (a) Si V y W son espacios vectoriales de dimensión finita, es claro que toda función lineal $f : V \rightarrow W$ es de tipo Fredholm. Más aún, tenemos que

$$\begin{aligned} \text{ind}(f) &= \dim \text{Nu}(f) - \text{codim}_W \text{Im}(f) \\ &= \dim \text{Nu}(f) - (\dim W - \dim \text{Im}(f)) && \text{por la Proposición 1.10.9} \\ &= \dim V + \dim W && \text{por el Teorema 2.4.1,} \end{aligned}$$

así que en esta situación el índice no depende realmente de la función f sino solamente de las dimensiones de su dominio y codominio.

- (b) Sea $n \in \mathbb{N}_0$ y sea $f_n : p \in \mathbb{C}[X] \mapsto p^{(n)} \in \mathbb{C}[X]$, donde $p^{(n)}$ denota la derivada n -ésima de p . Por supuesto, tenemos que

$$f_n(X^i) = \begin{cases} 0 & \text{si } i < n; \\ i(i-1)\cdots(i-n+1)X^{i-n} & \text{si } i \geq n. \end{cases}$$

Usando esto es inmediato verificar que $\text{Nu}(f_n) = \langle 1, X, \dots, X^{n-1} \rangle$ y que $\text{Im}(f) = \mathbb{C}[X]$, de manera que

$$\text{ind}(f_n) = \dim \text{Nu}(f_n) - \text{codim}_{\mathbb{C}[X]} \text{Im}(f) = n.$$

Por otro lado, sea $g_n : p \in \mathbb{C}[X] \mapsto X^n p \in \mathbb{C}[X]$ la función dada por la multiplicación por X^n . Es inmediato ahora que $\text{Nu}(g_n) = 0$ y que $\text{Im}(g_n) = \langle X^i : i \in \mathbb{N}_0, i \geq n \rangle$. En particular, la imagen $\text{Im}(g_n)$ tiene a $\langle 1, X, \dots, X^{n-1} \rangle$ como complemento en $\mathbb{C}[X]$ y, por lo tanto,

$$\text{ind}(g_n) = \dim \text{Nu}(g_n) - \text{codim}_{\mathbb{C}[X]} \text{Im}(g_n) = -n.$$

Este ejemplo muestra que el índice puede tomar cualquier valor entero. \diamond

- 2.10.13.** El siguiente resultado es conocido como el *Teorema de aditividad del índice*.

Proposición. Si $f : U \rightarrow V$ y $g : V \rightarrow W$ son dos funciones lineales de tipo Fredholm, entonces la composición $g \circ f : U \rightarrow W$ también es de tipo Fredholm y su índice es

$$\text{ind}(g \circ f) = \text{ind}(g) + \text{ind}(f).$$

Demostración. Sean $f : U \rightarrow V$ y $g : V \rightarrow W$ dos funciones lineales de tipo Fredholm. De acuerdo a la Proposición 2.10.10, hay una sucesión exacta

$$0 \rightarrow \text{Nu}(f) \rightarrow \text{Nu}(g \circ f) \rightarrow \text{Nu}(h) \rightarrow \text{Con}(f) \rightarrow \text{Con}(g \circ f) \rightarrow \text{Con}(g) \rightarrow 0$$

Como f y g son de tipo Fredholm, sus núcleos $\text{Nu}(f)$ y $\text{Nu}(g)$ tienen dimensión finita y la exactitud de esta sucesión exacta junto con la Proposición 2.10.4 nos permiten concluir que $\text{Nu}(g \circ f)$ tiene dimensión finita. De la misma forma, como $\text{Con}(f)$ y $\text{Con}(g)$ tienen dimensión finita, la exactitud de la sucesión exactitud en $\text{Con}(g \circ f)$ implica que este espacio también tiene dimensión finita. Concluimos con todo esto que la composición $g \circ f$ es de tipo Fredholm. Más aún, como todos los espacios vectoriales que aparecen la sucesión exacta de arriba tienen dimensión finita, la Proposición 2.10.6 nos dice que

$$\begin{aligned} \dim \text{Nu}(f) - \dim \text{Nu}(g \circ f) + \dim \text{Nu}(h) - \dim \text{Con}(f) \\ + \dim \text{Con}(g \circ f) - \dim \text{Con}(g) = 0 \end{aligned}$$

y, por lo tanto,

$$\begin{aligned} \text{ind}(g \circ f) &= \dim \text{Nu}(g \circ f) - \dim \text{Con}(g \circ f) \\ &= \dim \text{Nu}(f) - \dim \text{Con}(f) + \dim \text{Nu}(g) - \dim \text{Con}(f) \\ &= \text{ind}(f) + \text{ind}(g), \end{aligned}$$

como afirma la proposición. □

§11. Subespacios invariantes

2.11.1. Sea V un espacio vectorial y sea $f : V \rightarrow V$ un endomorfismo de V . Un subespacio W de V es ***f-invariante*** si $f(W) \subseteq W$. Cuando ése es el caso, podemos considerar la función lineal $f_W : x \in W \mapsto f(x) \in W$ que se obtiene de f restringiendo a la vez el dominio y el codominio a W . Llamamos al endomorfismo f_W de W la ***restricción de f a W***.

Por otro lado, si W es f -invariante y $\pi : V \rightarrow V/W$ es la proyección canónica al cociente V/W , entonces $W \subseteq \text{Nu}(\pi)$ y la Proposición 2.8.6 nos dice que existe exactamente una función lineal $f^W : V/W \rightarrow V/W$ tal que $f([x]) = [f(x)]$ para todo $x \in V$.

2.11.2. Proposición. Sea V un espacio vectorial, sea $f : V \rightarrow V$ un endomorfismo de V y sea W un subespacio de V que es f -invariante.

- (i) Si f_W y f^W son inyectivas, entonces f es inyectiva.
- (ii) Si f_W y f^W son sobreyectivas, entonces f es sobreyectiva.
- (iii) Si f es inyectiva, entonces f_W es inyectiva y, si además W tiene dimensión finita, f^W es inyectiva.
- (iv) Si f es sobreyectiva, entonces f^W es sobreyectiva y, si además W tiene codimensión finita en V , f_W es sobreyectiva.

Demostración. (i) Supongamos que f_W y f^W son inyectivas, y sea $x \in V$ tal que $f(x) = 0$. En ese caso tenemos que $f^W([x]) = [f(x)] = 0$ en V/W , así que la inyectividad de f^W implica que $[x] = 0$, esto es, que $x \in W$. Pero entonces $f_W(x) = f(x) = 0$ y, como f_W es inyectiva, tiene que ser $x = 0$. Esto muestra que f es inyectiva.

(ii) Supongamos ahora que f_W y f^W son sobreyectivas, y sea $y \in V$. Como f^W es sobreyectiva, existe $x \in V$ tal que $f^W([x]) = [f(x)] = [y]$, de manera que $y - f(x) \in W$. Como f_W es sobreyectiva, existe $z \in W$ tal que $f_W(z) = f(z) = y - f(x)$. Como entonces $y = f(x + z)$, esto nos dice que y está en la imagen de f y, en definitiva, que f es sobreyectiva.

(iii) Supongamos que la función f es inyectiva. Si $x \in W$ es tal que $f_W(x) = f(x) = 0$, entonces por supuesto es $x = 0$ y, por lo tanto, f_W es inyectiva.

Supongamos ahora además que W tiene dimensión finita. Como $f_W : W \rightarrow W$ es inyectiva, como acabamos de probar, es también sobreyectiva. Sea $u \in V/W$ tal que $f^W(u) = 0$. Si $x \in V$ es un representante de u en V , de manera que $u = [x]$, entonces $f^W(u) = [f(x)] = 0$ y, como este elemento de V/W es nulo, tenemos que $f(x) \in W$. Como sabemos que f_W es sobreyectiva, esto nos dice que existe $y \in W$ tal que $f_W(y) = f(y) = f(x)$. Ahora bien, esto implica que $f(y - x) = 0$ y, como f es inyectiva, que $y - x = 0$, esto es, que $x = y \in W$. Así, es $u = [x] = [0]$. Esto prueba que f^W es inyectiva.

(iv) Finalmente, supongamos que f es sobreyectiva. Si $u \in V/W$ y $y \in V$ es un representante de u en V , entonces existe $x \in V$ tal que $f(x) = y$ y, en consecuencia, $f^W([x]) = [f(x)] = [y] = u$: vemos de esta forma que f^W es sobreyectiva.

Supongamos, para terminar, que W tiene codimensión finita. De acuerdo a la Proposición 2.8.9(i), el espacio V/W tiene dimensión finita y esto implica que el endomorfismo f^W de V/W es inyectivo, ya que es sobreyectivo. Sea $y \in W$. Como f es sobreyectiva, existe $x \in V$ tal que $f(x) = y$. Se tiene entonces que $f^W([x]) = [f(x)] = [y] = 0$, ya que $y \in W$ y, como f^W es inyectiva, que $[x] = 0$, esto es, que $x \in W$. Luego $f_W(x) = f(x) = y$ y la función f_W es sobreyectiva, como queremos \square

2.11.3. Sin las hipótesis de finitud que aparecen en ellas, las últimas dos partes de la Proposición 2.11.2 dejan de ser ciertas. Veamos un ejemplo de esto: sea $\mathbb{k}^{\mathbb{Z}}$ es el espacio vectorial de todas las funciones $\mathbb{Z} \rightarrow \mathbb{k}$, sea $f : \mathbb{k}^{\mathbb{Z}} \rightarrow \mathbb{k}^{\mathbb{Z}}$ es la función lineal tal que $f(\phi)(n) = \phi(n+1)$ para cada $\phi \in \mathbb{k}^{\mathbb{Z}}$ y cada $n \in \mathbb{Z}$, y sea W es el subespacio de $\mathbb{k}^{\mathbb{Z}}$ de las funciones $\phi : \mathbb{Z} \rightarrow \mathbb{k}$ tales que $\phi(n) = 0$

si $n \in \mathbb{N}$. Tenemos entonces:

- El subespacio W de $\mathbb{k}^{\mathbb{Z}}$ es f -invariante. En efecto, si $\phi \in W$ y $n \in \mathbb{N}$, entonces

$$f(\phi)(n) = \phi(n+1) = 0,$$

ya que $n+1$ también está en \mathbb{N} .

- La función f es un automorfismo de $\mathbb{k}^{\mathbb{Z}}$ y, de hecho, su inversa es la función $g : \mathbb{k}^{\mathbb{Z}} \rightarrow \mathbb{k}^{\mathbb{Z}}$ tal que $g(\phi)(n) = \phi(n-1)$ para cada $\phi \in \mathbb{k}^{\mathbb{Z}}$ y cada $n \in \mathbb{Z}$.
- La función f_W no es sobreyectiva. En efecto, si $\xi : \mathbb{Z} \rightarrow \mathbb{k}$ es la función tal que $\xi(0) = 1$ y $\xi(n) = 0$ si $n \in \mathbb{Z} \setminus \{0\}$, entonces el único elemento ζ de $\mathbb{k}^{\mathbb{Z}}$ tal que $f(\zeta) = \xi$ es $g(\xi)$, pero calculándolo se ve inmediatamente que este elemento no pertenece a W .
- La función f^W no es inyectiva. Por ejemplo, si $\eta : \mathbb{Z} \rightarrow \mathbb{k}$ es la función tal que $\eta(1) = 1$ y $\eta(n) = 0$ para todo $n \in \mathbb{Z} \setminus \{1\}$, entonces $\eta \notin W$, de manera que $[\eta] \neq 0$ en $\mathbb{k}^{\mathbb{Z}}/W$, pero $f^W([\eta]) = [f(\eta)] = 0$ en ese espacio, ya que $f(\eta) \in W$.

Esto implica, claro, que W no tiene ni dimensión finita ni codimensión finita en $\mathbb{k}^{\mathbb{Z}}$.

Este ejemplo también muestra que no podemos eliminar la hipótesis de finitud en la segunda parte del siguiente resultado:

2.11.4. Corolario. *Sea V un espacio vectorial, sea $f : V \rightarrow V$ un endomorfismo de V y sea W un subespacio de V que es f -invariante.*

- (i) Si f_W y f^W son isomorfismos, entonces f es un isomorfismo.
- (ii) Si f es un isomorfismo y V tiene dimensión finita, entonces f_W y f^W son isomorfismos.

Demostración. La primera parte es consecuencia inmediata de las partes (i) y (ii) de la Proposición 2.11.2, y la segunda de las partes (iii) y (iv) de ésta, ya que si V tiene dimensión finita, entonces W tiene tanto dimensión finita como codimensión finita en V . \square

2.11.5. Proposición. *Sea V un espacio vectorial de dimensión finita n y sea $f : V \rightarrow V$ un endomorfismo de V . Si W es un subespacio f -invariante de V de dimensión r y $\mathcal{B} = (x_1, \dots, x_n)$ es una base ordenadas de V tal que $\mathcal{B}' = (x_1, \dots, x_r)$ es una base ordenada de W , entonces $\mathcal{B}'' = ([x_{r+1}], \dots, [x_n])$ es una base ordenada del cociente V/W y la matriz de f con respecto a \mathcal{B} tiene una descomposición en bloques triangular superior de la forma*

$$[f]_{\mathcal{B}}^{\mathcal{B}} = \begin{pmatrix} [f_W]_{\mathcal{B}'}^{\mathcal{B}'} & C \\ 0 & [f^W]_{\mathcal{B}''}^{\mathcal{B}''} \end{pmatrix}$$

con C una matriz de $M_{r, n-r}(\mathbb{k})$.

Demostración. **HACER.** \square

Capítulo 3

Dualidad

§1. El espacio dual

3.1.1. Si V es un espacio vectorial, el *espacio dual* de V es el espacio vectorial $V^* := \text{hom}(V, \mathbb{k})$. Muchas veces llamamos a los vectores de V^* *funcionales*, para distinguirlos de los elementos de V .

Todo espacio vectorial tiene su correspondiente espacio dual, pero esta construcción tiene mejores propiedades cuando partimos de un espacio vectorial de dimensión finita. Es por eso que en este capítulo varios de los resultados importantes se ocupan solamente de espacios de dimensión infinita.

3.1.2. Buena parte de las cosas que hacemos con espacios duales depende de la siguiente observación:

Proposición. Sea V un espacio vectorial.

- (i) Si x es un elemento no nulo de V , entonces existe un funcional $\phi \in V^*$ tal que $\phi(x) \neq 0$.
- (ii) Más generalmente, si x es un elemento de V y S un subespacio de V tal que $x \notin S$, entonces existe $\phi \in V^*$ tal que $\phi(x) \neq 0$ y $\phi(y) = 0$ para todo $y \in S$.

Demostración. Bastará que probemos la segunda parte, ya que la primera se deduce de esta tomando $S = 0$. Sean entonces x y S un elemento y un subespacio de V , respectivamente, y supongamos que $x \notin S$. Como $x \neq 0$, $\{x\}$ es una base de $\langle x \rangle$, así que existe una función lineal $\phi_0 : \langle x \rangle \rightarrow \mathbb{k}$ tal que $\phi_0(x) = 1$. Por otro lado, como $x \notin S$, los espacios $\langle x \rangle$ y S son independientes, así que la suma $\langle x \rangle + S$ es directa. De acuerdo a la Proposición 2.1.7, existe una función lineal $\phi_1 : \langle x \rangle \oplus S \rightarrow \mathbb{k}$ tal que $\phi_1(y) = \phi_0(y)$ para todo $y \in \langle x \rangle$ y $\phi_1(z) = 0$ para todo $z \in S$. Finalmente, la Proposición 2.1.9 nos dice que podemos extender ϕ_1 a una función lineal sobre todo V , esto es, que hay una función lineal $\phi : V \rightarrow \mathbb{k}$ tal que $\phi(y) = \phi_1(y)$ para todo $y \in \langle x \rangle \oplus S$. Esta función ϕ es un elemento de V^* que se anula sobre S y no sobre x , como queremos. \square

3.1.3. Si tenemos una base de un espacio vectorial de dimensión finita es fácil construir otra de su

espacio dual:

Proposición. *Sea V un espacio vectorial.*

- (i) *Si V tiene dimensión finita, entonces el espacio dual V^* también tiene dimensión finita y, de hecho, se tiene que $\dim V = \dim V^*$ y que $V \cong V^*$.*
- (ii) *Si $n = \dim V$ y $\mathcal{B} = (x_1, \dots, x_n)$ es una base ordenada de V , entonces para cada $i \in \llbracket n \rrbracket$ existe exactamente una función lineal $\phi_i : V \rightarrow \mathbb{k}$ tal que para cada $j \in \llbracket n \rrbracket$*

$$\phi_i(x_j) = \begin{cases} 1, & \text{si } i = j; \\ 0, & \text{en caso contrario.} \end{cases} \quad (1)$$

El conjunto ordenado $\mathcal{B}^ = (\phi_1, \dots, \phi_n)$ es una base ordenada de V^* .*

Llamamos a la base ordenada \mathcal{B}^* del espacio dual V^* descripta en esta proposición la **base dual** a la base \mathcal{B} de V . No hay un resultado análogo al de esta proposición para espacios vectoriales de dimensión infinita.

Demostración. De la Proposición 2.6.5 sabemos que

$$\dim V^* = \dim \hom(V, \mathbb{k}) = \dim V \cdot \dim \mathbb{k} = \dim V$$

y el Corolario 2.3.9 nos dice entonces que $V \cong V^*$. Que para cada $i \in \llbracket n \rrbracket$ existe exactamente una función lineal $\phi_i : V \rightarrow \mathbb{k}$ que satisface la condición (1) del enunciado se sigue de la Proposición 2.1.5. Veamos, para completar la prueba, que $\mathcal{B}^* = (\phi_1, \dots, \phi_n)$ es una base ordenada de V^* . Notemos que como \mathcal{B}^* tiene exactamente n elementos —ya que las funciones ϕ_1, \dots, ϕ_n son claramente dos a dos distintas— es suficiente para ello con mostrar que el conjunto \mathcal{B}^* genera a V^* .

Sea $\phi \in V^*$ y consideremos el elemento $\psi = \phi(x_1)\phi_1 + \dots + \phi(x_n)\phi_n$ de V^* . Si $i \in \llbracket n \rrbracket$, entonces

$$\psi(x_i) = (\phi(x_1)\phi_1 + \dots + \phi(x_n)\phi_n)(x_i) = \phi(x_1)\phi_1(x_i) + \dots + \phi(x_n)\phi_n(x_i) = \phi(x_i)$$

y esto nos dice que las funciones lineales $\phi, \psi : V \rightarrow \mathbb{k}$ coinciden sobre los elementos de la base \mathcal{B} : de acuerdo a la Proposición 2.1.5, entonces, es $\phi = \psi$. Como claramente ψ pertenece a $\langle \mathcal{B}^* \rangle$, vemos que \mathcal{B}^* genera a V^* , como queremos. \square

3.1.4. Cuando tenemos un par de bases ordenadas duales \mathcal{B} y \mathcal{B}^* para un espacio vectorial de dimensión finita y su dual, expresar en coordenadas vectores y funcionales es algo que puede hacerse de una forma muy sencilla:

Proposición. *Sea V un espacio vectorial de dimensión finita n . Si $\mathcal{B} = (x_1, \dots, x_n)$ es una base ordenada de \mathcal{B} y $\mathcal{B}^* = (\phi_1, \dots, \phi_n)$ es la base ordenada de V^* dual a \mathcal{B} , entonces para todo $x \in V$ se tiene que*

$$x = \phi_1(x)x_1 + \dots + \phi_n(x)x_n$$

y para todo $\phi \in V^*$ se tiene que

$$\phi = \phi(x_1)\phi_1 + \cdots + \phi(x_n)\phi_n.$$

La conclusión de esta proposición es que $[x]_{\mathcal{B}} = (\phi_1(x), \dots, \phi_n(x))^t$ para todo $x \in V$ y que $[\phi]_{\mathcal{B}^*} = (\phi(x_1), \dots, \phi(x_n))^t$ para todo $\phi \in V^*$.

Demostración. Sea $x \in V$. Como \mathcal{B} es una base, sabemos que existen escalares $a_1, \dots, a_n \in \mathbb{k}$ tales que $x = a_1x_1 + \cdots + a_nx_n$, y para cada $i \in \llbracket n \rrbracket$ se tiene que

$$a_i = a_1\phi_i(x_1) + \cdots + a_n\phi_i(x_n) = \phi_i(a_1x_1 + \cdots + a_nx_n) = \phi_i(x).$$

Esto prueba la primera afirmación. A la segunda la probamos en la demostración de la Proposición 3.1.3. \square

3.1.5. Ejemplo. La Proposición 3.1.4 nos permite explicitar los elementos de la base dual de una base de un espacio vectorial. Veamos un ejemplo de esto.

Fijemos $n \in \mathbb{N}$ y consideremos el espacio vectorial $V = \mathbb{k}^n$ y su base ordenada estándar $\mathcal{B} = (e_1, \dots, e_n)$. De acuerdo a la Proposición 3.1.3, el espacio dual V^* tiene la correspondiente dual $\mathcal{B}^* = (\phi_1, \dots, \phi_n)$. ¿Cuáles son exactamente estas funciones lineales $\phi_1, \dots, \phi_n : V \rightarrow \mathbb{k}$? Observemos que si $x = (x_1, \dots, x_n)$ es un elemento de V , entonces por un lado tenemos que

$$x = x_1e_1 + \cdots + x_ne_n$$

y, por otro, de acuerdo a la Proposición 3.1.4,

$$x = \phi_1(x)e_1 + \cdots + \phi_n(x)e_n.$$

Como \mathcal{B} es una base, esto nos dice que $\phi_i(x) = x_i$ para cada $i \in \llbracket n \rrbracket$. Así, las funciones $\phi_1, \dots, \phi_n : V \rightarrow \mathbb{k}$ son las **funciones coordenadas**.

Consideremos ahora la base ordenada $\mathcal{B}' = (f_1, \dots, f_n)$ de \mathbb{k}^n que tiene, para cada $i \in \llbracket n \rrbracket$,

$$f_i = (0, \dots, 1, \dots, 1),$$

con las primeras $i - 1$ componentes nulas y todas las demás iguales a 1, y sea $\mathcal{B}'^* = (\psi_1, \dots, \psi_n)$ la base de V^* dual a \mathcal{B}' . Si $x = (x_1, \dots, x_n)$ es un elemento de V , es fácil ver que

$$x = x_1f_1 + (x_2 - x_1)f_2 + \cdots + (x_n - x_{n-1})f_n$$

y, al mismo tiempo, la Proposición 3.1.4 nos dice que

$$x = \psi_1(x)f_1 + \psi_2(x)f_2 + \cdots + \psi_n(x)f_n,$$

así que para cada $i \in \llbracket n \rrbracket$ tenemos que

$$\psi_i(x) = \begin{cases} x_1 & \text{si } i = 1; \\ x_i - x_{i-1} & \text{si } 1 < i \leq n. \end{cases}$$

En términos de la base \mathcal{B}^* que encontramos antes, esto nos dice que

$$\psi_1 = \phi_1 \quad y \quad \psi_i = \phi_i - \phi_{i-1} \quad \text{para cada } i \in \llbracket 2, n \rrbracket.$$

◇

3.1.6. La base ordenada dual a una base ordenada de un espacio vectorial depende claramente de ésta última: al cambiar de base cambiamos de base dual, y la siguiente proposición nos dice exactamente cómo.

Proposición. *Sea V un espacio vectorial de dimensión finita. Si \mathcal{B} y \mathcal{B}' son dos bases ordenadas de V y \mathcal{B}^* y \mathcal{B}'^* son las correspondientes bases ordenadas duales, entonces*

$$C(\mathcal{B}'^*, \mathcal{B}^*) = C(\mathcal{B}, \mathcal{B}')^t.$$

Aquí $C(\mathcal{B}, \mathcal{B}')^t$ denota, como siempre, la matriz transpuesta de $C(\mathcal{B}, \mathcal{B}')$.

Demostración. Sea $n = \dim V$ y sean $\mathcal{B} = (x_1, \dots, x_n)$, $\mathcal{B}' = (y_1, \dots, y_n)$, $\mathcal{B}^* = (\phi_1, \dots, \phi_n)$ y $\mathcal{B}'^* = (\psi_1, \dots, \psi_n)$. Supongamos además que $C(\mathcal{B}, \mathcal{B}') = (c_{i,j})$, de manera que para cada $j \in \llbracket n \rrbracket$ se tiene que $x_j = c_{1,j}y_1 + \dots + c_{n,j}y_n$. Si $k \in \llbracket n \rrbracket$, entonces para cada $j \in \llbracket n \rrbracket$ es

$$\psi_k(x_j) = \psi_k(c_{1,j}y_1 + \dots + c_{n,j}y_n) = c_{1,j}\psi_k(y_1) + \dots + c_{n,j}\psi_k(y_n) = c_{k,j}$$

y usando la última afirmación de la Proposición 3.1.4 vemos que

$$\psi_k = \psi_k(x_1)\phi_1 + \dots + \psi_k(x_n)\phi_n = c_{k,1}\phi_1 + \dots + c_{k,n}\phi_n.$$

Si $C(\mathcal{B}'^*, \mathcal{B}^*) = (d_{i,j})$, esto nos dice que $d_{i,j} = c_{j,i}$ para cada $i, j \in \llbracket n \rrbracket$ y, por lo tanto, prueba la proposición. □

3.1.7. Sea V un espacio vectorial. Para cada $x \in V$, hay una función

$$\Phi(x) : \phi \in V^* \mapsto \phi(x) \in \mathbb{k}$$

dada por la evaluación de los funcionales de V^* en x , y es fácil ver que se trata de una función lineal, de manera que $\Phi(x) \in V^{**}$, el **espacio bidual** de V , esto es, el espacio dual del espacio dual de V . Obtenemos de esta forma una función

$$\Phi : V \rightarrow V^{**},$$

a la que llamamos el **morfismo canónico**.

Proposición. *Sea V un espacio vectorial. El morfismo canónico $\Phi : V \rightarrow V^{**}$ es lineal e inyectivo. Si V tiene dimensión finita, este morfismo Φ es un isomorfismo.*

Demostración. Veamos primero que Φ es lineal: si $x, y \in V$ y $\alpha, \beta \in \mathbb{k}$, entonces para cada

funcional $\phi \in V^*$ se tiene que

$$\begin{aligned}\Phi(\alpha x + \beta y)(\phi) &= \phi(\alpha x + \beta y) \\ &= \alpha\phi(x) + \beta\phi(y) \\ &= \alpha\Phi(x)(\phi) + \beta\Phi(y)(\phi) \\ &= (\alpha\Phi(x) + \beta\Phi(y))(\phi)\end{aligned}$$

y esto significa que $\Phi(\alpha x + \beta y) = \alpha\Phi(x) + \beta\Phi(y)$.

Sea x un elemento no nulo de V . De acuerdo a la Proposición 3.1.2, existe un funcional $\phi \in V^*$ tal que $\phi(x) \neq 0$, y entonces $\Phi(x) \neq 0$, ya que $\Phi(x)(\phi) = \phi(x) \neq 0$: esto nos dice que $x \notin \text{Nu}(\Phi)$ y prueba que la función Φ es inyectiva.

Finalmente, si V tiene dimensión finita, entonces V^{**} tiene la misma dimensión y la función $\Phi : V \rightarrow V^{**}$, que es inyectiva, es necesariamente un isomorfismo. \square

3.1.8. Una aplicación de la Proposición 3.1.7 es la construcción recíproca a la de las bases duales:

Proposición. Sea V un espacio vectorial de dimensión finita. Si \mathcal{B}^* es una base ordenada del espacio dual V^* , entonces existe una base \mathcal{B} ordenada de V tal que \mathcal{B}^* es la base ordenada dual a \mathcal{B} .

Llamamos a la base ordenada \mathcal{B} de V la base **predual** a \mathcal{B}^* .

Demostración. Sea $n = \dim V$, supongamos que $\mathcal{B}^* = (\phi_1, \dots, \phi_n)$ y sea $\mathcal{B}^{**} = (\psi_1, \dots, \psi_n)$ la base de V^{**} dual a \mathcal{B}^* . La función $\Phi : V \rightarrow V^{**}$ de la Proposición 3.1.7 es un isomorfismo así que su función inversa $\Phi^{-1} : V^{**} \rightarrow V$ también lo es. Para cada $i \in [n]$ sea $x_i = \Phi^{-1}(\psi_i)$ y sea $\mathcal{B} = (x_1, \dots, x_n)$. Como Φ^{-1} es un isomorfismo, la Proposición 2.3.6 nos dice que \mathcal{B} es una base ordenada de V . Veamos que, además, su base dual es precisamente la base \mathcal{B}^* de V^* con la que empezamos. En efecto, si i y j son elementos de $[n]$, entonces $\Phi(x_i) = \psi_i$ y, por lo tanto,

$$\phi_i(x_j) = \Phi(x_j)(\phi_i) = \psi_j(\phi_i) = \begin{cases} 1 & \text{si } i = j; \\ 0 & \text{si no.} \end{cases}$$

La prueba de la proposición queda con esto completa. \square

3.1.9. Como observamos arriba, la Proposición 3.1.3 nos dice que el dual de un espacio vectorial de dimensión finita tiene dimensión finita y, de hecho, que ambos espacios tienen la misma dimensión, de manera que son isomorfos. Es posible probar que

si V es un espacio vectorial de dimensión infinita, entonces $V \neq V^*$.

Una discusión de esto puede encontrarse en [Gin10]. Allí se presentan varias demostraciones de que cuando V es un espacio vectorial de dimensión infinita no solamente $V \neq V^*$ sino que, además, el espacio dual V^* tiene dimensión infinita y *estrictamente más grande* que la de V — para hacer esto preciso es necesario definir qué es la dimensión de un espacio vectorial de dimensión infinita, por supuesto, y eso no lo hicimos en estas notas. Como consecuencia de esto, el bidual

V^{**} tiene también dimensión estrictamente más grande que la de V^* y, en particular, también $V \not\cong V^{**}$ en este caso. Por supuesto, esto nos dice, en particular, que la función $\Phi : V \rightarrow V^{**}$ de la Proposición 3.1.7 no es un isomorfismo cuando V tiene dimensión infinita.

3.1.10. Ejemplo. Veamos un ejemplo en el que podemos describir explícitamente —a menos de un isomorfismo— el espacio dual de un espacio de dimensión infinita.

Digamos que una función $f : \mathbb{N} \rightarrow \mathbb{k}$ tiene **soporte finito** si existe $n_0 \in \mathbb{N}$ tal que $f(n) = 0$ siempre que $n > n_0$ —decimos también que f se anula en *casi todos* los elementos de \mathbb{N} , esto es, en todos salvo un número finito de ellos— y sea \mathcal{B} el subconjunto de $\mathbb{k}^{\mathbb{N}}$ de todas las funciones $\mathbb{N} \rightarrow \mathbb{k}$ que tienen soporte finito. Es inmediato verificar que se trata de un subespacio vectorial de $\mathbb{k}^{\mathbb{N}}$. Para cada $k \in \mathbb{N}$ sea $\delta_k : \mathbb{N} \rightarrow \mathbb{k}$ la función tal que

$$\delta_k(n) = \begin{cases} 1 & \text{si } n = k; \\ 0 & \text{si no.} \end{cases}$$

Es claro que el conjunto $\mathcal{B} = \{\delta_k : k \in \mathbb{N}\}$ está contenido en V y sabemos que es linealmente independiente —vimos en el Ejemplo 1.5.7(b) que es linealmente independiente de $\mathbb{k}^{\mathbb{N}}$. Más aún se trata de una base de V . En efecto, si $f : \mathbb{N} \rightarrow \mathbb{k}$ es un elemento de V , de manera que existe $n_0 \in \mathbb{N}$ con $f(n) = 0$ siempre que $n > n_0$, entonces es

$$f = \sum_{k=1}^{n_0} f(k) \delta_k \in \langle \mathcal{B} \rangle.$$

Si $\phi : V \rightarrow \mathbb{k}$ es un elemento del espacio dual V^* , podemos considerar la función $T(\phi) : \mathbb{N} \rightarrow \mathbb{k}$ tal que $T(\phi)(n) = \phi(\delta_n)$ para todo $n \in \mathbb{N}$: se trata claramente de un elemento de $\mathbb{k}^{\mathbb{N}}$. Obtenemos de esta forma una función

$$T : \phi \in V^* \mapsto T(\phi) \in \mathbb{k}^{\mathbb{N}}.$$

Esta función es un isomorfismo. Primero, si ϕ y ψ son elementos de V^* y α y β escalares de \mathbb{k} , entonces para cada $n \in \mathbb{N}$ se tiene que

$$\begin{aligned} T(\alpha\phi + \beta\psi)(n) &= (\alpha\phi + \beta\psi)(\delta_n) = \alpha\phi(\delta_n) + \beta\psi(\delta_n) \\ &= \alpha T(\phi)(n) + \beta T(\psi)(n) = (\alpha T(\phi) + \beta T(\psi))(n). \end{aligned}$$

Esto nos dice que $T(\alpha\phi + \beta\psi) = \alpha T(\phi) + \beta T(\psi)$, esto es, que T es lineal. Si $\phi \in V^*$ es tal que $T(\phi) = 0$, entonces para todo $n \in \mathbb{N}$ tenemos que $\phi(n) = T(\phi)(n) = 0$ y, por lo tanto, la función lineal $\phi : V \rightarrow \mathbb{k}$ se anula en cada elemento de la base \mathcal{B} de su dominio. Esto implica, como sabemos, que $\phi = 0$ y, por lo tanto, que la función T es inyectiva.

Finalmente, supongamos que $f : \mathbb{N} \rightarrow \mathbb{k}$ es un elemento de $\mathbb{k}^{\mathbb{N}}$. Como el conjunto \mathcal{B} es una base de V , sabemos que existe una función lineal $\phi : V \rightarrow \mathbb{k}$ tal que $\phi(\delta_n) = f(n)$ y es claro que $T(\phi) = f$. Vemos así que la función T también es sobreyectiva. Observemos que esto junto nuestra discusión de 3.1.9 implica que no existe ningún isomorfismo de V a $\mathbb{k}^{\mathbb{N}}$. \diamond

§2. Anuladores

Anulador a izquierda

3.2.1. Sea V es un espacio vectorial y sea V^* su espacio dual. Si S es un subconjunto de V^* , el **anulador a izquierda** de S es el conjunto

$${}^\circ S = \{x \in V : \phi(x) = 0 \text{ para todo } \phi \in S\}.$$

3.2.2. Ejemplo. Sea $n \in \mathbb{N}$ y $V = \mathbb{k}^n$. Si $\phi \in V^*$, de manera que $\phi : V \rightarrow \mathbb{k}$ es una función lineal, sabemos que existen escalares $a_1, \dots, a_n \in \mathbb{k}$, bien determinados por ϕ , tales que

$$\phi(x) = a_1x_1 + \dots + a_nx_n$$

para cada vector $x = (x_1, \dots, x_n) \in V$. Esto significa que $x = (x_1, \dots, x_n)$ es tal que $\phi(x) = 0$ si y solamente si x es solución de la ecuación lineal homogénea

$$a_1X_1 + \dots + a_nX_n = 0.$$

Así, un subconjunto S de V^* puede verse en este ejemplo como un conjunto de ecuaciones lineales y su anulador a izquierda ${}^\circ S$ es el conjunto de todas las soluciones comunes de esas ecuaciones. \diamond

3.2.3. Proposición. Sea V un espacio vectorial y sea V^* su espacio dual.

- (i) Es ${}^\circ 0 = V$ y ${}^\circ(V^*) = 0$.
- (ii) Si S y T son subconjuntos de V^* y $S \subseteq T$, entonces ${}^\circ S \supseteq {}^\circ T$.
- (iii) Si S es un subconjunto de V^* , entonces ${}^\circ S$ es un subespacio de V y coincide con ${}^\circ\langle S \rangle$.
- (iv) Si S y T son subespacios de V^* , entonces

$${}^\circ(S + T) = {}^\circ S \cap {}^\circ T, \quad {}^\circ S + {}^\circ T \subseteq {}^\circ(S \cap T). \quad (2)$$

La inclusión de (2) es de hecho una igualdad cuando V tiene dimensión finita —la Proposición 3.2.5 de más abajo afirma precisamente eso— pero esto no es cierto en general. Daremos un ejemplo que muestra esto en 3.2.8.

Demostración. (i) El único elemento del subespacio nulo de V^* es la función nula y ésta que se anula en todo elemento de V : esto significa que ${}^\circ 0 = V$. Por otro lado, si x es un elemento no nulo de V , de acuerdo a la Proposición 3.1.2 existe $\phi \in V^*$ tal que $\phi(x) \neq 0$ y, por lo tanto, tenemos que $x \notin {}^\circ(V^*)$. Como claramente $0 \in {}^\circ(V^*)$, vemos que ${}^\circ(V^*) = 0$.

(ii) Sea S y T subconjuntos de V^* tales que $S \subseteq T$ y sea $x \in {}^\circ T$. Si $\phi \in S$, entonces $\phi \in T$ y, por lo tanto, $\phi(x) = 0$: esto nos dice que $x \in {}^\circ S$.

(iii) Sea S un subconjunto de V^* . La definición de ${}^\circ S$ implica inmediatamente que

$${}^\circ S = \bigcap_{\phi \in S} \text{Nu}(\phi),$$

así que ${}^{\circ}S$ es un subespacio de V por la Proposición 1.3.5(ii). Como $S \subseteq \langle S \rangle$, la parte (ii) implica que ${}^{\circ}\langle S \rangle \subseteq {}^{\circ}S$. Recíprocamente, si $x \in {}^{\circ}S$, entonces para cada $\phi \in \langle S \rangle$ existen $k \in \mathbb{N}_0$, $\phi_1, \dots, \phi_k \in S$ y $a_1, \dots, a_k \in \mathbb{k}$ tales que $\phi = a_1\phi_1 + \dots + a_k\phi_k$ y, por lo tanto,

$$\phi(x) = (a_1\phi_1 + \dots + a_k\phi_k)(x) = a_1\phi_1(x) + \dots + a_k\phi_k(x) = 0,$$

de manera que $x \in {}^{\circ}\langle S \rangle$.

(iv) Sean S y T subespacios de V^* . Como $S \subseteq S + T$ y $T \subseteq S + T$, lo que ya probamos implica que ${}^{\circ}(S + T) \subseteq {}^{\circ}S$ y ${}^{\circ}(S + T) \subseteq {}^{\circ}T$ y, por lo tanto, que ${}^{\circ}(S + T) \subseteq {}^{\circ}S \cap {}^{\circ}T$. Por otro lado, si $x \in {}^{\circ}S \cap {}^{\circ}T$, entonces para cada $\phi \in S + T$ existen $\sigma \in S$ y $\tau \in T$ tales que $\phi = \sigma + \tau$ y

$$\phi(x) = \sigma(x) + \tau(x) = 0,$$

de manera que $x \in {}^{\circ}S$. Vemos así que vale la igualdad de (2).

Como $S \cap T \subseteq S$ y $S \cap T \subseteq T$, tenemos que ${}^{\circ}S \subseteq {}^{\circ}(S \cap T)$ y ${}^{\circ}T \subseteq {}^{\circ}(S \cap T)$. Como ${}^{\circ}(S \cap T)$ es un subespacio de V , esto implica que ${}^{\circ}S + {}^{\circ}T \subseteq {}^{\circ}(S \cap T)$. \square

3.2.4. Cuando el espacio ambiente V tiene dimensión finita, hay una relación muy sencilla entre la dimensión de un subespacio de V^* y la de su anulador a izquierda:

Proposición. *Sea V un espacio vectorial de dimensión finita. Si S es un subespacio de V^* , entonces tanto S como su anulador a izquierda ${}^{\circ}S$ tienen dimensión finita y*

$$\dim S + \dim {}^{\circ}S = \dim V. \tag{3}$$

Demostración. Sea S un subespacio de V^* . Como S y ${}^{\circ}S$ son subespacios de V^* y de V , respectivamente, que tienen dimensión finita, sabemos que S y ${}^{\circ}S$ tienen ellos también dimensión finita. Sean $n = \dim V$ y $s = \dim S$. Sea (ϕ_1, \dots, ϕ_s) una base ordenada de S y sean $\phi_{s+1}, \dots, \phi_n \in V^*$ tales que $\mathcal{B}^* = (\phi_1, \dots, \phi_n)$ es una base ordenada de V^* . De acuerdo a la Proposición 3.1.8, existe una base ordenada $\mathcal{B} = (x_1, \dots, x_n)$ de V que tiene a \mathcal{B}^* por base dual. Para ver que vale la igualdad (3) del enunciado, es suficiente con que mostremos que $\mathcal{B}' = (x_{s+1}, \dots, x_n)$ es una base ordenada de ${}^{\circ}S$. Como \mathcal{B}' es un subconjunto de \mathcal{B} , es linealmente independiente, y sólo queda mostrar que genera a ${}^{\circ}S$.

Sea para ello $x \in {}^{\circ}S$. Como \mathcal{B} es una base de V , existen escalares $a_1, \dots, a_n \in \mathbb{k}$ tales que

$$x = a_1x_1 + \dots + a_nx_n.$$

Si $i \in [s]$, el funcional ϕ_i está en S y, por lo tanto, se anula en x , así que

$$0 = \phi_i(x) = \phi_i(a_1x_1 + \dots + a_nx_n) = a_1\phi_i(x_1) + \dots + a_n\phi_i(x_n) = a_i.$$

Esto nos dice que, de hecho, $x = a_{s+1}x_{s+1} + \dots + a_nx_n \in \langle \mathcal{B}' \rangle$ y completa la prueba. \square

3.2.5. Podemos ahora mejorar la última afirmación de la Proposición 3.2.3:

Proposición. *Sea V un espacio vectorial de dimensión finita. Si S y T son dos subespacios de V , entonces ${}^{\circ}S + {}^{\circ}T = {}^{\circ}(S \cap T)$.*

Demostración. Sabemos de la Proposición 3.2.3 que ${}^{\circ}S + {}^{\circ}T \subseteq {}^{\circ}(S \cap T)$, así que como ambos espacios tienen dimensión finita para probar el corolario es suficiente que mostremos que tienen a misma dimensión. Usando las Proposiciones 3.2.4 y 1.8.4 vemos que

$$\begin{aligned}\dim {}^{\circ}(S + T) &= \dim V - \dim(S + T) \\ &= \dim V - \dim S - \dim T + \dim S \cap T\end{aligned}$$

así que

$$\begin{aligned}\dim({}^{\circ}S + {}^{\circ}T) &= \dim {}^{\circ}S + \dim {}^{\circ}T - \dim({}^{\circ}S \cap {}^{\circ}T) \\ &= (\dim V - \dim S) + (\dim V - \dim T) - (\dim V - \dim S - \dim T + \dim S \cap T) \\ &= \dim V - \dim S \cap T \\ &= \dim {}^{\circ}(S \cap T),\end{aligned}$$

como queremos. \square

3.2.6. Ejemplo. Demos un ejemplo sencillo de para qué sirve la Proposición 3.2.4.

Sea $n \in \mathbb{N}$ y sea $V := \mathbb{k}^n$. Si $m \in \mathbb{N}$ y ϕ_1, \dots, ϕ_m son elementos de V^* , podemos considerar el subconjunto $E := \{\phi_1, \dots, \phi_m\}$ de V^* y el subespacio $S := \langle E \rangle$ que genera. Para cada $i \in [m]$ existen escalares $a_{i,1}, \dots, a_{i,n} \in \mathbb{k}$ tales que

$$\phi_i(x) = a_{i,1}x_1 + \dots + a_{i,n}x_n$$

siempre que $x = (x_1, \dots, x_n) \in \mathbb{k}^n$. Sabemos que ${}^{\circ}E = {}^{\circ}S$ y es claro que ${}^{\circ}E$ es el conjunto de las soluciones $x = (x_1, \dots, x_n)$ al sistema de m ecuaciones lineales homogéneas

$$\left\{ \begin{array}{l} a_{1,1}X_1 + \dots + a_{1,n}X_n = 0, \\ a_{2,1}X_1 + \dots + a_{2,n}X_n = 0, \\ \vdots \\ a_{m,1}X_1 + \dots + a_{m,n}X_n = 0. \end{array} \right.$$

La Proposición 3.2.4 nos permite calcular la dimensión del conjunto ${}^{\circ}S$ de las soluciones de este sistema: es

$$\dim {}^{\circ}S = \dim V - \dim S = n - \dim S.$$

Para calcular esto precisamente necesitamos, claro, conocer la dimensión de S . Pero, por ejemplo, como S está generado por m elementos, sabemos que $\dim S \leq m$ y, por lo tanto, que $\dim {}^{\circ}S \geq n - m$: esto nos dice que

el sistema de ecuaciones tiene soluciones no triviales —es decir, no nulas— si $m < n$.

En otras palabras, si hay más incógnitas que ecuaciones hay soluciones no triviales. \diamond

3.2.7. Ejemplo. Veamos ahora una aplicación bien típica de la Proposición 3.2.4 a la geometría. Sea V el espacio de los polinomio de $\mathbb{k}[X, Y]$ nulos o de grado a lo sumo 2, esto es, los polinomio de la forma

$$a_1 + a_2 X + a_3 Y + a_4 X^2 + a_5 XY + a_6 Y^2$$

con $a_1, \dots, a_6 \in \mathbb{k}$. Es inmediato verificar que el conjunto $\{1, X, Y, X^2, XY, Y^2\}$ es una base de V , así que $\dim V = 6$.

Si $P = (x, y)$ es un punto de \mathbb{k}^2 , la función

$$\phi_P : f \in V \mapsto f(x, y) \in \mathbb{k}$$

es lineal, así que se trata de un elemento del espacio dual V^* . Si $f \in V$, entonces $\phi_P(f) = 0$ exactamente cuando $f(x, y) = 0$, y esto ocurre si y solamente si el punto P pertenece al lugar geométrico de los ceros de f en \mathbb{k}^2 ,

$$Z(f) := \{(x, y) \in \mathbb{k}^2 : f(x, y) = 0\}.$$

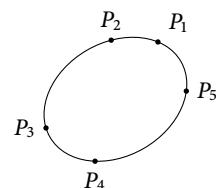
Observemos que si f no es el polinomio nulo este lugar geométrico es una recta de \mathbb{k}^2 , si es que f tiene grado 1, o una cónica (posiblemente degenerada), si el grado de f es 2.

Sean ahora P_1, \dots, P_5 cinco puntos distintos dos a dos de \mathbb{k}^2 tales que no hay tres de ellos sobre una misma recta y llamemos $S = \langle \phi_{P_1}, \dots, \phi_{P_5} \rangle$ al subespacio de V^* que generan los funcionales asociados. Es $\dim S \leq 5$ y, por lo tanto,

$$\dim {}^\circ S = \dim V - \dim S \geq 6 - 5 = 1.$$

Vemos así que existe un polinomio no nulo f de V que pertenece a ${}^\circ S$, y esto significa que el lugar geométrico $Z(f)$ contiene a los cinco puntos P_1, \dots, P_5 . Este polinomio no puede ser constante, porque se anula en P_1 y no es nulo, y no puede tener grado 1, porque sino los cinco puntos estarían alineados: trata por lo tanto de un polinomio de grado 2. Por otro lado, el polinomio no puede ser reducible: si lo fuera, sería igual a un producto gh de dos polinomios de grado 1 y los 5 puntos estarían sobre la unión de las dos rectas $Z(g)$ y $Z(h)$, pero esto es imposible porque necesariamente alguna de esas dos rectas debería contener a al menos tres de los cinco puntos. La conclusión de esto es que

por cinco puntos de \mathbb{k}^2 entre los cuales
no hay tres alineados pasa al menos una
cónica irreducible



y no es difícil ver que, de hecho, hay exactamente *una* tal cónica¹

¹Para verlo, observemos primero que el cuerpo \mathbb{k} tiene que tener al menos tres elementos, porque si tiene solo

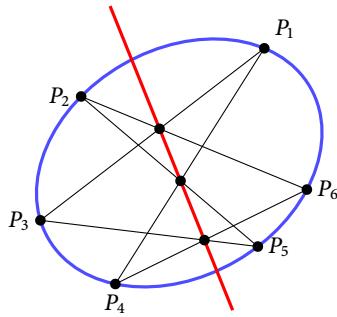


Figura 3.1. El Teorema de Pascal afirma que si P_1, \dots, P_6 son seis puntos distintos dos a dos sobre una cónica, entonces los tres puntos $\overline{P_1P_3} \cap \overline{P_2P_5}$, $\overline{P_1P_4} \cap \overline{P_2P_6}$ y $\overline{P_3P_5} \cap \overline{P_4P_6}$ están alineados. La fuente original del teorema es el manuscrito [Pas40] de *Blaise Pascal*, escrito cuando Pascal tenía 16 años.. Una demostración sencilla puede encontrarse en [vYz93].

Si en lugar de tener cinco puntos tenemos seis, entonces el subespacio S de V^* generado por las correspondientes funcionales puede tener en principio dimensión 6, y cuando ese es el caso la dimensión de ${}^\circ S$ es 0: en ese caso no hay ninguna cónica que pase por los seis puntos. De hecho, el Teorema de Pascal, que ilustramos en la Figura 3.1, impone una condición no trivial que satisfacen seis puntos de una cónica: en general, seis puntos arbitrarios del plano no la satisfacen, y cuando eso ocurre el espacio ${}^\circ S$ tiene dimensión nula. \diamond

3.2.8. Ejemplo. Mostremos que la igualdad de la Proposición 3.2.5 en general no vale. Sea V el espacio de las funciones $\mathbb{N} \rightarrow \mathbb{k}$ de soporte finito, como en el Ejemplo 3.1.10. Para cada $n \in \mathbb{N}$ consideremos la función

$$\phi_n : f \in V \mapsto f(n+1) - f(1) \in \mathbb{k},$$

que claramente es un elemento de V^* , y el subespacio $S = \langle \phi_n : n \in \mathbb{N} \rangle$. Si $f \in V$ es un elemento de ${}^\circ S$, entonces para cada $n \in \mathbb{N}$ tenemos que $0 = \phi_n(f) = f(n+1) - f(1)$, así que la función f es constante de valor $f(1)$. Como además tiene soporte finito, es claro que es idénticamente nula. Vemos así que ${}^\circ S = 0$.

dos no hay cinco puntos distintos dos a dos en \mathbb{k}^2 . Podemos suponer sin pérdida de generalidad que $P_1 = (0, 0)$ y que $P_2 = (\alpha, 0)$ para algún escalar α no nulo — en efecto, si ese no es el caso, podemos hacer un cambio lineal de coordenadas que nos lleve a esa situación. Supongamos que f y g son dos elementos de V linealmente independientes que se anulan en los cinco puntos y elijamos $\beta \in \mathbb{k} \setminus \{0, \alpha\}$ y $Q = (\beta, 0)$; observemos que podemos elegir β precisamente porque \mathbb{k} tiene más que dos elementos. Es fácil ver que hay dos escalares a y b en \mathbb{k} tal que el polinomio $h = af + bg$ se anula en Q . El polinomio $h(X, 0) \in \mathbb{k}[X]$ es o nulo o tiene grado a lo sumo 2: como h se anula en P_1 , en P_2 y en Q , el polinomio $h(X, 0)$ se anula en 0, en α y en β . Vemos así que $h(X, 0)$ tiene que ser idénticamente nulo, y esto implica que $h(X, Y)$ es igual a un producto $k(X, Y)Y$, con $k(X, Y)$ un polinomio de grado a lo sumo 1 de $\mathbb{k}[X, Y]$, y esto es imposible porque entre nuestros cinco puntos no hay tres alineados. De esta forma concluimos que el espacio ${}^\circ S$ tiene dimensión 1 y, por lo tanto, que hay una *única* cónica que pasa por los cinco puntos.

Si $f \in V$, entonces el conjunto $Z_f := \{n \in \mathbb{N} : f(n) = 0 \text{ para todo } m \geq n\}$ no es vacío, así que podemos considerar el entero $n_f := \min Z_f$. Sea

$$\tau : f \in V \mapsto \sum_{n=1}^{n_f} f(n) \in \mathbb{k}.$$

Observemos que, de hecho, para cada $m \geq n_f$ se tiene que

$$\tau(f) = \sum_{n=1}^m f(n) \in \mathbb{k}.$$

Esta función τ es lineal: si f y g son elementos de V y α y β escalares de \mathbb{k} , entonces claramente $n_{\alpha f + \beta g} \leq \max\{n_f, n_g\}$, así que

$$\begin{aligned} \tau(\alpha f + \beta g) &= \sum_{n=1}^{n_{\alpha f + \beta g}} (\alpha f(n) + \beta g(n)) = \sum_{n=1}^{\max\{n_f, n_g\}} (\alpha f(n) + \beta g(n)) \\ &= \alpha \sum_{n=1}^{\max\{n_f, n_g\}} f(n) + \beta \sum_{n=1}^{\max\{n_f, n_g\}} g(n) = \alpha \tau(f) + \beta \tau(g). \end{aligned}$$

Esto nos dice que $\tau \in V^*$. Sea $T = \langle \tau \rangle$.

Sea $\psi \in S \cap T$. Existe entonces, por un lado, $\lambda \in \mathbb{k}$ tal que $\psi = \lambda \tau$ y, por otro, $k \in \mathbb{N}_0$, $n_1, \dots, n_k \in \mathbb{N}$ y $\alpha_1, \dots, \alpha_k \in \mathbb{k}$ tales que $\psi = \alpha_1 \phi_{n_1} + \dots + \alpha_k \phi_{n_k}$. Si ponemos

$$m = 2 + \max\{1, n_1, \dots, n_k\}$$

y $g \in V$ es la función tal que $g(m) = 1$ y $g(r) = 0$ si $r \in \mathbb{N} \setminus \{m\}$, entonces tenemos que $n_g = m$, que

$$\psi(g) = (\lambda \tau)(g) = \lambda \sum_{n=1}^m g(n) = \lambda$$

y que

$$\begin{aligned} \psi(g) &= (\alpha_1 \phi_{n_1} + \dots + \alpha_k \phi_{n_k})(g) \\ &= \alpha_1(g(n_1 + 1) - g(n_1)) + \dots + \alpha_k(g(n_k + 1) - g(n_k)) \\ &= 0, \end{aligned}$$

ya que g se anula en todo entero menor que m . Vemos así que $\lambda = 0$ y, por lo tanto, que $\psi = 0$. Esto nos dice que $S \cap T = 0$.

Por otro lado, si $h : \mathbb{N} \rightarrow \mathbb{Z}$ es la función tal que $h(1) = 1$ y $h(n) = 0$ para todo entero $n \geq 2$, entonces $h \notin {}^\circ T$, ya que $\tau(h) = 1 \neq 0$, y entonces también $h \notin {}^\circ S + {}^\circ T$, porque ${}^\circ S = 0$. Como ${}^\circ(S \cap T) = V$, vemos de esta forma que

$${}^\circ S + {}^\circ T \subsetneq {}^\circ(S + T).$$

Esto muestra que la inclusión (2) de la Proposición 3.2.3 es en general estricta. \diamond

Anulador a derecha

3.2.9. Sea V es un espacio vectorial y sea V^* su espacio dual. Si S es un subconjunto de V , llamamos **anulador a derecha** de S al conjunto

$$S^\circ = \{\phi \in V^* : \phi(x) = 0 \text{ para todo } x \in S\}.$$

3.2.10. Ejemplo. Sea $n \in \mathbb{N}$, consideremos el espacio vectorial $V = \mathbb{k}^n$ y sea S un subconjunto de V . Vimos en el Ejemplo 3.2.2 que los elementos de V^* se corresponden con ecuaciones lineales homogéneas en V . Si $\phi : V \rightarrow \mathbb{k}$ es un elemento de V^* , entonces existen escalares $a_1, \dots, a_n \in \mathbb{k}$ tales que

$$\phi(x) = a_1x_1 + \dots + a_nx_n, \quad \text{para cada } x = (x_1, \dots, x_n)^t \in \mathbb{k}^n,$$

esos escalares están bien determinados por ϕ , y la función ϕ pertenece a S° si y solamente si para cada $x = (x_1, \dots, x_n)^t \in S$ se tiene que

$$\phi(x) = a_1x_1 + \dots + a_nx_n = 0,$$

de manera que todos los elementos de S son soluciones de la ecuación lineal

$$a_1X_1 + \dots + a_nX_n = 0.$$

Así, podemos pensar al anulador a derecha S° de S como el conjunto de las ecuaciones lineales que son satisfechas simultáneamente por todos los elementos de S . Esta situación es la recíproca a la del Ejemplo 3.2.2: en lugar de tener algún número de ecuaciones lineales y querer encontrar los vectores de V que son sus soluciones comunes, empezamos con subespacio de V y buscamos el sistema más grande de ecuaciones lineales que lo tiene como espacio de soluciones. \diamond

3.2.11. Proposición. Sea V un espacio vectorial y sea V^* su espacio dual.

- (i) Se tiene que $0^\circ = V$ y $V^\circ = 0$.
- (ii) Si S y T son subconjuntos de V y $S \subseteq T$, entonces $S^\circ \supseteq T^\circ$.
- (iii) Si S es un subconjunto de V , entonces S° es un subespacio de V^* y coincide con $\langle S \rangle^\circ$.
- (iv) Si S y T son subespacios de V , entonces

$$(S + T)^\circ = S^\circ \cap T^\circ, \quad S^\circ + T^\circ \subseteq (S \cap T)^\circ. \quad (4)$$

De hecho, la inclusión de (4) es siempre una igualdad. Más abajo probaremos la Proposición 3.2.13 que dice que eso es así cuando V tiene dimensión finita. El caso general depende del Teorema 1.6.7.

Demostración. (i) Todo elemento de V^* se anula en 0, así que $0^\circ = V^*$. Por otro lado, si un elemento de V^* se anula en todo V , entonces se trata necesariamente de la función nula: esto significa que $V^\circ = 0$.

(ii) Sean S y T subconjuntos de V tales que $S \subseteq T$ y sea $\phi \in T^\circ$. Si $x \in S$, entonces $x \in T$ y, como $\phi \in T^\circ$, es $\phi(x) = 0$. Esto muestra que $\phi \in S^\circ$.

(iii) Sea S un subconjunto de V . Si ϕ y ψ son elementos de S° y α y β son escalares de \mathbb{k} , entonces para cada $x \in S$ tenemos que

$$(\alpha\phi + \beta\psi)(x) = \alpha\phi(x) + \beta\psi(x) = 0,$$

de manera que $\alpha\phi + \beta\psi \in S^\circ$. Como además claramente el elemento cero de V^* pertenece a S° , esto nos dice que S° es un subespacio de V^* .

Es $S \subseteq \langle S \rangle$, así que la parte (ii) implica que $\langle S \rangle^\circ \subseteq S^\circ$. Veamos la inclusión recíproca. Sea $\phi \in S^\circ$ y sea $x \in \langle S \rangle$, de manera que existen $n \in \mathbb{N}_0$, $x_1, \dots, x_n \in S$ y $a_1, \dots, a_n \in \mathbb{k}$ tales que $x = a_1x_1 + \dots + a_nx_n$: se tiene entonces que

$$\phi(x) = \phi(a_1x_1 + \dots + a_nx_n) = a_1\phi(x_1) + \dots + a_n\phi(x_n) = 0,$$

ya que $\phi(x_i) = 0$ para todo $i \in \llbracket n \rrbracket$. Esto muestra que $x \in \langle S \rangle^\circ$ y, en definitiva, que $S^\circ \subseteq \langle S \rangle^\circ$, como queríamos.

(iv) Sean S y T subespacios de V . Como $S \subseteq S + T$ y $T \subseteq S + T$, la parte (ii) implica que $(S + T)^\circ \subseteq S^\circ$ y $(S + T)^\circ \subseteq T^\circ$. Vemos así que, de hecho, $(S + T)^\circ \subseteq S^\circ \cap T^\circ$. Por otro lado, sea $\phi \in S^\circ \cap T^\circ$ y sea $x \in S + T$, de manera que existen $s \in S$ y $t \in T$ con $x = s + t$. Entonces $\phi(x) = \phi(s) + \phi(t) = 0$ y vemos que $\phi \in (S + T)^\circ$. Esto prueba la igualdad de (4)

Si ahora $\phi \in S^\circ + T^\circ$, de manera que existen $\sigma \in S^\circ$ y $\tau \in T^\circ$ con $\phi = \sigma + \tau$, y $x \in S \cap T$, entonces $\phi(x) = \sigma(x) + \tau(x) = 0$: esto nos dice que $\phi \in (S \cap T)^\circ$ y, en definitiva, que $S^\circ + T^\circ \subseteq (S \cap T)^\circ$. \square

3.2.12. De forma similar a lo que ocurre con el anulador a izquierda, cuando tenemos un espacio de dimensión finita la dimensión del anulador a derecha de un subespacios tiene una relación simple con la de este:

Proposición. *Sea V un espacio vectorial de dimensión finita. Si S es un subespacio de V , entonces tanto S como su anulador a derecha S° tienen dimensión finita y*

$$\dim S + \dim S^\circ = \dim V. \tag{5}$$

Demostración. Sea S un subespacio de V . Como S y S° son subespacios de V y de V^* , respectivamente, que tienen dimensión finita, sabemos que S y S° tienen ellos también dimensión finita. Sean $n = \dim V$ y $s = \dim S$. Sea (x_1, \dots, x_s) una base ordenada de S y sean $x_{s+1}, \dots, x_n \in V$ tales que $\mathcal{B} = (x_1, \dots, x_n)$ es una base ordenada de V . Sea, finalmente, $\mathcal{B}^* = (\phi_1, \dots, \phi_n)$ la base ordenada de V^* dual a \mathcal{B} . Para ver que vale la igualdad (5) del enunciado, es suficiente con que mostremos que $\mathcal{B}' = (\phi_{s+1}, \dots, \phi_n)$ es una base ordenada de S° . Como es un subconjunto de \mathcal{B}^* , es linealmente independiente, y sólo queda mostrar que genera a S° .

Sea $\phi \in S^\circ$. Como el conjunto \mathcal{B}^* es una base de V^* , existen escalares $a_1, \dots, a_n \in \mathbb{k}$ tales que $\phi = a_1\phi_1 + \dots + a_n\phi_n$. Si $i \in \llbracket s \rrbracket$, el vector x_i está en S y entonces

$$0 = \phi(x_i) = (a_1\phi_1 + \dots + a_n\phi_n)(x_i) = a_1\phi_1(x_i) + \dots + a_n\phi_n(x_i) = a_i.$$

Esto nos dice que, de hecho, $\phi = a_{s+1}\phi_{s+1} + \cdots + a_n\phi_n \in \langle \mathcal{B}' \rangle$ y completa la prueba. \square

3.2.13. Usando la Proposición 3.2.12 podemos mostrar que la inclusión (4) de la Proposición 3.2.11 es una igualdad cuando el espacio ambiente tiene dimensión finita. Como dijimos ahí, esto es así en general, sin necesidad de esa hipótesis, pero para mostrarlo necesitamos el Teorema 1.6.7.

Proposición. *Sea V un espacio vectorial de dimensión finita. Si S y T son subespacios de V , entonces $S^\circ + T^\circ = (S \cap T)^\circ$.*

Demostración. Sean S y T subespacios de V . Tenemos que $S^\circ \cap T^\circ = (S + T)^\circ$, así que

$$\begin{aligned}\dim(S^\circ \cap T^\circ) &= \dim(S + T)^\circ \\ &= \dim V - \dim(S + T) \\ &= \dim V - \dim S - \dim T + \dim(S \cap T),\end{aligned}$$

y usando esto vemos que

$$\begin{aligned}\dim(S^\circ + T^\circ) &= \dim S^\circ + \dim T^\circ - \dim(S^\circ \cap T^\circ) \\ &= (\dim V - \dim S) + (\dim V - \dim T) - (\dim V - \dim S - \dim T + \dim(S \cap T)) \\ &= \dim V - \dim(S \cap T) \\ &= \dim(S \cap T)^\circ.\end{aligned}$$

Por otro lado, de acuerdo a la Proposición 3.2.11 tenemos que $S^\circ + T^\circ \subseteq (S \cap T)^\circ$: estas dos cosas nos permiten concluir que la igualdad del enunciado vale. \square

Dualidad

3.2.14. Queremos ahora estudiar la interacción entre las operaciones ${}^\circ(-)$ y $(-)^{\circ}$. Bajo hipótesis adecuadas resultan ser mutuamente inversas: ese fenómeno es conocido como una dualidad.

Empezamos con una observación sencilla aunque no evidente:

Proposición. *Sea V un espacio vectorial.*

- (i) *Si S es un subconjunto de V , entonces $S \subseteq {}^\circ(S^\circ)$ y $S^\circ = ({}^\circ(S^\circ))^\circ$.*
- (ii) *Si T es un subconjunto de V^* , entonces $T \subseteq ({}^\circ T)^\circ$ y ${}^\circ T = ({}^\circ({}^\circ T))^\circ$.*

Demostración. Sea S un subconjunto de V . Si $x \in S$ y $\phi \in S^\circ$, entonces $\phi(x) = 0$: vemos así que $x \in {}^\circ(S^\circ)$ y, en definitiva, que $S \subseteq {}^\circ(S^\circ)$. Esto prueba la primera afirmación de (i). La primera afirmación de (ii) se prueba de exactamente la misma forma.

Sea otra vez $S \subseteq V$ un subconjunto de V . Lo que ya probamos de (i) nos dice que $S \subseteq {}^\circ(S^\circ)$ y esto implica, de acuerdo a la Proposición 3.2.11(ii), que $({}^\circ(S^\circ))^\circ \subseteq S^\circ$. Por otro lado, la primera parte de (i), cuando $T = S^\circ$, afirma que $S^\circ \subseteq ({}^\circ(S^\circ))^\circ$. Estas dos inclusiones prueban que

$S^\circ = (\circ(S^\circ))^\circ$, como afirma la segunda parte de la afirmación (i). La prueba de la segunda parte de (ii) puede hacerse de manera similar. \square

3.2.15. Si restringimos ahora nuestra atención a espacios vectoriales de dimensión finita, obtenemos fácilmente el resultado de dualidad:

Proposición. *Sea V un espacio vectorial.*

- (i) *Si S es un subespacio de V , entonces $S = \circ(S^\circ)$.*
- (ii) *Si V tiene dimensión finita y T es un subespacio de V^* , entonces $T = (\circ T)^\circ$.*

La segunda de estas afirmaciones es en general falsa si V no tiene dimensión finita. Damos en [3.2.18](#) un ejemplo de esto.

Demostración. (i) Sea S un subespacio de V . De la Proposición [3.2.14\(i\)](#) sabemos que $S \subseteq \circ(S^\circ)$. Si x es un vector de V que no está en S , entonces la Proposición [3.1.2](#) nos dice que existe $\phi \in V^*$ tal que $\phi(x) \neq 0$ y $\phi \in S^\circ$, y, por lo tanto, que $x \notin \circ(S^\circ)$. Esto muestra que $\circ(S^\circ) \subseteq S$.

(ii) Supongamos ahora que V tiene finita y que T es un subespacio de V^* . La Proposición [3.2.14\(ii\)](#) nos dice que $T \subseteq (\circ T)^\circ$ y como

$$\dim(\circ T)^\circ = \dim V - \dim(\circ T) = \dim V - (\dim V - \dim T) = \dim T,$$

vale, de hecho, la igualdad. \square

3.2.16. Corolario. *Sea V un espacio vectorial.*

- (i) *Si S y T son subespacios de V , entonces*

$$S = T \iff S^\circ = T^\circ.$$

- (ii) *Si V tiene dimensión finita y S y T son subespacios de V^* , entonces*

$$S = T \iff \circ S = \circ T.$$

La conclusión de la segunda parte de este corolario es falsa en general sin la hipótesis de dimensión finita de V . En [3.2.18](#) damos un ejemplo de esto.

Demostración. Si S y T son subespacios de V tales que $S^\circ = T^\circ$, entonces la primera parte de la Proposición [3.2.15](#) nos dice que $S = \circ(S^\circ) = \circ(T^\circ) = T$. Esto prueba una dirección de la equivalencia de la primera parte del corolario, y la otra es evidente. La segunda parte puede probarse de exactamente la misma forma, usando ahora la segunda parte de la Proposición [3.2.15](#). \square

3.2.17. Corolario. *Sea V un espacio vectorial.*

- (i) *Si S es un subconjunto de V , entonces $\langle S \rangle = \circ(S^\circ)$.*
- (ii) *Si V tiene dimensión finita y T es un subconjunto de V^* , entonces $\langle T \rangle = (\circ T)^\circ$.*

Como en los resultados anteriores, la conclusión de la segunda parte de este corolario no es cierta en general sin la hipótesis de dimensión finita, como mostramos en el Ejemplo 3.2.18.

Demostración. Si S es un subconjunto de S , entonces sabemos de la Proposición 3.2.14(i) que $S \subseteq {}^\circ(S^\circ)$, así que como ${}^\circ(S^\circ)$ es un subespacio de V , tenemos que $\langle S \rangle \subseteq {}^\circ(S^\circ)$. De esto, además, tenemos que $\langle S \rangle^\circ = ({}^\circ(S^\circ))^\circ = S^\circ$ y entonces, por el Corolario 3.2.16(ii), es $\langle S \rangle = S$. La segunda afirmación del corolario puede probarse razonando de la misma forma. \square

3.2.18. Ejemplo. La hipótesis de dimensión finita en las segundas partes de la Proposición 3.2.15 y el Corolario 3.2.16(ii) es necesaria en general. En el Ejemplo 3.2.8 construimos un espacio vectorial V , el de las funciones $\mathbb{N} \rightarrow \mathbb{k}$ de soporte finito, y un subespacio propio S de V^* tal que ${}^\circ S = 0$. Tenemos entonces, en primer lugar, que $S \not\subseteq V = ({}^\circ S)^\circ$, en segundo lugar, que ${}^\circ S = {}^\circ V^*$ y $S \neq V^*$ y, finalmente, que $\langle S \rangle = S \not\subseteq V = ({}^\circ S)^\circ$. \diamond

§3. Funciones transpuestas

3.3.1. Sean V y W espacios vectoriales y sea $f : V \rightarrow W$ una función lineal. Llamamos *función transpuesta* de f a la función

$$f^t : \phi \in W^* \mapsto \phi \circ f \in V^*.$$

Observemos que esto tiene sentido: si $\phi : W \rightarrow \mathbb{k}$ es una función lineal, entonces la composición $\phi \circ f : V \rightarrow \mathbb{k}$ es lineal, así que se trata de un elemento de V^* .

3.3.2. Proposición. Sean V y W espacios vectoriales.

- (i) Para cada función lineal $f : V \rightarrow W$ la función transpuesta $f^t : V^* \rightarrow W^*$ es también lineal.
- (ii) La función

$$\tau : f \in \text{hom}(V, W) \mapsto f^t \in \text{hom}(W^*, V^*)$$

es lineal.

Demostración. (i) Sea $f : V \rightarrow W$ una función lineal. Si $\phi, \psi \in W^*$ y $\lambda, \mu \in \mathbb{k}$, entonces

$$f^t(\lambda\phi + \mu\psi) = (\lambda\phi + \mu\psi) \circ f$$

y, como la composición es una función bilineal —como afirma la Proposición 2.5.6— esto es

$$= \lambda\phi \circ f + \mu\psi \circ f = \lambda f^t(\phi) + \mu f^t(\psi).$$

Vemos así que la función f^t es lineal.

(ii) Sean ahora $f, g \in \text{hom}(V, W)$ y $\lambda, \mu \in \mathbb{k}$. Como para cada $\phi \in W^*$ es

$$\begin{aligned}\tau(\lambda f + \mu g)(\phi) &= (\lambda f + \mu g)^t(\phi) \\ &= \phi \circ (\lambda f + \mu g) \\ &= \lambda \phi \circ f + \mu \circ g \\ &= \lambda f^t(\phi) + \mu g^t(\phi) \\ &= \lambda \tau(f)(\phi) + \mu \tau(g)(\phi) \\ &= (\lambda \tau(f) + \mu \tau(g))(\phi),\end{aligned}$$

vemos que $\tau(\lambda f + \mu g) = \lambda \tau(f) + \mu \tau(g)$, es decir, que la función τ del enunciado es lineal. \square

3.3.3. Proposición. Sean V y W dos espacios vectoriales.

- (i) Una función lineal $f : V \rightarrow W$ cuya función transpuesta $f^t : W^* \rightarrow V^*$ es nula es ella misma nula.
- (ii) La función lineal $\tau : f \in \text{hom}(V, W) \mapsto f^t \in \text{hom}(W^*, V^*)$ de la Proposición 3.3.2(ii) es inyectiva y, si V y W tienen dimensión finita, un isomorfismo.

Demostración. (i) Sea $f : V \rightarrow W$ una función lineal no nula, de manera que existe $v \in V$ tal que $f(v) \neq 0$. Existe entonces, de acuerdo a la Proposición 3.1.2, un funcional $\phi \in W^*$ tal que $\phi(f(v)) \neq 0$ y, por lo tanto, $f^t(\phi) \neq 0$: esto nos dice que $f^t \neq 0$.

(ii) Que la función lineal τ es inyectiva es consecuencia inmediata de la parte (i). Si V y W tienen dimensión finita, entonces esta función τ es un isomorfismo, ya que es inyectiva y su dominio y codominio tienen la misma dimensión finita:

$$\dim \text{hom}(V, W) = \dim V \cdot \dim W = \dim V^* \cdot \dim W^* = \dim(W^*, V^*).$$

\square

3.3.4. Proposición.

- (i) Si V es un espacio vectorial e $\text{id}_V : V \rightarrow V$ es la función identidad de V , entonces la función transpuesta $(\text{id}_V)^t : V^* \rightarrow V^*$ es la función identidad de V^*
- (ii) Si U, V y W son espacios vectoriales y $f : U \rightarrow V$ y $g : V \rightarrow W$ son funciones lineales, entonces $(g \circ f)^t = f^t \circ g^t$.

Demostración. (i) Sea V un espacio vectorial. Si $\phi \in V^*$, entonces

$$(\text{id}_V)^t(\phi) = \phi \circ \text{id}_V = \phi_V = \text{id}_{V^*}(\phi).$$

Esto significa que $(\text{id}_V)^t = \text{id}_{V^*}$.

(ii) Sean U, V y W espacios vectoriales y sean $f : U \rightarrow V$ y $g : V \rightarrow W$ funciones lineales. Para ver que $(g \circ f)^t = f^t \circ g^t$ basta observar que si $\phi \in W^*$, entonces

$$(g \circ f)^t(\phi) = \phi \circ (g \circ f) = (\phi \circ g) \circ f = g^t(\phi) \circ f = f^t(g^t(\phi)) = (g^t \circ f^t)(\phi),$$

como afirma la proposición. \square

3.3.5. Corolario. Sean V y W dos espacios vectoriales. Si una función lineal $f : V \rightarrow W$ es un isomorfismo, entonces la función transpuesta $f^t : W^* \rightarrow V^*$ es también un isomorfismo y su función inversa es $(f^{-1})^t : V^* \rightarrow W^*$.

Demostración. Sea $f : V \rightarrow W$ un isomorfismo. Para ver que $f^t : W^* \rightarrow V^*$ es un isomorfismo con inversa $(f^{-1})^t$ basta observar que, en vista de las partes (ii) y (i) de la Proposición 3.3.4, se tiene que

$$f^t \circ (f^{-1})^t = (f^{-1} \circ f)^t = (\text{id}_V)^t = \text{id}_{V^*}$$

y, de manera similar, que $(f^{-1})^t \circ f^t = \text{id}_{W^*}$. \square

3.3.6. La matriz de la función transpuesta de una función lineal tiene una expresión muy sencilla en términos de la matriz de esta última:

Proposición. Sean V y W espacios vectoriales de dimensión finita y sea $f : V \rightarrow W$ una función lineal. Si \mathcal{B} y \mathcal{B}' son bases ordenadas de V y de W , respectivamente, y \mathcal{B}^* y \mathcal{B}'^* son las bases de V^* y de W^* duales a \mathcal{B} y a \mathcal{B}' , entonces

$$[f^t]_{\mathcal{B}^*}^{\mathcal{B}'^*} = ([f]_{\mathcal{B}'}^{\mathcal{B}})^t.$$

Así, la matriz de la función transpuesta de f es la matriz transpuesta de la matriz de f .

Demostración. Sean $m = \dim V$ y $n = \dim W$, y sean $\mathcal{B} = (x_1, \dots, x_m)$, $\mathcal{B}' = (y_1, \dots, y_n)$, $\mathcal{B}^* = (\phi_1, \dots, \phi_m)$ y $\mathcal{B}'^* = (\psi_1, \dots, \psi_m)$. Sea $(a_{i,j}) = [f]_{\mathcal{B}'}^{\mathcal{B}}$ la matriz de f , de manera que para cada $i \in \llbracket m \rrbracket$ es

$$f(x_i) = a_{1,i}y_1 + \dots + a_{n,i}y_n.$$

Sea $j \in \llbracket n \rrbracket$. Como para cada $i \in \llbracket m \rrbracket$,

$$\begin{aligned} f^t(\psi_j)(x_i) &= (\psi_j \circ f)(x_i) \\ &= \psi_j(f(x_i)) \\ &= \psi_j(a_{1,i}y_1 + \dots + a_{n,i}y_n) \\ &= a_{1,i}\psi_j(y_1) + \dots + a_{n,i}\psi_j(y_n) \\ &= a_{j,i}, \end{aligned}$$

vemos que $f^t(\psi_j) = a_{j,1}\phi_1 + \dots + a_{j,m}\phi_m$. Esto significa que

$$[f^t]_{\mathcal{B}^*}^{\mathcal{B}'^*} = \begin{pmatrix} a_{1,1} & \cdots & a_{n,1} \\ \vdots & \ddots & \vdots \\ a_{1,m} & \cdots & a_{n,m} \end{pmatrix}$$

y esta es precisamente la matriz transpuesta de $[f]_{\mathcal{B}'}^{\mathcal{B}}$. \square

Monomorfismos, epimorfismos e isomorfismos

3.3.7. El núcleo y la imagen de una función lineal están en dualidad con la imagen y el núcleo de su función transpuesta:

Proposición. *Sean V y W espacios vectoriales. Si $f : V \rightarrow W$ es una función lineal entonces $\text{Nu}(f^t) = \text{Im}(f)^\circ$ e $\text{Im}(f^t) \subseteq \text{Nu}(f)^\circ$.*

Demostración. Sea $f : V \rightarrow W$ una función lineal. Si $\phi \in W^*$, es

$$\begin{aligned} \phi \in \text{Nu}(f^t) &\iff f^t(\phi) = 0 \\ &\iff \phi \circ f = 0 \\ &\iff \text{para todo } x \in V \text{ es } \phi(f(x)) = 0 \\ &\iff \text{para todo } y \in \text{Im}(f) \text{ es } \phi(y) = 0 \\ &\iff \phi \in \text{Im}(f)^\circ. \end{aligned}$$

Esto muestra que $\text{Nu}(f^t) = \text{Im}(f)^\circ$.

Por otro lado, si $\phi \in \text{Im}(f^t)$, de manera que existe $\psi \in W^*$ tal que $\phi = f^t(\psi) = \psi \circ f$, para cada $x \in \text{Nu}(f)$ se tiene que $\phi(x) = \psi(f(x)) = 0$: esto nos dice que $\phi \in \text{Nu}(f)^\circ$ y, por lo tanto, que $\text{Im}(f^t) \subseteq \text{Nu}(f)^\circ$.

Probemos la inclusión recíproca. Sea $\phi \in \text{Nu}(f)^\circ$. Si $y \in \text{Im}(f)$, existe $x \in V$ tal que $y = f(x)$ y $\phi(x)$ depende solamente de y y no de la elección hecha de x : en efecto, si x' es otro elemento de V tal que $y = f(x')$, entonces $0 = y - y = f(x) - f(x') = f(x - x')$, así que $x - x' \in \text{Nu}(f)$ y, por lo tanto, $\phi(x) - \phi(x') = \phi(x - x') = 0$ porque ϕ se anula sobre $\text{Nu}(f)$. Una consecuencia de esto es que hay una función $\bar{\phi} : \text{Im}(f) \rightarrow \mathbb{k}$ tal que para todo $x \in V$ se tiene que $\phi(x) = \bar{\phi}(f(x))$. En vista de la Proposición 2.1.9, hay una función lineal $\psi : W \rightarrow \mathbb{k}$ que extiende a $\bar{\phi}$ de su dominio $\text{Im}(f)$ a W , esto es, tal que $\psi(x) = \bar{\phi}(x)$ para todo $x \in \text{Im}(f)$. De esta construcción se sigue que $\phi = f^t(\psi) \in \text{Im}(f^t)$. Así, tenemos que $\text{Nu}(f)^\circ \subseteq \text{Im}(f^t)$. \square

3.3.8. Corolario. *Sean V y W dos espacios vectoriales. Una función lineal $f : V \rightarrow W$ es un monomorfismo, un epimorfismo, o un isomorfismo si y solamente si la correspondiente función transpuesta $f^t : W^* \rightarrow V^*$ es un epimorfismo, un monomorfismo, o un isomorfismo, respectivamente.*

Demostración. Como $\text{Nu}(f^t) = \text{Im}(f)^\circ$, tenemos que

$$\text{Nu}(f^t) = 0 \iff \text{Im}(f)^\circ = 0 = V^\circ \iff \text{Im}(f) = V, \quad (6)$$

así que f^t es un monomorfismo si y solamente si f es un epimorfismo. Por otro lado, como $\text{Im}(f^t) = \text{Nu}(f)^\circ$ vale que

$$\text{Im}(f^t) = V^* \iff \text{Nu}(f)^\circ = V^* = 0^\circ \iff \text{Nu}(f) = 0. \quad (7)$$

Esto nos dice que f^t es un epimorfismo si y solamente si f es un monomorfismo. Esto prueba dos de las afirmaciones del corolario, y la tercera se deduce inmediatamente de estas dos. Observemos que en las últimas equivalencias de (6) y de (7) están garantizadas por el Corolario 3.2.16(i). \square

3.3.9. Corolario. Consideremos un diagrama

$$U \xrightarrow{f} V \xrightarrow{g} W \quad (8)$$

de espacios vectoriales de dimensión finita y funciones lineales y su correspondiente diagrama dual

$$W^* \xrightarrow{g^t} V^* \xrightarrow{f^t} U^* \quad (9)$$

El diagrama (8) es un complejo si y solamente si el diagrama (9) lo es, y uno es exacto si y solamente si el otro lo es.

Demostración. Si el diagrama (8) es un complejo, de manera que $g \circ f = 0$, entonces

$$f^t \circ g^t = (g \circ f)^t = 0$$

y el diagrama (9) también es un complejo. Recíprocamente, si el diagrama (9) es un complejo, entonces $0 = f^t \circ g^t = (g \circ f)^t$. Como los tres espacios tienen dimensión finita, tenemos que $\text{Im}(g^t) = \text{Nu}(g)^\circ$ y $\text{Nu}(f^t) = \text{Im}(f)^\circ$, y entonces, gracias al Corolario 3.2.16(i), tenemos que

$$\text{Im}(g^t) = \text{Nu}(f^t) \iff \text{Nu}(g)^\circ = \text{Im}(f)^\circ \iff \text{Nu}(g) = \text{Im}(f).$$

\square

3.3.10. Proposición. Sea V un espacio vectorial y sea W un subespacio de V .

- (i) Si $\pi : V \rightarrow V/W$ es la proyección canónica al cociente, entonces la imagen de la función transpuesta $\pi^t : (V/W)^* \rightarrow V^*$ está contenida en W° y la correstricción $p : (V/W)^* \rightarrow W^\circ$ de π^t a W° es un isomorfismo.
- (ii) Si $\iota : W \rightarrow V$ es la inclusión, entonces hay un isomorfismo $q : V^*/W^\circ \rightarrow W^*$ tal que $q([\phi]) = \iota^t(\phi)$ para toda $\phi \in V^*$.

Demostración. **HACER.**

\square

3.3.11. Proposición. Sea V un espacio vectorial, sea $f : V \rightarrow V$ un endomorfismo de V y sea W un subespacio f -invariante de V . El subespacio W° de V^* es f^t -invariante y si $p : (V/W)^* \rightarrow W^\circ$ y $q : V^*/W^\circ \rightarrow W^*$ son los isomorfismos de la Proposición 3.3.10, entonces comutan los diagramas

$$\begin{array}{ccc} (V/W)^* & \xrightarrow{(f^W)^t} & (V/W)^* \\ p \downarrow & & \downarrow p \\ W^\circ & \xrightarrow{(f^t)_{W^\circ}} & W^\circ \end{array} \quad \begin{array}{ccc} V^*/W^\circ & \xrightarrow{(f^t)^{W^\circ}} & V^*/W^\circ \\ q \downarrow & & \downarrow q \\ W^* & \xrightarrow{(f_W)^t} & W^* \end{array}$$

Demostración. **HACER.**

□

3.3.12. En la Proposición 3.1.7 exhibimos, para cada espacio vectorial V de dimensión finita, un isomorfismo $\Phi : V \rightarrow V^{**}$, al que llamamos morfismo canónico. La razón de este nombre es el siguiente resultado:

Proposición. *Sea V un espacio vectorial de dimensión finita. El isomorfismo canónico $\Phi : V \rightarrow V^{**}$ es el único isomorfismo de V a V^{**} tal que para toda función lineal $f : V \rightarrow V$ conmuta el diagrama*

$$\begin{array}{ccc} V & \xrightarrow{\Phi} & V^{**} \\ f \downarrow & & \downarrow f^{\text{tt}} \\ V & \xrightarrow{\Phi} & V^{**} \end{array}$$

Aquí f^{tt} es la función transpuesta de $f^t : W^* \rightarrow V^*$.

Demostración. Si $f : V \rightarrow V$ es una función lineal, para cada $v \in V$ se tiene que

$$\begin{aligned} (f^{\text{tt}}(\Phi(x)))(\phi) &= (\Phi(x) \circ f^t)(\phi) = \Phi(x)(f^t(\phi)) = \Phi(x)(\phi \circ f) = (\phi \circ f)(x) \\ &= \phi(f(x)) = \Phi(f(x))(\phi) \end{aligned}$$

cada $\phi \in V^*$, de manera que $f^{\text{tt}}(\Phi(x)) = \Phi(f(x))$ y, por lo tanto, $f^{\text{tt}} \circ \Phi = \Phi \circ f$. Esto nos dice que el diagrama del enunciado conmuta. **HACER.** □

§4. El rango de una matriz

3.4.1. Si $m, n \in \mathbb{N}$ y $A \in M_{m,n}(\mathbb{k})$ es una matriz de m filas y n columnas con entradas en \mathbb{k} , llamamos **rango por columnas** de A y escribimos $\text{rg}(A)$ a la dimensión del subespacio $\text{col}(A)$ de \mathbb{k}^m generado por las n columnas de A .

3.4.2. El rango de una matriz posee una caracterización muy sencilla:

Proposición. *Sean $m, n \in \mathbb{N}$. El rango por columnas de una matriz $A \in M_{m,n}(\mathbb{k})$ es igual a la dimensión de la imagen $\text{Im}(f)$ de la función lineal $f : x \in \mathbb{k}^n \mapsto Ax \in \mathbb{k}^m$.*

Demostración. Sea (e_1, \dots, e_n) la base ordenada estándar de \mathbb{k}^n . La imagen de la función f está generada por el conjunto $\{f(e_1), \dots, f(e_n)\}$ y, como para cada $i \in [n]$ el vector $f(e_i) = Ae_i$ es precisamente la columna i -ésima de la matriz A , lo que afirma la proposición es inmediato. □

3.4.3. Proposición. *Sean $m, n \in \mathbb{N}$, sea $A \in M_{m,n}(\mathbb{k})$ y sea $f : x \in \mathbb{k}^n \mapsto Ax \in \mathbb{k}^m$.*

(i) Es $0 \leq \text{rg}(A) \leq \min\{m, n\}$ y el rango por columnas de A es nulo si y solamente si A es la

matriz nula.

- (ii) La función f es inyectiva si y solamente si $\text{rg}(A) = n$.
- (iii) La función f es sobreyectiva si y solamente si $\text{rg}(A) = m$.

Cuando el rango de A es igual a $\min\{m, n\}$ decimos que tiene **rango máximo**. En ese caso la función f es o inyectiva o sobreyectiva.

Demostración. Como la imagen de f es un subespacio del espacio \mathbb{k}^m , que tiene dimensión m , es $\dim \text{Im}(f) \leq m$. Por otro lado, del Teorema 2.4.1 es inmediato que $\dim \text{Im}(f) \leq \dim \mathbb{k}^n = n$. Estas dos desigualdades prueban la primera afirmación de (i), y la segunda es evidente.

Del Teorema 2.4.3 tenemos que $n = \dim \text{Nu}(f) + \text{rg}(A)$, y entonces es claro que $\text{Nu}(f) = 0$ si y solamente si $\text{rg}(A) = n$, como afirma la parte (ii). Por otro lado, como $\text{Im}(f)$ es un subespacio de \mathbb{k}^n , la función es sobreyectiva si y solamente si $\dim \text{Im}(f) = \dim \mathbb{k}^n = n$: esto prueba (iii). \square

3.4.4. Proposición. Sean $m, n \in \mathbb{N}$ y sea $A \in M_{m,n}(\mathbb{k})$.

- (i) Si $p \in \mathbb{N}$ y $B \in M_{n,p}(\mathbb{k})$, entonces

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

y si $\text{rg}(B) = n$ entonces, de hecho,

$$\text{rg}(AB) = \text{rg}(A).$$

- (ii) Si $k \in \mathbb{N}$ y $C \in M_{k,m}(\mathbb{k})$ es una matriz con $\text{rg}(C) = m$, entonces

$$\text{rg}(CA) = \text{rg}(A).$$

Demostración. **HACER.** \square

3.4.5. El siguiente resultado es conocido como la *desigualdad de Sylvester*, por James Joseph Sylvester (1814–1897, Inglaterra):

Proposición. Sean $m, n, p \in \mathbb{N}$. Si $A \in M_{m,n}(\mathbb{k})$ y $B \in M_{n,p}(\mathbb{k})$, entonces

$$\text{rg}(A) + \text{rg}(B) - n \leq \text{rg}(AB).$$

Demostración. Consideremos las funciones lineales $f : x \in \mathbb{k}^p \mapsto Bx \in \mathbb{k}^n$ y $g : x \in \mathbb{k}^n \mapsto Ax \in \mathbb{k}^m$. Del Teorema 2.4.1 y la Proposición 3.4.2 sabemos que

$$p = \dim \text{Nu}(f) + \text{rg}(B), \quad n = \dim \text{Nu}(g) + \text{rg}(A), \quad p = \dim \text{Nu}(g \circ f) + \text{rg}(AB)$$

y entonces

$$\text{rg}(AB) + n = \text{rg}(A) + \text{rg}(B) + [\dim \text{Nu}(f) + \dim \text{Nu}(g) - \dim \text{Nu}(g \circ f)].$$

Bastará entonces que mostremos que la expresión entre corchetes es no negativa.

Si $x \in \text{Nu}(g \circ f)$ entonces claramente $f(x) \in \text{Nu}(g)$, así que podemos considerar la función lineal

$$h : x \in \text{Nu}(g \circ f) \mapsto f(x) \in \text{Nu}(g).$$

Como $\text{Im}(h) \subseteq \text{Nu}(g)$, es $\dim \text{Im}(h) \leq \dim \text{Nu}(g)$. Por otro lado, es claro que $\text{Nu}(h) \subseteq \text{Nu}(f)$, así que $\dim \text{Nu}(h) \leq \dim \text{Nu}(f)$. Aplicando el Teorema 2.4.1 a la función h vemos entonces que

$$\dim \text{Nu}(g \circ f) = \dim \text{Nu}(h) + \dim \text{Im}(h) \leq \dim \text{Nu}(f) + \dim \text{Nu}(g),$$

y esto prueba lo que queremos. \square

3.4.6. Proposición. Sean $m, n \in \mathbb{N}$.

(i) Si A y B son elementos de $M_{m,n}(\mathbb{k})$, entonces

$$\text{rg}(A + B) \leq \text{rg}(A) + \text{rg}(B).$$

(ii) Si A es un elemento de $M_{m,n}(\mathbb{k})$, el rango de A es el menor número $r \in \mathbb{N}_0$ tal que existen matrices $B_1, \dots, B_r \in M_{m,n}(\mathbb{k})$ de rango 1 tales que $A = B_1 + \dots + B_r$.

Demostración. (i) Sean $A, B \in M_{m,n}(\mathbb{k})$ y consideremos las funciones $f : x \in \mathbb{k}^n \mapsto Ax \in \mathbb{k}^m$ y $g : x \in \mathbb{k}^n \mapsto Bx \in \mathbb{k}^m$. Si $x \in \mathbb{k}^n$, entonces $(f + g)(x) = f(x) + g(x) \in \text{Im}(f) + \text{Im}(g)$: esto muestra que $\text{Im}(f + g) \subseteq \text{Im}(f) + \text{Im}(g)$ y, en vista de la Proposición 1.8.4, que

$$\begin{aligned} \text{rg}(A + B) &= \dim \text{Im}(f + g) \\ &\leq \dim(\text{Im}(f) + \text{Im}(g)) \\ &= \dim \text{Im}(f) + \dim \text{Im}(g) - \dim(\text{Im}(f) \cap \text{Im}(g)) \\ &\leq \dim \text{Im}(f) + \dim \text{Im}(g) \\ &= \text{rg}(A) + \text{rg}(B). \end{aligned}$$

(ii) Sea $A \in M_{m,n}(\mathbb{k})$ y sea $r = \text{rg}(A)$. Si $s \in \mathbb{N}_0$ es el menor número tal que hay matrices $B_1, \dots, B_s \in M_{m,n}(\mathbb{k})$ de rango 1 con $A = B_1 + \dots + B_s$, la primera parte de la proposición implica que

$$r = \text{rg}(A) = \text{rg}(B_1 + \dots + B_s) \leq \text{rg}(B_1) + \dots + \text{rg}(B_s) = s.$$

Sea, por otro lado, $\{x_1, \dots, x_r\}$ una base del subespacio $\langle Ae_1, \dots, Ae_n \rangle$ de \mathbb{k}^m generado por las columnas de A . Para cada $i \in \llbracket n \rrbracket$ existen entonces escalares $b_{i,1}, \dots, b_{i,r} \in \mathbb{k}$ tales que

$$Ae_i = b_{i,1}x_1 + \dots + b_{i,r}x_r.$$

Si $j \in \llbracket r \rrbracket$, sea B_j la matriz de $M_{m,n}$ cuyas columnas son $b_{1,j}x_1, \dots, b_{n,j}x_j$. Esta matriz tiene rango a lo sumo 1 porque todas sus columnas son múltiplos del vector x_j . Es fácil verificar que $A = B_1 + \dots + B_r$, así que $s \leq r$. Esto, junto con la desigualdad que obtuvimos antes, implica el resultado que buscamos. \square

3.4.7. HACER: FIX THIS: De manera similar, el *rango por filas* de A , al que escribimos $\text{rg}_F(A)$, es la dimensión del subespacio de \mathbb{k}^m generado por las n filas de A .

Un resultado importante es que estos dos rangos son de hecho iguales:

Proposición. Sean $n, m \in \mathbb{N}$. Si $A \in M_{n,m}(\mathbb{k})$, entonces $\text{rg}(A) = \text{rg}_F(A)$.

En vista de esto, llamamos simplemente *rango* de la matriz A y escribimos $\text{rg}(A)$ al valor común de $\text{rg}_C(A)$ y $\text{rg}_F(A)$.

Demostración. Consideremos la función lineal $f : x \in \mathbb{k}^m \mapsto Ax \in \mathbb{k}^n$. Tenemos que

$$\begin{aligned}\dim \text{Im}(f) &= n - \dim \text{Im}(f)^\circ && \text{por la Proposición 3.2.12} \\ &= n - \dim \text{Nu}(f^t) && \text{por la Proposición 3.3.7} \\ &= n - (n - \dim \text{Im}(f^t)) && \text{otra vez por la Proposición 3.2.12} \\ &= \dim \text{Im}(f^t).\end{aligned}$$

Si (e_1, \dots, e_m) es la base estándar de \mathbb{k}^m , entonces la imagen $\text{Im}(f)$ de f está generada por el conjunto $\{f(e_1), \dots, f(e_m)\}$, esto es, por las columnas de A , y entonces $\text{rg}_C(A) = \dim \text{Im}(f)$. Por otro lado, sabemos de la Proposición 3.3.6 que la matriz de la función transpuesta f^t con respecto a las bases duales de las bases estándares de \mathbb{k}^n y \mathbb{k}^m es precisamente la matriz transpuesta A^t , así que un razonamiento similar nos dice que $\text{rg}_C(A^t) = \dim \text{Im}(f^t)$. Como claramente el rango por columnas de A^t coincide con el rango por filas de A , esto prueba la proposición. \square

3.4.8. Nuestro último resultado sobre el rango de matrices se refiere a matrices reales y complejas:

Proposición. Sean $m, n \in \mathbb{N}$.

(i) Si $A \in M_{m,n}(\mathbb{R})$, entonces

$$\text{rg}(AA^t) = \text{rg}(A^t A) = \text{rg}(A^t) = \text{rg}(A).$$

(ii) Si $A \in M_{m,n}(\mathbb{C})$, entonces

$$\text{rg}(AA^*) = \text{rg}(A^* A) = \text{rg}(\overline{A}) = \text{rg}(A^t) = \text{rg}(A^*) = \text{rg}(A).$$

En la segunda parte de esta proposición, \overline{A} denota a la *matriz conjugada* de A , esto es, a la matriz de $M_{m,n}(\mathbb{C})$ cuyas entradas se obtienen conjugando a las de A , y A^* a la *matriz adjunta* de A , que es la matriz transpuesta de \overline{A} .

Demostración. **HACER.** \square

Capítulo 4

Determinantes

§1. Funciones multilineales alternantes

4.1.1. Sea V un espacio vectorial. Si $d \in \mathbb{N}$, escribimos

$$V^d = \underbrace{V \times \cdots \times V}_{d \text{ factores}}$$

y decimos que una función $f : V^d \rightarrow \mathbb{k}$ es **d -multilineal** si para cada $i \in \llbracket d \rrbracket$ y cada elección de $d - 1$ vectores $x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_d$ de V la función

$$x \in V \mapsto f(x_1, \dots, x_{i-1}, x, x_{i+1}, \dots, x_d) \in \mathbb{k}$$

es lineal, esto es, si la función f es lineal en cada uno de sus d argumentos *separadamente*, y, en ese caso, decimos que el número d es el **grado** de f . Si d es 2 o 3, decimos que f es **bilineal** o **trilineal** en lugar de 2-multilineal y 3-multilineal.

4.1.2. Ejemplo. Sea $n \in \mathbb{N}$, sea $V = \mathbb{k}^n$ y consideremos la función $f : V \times V \rightarrow \mathbb{k}$ tal que

$$f(x, y) = x_1y_1 + \cdots + x_ny_n$$

cada vez que $x = (x_1, \dots, x_n)$ e $y = (y_1, \dots, y_n)$ son elementos de V . Es inmediato verificar que para cada $y \in V$ la función $x \in V \mapsto f(x, y) \in \mathbb{k}$ es lineal y que para cada $x \in V$ la función $y \in V \mapsto f(x, y) \in \mathbb{k}$ es lineal, así que f es 2-lineal.

Observemos que el conjunto $V \times V$ tiene una estructura natural de espacio vectorial: con respecto a esta estructura la función f *no* es lineal salvo que el cuerpo \mathbb{k} tenga exactamente dos elementos. En efecto, si (x, y) es un elemento de $V \times V$ y $\lambda \in \mathbb{k}$, entonces el producto $\lambda \cdot (x, y)$ es $(\lambda x, \lambda y)$ en esa estructura de espacio vectorial de $V \times V$, y tenemos que $f(\lambda x, \lambda y) = \lambda^2 f(x, y)$, que es en general distinto de $\lambda f(x, y)$, salvo que se tenga que $\lambda = \lambda^2$ para todo elemento de \mathbb{k} , y no es difícil ver que esto ocurre exactamente cuando el cuerpo \mathbb{k} tiene exactamente dos elementos. ◇

4.1.3. Una función $f : V^d \rightarrow \mathbb{k}$ es **alternante** si tiene la propiedad de que

para toda elección de vectores $x_1, \dots, x_d \in V$ tales que existen $i, j \in \llbracket d \rrbracket$ con $i \neq j$
 $y x_i = x_j$ se tiene que $f(x_1, \dots, x_d) = 0$.

Normalmente solo consideraremos funciones alternante que son multilineales. Cuando ese es el caso, tenemos la siguiente caracterización alternativa extremadamente útil:

Proposición. Sea V un espacio vectorial y sea $d \in \mathbb{N}$. Una función d -multilineal $f : V^d \rightarrow \mathbb{k}$ es alternante si y solamente si $f(x_1, \dots, x_d) = 0$ cada vez que x_1, \dots, x_d son elementos linealmente dependientes de V .

Demostración. La suficiencia de la condición es inmediata: si $f : V^d \rightarrow \mathbb{k}$ es d -multilineal satisface la condición, entonces cada vez que entre d -vectores x_1, \dots, x_d de V hay dos repetidos se tiene que $f(x_1, \dots, x_d) = 0$, ya que los d vectores son linealmente dependientes.

Probemos ahora la necesidad de la condición. Sea $f : V^d \rightarrow \mathbb{k}$ una función d -multilineal alternante y sean x_1, \dots, x_d elementos linealmente dependientes de V , de manera que existen escalares a_1, \dots, a_n en \mathbb{k} no todos nulos tales que $a_1x_1 + \dots + a_dx_d = 0$. Como los escalares no son todos nulos, existe $i \in \llbracket d \rrbracket$ tal que $a_i \neq 0$ y, por lo tanto,

$$x_i = (-a_i^{-1}a_1)x_1 + \dots + \widehat{(-a_i^{-1}a_i)x_i} + \dots + (-a_i^{-1}a_d)x_d.$$

Usando esto y la linealidad de la función f con respecto a su i -ésimo argumento, vemos que

$$\begin{aligned} f(x_1, \dots, x_d) &= f(x_1, \dots, x_{i-1}, \underbrace{(-a_i^{-1}a_1)x_1 + \dots + \widehat{(-a_i^{-1}a_i)x_i} + \dots + (-a_i^{-1}a_d)x_d}_{i}, x_{i+1}, \dots, x_d) \\ &= (-a_i^{-1}a_1)f(x_1, \dots, x_{i-1}, \underbrace{x_1}_{i}, x_{i+1}, \dots, x_d) + \dots \\ &\quad + \overbrace{(-a_i^{-1}a_i)f(x_1, \dots, x_{i-1}, \underbrace{x_i}_{i}, x_{i+1}, \dots, x_d)} \\ &\quad + \dots + (-a_i^{-1}a_d)f(x_1, \dots, x_{i-1}, \underbrace{x_d}_{i}, x_{i+1}, \dots, x_d) \\ &= 0, \end{aligned}$$

ya que en cada uno de los $d - 1$ sumandos del último miembro de esta cadena de igualdades la función f aparece evaluada en d vectores de V entre los que hay dos iguales. \square

4.1.4. Una consecuencia inmediata de esta proposición es la siguiente:

Corolario. Sea V un espacio vectorial, sea $d \in \mathbb{N}$ y sea $f : V^d \rightarrow \mathbb{k}$ una función d -multilineal

alternante. Si $x_1, \dots, x_d \in V$, $i \in \llbracket d \rrbracket$ e $y \in \langle x_1, \dots, \widehat{x_i}, \dots, x_d \rangle$, entonces

$$f(x_1, \dots, x_{i-1}, x_i + y, x_{i+1}, \dots, x_d) = f(x_1, \dots, x_d).$$

Así, el valor de una función multilineal alternada no cambia si a uno de sus argumentos le sumamos una combinación lineal de los otros.

Demostración. Sean $x_1, \dots, x_d \in V$, sea $i \in \llbracket d \rrbracket$ y sea $y \in \langle x_1, \dots, \widehat{x_i}, \dots, x_d \rangle$. Como la función f es lineal en su i -ésimo argumento, tenemos que

$$\begin{aligned} & f(x_1, \dots, x_{i-1}, x_i + y, x_{i+1}, \dots, x_d) \\ &= f(x_1, \dots, x_{i-1}, x_i, x_{i+1}, \dots, x_d) + f(x_1, \dots, x_{i-1}, y, x_{i+1}, \dots, x_d) \end{aligned}$$

y, como los vectores $x_1, \dots, x_{i-1}, y, x_{i+1}, \dots, x_d$ son linealmente dependientes, el segundo sumando se anula y esto es

$$= f(x_1, \dots, x_{i-1}, x_i, x_{i+1}, \dots, x_d),$$

como queremos. \square

4.1.5. El siguiente criterio simplifica la verificación de que una función multilineal satisface la condición de alternancia.

Proposición. Sea V un espacio vectorial y sea $d \in \mathbb{N}$. Una función d -multilineal $f : V^d \rightarrow \mathbb{k}$ es alternante si y solamente si

cada vez que x_1, \dots, x_n son elementos de V tales que existe $i \in \llbracket d-1 \rrbracket$ con $x_i = x_{i+1}$ se tiene que $f(x_1, \dots, x_d) = 0$.

Esto nos dice que es suficiente para ver que f es alternante con verificar que se anula cuando dos argumentos *contiguos* son iguales.

Demostración. Sea $f : V^d \rightarrow \mathbb{k}$ una función d -multilineal. Es claro que si f es alternante, entonces f satisface la condición del enunciado, así que es suficiente que mostremos que esta condición es también suficiente para que f sea alternante. Para cada $n \in \llbracket d-1 \rrbracket$ sea $P(n)$ la afirmación

cada vez que x_1, \dots, x_d son d elementos de V tales que existen $m \in \llbracket n \rrbracket$ e $i \in \llbracket d-m \rrbracket$ con $x_i = x_{i+m}$ se tiene que $f(x_1, \dots, x_d) = 0$.

En otras palabras, la afirmación $P(n)$ dice que $f(x_1, \dots, x_d)$ se anula siempre que dos de los argumentos que están separados por menos que n lugares son iguales. La condición del enunciado nos dice precisamente que la afirmación $P(1)$ vale, y probar que la función f es alternante es lo mismo que probar que la afirmación $P(d-1)$ vale. Procedemos por inducción.

Sea $k \in \llbracket d-2 \rrbracket$, supongamos que la afirmación $P(k)$ vale y mostremos que entonces la afirmación $P(k+1)$ también vale. Esto establecerá la validez de la proposición. Para verificar

la afirmación $P(k+1)$, sean x_1, \dots, x_d elementos de V , $m \in \llbracket k+1 \rrbracket$ e $i \in \llbracket d-k-1 \rrbracket$ tales que $x_i = x_{i+m}$. Es

$$0 = f(x_1, \dots, \underbrace{x_i, \dots, x_{i+m-2}}_i, \underbrace{x_{i+m-1} + x_{i+m}}_{i+m-1}, \underbrace{x_{i+m-1} + x_{i+m}}_{i+m}, x_{i+m+1}, \dots, x_d)$$

porque vale $P(1)$, y usando la multilinealidad de f vemos que esto es

$$\begin{aligned} &= f(x_1, \dots, \underbrace{x_i, \dots, x_{i+m-2}}_i, \underbrace{x_{i+m-1}}_{i+m-1}, \underbrace{x_{i+m-1}}_{i+m}, x_{i+m+1}, \dots, x_d) \\ &\quad + f(x_1, \dots, \underbrace{x_i, \dots, x_{i+m-2}}_i, \underbrace{x_{i+m-1}}_{i+m-1}, \underbrace{x_{i+m}}_{i+m}, x_{i+m+1}, \dots, x_d) \\ &\quad + f(x_1, \dots, \underbrace{x_i, \dots, x_{i+m-2}}_i, \underbrace{x_{i+m}}_{i+m-1}, \underbrace{x_{i+m-1} + x_{i+m}}_{i+m}, x_{i+m+1}, \dots, x_d). \end{aligned}$$

El primero de estos tres sumandos se anula porque en él f tiene sus argumentos $(i+m-1)$ -ésimo e $(i+m)$ -ésimos iguales y vale $P(1)$, y el tercero se anula porque en él f tiene sus argumentos i -ésimo e $(i+m-1)$ -ésimo iguales, que están separados por $m-1$ lugares, es $m-1 \leq k$ y vale la afirmación $P(k)$. Concluimos de esta forma que $f(x_1, \dots, x_d) = 0$, y esto prueba que vale $P(k+1)$, como queremos. \square

Funciones anti-simétricas

4.1.6. Sea V un espacio vectorial. Si $d \in \mathbb{N}$ y $f : V^d \rightarrow \mathbb{k}$ es una función d -multilineal, decimos que f es **anti-simétrica** si para cada $x_1, \dots, x_d \in V$ y cada elección de $i, j \in \llbracket d \rrbracket$ con $i < j$ se tiene que

$$f(x_1, \dots, \underbrace{x_i, \dots, x_j, \dots, x_d}_i, \dots, \underbrace{x_j, \dots, x_i, \dots, x_d}_j) = -f(x_1, \dots, x_d),$$

esto es, si al intercambiar dos de sus argumentos el valor de la función se multiplica por -1 .

4.1.7. Hay una relación estrecha entre las nociones de alternancia y de anti-simetría:

Proposición. *Sea V un espacio vectorial, sea $d \in \mathbb{N}$ y sea $f : V^d \rightarrow \mathbb{k}$ una función d -multilineal.*

- (i) *Si f es alternante, entonces f es anti-simétrica.*
- (ii) *Si f es anti-simétrica y en el cuerpo \mathbb{k} se tiene que $2 \neq 0$, entonces f es alternante.*

En un cuerpo general \mathbb{k} denotamos 2 al elemento $1+1$. Recordemos que $2 \neq 0$ en \mathbb{k} precisamente cuando \mathbb{k} tiene característica distinta de 2 .

Demostración. Sean $x_1, \dots, x_d \in V$ e $i, j \in \llbracket d \rrbracket$ tales que $i < j$. Como la función f es multilineal, se tiene que

$$f(x_1, \dots, x_{i-1}, \underbrace{x_i + x_j}_i, x_{i+1}, \dots, x_{j-1}, \underbrace{x_i + x_j}_j, x_{j+1}, \dots, x_d)$$

$$\begin{aligned}
&= f(x_1, \dots, x_{i-1}, \underbrace{x_i}_{i}, x_{i+1}, \dots, x_{j-1}, \underbrace{x_i}_{j}, x_{j+1}, \dots, x_d) \\
&\quad + f(x_1, \dots, x_{i-1}, \underbrace{x_i}_{i}, x_{i+1}, \dots, x_{j-1}, \underbrace{x_j}_{j}, x_{j+1}, \dots, x_d) \\
&\quad + f(x_1, \dots, x_{i-1}, \underbrace{x_j}_{i}, x_{i+1}, \dots, x_{j-1}, \underbrace{x_i}_{j}, x_{j+1}, \dots, x_d) \\
&\quad + f(x_1, \dots, x_{i-1}, \underbrace{x_j}_{i}, x_{i+1}, \dots, x_{j-1}, \underbrace{x_j}_{j}, x_{j+1}, \dots, x_d)
\end{aligned} \tag{1}$$

Si la función f es alternante, entonces el lado izquierdo de esta igualdad y el primero y el cuarto de los sumandos del lado derecho se anulan, y entonces

$$\begin{aligned}
&f(x_1, \dots, x_{i-1}, \underbrace{x_i}_{i}, x_{i+1}, \dots, x_{j-1}, \underbrace{x_i}_{j}, x_{j+1}, \dots, x_d) \\
&\quad + f(x_1, \dots, x_{i-1}, \underbrace{x_j}_{i}, x_{i+1}, \dots, x_{j-1}, \underbrace{x_i}_{j}, x_{j+1}, \dots, x_d) = 0.
\end{aligned}$$

Esto nos dice que en ese caso f es anti-simétrica y prueba la primera parte de la proposición.

Para ver la segunda, supongamos que f es anti-simétrica y que $2 \neq 0$ en el cuerpo \mathbb{k} . Si en la igualdad (1) de arriba es $x_i = x_j$, entonces el lado izquierdo es

$$\begin{aligned}
&f(x_1, \dots, x_{i-1}, \underbrace{2x_i}_{i}, x_{i+1}, \dots, x_{j-1}, \underbrace{2x_i}_{j}, x_{j+1}, \dots, x_d) \\
&= 2 \cdot 2 \cdot f(x_1, \dots, x_{i-1}, \underbrace{x_i}_{i}, x_{i+1}, \dots, x_{j-1}, \underbrace{x_i}_{j}, x_{j+1}, \dots, x_d)
\end{aligned}$$

mientras que el lado derecho —ya que la suma de su segundo y su tercer término es nula por la hipótesis hecha sobre f — es

$$2 \cdot f(x_1, \dots, x_{i-1}, \underbrace{x_i}_{i}, x_{i+1}, \dots, x_{j-1}, \underbrace{x_i}_{j}, x_{j+1}, \dots, x_d).$$

La igualdad 1, en consecuencia, nos dice que

$$\begin{aligned}
&2 \cdot 2 \cdot f(x_1, \dots, x_{i-1}, \underbrace{x_i}_{i}, x_{i+1}, \dots, x_{j-1}, \underbrace{x_i}_{j}, x_{j+1}, \dots, x_d) \\
&= 2 \cdot f(x_1, \dots, x_{i-1}, \underbrace{x_i}_{i}, x_{i+1}, \dots, x_{j-1}, \underbrace{x_i}_{j}, x_{j+1}, \dots, x_d).
\end{aligned}$$

Por supuesto, esto implica que

$$2 \cdot f(x_1, \dots, x_{i-1}, \underbrace{x_i}_{i}, x_{i+1}, \dots, x_{j-1}, \underbrace{x_i}_{j}, x_{j+1}, \dots, x_d) = 0$$

y, como $2 \neq 0$ en \mathbb{k} , esto a su vez nos dice que

$$f(x_1, \dots, x_{i-1}, \underbrace{x_i}_{i}, x_{i+1}, \dots, x_{j-1}, \underbrace{x_i}_{j}, x_{j+1}, \dots, x_d) = 0.$$

Vemos así que f es alternante. \square

4.1.8. Ejemplo. La condición de que 2 sea distinto de 0 en el cuerpo \mathbb{k} que aparece en la segunda parte de la Proposición 4.1.7 es necesaria. Para ver un ejemplo, supongamos que $\mathbb{k} = \mathbb{F}_2$, el cuerpo de dos elementos que describimos en el Ejemplo 1.1.3(c), sea $V = \mathbb{k}$ y consideremos la función $f : (x, y) \in V \times V \mapsto xy \in V$ dada por la multiplicación de \mathbb{k} . Que f es bilineal es inmediato y si $x, y \in V$, entonces $f(x, y) = f(y, x)$, así que $f(x, y) + f(y, x) = 2f(x, y) = 0$. Esta función es por lo tanto bilineal y anti-simétrica. Sin embargo, no es alternante: $f(1, 1) = 1 \neq 0$. \diamond

4.1.9. El siguiente resultado es un resultado análogo al de la Proposición 4.1.5 pero para funciones multilineales anti-simétricas en lugar de alternantes.

Proposición. Sea V un espacio vectorial y sea $d \in \mathbb{N}$. Una función f -multilineal $f : V^d \rightarrow \mathbb{k}$ es anti-simétrica si y solamente si para toda elección de d vectores x_1, \dots, x_d en V y de $i \in \llbracket d-1 \rrbracket$ se tiene que

$$f(x_1, \dots, x_{i-1}, \underbrace{x_i}_{i}, \underbrace{x_{i+1}}_{i+1}, x_{i+2}, \dots, x_d) = -f(x_1, \dots, x_{i-1}, \underbrace{x_{i+1}}_{i}, \underbrace{x_i}_{i+1}, x_{i+2}, \dots, x_d).$$

En otras palabras, para que la función f sea anti-simétrica es suficiente con que al intercambiar dos de sus argumentos *contiguos* su valor se multiplique por -1 . La demostración que daremos es totalmente similar a la que dimos de la Proposición 4.1.5.

Demostración. Sea $f : V^d \rightarrow \mathbb{k}$ una función d -multilineal. Es claro que si f es anti-simétrica satisface la condición del enunciado, así que esta es necesaria. Vamos que es además suficiente.

Supongamos que la función f satisface la condición del enunciado, y para cada $n \in \llbracket d-1 \rrbracket$ sea $P(n)$ la afirmación

cada vez que x_1, \dots, x_d son elementos de V , $m \in \llbracket n \rrbracket$ e $i \in \llbracket d-n \rrbracket$ se tiene que

$$f(x_1, \dots, \underbrace{x_i}_{i}, \dots, \underbrace{x_{i+m}}_{i+m}, \dots, x_d) = -f(x_1, \dots, \underbrace{x_{i+m}}_{i}, \dots, \underbrace{x_i}_{i+m}, \dots, x_d).$$

En otras palabras, la afirmación $P(n)$ dice que cuando a f le intercambiamos dos argumentos que están separados por menos que n lugares el valor se multiplica por -1 . Que f satisface la condición del enunciado significa precisamente que vale la afirmación $P(1)$, y que la función f sea anti-simétrica es equivalente a que valga la afirmación $P(d-1)$, así que podemos proceder por inducción: para probar la proposición será suficiente que mostremos que para cada $k \in \llbracket d-1 \rrbracket$ vale $P(k) \implies P(k+1)$.

Fijemos entonces $k \in \llbracket d-1 \rrbracket$, supongamos que la afirmación $P(k)$ vale, y sean x_1, \dots, x_d elementos de V , $m \in \llbracket k+1 \rrbracket$ e $i \in \llbracket d-k-1 \rrbracket$. Si $m < k+1$, entonces que vale $P(k)$ implica

inmediatamente que

$$f(x_1, \dots, \underbrace{x_i, \dots, x_{i+m}}_i, \dots, x_d) = -f(x_1, \dots, \underbrace{x_{i+m}, \dots, x_i}_{i+m}, \dots, x_d).$$

Supongamos entonces que $m = k + 1$. Intercambiando el argumento i -ésimo y el $(i+1)$ -ésimo, vemos que

$$f(x_1, \dots, \underbrace{x_i, x_{i+1}, \dots, x_{i+m}}_{i+1}, \dots, x_d) = -f(x_1, \dots, \underbrace{x_{i+1}, x_i, \dots, x_{i+m}}_{i+1}, \dots, x_d)$$

porque vale la condición del enunciado, y esto es

$$= f(x_1, \dots, \underbrace{x_{i+1}, x_{i+m}, \dots, x_i}_{i+1}, \dots, x_d)$$

porque estamos suponiendo que vale $P(k)$ e intercambiamos el argumento $(i+1)$ -ésimo con el $(i+m)$ -ésimo, del que está separado por $k-1$ lugares, y finalmente esto es

$$= -f(x_1, \dots, \underbrace{x_{i+m}, x_{i+1}, \dots, x_i}_{i+1}, \dots, x_d),$$

ya que intercambiamos ahora el i -ésimo argumento con el $(i+1)$ -ésimo. Esto completa la inducción y, por lo tanto, la prueba de la proposición. \square

§2. Funciones multilineales alternantes de grado máximo

4.2.1. Sea V un espacio vectorial. Para cada $d \in \mathbb{N}$, escribimos $\text{Alt}^d(V)$ al conjunto de todas las funciones $V^d \rightarrow \mathbb{k}$ que son d -multilineales y alternantes. Se trata de un subconjunto del espacio vectorial de todas las funciones $V^d \rightarrow \mathbb{k}$ y, de hecho, es fácil verificar que se trata de un subespacio. Veremos entonces a $\text{Alt}^d(V)$ como un espacio vectorial en todo lo que sigue.

4.2.2. Una consecuencia inmediata de la Proposición 4.1.3 es la siguiente:

Proposición. *Sea V un espacio vectorial. Si V tiene dimensión finita y $d > \dim V$, entonces toda función d -multilineal alternante $f : V^d \rightarrow \mathbb{k}$ es idénticamente nula y, en consecuencia, $\text{Alt}^d(V) = 0$.*

Demostración. En efecto, si V tiene dimensión finita, $d > \dim V$ y $f : V^d \rightarrow \mathbb{k}$ es d -multilineal alternante, entonces cada vez que elegimos d vectores x_1, \dots, x_d en V estos son linealmente dependientes y la Proposición 4.1.3 nos dice que $f(x_1, \dots, x_d) = 0$. \square

4.2.3. Como consecuencia de esto, el grado máximo en el que podemos esperar encontrar funciones multilineales alternantes no nulas sobre un espacio vectorial de dimensión finita n es precisamente n . La siguiente observación que hacemos es que si las hay no hay muchas:

Proposición. Sea V un espacio vectorial de dimensión finita n . Si $\mathcal{B} = (e_1, \dots, e_n)$ es una base ordenada de V , entonces la función

$$\Phi : f \in \text{Alt}^n(V) \mapsto f(e_1, \dots, e_n) \in \mathbb{k}$$

es lineal e inyectiva. En particular, el espacio vectorial $\text{Alt}^n(V)$ tiene dimensión finita y, de hecho,

$$\dim \text{Alt}^n(V) \leq 1.$$

Demostración. Fijemos una base ordenada $\mathcal{B} = (e_1, \dots, e_n)$ de V . Que la función Φ es lineal es claro. Probaremos que es inyectiva. Hecho esto, la última afirmación del enunciado seguirá inmediatamente de la Proposición 2.4.3(ii).

Sea entonces f un elemento de $\text{Alt}^n(V)$ tal que $\Phi(f) = 0$ y mostremos que f es necesariamente la función nula. Haremos esto en dos pasos.

PRIMER PASO. Empezamos por probar que

$$\text{si } i_1, \dots, i_n \in \llbracket n \rrbracket, \text{ entonces } f(e_{i_1}, \dots, e_{i_n}) = 0, \quad (2)$$

es decir, que f se anula cuando todos sus argumentos pertenecen a la base \mathcal{B} .

Para ello, sean i_1, \dots, i_n elementos de $\llbracket n \rrbracket$ y para cada $k \in \llbracket 0, n \rrbracket$ sea $P(k)$ la afirmación

$$\text{si } i_{k+1}, \dots, i_n \in \llbracket n \rrbracket, \text{ entonces } f(e_1, \dots, e_k, e_{i_{k+1}}, \dots, e_{i_n}) = 0.$$

Observemos que sabemos que $P(n)$ vale, ya que $f(e_1, \dots, e_n) = 0$ por hipótesis, y que $P(0)$ es la afirmación de que vale la afirmación (2) que queremos. Podemos entonces proceder por inducción descendente: basta que probemos que si $k \in \llbracket n \rrbracket$ y vale $P(k)$, entonces también vale $P(k - 1)$.

Sea entonces $k \in \llbracket n \rrbracket$, supongamos que vale $P(k)$ y que $i_k, \dots, i_n \in \llbracket n \rrbracket$. Tenemos ahora dos posibilidades:

- Puede ser que $k \notin \{i_k, \dots, i_n\}$ y entonces $f(e_1, \dots, e_{k-1}, e_{i_k}, \dots, e_{i_n}) = 0$, ya que los argumentos que tiene aquí f son n y pertenecen al subespacio generado por $\mathcal{B} \setminus \{e_k\}$, que tiene dimensión $n - 1$, así que son linealmente dependientes — recordemos que f es alternante.
- Puede ser, por el contrario, que k sí sea un elemento de $\{i_k, \dots, i_n\}$, de manera que existe $r \in \llbracket k, n \rrbracket$ tal que $i_r = k$ y, por lo tanto,

$$\begin{aligned} f(e_1, \dots, e_{k-1}, \underbrace{e_{i_k}}_k, \dots, \underbrace{e_{i_r}}_r, \dots, e_{i_n}) &= -f(e_1, \dots, e_{k-1}, \underbrace{e_{i_r}}_k, \dots, \underbrace{e_{i_k}}_r, \dots, e_{i_n}) \\ &= -f(e_1, \dots, e_{k-1}, \underbrace{e_k}_k, \dots, \underbrace{e_{i_k}}_r, \dots, e_{i_n}) \\ &= 0 \end{aligned}$$

ya f es anti-simétrica y estamos suponiendo que vale $P(k)$.

En cualquiera de los dos casos tenemos entonces que $f(e_1, \dots, e_{k-1}, e_{i_k}, \dots, e_{i_n}) = 0$: esto prueba que vale $P(k - 1)$ y completa la inducción. Hemos verificado así que la afirmación (2) es cierta.

SEGUNDO PASO. Queremos probar ahora que f es la función nula, esto es, que

$$\text{si } x_1, \dots, x_n \in V, \text{ entonces } f(x_1, \dots, x_n) = 0. \quad (3)$$

Para hacerlo, procederemos de manera similar a lo que hicimos recién: para cada $k \in \llbracket 0, n \rrbracket$ sea $Q(k)$ la afirmación

$$\text{si } x_1, \dots, x_k \in \mathcal{B} \text{ y } x_{k+1}, \dots, x_n \in V, \text{ entonces } f(x_1, \dots, x_n) = 0.$$

La afirmación (3) que queremos probar es claramente equivalente a $Q(0)$ y sabemos que $Q(n)$ vale, ya que eso es precisamente lo que nos dice (2). Es suficiente entonces que, procediendo inductivamente, mostremos que si $k \in \llbracket n \rrbracket$ y vale $Q(k)$ entonces también vale $Q(k-1)$. Esto completará la prueba de la proposición.

Fijemos entonces $k \in \llbracket n \rrbracket$, supongamos que vale la afirmación $Q(k)$ y sean $x_1, \dots, x_{k-1} \in \mathcal{B}$ y $x_k, \dots, x_n \in V$. Como \mathcal{B} es una base, hay escalares $a_1, \dots, a_n \in \mathbb{k}$ tales que $x_k = a_1 e_1 + \dots + a_n e_n$ y, como la función f es multilíneaal,

$$\begin{aligned} f(x_1, \dots, \underbrace{x_k, \dots, x_n}_{k}, \dots, x_n) &= f(x_1, \dots, \underbrace{a_1 e_1 + \dots + a_n e_n}_{k}, \dots, x_n) \\ &= a_1 f(x_1, \dots, \underbrace{e_1, \dots, x_n}_k) + \dots + a_n f(x_1, \dots, \underbrace{e_n, \dots, x_n}_k). \end{aligned}$$

Cada uno de los n sumandos del último miembro de esta igualdad tiene como factor al resultado de evaluar a f en una n -upla de elementos de V cuyos primeros k elementos están en la base \mathcal{B} y entonces, como estamos suponiendo que vale $Q(k)$, cada uno de ellos se anula. Así, vemos que $f(x_1, \dots, x_n) = 0$ y, en definitiva, podemos concluir que vale la afirmación $Q(k-1)$, completando la inducción del segundo y último paso de la prueba. \square

4.2.4. Lo que nos queda por hacer para describir el espacio de funciones multilíneaales alternantes de grado máximo es mostrar que existe alguna no nula. Empezamos haciéndolo en un caso especial, en el que $V = \mathbb{k}^n$ para algún entero positivo n :

Proposición. Para cada $n \in \mathbb{N}$ existe un único elemento $D_n \in \text{Alt}^n(\mathbb{k}^n)$ tal que

$$D_n(e_1, \dots, e_n) = 1$$

si (e_1, \dots, e_n) es la base ordenada estándar de \mathbb{k}^n y, en particular, es

$$\dim \text{Alt}^n(\mathbb{k}^n) = 1.$$

Demostración. Es suficiente que nos ocupemos de la afirmación de existencia, ya que la unicidad y la afirmación sobre la dimensión de $\text{Alt}^n(V)$ se deducen de la Proposición 4.2.3. Procederemos por inducción en n , observando que como la función $D_1 : x \in \mathbb{k}^1 \mapsto x \in \mathbb{k}$ es 1-multilíneaal, alternante y $D_1(1) = 1$, la proposición vale si $n = 1$.

Supongamos que $n \geq 2$ e, inductivamente, que existe una función $(n-1)$ -multilineal alternante $D_{n-1} : (\mathbb{k}^{n-1})^{n-1} \rightarrow \mathbb{k}$ sobre \mathbb{k}^{n-1} tal que $D_{n-1}(e_1, \dots, e_{n-1}) = 0$ si (e_1, \dots, e_{n-1}) es la base ordenada estándar de \mathbb{k}^{n-1} .

Si $u = (u_1, \dots, u_n)^t$ es un elemento de \mathbb{k}^n , convengamos en escribir u' al escalar u_1 y u'' al vector $(u_2, \dots, u_n)^t$ de \mathbb{k}^{n-1} . Es claro que si $y, z \in \mathbb{k}^n$ y $a, b \in \mathbb{k}$ se tiene que

$$(ay + bz)' = ay' + bz', \quad (ay + bz)'' = ay'' + bz''. \quad (4)$$

Definimos una función $D_n : (\mathbb{k}^n)^n \rightarrow \mathbb{k}$ de la siguiente manera: si $x_1, \dots, x_n \in \mathbb{k}^n$, ponemos

$$D_n(x_1, \dots, x_n) = \sum_{l=1}^n (-1)^{l+1} x'_l D_{n-1}(x''_1, \dots, \widehat{x''_l}, \dots, x''_n).$$

Mostremos que D_n es n -multilineal alternante y, si (e_1, \dots, e_n) es la base ordenada estándar de \mathbb{k}^n , que $D_n(e_1, \dots, e_n) = 1$: esto probará la proposición.

- Se tiene que $e'_1 = 1$ y que $e'_l = 0$ si $l \in [2, n]$, así que

$$D_n(e_1, \dots, e_n) = \sum_{l=1}^n (-1)^{l+1} e'_l D_{n-1}(e''_1, \dots, \widehat{e''_l}, \dots, e''_n) = D_{n-1}(e''_2, \dots, e''_n)$$

porque en la suma todos los términos son nulos salvo posiblemente el que corresponde a $l = 1$. Como (e''_2, \dots, e''_n) es precisamente la base ordenada estándar de \mathbb{k}^{n-1} , la hipótesis inductiva nos dice que $D_n(e_1, \dots, e_n) = 1$.

- Sea $i \in [n]$, sean $x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n, y, z \in \mathbb{k}^n$ y $a, b \in \mathbb{k}$. De la definición de la función D_n tenemos que

$$\begin{aligned} D_n(x_1, \dots, x_{i-1}, ay + bz, x_{i+1}, \dots, x_n) \\ = \sum_{l=1}^{i-1} (-1)^{l+1} x'_l D_{n-1}(x''_1, \dots, \widehat{x''_l}, \dots, x''_{i-1}, (ay + bz)'', x''_{i+1}, \dots, x''_n) \\ + (-1)^{i+1} (ay + bz)' D_{n-1}(x''_1, \dots, x''_{i-1}, x''_{i+1}, \dots, x''_n) \\ + \sum_{l=i+1}^n (-1)^{l+1} x'_l D_{n-1}(x''_1, \dots, x''_{i-1}, (ay + bz)'', x''_{i+1}, \dots, \widehat{x''_l}, \dots, x''_n). \end{aligned}$$

Usando primero las igualdades (4) y después la multilinealidad de la función D_{n-1} , podemos reescribir esto en la forma

$$\begin{aligned} & \sum_{l=1}^{i-1} (-1)^{l+1} x'_l \left(a D_{n-1}(x''_1, \dots, \widehat{x''_l}, \dots, x''_{i-1}, y'', x''_{i+1}, \dots, x''_n) \right. \\ & \quad \left. + b D_{n-1}(x''_1, \dots, \widehat{x''_l}, \dots, x''_{i-1}, z'', x''_{i+1}, \dots, x''_n) \right) \\ & + (-1)^{i+1} a y' D_{n-1}(x''_1, \dots, x''_{i-1}, x''_i, \dots, x''_n) \\ & + (-1)^{i+1} b z' D_{n-1}(x''_1, \dots, x''_{i-1}, x''_i, \dots, x''_n) \end{aligned}$$

$$+ \sum_{l=i+1}^n (-1)^{l+1} x'_l \left(a D_{n-1}(x''_1, \dots, x''_{i-1}, y'', x''_{i+1}, \dots, \widehat{x''_l}, \dots, x''_n) \right. \\ \left. + b D_{n-1}(x''_1, \dots, x''_{i-1}, z'', x''_{i+1}, \dots, \widehat{x''_l}, \dots, x''_n) \right).$$

Por otro lado,

$$a D_n(x_1, \dots, x_{i-1}, y, x_{i+1}, \dots, x_n) + b D_n(x_1, \dots, x_{i-1}, z, x_{i+1}, \dots, x_n) \\ = a \sum_{l=1}^{i-1} (-1)^{l+1} x'_l D_{n-1}(x''_1, \dots, \widehat{x''_l}, \dots, x''_{i-1}, y'', x''_{i+1}, \dots, x''_n) \\ + a (-1)^{i+1} y' D_{n-1}(x''_1, \dots, x''_{i-1}, x''_i, \dots, x''_n) \\ + a \sum_{l=i+1}^n (-1)^{l+1} x'_l D_{n-1}(x''_1, \dots, x''_{i-1}, y'', x''_{i+1}, \dots, \widehat{x''_l}, \dots, x''_n) \\ + b \sum_{l=1}^{i-1} (-1)^{l+1} x'_l D_{n-1}(x''_1, \dots, \widehat{x''_l}, \dots, x''_{i-1}, z'', x''_{i+1}, \dots, x''_n) \\ + b (-1)^{i+1} y' D_{n-1}(x''_1, \dots, x''_{i-1}, x''_i, \dots, x''_n) \\ + b \sum_{l=i+1}^n (-1)^{l+1} x'_l D_{n-1}(x''_1, \dots, x''_{i-1}, z'', x''_{i+1}, \dots, \widehat{x''_l}, \dots, x''_n)$$

y esto es igual a la expresión anterior. Vemos así que la función D_n es lineal con respecto a su i -ésimo argumento.

- Finalmente, mostremos que la función F_n es alternante usando el criterio que nos da la Proposición 4.1.5. Sean finalmente $x_1, \dots, x_n \in V$ y $r \in \llbracket n-1 \rrbracket$ tales que $x_r = x_{r+1}$. De la definición de D_n tenemos que

$$D_n(x_1, \dots, x_n) = \sum_{l=1}^n (-1)^{l+1} x'_l D_{n-1}(x''_1, \dots, \widehat{x''_l}, \dots, x''_n).$$

Si el índice $l \in \llbracket n \rrbracket$ es distinto de r y de $r+1$, el término l -ésimo de esta suma tiene a $D_{n-1}(x''_1, \dots, \widehat{x''_l}, \dots, x''_n)$ como factor, y este escalar es nulo porque D_{n-1} es una función alternante: entre sus argumentos aparecen los vectores x''_r y x''_{r+1} , que son iguales. Vemos así que

$$D_n(x_1, \dots, x_n) \\ = (-1)^{r+1} x'_r D_{n-1}(x''_1, \dots, \widehat{x''_r}, \dots, x''_n) + (-1)^{r+2} x'_{r+1} D_{n-1}(x''_1, \dots, \widehat{x''_{r+1}}, \dots, x''_n).$$

Como $x_r = x_{r+1}$, tenemos por supuesto que $x'_r = x'_{r+1}$ y $x''_r = x''_{r+1}$ y, en particular, las secuencias de $n-1$ vectores

$$x''_1, \dots, \widehat{x''_r}, x''_{r+1}, \dots, x''_n \quad \text{y} \quad x''_1, \dots, x''_r, \widehat{x''_{r+1}}, \dots, x''_n$$

coinciden. Usando esto en la expresión para $D_n(x_1, \dots, x_n)$ que obtuvimos arriba, podemos concluir que

$$D_n(x_1, \dots, x_n) = ((-1)^{r+1} + (-1)^{r+2}) \cdot x'_r D_{n-1}(x''_1, \dots, \widehat{x''_r}, \dots, x''_n) = 0,$$

ya que el escalar entre paréntesis es nulo.

La proposición queda así probada. \square

4.2.5. Para describir el espacio $\text{Alt}^n(V)$ cuando V es un espacio de dimensión finita n arbitrario necesitamos el siguiente resultado:

Proposición. Sean V y W dos espacios vectoriales y sea $d \in \mathbb{N}$.

(i) Si $\phi : V \rightarrow W$ es una función lineal y $f : W^d \rightarrow \mathbb{k}$ es una función d -multilineal y alternante, entonces la función

$$\phi^*(f) : (x_1, \dots, x_d) \in V^d \mapsto f(\phi(x_1), \dots, \phi(x_d)) \in \mathbb{k}$$

es d -multilineal y alternante.

(ii) Si $\phi : V \rightarrow W$ es una función lineal, entonces

$$\phi^* : f \in \text{Alt}^d(W) \mapsto \phi^*(f) \in \text{Alt}^d(V)$$

es también una función lineal.

(iii) Si $\text{id}_V : V \rightarrow V$ es la función identidad de V , entonces

$$\text{id}_V^* : \text{Alt}^d(V) \rightarrow \text{Alt}^d(V)$$

es la función identidad de $\text{Alt}^d(V)$.

(iv) Si $\phi : U \rightarrow V$ y $\psi : V \rightarrow W$ son funciones lineales, entonces

$$(\psi \circ \phi)^* = \phi^* \circ \psi^* : \text{Alt}^d(W) \rightarrow \text{Alt}^d(U).$$

(v) Si $\phi : V \rightarrow W$ es un isomorfismo, entonces la función

$$\phi^* : \text{Alt}^d(W) \rightarrow \text{Alt}^d(V)$$

es un isomorfismo.

Observemos que la función ϕ^* de la segunda parte de la proposición tiene sentido precisamente por lo que afirma la primera.

Demostración. (i) Sea $\phi : V \rightarrow W$ una función lineal, sea $f : W^d \rightarrow \mathbb{k}$ una función d -multilineal y alternante, y sea $g : V^d \rightarrow \mathbb{k}$ la función tal que

$$g(x_1, \dots, x_d) = f(\phi(x_1), \dots, \phi(x_d))$$

siempre que $x_1, \dots, x_d \in V$. Esta es la función que en el enunciado del lema aparece nombrada como $\phi^*(f)$: tenemos que probar que g es d -multilineal y alternante.

- Sea primero $i \in [\![d]\!]$. Si $x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_d, y, z \in V$ y $a, b \in \mathbb{k}$, entonces

$$g(x_1, \dots, x_{i-1}, ay + bz, x_{i+1}, \dots, x_d)$$

$$\begin{aligned}
&= f(\phi(x_1), \dots, \phi(x_{i-1}), \phi(ay + bz), \phi(x_{i+1}), \dots, \phi(x_d)) \\
&= f(\phi(x_1), \dots, \phi(x_{i-1}), a\phi(y) + b\phi(z), \phi(x_{i+1}), \dots, \phi(x_d)) \\
&= af(\phi(x_1), \dots, \phi(x_{i-1}), \phi(y), \phi(x_{i+1}), \dots, \phi(x_d)) \\
&\quad + bf(\phi(x_1), \dots, \phi(x_{i-1}), \phi(z), \phi(x_{i+1}), \dots, \phi(x_d)) \\
&= ag(x_1, \dots, x_{i-1}, y, x_{i+1}, \dots, x_d) + bg(x_1, \dots, x_{i-1}, z, x_{i+1}, \dots, x_d),
\end{aligned}$$

Esto nos dice que la función g es lineal con respecto a su i -ésimo argumento y, como i es cualquier elemento de $\llbracket d \rrbracket$, que es d -multilineal.

- Por otro lado, si x_1, \dots, x_d son elementos de V entre los que hay dos iguales, entonces

$$g(x_1, \dots, x_d) = f(\phi(x_1), \dots, \phi(x_d)) = 0$$

porque entre los d vectores $\phi(x_1), \dots, \phi(x_d)$ hay dos que son iguales y f es alternante. Vemos así que la función g es alternante.

(ii) Sea $\phi : V \rightarrow W$ una función lineal. De acuerdo a lo que ya probamos, para cada $f \in \text{Alt}^d(W)$ la función $\phi^*(f) : V^d \rightarrow \mathbb{k}$ es un elemento de $\text{Alt}^d(V)$, así que efectivamente hay una función

$$\phi^* : f \in \text{Alt}^d(W) \mapsto \phi^*(f) \in \text{Alt}^d(V).$$

Mostremos que es lineal: sean f y g dos elementos de $\text{Alt}^d(W)$ y sean a y b dos elementos de \mathbb{k} . Si x_1, \dots, x_d son elementos de V , entonces

$$\begin{aligned}
\phi^*(af + bg)(x_1, \dots, x_d) &= (af + bg)(\phi(x_1), \dots, \phi(x_d)) \\
&= af(\phi(x_1), \dots, \phi(x_d)) + bg(\phi(x_1), \dots, \phi(x_d)) \\
&= a\phi^*(f)(x_1, \dots, x_d) + b\phi^*(g)(x_1, \dots, x_d) \\
&= (a\phi^*(f) + b\phi^*(g))(x_1, \dots, x_d),
\end{aligned}$$

y esto nos dice que, como queremos, $\phi^*(af + bg) = a\phi^*(f) + b\phi^*(g)$.

(iii) Para ver que $\text{id}_V^* : \text{Alt}^d(V) \rightarrow \text{Alt}^d(V)$ es la función identidad de $\text{Alt}^d(V)$ simplemente calculamos: si $f \in \text{Alt}^d(V)$, entonces cada vez que $x_1, \dots, x_d \in V$ tenemos que

$$\text{id}_V^*(f)(x_1, \dots, x_d) = f(\text{id}_V(x_1), \dots, \text{id}_V(x_d)) = f(x_1, \dots, x_d)$$

y, por lo tanto, $\text{id}_V^*(f) = f = \text{id}_{\text{Alt}^d(V)}(f)$.

(iv) Sean $\phi : U \rightarrow V$ y $\psi : V \rightarrow W$ funciones lineales. Si $f \in \text{Alt}^d(W)$, entonces cada vez que $x_1, \dots, x_d \in V$ se tiene que

$$\begin{aligned}
(\psi \circ \phi)^*(f)(x_1, \dots, x_d) &= f(\psi(\phi(x_1)), \dots, \psi(\phi(x_d))) \\
&= \psi^*(f)(\phi(x_1), \dots, \phi(x_d)) \\
&= \phi^*(\psi^*(f))(x_1, \dots, x_d) \\
&= (\phi^* \circ \psi^*)(f)(x_1, \dots, x_d),
\end{aligned}$$

de manera que $(\psi \circ \phi)^*(f) = (\phi^* \circ \psi^*)(f)$. Como esto es así cualquiera sea $f \in \text{Alt}^d(W)$, vemos que, en definitiva, $(\psi \circ \phi)^* = \phi^* \circ \psi^*$, como afirma el enunciado.

(v) Sea finalmente $\phi : V \rightarrow W$ un isomorfismo y sea $\psi : W \rightarrow V$ el isomorfismo inverso de ϕ . Podemos considerar entonces las funciones lineales $\phi^* : \text{Alt}^d(W) \rightarrow \text{Alt}^d(V)$ y $\psi^* : \text{Alt}^d(V) \rightarrow \text{Alt}^d(W)$, y de acuerdo a lo que probamos en (iii) y (iv) tenemos que

$$\psi^* \circ \phi^* = (\phi \circ \psi)^* = \text{id}_W^* = \text{id}_{\text{Alt}^d(W)}$$

y, de manera similar,

$$\phi^* \circ \psi^* = \text{id}_{\text{Alt}^d(V)}.$$

Vemos así que ϕ^* y ψ^* son isomorfismo mutuamente inversos. \square

4.2.6. Podemos ahora por fin describir el espacio $\text{Alt}^n(V)$ para un espacio vectorial V de dimensión finita n arbitrario:

Corolario. *Sea V un espacio vectorial de dimensión finita, sea $n = \dim V$ y sea $\mathcal{B} = (x_1, \dots, x_n)$ una base ordenada de V .*

- (i) *Hay exactamente una función $D_{\mathcal{B}} : V^n \rightarrow \mathbb{k}$ que es n -multilineal y alternante y tal que $D_{\mathcal{B}}(x_1, \dots, x_n) = 1$.*
- (ii) *Se tiene que $\dim \text{Alt}^n(V) = 1$.*

Demostración. Sea (e_1, \dots, e_n) la base ordenada estándar de \mathbb{k}^n . Sabemos que hay un isomorfismo $\phi : V \rightarrow \mathbb{k}^n$ tal que $\phi(x_i) = e_i$ y de acuerdo a la última parte de la Proposición 4.2.5 a partir de ϕ podemos construir un isomorfismo $\phi^* : \text{Alt}^n(\mathbb{k}^n) \rightarrow \text{Alt}^n(V)$. La Proposición 4.2.4 nos dice que el espacio $\text{Alt}^n(\mathbb{k}^n)$ tiene dimensión 1, así que $\text{Alt}^n(V)$ también tiene dimensión 1. Más aún, tenemos el elemento $D_{\mathcal{B}} = \phi^*(D_n)$ en $\text{Alt}^n(V)$, para el cual se tiene que

$$D_{\mathcal{B}}(x_1, \dots, x_n) = \phi^*(D_n)(x_1, \dots, x_n) = D_n(\phi(x_1), \dots, \phi(x_n)) = D_n(e_1, \dots, e_n) = 1.$$

Finalmente, si D es un elemento arbitrario de $\text{Alt}^n(V)$, sabemos que existe un escalar $\lambda \in \mathbb{k}$ tal que $D = \lambda D_{\mathcal{B}}$, porque $D_{\mathcal{B}}$ es un elemento no nulo de espacio 1-dimensional $\text{Alt}^n(V)$, y entonces

$$D(x_1, \dots, x_n) = \lambda D_{\mathcal{B}}(x_1, \dots, x_n) = \lambda,$$

que es 1 si y solamente si $\lambda = 1$, esto es, si y solamente si $D = D_{\mathcal{B}}$. \square

§3. El determinante de una matriz

4.3.1. Sea $n \in \mathbb{N}$ y sea (e_1, \dots, e_n) la base ordenada estándar de \mathbb{k}^n . Si $A \in M_n(\mathbb{k})$ es una matriz cuadrada de n filas y n columnas con entradas en \mathbb{k} , el **determinante** de A , al que escribimos $\det(A)$ o $|A|$, es el escalar

$$\det(A) := D_n(Ae_1, \dots, Ae_n)$$

que se obtiene al evaluar la función $D_n : (\mathbb{k}^n)^n \rightarrow \mathbb{k}$ de la Proposición 4.2.4 en los n vectores Ae_1, \dots, Ae_n de \mathbb{k}^n . Observemos que estos vectores son precisamente las columnas de la matriz A . Si A es la matriz

$$\begin{pmatrix} a_{1,1} & \cdots & a_{1,n} \\ \vdots & \ddots & \vdots \\ a_{n,1} & \cdots & a_{n,n} \end{pmatrix}$$

escribimos muchas veces al determinante de A usando la notación

$$\begin{vmatrix} a_{1,1} & \cdots & a_{1,n} \\ \vdots & \ddots & \vdots \\ a_{n,1} & \cdots & a_{n,n} \end{vmatrix}.$$

La prueba de la Proposición 4.2.4 da una construcción recursiva de la función D_n , que la expresa en términos de D_{n-1} : siguiendo esa construcción, vemos que

- Si $n = 1$, entonces

$$|a_{1,1}| = D_1((a_{1,1})) = a_{1,1}.$$

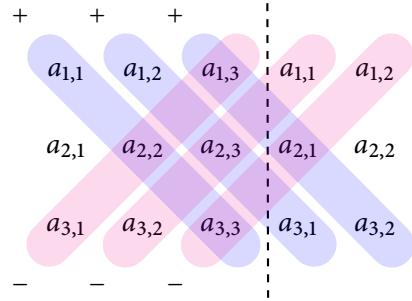
- Si $n = 2$, entonces

$$\begin{aligned} \begin{vmatrix} a_{1,1} & a_{1,2} \\ a_{2,1} & a_{2,2} \end{vmatrix} &= D_2\left(\begin{pmatrix} a_{1,1} \\ a_{2,1} \end{pmatrix}, \begin{pmatrix} a_{1,2} \\ a_{2,2} \end{pmatrix}\right) \\ &= a_{1,1}D_1((a_{2,2})) - a_{1,2}D_1((a_{2,1})) \\ &= a_{1,1}a_{2,2} - a_{1,2}a_{2,1}. \end{aligned}$$

- Si $n = 3$, entonces

$$\begin{aligned} \begin{vmatrix} a_{1,1} & a_{1,2} & a_{1,3} \\ a_{2,1} & a_{2,2} & a_{2,3} \\ a_{3,1} & a_{3,2} & a_{3,3} \end{vmatrix} &= D_3\left(\begin{pmatrix} a_{1,1} \\ a_{2,1} \\ a_{3,1} \end{pmatrix}, \begin{pmatrix} a_{1,2} \\ a_{2,2} \\ a_{3,2} \end{pmatrix}, \begin{pmatrix} a_{1,3} \\ a_{2,3} \\ a_{3,3} \end{pmatrix}\right) \\ &= a_{1,1}D_2\left(\begin{pmatrix} a_{2,2} \\ a_{3,2} \end{pmatrix}, \begin{pmatrix} a_{2,3} \\ a_{3,3} \end{pmatrix}\right) - a_{1,2}D_2\left(\begin{pmatrix} a_{2,1} \\ a_{3,1} \end{pmatrix}, \begin{pmatrix} a_{2,3} \\ a_{3,3} \end{pmatrix}\right) + a_{1,3}D_2\left(\begin{pmatrix} a_{2,1} \\ a_{3,1} \end{pmatrix}, \begin{pmatrix} a_{2,2} \\ a_{3,2} \end{pmatrix}\right) \\ &= a_{1,1}(a_{2,2}a_{3,3} - a_{3,2}a_{2,3}) - a_{1,2}(a_{2,1}a_{3,3} - a_{3,1}a_{2,3}) + a_{1,3}(a_{2,1}a_{3,2} - a_{3,1}a_{2,2}) \\ &= a_{1,1}a_{2,2}a_{3,3} - a_{1,1}a_{3,2}a_{2,3} - a_{1,2}a_{2,1}a_{3,3} + a_{1,2}a_{3,1}a_{2,3} + a_{1,3}a_{2,1}a_{3,2} - a_{1,3}a_{3,1}a_{2,2}. \end{aligned}$$

Una forma de recordar esta última expresión es la siguiente *regla de Sarrus*, por *Pierre Frédéric Sarrus*: si copiamos las dos primeras columnas a la derecha de la matriz, para obtener una matriz de 3 filas y 5 columnas,



entonces el determinante de la matriz es la suma de los tres productos correspondientes a las diagonales azules del dibujo menos la suma de los tres productos correspondientes a las rojas.

- Si $n = 4$, entonces

$$\begin{aligned}
 \begin{vmatrix} a_{1,1} & a_{1,2} & a_{1,3} & a_{1,4} \\ a_{2,1} & a_{2,2} & a_{2,3} & a_{2,4} \\ a_{3,1} & a_{3,2} & a_{3,3} & a_{3,4} \\ a_{4,1} & a_{4,2} & a_{4,3} & a_{4,4} \end{vmatrix} &= D_4 \left(\begin{pmatrix} a_{1,1} \\ a_{2,1} \\ a_{3,1} \\ a_{4,1} \end{pmatrix}, \begin{pmatrix} a_{1,2} \\ a_{2,2} \\ a_{3,2} \\ a_{4,2} \end{pmatrix}, \begin{pmatrix} a_{1,3} \\ a_{2,3} \\ a_{3,3} \\ a_{4,3} \end{pmatrix}, \begin{pmatrix} a_{1,4} \\ a_{2,4} \\ a_{3,4} \\ a_{4,4} \end{pmatrix} \right) \\
 &= a_{1,1} D_3 \left(\begin{pmatrix} a_{2,2} \\ a_{3,2} \\ a_{4,2} \end{pmatrix}, \begin{pmatrix} a_{2,3} \\ a_{3,3} \\ a_{4,3} \end{pmatrix}, \begin{pmatrix} a_{2,4} \\ a_{3,4} \\ a_{4,4} \end{pmatrix} \right) - a_{1,2} D_3 \left(\begin{pmatrix} a_{2,1} \\ a_{3,1} \\ a_{4,1} \end{pmatrix}, \begin{pmatrix} a_{2,3} \\ a_{3,3} \\ a_{4,3} \end{pmatrix}, \begin{pmatrix} a_{2,4} \\ a_{3,4} \\ a_{4,4} \end{pmatrix} \right) \\
 &\quad + a_{1,3} D_3 \left(\begin{pmatrix} a_{2,1} \\ a_{3,1} \\ a_{4,1} \end{pmatrix}, \begin{pmatrix} a_{2,2} \\ a_{3,2} \\ a_{4,2} \end{pmatrix}, \begin{pmatrix} a_{2,4} \\ a_{3,4} \\ a_{4,4} \end{pmatrix} \right) - a_{1,4} D_3 \left(\begin{pmatrix} a_{2,1} \\ a_{3,1} \\ a_{4,1} \end{pmatrix}, \begin{pmatrix} a_{2,2} \\ a_{3,2} \\ a_{4,2} \end{pmatrix}, \begin{pmatrix} a_{2,3} \\ a_{3,3} \\ a_{4,3} \end{pmatrix}, \right) \\
 &= a_{1,1}, a_{2,2}, a_{3,3}, a_{4,4} - a_{1,1}, a_{2,2}, a_{4,3}, a_{3,4} - a_{1,1}, a_{3,2}, a_{2,3}, a_{4,4} \\
 &\quad + a_{1,1}, a_{3,2}, a_{4,3}, a_{2,4} + a_{1,1}, a_{4,2}, a_{2,3}, a_{3,4} - a_{1,1}, a_{4,2}, a_{3,3}, a_{2,4} \\
 &\quad - a_{2,1}, a_{1,2}, a_{3,3}, a_{4,4} + a_{2,1}, a_{1,2}, a_{4,3}, a_{3,4} + a_{2,1}, a_{3,2}, a_{1,3}, a_{4,4} \\
 &\quad - a_{2,1}, a_{3,2}, a_{4,3}, a_{1,4} - a_{2,1}, a_{4,2}, a_{1,3}, a_{3,4} + a_{2,1}, a_{4,2}, a_{3,3}, a_{1,4} \\
 &\quad + a_{3,1}, a_{1,2}, a_{2,3}, a_{4,4} - a_{3,1}, a_{1,2}, a_{4,3}, a_{2,4} - a_{3,1}, a_{2,2}, a_{1,3}, a_{4,4} \\
 &\quad + a_{3,1}, a_{2,2}, a_{4,3}, a_{1,4} + a_{3,1}, a_{4,2}, a_{1,3}, a_{2,4} - a_{3,1}, a_{4,2}, a_{2,3}, a_{1,4} \\
 &\quad - a_{4,1}, a_{1,2}, a_{2,3}, a_{3,4} + a_{4,1}, a_{1,2}, a_{3,3}, a_{2,4} + a_{4,1}, a_{2,2}, a_{1,3}, a_{3,4} \\
 &\quad - a_{4,1}, a_{2,2}, a_{3,3}, a_{1,4} - a_{4,1}, a_{3,2}, a_{1,3}, a_{2,4} + a_{4,1}, a_{3,2}, a_{2,3}, a_{1,4}.
 \end{aligned}$$

Podemos continuar de esta forma, pero rápidamente el tamaño de las fórmulas crece: puede verse en las que acabamos de dar que la cantidad de términos en el determinante de una matriz

cuadrada de tamaño 1, 2, 3 y 4 es, respectivamente, 1, 2, 6 y 24. Más generalmente, es fácil deducir inductivamente de la prueba de la Proposición 4.2.4 que tenemos una expresión con a lo sumo $n!$ términos —cada uno de los cuales es un producto de n entradas de la matriz— para el determinante de una matriz cuadrada de tamaño n .

4.3.2. Definimos el determinante de una matriz usando la función D_n que construimos en la sección anterior y, por supuesto, las propiedades de esta nos dan propiedades de aquel.

Proposición. Sea $n \in \mathbb{N}$. La función $\det : M_n(\mathbb{k}) \rightarrow \mathbb{k}$ tiene las siguientes propiedades:

- (i) $\det(A)$ es una función multilínea alternante de las columnas de la matriz A .
- (ii) Si $A \in M_n(\mathbb{k})$ tiene dos columnas iguales o, más generalmente, si las columnas de A son linealmente dependientes, entonces $\det(A) = 0$.
- (iii) Si A y B son elementos de $M_n(\mathbb{k})$ y B se obtiene de A intercambiando dos columnas, entonces $\det(B) = -\det(A)$.
- (iv) Si A y B son elementos de $M_n(\mathbb{k})$ y B se obtiene de A sumando a una de sus columnas una combinación lineal de las demás, entonces $\det(B) = \det(A)$.

Demostración. La primera afirmación es consecuencia inmediata de la definición de \det y de la Proposición 4.2.4 y las demás se deducen de ella. \square

4.3.3. Usando el hecho de que la función D_n genera al espacio vectorial $\text{Alt}^n(\mathbb{k}^n)$ podemos obtener una de las propiedades más importante del determinante, que es la segunda de las afirmaciones de la siguiente proposición.

Proposición. Sea $n \in \mathbb{N}$.

- (i) Si $I_n \in M_n(\mathbb{k})$ es la matriz identidad, entonces $\det(I_n) = 1$.
- (ii) Si A y B son matrices de $M_n(\mathbb{k})$, entonces $\det(AB) = \det(A) \cdot \det(B)$.
- (iii) Si $A \in M_n(\mathbb{k})$ es inversible, entonces $\det(A) \neq 0$ y, de hecho, $\det(A^{-1}) = (\det(A))^{-1}$.

Demostración. (i) De acuerdo a nuestras definiciones y la elección de la función D_n , es

$$\det(I_n) = D_n(I_n e_1, \dots, I_n e_n) = D_n(e_1, \dots, e_n) = 1.$$

(ii) Sean A y B dos elementos de $M_n(\mathbb{k})$ y sea $\tilde{D} : (\mathbb{k}^n)^n \rightarrow \mathbb{k}$ la función tal que cada vez que x_1, \dots, x_n son elementos de \mathbb{k}^n es

$$\tilde{D}(x_1, \dots, x_n) = D_n(Ax_1, \dots, Ax_n).$$

Como $D_n : (\mathbb{k}^n)^n \rightarrow \mathbb{k}$ es una función n -multilínea y alternante, es inmediato verificar que lo mismo es cierto de \tilde{D} , esto es, que \tilde{D} es un elemento del espacio vectorial $\text{Alt}^n(\mathbb{k}^n)$. Sabemos que este espacio vectorial tiene dimensión 1 y que $\{D_n\}$ es una de sus bases: esto significa que existe un escalar $\alpha \in \mathbb{k}$ tal que $\tilde{D} = \alpha D_n$ y, de hecho, si (e_1, \dots, e_n) es la base ordenada estándar de \mathbb{k}^n

tenemos que

$$\alpha = \alpha D_n(e_1, \dots, e_n) = \tilde{D}(e_1, \dots, e_n) = D_n(Ae_1, \dots, Ae_n) = \det(A).$$

Esto implica que

$$\tilde{D}(Be_1, \dots, Be_n) = \det(A \cdot D(Be_1, \dots, Be_n)) = \det(A) \cdot \det(B)$$

mientras que la definición misma de la función \tilde{D} nos dice, por otro lado, que

$$\tilde{D}(Be_1, \dots, Be_n) = D_n(ABe_1, \dots, ABe_n) = \det(AB).$$

Esto prueba la afirmación (ii) de la proposición.

(iii) Sea $A \in M_n(\mathbb{k})$ una matriz inversible. Usando las partes (ii) y (i) de la proposición, vemos que

$$\det(A) \cdot \det(A^{-1}) = \det(AA^{-1}) = \det(I_n) = 1,$$

así que claramente $\det(A) \neq 0$ y $\det(A^{-1}) = (\det(A))^{-1}$, que es lo que queríamos probar. \square

4.3.4. Un corolario inmediato pero importante de la Proposición 4.3.3 es el siguiente:

Corolario. Si A y B son elementos de $M_n(\mathbb{k})$ para los que existe una matriz $C \in M_n(\mathbb{k})$ inversible tal que $A = CBC^{-1}$, entonces $\det(A) = \det(B)$.

Demostración. En efecto, en ese caso la Proposición 4.3.3 nos dice que $\det(C) \neq 0$ y que

$$\det(A) = \det(CBC^{-1}) = \det(C) \cdot \det(B) \cdot \det(C^{-1}) = \det(C) \cdot \det(B) \cdot \det(C)^{-1} = \det(B),$$

como afirma el corolario. \square

4.3.5. El Corolario 4.3.4 tiene una aplicación inmediata a la definición del determinante de un endomorfismo de un espacio vectorial. Terminemos esta sección dando los detalles de esto.

Si V es un espacio vectorial de dimensión finita, \mathcal{B} una base ordenada de V y $f : V \rightarrow V$ un endomorfismo de V , llamamos **determinante de f** al determinante de la matriz $[f]_{\mathcal{B}}^{\mathcal{B}}$. Esta definición tiene sentido porque este escalar no depende de la elección de la base \mathcal{B} sino solamente de f : esto es el contenido de la siguiente proposición.

Proposición. Sea V un espacio vectorial de dimensión finita y sean \mathcal{B} y \mathcal{B}' dos bases ordenadas de V . Si $f : V \rightarrow V$ es un endomorfismo de V , entonces

$$\det([f]_{\mathcal{B}}^{\mathcal{B}'}) = \det([f]_{\mathcal{B}'}^{\mathcal{B}}).$$

Demostración. Sabemos de la Proposición 2.6.4 que

$$[f]_{\mathcal{B}'}^{\mathcal{B}'} = C(\mathcal{B}, \mathcal{B}') \cdot [f]_{\mathcal{B}}^{\mathcal{B}} \cdot C(\mathcal{B}', \mathcal{B})$$

y de acuerdo a la Proposición 1.11.4(iii) las matrices $C(\mathcal{B}, \mathcal{B}')$ y $C(\mathcal{B}', \mathcal{B})$ son mutuamente inversas, así que el Corolario 4.3.4 nos permite concluir lo que queremos. \square

4.3.6. Proposición. *Sea V un espacio vectorial de dimensión finita.*

- (i) *Si $\text{id}_V : V \rightarrow V$ es el endomorfismo identidad de V , entonces $\det(\text{id}_V) = 1$.*
- (ii) *Si $f, g : V \rightarrow V$ son dos endomorfismos de V , entonces $\det(f \circ g) = \det(f) \cdot \det(g)$.*
- (iii) *Si $h : V \rightarrow V$ es un endomorfismo de V , entonces h es un automorfismo exactamente cuando $\det(h) \neq 0$ y en ese caso se tiene que $\det(h^{-1}) = (\det(h))^{-1}$.*

Demostración. Sea $n = \dim V$ y sea \mathcal{B} una base ordenada de V . Sabemos que $[\text{id}_V] = I_n$, la matriz identidad de $M_n(\mathbb{k})$, así que $\det(\text{id}_V) = \det(I_n) = 1$: esto prueba (i).

Si $f, g : V \rightarrow V$ son endomorfismos de V , entonces

$$\det(f \circ g) = \det([f \circ g]_{\mathcal{B}}^{\mathcal{B}}) = \det([f]_{\mathcal{B}}^{\mathcal{B}} \cdot [g]_{\mathcal{B}}^{\mathcal{B}}) = \det([f]_{\mathcal{B}}^{\mathcal{B}}) \cdot \det([g]_{\mathcal{B}}^{\mathcal{B}}) = \det(f) \cdot \det(g),$$

como afirma (ii). Por otro lado, si $h : V \rightarrow V$ es un endomorfismo, sabemos de la Proposición 2.6.9 que h es un isomorfismo si y solamente si la matriz $[h]_{\mathcal{B}}^{\mathcal{B}}$ es inversible y, de acuerdo a la Proposición 4.6.5(ii), esto ocurre exactamente cuando el escalar $\det(h) = \det([h]_{\mathcal{B}}^{\mathcal{B}})$ es no nulo. Si ése es el caso, entonces

$$\det(h) \cdot \det(h^{-1}) = \det(h \circ h^{-1}) = \det(\text{id}_V) = 1$$

y, por supuesto, se sigue que $\det(h^{-1}) = (\det(h))^{-1}$. \square

§4. Permutaciones y sus signos

4.4.1. Una aplicación sencilla de la existencia del determinante es la construcción del signo de las permutaciones. Dedicamos esta sección a estudiar en detalle esto — por un lado, porque es un ejemplo ilustrativo e importante y, por otro, porque necesitaremos estos resultados para dar en la Sección 4.5 una fórmula explícita para el determinante.

4.4.2. Si $n \in \mathbb{N}$, escribimos, como siempre, $\llbracket n \rrbracket$ al conjunto $\{1, \dots, n\}$. Una **permutación** de $\llbracket n \rrbracket$ es una función biyectiva $\sigma : \llbracket n \rrbracket \rightarrow \llbracket n \rrbracket$; a la función identidad $\text{id}_n : \llbracket n \rrbracket \rightarrow \llbracket n \rrbracket$ la llamamos **permutación identidad** y la escribimos id_n . Finalmente, escribimos S_n al conjunto de todas las permutaciones de $\llbracket n \rrbracket$.

Proposición. *Sea $n \in \mathbb{N}$.*

- (i) *El conjunto S_n tiene exactamente $n!$ elementos.*
- (ii) *Si σ y τ son dos permutaciones de $\llbracket n \rrbracket$, entonces la composición $\sigma \circ \tau$ también lo es.*
- (iii) *Si σ es una permutación de $\llbracket n \rrbracket$, entonces la función inversa σ^{-1} —que existe porque σ es*

biyectiva—también es una permutación de $\llbracket n \rrbracket$.

Demostración. Todas las afirmaciones son inmediatas. \square

4.4.3. Cuando n es pequeño, escribimos a veces a una permutación $\sigma \in S_n$ en la forma de una matriz de 2 filas y n columnas

$$\begin{pmatrix} 1 & 2 & \cdots & n \\ \sigma(1) & \sigma(2) & \cdots & \sigma(n) \end{pmatrix}$$

listando en la primera fila los elementos de $\llbracket n \rrbracket$ y en la segunda las correspondientes imágenes por σ . De manera todavía más compacta, escribimos a σ también como la secuencia ordenada de sus valores

$$\sigma(1) \sigma(2) \cdots \sigma(n).$$

Así, por 35142 y por $(\begin{smallmatrix} 1 & 2 & 3 & 4 & 5 \\ 3 & 5 & 1 & 4 & 2 \end{smallmatrix})$ denotamos a la permutación de $\llbracket 5 \rrbracket$ que manda $1, 2, 3, 4$ y 5 a $3, 5, 1, 4$ y 2 , respectivamente.

4.4.4. Sea $n \in \mathbb{N}$ y sean r y s dos elementos distintos de $\llbracket n \rrbracket$. Escribimos (rs) a la permutación de $\llbracket n \rrbracket$ tal que para cada $i \in \llbracket n \rrbracket$ es

$$(rs)(i) = \begin{cases} s, & \text{si } i = r; \\ r, & \text{si } i = s; \\ i, & \text{si } i \neq \{r, s\}. \end{cases}$$

Observemos que (rs) y (sr) denotan la misma permutación.

Una permutación σ de $\llbracket n \rrbracket$ es una **transposición** si el conjunto $\{i \in \llbracket n \rrbracket : \sigma(i) \neq i\}$ tiene exactamente dos elementos; si ése es el caso y r y s son esos dos elementos, es claro que $\sigma = (rs)$.

4.4.5. Toda permutación es composición de transposiciones: en ese sentido, pueden verse como permutaciones elementales.

Proposición. Sea $n \in \mathbb{N}$. Si σ es una permutación de $\llbracket n \rrbracket$, entonces existen $l \in \mathbb{N}_0$ y transposiciones τ_1, \dots, τ_l de S_n tales que $\sigma = \tau_1 \circ \cdots \circ \tau_l$.

Demostración. Para cada $\sigma \in S_n$ sea $|\sigma|$ el conjunto $\{i \in \llbracket n \rrbracket : \sigma(i) \neq i\}$ y para cada $k \in \llbracket 0, n \rrbracket$ sea $P(k)$ la afirmación

$$\begin{aligned} &\text{si } \sigma \in S_n \text{ y el conjunto } |\sigma| \text{ tiene a lo sumo } k \text{ elementos, entonces existen } l \in \mathbb{N}_0 \text{ y} \\ &\text{transposiciones } \tau_1, \dots, \tau_l \text{ de } S_n \text{ tales que } \sigma = \tau_1 \circ \cdots \circ \tau_n. \end{aligned} \tag{5}$$

Como cualquiera sea $\sigma \in S_n$ el conjunto $|\sigma|$ tiene a lo sumo n elementos, para probar a proposición es suficiente con que probemos que la afirmación $P(n)$ vale. Procedemos inductivamente.

Consideremos primero la afirmación $P(0)$. Si σ es una permutación tal que $|\sigma|$ tiene a lo sumo 0 elementos, entonces por supuesto $|\sigma| = \emptyset$ y, en consecuencia, $\sigma(i) = i$ para cada $i \in \llbracket n \rrbracket$:

vemos así que $\sigma = \text{id}_n$ y podemos tomar $l = 0$ en (5), ya que la composición de cero permutaciones es igual a id_n . Vemos que de esta manera que la afirmación $P(0)$ vale.

Supongamos ahora que $k \in \llbracket n \rrbracket$ y que sabemos que $P(k - 1)$ vale, y sea σ una permutación de $\llbracket n \rrbracket$ tal que $|\sigma|$ tiene a lo sumo k elementos. Si tiene *menos* que k elementos, entonces tiene a lo sumo $k - 1$ y, como estamos suponiendo que vale $P(k - 1)$, existen $l \in \mathbb{N}_0$ y transposiciones τ_1, \dots, τ_l de $\llbracket n \rrbracket$ tales que $\sigma = \tau_1 \circ \dots \circ \tau_l$. Esto muestra que $P(k)$ vale en ese caso.

Nos resta ocuparnos del caso en el que el conjunto $|\sigma|$ tiene exactamente k elementos y, en particular, no es vacío. Sea $i_0 \in |\sigma|$ e $i_1 = \sigma(i_0)$, que es distinto de i_0 . Observemos que $i_1 \in |\sigma|$: tendríamos si no que $\sigma(i_1) = i_1$, de manera que $i_1 = \sigma^{-1}(i_1) = i_0$, lo que es absurdo.

Sea $\tau = (i_0 \ i_1) \circ \sigma$. Si $i \in \llbracket n \rrbracket \setminus |\sigma|$, entonces

$$\tau(i) = (i_0 \ i_1)(\sigma(i)) = (i_0 \ i_1)(i) = i,$$

ya que $i \notin \{i_0, i_1\}$. Esto significa que $|\tau| \subseteq |\sigma|$. Por otro lado, es

$$\tau(i_0) = (i_0 \ i_1)(\sigma(i_0)) = (i_0 \ i_1)(i_1) = i_0,$$

así que $i_0 \notin |\tau|$. Vemos de esta forma que $|\tau|$ está estrictamente contenido en $|\sigma|$ y que, en particular, tiene a lo sumo $k - 1$ elementos. Como estamos suponiendo que vale la afirmación $P(k - 1)$, esto implica que existen $l \in \mathbb{N}$ y transposiciones τ_1, \dots, τ_l de $\llbracket n \rrbracket$ tales que $\tau = \tau_1 \circ \dots \circ \tau_l$. Como $(i_0 \ i_1) \circ (i_0 \ i_1) = \text{id}$, se tiene entonces que

$$\sigma = (i_0 \ i_1) \circ (i_0 \ i_1) \circ \sigma = (i_0 \ i_1) \circ \tau = (i_0 \ i_1) \circ \tau_1 \circ \dots \circ \tau_l,$$

esto es, que σ es igual a una composición de transposiciones. Esto prueba que la afirmación $P(k)$ también vale en este caso, y completa la demostración de la proposición. \square

4.4.6. La Proposición 4.4.5 nos dice que toda permutación puede escribirse como una composición de transposiciones. Es importante observar que esta escritura no es única: ni las transposiciones que aparecen en ella ni su cantidad están determinadas por la permutación. Es fácil verificar, por ejemplo, que en S_4 se tiene que

$$(13) = (12) \circ (23) \circ (12) = (23) \circ (12) \circ (23) = (23) \circ (14) \circ (24) \circ (23) \circ (14).$$

4.4.7. Si $n \in \mathbb{N}$ y $\sigma \in S_n$ es una permutación, la **matriz de permutación** asociada a σ es la matriz $A(\sigma) = (a_{i,j}) \in M_n(\mathbb{K})$ tal que para cada $i, j \in \llbracket n \rrbracket$ es

$$a_{i,j} = \begin{cases} 1, & \text{si } \sigma(j) = i; \\ 0, & \text{en caso contrario.} \end{cases}$$

Por ejemplo, la matriz de permutación correspondiente a la permutación 3 5 1 4 2 de $\llbracket 5 \rrbracket$ es

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

La propiedad útil de las matrices de permutación es la siguiente:

Proposición. *Sea $n \in \mathbb{N}$ y sea σ una permutación de $[\![n]\!]$. Para cada $i \in [\![n]\!]$ se tiene que*

$$A(\sigma)e_i = e_{\sigma(i)}.$$

Demostración. Sea $i \in [\![n]\!]$ y sea $A(\sigma) = (a_{i,j})$. Si $k \in [\![n]\!]$, la componente k -ésima del vector $A(\sigma)e_i$ es $a_{i,k}$, y esto es 1 si $k = \sigma(i)$ y 0 en caso contrario. Esto nos dice que $A(\sigma)e_i$ es el vector $e_{\sigma(i)}$, como afirma la proposición. \square

4.4.8. Por otro lado, la regla que asigna a una permutación su matriz de permutación es compatible con las operaciones de composición e invención:

Proposición. *Sea $n \in \mathbb{N}$.*

- (i) *La matriz de permutación $A(\text{id}_n)$ correspondiente a la permutación identidad $\text{id}_n \in S_n$ es la matriz identidad I_n de $M_n(\mathbb{k})$.*
- (ii) *Si σ y τ son permutaciones de $[\![n]\!]$, entonces*

$$A(\sigma \circ \tau) = A(\sigma) \cdot A(\tau).$$

- (iii) *Si σ es una permutación de $[\![n]\!]$, entonces la matriz $A(\sigma)$ es inversible y*

$$A(\sigma)^{-1} = A(\sigma^{-1}) = A(\sigma)^t.$$

Demostración. (i) Sea $A = (a_{i,j})$ la matriz de permutación de $\text{id}_n \in S_n$. De acuerdo a la definición, para cada $i, j \in [\![n]\!]$ el escalar $a_{i,j}$ es igual a 1 si $j = i$ y a 0 en caso contrario: esto significa precisamente que A es a matriz identidad de $M_n(\mathbb{k})$.

(ii) Sean σ y τ permutaciones de $[\![n]\!]$ y sean $A(\sigma) = (a_{i,j})$ y $A(\tau) = (b_{i,j})$ las correspondientes matrices de permutación. Sea $(c_{i,j})$ la matriz producto $A(\sigma) \cdot A(\tau)$. Si $i, j \in [\![n]\!]$, entonces

$$c_{i,j} = \sum_{k=1}^n a_{i,k} b_{k,j}.$$

El escalar $b_{k,j}$ es nulo salvo si $k = \tau(j)$, y en ese caso es igual a 1: vemos así que, de hecho, $c_{i,j} = a_{i,\tau(j)}$, y esto nos dice que $c_{i,j} = 1$ si $i = \sigma(\tau(j))$ y que $c_{i,j} = 0$ en caso contrario. Por supuesto, esto significa que la matriz $(c_{i,j})$ es precisamente la matriz de permutación correspondiente a la permutación $\sigma \circ \tau$, y esto es lo que se afirma en la proposición.

(iii) Sea σ una permutación de $[\![n]\!]$. De acuerdo a las partes (ii) y (i) de la proposición, que ya probamos, tenemos que

$$A(\sigma^{-1}) \cdot A(\sigma) = A(\sigma^{-1} \circ \sigma) = A(\text{id}_n) = I_n,$$

la matriz identidad de $M_n(\mathbb{k})$ y, de manera similar, que $A(\sigma) \cdot A(\sigma^{-1}) = I_n$. Así, $A(\sigma)$ y $A(\sigma^{-1})$ son matrices inversas y, en particular, la matriz $A(\sigma)$ es inversible.

Por otro lado, supongamos que $A(\sigma) = (a_{i,j})$ y sea $(d_{i,j}) = A(\sigma) \cdot A(\sigma)^t$. Si $i, j \in \llbracket n \rrbracket$, entonces

$$d_{i,j} = \sum_{k=1}^n a_{i,k} a_{j,k}$$

Ahora bien: para cada $k \in \llbracket n \rrbracket$ es escalar $a_{j,k}$ es nulo salvo si $j = \sigma(k)$ y $a_{i,k}$ es nulo salvo si $i = \sigma(k)$ y, por lo tanto, hay dos casos: o bien $i = j$ y en ese caso $d_{i,j} = 1$, o bien $i \neq j$ y en ese caso $d_{i,j} = 0$. Vemos de esta forma que $A(\sigma) \cdot A(\sigma)^t = I_n$ y un razonamiento similar muestra que $A(\sigma)^t \cdot A(\sigma) = I_n$: podemos concluir entonces, como queremos, que $A(\sigma)^t = A(\sigma)^{-1}$. \square

4.4.9. Llamamos *signo* de una permutación $\sigma \in S_n$, y escribimos $\text{sgn}(\sigma)$, al valor del determinante de la matriz de permutación asociada a σ , esto es,

$$\text{sgn}(\sigma) := \det(A(\sigma)).$$

Notemos que, de acuerdo a esta definición, el signo de una permutación es un elemento de \mathbb{k} .

4.4.10. Proposición. Sea $n \in \mathbb{N}$.

(i) El signo de la permutación identidad es $\text{sgn}(\text{id}_n) = 1$.

(ii) Si σ y τ son permutaciones de $\llbracket n \rrbracket$, entonces

$$\text{sgn}(\sigma \circ \tau) = \text{sgn}(\sigma) \cdot \text{sgn}(\tau).$$

(iii) Si τ es una transposición de $\llbracket n \rrbracket$, entonces $\text{sgn}(\tau) = -1$.

(iv) Si σ es una permutación de $\llbracket n \rrbracket$, entonces $\text{sgn}(\sigma) \in \{+1, -1\}$ y

$$\text{sgn}(\sigma^{-1}) = \text{sgn}(\sigma).$$

Demostración. Como la matriz de permutación correspondiente a la permutación identidad id de S_n es la matriz identidad I_n , es $\text{sgn}(\text{id}) = \det(I_n) = 1$. Por otro lado, si σ y τ son elementos de S_n , entonces

$$\text{sgn}(\sigma \circ \tau) = \det(A(\sigma \circ \tau)) = \det(A(\sigma) \cdot A(\tau)) = \det(A(\sigma)) \cdot \det(A(\tau)) = \text{sgn}(\sigma) \cdot \text{sgn}(\tau),$$

por la Proposición 4.4.8(ii) y la Proposición 4.3.3(ii).

Sean entonces r y s dos elementos de $\llbracket n \rrbracket$ tales que $r < s$ y sea $\tau = (rs)$ la transposición que los intercambia. Para cada $i \in \llbracket n \rrbracket$ es

$$A(\tau)e_i = e_{\tau(i)} = \begin{cases} e_i & \text{si } i \notin \{r, s\}; \\ e_s & \text{si } i = r; \\ e_r & \text{si } i = s. \end{cases}$$

Nuestras definiciones nos dicen entonces que

$$\begin{aligned}\operatorname{sgn}(\tau) &= \det(A(\tau)) = D_n(A(\tau)e_1, \dots, A(\tau)e_n) = D_n(e_1, \dots, \underbrace{e_s, \dots, e_r}_{r}, \dots, e_n) \\ &= -D_n(e_1, \dots, e_n) = -1,\end{aligned}$$

como queremos.

Finalmente, si σ es una permutación de $\llbracket n \rrbracket$, sabemos que existen $l \in \mathbb{N}$ y transposiciones τ_1, \dots, τ_l de $\llbracket n \rrbracket$ tales que $\sigma = \tau_1 \circ \dots \circ \tau_l$, y por lo que ya probamos es

$$\operatorname{sgn}(\sigma) = \operatorname{sgn}(\tau_1 \circ \dots \circ \tau_l) = \operatorname{sgn}(\tau_1) \cdots \operatorname{sgn}(\tau_l) \in \{+1, -1\},$$

porque cada uno de los factores está en $\{+1, -1\}$. En particular, como

$$\operatorname{sgn}(\sigma) \cdot \operatorname{sgn}(\sigma^{-1}) = \operatorname{sgn}(\sigma \circ \sigma^{-1}) = \operatorname{sgn}(\operatorname{id}_n) = 1,$$

los dos elementos $\operatorname{sgn}(\sigma)$ y $\operatorname{sgn}(\sigma^{-1})$ de $\{+1, -1\}$ tienen producto 1, así que son iguales. \square

4.4.11. Una consecuencia inmediata de la proposición que acabamos de probar es:

Proposición. Sea $n \in \mathbb{N}$ y sea σ una permutación de $\llbracket n \rrbracket$. Si $l \in \mathbb{N}_0$ y τ_1, \dots, τ_l son transposiciones de $\llbracket n \rrbracket$ tales que $\sigma = \tau_1 \circ \dots \circ \tau_l$, entonces $\operatorname{sgn}(\sigma) = (-1)^l$.

Demostración. Sea σ una permutación de $\llbracket n \rrbracket$ y sean $l \in \mathbb{N}$ y τ_1, \dots, τ_l son transposiciones de $\llbracket n \rrbracket$ tales que $\sigma = \tau_1 \circ \dots \circ \tau_l$. De acuerdo a la Proposición 4.4.10(ii), tenemos que

$$\operatorname{sgn}(\sigma) = \operatorname{sgn}(\tau_1) \cdots \operatorname{sgn}(\tau_l) = (-1)^l,$$

ya que todos los factores son iguales a -1 . \square

4.4.12. Un corolario importante de la Proposición 4.4.11 es el siguiente:

Corolario. Sea $n \in \mathbb{N}$. Si k y l son enteros no negativos y τ_1, \dots, τ_k y ρ_1, \dots, ρ_l son transposiciones de S_n tales que $\tau_1 \circ \dots \circ \tau_k = \rho_1 \circ \dots \circ \rho_l$, entonces los números k y l tienen la misma paridad.

Decimos que una permutación σ de $\llbracket n \rrbracket$ es **par** si es igual a la composición de un número par de transposiciones y que es **impar** en caso contrario: el corolario nos dice que esta definición tiene sentido, ya que la paridad del número de factores en una escritura de σ como producto de transposiciones depende solamente de σ y no de la escritura elegida. Así, la permutación 3 5 1 4 2 de $\llbracket 5 \rrbracket$ es par, porque es igual a $(1 3) \circ (2 5)$, mientras que la permutación 2 4 1 3 5 es impar, porque es igual a $(1 2) \circ (2 4) \circ (3 4)$.

Demostración. Elijamos como nuestro cuerpo \mathbb{k} al cuerpo \mathbb{Q} de los números racionales. Si $k, l, \tau_1, \dots, \tau_k$ y ρ_1, \dots, ρ_l son como en el enunciado, la Proposición 4.4.11 nos dice que $(-1)^k = (-1)^l$ en \mathbb{Q} y el corolario sigue inmediatamente de esto, ya que en \mathbb{Q} se tiene que $1 \neq -1$. \square

4.4.13. Para terminar esta sección sobre las permutaciones y sus signos, establezcamos una conexión con las funciones anti-simétricas de la Sección 4.1.

Proposición. *Sea V un espacio vectorial y sea $d \in \mathbb{N}$. Una función d -multilineal $f : V^d \rightarrow \mathbb{k}$ es anti-simétrica si y solamente si para cada elección de d vectores x_1, \dots, x_d en V y cada permutación σ de $\llbracket d \rrbracket$ se tiene que*

$$f(x_{\sigma(1)}, \dots, x_{\sigma(d)}) = \text{sgn}(\sigma) \cdot f(x_1, \dots, x_d). \quad (6)$$

Demostración. Sea $f : V^d \rightarrow \mathbb{k}$ una función d -multilineal, supongamos primero que f satisface la condición del enunciado y mostremos que entonces f es anti-simétrica. Sean $x_1, \dots, x_d \in V$ y sean $i, j \in \llbracket d \rrbracket$ tales que $i < j$. Sabemos que la transposición $\tau = (i \ j)$ tiene signo $\text{sgn}(\tau) = -1$ y, en vista de la definición de τ y la hipótesis, tenemos que

$$\begin{aligned} f(x_1, \dots, \underbrace{x_j, \dots, x_i}_{i}, \dots, x_d) &= f(x_{\tau(1)}, \dots, x_{\tau(d)}) \\ &= \text{sgn}(\tau) \cdot f(x_1, \dots, x_d) \\ &= -f(x_1, \dots, x_d). \end{aligned}$$

Esto nos dice que la función f es anti-simétrica.

Para probar la implicación recíproca, supongamos ahora que f es antisimétrica y sea σ una permutación de $\llbracket d \rrbracket$. De acuerdo a la Proposición 4.4.5, existen $l \in \mathbb{N}_0$ y transposiciones τ_1, \dots, τ_l de $\llbracket d \rrbracket$ tales que $\sigma = \tau_1 \circ \dots \circ \tau_l$. Probaremos que vale la igualdad (6) haciendo inducción con respecto a l . Para empezar, observemos que cuando $l = 0$ la permutación σ es la identidad id_n , su signo es $\text{sgn}(\sigma) = 1$ y la igualdad vale trivialmente, ya que ambos lados del signo igual son idénticos.

Sea $l \in \mathbb{N}$ y supongamos que para toda permutación σ de $\llbracket d \rrbracket$ que puede escribirse como composición de *menos* que l transposiciones vale la igualdad (6) y sea μ una permutación de $\llbracket d \rrbracket$ que es igual a una composición $\tau_1 \circ \dots \circ \tau_l$ de l transposiciones. Sea $\rho = \tau_1 \circ \dots \circ \tau_{l-1}$. Como $\sigma = \rho \circ \tau_l$, es

$$f(x_{\sigma(1)}, \dots, x_{\sigma(d)}) = f(x_{\rho(\tau_l(1))}, \dots, x_{\rho(\tau_l(d))})$$

y si $r, s \in \llbracket d \rrbracket$ son tales que $r < s$ y $\tau_l = (r \ s)$, esto es

$$= f(x_{\rho(1)}, \dots, \underbrace{x_{\rho(s)}, \dots, x_{\rho(r)}}, \dots, x_{\rho(d)}),$$

que, como f es anti-simétrica, intercambiando el argumento r -ésimo y el s -ésimo, es

$$= -f(x_{\rho(1)}, \dots, x_{\rho(d)}).$$

Finalmente, como la permutación ρ puede escribirse como composición de menos que l transposiciones, tenemos que

$$f(x_{\rho(1)}, \dots, x_{\rho(d)}) = \text{sgn}(\rho) \cdot f(x_1, \dots, x_d)$$

y si juntamos todo vemos que

$$f(x_{\sigma(1)}, \dots, x_{\sigma(d)}) = -\text{sgn}(\rho) \cdot f(x_1, \dots, x_d) = \text{sgn}(\sigma) \cdot f(x_1, \dots, x_d),$$

como queremos, ya que $\text{sgn}(\sigma) = \text{sgn}(\rho \circ \tau_l) = \text{sgn}(\rho) \text{sgn}(\tau_l) = -\text{sgn}(\rho)$. □

§5. La fórmula de Leibniz

4.5.1. Estamos por fin en posición dar una fórmula explícita para el determinante de una matriz:

Proposición. Sea $n \in \mathbb{N}$. Si $A = (a_{i,j}) \in M_n(\mathbb{k})$, entonces

$$\det(A) = \sum_{\sigma \in S_n} \text{sgn}(\sigma) a_{\sigma(1),1} \cdots a_{\sigma(n),n}.$$

Esta expresión para el determinante de una matriz se llama la *fórmula de Leibniz*, por *Gottfried Wilhelm Leibniz*. Observemos que la suma tiene $n!$ términos, uno por cada permutación de $\llbracket n \rrbracket$.

Demostración. Sea $A = (a_{i,j})$ un elemento de $M_n(\mathbb{k})$ y sea (e_1, \dots, e_n) la base ordenada estándar de \mathbb{k}^n . Para cada $i \in \llbracket n \rrbracket$ es $Ae_i = a_{1,i}e_1 + \cdots + a_{n,i}e_n$, así que

$$\det(A) = D_n(a_{1,1}e_1 + \cdots + a_{n,1}e_n, \dots, a_{1,n}e_1 + \cdots + a_{n,n}e_n)$$

y, como la función D_n es multilineal, esto es igual a

$$= \sum_{i_1, \dots, i_n \in \llbracket n \rrbracket} a_{i_1,1} \cdots a_{i_n,n} D_n(e_{i_1}, \dots, e_{i_n}), \quad (7)$$

con la suma tomada sobre todas las formas de elegir los n índices i_1, \dots, i_n en el conjunto $\llbracket n \rrbracket$.

Ahora bien: si i_1, \dots, i_n son elementos de $\llbracket n \rrbracket$ que no distintos dos a dos, el hecho de que la función D_n es alternante implica que $D_n(e_{i_1}, \dots, e_{i_n}) = 0$. Esto significa que en la suma (7) basta considerar sólo los términos que corresponden a elecciones de índices i_1, \dots, i_n en $\llbracket n \rrbracket$ que distintos dos a dos, esto es, tales que $\begin{pmatrix} 1 & \cdots & n \\ i_1 & \cdots & i_n \end{pmatrix}$ sea una permutación de $\llbracket n \rrbracket$, y entonces vemos que

$$\det(A) = \sum_{\sigma \in S_n} a_{\sigma(1),1} \cdots a_{\sigma(n),n} D_n(e_{\sigma(1)}, \dots, e_{\sigma(n)}). \quad (8)$$

Para cada $\sigma \in S_n$ sabemos que $e_{\sigma(j)} = A(\sigma)e_j$ si $j \in \llbracket n \rrbracket$, así que

$$D_n(e_{\sigma(1)}, \dots, e_{\sigma(n)}) = D_n(A(\sigma)e_1, \dots, A(\sigma)e_n) = \det(A(\sigma)) = \text{sgn}(\sigma).$$

Usando esto en el lado derecho de la igualdad (8) obtenemos la fórmula que aparece en el enunciado de la proposición. \square

4.5.2. Un corolario casi inmediato de la fórmula de Leibniz es el siguiente:

Corolario. Sea $n \in \mathbb{N}$

- (i) Si $A \in M_n(\mathbb{Q})$ y todas las entradas de A están en \mathbb{Z} , entonces $\det(A) \in \mathbb{Z}$.
- (ii) Si $A \in M_n(\mathbb{k}(X))$ y todas las entradas de A están en $\mathbb{k}[X]$, entonces $\det(A) \in \mathbb{k}[X]$. Más aún, si $d \in \mathbb{N}_0$ y todas las entradas de A son polinomios o nulos o de grado a lo sumo d , entonces su determinante $\det(A)$ es o nulo o un polinomio de grado a lo sumo nd .

Demostración. Sea $A = (a_{i,j})$ una matriz de $M_n(\mathbb{Q})$ que tiene todas sus entradas en \mathbb{Z} . De acuerdo a la proposición que acabamos de probar, es

$$\det(A) = \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) a_{\sigma(1),1} \cdots a_{\sigma(n),n}. \quad (9)$$

Como las entradas de A son enteros, cada uno de los $n!$ términos de esta suma —un producto de $n+1$ enteros— es un entero y, por supuesto, también lo es la suma. Esto prueba la primera afirmación del corolario.

Supongamos ahora que $A = (a_{i,j})$ es una matriz de $\mathbb{k}[X]$, que $d \in \mathbb{N}_0$ y que cada una de sus entradas es o nula o un polinomio de grado a lo sumo d . Cada uno de los $n!$ términos de la suma (9) es entonces un producto de n entradas de A , así que es o nulo o un polinomio de grado a lo sumo nd : se sigue de esto inmediatamente que $\det(A)$ es un polinomio o nulo o de grado a lo sumo nd . \square

4.5.3. Usando la fórmula de Leibniz podemos obtener fácilmente el siguiente resultado, que es importante tanto teóricamente como en la práctica: una matriz tiene el mismo determinante que su transpuesta.

Proposición. Sea $n \in \mathbb{N}$. Para cada $A \in M_n(\mathbb{k})$ se tiene que $\det(A) = \det(A^t)$.

Demostración. Sea $A = (a_{i,j}) \in M_n(\mathbb{k})$. De acuerdo a la Proposición 4.5.1, es

$$\det(A) = \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) a_{\sigma(1),1} \cdots a_{\sigma(n),n}.$$

La función $\sigma \in S_n \mapsto \sigma^{-1} \in S_n$ es una biyección —es su propia función inversa, de hecho— así que podemos reescribir esta suma en la forma

$$\sum_{\tau \in S_n} \operatorname{sgn}(\tau^{-1}) a_{\tau^{-1}(1),1} \cdots a_{\tau^{-1}(n),n}. \quad (10)$$

Si τ es una permutación de $\llbracket n \rrbracket$ tenemos que $\operatorname{sgn}(\tau^{-1}) = \operatorname{sgn}(\tau)$ y el producto

$$a_{\tau^{-1}(1),1} \cdots a_{\tau^{-1}(n),n}$$

coincide con

$$a_{1,\tau(1)} \cdots a_{n,\tau(n)},$$

ya que tiene exactamente los mismos n factores aunque en otro orden. La suma (10) tiene el mismo valor, entonces, que

$$\sum_{\tau \in S_n} \operatorname{sgn}(\tau) a_{1,\tau(1)} \cdots a_{n,\tau(n)}.$$

Si $(b_{i,j})$ es la matriz transpuesta A^t , de manera que $b_{i,j} = a_{j,i}$ para cada elección de $i, j \in \llbracket n \rrbracket$, podemos escribir esto en la forma

$$\sum_{\tau \in S_n} \operatorname{sgn}(\tau) b_{\tau(1),1} \cdots b_{\tau(n),n}$$

y es claro, ahora, que esto es igual a $\det(A^t)$. Esto prueba la proposición. \square

4.5.4. Una consecuencia de la Proposición 4.5.3 es que si en el enunciado de la Proposición 4.3.2 reemplazamos la palabra *columna* por *fila* obtenemos también una afirmación cierta:

Proposición. Sea $n \in \mathbb{N}$. La función $\det : M_n(\mathbb{k}) \rightarrow \mathbb{k}$ tiene las siguientes propiedades:

- (i) $\det(A)$ es una función multilinear alternante de las filas de la matriz A .
- (ii) Si $A \in M_n(\mathbb{k})$ tiene dos filas iguales o, más generalmente, si las columnas de A son linealmente dependientes, entonces $\det(A) = 0$.
- (iii) Si A y B son elementos de $M_n(\mathbb{k})$ y B se obtiene de A intercambiando dos filas, entonces $\det(B) = -\det(A)$.
- (iv) Si A y B son elementos de $M_n(\mathbb{k})$ y B se obtiene de A sumando a una de sus filas una combinación lineal de las demás, entonces $\det(B) = \det(A)$.

Demostración. Estas afirmaciones se deducen de las correspondientes en la Proposición 4.3.2 usando el hecho de que el determinante de una matriz coincide con el de su matriz transpuesta, y la observación evidente de que las filas de aquella son las filas de esta, a menos de transponer. \square

§6. La fórmula de Laplace y la regla de Cramer

4.6.1. Si $m, n \in \mathbb{N}$ y $A \in M_{m,n}(\mathbb{k})$ es una matriz, llamamos *menor* de A a toda matriz que se obtiene de A eliminando algún número de filas y de columnas. Cuando $m, n \geq 2, r \in \llbracket m \rrbracket$ y $s \in \llbracket n \rrbracket$, escribimos $A^{(r,s)}$ al menor de A que se obtiene eliminando la fila r -ésima y la columna s -ésima, y que es un elemento de $M_{m-1,n-1}(\mathbb{k})$.

4.6.2. Proposición. Sea $n \in \mathbb{N}$ y sea $A = (a_{i,j})$ un elemento de $M_n(\mathbb{k})$.

(i) Para cada $i \in \llbracket n \rrbracket$ se tiene que

$$\det(A) = \sum_{l=1}^n (-1)^{i+l} a_{i,l} \cdot \det(A^{(i,l)}).$$

(ii) Para cada $j \in \llbracket n \rrbracket$ se tiene que

$$\det(A) = \sum_{l=1}^n (-1)^{j+l} a_{l,j} \cdot \det(A^{(l,j)}).$$

Estas fórmulas para el determinante de una matriz se llaman *fórmulas de Laplace*, por *Pierre-Simon Laplace*, o *desarrollos de $\det(A)$ a lo largo de la fila i -ésima de A y de la columna j -ésima*, respectivamente.

Demostración. Basta que probemos la primera parte de la proposición, ya que la segunda se obtiene inmediatamente de la primera considerando la matriz transpuesta de A .

Ahora bien, que vale (i) cuando i es 1 es consecuencia de la forma en que definimos la función D_n en la prueba de la Proposición 4.2.4. Si i no es 1, entonces podemos considerar la matriz B que se obtiene de A intercambiando su primera fila y la i -ésima. De la Proposición 4.5.4(iii) sabemos que $\det(B) = -\det(A)$. Por otro lado, para cada $l \in \llbracket n \rrbracket$ el menor $B^{(1,l)}$ es la matriz que se obtiene del menor $A^{(i,l)}$ intercambiando su primera fila con la segunda, luego la segunda con la tercera y así hasta intercambiar la $(i-2)$ -ésima por la $(i-1)$ -ésima: estos son $i-2$ intercambios, así que $\det(B^{(1,l)}) = (-1)^i \det(A^{(i,l)})$. Teniendo todo esto en cuenta, concluimos que

$$\begin{aligned} \det(A) &= -\det(B) = -\sum_{l=1}^n (-1)^{l+1} b_{1,l} \cdot \det(B^{(1,l)}) = -\sum_{l=1}^n (-1)^{l+i+1} a_{i,l} \cdot \det(A^{(1,l)}) \\ &= \sum_{l=1}^n (-1)^{l+i} a_{i,l} \cdot \det(A^{(1,l)}), \end{aligned}$$

como queremos. \square

4.6.3. Ejemplo. Desarrollando por cada una de las tres filas de una matriz de 3×3 obtenemos las siguientes expresiones:

$$\begin{aligned} \begin{vmatrix} a_{1,1} & a_{1,2} & a_{1,3} \\ a_{2,1} & a_{2,2} & a_{2,3} \\ a_{3,1} & a_{3,2} & a_{3,3} \end{vmatrix} &= a_{1,1} \begin{vmatrix} a_{2,2} & a_{2,3} \\ a_{3,2} & a_{3,3} \end{vmatrix} + a_{1,2} \begin{vmatrix} a_{2,1} & a_{2,3} \\ a_{3,1} & a_{3,3} \end{vmatrix} - a_{1,3} \begin{vmatrix} a_{2,1} & a_{2,2} \\ a_{3,1} & a_{3,2} \end{vmatrix} \\ &= a_{2,1} \begin{vmatrix} a_{1,2} & a_{1,3} \\ a_{3,2} & a_{3,3} \end{vmatrix} + a_{2,2} \begin{vmatrix} a_{1,1} & a_{1,3} \\ a_{3,1} & a_{3,3} \end{vmatrix} - a_{2,3} \begin{vmatrix} a_{1,1} & a_{1,2} \\ a_{3,1} & a_{3,2} \end{vmatrix} \\ &= a_{3,1} \begin{vmatrix} a_{1,2} & a_{1,3} \\ a_{2,2} & a_{2,3} \end{vmatrix} + a_{3,2} \begin{vmatrix} a_{1,1} & a_{1,3} \\ a_{2,1} & a_{2,3} \end{vmatrix} - a_{3,3} \begin{vmatrix} a_{1,1} & a_{1,2} \\ a_{2,1} & a_{2,2} \end{vmatrix}. \end{aligned}$$

Por otro lado, desarrollando el determinante por la segunda columna vemos que

$$\begin{vmatrix} a_{1,1} & a_{1,2} & a_{1,3} & a_{1,4} \\ a_{2,1} & a_{2,2} & a_{2,3} & a_{2,4} \\ a_{3,1} & a_{3,2} & a_{3,3} & a_{3,4} \\ a_{4,1} & a_{4,2} & a_{4,3} & a_{4,4} \end{vmatrix} = -a_{1,2} \begin{vmatrix} a_{2,1} & a_{2,3} & a_{2,4} \\ a_{3,1} & a_{3,3} & a_{3,4} \\ a_{4,1} & a_{4,3} & a_{4,4} \end{vmatrix} + a_{2,2} \begin{vmatrix} a_{1,1} & a_{1,3} & a_{1,4} \\ a_{3,1} & a_{3,3} & a_{3,4} \\ a_{4,1} & a_{4,3} & a_{4,4} \end{vmatrix} \\ -a_{3,2} \begin{vmatrix} a_{1,1} & a_{1,3} & a_{1,4} \\ a_{2,1} & a_{2,3} & a_{2,4} \\ a_{4,1} & a_{4,3} & a_{4,4} \end{vmatrix} + a_{4,2} \begin{vmatrix} a_{1,1} & a_{1,3} & a_{1,4} \\ a_{2,1} & a_{2,3} & a_{2,4} \\ a_{3,1} & a_{3,3} & a_{3,4} \end{vmatrix}$$

Notemos que los signos de los sumandos en esos desarrollos van alternándose y que el primer es $+ o -$ de acuerdo a que el desarrollo sea por una columna o fila de índice impar o par. \diamond

4.6.4. Como consecuencia de las fórmulas de Laplace y la alternancia del determinante, obtenemos el siguiente corolario que usaremos para probar la Proposición 4.6.5:

Corolario. Sea $n \in \mathbb{N}$ y sea $A = (a_{i,j})$ un elemento de $M_n(\mathbb{k})$.

(i) Si $i, i' \in \llbracket n \rrbracket$ son distintos, entonces

$$\sum_{l=1}^n (-1)^{i+l} a_{i,l} \cdot \det(A^{(i',l)}) = 0.$$

(ii) Si $j, j' \in \llbracket n \rrbracket$ son distintos, entonces

$$\sum_{l=1}^n (-1)^{j+l} a_{l,j} \cdot \det(A^{(l,j')}) = 0.$$

Demostración. Otra vez, gracias a la Proposición 4.5.3, es suficiente que probemos la primera parte. Sean i e i' elementos distintos de $\llbracket n \rrbracket$ y sea $B = (b_{i,j})$ la matriz que se obtiene de la matriz A reemplazando su fila i' -ésima por una copia de su fila i -ésima. Como B tiene dos filas iguales, sabemos que $\det(B) = 0$ y entonces el desarrollo de Laplace para el determinante de B a lo largo de la fila i' -ésima de esta matriz nos dice que

$$\sum_{l=1}^n (-1)^{i+l} b_{i',l} \cdot \det(B^{(i',l)}) = 0. \quad (11)$$

Si $l \in \llbracket n \rrbracket$, es $b_{i',l} = a_{i,l}$ y el menor $B^{(i',l)}$ coincide con el menor $A^{(i',l)}$: la igualdad (11) implica entonces que

$$\sum_{l=1}^n (-1)^{i+l} a_{i,l} \cdot \det(A^{(i',l)}) = 0,$$

como afirma la proposición. \square

4.6.5. La Proposición 4.3.3(iii) nos dice que una condición necesaria para que una matriz sea inversible es que su determinante sea no nulo. Podemos probar ahora que esta condición es también suficiente:

Proposición. Sea $n \in \mathbb{N}$, sea $A = (a_{i,j}) \in M_n(\mathbb{k})$ y consideremos la matriz $\text{adj } A = (b_{i,j})$ tal que $b_{i,j} = (-1)^{i+j} \det(A^{(j,i)})$ para cada $i, j \in \llbracket n \rrbracket$.

(i) Se tiene que

$$A \cdot \text{adj } A = \text{adj } A \cdot A = \det(A) \cdot I_n.$$

(ii) La matriz A es inversible si y solamente si su determinante $\det(A)$ es no nulo, y en ese caso vale que

$$A^{-1} = \frac{1}{\det(A)} \text{adj } A. \quad (12)$$

La matriz $\text{adj } A$ aquí descripta es la **matriz de cofactores** de A .

Demostración. Si $i, j \in \llbracket n \rrbracket$, la entrada (i, j) -ésima del producto $A \cdot \text{adj } A$ es

$$\sum_{l=1}^n a_{i,l} b_{l,j} = \sum_{l=1}^n a_{i,l} (-1)^{l+j} \det(A^{(j,l)})$$

y de acuerdo a las Proposiciones 4.6.2(i) y 4.6.4(i), esto es igual a $\det(A)$ si $i = j$ y a 0 si $i \neq j$. Esto significa que $A \cdot \text{adj } A = \det(A) \cdot I_n$. De la misma forma, a partir de las Proposiciones 4.6.2(ii) y 4.6.4(ii) podemos ver que $\text{adj } A \cdot A = \det(A) \cdot I_n$. Esto prueba la afirmación (i) de la proposición.

Ya sabemos de la Proposición 4.3.3(iii) que es necesario para que la matriz A sea inversible que su determinante sea no nulo. Recíprocamente, si tenemos que $\det(A) \neq 0$, entonces en la igualdad de (i) que acabamos de probar podemos dividir por el escalar $\det(A)$ para ver que

$$A \cdot \frac{\text{adj } A}{\det(A)} = \frac{\text{adj } A}{\det(A)} \cdot A = I_n$$

y concluir que A es inversible y que su matriz inversa es la descripta en (12). □

4.6.6. Ejemplo. Cuando tenemos una matriz inversible de 2×2 , los menores que aparecen en la fórmula para la matriz inversa que nos da esta proposición son de 1×1 , así que sus determinantes se calculan inmediatamente: obtenemos así para la inversa de una matriz 2×2 inversible la expresión

$$\begin{pmatrix} a_{1,1} & a_{1,2} \\ a_{2,1} & a_{2,2} \end{pmatrix}^{-1} = \frac{1}{a_{1,1}a_{2,2} - a_{1,2}a_{2,1}} \begin{pmatrix} a_{2,2} & -a_{1,2} \\ -a_{2,1} & a_{1,1} \end{pmatrix}.$$

Por otro lado, si

$$A = \begin{pmatrix} a_{1,1} & a_{1,2} & a_{1,3} \\ a_{2,1} & a_{2,2} & a_{2,3} \\ a_{3,1} & a_{3,2} & a_{3,3} \end{pmatrix}$$

es una matriz inversible, entonces la matriz inversa de A es

$$A^{-1} = \frac{1}{\det(A)} \begin{pmatrix} \left| \begin{array}{cc} a_{2,2} & a_{2,3} \\ a_{3,2} & a_{3,3} \end{array} \right| - \left| \begin{array}{cc} a_{1,2} & a_{1,3} \\ a_{3,2} & a_{3,3} \end{array} \right| & \left| \begin{array}{cc} a_{1,2} & a_{1,3} \\ a_{2,2} & a_{2,3} \end{array} \right| \\ - \left| \begin{array}{cc} a_{2,1} & a_{2,3} \\ a_{3,1} & a_{3,3} \end{array} \right| & \left| \begin{array}{cc} a_{1,1} & a_{1,3} \\ a_{3,1} & a_{3,3} \end{array} \right| - \left| \begin{array}{cc} a_{1,1} & a_{1,3} \\ a_{2,1} & a_{2,3} \end{array} \right| \\ \left| \begin{array}{cc} a_{2,1} & a_{2,2} \\ a_{3,1} & a_{3,2} \end{array} \right| - \left| \begin{array}{cc} a_{1,1} & a_{1,2} \\ a_{3,1} & a_{3,2} \end{array} \right| & \left| \begin{array}{cc} a_{1,1} & a_{1,2} \\ a_{2,1} & a_{2,2} \end{array} \right| \end{pmatrix}.$$

Casi nunca calculamos la matriz inversa de una matriz usando la expresión de la Proposición 4.6.5: en cuanto el tamaño de la matriz es mayor que 2 hay formas mucho más eficientes de hacerlo. \diamond

4.6.7. Usando la expresión para la inversa de una matriz inversible que nos da la Proposición 4.6.5 podemos obtener una expresión cerrada para solución de un sistema de ecuaciones con la misma cantidad de incógnitas que de ecuaciones con matriz de coeficientes inversible. Llamamos a este resultado la *regla de Cramer*, por *Daniel Cramer*.

Proposición. Sean $n \in \mathbb{N}$, $A = (a_{i,j}) \in M_n(\mathbb{k})$ y $b = (b_i) \in \mathbb{k}^n$. Si la matriz A es inversible, de manera que $\det(A) \neq 0$, entonces existe exactamente un vector $x = (x_i) \in \mathbb{k}^n$ tal que $Ax = b$ y para cada $i \in [n]$ la componente i -ésima de x es

$$x_i = \frac{\left| \begin{array}{ccccccc} a_{1,1} & \cdots & a_{1,i-1} & b_1 & a_{1,i+1} & \cdots & a_{1,n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{n,1} & \cdots & a_{n,i-1} & b_n & a_{n,i+1} & \cdots & a_{n,n} \end{array} \right|}{\det(A)}.$$

El numerador de este cociente es el determinante de la matriz que se obtiene de A reemplazando su columna i -ésima por el vector b .

Demostración. Supongamos que la matriz A es inversible. La función lineal $f : x \in \mathbb{k}^n \mapsto Ax \in \mathbb{k}^n$ es entonces un isomorfismo y su inversa es la función $g : x \in \mathbb{k}^n \mapsto A^{-1}x \in \mathbb{k}^n$. En particular, como f es biyectiva, existe exactamente un elemento x de \mathbb{k}^n tal que $Ax = f(x) = b$ y es $x = g(b) = A^{-1}b$. En vista de la Proposición 4.6.5(ii), esto significa que

$$x = \frac{1}{\det(A)} \operatorname{adj} A \cdot b.$$

Si $i \in [n]$ y escribimos $\operatorname{adj} A = (c_{i,j})$, la componente i -ésima de este vector es

$$\frac{1}{\det(A)} \sum_{l=1}^n c_{i,l} b_l = \sum_{l=1}^n (-1)^{i+l} b_l \det(A^{(l,i)})$$

y la suma que aparece aquí es el desarrollo de Laplace a lo largo de la columna i -ésima del determinante de la matriz que se obtiene de A reemplazando la columna i -ésima por el vector b . Esto es precisamente lo que afirma la proposición. \square

§7. El rango de una matriz

4.7.1. El determinante nos da un criterio muy compacto para decidir la independencia lineal de n vectores de \mathbb{k}^n :

Proposición. *Sea $n \in \mathbb{N}$ y sean $x_1, \dots, x_n \in \mathbb{k}^n$. Los vectores x_1, \dots, x_n son linealmente independientes si y solamente si la matriz de $M_n(\mathbb{k})$ que los tiene por columnas tiene determinante no nulo.*

Demostración. Sea $A \in M_n(\mathbb{k})$ la matriz que tiene a los vectores x_1, \dots, x_n por columnas. Sabemos que si los vectores son linealmente dependientes, entonces el determinante de A es nulo. Recíprocamente, si son linealmente independientes, entonces la función $f : x \in \mathbb{k}^n \mapsto Ax \in \mathbb{k}^n$ es inversible y, de acuerdo a la Proposición 2.6.9, la matriz $[f]_{\mathcal{B}}^{\mathcal{B}}$ de f con respecto a la base estándar \mathcal{B} de \mathbb{k}^n , que coincide con A , como vimos en el Ejemplo 2.6.2(a), es inversible. En vista de la Proposición 4.6.5(ii), entonces, es $\det(A) \neq 0$. \square

4.7.2. Usando este resultado podemos obtener una nueva descripción del rango de una matriz, esta vez en términos del tamaño de sus menores cuadrados no singulares:

Proposición. *Sean $m, n \in \mathbb{N}$ y sea $A \in M_{m,n}(\mathbb{k})$. El rango de A es igual al tamaño máximo de los menores cuadrados de A que tienen determinante no nulo.*

Demostración. Sea r el rango de la matriz A y sea s el tamaño máximo de un menor cuadrado de A con determinante no nulo. Como r es la dimensión del subespacio de \mathbb{k}^m generado por las columnas de A , sabemos que hay r columnas en A que son linealmente independientes: sea B el menor de A que se obtiene eliminando todas las otras $n - r$ columnas, de manera que $B \in M_{m,r}(\mathbb{k})$. Por construcción, las columnas de B son linealmente independientes, así que el rango de B es r . De acuerdo a la Proposición 3.4.7, el rango por filas de B también es r y, en consecuencia, hay r filas en B que son linealmente independientes. Sea C el menor de B que se obtiene eliminando las otras $m - r$ filas. Es claro que C es también un menor de la matriz A con la que empezamos, se trata de un menor cuadrado de tamaño r y, como sus filas son linealmente independientes, su determinante es no nulo: la elección de s , en consecuencia, implica que es $s \geq r$.

Por otro lado, si D es un menor cuadrado de A de tamaño *mayor* que r , entonces las filas de A que tocan a D son linealmente dependientes —ya que A no posee $r + 1$ columnas linealmente independientes— y entonces las columnas de D son también linealmente dependientes: esto implica que $\det(D) = 0$. Así, todo menor cuadrado de A de tamaño mayor que r tiene determinante nulo y, por lo tanto, $s \leq r$. \square

§8. Tres determinantes

Matrices triangulares

4.8.1. Decimos que una matriz $A = (a_{i,j}) \in M_n(\mathbb{k})$ es *triangular superior* si cada vez que $i, j \in \llbracket n \rrbracket$ son tales que $i > j$ se tiene que $a_{i,j} = 0$, esto es, si todas las entradas de la matriz que están por debajo de la diagonal principal son nulas. De forma simétrica, decimos que A es *triangular inferior* si cada vez que $i, j \in \llbracket n \rrbracket$ son tales que $i < j$ se tiene que $a_{i,j} = 0$

Proposición. *Sea $n \in \mathbb{N}$. Si $A = (a_{i,j}) \in M_n(\mathbb{k})$ es una matriz triangular superior o triangular inferior, entonces el determinante de A es el producto de las entradas que aparecen en su diagonal, esto es,*

$$\det(A) = a_{1,1} \cdots a_{n,n}.$$

Demostración. Basta que consideremos el caso en que A es triangular superior, ya que si A es triangular inferior entonces la matriz transpuesta A^t es triangular superior y, como sabemos, $\det(A) = \det(A^t)$. Supongamos entonces que la matriz A es triangular superior. Según la Proposición 4.5.1, es

$$\det(A) = \sum_{\sigma \in S_n} \text{sgn}(\sigma) a_{\sigma(1),1} \cdots a_{\sigma(n),n}. \quad (13)$$

Si σ es una permutación de $\llbracket n \rrbracket$, entonces el producto $a_{\sigma(1),1} \cdots a_{\sigma(n),n}$ es nulo si cualquiera de sus n factores es nulo, y esto ocurre si existe $i \in \llbracket n \rrbracket$ tal que $\sigma(i) > i$. Vemos así que la suma (13) no cambia si solamente sumamos los términos que corresponden a permutaciones $\sigma \in S_n$ tales que $\sigma(i) \leq i$ para cada $i \in \llbracket n \rrbracket$. Ahora bien: hay exactamente una permutación que satisface esta condición, la permutación identidad id_n , y, en consecuencia, tenemos que

$$\det(A) = \text{sgn}(\text{id}_n) a_{\text{id}_n(1),1} \cdots a_{\text{id}_n(n),n} = a_{1,1} \cdots a_{n,n},$$

como afirma la proposición. \square

4.8.2. Más generalmente, podemos describir el determinante de una matriz triangular por bloques:

Proposición. *Sean $n, r \in \mathbb{N}$, sean $n_1, \dots, n_r \in \mathbb{N}$ tales que $n = n_1 + \cdots + n_r$ y supongamos que para cada $i, j \in \llbracket r \rrbracket$ con $i \leq j$ tenemos una matriz $A_{i,j} \in M_{n_i, n_j}(\mathbb{k})$. Si A es la matriz de bloques*

$$\begin{pmatrix} A_{1,1} & A_{1,2} & \cdots & A_{1,r-1} & A_{1,r} \\ 0 & A_{2,2} & \cdots & A_{2,r-1} & A_{2,r} \\ \vdots & \ddots & \vdots & \vdots & \vdots \\ \vdots & & A_{r-1,r-1} & A_{r-1,r} & \\ 0 & \cdots & 0 & 0 & A_{r,r} \end{pmatrix},$$

entonces el determinante de A es

$$\det(A) = \det(A_{1,1}) \cdots \det(A_{r,r}).$$

Demostración. Como podemos hacer una inducción evidente con respecto a r , es suficiente que consideremos el caso en que $r = 2$. Sean entonces $n, n_1, n_2 \in \mathbb{N}$ tales que $n = n_1 + n_2$, sean $A_{1,1} \in M_{n_1, n_1}(\mathbb{k})$, $A_{1,2} \in M_{n_1, n_2}(\mathbb{k})$ y $A_{2,2} \in M_{n_2, n_2}(\mathbb{k})$, consideremos la matriz de bloques

$$A = \begin{pmatrix} A_{1,1} & A_{1,2} \\ 0 & A_{2,2} \end{pmatrix}.$$

y mostremos que

$$\det(A) = \det(A_{1,1}) \cdot \det(A_{2,2}). \quad (14)$$

Si la matriz $A_{1,1}$ tiene determinante nulo, sus columnas son linealmente dependientes y entonces claramente las primeras n_1 columnas de la matriz A son linealmente dependientes: vemos que en este caso $\det(A) = 0$ y que, en consecuencia, vale la igualdad (14).

Supongamos ahora que $A_{1,1}$ tiene determinante no nulo, de manera que es inversible. Vale entonces que

$$A = \begin{pmatrix} A_{1,1} & 0 \\ 0 & I_{n_2} \end{pmatrix} \cdot \begin{pmatrix} I_{n_1} & A_{1,1}^{-1}A_{1,2} \\ 0 & A_{2,2} \end{pmatrix}$$

y, en consecuencia, que

$$\det(A) = \det\begin{pmatrix} A_{1,1} & 0 \\ 0 & I_{n_2} \end{pmatrix} \cdot \det\begin{pmatrix} I_{n_1} & A_{1,1}^{-1}A_{1,2} \\ 0 & A_{2,2} \end{pmatrix}. \quad (15)$$

Ahora bien, desarrollando el determinante por la última fila vemos inmediatamente que

$$\det\begin{pmatrix} A_{1,1} & 0 \\ 0 & I_{n_2} \end{pmatrix} = \det\begin{pmatrix} A_{1,1} & 0 \\ 0 & I_{n_2-1} \end{pmatrix}$$

y entonces una inducción evidente muestra que, de hecho,

$$\det\begin{pmatrix} A_{1,1} & 0 \\ 0 & I_{n_2} \end{pmatrix} = \det(A_{1,1}).$$

De manera similar, desarrollando el determinante por la primera columna y haciendo inducción en n_1 vemos que

$$\det\begin{pmatrix} I_{n_1} & A_{1,1}^{-1}A_{1,2} \\ 0 & A_{2,2} \end{pmatrix} = \det(A_{2,2})$$

Estas últimas dos igualdades y (15) implican que vale la igualdad (14), como queremos. \square

Matrices compañeras

4.8.3. Si $n \in \mathbb{N}$ y $p = c_0 + c_1X + \dots + c_{n-1}X^{-1} + X^n \in \mathbb{k}[X]$ es un polinomio mónico de grado n , la *matriz compañera* de p es la matriz

$$C(p) = \begin{pmatrix} 0 & 0 & \dots & 0 & -c_0 \\ 1 & 0 & \dots & 0 & -c_1 \\ 0 & 1 & \dots & 0 & -c_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & 1 & -c_{n-1} \end{pmatrix} \in M_n(\mathbb{k}),$$

de manera que si $C(p) = (a_{i,j})$, entonces para cada $i, j \in \llbracket n \rrbracket$ es

$$a_{i,j} = \begin{cases} 1, & \text{si } i = j + 1; \\ -c_{i-1}, & \text{si } j = n; \\ 0, & \text{en cualquier otro caso.} \end{cases}$$

Esta matriz $C(p)$ tiene sus entradas en \mathbb{k} , y podemos considerar la matriz

$$X \cdot I_n - C(p)$$

que tiene ahora sus entradas en el cuerpo $\mathbb{k}(X)$.

4.8.4. Ejemplo. Si $p = c_0 + c_1X + c_2X^2 + X^3$ es un polinomio cúbico mónico, entonces la matriz compañera de p es

$$C(p) = \begin{pmatrix} 0 & 0 & -c_0 \\ 1 & 0 & -c_1 \\ 0 & 1 & -c_2 \end{pmatrix}$$

y la matriz

$$X \cdot I_3 - C(p) = \begin{pmatrix} X & 0 & c_0 \\ -1 & X & c_1 \\ 0 & -1 & X + c_2 \end{pmatrix}$$

tiene determinante $\det(X \cdot I_3 - C(p)) = c_0 + c_1X + c_2X^2 + X^3 = p$. ◇

4.8.5. El resultado de este ejemplo es completamente general:

Proposición. Sea $n \in \mathbb{N}$ y sea $p = c_0 + c_1X + \dots + c_{n-1}X^{-1} + X^n \in \mathbb{k}[X]$ un polinomio mónico de grado n . El determinante de la matriz $X \cdot I_n - C(p) \in M_n(\mathbb{k}(X))$ es

$$\det(X \cdot I_n - C(p)) = p.$$

Demostración. Demostraremos la proposición haciendo inducción en n , observando que cuando $n = 1$ el resultado es inmediato. Supongamos entonces que $n \geq 2$. Tenemos que calcular el determinante de la matriz

$$X \cdot I_n - C(p) = \begin{pmatrix} X & 0 & \dots & 0 & c_0 \\ -1 & X & \dots & 0 & c_1 \\ 0 & -1 & \dots & 0 & c_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & -1 & X + c_{n-1} \end{pmatrix}.$$

Desarrollándolo a lo largo de la primera fila de la matriz, en la que hay solamente dos entradas no nulas, vemos que

$$\det(X \cdot I_n - C(p)) = X \begin{vmatrix} X & 0 & \dots & 0 & c_1 \\ -1 & X & \dots & 0 & c_2 \\ 0 & -1 & \dots & 0 & c_3 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & -1 & X + c_{n-1} \end{vmatrix} + (-1)^{n+1} c_0 \begin{vmatrix} -1 & X & \dots & 0 \\ 0 & -1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & -1 \end{vmatrix}. \quad (16)$$

El primer determinante que aparece en el último miembro de esta igualdad es el de la matriz $X \cdot I_{n-1} - C(q) \in M_n(\mathbb{k}(X))$, con q el polinomio

$$c_1 + c_2 X + \dots + c_{n-1} X^{n-2} + X^{n-1} \in \mathbb{k}(X),$$

que es mónico y de grado $n - 1$, así que la hipótesis inductiva nos dice que es igual a q . El segundo determinante, por otro lado, es el de una matriz triangular inferior y de la Proposición 4.8.1 sabemos que vale $(-1)^{n-1}$. Volviendo con todo esto a la igualdad (16) vemos que

$$\det(X \cdot I_n - C(p)) = Xq + c_0 = p,$$

como queremos. \square

Matrices de Vandermonde

4.8.6. Si $n \in \mathbb{N}$ y $\alpha_1, \dots, \alpha_n \in \mathbb{k}$ son escalares, la **matriz de Vandermonde** para $\alpha_1, \dots, \alpha_n$ es la matriz $V(\alpha_1, \dots, \alpha_n) = (a_{i,j}) \in M_n(\mathbb{k})$ que tiene $a_{i,j} = \alpha_j^{i-1}$ para cada $i, j \in \llbracket n \rrbracket$.

$$V(\alpha_1, \dots, \alpha_n) = \begin{pmatrix} 1 & 1 & \dots & 1 \\ \alpha_1 & \alpha_2 & \dots & \alpha_n \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_1^{n-1} & \alpha_2^{n-1} & \dots & \alpha_n^{n-1} \end{pmatrix}$$

El nombre recuerda a *Alexandre-Théophile Vandermonde*, a pesar de que no hay ningún registro de que él lo haya considerado; se puede ver, por ejemplo, el artículo [Yca13] para una discusión de esto.

Vandermonde, sin embargo, es considerado el fundador de la teoría moderna de los determinantes: por ejemplo, es el primero en haber observado el efecto que tiene sobre el determinante de una matriz intercambiar dos de las columnas de ésta y que el determinante es en consecuencia nulo cuando dos de ellas son iguales.

4.8.7. Proposición. *Sea $n \in \mathbb{N}$. Si $\alpha_1, \dots, \alpha_n \in \mathbb{k}$, entonces*

$$\det(V(\alpha_1, \dots, \alpha_n)) = \prod_{1 \leq i < j \leq n} (\alpha_j - \alpha_i) \quad (17)$$

Demostración. Hagamos inducción en n , observando que si $n = 1$ no hay nada que probar.

Escribamos V en lugar de $V(\alpha_1, \dots, \alpha_n)$, por simplicidad. Es claro que si existen i y j en $\llbracket n \rrbracket$ distintos tales que $\alpha_i = \alpha_j$, entonces $\det(V) = 0$, ya que en ese caso la matriz tiene dos columnas iguales, y vale evidentemente la igualdad (17). Queda entonces considerar el caso en el que los escalares $\alpha_1, \dots, \alpha_n$ son distintos dos a dos.

Si todos los escalares $\alpha_1, \dots, \alpha_n$ son no nulos pongamos $k = 1$ y si no sea $k \in \llbracket n \rrbracket$ tal que $\alpha_k = 0$. Según la Proposición 4.6.2(ii), desarrollando el determinante de V a lo largo de su columna k -ésima, es

$$\det(V) = \sum_{l=1}^n (-1)^{l+k} \alpha_k^{l-1} \det(V^{(l,k)}).$$

Consideremos el polinomio

$$p = \sum_{l=1}^n (-1)^{l+k} X^{l-1} \det(V^{(l,k)}) \in \mathbb{k}[X].$$

El grado de p es a lo sumo $n - 1$ y el coeficiente de X^{n-1} en p es $(-1)^{n+k} \det(V^{(n,k)})$. Como $V^{(n,k)}$ es la matriz de Vandermonde $V(\alpha_1, \dots, \widehat{\alpha_k}, \dots, \alpha_n)$, usando la hipótesis inductiva tenemos que

$$\det(V^{(n,k)}) = \det(V(\alpha_1, \dots, \widehat{\alpha_k}, \dots, \alpha_n)) = \prod_{\substack{1 \leq i < j \leq n \\ i, j \neq k}} (\alpha_j - \alpha_i) \neq 0.$$

Concluimos de esta forma que el polinomio p tiene grado exactamente $n - 1$.

Si $i \in \{1, \dots, \widehat{k}, \dots, n\}$, entonces

$$p(\alpha_i) = \sum_{l=1}^n (-1)^{l+k} \alpha_i^{l-1} \det(V^{(l,1)})$$

y esto es precisamente el valor del determinante de la matriz

$$V(\alpha_1, \dots, \underbrace{\alpha_i, \dots, \alpha_n}_k),$$

que tiene dos columnas iguales, la i -ésima y la k -ésima: vemos así que $p(\alpha_i) = 0$. El polinomio p , que tiene grado $n - 1$, tiene entonces a los $n - 1$ escalares $\alpha_1, \dots, \widehat{\alpha_k}, \dots, \alpha_n$, que son distintos dos a

dos, como raíces y, en consecuencia, existe un escalar $\beta \in \mathbb{k}$ tal que

$$p = \beta(X - \alpha_1) \cdots \widehat{(X - \alpha_k)} \cdots (X - \alpha_n).$$

De esta igualdad se sigue, en particular, que

$$p(0) = (-1)^{n-1} \beta \alpha_1 \cdots \widehat{\alpha_k} \cdots \alpha_n. \quad (18)$$

Por otro lado, de la definición de p es inmediato que

$$p(0) = (-1)^{k+1} \det(V^{(1,k)}) = (-1)^{k+1} \begin{vmatrix} \alpha_1 & \cdots & \alpha_{k-1} & \alpha_{k+1} & \cdots & \alpha_n \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \alpha_1^{n-1} & \cdots & \alpha_{k-1}^{n-1} & \alpha_{k+1}^{n-1} & \cdots & \alpha_n^{n-1} \end{vmatrix} \quad (19)$$

y, usando la homogeneidad del determinante de una matriz con respecto a las columnas de ésta, este último determinante es

$$\begin{aligned} \begin{vmatrix} \alpha_1 & \cdots & \alpha_{k-1} & \alpha_{k+1} & \cdots & \alpha_n \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \alpha_1^{n-1} & \cdots & \alpha_{k-1}^{n-1} & \alpha_{k+1}^{n-1} & \cdots & \alpha_n^{n-1} \end{vmatrix} &= \alpha_1 \cdots \widehat{\alpha_k} \cdots \alpha_n \begin{vmatrix} 1 & \cdots & 1 & 1 & \cdots & 1 \\ \alpha_1 & \cdots & \alpha_{k-1} & \alpha_{k+1} & \cdots & \alpha_n \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \alpha_1^{n-2} & \cdots & \alpha_{k-1}^{n-2} & \alpha_{k+1}^{n-2} & \cdots & \alpha_n^{n-2} \end{vmatrix} \\ &= \alpha_1 \cdots \widehat{\alpha_k} \cdots \alpha_n \det(V(\alpha_1, \dots, \widehat{\alpha_k}, \dots, \alpha_n)). \end{aligned}$$

Comparando las expresiones (18) y (19) para $p(0)$ concluimos entonces que

$$(-1)^{n-1} \beta \alpha_1 \cdots \widehat{\alpha_k} \cdots \alpha_n = (-1)^k \alpha_1 \cdots \widehat{\alpha_k} \cdots \alpha_n \det(V(\alpha_1, \dots, \widehat{\alpha_k}, \dots, \alpha_n))$$

Como $\alpha_1 \cdots \widehat{\alpha_k} \cdots \alpha_n \neq 0$ por la forma en que elegimos el índice k , esto y la hipótesis inductiva nos dicen que

$$\beta = (-1)^{n-k-1} \det(V(\alpha_1, \dots, \widehat{\alpha_k}, \dots, \alpha_n)) = (-1)^{n-k-1} \prod_{\substack{1 \leq i < j \leq n \\ i, j \neq k}} (\alpha_j - \alpha_i)$$

y, en definitiva, que

$$p = (-1)^{n-k-1} \prod_{\substack{1 \leq i < j \leq n \\ i, j \neq k}} (\alpha_j - \alpha_i) \cdot (X - \alpha_1) \cdots \widehat{(X - \alpha_k)} \cdots (X - \alpha_n).$$

En particular, evaluando esto en α_k vemos que

$$p(\alpha_k) = (-1)^{n-k-1} \prod_{\substack{1 \leq i < j \leq n \\ i, j \neq k}} (\alpha_j - \alpha_i) \cdot (\alpha_k - \alpha_1) \cdots \widehat{(\alpha_k - \alpha_k)} \cdots (\alpha_k - \alpha_n) = \prod_{\substack{1 \leq i < j \leq n \\ i, j \neq k}} (\alpha_j - \alpha_i).$$

Esto completa la inducción, ya que $p(\alpha_k) = \det(V(\alpha_1, \dots, \alpha_k))$. \square

4.8.8. Una aplicación casi inmediata del cálculo del determinante de las matrices de Vandermonde es la existencia de polinomios interpoladores:

Proposición. Sea $n \in \mathbb{N}$ y sean $\alpha_1, \dots, \alpha_n$ elementos distintos dos a dos de \mathbb{k} . Si $\beta_1, \dots, \beta_n \in \mathbb{k}$ son n escalares, entonces existe exactamente un polinomio $p \in \mathbb{k}[X]$ de grado menor que n tal que $p(\alpha_i) = \beta_i$ para cada $i \in [\![n]\!]$.

Demostración. Sea $V = \mathbb{k}[X]_{<n}$ el espacio vectorial de los polinomios de $\mathbb{k}[X]$ de grado menor que n . Sabemos que $\dim V = n$ y que, de hecho, $\mathcal{B} = (1, X, \dots, X^{n-1})$ es una base de V . Es inmediato verificar que la función

$$f : p \in V \mapsto (p(\alpha_1), \dots, p(\alpha_n))^t \in \mathbb{k}^n$$

es lineal. Más aún, si $\mathcal{B}' = (e_1, \dots, e_n)$ es la base ordenada estándar de \mathbb{k}^n , claramente tenemos que $f(X^i) = \alpha_1^i e_1 + \dots + \alpha_n^i e_n$ cualquiera sea $i \in [\![n-1]\!]$, así que la matriz $[f]_{\mathcal{B}'}^{\mathcal{B}}$ de f con respecto a las bases \mathcal{B} y \mathcal{B}' de V y de \mathbb{k}^n es precisamente la matriz transpuesta de la matriz de Vandermonde $V(\alpha_1, \dots, \alpha_n)$. Como los escalares $\alpha_1, \dots, \alpha_n$ son distintos dos a dos, la Proposición 4.8.7 nos dice que $[f]_{\mathcal{B}'}^{\mathcal{B}}$ tiene determinante no nulo y, por lo tanto, que f es un isomorfismo. En particular, se trata de una biyectiva y para cada elección de n escalares β_1, \dots, β_n en \mathbb{k} existe exactamente un polinomio $p \in V$ tal que $f(p) = (p(\alpha_1), \dots, p(\alpha_n))^t = (\beta_1, \dots, \beta_n)^t$. Esto claramente prueba la proposición. \square

§9. Espacios de funciones multilineales alternantes

4.9.1. En la Sección 4.2 describimos completamente el espacio vectorial de las funciones multilineales y alternantes de grado máximo sobre un espacio vectorial de dimensión finita. Usando ese resultado y después de algunos cálculos bastante laboriosos, en esta sección describiremos *todos* los espacios de funciones multilineales y alternantes.

4.9.2. Sea V un espacio vectorial de dimensión finita y sea $n = \dim V$. Sea x un vector no nulo de V y sea W un complemento para el subespacio $\langle x \rangle$ en V , de manera que se tenga $V = W \oplus \langle x \rangle$. Podemos considerar la función inclusión $\iota : W \rightarrow V$ y la función lineal $\pi : V \rightarrow W$ tal que $\pi(w) = w$ para todo $w \in W$ y $\pi(x) = 0$. Es inmediato que $\pi \circ \iota = \text{id}_W$. Por otro lado, hay una única función lineal $\lambda : V \rightarrow \mathbb{k}$ tal que $\lambda(x) = 1$ y $\lambda(w) = 0$ para todo $w \in W$. Cualquiera sea $v \in V$ se tiene que

$$v = \pi(v) + \lambda(v)x.$$

Fijemos ahora $d \in \mathbb{N}$ tal que $d \geq 2$. La Proposición 4.2.5 nos permite construir a partir de ι y de π funciones lineales

$$\iota^* : \text{Alt}^d(V) \rightarrow \text{Alt}^d(W), \quad \pi^* : \text{Alt}^d(W) \rightarrow \text{Alt}^d(V).$$

De hecho, si $f : V^d \rightarrow \mathbb{k}$ es un elemento de $\text{Alt}^d(V)$, entonces $\iota^*(f) : W^d \rightarrow \mathbb{k}$ es simplemente la restricción de f a W^d , que es un subconjunto de V^d . Usando otra vez la Proposición 4.2.5, vemos que

$$\iota^* \circ \pi^* = (\pi \circ \iota)^* = \text{id}_W^* = \text{id}_{\text{Alt}^d(W)}$$

y, en particular, esto nos dice que la función ι^* es sobreyectiva. Queremos describir su núcleo.

4.9.3. Lema. Si $g : W^{d-1} \rightarrow \mathbb{k}$ un elemento de $\text{Alt}^{d-1}(W)$, entonces la función $\tilde{g} : V^d \rightarrow \mathbb{k}$ tal que

$$\tilde{g}(v_1, \dots, v_d) = \sum_{i=1}^d (-1)^{i+1} \lambda(v_i) g(\pi(v_1), \dots, \widehat{\pi(v_i)}, \dots, \pi(v_d))$$

cada vez que v_1, \dots, v_n son elementos de V es d -multilineal y alternante.

Demostración. Sea $g : W^{d-1} \rightarrow \mathbb{k}$ un elemento de $\text{Alt}^{d-1}(W)$ y para cada $i \in \llbracket d \rrbracket$ consideremos la función $\tilde{g}_i : V^d \rightarrow \mathbb{k}$ tal que

$$\tilde{g}_i(v_1, \dots, v_d) = \lambda(v_i) g(\pi(v_1), \dots, \widehat{\pi(v_i)}, \dots, \pi(v_d))$$

para cada elección de v_1, \dots, v_d en V .

PRIMER PASO. Empecemos viendo que estas d funciones $\tilde{g}_1, \dots, \tilde{g}_d$ son d -multilineales. Sean $i \in \llbracket d \rrbracket$ y $k \in \llbracket d \rrbracket$, y probemos que la función \tilde{g}_i es lineal con respecto a su k -ésimo argumento. Sean $v_1, \dots, v_{k-1}, v_{k+1}, \dots, v_d, y, z \in V$ y $a, b \in \mathbb{k}$. Consideramos tres casos, de acuerdo a que i sea menor, igual o mayor que k .

- Supongamos primero que $i < k$, de manera que

$$\begin{aligned} \tilde{g}_i(v_1, \dots, v_{k-1}, ay + bz, v_{k+1}, \dots, v_d) \\ = \lambda(v_i) g(\pi(v_1), \dots, \widehat{\pi(v_i)}, \dots, \pi(v_{k-1}), \pi(ay + bz), \pi(v_{k+1}), \dots, \pi(v_d)). \end{aligned}$$

Como la función π es lineal y g es lineal con respecto a su $(k-1)$ -ésimo argumento¹, esto es lo mismo que

$$\begin{aligned} a\lambda(v_i)g(\pi(v_1), \dots, \widehat{\pi(v_i)}, \dots, \pi(v_{k-1}), \pi(y), \pi(v_{k+1}), \dots, \pi(v_d)) \\ + b\lambda(v_i)g(\pi(v_1), \dots, \widehat{\pi(v_i)}, \dots, \pi(v_{k-1}), \pi(z), \pi(v_{k+1}), \dots, \pi(v_d)) \end{aligned}$$

y de acuerdo a la definición de \tilde{g}_i esto es

$$= a\tilde{g}_i(v_1, \dots, v_{k-1}, y, v_{k+1}, \dots, v_d) + b\tilde{g}_i(v_1, \dots, v_{k-1}, z, v_{k+1}, \dots, v_d).$$

- Si en cambio es $i = k$, entonces tenemos que

$$\tilde{g}_i(v_1, \dots, v_{k-1}, ay + bz, v_{k+1}, \dots, v_d)$$

¹Notemos que como en esta situación es $i < k \leq d$ y $i \geq 1$, tenemos que $1 \leq k-1 \leq d-1$ y, por lo tanto, tiene sentido hablar del $(k-1)$ -ésimo argumento de g .

$$\begin{aligned}
&= \lambda(ay + bz)g(\pi(v_1), \dots, \pi(v_{k-1}), \pi(v_{k+1}), \dots, \pi(v_d)) \\
&= a\lambda(y)g(\pi(v_1), \dots, \pi(v_{k-1}), \pi(v_{k+1}), \dots, \pi(v_d)) \\
&\quad + b\lambda(z)g(\pi(v_1), \dots, \pi(v_{k-1}), \pi(v_{k+1}), \dots, \pi(v_d))
\end{aligned}$$

porque la función λ es lineal, y esto es

$$= a\tilde{g}_i(v_1, \dots, v_{k-1}, y, v_{k+1}, \dots, v_d) + b\tilde{g}_i(v_1, \dots, v_{k-1}, z, v_{k+1}, \dots, v_d).$$

- Finalmente, si $i > k$, entonces

$$\begin{aligned}
&\tilde{g}_i(v_1, \dots, v_{k-1}, ay + bz, v_{k+1}, \dots, v_d) \\
&= \lambda(v_i)g(\pi(v_1), \dots, \pi(v_{k-1}), \pi(ay + bz), \pi(v_{k+1}), \dots, \widehat{\pi(v_i)}, \dots, \pi(v_d))
\end{aligned}$$

que, como π es lineal y g es lineal con respecto a su k -ésimo argumento², es

$$\begin{aligned}
&= a\lambda(v_i)g(\pi(v_1), \dots, \pi(v_{k-1}), \pi(y), \pi(v_{k+1}), \dots, \widehat{\pi(v_i)}, \dots, \pi(v_d)) \\
&\quad + b\lambda(v_i)g(\pi(v_1), \dots, \pi(v_{k-1}), \pi(z), \pi(v_{k+1}), \dots, \widehat{\pi(v_i)}, \dots, \pi(v_d)) \\
&= a\tilde{g}_i(v_1, \dots, v_{k-1}, y, v_{k+1}, \dots, v_d) + b\tilde{g}_i(v_1, \dots, v_{k-1}, z, v_{k+1}, \dots, v_d).
\end{aligned}$$

Así, en cualquier caso tenemos que

$$\begin{aligned}
&\tilde{g}_i(v_1, \dots, v_{k-1}, ay + bz, v_{k+1}, \dots, v_d) \\
&= a\tilde{g}_i(v_1, \dots, v_{k-1}, y, v_{k+1}, \dots, v_d) + b\tilde{g}_i(v_1, \dots, v_{k-1}, z, v_{k+1}, \dots, v_d).
\end{aligned}$$

SEGUNDO PASO. Observemos que de acuerdo a la forma en que definimos las funciones $\tilde{g}_1, \dots, \tilde{g}_d$, tenemos que la función \tilde{g} del enunciado del lema es tal que

$$\tilde{g}(v_1, \dots, v_d) = \sum_{i=1}^d (-1)^{i+1} \tilde{g}_i(v_1, \dots, v_d)$$

para cualquier elección de vectores v_1, \dots, v_d en V . Usando esto, podemos probar fácilmente que la función \tilde{g} es d -multilineal: si $k \in \llbracket d \rrbracket$, $v_1, \dots, v_{k-1}, v_{k+1}, \dots, v_d, y, z \in V$ y $a, b \in \mathbb{k}$, entonces

$$\begin{aligned}
&\tilde{g}(v_1, \dots, v_{k-1}, ay + bz, v_{k+1}, \dots, v_d) \\
&= \sum_{i=1}^d (-1)^{i+1} \tilde{g}_i(v_1, \dots, v_{k-1}, ay + bz, v_{k+1}, \dots, v_d) \\
&= \sum_{i=1}^d (-1)^{i+1} (a\tilde{g}_i(v_1, \dots, v_{k-1}, y, v_{k+1}, \dots, v_d) + b\tilde{g}_i(v_1, \dots, v_{k-1}, z, v_{k+1}, \dots, v_d)) \\
&= a \sum_{i=1}^d (-1)^{i+1} \tilde{g}_i(v_1, \dots, v_{k-1}, y, v_{k+1}, \dots, v_d) \\
&\quad + b \sum_{i=1}^d (-1)^{i+1} \tilde{g}_i(v_1, \dots, v_{k-1}, z, v_{k+1}, \dots, v_d)
\end{aligned}$$

$$= a\tilde{g}(v_1, \dots, v_{k-1}, y, v_{k+1}, \dots, v_d) + b\tilde{g}(v_1, \dots, v_{k-1}, z, v_{k+1}, \dots, v_d).$$

TERCER PASO. Para terminar, mostremos que la función \tilde{g} es alternante usando el criterio que nos da la Proposición 4.1.5. Sean v_1, \dots, v_k vectores de V y supongamos que $k \in \llbracket d-1 \rrbracket$ es tal que $v_k = v_{k+1}$. Si $i \in \llbracket d \rrbracket$ es tal que $i \neq k$ e $i \neq k+1$, entonces

$$\tilde{g}_i(v_1, \dots, v_d) = \lambda(v_i)g(\pi(v_1), \dots, \widehat{\pi(v_i)}, \dots, \pi(v_d)) = 0$$

porque la función g es alternante y entre los $d-1$ argumentos $\pi(v_1), \dots, \widehat{\pi(v_i)}, \dots, \pi(v_d)$ en los que aparece aquí evaluada hay dos iguales. Esto nos dice que

$$\begin{aligned} \tilde{g}(v_1, \dots, v_d) &= \sum_{i=1}^d (-1)^{i+1} \tilde{g}_i(v_1, \dots, v_d) \\ &= (-1)^{k+1} \tilde{g}_k(v_1, \dots, v_d) + (-1)^{k+2} \tilde{g}_{k+1}(v_1, \dots, v_d) \\ &= (-1)^{k+1} \lambda(v_k) g(\pi(v_1), \dots, \widehat{\pi(v_k)}, \dots, \pi(v_d)) \\ &\quad + (-1)^{k+2} \lambda(v_{k+1}) g(\pi(v_1), \dots, \widehat{\pi(v_{k+1})}, \dots, \pi(v_d)) \\ &= 0 \end{aligned}$$

simplemente porque los dos términos de esta última suma son opuestos, ya que $\lambda(v_k) = \lambda(v_{k+1})$ y las dos $(d-1)$ -uplas de vectores

$$(\pi(v_1), \dots, \widehat{\pi(v_k)}, \dots, \pi(v_d)) \quad \text{y} \quad (\pi(v_1), \dots, \widehat{\pi(v_{k+1})}, \dots, \pi(v_d))$$

en las que aparece la función g evaluada son, de hecho, iguales. \square

4.9.4. El lema que acabamos de probar nos permite construir para cada $g \in \text{Alt}^{d-1}(W)$ una nueva función $\tilde{g} \in \text{Alt}^d(V)$. Mostremos que esta construcción es lineal y fiel:

Lema. La función

$$\varepsilon : g \in \text{Alt}^{d-1}(W) \mapsto \tilde{g} \in \text{Alt}^d(V)$$

es lineal e inyectiva.

Demostración. Sean g y h dos elementos de $\text{Alt}^{d-1}(W)$ y sean a y b dos escalares de \mathbb{k} . Si

$v_1, \dots, v_d \in V$, entonces

$$\begin{aligned}
\varepsilon(ag + bh)(v_1, \dots, v_d) &= \sum_{i=1}^d (-1)^{i+1} \lambda(v_i)(ag + bh)(\pi(v_1), \dots, \widehat{\pi(v_i)}, \dots, \pi(v_d)) \\
&= \sum_{i=1}^d (-1)^{i+1} \lambda(v_i) \left(ag(\pi(v_1), \dots, \widehat{\pi(v_i)}, \dots, \pi(v_d)) \right. \\
&\quad \left. + bh(\pi(v_1), \dots, \widehat{\pi(v_i)}, \dots, \pi(v_d)) \right) \\
&= a \sum_{i=1}^d (-1)^{i+1} \lambda(v_i) g(\pi(v_1), \dots, \widehat{\pi(v_i)}, \dots, \pi(v_d)) \\
&\quad + b \sum_{i=1}^d (-1)^{i+1} \lambda(v_i) h(\pi(v_1), \dots, \widehat{\pi(v_i)}, \dots, \pi(v_d)) \\
&= a\varepsilon(g)(v_1, \dots, v_d) + b\varepsilon(h)(v_1, \dots, v_d) \\
&= (a\varepsilon(g) + b\varepsilon(h))(v_1, \dots, v_d),
\end{aligned}$$

de manera que $\varepsilon(ag + bh) = a\varepsilon(g) + b\varepsilon(h)$. Esto prueba que la función ε es inyectiva.

Supongamos ahora que $g \in \text{Alt}^{d-1}(V)$ es tal que $\varepsilon(g) = 0$ y sean w_1, \dots, w_{d-1} vectores de W . Recordemos de 4.9.2 que el vector x es tal que $V = W \oplus \langle x \rangle$: tenemos que

$$\begin{aligned}
0 &= \varepsilon(g)(w_1, \dots, w_{d-1}, x) \\
&= \sum_{i=1}^{d-1} (-1)^{i+1} \lambda(w_i) g(\pi(w_1), \dots, \widehat{\pi(w_i)}, \dots, \pi(w_{d-1}), \pi(x)) \\
&\quad + (-1)^{d+1} \lambda(x) g(\pi(w_1), \dots, \pi(w_{d-1})) \\
&= (-1)^{d+1} g(w_1, \dots, w_{d-1}),
\end{aligned}$$

ya que $\pi(x) = 0$, $\lambda(x) = 1$ y $\pi(w_i) = w_i$ para cada $i \in \llbracket d-1 \rrbracket$. Vemos así que $g = 0$ y, en definitiva, que la función ε es inyectiva. \square

4.9.5. Podemos por fin describir el núcleo del morfismo ι^* :

Lema. La imagen de la función $\varepsilon : \text{Alt}^{d-1}(W) \rightarrow \text{Alt}^d(V)$ es precisamente el núcleo de la función $\iota^* : \text{Alt}^d(V) \rightarrow \text{Alt}^d(W)$ de 4.9.2, de manera que tenemos una sucesión exacta corta

$$0 \longrightarrow \text{Alt}^{d-1}(W) \xrightarrow{\varepsilon} \text{Alt}^d(V) \xrightarrow{\iota^*} \text{Alt}^d(W) \longrightarrow 0$$

Demostración. Si $g \in \text{Alt}^{d-1}(V)$, entonces para cada elección de w_1, \dots, w_d en W tenemos que

$$\iota^*(\varepsilon(g))(w_1, \dots, w_d) = \varepsilon(g)(w_1, \dots, w_d) = \sum_{i=1}^d (-1)^{i+1} \lambda(w_i) g(w_1, \dots, \widehat{w_i}, \dots, w_d) = 0,$$

ya que $\lambda(w_i) = 0$ para todo $i \in \llbracket d \rrbracket$, y esto nos dice que $\iota^*(\varepsilon(g)) = 0$, es decir, que $\varepsilon(g)$ está en el núcleo de ι^* . Vemos de esta forma que $\text{Im}(\varepsilon) \subseteq \text{Nu}(\iota^*)$. Veamos la inclusión recíproca.

Sea $f \in \text{Alt}^d(V)$ tal que $\iota^*(f) = 0$ y definamos una función $g : W^{d-1} \rightarrow \mathbb{k}$ poniendo, para cada elección de vectores w_1, \dots, w_{d-1} en W ,

$$g(w_1, \dots, w_{d-1}) := f(w_1, \dots, w_{d-1}, x).$$

Esta función g es $(d-1)$ -multilineal. En efecto, si $k \in \llbracket d-1 \rrbracket$, $w_1, \dots, w_{k-1}, w_{k+1}, \dots, w_{d-1}, y, z \in W$ y $a, b \in \mathbb{k}$, entonces

$$\begin{aligned} g(w_1, \dots, w_{k-1}, ay + bz, w_{k+1}, \dots, w_{d-1}) &= f(w_1, \dots, w_{k-1}, ay + bz, w_{k+1}, \dots, w_{d-1}, x) \\ &= af(w_1, \dots, w_{k-1}, y, w_{k+1}, \dots, w_{d-1}, x) + bf(w_1, \dots, w_{k-1}, z, w_{k+1}, \dots, w_{d-1}, x) \end{aligned}$$

porque f es lineal con respecto a su k -ésimo argumento, y esto es

$$= ag(w_1, \dots, w_{k-1}, y, w_{k+1}, \dots, w_{d-1}) + bg(w_1, \dots, w_{k-1}, z, w_{k+1}, \dots, w_{d-1}).$$

Por otro lado, la función g es alternante: si w_1, \dots, w_{d-1} son $d-1$ vectores de W entre los que hay dos iguales, entonces

$$g(w_1, \dots, w_{d-1}) := f(w_1, \dots, w_{d-1}, x) = 0$$

porque la función f es alternante y está evaluada aquí en d vectores entre los cuales, claro, hay dos que son iguales.

Juntando todo, vemos que g es un elemento de $\text{Alt}^{d-1}(W)$. Mostremos ahora que $f = \varepsilon(g)$: hecho esto, tendremos que $f \in \text{Im}(\varepsilon)$ y, en definitiva, que $\text{Nu}(\iota^*) \subseteq \text{Im}(\varepsilon)$, como queremos.

Para ver que $f = \varepsilon(g)$, consideremos d vectores v_1, \dots, v_d de V y mostramos que

$$f(v_1, \dots, v_d) = \varepsilon(g)(v_1, \dots, v_d). \quad (20)$$

Para cada $i \in \llbracket d \rrbracket$, pongamos $w_{i,1} := \pi(v_i)$ y $w_{i,2} := \lambda(v_i)x$, de manera que $w_{i,1} \in W$, $w_{i,2} \in \langle x \rangle$ y $v_i = w_{i,1} + w_{i,2}$. Tenemos que

$$f(v_1, \dots, v_d) = f(w_{1,1} + w_{1,2}, \dots, w_{d,1} + w_{d,2}).$$

Como f es multilineal, esto es igual a la suma

$$\sum_{k_1, \dots, k_d \in \{1, 2\}} f(w_{1,k_1}, \dots, w_{d,k_d}), \quad (21)$$

con un sumando por cada forma de elegir índices k_1, \dots, k_d en $\{1, 2\}$.

Consideremos un sumando de la suma, esto es fijemos k_1, \dots, k_d en $\{1, 2\}$.

- Si tenemos $k_i = 1$ para todo $i \in \llbracket d \rrbracket$, entonces los d vectores $w_{1,k_1}, \dots, w_{d,k_d}$ están en W y, como $\iota^*(f) = 0$, tenemos que $f(w_{1,k_1}, \dots, w_{d,k_d}) = 0$.
- Por otro lado, si hay al menos dos de los índices k_1, \dots, k_d que son iguales a 2, entonces los d vectores $w_{1,k_1}, \dots, w_{d,k_d}$ son linealmente dependientes —ya que dos de ellos son múltiplos escalares de x — así que otra vez $f(w_{1,k_1}, \dots, w_{d,k_d}) = 0$ porque la función f es alternante.

Vemos así que en la suma de (21) todos los términos se anulan salvo posiblemente aquellos que corresponden a una elección de los índices k_1, \dots, k_d en la que todos son iguales son 1 salvo uno, que es igual a 2. Esto significa que podemos reescribir esa suma en la forma

$$\sum_{i=1}^d f(w_{1,1}, \dots, w_{i-1,1}, w_{i,2}, w_{i+1,1}, \dots, w_{d,1})$$

y, en vista de las definiciones de los vectores que aparecen aquí, esto es

$$\sum_{i=1}^d f(\pi(v_1), \dots, \pi(v_{i-1}), \lambda(v_i)x, \pi(v_{i+1}), \dots, \pi(v_d)). \quad (22)$$

Si $i \in \llbracket d \rrbracket$, entonces como f es multilineal tenemos que

$$\begin{aligned} &f(\pi(v_1), \dots, \pi(v_{i-1}), \lambda(v_i)x, \pi(v_{i+1}), \dots, \pi(v_d)) \\ &= \lambda(v_i) f(\pi(v_1), \dots, \pi(v_{i-1}), x, \pi(v_{i+1}), \dots, \pi(v_d)) \end{aligned} \quad (23)$$

y como f es anti-simétrica, esto es

$$= (-1)^{d-i} \lambda(v_i) f(\pi(v_1), \dots, \pi(v_{i-1}), \pi(v_{i+1}), \dots, \pi(v_d), x),$$

ya que esta última expresión se obtiene de (23) haciendo exactamente $d - i$ transposiciones de los argumentos de f — podemos decir que la x tiene que «saltarse» a los $d - i$ argumentos que lo siguen hasta llegar al final. Usando esto para reescribir cada uno de los sumandos de la suma (22) concluimos que

$$f(v_1, \dots, v_d) = \sum_{i=1}^d (-1)^{d-i} \lambda(v_i) f(\pi(v_1), \dots, \pi(v_{i-1}), \pi(v_{i+1}), \dots, \pi(v_d), x). \quad (24)$$

Por otro lado, tenemos que

$$\varepsilon(g)(v_1, \dots, v_{d-1}) = \sum_{i=1}^d (-1)^{i+1} \lambda(v_i) g(\pi(v_1), \dots, \widehat{\pi(v_i)}, \dots, \pi(v_d))$$

y, de acuerdo a la definición de g , esto es

$$= \sum_{i=1}^d (-1)^{i+1} \lambda(v_i) f(\pi(v_1), \dots, \widehat{\pi(v_i)}, \dots, \pi(v_d), x).$$

Comparando esto con (24), vemos que vale la igualdad (20) que buscábamos. \square

4.9.6. Usando el Lema 4.9.5 podemos dar una expresión muy sencilla para las dimensiones de los espacios de funciones multilineales alternantes.

Proposición. *Sea $n \in \mathbb{N}$. Si V es un espacio vectorial con $\dim V = n$, entonces para cada $d \in \llbracket n \rrbracket$ se*

tiene que

$$\dim \text{Alt}^d(V) = \binom{n}{d}.$$

Demostración. Para cada $n \in \mathbb{N}$ sea $P(n)$ la afirmación del enunciado y procedamos haciendo inducción con respecto n . La afirmación $P(1)$ dice que si V es un espacio vectorial de dimensión 1 entonces $\dim \text{Alt}^1(V) = 1$: esto es consecuencia del Corolario 4.2.6.

Supongamos entonces que $n \in \mathbb{N}$ es tal que $n \geq 2$ y que la afirmación $P(n - 1)$ es cierta, y sea V un espacio vectorial de dimensión n . El espacio $\text{Alt}^1(V)$ es el espacio dual de V , y sabemos entonces que $\dim \text{Alt}^1(V) = n = \binom{n}{1}$. Por otro lado, el Corolario 4.2.6 nos dice que $\dim \text{Alt}^n(V) = 1 = \binom{n}{n}$. Nos queda calcular la dimensión de $\text{Alt}^d(V)$ cuando $1 < d < n$.

Fijemos entonces $d \in [2, n - 1]$, sea x un elemento no nulo de V y sea W un complemento de $\langle x \rangle$ en V , de manera que $V = W \oplus \langle x \rangle$. Tenemos la función inclusión $\iota : W \rightarrow V$, que induce como en la Proposición 4.2.5 una función lineal $\iota^* : \text{Alt}^d(V) \rightarrow \text{Alt}^d(W)$. Como observamos en 4.9.2, esta función ι^* es sobreyectiva. Por otro lado, el Lema 4.9.5 implica inmediatamente que el núcleo de ι^* es isomorfo al espacio $\text{Alt}^{d-1}(W)$. De acuerdo al Teorema 2.4.1, tenemos que

$$\dim \text{Alt}^d(V) = \dim \text{Alt}^{d-1}(W) + \dim \text{Alt}^d(W)$$

y como $\dim W = n - 1$ y estamos suponiendo que vale la afirmación $P(n - 1)$, esto es

$$= \binom{n-1}{d-1} + \binom{n-1}{d} = \binom{n}{d}.$$

La inducción queda así completa y, por lo tanto, también la prueba de la proposición. \square

4.9.7. Fijemos $n \in \mathbb{N}$ y $d \in [n]$. Nos proponemos ahora exhibir explícitamente una base del espacio vectorial $\text{Alt}^d(\mathbb{k}^n)$.

Si x_1, \dots, x_d son elementos de \mathbb{k}^n , escribamos $A(x_1, \dots, x_d)$ a la matriz de $M_{n,d}(\mathbb{k})$ que tiene por columnas, en orden, a los vectores x_1, \dots, x_d . Más aún, para cada subconjunto I de $[n]$ con exactamente d elementos, escribamos $A_I(x_1, \dots, x_d)$ al menor de la matriz $A(x_1, \dots, x_d)$ que se obtiene eliminando las filas cuyo índice *no* pertenece al conjunto I — en otras palabras, las filas de $A_I(x_1, \dots, x_d)$ son las de $A(x_1, \dots, x_d)$ cuyo índice está en I : claramente $A_I(x_1, \dots, x_n)$ es una matriz cuadrada de tamaño d , así que tiene sentido calcular su determinante.

Escribamos $\mathcal{P}_d(n)$ al conjunto de todos los subconjuntos I de $[n]$ de exactamente d elementos, y para cada elemento I de $\mathcal{P}_d(n)$ consideremos la función

$$f_I : (x_1, \dots, x_d) \in (\mathbb{k}^n)^d \mapsto \det(A_I(x_1, \dots, x_d)) \in \mathbb{k}.$$

Sabemos que en $\mathcal{P}_d(n)$ hay $\binom{n}{d}$ elementos, así que de esta forma obtenemos ese número de funciones $(\mathbb{k}^n)^d \rightarrow \mathbb{k}$.

Proposición. *El conjunto $\mathcal{B} = \{f_I : I \in \mathcal{P}_d(n)\}$ es una base del espacio vectorial $\text{Alt}^d(\mathbb{k}^n)$.*

Demostración. Mostraremos, primero, que cualquiera sea I en $\mathcal{P}_d(n)$ la función $f_I : (\mathbb{k}^n)^d \rightarrow \mathbb{k}$ es d -multilineal y alternante, de manera que el conjunto \mathcal{B} está contenido en $\text{Alt}^d(\mathbb{k}^n)$. En segundo lugar, veremos que si $I, J \in \mathcal{P}_d(n)$ son distintos, entonces $f_I \neq f_J$, de manera que \mathcal{B} tiene exactamente $\binom{n}{d}$ elementos. Finalmente, probaremos que el conjunto \mathcal{B} es linealmente independiente: como ya sabemos que $\text{Alt}^d(\mathbb{k}^n)$ tiene dimensión $\binom{n}{d}$, resultará entonces que \mathcal{B} es una base de este espacio.

Sea $I \in \mathcal{P}_d(n)$ y mostremos que $f_I : (\mathbb{k}^n)^d \rightarrow \mathbb{k}$ es d -multilineal y alternante.

- Sea $k \in [\![d]\!]$, sea $x_1, \dots, x_{k-1}, x_{k+1}, \dots, x_d, y, z \in \mathbb{k}^n$ y sean $a, b \in \mathbb{k}$. Es

$$f_I(x_1, \dots, x_{k-1}, ay + bz, x_{k+1}, \dots, x_d) = \det(A_I(x_1, \dots, x_{k-1}, ay + bz, x_{k+1}, \dots, x_d))$$

y como el determinante de una matriz es una función multilineal de las columnas de esta, esto es

$$\begin{aligned} &= a \det(A_I(x_1, \dots, x_{k-1}, y, x_{k+1}, \dots, x_d)) \\ &\quad + b \det(A_I(x_1, \dots, x_{k-1}, z, x_{k+1}, \dots, x_d)) \\ &= a f_I(x_1, \dots, x_{k-1}, y, x_{k+1}, \dots, x_d) + b f_I(x_1, \dots, x_{k-1}, z, x_{k+1}, \dots, x_d). \end{aligned} \quad \blacksquare$$

Vemos así que la función f_I es lineal con respecto a su k -ésimo argumento.

- Sean ahora x_1, \dots, x_d vectores de \mathbb{k}^n entre los cuales hay dos iguales. Por supuesto, la matriz $A(x_1, \dots, x_d)$ tiene dos de sus columnas iguales y, en consecuencia, lo mismo ocurre con su menor $A_I(x_1, \dots, x_d)$: se tiene entonces que

$$f_I(x_1, \dots, x_n) = \det(A_I(x_1, \dots, x_d)) = 0.$$

La función f_I es, por lo tanto, alternante.

Sean ahora I y J dos elementos de $\mathcal{P}_d(n)$ y sean j_1, \dots, j_d los elementos de J listados en orden creciente. Si $I \neq J$, entonces como I y J tienen el mismo número de elementos existe $i \in I \setminus J$ y es claro que la fila i -ésima de la matriz $A(e_{j_1}, \dots, e_{j_d})$ es nula: esa fila es una de las filas del menor $A_I(e_{j_1}, \dots, e_{j_d})$, así que

$$f_I(e_{j_1}, \dots, e_{j_n}) = \det(A_I(e_{j_1}, \dots, e_{j_d})) = 0.$$

Si en cambio $I = J$, entonces el menor $A_I(e_{j_1}, \dots, e_{j_d})$ es la matriz identidad I_d de tamaño d , así que en este caso es

$$f_I(e_{j_1}, \dots, e_{j_n}) = \det(A_I(e_{j_1}, \dots, e_{j_d})) = \det(I_d) = 1.$$

Usando esto podemos probar la segunda afirmación que anunciamos al principio de la prueba: si I y J son dos elementos distintos de $\mathcal{P}_d(n)$ y j_1, \dots, j_d son los elementos de J listados en orden creciente, las observaciones que acabamos de hacer nos dicen que

$$f_I(e_{j_1}, \dots, e_{j_d}) = 0 \neq 1 = f_J(e_{j_1}, \dots, e_{j_d}),$$

así que $f_I \neq f_J$.

Terminemos probando que el conjunto \mathcal{B} es linealmente independiente. Supongamos que para cada $I \in \mathcal{P}_d(n)$ tenemos un escalar $a_I \in \mathbb{k}$ de manera tal que

$$\sum_{I \in \mathcal{P}_d(n)} a_I f_I = 0$$

es una relación de dependencia lineal entre los elementos de \mathcal{B} . Si J es un elemento de $\mathcal{P}_d(n)$ y j_1, \dots, j_d son los elementos de J listados en orden creciente, entonces tenemos que

$$0 = \sum_{I \in \mathcal{P}_d(n)} a_I f_I(e_{j_1}, \dots, d_{j_n}) = a_J,$$

otra vez usando nuestras observaciones de arriba. Vemos así que la relación de dependencia lineal es necesariamente trivial y, por lo tanto, que el conjunto \mathcal{B} es linealmente independiente, como queremos. \square

4.9.8. Ejemplo. Consideremos el caso en que $n = 3$ y $d = 2$. Es $\mathcal{P}_2(3) = \{\{1, 2\}, \{1, 3\}, \{2, 3\}\}$, las funciones 2-lineales alternantes $f_{\{1, 2\}}, f_{\{1, 3\}}, f_{\{2, 3\}} : (\mathbb{k}^3)^2 \rightarrow \mathbb{k}$ están dadas por

$$\begin{aligned} f_{\{1, 2\}} \left(\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}, \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} \right) &= \begin{vmatrix} x_1 & y_1 \\ x_2 & y_2 \end{vmatrix} = x_1 y_2 - x_2 y_1, \\ f_{\{1, 3\}} \left(\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}, \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} \right) &= \begin{vmatrix} x_1 & y_1 \\ x_3 & y_3 \end{vmatrix} = x_1 y_3 - x_3 y_1, \\ f_{\{2, 3\}} \left(\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}, \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} \right) &= \begin{vmatrix} x_2 & y_2 \\ x_3 & y_3 \end{vmatrix} = x_2 y_3 - x_3 y_2, \end{aligned}$$

y la proposición nos dice que $\{f_{\{1, 2\}}, f_{\{1, 3\}}, f_{\{2, 3\}}\}$ es una base de $\text{Alt}^2(\mathbb{k}^3)$. \diamond

4.9.9. Terminemos esta sección con una aplicación importante de los resultados que obtuvimos en ella.

Proposición. Sean $n \in \mathbb{N}$ y $d \in [\![d]\!]$. Sean U y U' dos subespacios de \mathbb{k}^n de dimensión d y sean $\mathcal{B} = (x_1, \dots, x_d)$ y $\mathcal{B}' = (y_1, \dots, y_d)$ bases ordenadas de U y de U' , respectivamente. Es $U = U'$ si y solamente si existe un escalar no nulo $\lambda \in \mathbb{k}$ tal que para cada $I \in \mathcal{P}_d(n)$ es

$$f_I(x_1, \dots, x_d) = \lambda f_I(y_1, \dots, y_d).$$

Demostración. Supongamos primero que $U = U'$, de manera que las bases \mathcal{B} y \mathcal{B}' son dos bases del subespacio U . Sea $C = C(\mathcal{B}, \mathcal{B}') = (c_{i,j})$ la matriz de cambio de base de \mathcal{B} a \mathcal{B}' , de manera que $x_i = c_{1,i} y_1 + \dots + c_{d,i} y_d$ para cada $i \in [\![d]\!]$ o, equivalentemente,

$$A(x_1, \dots, x_d) = A(y_1, \dots, y_d) \cdot C.$$

Como la matriz C es inversible, su determinante $\lambda := \det(C)$ es un escalar no nulo. Si ahora I es un elemento de $\mathcal{P}_d(n)$, entonces esta igualdad nos dice que

$$A_I(x_1, \dots, x_d) = A_I(y_1, \dots, y_d) \cdot C$$

y, por lo tanto, que

$$f_I(x_1, \dots, x_d) = \det(A_I(x_1, \dots, x_d)) = \det(A_I(y_1, \dots, y_d)) \cdot \det(C) = \lambda \cdot f_I(y_1, \dots, y_d).$$

Esto prueba que la condición del enunciado es necesaria para que los subespacio U y U' coincidan.

Probaremos ahora que esa condición también es suficiente. Notemos que si $d = n$ no hay nada que probar, ya que \mathbb{k}^n tiene exactamente un subespacio de dimensión n , así que en todo lo que sigue podemos suponer que $d < n$.

Empezaremos probando la siguiente igualdad:

$$U = \{x \in \mathbb{k}^n : f_J(x_1, \dots, x_d, x) = 0 \text{ para todo } J \in \mathcal{P}_{d+1}(n)\}$$

Sea \bar{U} el conjunto que aparece a la derecha. Si $x \in U$, para cada $J \in \mathcal{P}_{d+1}(n)$ tenemos que $f_J(x_1, \dots, x_d, x) = 0$ porque la función f_J es alternante y los $d + 1$ vectores x_1, \dots, x_d, x son linealmente dependiente: esto nos dice que $x \in \bar{U}$ y, en definitiva, que $U \subseteq \bar{U}$.

Sea ahora $x \in \bar{U}$ y escribamos (e_1, \dots, e_n) a la base ordenada estándar de \mathbb{k}^n . Como los vectores x_1, \dots, x_d son linealmente independientes, la matriz $A(x_1, \dots, x_d)$ tiene rango d y posee, por lo tanto, d filas linealmente independientes o, equivalentemente, un menor cuadrado de tamaño d con determinante no nulo: existe entonces $J \in \mathcal{P}_d(n)$ tal que $\det(A_J(x_1, \dots, x_d)) \neq 0$.

Sean i_{d+1}, \dots, i_n los elementos de $\llbracket n \rrbracket \setminus J$ listados en orden creciente. No es difícil ver —desarrollando el determinante por columnas, por ejemplo— que

$$\det(A(x_1, \dots, x_d, e_{i_{d+1}}, \dots, e_{i_n})) = \pm \det(A_J(x_1, \dots, x_d)) \neq 0,$$

así que $(x_1, \dots, x_d, e_{i_{d+1}}, \dots, e_{i_n})$ es una base ordenada de \mathbb{k}^n . Hay entonces escalares $a_1, \dots, a_n \in \mathbb{k}$ tales que

$$x = a_1 x_1 + \dots + a_d x_d + a_{d+1} e_{i_{d+1}} + \dots + a_n e_{i_n}. \tag{25}$$

Sea $r \in \llbracket d+1, n \rrbracket$. Es $i_r \notin J$, así que el conjunto $K := J \cup \{i_r\}$ tiene exactamente $d + 1$ elementos. Como $x \in \bar{U}$, tenemos que

$$0 = f_K(x_1, \dots, x_d, x) = f_K(x_1, \dots, x_d, a_1 x_1 + \dots + a_d x_d + a_{d+1} e_{i_{d+1}} + \dots + a_n e_{i_n})$$

y, como la función f_K es alternante, el Corolario 4.1.4 nos dice que esto es

$$\begin{aligned} &= f_K(x_1, \dots, x_d, a_{d+1} e_{i_{d+1}} + \dots + a_n e_{i_n}) \\ &= \sum_{l=d+1}^n a_l f_K(x_1, \dots, x_d, e_{j_l}). \end{aligned} \tag{26}$$

Si $l \in \llbracket d+1, n \rrbracket$ es distinto de r , entonces la ultima columna de la matriz $A_K(x_1, \dots, x_d, e_{j_l})$ es nula, así que su determinante $f_K(x_1, \dots, x_d, e_{j_l})$ es nulo. Vemos así que en la suma (26) todos los términos salvo posiblemente el que tiene $l = r$ se anulan y, por lo tanto, que

$$0 = a_r f_K(x_1, \dots, x_d, e_{j_r}) \quad (27)$$

Si desarrollamos el determinante de la matriz $A_K(x_1, \dots, x_d, e_{j_r})$ por la última columna, vemos que coincide, a menos de un factor ± 1 , del determinante de $A_J(x_1, \dots, x_d)$, que no es nulo: la igualdad (27) implica entonces que $a_r = 0$.

Hemos probado que en la igualdad (25) los escalares a_{d+1}, \dots, a_n son todos nulos y, por lo tanto, que $x \in U$. Esto prueba que $\bar{U} \subseteq U$, como queríamos, y completa la verificación de la igualdad 4.9.9.

Podemos ahora probar que suficiencia de la condición del enunciado. Supongamos que hay un escalar no nulo $\lambda \in \mathbb{k}$ tal que $f_I(x_1, \dots, x_d) = \lambda f_I(y_1, \dots, y_d)$ para todo $I \in \mathcal{P}_d(n)$. Como el conjunto $\{f_I : I \in \mathcal{P}_d(n)\}$ es una base de $\text{Alt}^d(\mathbb{k}^n)$, se deduce inmediatamente de esta hipótesis que, de hecho,

$$f(x_1, \dots, x_d) = \lambda f(y_1, \dots, y_d)$$

cualquiera sea $f \in \text{Alt}^d(\mathbb{k}^n)$.

Sea $z \in V$. Si $J \in \mathcal{P}_d(n)$, entonces la función

$$g_{J,z} : (v_1, \dots, v_d) \in (\mathbb{k}^n)^d \mapsto f_J(v_1, \dots, v_d, z) \in \mathbb{k}$$

es d -multilineal y alternante, es decir, es un elemento de $\text{Alt}^d(\mathbb{k}^n)$, así que tenemos que

$$f_J(x_1, \dots, x_d, z) = g_{J,x}(x_1, \dots, x_d) = \lambda g_{J,x}(y_1, \dots, y_d) = \lambda f_J(y_1, \dots, y_d, z)$$

y, como $\lambda \neq 0$, que

$$f_J(x_1, \dots, x_d, z) = 0 \iff f_J(y_1, \dots, y_d, z) = 0.$$

Esto implica que

$$\begin{aligned} U &= \{z \in \mathbb{k}^n : f_J(x_1, \dots, x_d, z) = 0 \text{ para todo } J \in \mathcal{P}_{d+1}(n)\} \\ &= \{z \in \mathbb{k}^n : f_J(y_1, \dots, y_d, z) = 0 \text{ para todo } J \in \mathcal{P}_{d+1}(n)\} = U' \end{aligned}$$

y prueba lo que queremos. \square

4.9.10. Esta proposición nos dice que un subespacio U de \mathbb{k}^n de dimensión d que tiene una base ordenada (x_1, \dots, x_d) queda completamente determinado por la familia de $\binom{n}{d}$ escalares

$$(f_I(x_1, \dots, x_n))_{I \in \mathcal{P}_d(n)}$$

indexada por los elementos de $\mathcal{P}_d(n)$. Las componentes de esta familia son conocidas como **coordenadas de Plücker** del subespacio U , por Julius Plücker.

4.9.11. Ejemplo. Tomemos $n = 4$ y $d = 2$. Si U es un subespacio de dimensión 2 de \mathbb{k}^4 , podemos encontrar una base ordenada (x_1, x_2) de U , con $x_1 = (x_{1,1}, x_{2,1}, x_{3,1}, x_{4,1})^t$ y $x_2 = (x_{1,2}, x_{2,2}, x_{3,2}, x_{4,2})^t$. Si escribimos ij a un subconjunto $\{i, j\}$ de $\llbracket 4 \rrbracket$ para simplificar la notación, las coordenadas de Plücker de U son

$$\begin{aligned} f_{12} &= x_{1,1}x_{2,2} - x_{2,1}x_{1,2}, & f_{23} &= x_{2,1}x_{3,2} - x_{3,1}x_{2,2}, \\ f_{13} &= x_{1,1}x_{3,2} - x_{3,1}x_{1,2}, & f_{24} &= x_{2,1}x_{4,2} - x_{4,1}x_{2,2}, \\ f_{14} &= x_{1,1}x_{4,2} - x_{4,1}x_{1,2}, & f_{34} &= x_{3,1}x_{4,2} - x_{4,1}x_{3,2}. \end{aligned}$$

Observemos que

$$\begin{aligned} &f_{12}f_{34} - f_{13}f_{24} + f_{14}f_{23} & (28) \\ &= (x_{1,1}x_{2,2} - x_{2,1}x_{1,2}) \begin{vmatrix} x_{3,1} & x_{3,2} \\ x_{4,1} & x_{4,2} \end{vmatrix} - (x_{1,1}x_{3,2} - x_{3,1}x_{1,2}) \begin{vmatrix} x_{2,1} & x_{2,2} \\ x_{4,1} & x_{4,2} \end{vmatrix} \\ &\quad + (x_{1,1}x_{4,2} - x_{4,1}x_{1,2}) \begin{vmatrix} x_{2,1} & x_{2,2} \\ x_{3,1} & x_{3,2} \end{vmatrix} \\ &= x_{1,1} \left(x_{2,2} \begin{vmatrix} x_{3,1} & x_{3,2} \\ x_{4,1} & x_{4,2} \end{vmatrix} - x_{3,2} \begin{vmatrix} x_{2,1} & x_{2,2} \\ x_{4,1} & x_{4,2} \end{vmatrix} + x_{4,2} \begin{vmatrix} x_{2,1} & x_{2,2} \\ x_{3,1} & x_{3,2} \end{vmatrix} \right) \\ &\quad - x_{1,2} \left(x_{2,1} \begin{vmatrix} x_{3,1} & x_{3,2} \\ x_{4,1} & x_{4,2} \end{vmatrix} - x_{3,1} \begin{vmatrix} x_{2,1} & x_{2,2} \\ x_{4,1} & x_{4,2} \end{vmatrix} + x_{4,1} \begin{vmatrix} x_{2,1} & x_{3,2} \\ x_{3,1} & x_{4,2} \end{vmatrix} \right) \\ &= x_{1,1} \begin{vmatrix} x_{2,1} & x_{2,2} & x_{2,2} \\ x_{3,1} & x_{3,2} & x_{3,2} \\ x_{4,1} & x_{4,2} & x_{4,2} \end{vmatrix} - x_{1,2} \begin{vmatrix} x_{2,1} & x_{2,1} & x_{2,2} \\ x_{3,1} & x_{3,1} & x_{3,2} \\ x_{4,1} & x_{4,1} & x_{4,2} \end{vmatrix} \\ &= 0. \end{aligned}$$

Esto nos dice que las 6 coordenadas de Plücker de un subespacio de dimensión 2 en \mathbb{k}^4 no son independientes entre sí, ya que siempre satisfacen la relación (28) — esta es la llamada **ecuación de Klein**, por **Felix Klein**. Es posible probar, de hecho, que si $(a_{12}, a_{13}, a_{14}, a_{23}, a_{24}, a_{34})$ es una 6-tupla de escalares tales que $a_{12}a_{34} - a_{13}a_{24} + a_{14}a_{23} = 0$, entonces existe exactamente un subespacio U de dimensión 2 en \mathbb{k}^4 y una base ordenada de (x_1, x_2) de U tales que $a_{ij} = f_{ij}(x_1, x_2)$ para cada $ij \in \mathcal{P}_d(4)$. Esto nos dice que la relación (28) que encontramos arriba es la **única** que hay entre las coordenadas de Plücker de U .

En general, las $\binom{n}{d}$ coordenadas de Plücker de un subespacio de dimensión d en \mathbb{k}^n satisfacen toda una serie de relaciones cuadráticas conocidas como **relaciones de Plücker**, todas muy similares a la de Klein. El estudio de estas relaciones es de enorme importancia en geometría algebraica y álgebra comutativa. \diamond

Capítulo 5

Autovectores y autovalores

§1. Autovectores y autovalores

5.1.1. Sea $f : V \rightarrow V$ un endomorfismo de un espacio vectorial V . Un vector no nulo x de V es un **autovector** de f si existe un escalar $\lambda \in \mathbb{k}$ tal que $f(x) = \lambda x$ y en ese caso decimos que λ es el **autovalor** correspondiente a x o que x es un autovector de autovalor λ . Por otro lado, un escalar $\lambda \in \mathbb{k}$ es un **autovalor** de f si existe un autovector de f que tenga a λ por autovalor. El **espectro** de f es el conjunto $\text{Spec}(f)$ de sus autovalores.

5.1.2. El conjunto de autovalores de un endomorfismo de un espacio vectorial correspondientes a un autovalor fijo *casi* es un subespacio:

Proposición. *Sea V un espacio vectorial y sea $f : V \rightarrow V$ un endomorfismo de V . Para cada $\lambda \in \mathbb{k}$, el subconjunto $E_\lambda(f) = \{x \in V : f(x) = \lambda x\}$ de V es un subespacio de V .*

Los elementos no nulos de $E_\lambda(f)$ son precisamente los autovectores de f de autovalor λ y bien puede no haber ninguno, por supuesto. Llamamos al subespacio $E_\lambda(f)$ el **autoespacio** de f correspondiente a λ . Si tiene dimensión finita, llamamos al número $\dim E_\lambda(f)$ la **multiplicidad geométrica** de λ como autovalor de f . Es claro que esta multiplicidad es positiva si y solamente si λ es un autovalor de f .

Demostración. La afirmación de la proposición es consecuencia de que $E_\lambda(f)$ es el núcleo de la función lineal $\lambda \text{id}_V - f : V \rightarrow V$. \square

5.1.3. Una observación importante es que autovectores correspondientes a autovalores distintos dos a dos son linealmente independientes.

Proposición. *Sea V un espacio vectorial y sea $f : V \rightarrow V$ un endomorfismo de V . Si $x_1, \dots, x_n \in V$ son autovectores de f correspondientes a autovalores $\lambda_1, \dots, \lambda_n \in \mathbb{k}$ y estos últimos son distintos dos*

a dos, entonces los vectores x_1, \dots, x_n son linealmente independientes.

Demostración. Sean $x_1, \dots, x_n \in V$ autovectores de f correspondientes a autovalores $\lambda_1, \dots, \lambda_n \in \mathbb{k}$ distintos dos a dos y sean $a_1, \dots, a_n \in \mathbb{k}$ escalares tales que $a_1x_1 + \dots + a_nx_n = 0$. Para cada $j \in \llbracket n \rrbracket$ podemos aplicar la función f^{j-1} a ambos lados de esta igualdad y ver que

$$\sum_{k=1}^n a_k \lambda_k^{j-1} x_k = 0. \quad (1)$$

Ahora bien, como los escalares $\lambda_1, \dots, \lambda_n$ son distintos dos a dos, es una consecuencia inmediata de la Proposición 4.8.6 que la matriz de Vandermonde

$$V(\lambda_1, \dots, \lambda_n) = \begin{pmatrix} 1 & 1 & \cdots & 1 \\ \lambda_1 & \lambda_2 & \cdots & \lambda_n \\ \vdots & \vdots & \ddots & \vdots \\ \lambda_1^{n-1} & \lambda_2^{n-1} & \cdots & \lambda_n^{n-1} \end{pmatrix}$$

es inversible. Supongamos que su matriz inversa es $V(\lambda_1, \dots, \lambda_n)^{-1} = (\mu_{i,j}) \in M_n(\mathbb{k})$, de manera que, en particular, para cada $i, k \in \llbracket n \rrbracket$ se tiene que

$$\sum_{j=1}^n \mu_{i,j} \lambda_k^{j-1} = \begin{cases} 1, & \text{si } i = j; \\ 0, & \text{en caso contrario.} \end{cases} \quad (2)$$

Para cada $i \in \llbracket n \rrbracket$ vemos entonces —usando primero la igualdad (1) y luego la (2)— que

$$0 = \sum_{j=1}^n \mu_{i,j} \left(\sum_{k=1}^n a_k \lambda_k^{j-1} x_k \right) = \sum_{k=1}^n a_k \left(\sum_{j=1}^n \mu_{i,j} \lambda_k^{j-1} \right) x_k = a_i x_i$$

y, por lo tanto, que los escalares a_1, \dots, a_n son todos nulos, ya que ninguno de los vectores x_1, \dots, x_n lo es. Esto prueba la proposición. \square

5.1.4. El siguiente corolario de la Proposición 5.1.3 es inmediato:

Corolario. *Sea V un espacio vectorial y sea $f : V \rightarrow V$ un endomorfismo de V . Si V tiene dimensión finita, entonces f posee un número finito de autovalores y, de hecho, el número de estos es a lo sumo igual a $\dim V$.*

Demostración. En efecto, si $\lambda_1, \dots, \lambda_n$ son autovalores de f distintos dos a dos, existen vectores no nulos $x_1, \dots, x_n \in V$ tales que $f(x_i) = \lambda_i x_i$ para cada $i \in \llbracket n \rrbracket$ y la Proposición 5.1.3 implica que estos n vectores son linealmente independientes. Por supuesto, de esto se deduce inmediatamente que $n \leq \dim V$. \square

5.1.5. La Proposición 5.1.3 tiene, por otro lado, la siguiente consecuencia importante:

Corolario. Sea V un espacio vectorial y sea $f : V \rightarrow V$ un endomorfismo de V . Si $\lambda_1, \dots, \lambda_n \in \mathbb{k}$ son escalares distintos dos a dos, entonces los subespacios $E_{\lambda_1}(f), \dots, E_{\lambda_n}(f)$ de V son independientes.

Demostración. Podemos usar la caracterización de la independencia de subespacios que da la tercera condición de la Proposición 1.9.4: basta mostrar que si x_1, \dots, x_n son vectores de V tales que para cada $i \in [n]$ es $x_i \in E_{\lambda_i}(f)$ y se tiene que $x_1 + \dots + x_n = 0$, entonces cada uno de los vectores x_1, \dots, x_n es nulo. Eso se sigue inmediatamente de la Proposición 5.1.3: en efecto, si hubiera entre ellos algunos no nulos la igualdad $x_1 + \dots + x_n = 0$ nos diría que esos vectores son linealmente dependientes, contra lo que afirma esa proposición. \square

5.1.6. Sea $n \in \mathbb{N}$, sea $A \in M_n(\mathbb{k})$ una matriz cuadrada de tamaño n con entradas en \mathbb{k} y considéremos el endomorfismo $f_A : x \in \mathbb{k}^n \mapsto Ax \in \mathbb{k}^n$ de \mathbb{k}^n . Aprovechando la estrecha relación que hay entre la matriz A y la función f_A , podemos adaptar las nociones presentadas en esta sección a la matriz A de la siguiente manera:

- Si $x \in \mathbb{k}^n$ y $\lambda \in \mathbb{k}$, decimos que x es un **autovector de A de autovalor λ** si x es un autovector de f_A de autovalor λ .
- Un escalar $\lambda \in \mathbb{k}$ es un **autovalor de A** si la matriz A posee un autovector de autovalor λ , y el **espectro** de A es el conjunto $\text{Spec}(A)$ de sus autovalores.
- Si $\lambda \in \mathbb{k}$, entonces el **autoespacio de A correspondiente a λ** es el subespacio $E_\lambda(f_A)$, que escribimos más simplemente en la forma $E_\lambda(A)$.

Es inmediato, en vista de la forma de estas definiciones, que valen para matrices y sus autovectores y autovalores propiedades análogas a las enunciadas arriba para endomorfismos: en todos los casos, estos resultados se obtienen para la matriz A de considerar el resultado correspondiente para el endomorfismo f_A . Esto mismo puede decirse de casi todos los enunciados que siguen.

5.1.7. En muchas situaciones necesitamos conocer los autovalores de un endomorfismo que se obtiene de otro de alguna forma. Un primer ejemplo de cómo hacer esto es el siguiente:

Lema. Sea $f : V \rightarrow V$ un endomorfismo de un espacio vectorial y sean a y b dos escalares de \mathbb{k} con $a \neq 0$. Un escalar λ es un autovalor de f si y solamente si $a\lambda + b$ es un autovalor de $af + b\text{id}_V$. Así,

$$\text{Spec}(af + b\text{id}_V) = \{a\lambda + b : \lambda \in \text{Spec}(f)\}$$

La Proposición 5.5.3 da una generalización de esto.

Demostración. Sea $\lambda \in \mathbb{k}$. Si $x \in V$ es un autovector de f de autovalor λ , entonces

$$(af + b\text{id}_V)(x) = af(x) + bx = (a\lambda + b)x,$$

de manera que x es un autovector de $af + b\text{id}_V$ de autovalor $a\lambda + b$. Recíprocamente, si x es un autovector de $af + b\text{id}_V$ de autovalor $a\lambda + b$, entonces

$$af(x) + bx = (af + b\text{id}_V)(x) = (a\lambda + b)x = a\lambda x + bx$$

y, como $a \neq 0$, vemos que $f(x) = \lambda x$, así que λ es un autovalor de f . \square

5.1.8. De manera similar, tenemos las siguientes descripciones de los autovalores de el inverso de un isomorfismo y de la composición de dos endomorfismos:

Proposición. Sea V un espacio vectorial de dimensión finita.

- (i) Un endomorfismo $f : V \rightarrow V$ es un automorfismo de V si y solamente si $0 \notin \text{Spec}(f)$, y en ese caso es $\text{Spec}(f^{-1}) = \{\lambda^{-1} : \lambda \in \text{Spec}(f)\}$.
- (ii) Si $f, g : V \rightarrow V$ son endomorfismos de V , entonces $\text{Spec}(f \circ g) = \text{Spec}(g \circ f)$.

Demostración. (i) Sea $f : V \rightarrow V$ un endomorfismo de V . El escalar 0 pertenece a $\text{Spec}(f)$ si y solamente si hay un vector no nulo $v \in V$ tal que $f(v) = 0$, esto es, si f no es inyectiva, y esto ocurre si y solamente si f no es un isomorfismo: esto prueba la primera afirmación.

Supongamos ahora que f es un isomorfismo y sea $\lambda \in \text{Spec}(f)$. Hay entonces un vector no nulo $v \in V$ tal que $f(v) = \lambda v$ y, por lo tanto, $v = \lambda f^{-1}(v)$, es decir, $f^{-1}(v) = \lambda^{-1}v$: vemos así que $\{\lambda^{-1} : \lambda \in \text{Spec}(f)\} \subseteq \text{Spec}(f^{-1})$. Como f^{-1} también es un isomorfismo, esto mismo nos dice que $\{\lambda^{-1} : \lambda \in \text{Spec}(f^{-1})\} \subseteq \text{Spec}((f^{-1})^{-1})$ y, como $(f^{-1})^{-1} = f$, esto prueba la segunda afirmación del enunciado.

(ii) Sean $f, g : V \rightarrow V$ dos endomorfismos de V , sea $\lambda \in \mathbb{k} \setminus \text{Spec}(f \circ g)$ y mostremos que $\lambda \in \mathbb{k} \setminus \text{Spec}(g \circ f)$. Esto implicará, por supuesto, que $\text{Spec}(g \circ f) \subseteq \text{Spec}(f \circ g)$ y, por simetría, que ambos espectros son iguales. Consideramos dos casos:

- Supongamos primero que $\lambda = 0$, de manera que, de acuerdo a la parte (i) de la proposición, el endomorfismo $f \circ g$ es un automorfismo de V . Esto implica que f es sobreinyectivo y que g es inyectivo y, como V tiene dimensión finita, que tanto f como g son automorfismos. Se sigue de esto que $g \circ f$ es un automorfismo y, en definitiva, que $\lambda \notin \text{Spec}(g \circ f)$.
- Supongamos, en segundo lugar, que $\lambda \neq 0$. Como λ no es un autovalor de $f \circ g$, el endomorfismo $\lambda \text{id}_V - f \circ g$ es inversible y existe un endomorfismo $h : V \rightarrow V$ tal que

$$h \circ (\lambda \text{id}_V - f \circ g) = (\lambda \text{id}_V - f \circ g) \circ h = \text{id}_V.$$

Afirmamos que el endomorfismo $k = \lambda^{-1}(\text{id}_V + g \circ h \circ f)$ es inverso de $\lambda \text{id}_V - g \circ f$. En efecto,

$$\begin{aligned} k \circ (\lambda \text{id}_V - g \circ f) &= \lambda^{-1}(\text{id}_V + g \circ h \circ f) \circ (\lambda \text{id}_V - g \circ f) \\ &= \text{id}_V - \lambda^{-1}g \circ f + g \circ h \circ f - \lambda^{-1}g \circ h \circ f \circ g \circ f \\ &= \text{id}_V - \lambda^{-1}g \circ f + \underbrace{\lambda^{-1}g \circ h \circ (\lambda \text{id}_V - f \circ g) \circ f}_{\text{id}_V} \end{aligned}$$

y, como la expresión marcada es igual a id_V , esto es

$$= \text{id}_V - \lambda^{-1}g \circ f + \lambda^{-1}g \circ f = \text{id}_V,$$

y un cálculo similar muestra que $(\lambda \text{id}_V - g \circ f) \circ k = \text{id}_V$. En particular, como el endomorfismo $\lambda \text{id}_V - g \circ f$ es inversible, es inyectivo y, por lo tanto, λ no es un autovalor de $g \circ f$.

Así, en cualquiera de los dos casos vemos que $\lambda \in \mathbb{k} \setminus \text{Spec}(g \circ f)$, como queríamos. \square

§2. Ejemplos

5.2.1. Ejemplos.

- (a) Si V es un espacio vectorial y $\lambda \in \mathbb{k}$ es un escalar, entonces todo vector no nulo de V es un autovector de la función lineal $f : x \in V \mapsto \lambda x \in V$ y el escalar λ es el único autovalor de f . Es $E_\lambda(f) = V$ y la multiplicidad geométrica de λ como autovalor de f es $\dim V$.
- (b) Si V es un espacio vectorial y $f : V \rightarrow V$ es un endomorfismo, los autovectores de f que tienen a 0 como autovalor son precisamente los elementos no nulos del núcleo de f y, en consecuencia, el autoespacio de f correspondiente al autovalor nulo es $E_0(f) = \text{Nu}(f)$.
- (c) Consideremos la función $f : (x, y)^t \in \mathbb{R}^2 \mapsto (y, -x)^t \in \mathbb{R}^2$, endomorfismo del espacio vectorial real \mathbb{R}^2 . Si $v = (x, y)^t$ es un elemento no nulo de \mathbb{R}^2 , entonces los vectores v y $f(v)$ son linealmente independientes: el determinante de la matriz $\begin{pmatrix} x & y \\ y & -x \end{pmatrix}$ que los tiene por columnas es $x^2 + y^2$ y este escalar no es nulo. Esto implica, en particular, que $f(v)$ no es un múltiplo escalar de v y, entonces, que v no es un autovector de f . Vemos así que esta función lineal no posee ningún autovector ni ningún autovalor.

En cambio, el endomorfismo $g : (x, y)^t \in \mathbb{C}^2 \mapsto (y, -x)^t \in \mathbb{C}^2$ del espacio vectorial complejo \mathbb{C}^2 tiene a los vectores $(a, ia)^t$ y $(a, -ia)^t$, cualquiera sea $a \in \mathbb{C} \setminus 0$, como autovectores, con autovalores correspondientes los escalares i y $-i$, ambos de multiplicidad geométrica 1.

- (d) Sea $C^\infty(\mathbb{R})$ el espacio vectorial real de las funciones $\mathbb{R} \rightarrow \mathbb{R}$ que son derivables infinitas veces y sea $D : f \in C^\infty(\mathbb{R}) \mapsto f' \in C^\infty(\mathbb{R})$. Supongamos que $f \in C^\infty(\mathbb{R})$ es un autovector de D de autovalor $\lambda \in \mathbb{R}$, de manera que

$$f' = \lambda f. \tag{3}$$

La función $g : t \in \mathbb{R} \mapsto e^{-\lambda t} f(t) \in \mathbb{R}$ es claramente derivable con continuidad en todo su dominio y tiene en cada $t \in \mathbb{R}$ derivada

$$g'(t) = -\lambda e^{-\lambda t} f(t) + e^{-\lambda t} f'(t) = 0,$$

en vista de la relación (3). Esto implica que la función g es constante y, como $g(0) = f(0)$, que $e^{-\lambda t} f(t) = f(0)$ para todo $t \in \mathbb{R}$, esto es, que $f(t) = f(0)e^{-\lambda t}$. Vemos así que f es un múltiplo escalar de la función exponencial

$$h_\lambda : x \in \mathbb{R} \mapsto e^{-\lambda t} \in \mathbb{R}.$$

Un cálculo directo muestra que esta función h_λ es en efecto un autovector de D de autovalor λ , junto con todos sus múltiplos escalares, así que podemos concluir con esto que

para cada $\lambda \in \mathbb{R}$ los autovectores de $D : f \in C^\infty(\mathbb{R}) \rightarrow f' \in C^\infty(\mathbb{R})$ de autovalor λ son precisamente los múltiplos escalares no nulos de la función h_λ .

Todo elemento de \mathbb{R} es un autovalor de la función D de multiplicidad geométrica 1.

- (e) Podemos considerar también la función dada por la derivación $D : p \in \mathbb{R}[X] \mapsto p' \in \mathbb{R}[X]$ pero ahora definida sobre el espacio vectorial $\mathbb{R}[X]$ de los polinomios con coeficientes reales.

Supongamos que $p \in \mathbb{R}[X]$ es un autovector de D de autovalor λ , de manera que $p' = \lambda p$. Como p no es nulo, tiene un grado d bien definido. Si $d > 0$ y $\lambda \neq 0$, entonces λp tiene grado d y p' tiene grado $d - 1$: esto es imposible, ya que estos dos polinomios son iguales. Por otro lado, si $\lambda = 0$, entonces p' es el polinomio nulo y que, por lo tanto, p es constante. Vemos así que tiene que ser $d = 0$, que el polinomio p es constante y que $\lambda = 0$.

Hemos mostrado, en definitiva, que

el único autovalor de $D : p \in \mathbb{R}[X] \mapsto p' \in \mathbb{R}[X]$ es $\lambda = 0$, los autovectores correspondientes son los polinomios constantes no nulos, y la multiplicidad geométrica de 0 es 1.

La situación es por lo tanto bien distinta a la del ejemplo anterior.

- (f) Sea ahora $Q : p \in \mathbb{R}[X] \mapsto Xp' \in \mathbb{R}[X]$ y supongamos que $p \in \mathbb{R}[X]$ y $\lambda \in \mathbb{R}$ son tales que p no es nulo y $Q(p) = \lambda p$. Si d es el grado de p , hay escalares $a_0, \dots, a_d \in \mathbb{R}$ tales que $a_d \neq 0$ y $p = \sum_{i=0}^d a_i X^i$ y entonces

$$\sum_{i=0}^n \lambda a_i X^i = \lambda p = Q(p) = \sum_{i=0}^d i a_i X^i.$$

Esto implica que $(\lambda - i)a_i = 0$ para todo $i \in \llbracket 0, d \rrbracket$ y, en particular, como $a_d \neq 0$, nos dice que $\lambda = d$. Por otro lado, si $i \in \llbracket 0, d - 1 \rrbracket$, entonces $\lambda - i = d - i \neq 0$, así que tenemos que $a_i = 0$. Vemos de esta forma que $p = a_d X^d$.

La conclusión de esto es que $\text{Spec}(Q) = \mathbb{N}_0$, que todos los autovalores tienen multiplicidad geométrica 1, y que para cada $d \in \mathbb{N}_0$ los autovectores correspondientes a d son los múltiplos no nulos de X^d . \diamond

5.2.2. Ejemplo. Sea ahora V el subespacio de $C^\infty(\mathbb{R})$ de las funciones que son periódicas de periodo 2π , esto es, las funciones $f : \mathbb{R} \rightarrow \mathbb{R}$ infinitamente diferenciables tales que

$$f(t + 2\pi) = f(t) \quad \text{para todo } t \in \mathbb{R}.$$

Si f es un elemento de V , entonces un cálculo inmediato muestra que f' y f'' también están en V , así que en particular tenemos una función $L : f \in V \mapsto f'' \in V$, que claramente es lineal.

Supongamos que $f \in V$ y $\lambda \in \mathbb{R}$ son tales que $L(f) = \lambda f$, de manera que $f'' = \lambda f$. Se tiene entonces que

$$\frac{d}{dt} (f'^2 - \lambda f^2) = 2f' f'' - 2\lambda f f' = 0$$

y, por lo tanto, hay una constante $c \in \mathbb{R}$ tal que

$$f'^2 - \lambda f^2 = c. \quad (4)$$

Supongamos por un momento que $\lambda > 0$. Como la función f es continua en $[0, 2\pi]$ y este intervalo es cerrado y acotado, existe $\xi \in [0, 2\pi]$ tal que $f(\xi) \geq f(t)$ cualquiera sea $t \in [0, 2\pi]$. Más aún, como f es periódica, se sigue inmediatamente de esto que $f(\xi) \geq f(t)$ cualquiera sea $t \in \mathbb{R}$ y, por lo tanto, que f tiene un máximo en ξ y que $f'(\xi) = 0$. De (4), entonces, vemos que $c = -\lambda f(\xi)^2 \leq 0$. Esta misma relación implica ahora que para todo $t \in \mathbb{R}$ es $\lambda f(t)^2 + c = f'(t)^2 \geq 0$, de manera que

$$f(t)^2 \geq -c/\lambda \quad \text{para todo } t \in \mathbb{R}.$$

Ahora bien, podemos hacer con la función f' , que es periódica y derivable, lo mismo que hicimos con f recién, y ver que existe $\zeta \in \mathbb{R}$ tal que $f''(\zeta) = 0$. Pero entonces

$$0 = f''(\zeta)^2 = \lambda^2 f(\zeta)^2 \geq -c\lambda$$

y esto nos permite concluir que $c = 0$. Usando esto y (4) vemos que $f'^2 = \lambda f^2$, así que

$$\lambda \int_0^{2\pi} f(t)^2 dt = \int_0^{2\pi} f'(t)^2 dt = f(t)f'(t) \Big|_0^{2\pi} - \int_0^{2\pi} f(t)f''(t) dt = -\lambda \int_0^{2\pi} f(t)^2 dt,$$

ya que las funciones f y f' son periódicas de periodo 2π y $f'' = \lambda f$. Es entonces

$$2\lambda \int_0^{2\pi} f(t)^2 dt = 0$$

y, como es continua en $[0, 2\pi]$, que la función f se anula allí idénticamente. Por supuesto, como f es periódica de periodo 2π vemos inmediatamente que $f = 0$: así, la función lineal L no posee autovalores negativos.

Consideremos ahora el caso en el que $\lambda < 0$, de manera que existe $\omega \in \mathbb{R}$ bien determinado por λ tal que $\omega > 0$ y $\lambda = -\omega^2$. Sea $g : \mathbb{R} \rightarrow \mathbb{R}$ la función dada por

$$g(t) = f - f(0) \cos \omega t - \omega^{-1} f'(0) \sin \omega t$$

para todo $t \in \mathbb{R}$. Claramente $g \in C^\infty(\mathbb{R})$, calculando directamente vemos que $g(0) = g'(0) = 0$ y $g'' = -\omega^2 g$, y una inducción evidente a partir de estas tres igualdades prueba que

$$g^{(k)}(0) = 0 \text{ cualquier sea } k \in \mathbb{N}_0.$$

Como $g(0) = g'(0)$, usando el Teorema Fundamental del Cálculo primero y la fórmula de integración por partes después vemos que

$$g(x) = \int_0^x g'(t) dt = tg'(t) \Big|_0^x - \int_0^x tg''(t) dt = xg'(x) - \int_0^x tg''(t) dt,$$

y, otra vez gracias el Teorema Fundamental del Cálculo, esto es

$$= x \int_0^x g''(t) dt - \int_0^x t g(t) dt = \int_0^x (x-t) g''(t) dt = -\omega^2 \int_0^x (x-t) g(t) dt.$$

Afirmamos que, más generalmente, para todo $k \in \mathbb{N}$ se tiene que

$$g(x) = (-1)^k \omega^{2k} \int_0^x \frac{(x-t)^{2k-1}}{(2k-1)!} g(t) dt. \quad (5)$$

Que esto es cierto cuando $k = 1$ es lo que acabamos de probar, y para ver que es cierto para los demás valores de k procedemos inductivamente. Supongamos entonces que $k \in \mathbb{N}$ y que esta igualdad vale, de manera que

$$g(x) = (-1)^k \omega^{2k} \int_0^x \frac{(x-t)^{2k-1}}{(2k-1)!} g(t) dt$$

y, de acuerdo a la fórmula de integración por partes, esto es

$$= -(-1)^k \omega^{2k} \frac{(x-t)^{2k}}{(2k)!} g(t) \Big|_0^x + (-1)^k \omega^{2k} \int_0^x \frac{(x-t)^{2k}}{(2k)!} g'(t) dt.$$

El primer término se anula porque $g(0) = 0$, y podemos usar otra vez la fórmula de integración por partes para reescribir el segundo, concluyendo que

$$g(x) = -(-1)^k \omega^{2k} \frac{(x-t)^{2k+1}}{(2k+1)!} g'(t) \Big|_0^x + (-1)^k \omega^{2k} \int_0^x \frac{(x-t)^{2k+1}}{(2k+1)!} g''(t) dt.$$

Otra vez el primer término se anula, ahora porque $g'(0) = 0$, y recordando que $g'' = -\omega^2 g$, vemos que

$$g(x) = (-1)^{k+1} \omega^{2k+2} \int_0^x \frac{(x-t)^{2k+1}}{(2k+1)!} g(t) dt.$$

Esto completa la prueba de (5).

Sea ahora x un número real positivo. De acuerdo a lo que acabamos de probar, tenemos que para todo entero $k \in \mathbb{N}$ es

$$\begin{aligned} |g(x)| &= \left| \omega^{2k} \int_0^x \frac{(x-t)^{2k-1}}{(2k-1)!} g(t) dt \right| \leq \frac{\omega^{2k}}{(2k-1)!} \cdot |x| \cdot \max_{t \in [0,x]} |x-t|^{2k-1} \cdot \max_{t \in [0,x]} |g(t)| \\ &= \frac{\omega^{2k}}{(2k-1)!} \cdot |x|^{2k} \cdot \max_{t \in [0,x]} |g(t)|. \end{aligned}$$

Cuando k crece, esta última expresión converge a 0: vemos así que $g(x) = 0$. De manera completamente similar podemos mostrar que $g(x) = 0$ si x es un número real negativo, así que juntando

todo concluimos¹ que $g = 0$, esto es, que

$$f = f(0) \cos \omega t + \omega^{-1} f'(0) \sin \omega t$$

cualquiera sea $t \in \mathbb{R}$. Recordemos que f está en V , así que es periódica de periodo 2π . Esto implica que

$$f(0) \cos 2\pi\omega + \omega^{-1} f'(0) \sin 2\pi\omega = f(2\pi) = f(0)$$

y

$$-f(0)\omega \sin \pi\omega + f'(0) \cos 2\pi\omega = f'(2\pi) = f'(0),$$

así que

$$\begin{pmatrix} \cos 2\pi\omega - 1 & \omega^{-1} \sin 2\pi\omega \\ -\omega \sin 2\pi\omega & \cos 2\pi\omega - 1 \end{pmatrix} \begin{pmatrix} f(0) \\ f'(0) \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Vemos así que o bien el vector $\begin{pmatrix} f(0) \\ f'(0) \end{pmatrix}$ se nula, y en ese caso la función f es idénticamente nula, o bien la matriz que aparece en esta igualdad tiene determinante nulo, esto es, es

$$\begin{vmatrix} \cos 2\pi\omega - 1 & \omega^{-1} \sin 2\pi\omega \\ -\omega \sin 2\pi\omega & \cos 2\pi\omega - 1 \end{vmatrix} = 2(1 - \cos 2\pi\omega) = 0.$$

Esto es sólo posible si ω es un entero y, como es positivo, si $\omega \in \mathbb{N}$. Así, si ese no es el caso, entonces $\lambda = -\omega^2$ no es un autovalor de L . Si sí lo es, entonces hemos visto que los autovectores correspondientes son de la forma

$$t \in \mathbb{R} \mapsto a \cos \omega t + b \sin \omega t \in \mathbb{R}$$

con $a, b \in \mathbb{R}$, y es inmediato verificar que todas las funciones de esta forma con alguno de los dos escalares no nulos son autovectores.

Finalmente nos queda considerar el caso en el que $\lambda = 0$ y, por lo tanto, en el que $f'' = 0$. Por supuesto, esto implica que hay escalares c y d en \mathbb{R} tales que $f(t) = ct + d$ para todo $t \in \mathbb{R}$ y, como f es periódica, necesariamente $c = 0$: vemos así que f es constante.

La conclusión de todo esto es que el espectro de la función lineal $L : f \in V \mapsto f'' \in V$ es el conjunto

$$\text{Spec}(L) = \{-n^2 : n \in \mathbb{N}_0\}.$$

Más aún, si $n \in \mathbb{N}_0$, entonces el autoespacio $E_{-n^2}(L)$ es el conjunto

- de las funciones constantes, si $n = 0$, y

¹Podríamos haber llegado a la misma conclusión observando que como $g'' = -\omega^2 g$ y $g(0) = g'(0) = 0$ el teorema de existencia y unicidad de soluciones de un problema de valores iniciales para una ecuación diferencial ordinaria nos dice que g es idénticamente nula; esto está explicado, por ejemplo, en [CL55, I.6]. El argumento *ad hoc* que dimos arriba evita tener que recurrir al teorema.

- el de las funciones de la forma $t \in \mathbb{R} \mapsto a \cos \omega t + b \sin \omega t \in \mathbb{R}$ con a y b en \mathbb{R} , si $n > 0$.

La multiplicidad geométrica de 0 como autovalor de L es 1, mientras que todos demás elementos de $\text{Spec}(L)$ tienen multiplicidad geométrica 2. \diamond

5.2.3. Ejemplo. Sea $n \in \mathbb{N}$, sea $p(X) = c_0 + c_1X + \cdots + c_{n-1}X^{-1} + X^n \in \mathbb{k}[X]$ un polinomio mónico de grado n y sea $C(p)$ la matriz compañera de p ,

$$C(p) = \begin{pmatrix} 0 & 0 & \dots & 0 & -c_0 \\ 1 & 0 & \dots & 0 & -c_1 \\ 0 & 1 & \dots & 0 & -c_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & 1 & -c_{n-1} \end{pmatrix} \in M_n(\mathbb{k}).$$

Si $\lambda \in \mathbb{k}$ es una raíz de p , de manera que $p(\lambda) = 0$, sea $x_\lambda = (1, \lambda, \dots, \lambda^{n-1})^t \in \mathbb{k}^n$. Calculando vemos que

$$C(p)^t \cdot x_\lambda = \begin{pmatrix} \lambda \\ \lambda^2 \\ \lambda^3 \\ \vdots \\ \lambda^{n-1} \\ -c_0 - c_1\lambda - \cdots - c_{n-1}\lambda^{n-1} \end{pmatrix} = \begin{pmatrix} \lambda \\ \lambda^2 \\ \lambda^3 \\ \vdots \\ \lambda^{n-1} \\ \lambda^n \end{pmatrix} = \lambda x_\lambda,$$

de manera que x_λ es un autovector de $C(p)^t$ de autovalor λ . Así, el espectro de la matriz $C(p)^t$ contiene al conjunto de raíces de p en \mathbb{k} . Mostraremos más adelante que, de hecho, los dos conjuntos coinciden. \diamond

§3. Diagonalizabilidad

5.3.1. Decimos que un endomorfismo $f : V \rightarrow V$ de un espacio vectorial V es *diagonalizable* si existe una base \mathcal{B} de V cuyos elementos son autovectores de f . La motivación para la elección de este adjetivo viene del siguiente resultado:

Proposición. *Sea V un espacio vectorial de dimensión finita y sea $f : V \rightarrow V$ un endomorfismo de V . Las siguientes afirmaciones son equivalentes:*

- El endomorfismo f es diagonalizable.*
- Existe una base ordenada \mathcal{B} de V tal que la matriz $[f]_{\mathcal{B}}^{\mathcal{B}}$ es diagonal.*
- Existen escalares $\mu_1, \dots, \mu_r \in \mathbb{k}$ distintos dos a dos tales que $V = E_{\mu_1}(f) \oplus \cdots \oplus E_{\mu_r}(f)$.*

Demostración. Sea n la dimensión de V .

(a) \Rightarrow (b) Supongamos que el endomorfismo f es diagonalizable y sea $\mathcal{B} = (x_1, \dots, x_n)$ una base ordenada de V cuyos elementos son autovectores de f , de manera que existen escalares $\lambda_1, \dots, \lambda_n \in \mathbb{k}$ tales que para cada $i \in \llbracket n \rrbracket$ se tiene que $f(x_i) = \lambda_i x_i$. Si $[f]_{\mathcal{B}}^{\mathcal{B}} = (a_{i,j}) \in M_n(\mathbb{k})$ es la matriz de f con respecto a la base \mathcal{B} , vale entonces que cualesquiera sean i y j en $\llbracket n \rrbracket$ es

$$a_{i,j} = \begin{cases} \lambda_i, & \text{si } i = j; \\ 0, & \text{en caso contrario.} \end{cases}$$

Esto, por supuesto, nos dice que $[f]_{\mathcal{B}}^{\mathcal{B}}$ es una matriz diagonal.

(b) \Rightarrow (c) Supongamos ahora que $\mathcal{B} = (x_1, \dots, x_n)$ es una base ordenada de V tal que la matriz $[f]_{\mathcal{B}}^{\mathcal{B}}$ es diagonal y sean $\lambda_1, \dots, \lambda_n \in \mathbb{k}$ los escalares que aparecen en su diagonal, en orden. Estos n escalares no son necesariamente distintos: sean μ_1, \dots, μ_r los elementos del conjunto $\{\lambda_1, \dots, \lambda_n\}$ listados sin repeticiones, de manera que hay una función $\phi : \llbracket n \rrbracket \rightarrow \llbracket r \rrbracket$ tal que $\lambda_i = \mu_{\phi(i)}$ para cada $i \in \llbracket n \rrbracket$. Del Corolario 5.1.5 sabemos que los subespacios $E_{\mu_1}(f), \dots, E_{\mu_r}(f)$ son independientes, así que si llamamos V' a su suma, tenemos que $V' = E_{\mu_1}(f) \oplus \dots \oplus E_{\mu_r}(f)$. Para cada $i \in \llbracket n \rrbracket$ es $f(x_i) = \lambda_i x_i = \mu_{\phi(i)} x_i$, así que $x_i \in E_{\mu_{\phi(i)}}(f)$. Vemos de esta forma que $\mathcal{B} \subseteq V'$ y, por lo tanto, que $V = \langle \mathcal{B} \rangle \subseteq V'$: esto nos dice, claro, que $V = V'$ y prueba que vale (c).

(c) \Rightarrow (a) Supongamos que existen escalares $\mu_1, \dots, \mu_r \in \mathbb{k}$ distintos dos a dos tales que $V = E_{\mu_1}(f) \oplus \dots \oplus E_{\mu_r}(f)$. Para cada $i \in \llbracket r \rrbracket$ sea \mathcal{B}_i una base de $E_{\mu_i}(f)$. De la Proposición 1.9.5 sabemos que la unión $\mathcal{B} = \mathcal{B}_1 \cup \dots \cup \mathcal{B}_r$ es una base de V . Para terminar, basta ver que todos sus elementos son autovectores de f : esto es claro, ya que para cada $i \in \llbracket r \rrbracket$ los elementos de \mathcal{B}_i son autovectores de f de autovalor μ_i porque están en $E_{\mu_i}(f)$ y no son nulos. \square

5.3.2. Podemos dar una caracterización de naturaleza más algebraica de los endomorfismos diagonalizables usando la noción de sistemas completos de proyectores ortogonales dos a dos de 2.7.7.

Proposición. *Sea V un espacio vectorial de dimensión finita. Un endomorfismo $f : V \rightarrow V$ de V es diagonalizable si y solamente si hay un sistema completo (f_1, \dots, f_n) de proyectores ortogonales dos a dos de V tales que $f \in \langle f_1, \dots, f_n \rangle$.*

Demostración. Sea $f : V \rightarrow V$ un endomorfismo de V y supongamos primero que f es diagonalizable. De acuerdo a la Proposición 5.3.1, existen $r \in \mathbb{N}_0$ y escalares $\mu_1, \dots, \mu_r \in \mathbb{k}$ distintos dos a dos tales que $V = E_{\mu_1}(f) \oplus \dots \oplus E_{\mu_r}(f)$. La primera parte de la Proposición 2.7.6 nos dice entonces que hay proyectores $f_1, \dots, f_r : V \rightarrow V$ tales que (f_1, \dots, f_r) es un sistema completo de proyectores ortogonales dos a dos y $E_{\mu_i}(f) = \text{Im}(f_i)$ para cada $i \in \llbracket n \rrbracket$. Afirmamos que

$$f = \mu_1 f_1 + \dots + \mu_r f_r. \tag{6}$$

Para verlo, sea $i \in \llbracket r \rrbracket$ y sea $x \in E_{\mu_i}(f)$: como $E_{\mu_i}(f) = \text{Im}(f_i)$ y f_i es un proyector, es $f_i(x) = x$,

y entonces

$$\begin{aligned} (\mu_1 f_1 + \cdots + \mu_r f_r)(x) &= (\mu_1 f_1 + \cdots + \mu_r f_r)(f_i(x)) = \mu_1 f_1(f_i(x)) + \cdots + \mu_r f_r(f_i(x)) \\ &= \mu_i x = f(x) \end{aligned}$$

Esto nos dice que los endomorfismos que aparecen a ambos lados de la igualdad (6) coinciden en cada elemento de $E_{\mu_1}(f) \cup \cdots \cup E_{\mu_r}(f)$ y, entonces, que coinciden en todo V , ya que esa unión genera a este espacio. Hemos probado con esto que la condición que da la proposición es necesaria.

Para ver que también es suficiente, supongamos ahora que hay un sistema completo de proyectores ortogonales dos a dos (f_1, \dots, f_r) en V y escalares μ_1, \dots, μ_r en \mathbb{k} tales que $f = \mu_1 f_1 + \cdots + \mu_r f_r$. De acuerdo a la segunda parte de la Proposición 2.7.6 nos dice que $V = \text{Im}(f_1) \oplus \cdots \oplus \text{Im}(f_r)$ y entonces si para cada $i \in [r]$ elegimos una base \mathcal{B}_i de $\text{Im}(f_i)$ la unión $\mathcal{B} = \mathcal{B}_1 \cup \cdots \cup \mathcal{B}_r$ es por un lado disjunta y por otro una base de V . Más aún, si $x \in \mathcal{B}$, entonces existe exactamente un índice $i \in [r]$ tal que $x \in \mathcal{B}_i \subseteq \text{Im}(f_i)$ y, por lo tanto, hay un vector $y \in V$ tal que $x = f_i(y)$: esto implica que

$$\begin{aligned} f(x) &= (\mu_1 f_1 + \cdots + \mu_r f_r)(f_i(y)) = \mu_1 f_1(f_i(y)) + \cdots + \mu_r f_r(f_i(y)) \\ &= \mu_i f_i(f_i(y)) = \mu_i x \end{aligned}$$

y, en consecuencia que x es un autovector de f . Concluimos con esto que f es diagonalizable, como queremos. \square

5.3.3. Si $n \in \mathbb{N}$ y $A \in M_n(\mathbb{k})$ es una matriz cuadrada de tamaño n , decimos que A es **diagonalizable** si hay una base de \mathbb{k}^n cuyos elementos son autovectores de A . El resultado análogo a la Proposición 5.3.1 para matrices es el siguiente:

Proposición. Sea $n \in \mathbb{N}$ y sea $A \in M_n(\mathbb{k})$. Las siguientes afirmaciones son equivalentes:

- (a) La matriz A es diagonalizable.
- (b) El endomorfismo $f_A : x \in \mathbb{k}^n \mapsto Ax \in \mathbb{k}^n$ de \mathbb{k}^n es diagonalizable.
- (c) Existe una matriz inversible $C \in GL_n(\mathbb{k})$ tal que la matriz $C^{-1}AC$ es diagonal.

Si estas afirmaciones se cumplen, entonces una matriz C tiene la propiedad de (c) si y solamente si el conjunto de los vectores de \mathbb{k}^n dados por las columnas de C es una base de \mathbb{k}^n formada por autovectores de A .

Demostración. Como los autovectores de la matriz A y los autovectores del endomorfismo f_A coinciden, la equivalencia (a) \Leftrightarrow (b) es inmediata.

Sea \mathcal{B} la base ordenada estándar de \mathbb{k}^n , sea \mathcal{B}' una base ordenada cualquiera de \mathbb{k}^n y sea $C = C(\mathcal{B}', \mathcal{B})$ la matriz de cambio de base de \mathcal{B}' a \mathcal{B} . Sabemos que C es inversible, que $[f_A]_{\mathcal{B}}^{\mathcal{B}} = A$ y, según la Proposición 2.6.4, que

$$[f_A]_{\mathcal{B}'}^{\mathcal{B}'} = C(\mathcal{B}, \mathcal{B}') \cdot [f_A]_{\mathcal{B}}^{\mathcal{B}} \cdot C(\mathcal{B}', \mathcal{B}) = C^{-1}AC. \quad (7)$$

Si f_A es diagonalizable, entonces eligiendo en lo anterior a la base ordenada \mathcal{B}' de manera tal que la matriz $[f_A]_{\mathcal{B}'}^{\mathcal{B}'}$ sea diagonal, vemos que $C^{-1}AC$ es diagonal. Esto prueba que (b) \Rightarrow (c).

Recíprocamente, supongamos que vale (c), sea $C \in \mathrm{GL}_n(\mathbb{k})$ una matriz inversible tal que $C^{-1}AC$ es diagonal y elijamos en lo anterior como \mathcal{B}' a la base ordenada de \mathbb{k}^n tal que $C(\mathcal{B}', \mathcal{B}) = C$, cuya existencia garantiza la Proposición 1.11.5: la igualdad (7) nos dice entonces que la matriz $[f]_{\mathcal{B}'}^{\mathcal{B}'}$ es diagonal y, por lo tanto, que el endomorfismo f_A es diagonalizable, como afirma (b). Vemos de esta manera que (c) \Rightarrow (b), completando la demostración de la equivalencia de las tres afirmaciones del enunciado.

Supongamos ahora que estas afirmaciones se cumplen y sea $C \in \mathrm{GL}_n(\mathbb{k})$ una matriz inversible tal que $D = C^{-1}AC$ es diagonal. Para cada $i \in \llbracket n \rrbracket$ sea x_i el elemento de \mathbb{k}^n dado por la i -ésima columna de C y sea λ_i la entrada (i, i) -ésima de D . Como la matriz C es inversible, la Proposición 4.7.1 nos dice que (x_1, \dots, x_n) es una base ordenada de V . Por otro lado, si (e_1, \dots, e_n) es la base ordenada estándar de \mathbb{k}^n , para cada $i \in \llbracket n \rrbracket$ vale que $x_i = Ce_i$ y entonces

$$Ax_i = ACe_i = C^{-1}De_i = \lambda_i C^{-1}e_i = \lambda_i x_i.$$

Los elementos de la base (x_1, \dots, x_n) son por lo tanto autovectores de A .

Recíprocamente, supongamos que C es una matriz de $M_n(\mathbb{C})$ que tiene por columnas a los vectores de una base ordenada (x_1, \dots, x_n) de autovectores de A , en orden. Como sus columnas son linealmente independientes, esta matriz C es inversible. Sean $\lambda_1, \dots, \lambda_n$ los autovalores correspondientes a x_1, \dots, x_n y sea D la matriz diagonal que los tiene, en orden, a lo largo de la diagonal. Si $i \in \llbracket n \rrbracket$, es

$$CDe_i = \lambda_i Ce_i = \lambda_i x_i = Ax_i = ACe_i.$$

Esto nos dice que la función lineal $x \in \mathbb{k}^n \mapsto CDx - ACx \in \mathbb{k}^n$ se anula sobre los vectores de la base estándar de \mathbb{k}^n , así que es idénticamente nula y, por lo tanto, $CD = AC$. Como C es inversible, tenemos que $C^{-1}AC = D$, así que la matriz C tiene la propiedad de (c). \square

5.3.4. Ejemplo. Sea $n \in \mathbb{N}$, sea $p \in \mathbb{k}[X]$ un polinomio mónico de grado n y sea $C(p)$ la matriz compañera de p . En el Ejemplo 5.3.4 vimos que cada vez que $\lambda \in \mathbb{k}$ es una raíz de p el vector $x_\lambda = (1, \lambda, \dots, \lambda^{n-1})^t \in \mathbb{k}^n$ es un autovector de $C(p)^t$ de autovalor λ . Se sigue de esto que si p posee n raíces $\lambda_1, \dots, \lambda_n$ en \mathbb{k} distintas dos a dos, entonces el conjunto $\{x_{\lambda_1}, \dots, x_{\lambda_n}\}$ es linealmente independiente —de acuerdo a la Proposición 5.1.3— y, como tiene cardinal igual a la dimensión de \mathbb{k}^n , es una base de este espacio vectorial: la matriz $C(p)^t$ es por lo tanto diagonalizable. La matriz $C \in M_n(\mathbb{k})$ cuyas columnas son los vectores $x_{\lambda_1}, \dots, x_{\lambda_n}$, en ese orden, es precisamente la matriz de Vandermonde $V(\lambda_1, \dots, \lambda_n)$ de 4.8.6, y se sigue de la Proposición 5.3.3 que el producto $C^{-1} \cdot C(p) \cdot C$ es una matriz diagonal que tiene a lo largo de la diagonal a los escalares $\lambda_1, \dots, \lambda_n$, en ese orden. \diamond

5.3.5. Ejemplo. Es importante tener en mente que no todo endomorfismo ni toda matriz es diagonalizable. Por ejemplo, supongamos que la matriz

$$A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \in M_2(\mathbb{k})$$

es diagonalizable, de manera que existen una matriz inversible $C \in M_2(\mathbb{k})$ y escalares $\lambda_1, \lambda_2 \in \mathbb{k}$ tales que

$$C^{-1}AC = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}. \quad (8)$$

Como $A^2 = 0$, se tiene entonces que

$$0 = C^{-1}A^2C = C^{-1}AC \cdot C^{-1}AC = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}^2 = \begin{pmatrix} \lambda_1^2 & 0 \\ 0 & \lambda_2^2 \end{pmatrix}$$

y esto sólo es posible si $\lambda_1 = \lambda_2 = 0$. En vista de la igualdad (8), esto implica que $C^{-1}AC = 0$, lo que nos dice que $A = 0$: esto es absurdo.

Usando el mismo argumento, podemos probar que si $A \in M_n(\mathbb{k})$ es una matriz no nula tal que existe $k \in \mathbb{N}$ con $A^k = 0$, entonces A no es diagonalizable.

A pesar de que, como estos ejemplos muestran, no todas las matrices son diagonalizables, es posible mostrar que sobre los cuerpos \mathbb{R} y \mathbb{C} la «mayor parte» de las matrices son diagonalizables — hay que dar, por supuesto, un sentido preciso a esta afirmaciones. Otra forma en la que esto se manifiesta es la siguiente:

si $A = (a_{i,j}) \in M_n(\mathbb{R})$ y $\varepsilon > 0$, entonces existe una matriz diagonalizable $B = (b_{i,j}) \in M_n(\mathbb{R})$ tal que $|a_{i,j} - b_{i,j}| < \varepsilon$ cualesquiera sean i y j en $[\![n]\!]$.

En otras palabras, cambiando arbitrariamente poco las entradas de una matriz de $M_n(\mathbb{R})$ podemos obtener una matriz diagonalizable. \diamond

§4. El polinomio característico

El polinomio característico de una matriz

5.4.1. Si $n \in \mathbb{N}$ y $A \in M_n(\mathbb{k})$, podemos ver a A como un elemento de $M_n(\mathbb{k}(X))$, ya que $\mathbb{k} \subseteq \mathbb{k}(X)$, y considerar la matriz $X \cdot I_n - A$. Llamamos **polinomio característico de A** al polinomio

$$\chi_A = \det(X \cdot I_n - A).$$

Observemos que, en vista del Corolario 4.5.2(ii), esto es efectivamente un elemento de $\mathbb{k}[X]$ y no solamente de $\mathbb{k}(X)$, ya todas las entradas de la matriz $X \cdot I_n - A$ están en $\mathbb{k}[X]$.

5.4.2. Ese mismo corolario nos dice que χ_A es o nulo o de grado a lo sumo n , ya que todas las entradas de la matriz $X \cdot I_n - A$ son o nulas o polinomios de grado a lo sumo 1. Mirando más en detalle, podemos ver que nunca es nulo y determinar precisamente su grado de χ_A .

Proposición. Si $n \in \mathbb{N}$ y $A \in M_n(\mathbb{k})$, entonces el polinomio característico χ_A de A es mónico y tiene grado n .

Demostración. Sea $n \in \mathbb{N}$ y sea $A = (a_{i,j}) \in M_n(\mathbb{k})$. Hacemos inducción con respecto a n , notando que cuando $n = 1$ la afirmación de la proposición es evidente. Supongamos entonces que $n > 1$.

Escribamos $B = (b_{i,j})$ a la matriz $X \cdot I_n - A$, de manera que $b_{i,j} = \delta_{i,j}X - a_{i,j}$ para cada elección de i y j en $\llbracket n \rrbracket$. El desarrollo de Laplace del determinante de B a lo largo de su primera fila nos dice que

$$\chi_A = \sum_{i=1}^n (-1)^{i+1} b_{1,i} \det(B^{(1,i)}) = b_{1,1} \det(B^{(1,1)}) + \underbrace{\sum_{i=2}^n (-1)^{i+1} b_{1,i} \det(B^{(1,i)})}_{\text{(9)}}.$$

El menor $B^{(1,1)}$ de B coincide con la matriz $X \cdot I_n - A^{(1,1)}$, así que $\det(B^{(1,1)})$ es el polinomio característico de $A^{(1,1)}$: la hipótesis inductiva evidente implica entonces que $\det(B^{(1,1)})$ es mónico de grado $n-1$. El primer término del desarrollo de $\det(B)$ es entonces $b_{1,1} \det(B^{(1,1)}) = (X - a_{1,1}) \chi_{A^{(1,1)}}$, que es un polinomio mónico de grado n . Por otro lado, si $i \in \llbracket 2, n \rrbracket$, entonces el término i -ésimo del desarrollo (9) es $(-1)^{i+1} b_{1,i} \det(B^{(1,i)})$: por un lado, $b_{1,i}$ está en \mathbb{k} y, por otro, como cada componente de la matriz $B^{(1,i)}$ es o nula o es un polinomio de grado a lo sumo igual a 1, el determinante $\det(B^{(1,i)})$ es o nulo o un polinomio de grado a lo sumo igual a $n-1$. Vemos así que la suma marcada en (9) es o nula o un polinomio de grado a lo sumo $n-1$ y, juntando todo, que χ_A es mónico de grado n , como afirma la proposición. \square

5.4.3. Ejemplos.

(a) Si $A = (a) \in M_1(\mathbb{k})$, entonces

$$\chi_A = |X - a| = X - a.$$

Si $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in M_2(\mathbb{k})$, entonces

$$\chi_A = \begin{vmatrix} X - a & -b \\ -c & X - d \end{vmatrix} = (X - a)(X - b) - bc = X^2 - (a + d)X + (ad - db).$$

Observemos que los coeficientes de este último polinomio son $-\text{tr}(A)$ y $\det(A)$, el opuesto de la traza de A y el determinante de A . En la Proposición 5.4.17 generalizaremos esto a matrices de tamaño arbitrario.

(b) Sea $n \in \mathbb{N}$ y sea $A = (a_{i,j}) \in M_n(\mathbb{k})$ una matriz triangular superior, de manera que $a_{i,j} = 0$ si $i, j \in \llbracket n \rrbracket$ son tales que $i > j$. Como la matriz $X \cdot I_n - A \in M_n(\mathbb{k}(X))$ es ella también triangular superior, la Proposición 4.8.1 nos permite calcular inmediatamente su determinante y vemos que

$$\chi_A = \det(X \cdot I_n - A) = (X - a_{1,1}) \cdots (X - a_{n,n}).$$

- (c) Más generalmente, sean $n, r \in \mathbb{N}$, sean $n_1, \dots, n_r \in \mathbb{N}$ tales que $n = n_1 + \dots + n_r$ y supongamos que para cada $i, j \in \llbracket r \rrbracket$ con $i \leq j$ tenemos una matriz $A_{i,j} \in M_{n_i, n_j}(\mathbb{k})$. Si A es la matriz triangular superior por bloques

$$\begin{pmatrix} A_{1,1} & A_{1,2} & \cdots & A_{1,r-1} & A_{1,r} \\ 0 & A_{2,2} & \cdots & A_{2,r-1} & A_{2,r} \\ \vdots & \ddots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & A_{r-1,r-1} & A_{r-1,r} \\ 0 & 0 & \cdots & 0 & A_{r,r} \end{pmatrix},$$

entonces el polinomio característico de A es el producto de los polinomios característicos de las matrices que aparecen a lo largo de la diagonal, esto es,

$$\chi_A = \chi_{A_{1,1}} \cdots \chi_{A_{r,r}}.$$

Esto es consecuencia de la Proposición 4.8.2 y de que

$$X \cdot I_n - A = \begin{pmatrix} X \cdot I_{n_1} - A_{1,1} & -A_{1,2} & \cdots & -A_{1,r-1} & -A_{1,r} \\ 0 & X \cdot I_{n_2} - A_{2,2} & \cdots & -A_{2,r-1} & -A_{2,r} \\ \vdots & \ddots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & X \cdot I_{n_{r-1}} - A_{r-1,r-1} & -A_{r-1,r} \\ 0 & 0 & \cdots & 0 & X \cdot I_{n_r} - A_{r,r} \end{pmatrix}$$

también es una matriz triangular superior por bloques.

- (d) Sea $n \in \mathbb{N}$, sea $p \in \mathbb{k}[X]$ un polinomio mónico de grado n y sea $C(p) \in M_n(\mathbb{k})$ la matriz compañera de p . En 4.8.3 calculamos el determinante de la matriz $X \cdot I_n - C(p)$ y ese cálculo nos dice, precisamente, que el polinomio característico de $C(p)$ es

$$\chi_{C(p)} = \det(X \cdot I_n - C(p)) = p.$$

- (e) Si $n \in \mathbb{N}$ y $A \in M_n(\mathbb{k})$, entonces el polinomio característico de A y el de su transpuesta A^t son iguales. En efecto, como $X \cdot I_n - A^t = (X \cdot I_n - A)^t$, de la Proposición 4.5.3 vemos que

$$\chi_{A^t} = \det(X \cdot I_n - A^t) = \det(X \cdot I_n - A)^t = \det(X \cdot I_n - A) = \chi_A. \quad \diamond$$

5.4.4. Una observación importante es que dos matrices que son conjugadas tienen el mismo polinomio característico:

Proposición. *Sea $n \in \mathbb{N}$ y sean A y B dos matrices de $M_n(\mathbb{k})$. Si existe una matriz inversible $C \in GL_n(\mathbb{k})$ tal que $B = C^{-1}AC$, entonces $\chi_B = \chi_A$.*

Demostración. En efecto, si existe una tal matriz, tenemos en $M_n(\mathbb{k}(X))$ que

$$X \cdot I_n - B = X \cdot C^{-1}I_nC - C^{-1}AC = C^{-1}(X \cdot I_n - A)C$$

de manera que, según el Corolario 4.3.4, es

$$\chi_B = \det(X \cdot I_n - B) = \det(C^{-1}(X \cdot I_n - A)C) = \det(X \cdot I_n - A) = \chi_A,$$

como afirma la proposición. \square

El polinomio característico de un endomorfismo

5.4.5. Usando la invariancia del polinomio característico de una matriz por conjugación que nos da la Proposición 5.4.4 podemos extender la noción de polinomio característico, que definimos arriba para matrices, a endomorfismos: si $f : V \rightarrow V$ es un endomorfismo de un espacio vectorial V de dimensión finita n y \mathcal{B} es una base ordenada de V , llamamos *polinomio característico de f* y escribimos χ_f al polinomio característico de la matriz $[f]_{\mathcal{B}}$, de manera que

$$\chi_f = \det(X \cdot I_n - [f]_{\mathcal{B}}).$$

Por supuesto, para que esta definición tenga sentido el segundo miembro de esta igualdad tiene que depender solamente de f y no de la elección de la base ordenada \mathcal{B} . Para ver que esto es así, sea \mathcal{B}' otra base ordenada de V y sea $C = C(\mathcal{B}', \mathcal{B})$ la matriz de cambio de base de \mathcal{B}' a \mathcal{B} . Sabemos que

$$[f]_{\mathcal{B}'} = C(\mathcal{B}, \mathcal{B}') \cdot [f]_{\mathcal{B}} \cdot C(\mathcal{B}', \mathcal{B}) = C^{-1} \cdot [f]_{\mathcal{B}} \cdot C$$

y, por lo tanto, que las matrices $[f]_{\mathcal{B}}$ y $[f]_{\mathcal{B}'}$ son conjugadas: la Proposición 5.4.4 nos dice entonces que estas dos matrices tienen el mismo polinomio característico, que es lo que queremos.

Una consecuencia inmediata de la forma de esta definición es que si $n \in \mathbb{N}$ y $A \in M_n(\mathbb{k})$, entonces el polinomio característico χ_A de A coincide con el polinomio característico χ_{f_A} del endomorfismo $f_A : x \in \mathbb{k}^n \mapsto Ax \in \mathbb{k}^n$ de \mathbb{k}^n asociado a A . En efecto, del Ejemplo 2.6.2(a) sabemos si \mathcal{B} es la base ordenada estándar de \mathbb{k}^n , entonces $[f_A]_{\mathcal{B}} = A$.

5.4.6. Ejemplos.

(a) Sea V un espacio vectorial de dimensión finita y sea $f : V \rightarrow V$ un endomorfismo de V .

Si V_1, \dots, V_r son subespacios de V que son f -invariantes tales que $V = V_1 \oplus \dots \oplus V_r$ y para cada $i \in \llbracket r \rrbracket$ escribimos $f_{V_i} : V_i \rightarrow V_i$ a la restricción de V a V_i , entonces

$$\chi_f = \chi_{f_{V_1}} \cdots \chi_{f_{V_r}}.$$

Para verlo, para cada $i \in \llbracket r \rrbracket$ sea $m_i = \dim V_i$ y sea $\mathcal{B}_i = (x_{1,1}, \dots, x_{1,m_i})$ una base ordenada de V_i . Sabemos que $\mathcal{B} = (x_{1,1}, \dots, x_{1,m_1}, \dots, x_{r,1}, \dots, x_{r,m_r})$ es una base ordenada de V . La matriz de f con respecto a \mathcal{B} es la matriz diagonal por bloques

$$[f]_{\mathcal{B}} = \begin{pmatrix} [f_{V_1}]_{\mathcal{B}_1}^{\mathcal{B}_1} & & \\ & \ddots & \\ & & [f_{V_r}]_{\mathcal{B}_r}^{\mathcal{B}_r} \end{pmatrix},$$

así que el resultado es consecuencia de lo que vimos en el Ejemplo 5.4.3(c).

- (b) Si V es un espacio vectorial de dimensión finita y $f : V \rightarrow V$ es un endomorfismo de V , entonces el polinomio característico χ_f de f coincide con el polinomio característico χ_{f^t} de la función transpuesta $f^t : V^* \rightarrow V^*$, esto es,

$$\chi_{f^t} = \chi_f.$$

En efecto, si \mathcal{B} es una base ordenada de V y \mathcal{B}^* es la base de V^* dual a \mathcal{B} , entonces la Proposición 3.3.6 nos dice que la matriz $[f^t]_{\mathcal{B}^*}^{\mathcal{B}}$ es la transpuesta de $[f]_{\mathcal{B}}^{\mathcal{B}}$ y el Ejemplo 5.4.3(e) nos dice que, en consecuencia, $[f^t]_{\mathcal{B}^*}^{\mathcal{B}}$ y $[f]_{\mathcal{B}}^{\mathcal{B}}$ tienen el mismo polinomio característico. \diamond

Autovalores y el polinomio característico

5.4.7. El polinomio característico de un endomorfismo o de una matriz nos permite caracterizar de manera sencilla sus autovalores. Para probarlo necesitaremos la siguiente observación muy sencilla. Si p y q son elementos de $\mathbb{k}[X]$ y $a, b, \lambda \in \mathbb{k}$, sabemos que

$$(ap + bq)(\lambda) = ap(\lambda) + bq(\lambda), \quad (p \cdot q)(\lambda) = p(\lambda) \cdot q(\lambda).$$

Así, el resultado de evaluar una combinación lineal o un producto de polinomios en λ es el mismo que el de hacer esa combinación lineal o ese producto con los resultados de evaluar los polinomios en λ . Queremos extender esto al cálculo de determinantes.

Si $A = (p_{i,j}) \in M_n(\mathbb{k}[X])$ una matriz con entradas en $\mathbb{k}[X]$ y $\lambda \in \mathbb{k}$ es un escalar, entonces podemos construir la matriz $A(\lambda) := (p_{i,j}(\lambda)) \in M_n(\mathbb{k})$, cada una de cuyas entradas se obtiene evaluando la correspondiente entrada de A en λ . Decimos que la matriz $A(\lambda)$ se obtiene **evaluando** la matriz A en λ . El siguiente lema nos permite calcular su determinante:

Lema. Sea $A = (p_{i,j}) \in M_n(\mathbb{k}[X])$ una matriz con entradas en $\mathbb{k}[X]$ y sea $p = \det(A)$ su determinante, que es un elemento de $\mathbb{k}[X]$. Si $\lambda \in \mathbb{k}$, entonces $p(\lambda) = \det(A(\lambda))$.

Esto nos dice que calcular el determinante de una matriz A de polinomios y evaluar el resultado en λ produce el mismo resultado que evaluar primero la matriz A en λ y calcular el determinante de la matriz que obtenemos.

Demostración. Procedemos por inducción con respecto a n , observando que cuando $n = 1$ el resultado es evidente. Supongamos entonces que $n > 1$. Es claro que para cada $i, j \in \llbracket n \rrbracket$ la matriz $A^{(i,j)}(\lambda)$ que se obtiene evaluando en λ al (i, j) -ésimo menor $A^{(i,j)} \in M_{n-1}(\mathbb{k}[X])$ de A coincide con el (i, j) -ésimo menor $A(\lambda)^{(i,j)}$ de la matriz $A(\lambda)$, así que, por supuesto, ambas matrices tienen el mismo determinante. Si usamos la fórmula de Laplace para desarrollar el determinante de A a lo largo de su primera fila y ponemos $q_i = \det(A^{(1,i)})$ para cada $i \in \llbracket n \rrbracket$, vemos que

$$p = \det(A) = \sum_{i=1}^n (-1)^{i+1} p_{1,i} \cdot q_i.$$

La hipótesis inductiva evidente nos dice que que

$$q_i(\lambda) = (\det(A^{(1,i)}))(\lambda) = \det(A(\lambda)^{(1,i)}),$$

y entonces

$$p(\lambda) = \sum_{i=1}^n (-1)^{i+1} p_{1,i}(\lambda) \cdot q_i(\lambda) = \sum_{i=1}^n (-1)^{i+1} p_{1,i}(\lambda) \cdot \det(A(\lambda)^{(1,i)}) = \det(A(\lambda)),$$

ya que la última suma es precisamente el desarrollo de Laplace del determinante de $A(\lambda)$ a lo largo de la primera fila de la matriz. \square

5.4.8. Podemos ahora dar la caracterización de los autovalores de un endomorfismo o de una matriz que prometimos:

Proposición.

- (i) Sea $f : V \rightarrow V$ un endomorfismo de un espacio vectorial de dimensión finita. Un escalar $\lambda \in \mathbb{k}$ es autovalor de f si y solamente si es raíz del polinomio característico χ_f de f .
- (ii) Sea $n \in \mathbb{N}$ y sea $A \in M_n(\mathbb{k})$. Un escalar $\lambda \in \mathbb{k}$ es autovalor de A si y solamente si es raíz del polinomio característico χ_A de A .

Demostración. (i) Sea $\lambda \in \mathbb{k}$ y sea \mathcal{B} una base ordenada de V . El escalar λ es un autovalor de f si y solo si la función lineal $\lambda \text{id}_V - f : V \rightarrow V$ no es un isomorfismo, y esto ocurre si y solo si el determinante de la matriz $[\lambda \text{id}_V - f]_{\mathcal{B}}$ es nulo. Esa matriz es $\lambda I_n - [f]_{\mathcal{B}}$, y su determinante es el valor de polinomio característico de la matriz $[f]_{\mathcal{B}}$ en λ o, lo que es lo mismo, el valor del polinomio característico de f en λ . Esto prueba la afirmaciones del enunciado.

(ii) Si $f_A : x \in \mathbb{k}^n \mapsto Ax \in \mathbb{k}^n$, entonces los autovalores y el polinomio característico de A coinciden con los de f_A , así que esta segunda parte es consecuencia inmediata de la primera. \square

5.4.9. Llamamos **multiplicidad algebraica** de un escalar $\lambda \in \mathbb{k}$ como autovalor de un endomorfismo f de un espacio vectorial de dimensión finita o de una matriz A a la multiplicidad con que λ es raíz del polinomio característico χ_f o χ_A .

Proposición.

- (i) Sea $f : V \rightarrow V$ un endomorfismo de un espacio vectorial V de dimensión finita. La multiplicidad geométrica de λ como autovalor de f es a lo sumo igual a la multiplicidad algebraica de λ como autovalor de f .
- (ii) Sea $n \in \mathbb{N}$ y sea $A \in M_n(\mathbb{k})$. La multiplicidad geométrica de λ como autovalor de A es a lo sumo igual a la multiplicidad algebraica de λ como autovalor de A .

Demostración. (i) Sea $\lambda \in \mathbb{k}$ y supongamos que la multiplicidad geométrica de λ como autovalor de f es r , que bien puede ser nula. Sea (x_1, \dots, x_r) una base ordenada de $E_\lambda(f)$, sea $n = \dim V$, y sean x_{r+1}, \dots, x_n elementos de V tales que $\mathcal{B} = (x_1, \dots, x_n)$ es una base ordenada de V . Como los

primeros r vectores de la mase ordenada \mathcal{B} son autovectores de f de autovalor λ , la matriz $[f]_{\mathcal{B}}^{\mathcal{B}}$ es de la forma

$$[f]_{\mathcal{B}}^{\mathcal{B}} = \begin{pmatrix} \lambda I_r & P \\ 0 & Q \end{pmatrix}$$

para ciertas matrices $P \in M_{r,n-r}(\mathbb{k})$ y $Q \in M_{n-r}(\mathbb{k})$, así que el polinomio característico de f , que es el de $[f]_{\mathcal{B}}^{\mathcal{B}}$, es, de acuerdo al Ejemplo 5.4.3(c),

$$\chi_f = \chi_{\lambda I_r} \cdot \chi_Q = (X - \lambda)^r \cdot \chi_Q$$

y claramente tiene a λ como raíz con multiplicidad al menos r .

(ii) Si $f_A : x \in \mathbb{k}^n \mapsto Ax \in \mathbb{k}^n$, entonces $\chi_A = \chi_{f_A}$ y $E_\lambda(A) = E_\lambda(f_A)$, así que lo que afirma esta segunda parte se deduce inmediatamente de la primera. \square

5.4.10. De la Proposición 5.4.8 obtenemos fácilmente una condición suficiente para la diagonalizabilidad de un endomorfismo o una matriz:

Proposición. *Sea V un espacio vectorial de dimensión finita n y sea $f : V \rightarrow V$ un endomorfismo de V . Si el polinomio característico de f tiene n raíces simples en \mathbb{k} , entonces f es diagonalizable.*

La condición que da esta proposición no es necesaria: por ejemplo, la función identidad $\text{id}_{\mathbb{k}^n} : \mathbb{k}^n \rightarrow \mathbb{k}^n$ es diagonalizable pero su polinomio característico, $(X - 1)^n$, no tiene raíces simples salvo si $n = 1$.

Demostración. Supongamos que el polinomio característico χ_f tiene n raíces simples en \mathbb{k} . De acuerdo a la Proposición 5.4.8, cada una de esas raíces es un autovalor de f y entonces f posee n autovectores de autovalores distintos dos a dos. El conjunto de estos autovectores es, según la Proposición 5.1.3, linealmente independiente y se trata, por lo tanto, de una base de V : la función f es por lo tanto diagonalizable. \square

5.4.11. En general, por supuesto, el polinomio característico de un endomorfismo no tiene raíces simples y la Proposición 5.4.10 no se aplica. El siguiente resultado da un criterio de diagonalizabilidad que tiene en cuenta esa posibilidad.

Proposición. *Sea V un espacio vectorial de dimensión finita. Un endomorfismo $f : V \rightarrow V$ de V es diagonalizable si y solamente si su polinomio característico se factoriza como producto de factores lineales y la multiplicidad geométrica de cada autovalor de f coincide con su multiplicidad algebraica.*

En otras palabras, para que f sea diagonalizable es necesario y suficiente que haya suficientes autovalores como para factorizar el polinomio característico de f y que para cada autovalor de f haya *suficientes* autovalores de cada autovalor. Esta proposición se ocupa de la situación en la que la multiplicidad geométrica de cada autovalor es la máxima posible: en el Corolario 5.7.9 que veremos más adelante nos ocuparemos de la situación opuesta, en la que todos los autovalores tienen la mínima multiplicidad geométrica posible, uno.

Demostración. Pongamos $n := \dim V$. Supongamos primero que el endomorfismo f es diagonalizable, sea $\mathcal{B} = (x_1, \dots, x_n)$ una base ordenada de V de autovectores de f y sean $\lambda_1, \dots, \lambda_n$ los autovalores correspondientes a los vectores de \mathcal{B} , en orden. La matriz $[f]_{\mathcal{B}}^{\mathcal{B}}$ es diagonal, con los escalares $\lambda_1, \dots, \lambda_n$ a lo largo de la diagonal, así que de acuerdo al Ejemplo 5.4.3(b) su polinomio característico se factoriza como producto de polinomios mónicos con raíces precisamente $\lambda_1, \dots, \lambda_n$.

Sea λ es uno de estos escalares y pongamos $I := \{i \in \llbracket n \rrbracket : \lambda_i\}$. La multiplicidad algebraica de λ como autovalor de f es entonces $\#I$. Por otro lado, los vectores del conjunto $\{x_i : i \in I\}$ son todos autovectores de f de autovalor λ y son linealmente independientes: esto implica que la multiplicidad geométrica de λ como autovalor de f es $\dim E_{\lambda}(f) \geq \#I$. Como sabemos que además $\dim E_{\lambda}(f) \leq \#I$, tenemos la igualdad que queremos.

Supongamos ahora, para probar la afirmación recíproca, que el polinomio característico χ_f se factoriza como producto de factores lineales. Si μ_1, \dots, μ_r son las raíces de f en \mathbb{k} listadas sin repeticiones y m_1, \dots, m_r las correspondientes multiplicidades geométricas, entonces $m_1 + \dots + m_r = n$. Esto implica que $E_{\mu_1}(f) \oplus \dots \oplus E_{\mu_r}(f)$ tiene dimensión n y, por lo tanto, que coincide con V , así que f es diagonalizable. \square

5.4.12. Ejemplos.

- (a) Si $p \in \mathbb{k}[X]$ es un polinomio mónico de grado n que tiene n raíces simples en \mathbb{k} , entonces la matriz compañera $C(p)$ y su transpuesta $C(p)^t$ son diagonalizables: en efecto, el polinomio característico ambas precisamente p , así que esto sigue de la Proposición 5.4.10. Veremos en el Ejemplo 5.7.3 que, de hecho, esta condición es también necesaria para que la matriz $C(p)$ sea diagonalizable.

En el Ejemplo 5.3.4 ya habíamos visto que bajo esta hipótesis sobre p la matriz $C(p)^t$ es diagonalizable: ahí lo hicimos exhibiendo explícitamente una base de \mathbb{k}^n de autovectores. Notemos que aquí probamos que $C(p)$ es diagonalizable sin tener que hacer eso.

- (b) Fijemos ahora $n \in \mathbb{N}$ y consideremos la matriz $A_n = (a_{i,j}) \in M_n(\mathbb{k})$ que para cada $i, j \in \llbracket n \rrbracket$ tiene entrada (i, j) -ésima dada por

$$a_{i,j} = \begin{cases} -1 & \text{si } |i-j| = 1; \\ 0 & \text{en caso contrario.} \end{cases}$$

Por ejemplo, cuando n es 5 la matriz es

$$A_5 = \begin{pmatrix} 0 & -1 & 0 & 0 & 0 \\ -1 & 0 & -1 & 0 & 0 \\ 0 & -1 & 0 & -1 & 0 \\ 0 & 0 & -1 & 0 & -1 \\ 0 & 0 & 0 & -1 & 0 \end{pmatrix}.$$

Si $n > 2$, para calcular el polinomio característico de A_n podemos considerar el desarrollo

de Laplace a lo largo de la primera de las filas de la matriz $X \cdot I_n - A_n \in M_n(\mathbb{k}(X))$,

$$\begin{aligned} \chi_{A_n} &= \overbrace{\begin{vmatrix} X & 1 & 0 & 0 & \cdots & 0 & 0 \\ 1 & X & 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & X & 1 & \cdots & 0 & 0 \\ 0 & 0 & 1 & X & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & X & 1 \\ 0 & 0 & 0 & 0 & \cdots & 1 & X \end{vmatrix}}^n \\ &= X \underbrace{\begin{vmatrix} X & 1 & 0 & \cdots & 0 & 0 \\ 1 & X & 1 & \cdots & 0 & 0 \\ 0 & 1 & X & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & X & 1 \\ 0 & 0 & 0 & \cdots & 1 & X \end{vmatrix}}_{n-1} - \underbrace{\begin{vmatrix} 1 & 1 & 0 & \cdots & 0 & 0 \\ 0 & X & 1 & \cdots & 0 & 0 \\ 0 & 1 & X & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & X & 1 \\ 0 & 0 & 0 & \cdots & 1 & X \end{vmatrix}}_{n-1}. \end{aligned}$$

El primero de estos dos últimos determinantes es precisamente el que calcula el polinomio característico de la matriz A_{n-1} y, por otro lado, desarrollando el segundo a lo largo de su primera columna vemos que es igual al polinomio característico de A_{n-2} . Concluimos así que

$$\chi_{A_n} = X\chi_{A_{n-1}} - \chi_{A_{n-2}} \quad \text{si } n > 2. \quad (10)$$

Calculando directamente vemos además que

$$\chi_{A_1} = X, \quad \chi_{A_2} = X^2 - 1, \quad (11)$$

Estas dos igualdades y la relación de recurrencia (10) nos permiten calcular los polinomios χ_{A_n} recursivamente. En el Cuadro 5.1 de la página siguiente listamos los primeros.

Para cada $n \in \mathbb{N}$ consideremos el polinomio U_n tal que

$$U_n(X) = \chi_{A_n}(2X).$$

Como consecuencia de (10) y de (11) se tiene que

$$U_1 = 2X, \quad U_2 = 4X^2 - 1, \quad U_n = 2XU_{n-1} - U_{n-2} \quad \text{si } n > 2.$$

El polinomio U_n se llama el n -ésimo **polinomio de Chebyshev de segunda especie**, por Pafnuti Chebyshev; es usual iniciar la sucesión de estos polinomios con $U_0 = 1$, y en ese caso la relación de recurrencia vale para todo $n > 1$.

n	χ_{A_n}	U_n
0		1
1	X	$2X$
2	$X^2 - 1$	$4X^2 - 1$
3	$X^3 - 2X$	$8X^3 - 4X$
4	$X^4 - 3X^2 + 1$	$16X^4 - 12X^2 + 1$
5	$X^5 - 4X^3 + 3X$	$32X^5 - 32X^3 + 6X$
6	$X^6 - 5X^4 + 6X^2 - 1$	$64X^6 - 80X^4 + 24X^2 - 1$
7	$X^7 - 6X^5 + 10X^3 - 4X$	$128X^7 - 192X^5 + 80X^3 - 8X$
8	$X^8 - 7X^6 + 15X^4 - 10X^2 + 1$	$256X^8 - 448X^6 + 240X^4 - 40X^2 + 1$
9	$X^9 - 8X^7 + 21X^5 - 20X^3 + 5X$	$512X^9 - 1024X^7 + 672X^5 - 160X^3 + 10X$

Cuadro 5.1. Los polinomios característicos de las matrices A_n del ejemplo 5.4.12(b) y los polinomios de Chebyshev de segunda especie.

Supongamos desde ahora que el cuerpo de base \mathbb{k} con el que estamos trabajando es \mathbb{R} . Mostremos que para cada $n \in \mathbb{N}_0$ y cada $\theta \in \mathbb{R} \setminus \pi\mathbb{Z}$ es

$$U_n(\cos \theta) = \frac{\sin(n+1)\theta}{\sin \theta}. \quad (12)$$

Cuando $n = 0$, esto es evidente, y cuando $n = 1$ la igualdad es consecuencia de la fórmula de duplicación para la función seno,

$$\sin 2\theta = 2 \sin \theta \cos \theta.$$

Finalmente, si $n > 1$, entonces de acuerdo a la relación de recurrencia de los polinomios de Chebyshev es

$$U_n(\cos \theta) = 2 \cos \theta \cdot U_{n-1}(\cos \theta) - U_{n-2}(\cos \theta)$$

y usando la hipótesis inductiva evidente vemos que esto es

$$= 2 \cos \theta \frac{\sin n\theta}{\sin \theta} - \frac{\sin(n-1)\theta}{\sin \theta} = \frac{2 \cos \theta \sin n\theta - \sin(n-1)\theta}{\sin \theta} = \frac{\sin(n+1)\theta}{\sin \theta}$$

porque la fórmula de adición para la función seno nos dice que

$$\sin(n \pm 1)\theta = \sin n\theta \cos \theta \mp \cos n\theta \sin \theta.$$

Una consecuencia inmediata de la igualdad (12) es que si $k \in \llbracket n \rrbracket$ entonces

$$\chi_{A_n}\left(2 \cos \frac{k\pi}{n+1}\right) = U_n\left(\cos \frac{k\pi}{n+1}\right) = \frac{\sin k\pi}{\sin(k\pi/(n+1))} = 0,$$

de manera que los n números

$$2 \cos \frac{k\pi}{n+1} \quad \text{con } k \in \llbracket n \rrbracket$$

son raíces de χ_{A_n} . Estos n números son distintos dos a dos — ya que la función \cos es estrictamente decreciente en el intervalo $[0, \pi]$ — así que el polinomio característico χ_{A_n} tiene sus n raíces en \mathbb{R} y son todas simples. La Proposición 5.4.10 nos dice entonces que la matriz A_n es diagonalizable.

Sea $f : \mathbb{R} \rightarrow \mathbb{R}$ una función cuatro veces diferenciable y supongamos que $K > 0$ es tal que $|f^{(4)}(x)| \leq K$ para todo $x \in \mathbb{R}$. El teorema de Taylor nos dice que para cada $x \in \mathbb{R}$ y $h > 0$ existen $\xi_+ \in [x, x+h]$ y $\xi_- \in [x-h, x]$ tales que

$$f(x+h) = f(x) + hf'(x) + \frac{h^2}{2}f''(x) + \frac{h^3}{6}f'''(x) + \frac{h^4}{24}f^{(4)}(\xi_+)$$

y

$$f(x-h) = f(x) - hf'(x) + \frac{h^2}{2}f''(x) - \frac{h^3}{6}f'''(x) + \frac{h^4}{24}f^{(4)}(\xi_-),$$

y entonces

$$\left| f''(x) - \frac{f(x+h) - 2f(x) + f(x-h)}{h^2} \right| \leq \frac{h^4}{24} |f^{(4)}(\xi_+) + f^{(4)}(\xi_-)| \leq \frac{h^4 K}{12}.$$

Esto nos dice que podemos aproximar $f''(x)$ por el cociente

$$\Delta_h f(x) := \frac{f(x+h) - 2f(x) + f(x-h)}{h^2}$$

con un error del orden de h^4 .

Supongamos ahora que $f(0) = f(1) = 0$, elijamos $n \in \mathbb{N}$ y $h := 1/(n+1)$, y pongamos $x_i := i/(n+1)$ para cada $i \in \llbracket 0, n+1 \rrbracket$. Podemos «discretizar» a la función f en el intervalo $[0, 1]$ vía el vector $y = (f(x_1), \dots, f(x_n))^t \in \mathbb{R}^n$ de sus valores en los puntos x_1, \dots, x_n ; omitimos aquí a $f(x_0)$ y a $f(x_{n+1})$ porque estos dos números son nulos. Hecho esto, el correspondiente vector de las derivadas $(f''(x_1), \dots, f''(x_n))$ puede aproximarse con un error del orden de $1/(n+1)^4$ por el vector $z = (z_1, \dots, z_n)$ que tiene

$$z_i = (n+1)^2 (f(x_{i-1}) - 2f(x_i) + f(x_{i+1}))$$

para cada $i \in \llbracket n \rrbracket$, de manera que

$$z = (n+1)^2 (A_n - 2I_n) \cdot y.$$

Vemos así que si V es el espacio de las funciones $[0, 1] \rightarrow \mathbb{R}$ cuatro veces diferenciables que se anulan en los extremos del intervalo, podemos *aproximar* para cada $n \in \mathbb{N}$ a la función lineal $D : f \in V \mapsto f'' \in V$ por la función lineal

$$d_n : y \in \mathbb{R}^n \mapsto (n+1)^2 (A_n - 2I_n) \cdot y \in \mathbb{R}^n.$$

Notemos que como la matriz A_n tiene n autovalores distintos, el Lema 5.1.7 nos dice que d_n también tiene n autovalores distintos y es, por lo tanto, diagonalizable. Saber esto es frecuentemente útil en las aplicaciones.

- (c) Fijemos otra vez un entero positivo $n \in \mathbb{N}$ y trabajemos sobre el cuerpo \mathbb{R} de los números reales. Si p es un elemento del espacio vectorial $\mathbb{R}[X]_{\leq n}$, entonces $xp' - p''$ también lo es y podemos por lo tanto considerar la función

$$H : p \in \mathbb{R}[X]_{\leq n} \mapsto xp' - p'' \in \mathbb{R}[X]_{\leq n},$$

que claramente lineal. Calculando, vemos inmediatamente que para cada $i \in \llbracket 0, n \rrbracket$ es

$$H(X^i) = iX^i - i(i-1)X^{i-2},$$

así que si la matriz de H con respecto a la base ordenada $\mathcal{B} = (1, X, \dots, X^n)$ de $\mathbb{R}[X]_{\leq n}$ es la matriz triangular superior

$$[H]_{\mathcal{B}}^{\mathcal{B}} = \begin{pmatrix} 0 & 0 & -2 & 0 & \cdots & 0 & 0 \\ 1 & 0 & -6 & \ddots & & \vdots & \\ 2 & 0 & \ddots & & 0 & & \\ 3 & \ddots & \ddots & -n(n-1) & & & \\ & \ddots & & 0 & & & \\ & & & & n & & \end{pmatrix}.$$

De acuerdo al Ejemplo 5.4.3(b), entonces, los autovalores de f son los números $0, 1, \dots, n$. Estos autovalores son $n+1$ en número y son distintos dos a dos, así que, como $\dim \mathbb{R}[X]_{\leq n} = n+1$, la Proposición 5.4.10 nos dice que la función H es diagonalizable.

Sea $k \in \llbracket 0, n \rrbracket$. La multiplicidad algebraica de k como autovalor de H es 1, así que la multiplicidad geométrica de k es exactamente 1: existe entonces un único autovector mónico $p \in \mathbb{k}[X]_{\leq n}$ de H de autovalor k . Si d es el grado de p , entonces $p = X^d + \sum_{i=0}^{d-1} c_i X^i$ para ciertos escalares $c_1, \dots, c_{d-1} \in \mathbb{R}$ y

$$kX^d + k \underbrace{\sum_{i=0}^{d-1} c_i X^i}_{kp} = kp = H(p) = dX^d - d(d-1)X^{d-2} + H\left(\sum_{i=0}^{d-1} c_i X^i\right).$$

La suma de los términos marcados en el último miembro de esta cadena de igualdades es un polinomio o nulo o de grado menor que d : el coeficiente de X^d en este último miembro es entonces d y, comparando con el primer miembro de la igualdad, vemos que tiene que ser $d = k$. Así, hemos probado que

para cada $k \in \llbracket 0, n \rrbracket$ la función lineal H tiene exactamente un autovector mónico de autovalor k y grado k .

Llamamos a ese autovector el **k -ésimo polinomio de Hermite**, por *Charles Hermite*, y lo escribimos He_k . De acuerdo a lo anterior, se trata de la única solución de la ecuación diferencial

n	He_n
0	1
1	X
2	$X^2 - 1$
3	$X^3 - 3X$
4	$X^4 - 6X^2 + 3$
5	$X^5 - 10X^3 + 15X$
6	$X^6 - 15X^4 + 45X^2 - 15$
7	$X^7 - 21X^5 + 105X^3 - 105X$
8	$X^8 - 28X^6 + 210X^4 - 420X^2 + 105$
9	$X^9 - 36X^7 + 378X^5 - 1260X^3 + 945X$
10	$X^{10} - 45X^8 + 630X^6 - 3150X^4 + 4725X^2 - 945$

Cuadro 5.2. Los primeros diez polinomios de Hermite.

ordinaria

$$p'' - Xp' + kp = 0$$

que es un polinomio mónico. En el Cuadro 5.2 están tabulados los diez primeros. \diamond

5.4.13. La consecuencia más importante de la Proposición 5.4.8 es que nos permite probar, bajo hipótesis razonables sobre el cuerpo \mathbb{k} , que un endomorfismo de un espacio vectorial o una matriz posee autovalores y autovalores.

Proposición. Supongamos que el cuerpo \mathbb{k} es algebraicamente cerrado. Si $f : V \rightarrow V$ es un endomorfismo de un espacio vectorial V de dimensión finita, entonces f posee al menos un autovalor.

Demostración. Como el cuerpo \mathbb{k} es algebraicamente cerrado, el polinomio característico χ_f , que es un elemento de $\mathbb{k}[X]$, posee una raíz en \mathbb{k} , esto es, existe $\lambda \in \mathbb{k}$ tal que $\chi_f(\lambda) = 0$ y, de acuerdo a la Proposición 5.4.8, esto implica que λ es un autovalor de f . \square

5.4.14. Que el cuerpo sea algebraicamente cerrado es, de hecho, una condición necesaria para que valga la conclusión de la Proposición 5.4.13 con total generalidad. En efecto, si \mathbb{k} no es algebraicamente cerrado, entonces existe un polinomio $p \in \mathbb{k}[X]$ que no tiene ninguna raíz en \mathbb{k} y la Proposición 5.4.8 nos dice que la matriz compañera $C(p)$ no tiene ningún autovalor, ya que, de acuerdo al ejemplo 5.4.3(d), su polinomio característico es precisamente p . Por ejemplo, podemos tomar $p = X^2 - 2$ si $\mathbb{k} = \mathbb{Q}$, $p = X^2 + 1$ si $\mathbb{k} = \mathbb{R}$, o $p = X^2 + X + 1$ si $\mathbb{k} = \mathbb{F}_2$.

5.4.15. A partir de la existencia de autovectores que nos da la Proposición 5.4.13 podemos probar que toda función lineal es «triangularizable»:

Proposición. Supongamos que el cuerpo \mathbb{k} es algebraicamente cerrado y sea V un espacio vectorial de dimensión finita. Si $f : V \rightarrow V$ es un endomorfismo de V , entonces existe una base ordenada \mathcal{B} de V tal que la matriz $[f]_{\mathcal{B}}^{\mathcal{B}}$ de f con respecto a \mathcal{B} es triangular superior.

Demostración. Sea $f : V \rightarrow V$ un endomorfismo de V . Procedemos haciendo inducción con respecto a la dimensión de V , notando que cuando $\dim V \leq 1$ no hay nada que probar.

Supongamos que V es un espacio vectorial de dimensión $n > 1$. La función transpuesta $f^t : V^* \rightarrow V^*$ es un endomorfismo de un espacio vectorial de dimensión finita, así que, como el cuerpo \mathbb{k} es algebraicamente cerrado, la Proposición 5.4.13 nos dice que f^t posee un autovalor $\lambda \in \mathbb{k}$ y, por lo tanto, que existe un autovector $\phi \in V^*$ de f^t de autovalor λ , de manera que $f^t(\phi) = \lambda\phi$.

La función lineal $\phi : V \rightarrow \mathbb{k}$ no es nula, porque es un autovector de f^t , así que es sobreyectiva y, en consecuencia, su núcleo $V' = \text{Nu}(\phi)$ es un subespacio de V de dimensión $n - 1$. Si $x \in V'$, entonces

$$\phi(f(v)) = f^t(\phi)(v) = (\lambda\phi)(v) = \lambda\phi(v) = 0,$$

de manera que $f(v)$ también es un elemento de V' . Esto significa que la función f se restringe a un endomorfismo $f_{V'} : V' \rightarrow V'$ de V' . Como $\dim V' = n - 1$, podemos suponer inductivamente que existe una base ordenada $\mathcal{B}' = (x_1, \dots, x_{n-1})$ de V' tal que la matriz $[f_{V'}]_{\mathcal{B}'}^{\mathcal{B}'}$ es triangular superior. Por otro lado, sabemos que podemos completar \mathcal{B}' a una base ordenada de V , esto es, que existe un vector $x_n \in V$ tal que $\mathcal{B} = (x_1, \dots, x_{n-1}, x_n)$ es una base ordenada de V . Existen escalares $b_1, \dots, b_n \in \mathbb{k}$ tales que $f(x_n) = b_1x_1 + \dots + b_nx_n$, y si consideramos la matriz

$$B = \begin{pmatrix} b_1 \\ \vdots \\ b_{n-1} \end{pmatrix} \in M_{n-1,1}(\mathbb{k}),$$

entonces la matriz de f con respecto a la base \mathcal{B} tiene una descomposición en bloques de la forma

$$[f]_{\mathcal{B}}^{\mathcal{B}} = \begin{pmatrix} [f_{V'}]_{\mathcal{B}'}^{\mathcal{B}'} & B \\ 0 & b_n \end{pmatrix}.$$

Como $[f_{V'}]_{\mathcal{B}'}^{\mathcal{B}'}$ es triangular superior, la matriz $[f]_{\mathcal{B}}^{\mathcal{B}}$ es también triangular superior y esto prueba lo que queremos. \square

5.4.16. Usando la Proposición 5.4.15 podemos obtener la siguiente descomposición de la traza y el determinante de una función lineal o de una matriz.

Corolario. Supongamos que \mathbb{k} es un cuerpo algebraicamente cerrado.

- (i) Sea V un espacio vectorial de dimensión finita y sea $f : V \rightarrow V$ un endomorfismo de V . La traza $\text{tr}(f)$ y el determinante $\det(f)$ de f son, respectivamente, la suma y el producto de los autovalores de f , cada uno de ellos tomados tantas veces como indica su multiplicidad algebraica.

- (ii) Sea $n \in \mathbb{N}$ y sea $A \in M_n(\mathbb{k})$. La traza $\text{tr}(A)$ y el determinante $\det(A)$ de A son, respectivamente, la suma y el producto de los autovalores de A , cada uno de ellos tomados tantas veces como indica su multiplicidad algebraica.

Demostración. (i) De acuerdo a la Proposición 5.4.15, hay una base ordenada \mathcal{B} de V tal que la matriz $[f]_{\mathcal{B}}$ es triangular superior. Si $N = \dim V$ y a_1, \dots, a_n son los escalares que aparecen en esa matriz a lo largo de la diagonal, entonces sabemos por un lado que

$$\chi_f = (X - a_1) \cdots (X - a_n),$$

así que los autovalores de f , tomados cada uno de ellos con su multiplicidad algebraica, son precisamente los escalares a_1, \dots, a_n , y, por otro, que

$$\text{tr}(f) = \text{tr}([f]_{\mathcal{B}}) = a_1 + \cdots + a_n, \quad \det(f) = \det([f]_{\mathcal{B}}) = a_1 \cdots a_n.$$

Es claro entonces que las afirmaciones del corolario son ciertas.

(ii) Considerando la función lineal $f_A : x \in \mathbb{k}^n \mapsto Ax \in \mathbb{k}^n$ es inmediato deducir la segunda parte del corolario de la primera. \square

5.4.17. Cuando el cuerpo \mathbb{k} no es algebraicamente cerrado, las afirmaciones del corolario 5.4.16 claramente no pueden ser ciertas en general, ya que hay funciones lineales y matrices que no tienen ningún autovalor. De todas maneras, inclusive en esa situación tenemos el siguiente resultado:

Proposición.

- (i) Sea V un espacio vectorial de dimensión finita n y sea $f : V \rightarrow V$ un endomorfismo de V . Si $\chi_f = X^n + c_1X^{n-1} + \cdots + c_{n-1}X + c_n$ con $c_1, \dots, c_n \in \mathbb{k}$, entonces

$$c_1 = -\text{tr}(f), \quad c_n = (-1)^n \det(f).$$

- (ii) Sea $n \in \mathbb{N}$ y sea $A \in M_n(\mathbb{k})$. Si $\chi_A = X^n + c_1X^{n-1} + \cdots + c_{n-1}X + c_n$ con $c_1, \dots, c_n \in \mathbb{k}$, entonces

$$c_1 = -\text{tr}(A), \quad c_n = (-1)^n \det(A).$$

Demostración. (i) Sea \mathcal{B} una base ordenada de V y sea $[f]_{\mathcal{B}} = (a_{i,j})$ la matriz de f con respecto a \mathcal{B} . Si denotamos $(b_{i,j})$ a la matriz $X \cdot I_n - [f]_{\mathcal{B}}$, cuyas entradas son polinomios de $\mathbb{k}[X]$, la fórmula de Leibniz 4.5.1 nos dice que

$$\chi_f = \det(X \cdot I_n - [f]_{\mathcal{B}}) = \sum_{\sigma \in S_n} \text{sgn}(\sigma) b_{\sigma(1),1} \cdots b_{\sigma(n),n}. \quad (13)$$

Ahora bien, para cada $i, j \in \llbracket n \rrbracket$ es

$$b_{i,j} = \begin{cases} X - a_{i,j}, & \text{si } i = j; \\ -a_{i,j}, & \text{en caso contrario} \end{cases}$$

Así, el polinomio $b_{i,j}$ tiene grado 1 si $i = j$ y es 0 nulo o de grado 0 si $i \neq j$. Esto implica que para cada $\sigma \in S_n$ el término

$$\operatorname{sgn}(\sigma)b_{\sigma(1),1} \cdots b_{\sigma(n),n}$$

de la suma (13) correspondiente a σ es un polinomio que, si no es nulo, tiene grado igual a la cantidad de elementos del conjunto $F_\sigma = \{k \in \llbracket n \rrbracket : \sigma(k) = k\}$ y es fácil ver que este conjunto F_σ

- tiene a lo sumo n elementos,
- tiene n elementos si y solamente si σ es la permutación identidad id_n de $\llbracket n \rrbracket$, y
- si tiene menos que n elementos entonces tiene a lo sumo $n - 2$.

Escribiendo la suma (13) en la forma

$$\chi_f = b_{1,1} \cdots b_{n,n} + \sum_{\substack{\sigma \in S_n \\ \sigma \neq \operatorname{id}_n}} \operatorname{sgn}(\sigma)b_{\sigma(1),1} \cdots b_{\sigma(n),n},$$

resulta entonces que el segundo sumando en el miembro izquierdo de la igualdad es 0 nulo o que tiene grado estrictamente menor que $n - 1$, mientras que el primero tiene grado exactamente igual a n . Esto implica, por supuesto, que el coeficiente de X^{n-1} en χ_f coincide con el de X^{n-1} en el producto

$$b_{1,1} \cdots b_{n,n} = (X - a_{1,1}) \cdots (X - a_{n,n}),$$

que es $-(a_{1,1} + \cdots + a_{n,n})$: esto nos dice que el coeficiente con el que aparece X^{n-1} en él es precisamente $-\operatorname{tr}(f)$.

Por otro lado, el coeficiente c_0 es el valor $\chi_f(0)$ del polinomio característico χ_f en 0 y, como sabemos, es igual al determinante de la matriz que se obtiene de $X \cdot I_n - [f]_{\mathcal{B}}$ en 0, es decir, de la matriz $-[f]_{\mathcal{B}}$. Claramente, entonces, se tiene que $c_0 = (-1)^n \det(f)$. \square

§5. Homomorfismos de álgebras e ideales

5.5.1. En la sección siguiente necesitaremos tres resultados sencillos sobre álgebras de polinomios. Empezamos por la siguiente proposición, que nos da una forma de construir homomorfismos de álgebras $\mathbb{k}[X] \rightarrow A$ con dominio en un álgebra de polinomios:

Proposición. *Sea A un álgebra. Si $a \in A$, entonces existe exactamente un homomorfismo de álgebras $\varepsilon_a : \mathbb{k}[X] \rightarrow A$ tal que $\varepsilon_a(X) = a$.*

Observemos que si $f : \mathbb{k}[X] \rightarrow A$ es un homomorfismo de álgebras cualquiera, entonces la unicidad que afirma la proposición implica inmediatamente que $f = \varepsilon_{f(X)}$: vemos así que *todos* los homomorfismos de álgebras $\mathbb{k}[X] \rightarrow A$ son como los que se obtienen en la proposición.

Demostración. Sea $a \in A$. Como el conjunto $\{X^i : i \in \mathbb{N}_0\}$ es una base de $\mathbb{k}[X]$, sabemos que existe una función lineal $\varepsilon_a : \mathbb{k}[X] \rightarrow A$ tal que $\varepsilon_a(X^i) = a^i$ para cada $i \in \mathbb{N}_0$. Mostremos que ε_a es un homomorfismo de álgebras —esto probará la afirmación de existencia.

Que $\varepsilon_a(1_{\mathbb{k}[X]}) = 1_A$ es consecuencia de la definición de ε_a . Sean p y q dos elementos de $\mathbb{k}[X]$. Existen $m, n \in \mathbb{N}_0$ y escalares $\lambda_0, \dots, \lambda_m$ y μ_0, \dots, μ_n en \mathbb{k} tales que $p = \sum_{i=0}^m \lambda_i X^i$ y $q = \sum_{j=1}^n \mu_j X^j$, y entonces

$$p \cdot q = \sum_{i=0}^m \lambda_i X^i \cdot \sum_{j=0}^n \mu_j X^j = \sum_{k=0}^{m+n} \left(\sum_{\substack{0 \leq i \leq m \\ 0 \leq j \leq n \\ i+j=k}} \lambda_i \mu_j \right) X^k. \quad (14)$$

Por la forma en que definimos la función ε_a sabemos que $\varepsilon_a(p) = \sum_{i=0}^m \lambda_i a^i$ y $\varepsilon_a(q) = \sum_{j=0}^n \mu_j a^j$, y se sigue de esto que

$$\varepsilon_a(p) \cdot \varepsilon_a(q) = \sum_{i=0}^m \lambda_i a^i \cdot \sum_{j=0}^n \mu_j a^j = \sum_{k=0}^{m+n} \left(\sum_{\substack{0 \leq i \leq m \\ 0 \leq j \leq n \\ i+j=k}} \lambda_i \mu_j \right) a^k.$$

De manera similar, de la igualdad (14) vemos

$$\varepsilon_a(p \cdot q) = \sum_{k=0}^{m+n} \left(\sum_{\substack{0 \leq i \leq m \\ 0 \leq j \leq n \\ i+j=k}} \lambda_i \mu_j \right) a^k$$

y comparando esta expresión con la que obtuvimos para $\varepsilon_a(p) \cdot \varepsilon_a(q)$ es claro que estos dos elementos de A son iguales.

Supongamos ahora, para probar la unicidad, que $\varepsilon, \eta : \mathbb{k}[X] \rightarrow A$ son dos homomorfismos de álgebras y que $\varepsilon(X) = a = \eta(X)$. Es entonces $\varepsilon(1_{\mathbb{k}[X]}) = 1_A = \eta(1_{\mathbb{k}[X]})$ y, para cada $i \in \mathbb{N}$,

$$\varepsilon(X^i) = \varepsilon(X)^i = a^i = \eta(X)^i = \eta(X^i).$$

Vemos así que ε y η coinciden sobre todos los elementos de la base $\{X^i : i \in \mathbb{N}_0\}$ de $\mathbb{k}[X]$ y, por lo tanto, que son iguales, ya que son funciones lineales. \square

5.5.2. Cada vez que A es un álgebra asociativa y a un elemento de A , la Proposición 5.5.1 nos provee un homomorfismo de álgebras $\varepsilon_a : \mathbb{k}[X] \rightarrow A$ completamente determinado por la condición de que sea $\varepsilon_a(X) = a$. Si $p \in \mathbb{k}[X]$ escribimos siempre $p(a)$ en lugar de $\varepsilon_a(p)$ y decimos que $p(a)$ se obtiene *evaluando el polinomio p en a*.

5.5.3. Si $p \in \mathbb{k}[X]$ y $S \subseteq \mathbb{k}$ es un conjunto de escalares, escribimos $p(S)$ al conjunto $\{p(\lambda) : \lambda \in S\}$.

Proposición. Sea V un espacio vectorial, sea $f : V \rightarrow V$ un endomorfismo de V y sea $p \in \mathbb{k}[X]$.

(i) Si x es un autovector de f de autovalor λ , entonces x es un autovector de $p(f)$ de autovalor $p(\lambda)$.

(ii) Se tiene que $\text{Spec}(p(f)) \supseteq p(\text{Spec}(f))$ y si el cuerpo \mathbb{k} es algebraicamente cerrado entonces vale, de hecho, la igualdad.

Demostración. Si p es un polinomio constante, entonces $p(f)$ es un múltiplo escalar de id_V y todas las afirmaciones son inmediatas. Supongamos entonces que p no es constante, que $p = a_d X^d + \dots + a_0$, con $a_0, \dots, a_d \in \mathbb{k}$ y que $a_d \neq 0$, con lo que el grado de p es exactamente d .

(i) Si x es un autovector de f de autovector λ , entonces $f^i(x) = \lambda^i x$ para cada $i \in \mathbb{N}_0$ y

$$\begin{aligned} p(f)(x) &= (a_d f^d + \dots + a_1 f + a_0 \text{id}_V)(x) \\ &= a_d \lambda^d x + \dots + a_1 \lambda x + a_0 x \\ &= (a_d \lambda^d + \dots + a_1 \lambda + a_0)x \\ &= p(\lambda)x, \end{aligned}$$

de manera que x es un autovector de $p(f)$ de autovector $p(\lambda)$.

(ii) Que el conjunto $p(\text{Spec}(f))$ está contenido en $\text{Spec}(p(f))$ es consecuencia inmediata de la parte (i) de la proposición. Supongamos ahora que el cuerpo \mathbb{k} es algebraicamente cerrado y sea $\mu \in \text{Spec}(p(f))$. El polinomio $p - \mu$ se factoriza en $\mathbb{k}[X]$ como producto de factores lineales, de manera que existen $\alpha_0, \dots, \alpha_d \in \mathbb{k}$ tales que

$$p - \mu = \alpha_0(X - \alpha_1) \cdots (X - \alpha_d). \quad (15)$$

Observemos que como p no es constante, $\alpha_0 \neq 0$. Evaluando ambos lados de la igualdad en f vemos que

$$p(f) - \mu \text{id}_V = \alpha_0(f - \alpha_1 \text{id}_V) \cdots (f - \alpha_d \text{id}_V).$$

Como μ es un autovalor de $p(f)$, hay un vector no nulo $v \in V$ tal que

$$\alpha_0(f - \alpha_1 \text{id}_V) \cdots (f - \alpha_d \text{id}_V)(v) = (p(f) - \mu \text{id}_V)(v) = 0.$$

Como $\alpha_0 \neq 0$, el conjunto

$$I = \{i \in \llbracket d \rrbracket : (f - \alpha_i \text{id}_V) \cdots (f - \alpha_d \text{id}_V)(v) = 0\}$$

no es vacío. Sea $t = \max I$. Si $t < d$, entonces el vector $w = (f - \alpha_{t+1} \text{id}_V) \cdots (f - \alpha_d \text{id}_V)(v)$ no es nulo pero que $(f - \alpha_t \text{id}_V)(w)$ sí lo es y, por lo tanto, α_t es un autovalor de f . Si en cambio $t = d$, entonces $v \neq 0$ y $(f - \alpha_d)(v) = 0$, de manera que también en este caso α_t es un autovalor de f . Evaluando ahora ambos lados de la igualdad (15) en α_t vemos que $p(\alpha_t) - \mu = 0$, así que $\mu = p(\alpha_t) \in \{p(\lambda) : \lambda \in \text{Spec}(f)\}$. \square

5.5.4. El segundo resultado sobre álgebras de polinomios que necesitamos es un criterio simple para decidir si un subespacio de $\mathbb{k}[X]$ es un ideal.

Proposición. Un subespacio I de $\mathbb{k}[X]$ es un ideal si y solamente si cada vez que $f \in I$ se tiene que $Xf \in I$.

Demostración. Sea I un subespacio de $\mathbb{k}[X]$. Que la condición de la proposición es necesaria para que I sea un ideal es evidente. Probemos que también es suficiente.

Supongamos para ello que $Xf \in I$ siempre que $f \in I$ y sea $g \in \mathbb{k}[X]$ un polinomio. Una inducción evidente muestra, a partir de la hipótesis, que

$$X^i f \in I \text{ cualesquiera sean } i \in \mathbb{N}_0 \text{ y } f \in I. \quad (16)$$

Ahora bien, existen $d \in \mathbb{N}_0$ y $a_0, \dots, a_d \in \mathbb{k}$ tales que $g = a_0 + a_1 X + \dots + a_d X^d$, y entonces

$$fg = gf = a_0 f + a_1 Xf + \dots + a_d X^d f \in I,$$

porque vale (16) e I es un subespacio de $\mathbb{k}[X]$. Esto prueba lo que queremos. \square

5.5.5. Finalmente, terminamos esta sección de álgebras de polinomios con la siguiente descripción de todos los ideales del álgebra $\mathbb{k}[X]$.

Proposición. Si I es un ideal no nulo de $\mathbb{k}[X]$, entonces existe un único polinomio mónico $m \in I$ tal que para cada $p \in \mathbb{k}[X]$ se tiene que

$$p \text{ pertenece a } I \text{ si y solamente si } p \text{ es divisible por } m. \quad (17)$$

El grado de m es $\min\{\deg(p) : p \in I \setminus 0\}$ y m es el único elemento mónico de I de ese grado.

Llamamos a ese polinomio m el **generador mónico** del ideal I .

Demostración. Como hay elementos no nulos en I , tiene sentido considerar el número

$$d := \min\{\deg(p) : p \in I \setminus 0\},$$

ya que el conjunto del que estamos tomando el mínimo es un subconjunto no vacío de \mathbb{N}_0 , y existe en I un elemento m no nulo tal que $\deg(m) = d$. Más aún, podemos suponer que m es mónico: en efecto, si no lo es y $a \in \mathbb{k}$ es el coeficiente de X^d en m , entonces $a^{-1}m$ también es un elemento de I de grado d y sí es mónico, así que es suficiente reemplazar a m por $a^{-1}m$.

Sea $p \in \mathbb{k}[X]$ un polinomio. Si p es divisible por m , existe $q \in \mathbb{k}[X]$ tal que $p = mq$ y entonces p pertenece a I porque éste es un ideal de $\mathbb{k}[X]$. Para ver la recíproca, supongamos ahora que p pertenece a I . Como m no es nulo y tiene grado d , sabemos que existen polinomios s y r en $\mathbb{k}[X]$ tales que $p = sm + r$ y o bien $r = 0$ o bien $\deg(r) < d$. En particular, tenemos que $r = p - sm$ y, como I es un ideal, que $r \in I$. Esto implica —en vista de la forma en que elegimos el número d — que no puede ser que r no sea nulo, porque su grado sería en ese caso menor que d . Así, es $r = 0$, $p = sm$ y, en definitiva, m divide a p .

Supongamos ahora que \tilde{m} es otro elemento mónico de $\mathbb{k}[X]$ que tiene la propiedad (17) descripta en el enunciado. Como m y \tilde{m} tienen ambos esta propiedad y pertenecen a I , se dividen mutuamente y esto y el hecho de que ambos polinomios son monómicos implica que $\tilde{m} = m$.

Finalmente, si \tilde{m} es un elemento mónico de I de grado d , entonces $\tilde{m} - m$ es un elemento de I , porque I es un subespacio de $\mathbb{k}[X]$, que tiene grado menor que d : la forma en que elegimos a d implica entonces que $\tilde{m} - m = 0$, esto es, que $\tilde{m} = m$. \square

§6. El polinomio minimal

5.6.1. Con la preparación de la sección anterior estamos en condiciones de exhibir uno de los invariantes más importantes de un endomorfismo de un espacio vectorial de dimensión finita, su polinomio minimal.

Proposición. *Sea V un espacio vectorial de dimensión finita y sea $f : V \rightarrow V$ un endomorfismo de V .*

(i) *Existe un único polinomio mónico $m_f \in \mathbb{k}[X]$ tal que*

- $m_f(f) = 0$ en $\text{End}(V)$ y
- cada vez que $p \in \mathbb{k}[X]$ es tal que $p(f) = 0$ en $\text{End}(V)$, el polinomio m_f divide a p .

Si $I = \{p \in \mathbb{k}[X] : p(f) = 0\}$, entonces $I \neq 0$, el grado de m_f es

$$d := \min\{\deg(p) : p \in I \setminus 0\},$$

y m_f es el único elemento mónico de I que tiene grado d .

(ii) *El subespacio $S := \langle f^i : i \in \mathbb{N}_0 \rangle$ de $\text{End}(V)$ tiene dimensión d y $\{\text{id}_V, f, \dots, f^{d-1}\}$ es una base de S . Más aún, si $a_0, \dots, a_{d-1} \in \mathbb{k}$ son los escalares tales que $f^d = a_0\text{id}_V + a_1f + \dots + a_{d-1}f^{d-1}$, entonces*

$$m_f = X^d - a_{d-1}X^{d-1} - \dots - a_1X - a_0.$$

Hay enteros no negativos r tales que $f^r \in \langle \text{id}_V, f, \dots, f^{r-1} \rangle$ y d es el menor de ellos.

Llamamos al polinomio m_f el **polinomio minimal** de f . La segunda parte de esta proposición nos da una cota para su grado —implica que éste no es mayor que $\dim \text{End}(V) = (\dim V)^2$, ya que coincide con la dimensión de un subespacio de $\text{End}(V)$ — y una forma de determinar explícitamente el polinomio m_f .

Demostración. (i) Sea $\varepsilon_f : \mathbb{k}[X] \rightarrow \text{End}(V)$ el homomorfismo de álgebras tal que $\varepsilon_f(X) = f$. Su núcleo $\text{Nu}(\varepsilon_f)$, que es un ideal de $\mathbb{k}[X]$, claramente coincide con el subespacio I descripto en el enunciado de la proposición. Como $\text{End}(V)$ tiene dimensión finita y $\mathbb{k}[X]$ no, la función ε_f no

es inyectiva y el ideal I de $\mathbb{k}[X]$ no es nulo. De acuerdo a la Proposición 5.5.5 existe entonces un único polinomio mónico $m_f \in I$ tal que para cada $p \in \mathbb{k}[X]$ se tiene que $p \in I$ si y solamente si p es divisible por m_f , que es además el único elemento mónico de I de grado $\min\{\deg(p) : p \in I \setminus 0\}$. Esto significa que m_f tiene precisamente las propiedades descriptas en la parte (i) de la proposición y que es el único elemento de $\mathbb{k}[X]$ que las tiene.

(iii) Sean $a_0, \dots, a_{d-1} \in \mathbb{k}$ tales que $m_f = X^d - a_{d-1}X^{d-1} - \dots - a_1X - a_0$. Como $m_f(f) = 0$, vale que

$$f^d = a_0\text{id}_V + a_1f + \dots + a_{d-1}f^{d-1}. \quad (18)$$

Consideremos los subespacios $S = \langle f^i : i \in \mathbb{N}_0 \rangle$ y $F = \langle \text{id}_V, f, \dots, f^{d-1} \rangle$ de $\text{End}(V)$. Es claro que $F \subseteq S$. Mostremos que también $S \subseteq F$ y, para ello, que $f^i \in F$ para todo $i \in \mathbb{N}_0$. Si $i = 0$ esto es evidente. Supongamos entonces que $i \geq 1$ e, inductivamente, que $f^{i-1} \in F$, de manera que existen escalares $b_0, \dots, b_{d-1} \in \mathbb{k}$ tales que $f^{i-1} = b_0\text{id}_V + b_1f + \dots + b_{d-1}f^{d-1}$. Componiendo a ambos lados de esta igualdad con f y usando (18), vemos entonces que

$$\begin{aligned} f^i &= b_0f + b_1f^2 + \dots + b_{d-2}f^{d-1} + b_{d-1}f^d \\ &= b_0f + b_1f^2 + \dots + b_{d-2}f^{d-1} + b_{d-1}(a_0\text{id}_V + a_1f + \dots + a_{d-1}f^{d-1}) \in F, \end{aligned}$$

como queremos. Concluimos con esto que $S = F$.

Mostremos ahora que los d homomorfismos $\text{id}_V, f, \dots, f^{d-1}$ son elementos linealmente independientes de $\text{End}(V)$, de manera que el conjunto $\mathcal{B} = \{\text{id}_V, f, \dots, f^{d-1}\}$ es una base de F con d elementos y que entonces se tiene que $\dim S = \dim F = d$. Sean $b_0, \dots, b_{d-1} \in \mathbb{k}$ escalares tales que $b_0\text{id}_V + b_1f + \dots + b_{d-1}f^{d-1} = 0$. El polinomio $p = b_0 + b_1X + \dots + b_{d-1}X^{d-1} \in \mathbb{k}[X]$ es tal que $p(f) = 0$ y por lo tanto es, de acuerdo a la parte (i) de la proposición, divisible por m_f : como su grado es menor que el de m_f , esto implica que $p = 0$, esto es, que $b_0 = \dots = b_{d-1} = 0$, y esto prueba lo que queremos.

Lo que ya hicimos nos dice que $f^d \in S = F = \langle \text{id}_V, f, \dots, f^{d-1} \rangle$. Finalmente, si r es un entero no negativo tal que $f^r \in \langle \text{id}_V, f, \dots, f^{d-1} \rangle$, entonces existen escalares $c_0, \dots, c_{r-1} \in \mathbb{k}$ tales que $f^r = a_0\text{id}_V + a_1f + \dots + a_{r-1}f^{r-1}$ y el polinomio $q = c_0 + c_1X + \dots + c_{r-1}X^{r-1} - X^r$ es tal que $q(f) = 0$: la parte (i) de la proposición nos dice entonces que m_f divide a q y, como $q \neq 0$, que $r \geq d$. \square

5.6.2. Destaquemos dos situaciones extremas que pueden ocurrir en la Proposición 5.6.1:

- Si V es un espacio nulo y $f : V \rightarrow V$ es un endomorfismo —necesariamente la función nula, claro— entonces el polinomio minimal de f es $m_f = 1$, que tiene grado nulo.
- Por otro lado, si V es un espacio de dimensión finita y positiva y $f : V \rightarrow V$ es el endomorfismo nulo, entonces $m_f = X$, de grado 1.

5.6.3. Ejemplo. Es importante observar que la hipótesis de que el espacio V tiene dimensión finita que aparece en la Proposición 5.6.1 es necesaria para obtener la conclusión. Por ejemplo, si

$$f : q \in \mathbb{k}[X] \mapsto Xq \in \mathbb{k}[X]$$

es el endomorfismo del espacio vectorial $\mathbb{k}[X]$ dado por la multiplicación por X , entonces para todo $p \in \mathbb{k}[X]$ no nulo se tiene que $p(f) \neq 0$: es fácil verificar que $p(f)(1) = p$. Esto implica que el homomorfismo de álgebras $\varepsilon_f : \mathbb{k}[X] \rightarrow \text{End}(\mathbb{k}[X])$ es inyectivo en este caso y la demostración que hicimos no puede ni comenzar: de hecho, en este ejemplo no hay *ningún* polinomio mónico m tal que $m(f) = 0$. \diamond

5.6.4. Como siempre, hay una versión de la Proposición 5.6.1, que se ocupa de endomorfismos, para matrices. Podríamos repetir esencialmente la misma prueba que hicimos para 5.6.1, pero, para variar un poco, la obtendremos directamente de esa proposición.

Empezamos con la siguiente observación:

Lema. Sea $n \in \mathbb{N}$ y para cada matriz A en $M_n(\mathbb{k})$ consideremos la función $f_A : x \in \mathbb{k}^n \mapsto Ax \in \mathbb{k}^n$. Si $p \in \mathbb{k}[X]$ y $A \in M_n(\mathbb{k})$, entonces $f_{p(A)} = p(f_A)$.

Demostración. Sean $p \in \mathbb{k}[X]$ y $A \in M_n(\mathbb{k})$. Sabemos de la Proposición 2.6.8 que la función

$$\rho : B \in M_n(\mathbb{k}) \mapsto f_B \in \text{End}(\mathbb{k}^n)$$

es un isomorfismo de álgebras. Por otro lado, de la Proposición 5.5.1 sabemos que hay homomorfismos de álgebras $\varepsilon_A : \mathbb{k}[X] \rightarrow M_n(\mathbb{k})$ y $\varepsilon_{f_A} : \mathbb{k}[X] \rightarrow \text{End}(\mathbb{k}^n)$ bien determinados por la condición de que $\varepsilon_A(X) = X$ y $\varepsilon_{f_A}(X) = f_A$. Ahora bien, la composición $\rho \circ \varepsilon_A : \mathbb{k}[X] \rightarrow \text{End}(\mathbb{k}^n)$ es también un homomorfismo de álgebras y se tiene que

$$(\rho \circ \varepsilon_A)(X) = \rho(\varepsilon_A(X)) = \rho(A) = f_A$$

así que, de hecho, es $\rho \circ \varepsilon_A = \varepsilon_{f_A}$. Usando esto vemos que

$$f_{p(A)} = \rho(p(A)) = \rho(\varepsilon_A(p)) = \varepsilon_{f_A}(p) = p(f_A),$$

como afirma el lema. \square

5.6.5. El análogo de la Proposición 5.6.1 para matrices es el siguiente:

Proposición. Sea $n \in \mathbb{N}$ y sea $A \in M_n(\mathbb{k})$.

(i) Existe un único polinomio mónico $m_A \in \mathbb{k}[X]$ tal que

- $m_A(A) = 0$ en $M_n(\mathbb{k})$ y
- cada vez que $p \in \mathbb{k}[X]$ es tal que $p(A) = 0$ en $M_n(\mathbb{k})$, el polinomio m_A divide a p .

Si $I = \{p \in \mathbb{k}[X] : p(A) = 0\}$, entonces el grado de m_A es

$$d := \min\{\deg(p) : p \in I \setminus 0\},$$

y m_A es el único elemento mónico de I que tiene grado d . Finalmente, el polinomio m_A coincide con el polinomio minimal del endomorfismo $f_A : x \in \mathbb{k}^n \mapsto Ax \in \mathbb{k}^n$.

(ii) El subespacio $S := \langle A^i : i \in \mathbb{N}_0 \rangle$ de $M_n(\mathbb{k})$ tiene dimensión d y $\{I_n, A, \dots, A^{d-1}\}$ es una base de S . Más aún, si $a_0, \dots, a_{d-1} \in \mathbb{k}$ son los escalares tales que $A^d = a_0 I_n + a_1 A + \dots + a_{d-1} A^{d-1}$,

entonces

$$m_A = X^d - a_{d-1}X^{d-1} - \cdots - a_1X - a_0.$$

Hay enteros no negativos r tales que $A^r \in \langle I, A, \dots, A^{r-1} \rangle$ y d es el menor de ellos.

Demostración. Para cada $B \in M_n(\mathbb{k})$ sea $f_B : x \in \mathbb{k}^n \mapsto Bx \in \mathbb{k}^n$ y consideremos la función $\rho : B \in M_n(\mathbb{k}) \mapsto f_B \in \text{End}(\mathbb{k}^n)$, que es un isomorfismo de álgebras. Los subespacios

$$I(A) := \{p \in \mathbb{k}[X] : p(A) = 0\}, \quad I(f_A) := I(A) := \{p \in \mathbb{k}[X] : p(f_A) = 0\}.$$

de $\mathbb{k}[X]$ coinciden: en efecto, si $p \in \mathbb{k}[X]$, entonces $p(A) = 0$ si y solamente si

$$p(f_A) = f_{p(A)} = \rho(p(A)) = 0,$$

ya que la función ρ es un isomorfismo y vale el Lema 5.6.4.

El polinomio minimal m_{f_A} es el único elemento mónico de $I(f_A)$ que divide a todos los elementos de $I(f_A)$, y es el único elemento mónico de $I(f_A)$ de grado $\min\{\deg(p) : p \in I(f_A) \setminus 0\}$. Como $I(A) = I(f_A)$, esto nos dice inmediatamente que si ponemos $m_A := m_{f_A}$, entonces m_A tiene todas las propiedades descriptas en (i).

(ii) Sabemos que el subespacio $S(f_A) := \{f_A^i : i \in \mathbb{N}_0\}$ de $\text{End}(\mathbb{k}^n)$ tiene dimensión d y que tiene al conjunto $\mathcal{B}(f_A) = \{\text{id}_{\mathbb{k}^n}, f_A, \dots, f_A^{d-1}\}$ como base. Si ponemos $S(A) = \langle A^i : i \in \mathbb{N}_0 \rangle$ y $\mathcal{B}(A) = \{I_n, A, \dots, A^{d-1}\}$, entonces $\rho(S(A)) = S(f_A)$ y $\rho(\mathcal{B}(A)) = \mathcal{B}(f_A)$, ya que $f_A^i = f_{A^i}$ para todo $i \in \mathbb{N}_0$: como la función ρ es un isomorfismo de espacios vectoriales, esto nos dice que $\dim S(A) = d$ y que $\mathcal{B}(A)$ es una base de $S(A)$.

Sean a_0, \dots, a_{d-1} los escalares de \mathbb{k} tales que $A^d = a_0I_n + a_1A + \cdots + a_{d-1}A^{d-1}$. Aplicando la función ρ a cada lado de esta igualdad vemos que $f_A^d = a_0\text{id}_{\mathbb{k}^n} + a_1f_A + \cdots + a_{d-1}f_A^{d-1}$ y, por lo tanto, de acuerdo a la segunda parte de la Proposición 5.6.1, $m_A = m_{f_A} = X^d - a_{d-1}X^{d-1} - \cdots - a_1X - a_0$.

Finalmente, si r es un entero no negativo, entonces $A^r \in \langle I, A, \dots, A^{r-1} \rangle$ si y solamente si $f_A^r = \rho(A^r) \in \rho(\langle I, A, \dots, A^{r-1} \rangle) = \langle \text{id}_{\mathbb{k}^n}, f_A, \dots, f_A^{r-1} \rangle$, y sabemos que esto ocurre si y solamente si $r \geq d$. Con esto queda completa la prueba de la proposición. \square

5.6.6. La segunda parte de la Proposición 5.6.5 nos da un procedimiento efectivo para determinar el polinomio minimal de una matriz. Si $A \in M_n(\mathbb{k})$, entonces tenemos que buscar el menor entero no negativo d tal que $A^d \in \langle I_n, A, \dots, A^{d-1} \rangle$. Hecho eso, podemos buscar los escalares $a_0, \dots, a_d \in \mathbb{k}$, tales que $A^d = a_0I_n + a_1A + \cdots + a_{d-1}A^{d-1}$ y concluir que el polinomio minimal de A es

$$m_A = X^d - a_{d-1}X^{d-1} - \cdots - a_1X - a_0.$$

5.6.7. Ejemplo. Consideremos la matriz $A = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 2 \end{pmatrix}$. Calculando explícitamente vemos que las

primeras potencias de A son

$$\begin{array}{ccccc} I_4 & A & A^2 & A^3 & A^4 \\ \hline \left(\begin{array}{cccc} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{array} \right) & \left(\begin{array}{cccc} 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 2 \end{array} \right) & \left(\begin{array}{cccc} 1 & 2 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 4 \end{array} \right) & \left(\begin{array}{cccc} 1 & 3 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 7 \\ 0 & 0 & 0 & 8 \end{array} \right) & \left(\begin{array}{cccc} 1 & 4 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 15 \\ 0 & 0 & 0 & 16 \end{array} \right) \end{array}$$

y no es difícil ver que las primeras tres son linealmente independientes y que

$$A^3 = 4A^2 - 5A + 2I_n.$$

Se sigue de esto, como dijimos arriba, que el polinomio minimal de A es

$$m_A = X^3 - 4X^2 + 5X - 2.$$

◇

Polinomios minimales puntuales

5.6.8. El polinomio minimal de un endomorfismo f de un espacio vectorial V de dimensión finita es, como vimos, el polinomio mónico p de menor grado tal que $p(f) = 0$. De una forma similar, podemos construir para cada vector x de V un polinomio caracterizado por ser el polinomio mónico p de menor grado tal que $p(f)$ se anula en x .

5.6.9. Si $f : V \rightarrow V$ es un endomorfismo de un espacio vectorial V y W es un subespacio de V , decimos que W es **f -invariante** si $f(w) \in W$ para todo $w \in W$, esto es, si $f(W) \subseteq W$. Cuando ese es el caso, tiene sentido considerar la función $f_W : w \in W \mapsto f(w) \in W$, que claramente es lineal: la llamamos la **restricción** de f a W .

Lema. Sea V un espacio vectorial.

- (i) Si $f, g : V \rightarrow V$ son dos endomorfismos de V y W es un subespacio de V que es f -y g -invariante, entonces para cada $a, b \in \mathbb{k}$ el subespacio W es $(af + bg)$ -y $(f \circ g)$ -invariante, y

$$(af + bg)_W = af_W + bg_W, \quad (f \circ g)_W = f_W \circ g_W.$$

- (ii) Si $f : V \rightarrow V$ un endomorfismo de V y W es un subespacio f -invariante de V , entonces para todo $p \in \mathbb{k}[X]$ el subespacio W es $p(f)$ -invariante y

$$p(f)_W = p(f_W).$$

Demostración. (i) Sean $f, g : V \rightarrow W$ dos endomorfismos de V , sea W un subespacio de V que es f -y g -invariante, y sean $a, b \in \mathbb{k}$. Si $w \in W$, entonces $(af + bg)(w) = af(w) + bg(w) \in W$, ya que $f(w)$ y $g(w)$ están en W : esto nos dice que W es $(af + bg)$ -invariante. Más aún, para todo $w \in W$ tenemos que

$$\begin{aligned} (af + bg)_W(w) &= (af + bg)(w) = af(w) + bg(w) = af_W(w) + bg_W(w) \\ &= (af_W + bg_W)(w), \end{aligned}$$

así que $(af + bg)_W = af_W + bg_W$.

En segundo lugar, si $w \in W$, entonces $(f \circ g)(w) = f(g(w)) \in f(W) \subseteq W$ porque W es f -y g -invariante, así que W es $(f \circ g)$ -invariante, y

$$(f \circ g)_W(w) = (f \circ g)(w) = f(g(w))f(g_W(w)) = f_W(g_W(w)) = (f_W \circ g_W)(w),$$

de manera que $(f \circ g)_W = f_W \circ g_W$.

(ii) Observemos primero que para todo $i \in \mathbb{N}_0$ el subespacio W es f^i -invariante. En efecto, esto es evidente si $i = 0$ y si suponemos que $i \in \mathbb{N}_0$ es tal que $f^i(W) \subseteq W$, entonces

$$f^{i+1}(W) = f(f^i(W)) \subseteq f(W) \subseteq W,$$

así que nuestra afirmación sigue por inducción.

Sea ahora $p \in \mathbb{k}[x]$ y sean $n \in \mathbb{N}_0$ y $a_0, \dots, a_n \in \mathbb{k}$ tales que $p = a_n X^n + \dots + a_1 X + a_0$. Si $w \in W$, entonces

$$p(f)(w) = (a_n f^n + \dots + a_1 f + a_0 \text{id}_V)(w) = a_n f^n(w) + \dots + a_1 f(w) + a_0 w \in W,$$

ya que $f^i(w) \in f^i(W) \subseteq W$ para todo $i \in \mathbb{N}_0$. Esto muestra que W es $p(f)$ -invariante y, en particular, que tenemos la restricción $p(f)_W : W \rightarrow W$.

Recordemos que la función $\Psi : p \in \mathbb{k}[X] \mapsto p(f) \in \text{End}(V)$ es un homomorfismo de álgebras y mostremos que la función

$$\Phi : p \in \mathbb{k}[X] \mapsto p(f)_W \in \text{End}(W)$$

también lo es. Para verlo, sean primero $p, q \in \mathbb{k}[X]$ y $a, b \in \mathbb{k}$: calculamos que

$$\begin{aligned} \Phi(ap + bq) &= (ap + bq)(f)_W = (ap(f) + bq(f))_W = ap(f)_W + bq(f)_W \\ &= a\Phi(p) + b\Phi(q), \end{aligned}$$

y esto nos dice que la función Φ es lineal. Por otro lado,

$$\Phi(1) = (1(f))_W = (\text{id}_V)_W = \text{id}_W$$

y si p y q están en $\mathbb{k}[X]$, entonces

$$\Phi(pq) = (pq)(f)_W = (p(f) \circ q(f))_W = p(f)_W \circ q(f)_W.$$

Ahora bien, el morfismo de álgebras Φ tiene $\Phi(X) = (X(f))_W = f_W$, así que coincide con el morfismo de evaluación

$$\varepsilon_{f_W} : p \in \mathbb{k}[X] \mapsto p(f_W) \in \text{End}(W).$$

Esto significa precisamente que para todo $p \in \mathbb{k}[X]$ se tiene que $p(f)_W = p(f_W)$, como afirma el lema. \square

5.6.10. Proposición. Sea $f : V \rightarrow V$ un endomorfismo de un espacio vectorial V y sea $x \in V$.

(i) El subespacio $\langle x \rangle_f := \{f^i(x) : i \in \mathbb{N}_0\}$ es f -invariante, contiene a x y está contenido en todo subespacio f -invariante de V que contiene a x . Más aún, es

$$\langle x \rangle_f = \{p(f)(x) : p \in \mathbb{k}[X]\}.$$

(ii) Si escribimos $f_x : \langle x \rangle_f \rightarrow \langle x \rangle_f$ a la restricción $f_{\langle x \rangle_f}$ de f a $\langle x \rangle_f$, entonces para cada $p \in \mathbb{k}[X]$ se tiene que

$$p(f_x) = 0 \iff p(f)(x) = 0.$$

Llamamos al subespacio $\langle x \rangle_f$ de V el **subespacio f -cíclico generado por x en V** .

Demostración. (i) Para ver que el subespacio $\langle x \rangle_f$ es f -invariante es suficiente con observar que $f(f^i(x)) \in \langle x \rangle_f$ para todo $i \in \mathbb{N}_0$. Es evidente que $x \in \langle x \rangle_f$, y si W es un subespacio f -invariante de V que contiene a x , entonces claramente $f^i(x) \in W$ para todo $i \in \mathbb{N}_0$ y, por lo tanto, $\langle x \rangle_f \subseteq W$. Finalmente,

(ii) Sea $W := \langle x \rangle_f$, sea $f_W : W \rightarrow W$ la restricción de f a W y sea $p \in \mathbb{k}[X]$. Como $x \in W$, sabemos que $p(f)(x) = p(f)_W(x) = p(f_W)(x)$, así que si $p(f_W) = 0$ tenemos que $p(f)(x) = 0$. Probemos la implicación recíproca.

Supongamos que $p(f)(x) = 0$. Para ver que $p(f_W) = 0$ es suficiente que mostremos que $p(f_W)(f^i(x)) = 0$ para cada $i \in \mathbb{N}_0$, ya que $\{f^i(x) : i \in \mathbb{N}_0\}$ genera el dominio de $p(f_W)$. Sea entonces $i \in \mathbb{N}_0$. Si $q = X^i$, entonces como $f^i(x) \in W$, es

$$\begin{aligned} p(f_W)(f^i(x)) &= p(f)_W(f^i(x)) = p(f)(f^i(x)) = p(f)(q(f)(x)) = (p(f) \circ q(f))(x) \\ &= (pq)(f)(x) = (qp)(f)(x) = (q(f) \circ p(f))(x) = q(f)(p(f)(x)) = 0. \end{aligned}$$

Esto completa la prueba. \square

5.6.11. Proposición. Sea V un espacio vectorial, sea $f : V \rightarrow V$ un endomorfismo de V , sea $x \in V$, y sea $f_x : \langle x \rangle_f \rightarrow \langle x \rangle_f$ la restricción de f a $\langle x \rangle_f$. Si el espacio f -cíclico $\langle x \rangle_f$ tiene dimensión finita, entonces existe un único polinomio mónico $m_{f,x} \in \mathbb{k}[X]$ tal que

- $m_{f,x}(f)(x) = 0$ en V y
- cada vez que $p \in \mathbb{k}[X]$ es tal que $p(f)(x) = 0$ en V , el polinomio $m_{f,x}$ divide a p .

Más aún, en ese caso el grado de $m_{f,x}$ coincide con la dimensión de $\langle x \rangle_f$, y si d es ese grado, entonces $\mathcal{B} = (x, f(x), \dots, f^{d-1}(x))$ es una base ordenada de $\langle x \rangle_f$ tal que

$$[f_x]_{\mathcal{B}}^{\mathcal{B}} = C(m_{f,x}),$$

la matriz compañera del polinomio $m_{f,x}$. Finalmente, si $a_0, \dots, a_{d-1} \in \mathbb{k}$ son los escalares tales que $f^d(x) = a_0x + a_1f(x) + \dots + a_{d-1}f^{d-1}(x)$, entonces

$$m_{f,x} = X^x - a_{d-1}X^{d-1} - \dots - a_1X - a_0.$$

Llamamos a $m_{f,x}$ el *polinomio minimal de f relativo a x* .

Demuestra. Supongamos que el subespacio f -cíclico $\langle x \rangle_f$ tiene dimensión finita. De acuerdo a la Proposición 5.6.1, existe un único polinomio mónico m en $\mathbb{k}[X]$ tal que $m(f_x) = 0$ y que divide a todo polinomio $p \in \mathbb{k}[X]$ tal que $p(f_x) = 0$. En vista de la segunda parte de la Proposición 5.6.10, este polinomio m es el único elemento mónico de $\mathbb{k}[X]$ tal que $m(f)(x) = 0$ y que divide a todo polinomio $p \in \mathbb{k}[X]$ tal que $p(f)(x) = 0$.

Sea d el grado de m y sean $a_0, \dots, a_{d-1} \in \mathbb{k}$ tales que $m = X^d - a_{d-1}X^{d-1} - \dots - a_1X - a_0$. Como $m(f)(x) = 0$, tenemos que

$$f^d(x) = a_0x + a_1f(x) + \dots + a_{d-1}f^{d-1}(x).$$

Sea $S = \langle x, f(x), \dots, f^{d-1}(x) \rangle$. Afirmamos que $\langle x \rangle_f = S$. Como claramente S está contenido en $\langle x \rangle_f$, es suficiente probar la inclusión recíproca y, para ello, que $f^i(x) \in S$ para todo $i \in \mathbb{N}_0$. Esto es claro si $i = 0$, y si suponemos que $i \in \mathbb{N}_0$ es tal que $f^i(x) \in S$, entonces hay escalares $c_0, \dots, c_{d-1} \in \mathbb{k}$ tales que $f^i(x) = c_0x + c_1f(x) + \dots + c_{d-1}f^{d-1}(x)$ y, por lo tanto,

$$\begin{aligned} f^{i+1}(x) &= c_0f(x) + \dots + c_{d-2}f^{d-2}(x) + c_{d-1}f^d(x) \\ &= c_0f(x) + \dots + c_{d-2}f^{d-2}(x) + c_{d-1}(a_0x + a_1f(x) + \dots + a_{d-1}f^{d-1}(x)) \in S, \end{aligned}$$

lo que completa la inducción.

Veamos ahora que los vectores $x, f(x), \dots, f^{d-1}(x)$ son linealmente independientes: esto implicará que $\mathcal{B} = (x, f(x), \dots, f^{d-1}(x))$ es una base ordenada de $\langle x \rangle_f$ y, por lo tanto, que $\dim \langle x \rangle_f = d = \deg(m)$. Sean $b_0, \dots, b_{d-1} \in \mathbb{k}$ tales que $b_0x + b_1f(x) + \dots + b_{d-1}f^{d-1}(x) = 0$. El polinomio $p = b_0 + b_1X + \dots + b_{d-1}X^{d-1}$ es tal que $p(f)(x) = b_0x + b_1f(x) + \dots + b_{d-1}f^{d-1}(x) = 0$, así que m lo divide: como su grado es estrictamente menor que el de m , esto nos dice que $p = 0$ y, por lo tanto, que $b_0 = \dots = b_{d-1} = 0$. Esto prueba lo que queríamos.

Finalmente, como tenemos que

$$f_x(f^i(x)) = \begin{cases} f^{i+1}(x) & \text{si } 0 \leq i < d-1; \\ a_0x + a_1f(x) + \dots + a_{d-1}f^{d-1}(x) & \text{si } i = d-1; \end{cases}$$

la matriz de $f_x : \langle x \rangle_f \rightarrow \langle x \rangle_f$ es

$$[f_x]_{\mathcal{B}}^{\mathcal{B}} = \begin{pmatrix} 0 & 0 & 0 & \cdots & 0 & -a_0 \\ 1 & 0 & 0 & \cdots & 0 & -a_1 \\ 0 & 1 & 0 & \cdots & 0 & -a_2 \\ \vdots & \vdots & \ddots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \ddots & 0 & -a_{d-2} \\ 0 & 0 & 0 & \cdots & 1 & -a_{d-1} \end{pmatrix} = C(m_{f,x}),$$

la matriz compañera de $m_{f,x}$. □

5.6.12. Recordemos que decimos que un polinomio p de $\mathbb{k}[X]$ es el *mínimo común múltiplo* de una familia de polinomios $(p_i)_{i \in I}$ si

- p es mónico,
- p es divisible por p_i para cada $i \in I$, y
- si un polinomio $q \in \mathbb{k}[X]$ es divisible por p_i para todo $i \in I$, entonces q es divisible por p .

Toda familia finita de elementos de $\mathbb{k}[X]$ tiene un mínimo común múltiplo, pero hay familias infinitas de polinomios que no admiten uno: por ejemplo, los polinomios de la familia $\{X^i : i \in \mathbb{N}_0\}$ no poseen ni siquiera un múltiplo común.

5.6.13. Los polinomios minimales puntuales de un endomorfismo nos permiten «aproximar» al polinomio minimal de ese endomorfismo, en el sentido preciso de la siguiente proposición:

Proposición. *Sea $f : V \rightarrow V$ un endomorfismo de un espacio vectorial V de dimensión finita.*

- (i) *El polinomio minimal m_f es el mínimo común múltiplo de los polinomios $m_{f,x}$ con $x \in V$.*
- (ii) *Si V_1, \dots, V_n son subespacios f -invariantes de V y $V = V_1 + \dots + V_n$, entonces m_f es el mínimo común múltiplo de los polinomios minimales $m_{f_{V_1}}, \dots, m_{f_{V_n}}$ de las restricciones f_{V_1}, \dots, f_{V_n} .*
- (iii) *Si x_1, \dots, x_n son vectores de V tales que $V = \langle x_1 \rangle_f + \dots + \langle x_n \rangle_f$, entonces m_f es el mínimo común múltiplo de los polinomios $m_{f,x_1}, \dots, m_{f,x_n}$.*

Demostración. (i) Si x es un vector de V , entonces $m_f(f_x)(x) = m_f(f)_{\langle x \rangle_f}(x) = 0$, así que $m_{f,x}$ divide a m_f . Por otro lado, supongamos que $p \in \mathbb{k}[x]$ es un polinomio que es divisible por $m_{f,x}$ para cada $x \in V$. En ese caso, para cada $x \in V$ existe $q \in \mathbb{k}[X]$ tal que $p = qm_{f,x}$ y entonces

$$p(f)(x) = q(f)(m_{f,x}(f)(x)) = 0.$$

Vemos así que $p(f) = 0$ y la minimalidad de m_f implica que m_f divide a p .

(ii) Sean V_1, \dots, V_n subespacios f -invariantes de V y supongamos que $V = V_1 + \dots + V_n$. Si $i \in \llbracket n \rrbracket$, sabemos que $m_f(f_{V_i}) = m_f(f)_{V_i} = 0$, y entonces $m_{f_{V_i}}$ divide a m_f : esto nos dice que m_f es un múltiplo común de los polinomios minimales $m_{f_{V_1}}, \dots, m_{f_{V_n}}$. Sea, por otro lado, $p \in \mathbb{k}[X]$ un múltiplo común de estos polinomios. Si $x \in V$, entonces existen $x_1 \in V_1, \dots, x_n \in V_n$ tales que $x = x_1 + \dots + x_n$ y para cada $i \in \llbracket n \rrbracket$ es

$$p(f)(x_i) = p(f)_{V_i}(x_i) = p(f_{V_i})(x_i) = 0,$$

ya que $p(f_{V_i}) = 0$, y como consecuencia de esto tenemos que

$$p(f)(x) = p(f)(x_1) + \dots + p(f)(x_n) = 0.$$

Esto nos dice que, de hecho, $p(f) = 0$ y, por lo tanto, que m_f divide a p .

(iii) Sean finalmente x_1, \dots, x_n vectores de V tales que $V = \langle x_1 \rangle_f + \dots + \langle x_n \rangle_f$. Para cada $i \in \llbracket n \rrbracket$ sabemos que el subespacio $\langle x_i \rangle_f$ es f -invariante y que el polinomio minimal de la restricción de f a $\langle x_i \rangle_f$ es m_{f,x_i} . El resultado de (iii) es consecuencia de esto y de la parte (ii) que ya probamos. \square

5.6.14. La segunda parte de la Proposición 5.6.13 tiene la siguiente consecuencia útil:

Corolario. Sea V un espacio vectorial de dimensión finita y sea $f : V \rightarrow V$ un endomorfismo de V . Si W es un subespacio f -invariante de V , entonces el polinomio minimal de la restricción $f_W : W \rightarrow W$ divide al de f .

Demostración. En efecto, como V y W son subespacios f -invariantes de V y, por supuesto, $V = W + V$, la Proposición 5.6.13(ii) nos dice que el polinomio minimal m_f es el mínimo común múltiplo de los polinomios m_{f_W} y m_f y, en particular, que es divisible por el primero de éstos. \square

5.6.15. Por otro lado, un caso especial de la tercera parte de la Proposición 5.6.13 es muchas veces de interés:

Corolario. Sea V un espacio vectorial de dimensión finita y sea $f : V \rightarrow V$ un endomorfismo de V . Si existe un vector $x \in V$ tal que $V = \langle x \rangle_f$, entonces $m_f = m_{f,x}$.

Un vector x tal que $V = \langle x \rangle_f$, como en este corolario, es un **vector cíclico para f** . En general, un endomorfismo de un espacio vectorial de dimensión finita no admite ningún vectores. La Proposición 5.7.8 que probaremos más adelante da una condición necesaria y suficiente para que existan, que tiene como consecuencia el hecho de que si existe alguno entonces «casi todos» los elementos de V son cílicos.

Demostración. Esto es el caso en el que $n = 1$ y $x_1 = x$ de la Proposición 5.6.13(iii). \square

5.6.16. Necesitaremos el siguiente lema sobre polinomios:

Lema. Si p y q son dos polinomios de $\mathbb{k}[X]$, entonces existen polinomios $u_1, u_2, v_1, v_2 \in \mathbb{k}[X]$ tales que

$$p = u_1 u_2, \quad q = v_1 v_2, \quad \text{mcm}\{p, q\} = u_1 v_2, \quad \gcd\{u_1, v_2\} = 1.$$

Demostración. Sean p y q dos elementos de $\mathbb{k}[X]$. Si p_1, \dots, p_r son todos los polinomios monómicos irreducibles de $\mathbb{k}[X]$ que dividen a pq , entonces existen $\alpha_1, \dots, \alpha_r, \beta_1, \dots, \beta_r \in \mathbb{N}_0$ tales que $p = p_1^{\alpha_1} \cdots p_r^{\alpha_r}$ y $q = p_1^{\beta_1} \cdots p_r^{\beta_r}$. A menos de reindexar los polinomios p_1, \dots, p_r podemos suponer que existe $s \in [0, r]$ tal que para todo $i \in [r]$ se tiene que

$$0 \leq i \leq s \iff \alpha_i \geq \beta_i.$$

Si ponemos

$$u_1 := p_1^{\alpha_1} \cdots p_s^{\alpha_s}, \quad u_2 := p_{s+1}^{\alpha_{s+1}} \cdots p_r^{\alpha_r}, \quad v_1 := p_1^{\beta_1} \cdots p_s^{\beta_s}, \quad v_2 := p_{s+1}^{\beta_{s+1}} \cdots p_r^{\beta_r},$$

entonces es claro que $p = u_1 u_2$, $q = v_1 v_2$, $\text{mcm}\{p, q\} = u_1 v_2$ y $\gcd\{u_1, v_2\} = 1$. \square

5.6.17. Proposición. Sea $f : V \rightarrow V$ un endomorfismo de un espacio vectorial V de dimensión finita. Si x e y son dos vectores de V , entonces existe un tercer vector z tal que

$$m_{f,z} = \text{mcm}\{m_{f,x}, m_{f,y}\}.$$

Demostración. Sean x e y dos vectores de V y $m_{f,x}$ y $m_{f,y}$ sus correspondientes polinomios minimales con respecto a f , y sea $m := \text{mcm}\{m_{f,x}, m_{f,y}\}$ el mínimo común múltiplo de esos polinomios minimales. De acuerdo al Lema 5.6.16, existe polinomios $u_1, u_2, v_1, v_2 \in \mathbb{k}[X]$ tales que

$$m_{f,x} = u_1 u_2, \quad m_{f,y} = v_1 v_2, \quad m = u_1 v_2, \quad \gcd\{u_1, v_2\} = 1.$$

Consideremos el vector $z := u_2(f)(x) + v_1(f)(y)$.

- Como $mu_2 = v_2 m_{f,x}$ y $mv_1 = u_2 m_{f,y}$, tenemos que

$$\begin{aligned} m(f)(z) &= (mu_2)(f)(x) + (mv_1)(f)(x) \\ &= (v_2 m_{f,x})(f)(x) + (u_2 m_{f,y})(f)(x) \\ &= 0. \end{aligned}$$

- Sea $p \in \mathbb{k}[X]$ un polinomio tal que $p(f)(z) = 0$. Se sigue de esta igualdad que

$$0 = (v_2 p)(f)(z) = (v_2 u_2 p)(f)(x) + (v_1 v_2 p)(f)(y) = (u_2 v_2 p)(f)(x),$$

ya que $v_1 v_2 = m_{f,y}$, y entonces $m_{f,x} = u_1 u_2$ divide a $u_2 v_2 p$: como u_1 y v_2 son coprimos, vemos que u_1 divide a p . De manera similar, como $p(f)(z) = 0$, tenemos que

$$0 = (u_1 p)(f)(z) = (u_1 u_2 p)(f)(x) + (u_1 v_2 p)(f)(y) = (u_1 v_2 p)(f)(y),$$

porque $u_1 u_2 = m_{f,x}$, así que $m_{f,y} = v_1 v_2$ divide a $u_1 v_2 p$ y, como u_1 y v_2 son coprimos, u_1 divide a p . Finalmente, otra vez porque como u_1 y v_2 son coprimos, podemos concluir que $m = u_1 v_2$ divide a p .

En vista de estas dos observaciones, m es el polinomio minimal de z . □

5.6.18. Una consecuencia de la proposición que acabamos de probar es que siempre podemos realizar el polinomio minimal de un endomorfismo como el polinomio minimal de un vector.

Corolario. Si $f : V \rightarrow V$ es un endomorfismo de un espacio vectorial de dimensión finita, entonces existe un vector $x \in V$ tal que $m_{f,x} = m_f$.

Demostración. Sea $f : V \rightarrow V$ un endomorfismo de un espacio vectorial de dimensión finita. Existen $r \in \mathbb{N}$ y vectores $x_1, \dots, x_r \in V$ tales que $V = \langle x_1 \rangle_f + \dots + \langle x_r \rangle_f$, y probaremos el corolario haciendo inducción con respecto a r . Cuando $r = 1$, tenemos que $m_f = m_{f,x_1}$ de acuerdo al Corolario 5.6.15, así que la afirmación del corolario vale en este caso

Supongamos entonces que $r > 1$. El subespacio $V_1 := \langle x_1 \rangle_f$ de V es f -invariante y el polinomio minimal de la restricción $f_{V_1} : V_1 \rightarrow V_1$ de f a V_1 es, de acuerdo al Corolario 5.6.15,

$$m_{f_{V_1}} = m_{f_{V_1,x_1}} = m_{f,x_1}.$$

Por otro lado, el subespacio $V_2 := \langle x_1 \rangle_f + \cdots + \langle x_{r-1} \rangle_f$ de V es f -invariante y si $f_{V_2} : V_2 \rightarrow V_2$ es la restricción de f a V_2 , entonces —ya que V_2 es suma de $r - 1$ subespacios f_{V_2} -cíclicos— la hipótesis inductiva evidente nos dice que existe un vector $y \in V_2$ tal que

$$m_{f,y} = m_{f_{V_2},y} = m_{f_{V_2}}.$$

La Proposición 5.6.17 nos dice ahora que hay un vector x en V tal que

$$m_{f,x} = \text{mcm}\{m_{f,x_1}, m_{f,y}\} = \text{mcm}\{m_{f_{V_1}}, m_{f_{V_2}}\}$$

y, de acuerdo a la Proposición 5.6.13(ii), esto es

$$= m_f,$$

ya que $V = V_1 + V_2$. Esto prueba el corolario. \square

5.6.19. La Proposición 5.6.13 nos da una forma de calcular el polinomio minimal de un endomorfismo o una matriz de manera más eficiente que con el algoritmo descripto en 5.6.6. Si tenemos un endomorfismo $f : V \rightarrow V$ de un espacio vectorial V de dimensión finita y llevamos a cabo el siguiente procedimiento, al terminar el polinomio m es el polinomio minimal de f .

```

1   m ← 1, el polinomio constante
2   S ← 0, el subespacio nulo de V
3   mientras S ≠ V hacer
4       | Elegir un vector x en V \ S y calcular el polinomio minimal m_{f,x}
5       | m ← mcm{m, m_{f,x}}
6       | S ← S + ⟨x⟩_f
7   fin

```

En efecto, sean $m_0 = 1$ y $S_0 =$ los valores iniciales de m y de S , y para cada $i \in \mathbb{N}$ escribamos x_i al vector elegido en la iteración i del bucle que empieza en la línea 3 y m_i y S_i a los valores de las variables m y S al terminar esa iteración. Para todo $i \in \mathbb{N}$ es $x_i \notin S_{i-1}$, así que $S_{i-1} \subsetneq S_i$: esto implica que las dimensiones de los subespacios S_0, S_1, \dots de V crecen estrictamente y, como V tiene dimensión finita, que el número de iteraciones del bucle es finita y, por lo tanto, que el procedimiento a la larga termina. Sea n el número de veces que se repite el bucle. Como $S_i = S_{i-1} + \langle x_i \rangle_f$ y $m_i = \text{mcm}\{m_{i-1}, m_{f,x_{i-1}}\}$ para cada $i \in \llbracket d \rrbracket$, tenemos que $V = S_n = \langle x_1 \rangle_f + \cdots + \langle x_n \rangle_f$ y $m_n = \text{mcm}\{m_{f,x_1}, \dots, m_{f,x_n}\}$, de manera que m_n es el polinomio minimal de f .

Este procedimiento requiere en el peor caso $\dim V$ cálculos de polinomios minimales puntuales, sumas de dos subespacios y cálculo del mínimo común múltiplo de dos polinomios. La ventaja más importante de seguir este procedimiento y no el de 5.6.6 es que no tenemos que calcular con potencias de f sino todo el tiempo solamente con vectores de V .

El Teorema de Cayley–Hamilton

5.6.20. El siguiente resultado es conocido como el *Teorema de Cayley–Hamilton*. Fue enunciado por Arthur Cayley en su *A memoir on the theory of matrices* [Cay58] de 1858 pero sólo para matrices de 2×2 ; Cayley menciona allí que sabe probarlo para matrices de 3×3 pero no lo hace y dice al pasar que «no le parece necesario emprender la prueba del teorema general». William Rowan Hamilton había probado antes, en su libro [Ham53] de 1853 sobre los cuaterniones, que una matriz de 3×3 satisface a un polinomio de grado 3 y que una de 4×4 satisface uno de grado 4, pero no identificó ese polinomio con el polinomio característico. La primera prueba completa del teorema general fue dada por Frobenius en 1878.

Teorema. Sea V un espacio vectorial de dimensión finita y sea $f : V \rightarrow V$ un endomorfismo de V . Si χ_f es el polinomio característico de f , entonces $\chi_f(f) = 0$ y, en particular, χ_f es divisible por el polinomio minimal m_f .

Demostración. Sea $x \in V$ y sea W un complemento del subespacio cíclico $\langle x \rangle_f$ en V , de manera que $V = \langle x \rangle_f \oplus W$. Si $d = \dim \langle x \rangle_f$, sabemos que $\mathcal{B}_x = (x, f(x), \dots, f^{d-1}(x))$ es una base ordenada de $\langle x \rangle_f$; sea, por otro lado, $m = \dim W$ y sea (y_1, \dots, y_m) una base ordenada de W . En esta situación sabemos que $\mathcal{B} = (x, f(x), \dots, f^{d-1}(x), y_1, \dots, y_m)$ es una base ordenada de V y, como $\langle x \rangle_f$ es un subespacio f -invariante, es inmediato ver que la matriz de f con respecto a esta base ordenada \mathcal{B} tiene una descomposición en bloques de la forma

$$[f]_{\mathcal{B}}^{\mathcal{B}} = \begin{pmatrix} [f_x]_{\mathcal{B}_x}^{\mathcal{B}_x} & C \\ 0 & D \end{pmatrix}$$

para ciertas matrices $D \in M_m(\mathbb{k})$ y $C \in M_{d,m}(\mathbb{k})$. El polinomio característico de f es, entonces,

$$\chi_f = \chi_{f_x} \cdot \chi_C = m_{f,x} \cdot \chi_C.$$

Vemos así que el polinomio característico χ_f es divisible por cada uno de los polinomio $m_{f,x}$ con $x \in V$ y, en vista de la Proposición 5.6.13(iii), que es divisible por m_f . Esto implica, como sabemos, que $\chi_f(f) = 0$, como afirma el teorema. \square

5.6.21. Es una consecuencia de la Proposición 5.6.1 que el grado del polinomio minimal de un endomorfismo $f : V \rightarrow V$ de un espacio vectorial de dimensión n es a lo sumo n^2 , ya que coincide con la dimensión de un cierto subespacio de $\text{End}(V)$. Del teorema de Cayley–Hamilton obtenemos trivialmente una cota mucho mejor:

Corolario. Sea V un espacio vectorial de dimensión finita n y sea $f : V \rightarrow V$ un endomorfismo de V . El polinomio minimal de f tiene grado a lo sumo igual a n . Si ese grado es igual a n , entonces el polinomio minimal de f coincide con el polinomio característico de f .

En la Proposición 5.7.8 vemos que nos dice sobre el endomorfismo f saber que sus polinomios minimal y característico coinciden.

Demostración. La primera afirmación es consecuencia inmediata de que el polinomio minimal divide al característico y de que este último tiene grado n . La segunda, de que si los dos polinomios tienen el mismo grado que uno divida al otro implica —como son mónicos— que son iguales. \square

5.6.22. Del teorema se deduce fácilmente el siguiente resultado, que es muchas veces útil para determinar los autovalores de un endomorfismo:

Corolario. *Sea V un espacio vectorial de dimensión finita y sea $f : V \rightarrow V$ un endomorfismo. Los polinomios minimal y característico de f tienen las mismas raíces en \mathbb{k} y, en particular, todo autovalor de f es raíz del polinomio minimal de f .*

Más adelante, en la Proposición 5.7.6, daremos una mejora de este corolario.

Demostración. Si $\lambda \in \mathbb{k}$ es una raíz del polinomio característico χ_f , sabemos que existe un autovector $x \in V$ de autovalor λ . Esto implica, claramente, que $m_{f,x} = X - \lambda$, y la Proposición 5.6.13(i) nos dice que este polinomio divide al polinomio minimal m_f , así que λ es una raíz de m_f . Recíprocamente, como el Teorema de Cayley–Hamilton nos dice que m_f divide a χ_f , toda raíz del primero es raíz del segundo. \square

5.6.23. Una tercera consecuencia directa del teorema de Cayley–Hamilton es la siguiente expresión para el endomorfismo inverso de un endomorfismo inversible:

Corolario. *Sea V un espacio vectorial de dimensión finita n , sea $f : V \rightarrow V$ un endomorfismo y supongamos que el polinomio característico de f es $\chi_f = X^n + a_1X^{n-1} + \dots + a_{n-1}X + a_n$. Si f es inversible, entonces $a_n \neq 0$ y*

$$f^{-1} = -a_n^{-1}f^{n-1} - a_n^{-1}a_1f^{n-2} - \dots - a_n^{-1}a_{n-1}\text{id}_V.$$

Demostración. Del teorema de Cayley–Hamilton sabemos que $\chi_f(f) = 0$ y entonces

$$f^n + a_1f^{n-1} + \dots + a_{n-1}f + a_n\text{id}_V = 0.$$

Si f es inversible, la Proposición 4.3.6(iii) nos dice que $\det(f) \neq 0$ y, de acuerdo a la Proposición 5.4.17, es entonces $a_n \neq 0$. Componiendo a la derecha, por ejemplo, con f^{-1} a ambos lados de esta igualdad y multiplicando por a_n^{-1} , entonces, vemos que

$$f^{-1} = -a_n^{-1}f^{n-1} - a_n^{-1}a_1f^{n-2} - \dots - a_n^{-1}a_{n-1}\text{id}_V,$$

como afirma la proposición. \square

5.6.24. Ejemplo. Usando el Teorema de Cayley–Hamilton podemos determinar fácilmente el polinomio minimal de la matriz compañera de un polinomio. Sea $n \in \mathbb{N}$ y sea

$$p = X^n + c_{n-1}X^{n-1} + \cdots + c_1X + c_0 \in \mathbb{k}[X]$$

un polinomio mónico de grado n . Escribamos $C(p) \in M_n(\mathbb{k})$ a la matriz compañera de p , sea $f : x \in \mathbb{k}^n \mapsto C(p)x \in \mathbb{k}^n$ y sea $\mathcal{B} = (e_1, \dots, e_n)$ la base ordenada estándar de \mathbb{k}^n . Los vectores

$$e_1, \quad f(e_1) = e_2, \quad f^2(e_1) = e_3, \quad \dots, \quad f^{n-1}(e_1) = e_n$$

son linealmente independientes y

$$f^n(e_1) = -c_0e_1 - c_1f(e_1) - \cdots - c_{n-1}f^{n-1}(e_1),$$

así que $m_{f,e_1} = p$. Tenemos entonces que

$$p = m_{f,e_1} \mid m_f \mid \chi_f = p$$

y, como los polinomios m_f y χ_f son mónicos, vemos que

$$m_f = \chi_f = p.$$

Así, el polinomio minimal de $C(p)$ es precisamente p . \diamond

§7. Descomposición primaria y diagonalizabilidad

5.7.1. El siguiente resultado nos provee de la llamada *descomposición primaria* de un espacio vectorial con respecto a uno de sus endomorfismos:

Proposición. *Sea V un espacio vectorial de dimensión finita, sea $f : V \rightarrow V$ un endomorfismo de V y sea m_f su polinomio minimal. Si $p_1, \dots, p_n \in \mathbb{k}[X]$ son polinomios mónicos no constantes coprimos dos a dos tales que $m_f = p_1 \cdots p_n$ y ponemos $V_i = \text{Nu}(p_i(f))$ para cada $i \in \llbracket n \rrbracket$, entonces tenemos una descomposición en suma directa*

$$V = V_1 \oplus \cdots \oplus V_n. \tag{19}$$

Para cada $i \in \llbracket n \rrbracket$ el subespacio V_i es no nulo y f -invariante, el polinomio minimal de la restricción $f_{V_i} : V_i \rightarrow V_i$ de f a V_i es precisamente p_i , y si $\pi_i : V \rightarrow V$ es el proyector correspondiente a V_i para la descomposición en suma directa (19) de V , entonces hay un polinomio $v_i \in \mathbb{k}[X]$ tal que $\pi_i = v_i(f)$ y, en particular, π_i commuta con f .

Llamamos a la descomposición de V que nos da esta proposición la *descomposición primaria* de V relativa a f y el subespacio $V_i = \text{Nu}(p_i(f))$ es la *componente p -primaria* de V .

Demostración. Sean $p_1, \dots, p_n \in \mathbb{k}[X]$ polinomios no constantes coprimos dos a dos tales que $m_f = p_1 \cdots p_n$, como en el enunciado, y para cada $i \in \llbracket n \rrbracket$ consideremos el polinomio

$$q_i = p_1 \cdots \hat{p}_i \cdots p_n,$$

producto de todos ellos salvo el i -ésimo. Como los polinomios p_1, \dots, p_n son coprimos dos a dos,

$$\text{los polinomios } q_1, \dots, q_n \text{ no poseen divisores no triviales comunes.} \quad (20)$$

Para probar esto, supongamos que $t \in \mathbb{k}[X]$ es un divisor común no trivial de esos polinomios y, sin pérdida de generalidad, que es irreducible. Como t divide a q_1 , existe $k \in \llbracket 2, r \rrbracket$ tal que t divide a p_k , y como t también divide a q_k , existe $s \in \llbracket n \rrbracket \setminus \{k\}$ tal que t divide a p_s : esto es absurdo, ya que p_k y p_s no tienen divisores comunes.

Como consecuencia de (20) existen polinomios $r_1, \dots, r_n \in \mathbb{k}[X]$ tales que

$$r_1 q_1 + \cdots + r_n q_n = 1. \quad (21)$$

Para cada $i \in \llbracket n \rrbracket$ sea $\pi_i = r_i(f)q_i(f) : V \rightarrow V$. Afirmamos que

$$\pi_1 + \cdots + \pi_n = \text{id}_V \quad (22)$$

y que para cada $i, j \in \llbracket n \rrbracket$ es

$$\pi_i \circ \pi_j = \begin{cases} \pi_i, & \text{si } i = j; \\ 0, & \text{en caso contrario.} \end{cases} \quad (23)$$

Probemos esto:

- La igualdad (22) es consecuencia inmediata de (21) y de la definición de los endomorfismos π_1, \dots, π_n .
- Si i y j son dos elementos distintos de $\llbracket n \rrbracket$, entonces p_j divide a q_i , así que existe $s \in \mathbb{k}[X]$ con $q_i = s p_j$, y por lo tanto el polinomio $r_i q_i r_j q_j$ es divisible por m_f , ya que es igual a $r_i r_j s p_j q_j$. Como consecuencia de esto,

$$\pi_i \circ \pi_j = (r_i q_i r_j q_j)(f) = 0.$$

- Por otro lado, si $i \in \llbracket n \rrbracket$, componiendo ambos miembros de la igualdad (22) con π_i vemos que

$$\pi_i \circ \pi_1 + \cdots + \pi_i \circ \pi_n = \pi_i$$

y, en vista de lo que ya probamos, el miembro izquierdo de esta ecuación es igual a $\pi_i \circ \pi_i$. Esto nos dice que $\pi_i \circ \pi_i = \pi_i$ y prueba (23).

Ahora bien, como consecuencia de (22) y (23), el Corolario 2.7.7 nos dice que hay una descomposición en suma directa

$$V = \text{Im}(\pi_1) \oplus \cdots \oplus \text{Im}(\pi_n)$$

cuyos proyectores asociados son precisamente los endomorfismos π_1, \dots, π_n , cada uno de los cuales se obtiene evaluando un polinomio de $\mathbb{k}[X]$ en f .

Mostremos ahora que para cada $i \in \llbracket n \rrbracket$ se tiene que

$$\text{Nu}(p_i(f)) = \text{Im}(\pi_i). \quad (24)$$

- Si x es un vector de en $\text{Im}(\pi_i)$, existe $y \in V$ tal que $x = \pi_i(y) = (r_i q_i)(f)(y)$ y entonces

$$p_i(f)(x) = (p_i r_i q_i)(f)(y) = (r_i m_f)(f)(y) = 0,$$

es decir, x pertenece a $\text{Nu}(p_i(f))$

- Recíprocamente, si x está en $\text{Nu}(p_i(f))$, se tiene que para cada para cada $j \in \llbracket n \rrbracket \setminus \{i\}$ es $q_j(f)(x) = 0$, ya que q_j es divisible por p_i , y, en consecuencia, $\pi_j(x) = 0$: recordando la igualdad (22), tenemos entonces que $x = \pi_1(x) + \dots + \pi_n(x) = \pi_i(x) \in \text{Im}(\pi_i)$. Esto prueba (24).

Para cada $i \in \llbracket n \rrbracket$ escribamos V_i al subespacio $\text{Nu}(p_i(f))$, de manera que

$$V = V_1 \oplus \dots \oplus V_n. \quad (25)$$

Como el endomorfismo $p_i(f)$ commuta con f , para cada $x \in V_i$ es

$$p_i(f)(f(x)) = f(p_i(f)(x)) = f(0) = 0,$$

de manera que $f(x)$ también está en V_i : esto nos dice que V_i es f -invariante. Podemos entonces considerar la restricción $f_{V_i} : V_i \rightarrow V_i$. Como $p_i(f_{V_i}) = p_i(f)_{V_i} = 0$, el polinomio $m_{f_{V_i}}$ divide a p_i y, por lo tanto, existe un polinomio mónico $s_i \in \mathbb{k}[X]$ tal que

$$p_i = s_i m_{f_{V_i}}. \quad (26)$$

Se sigue de esto, en particular, que los polinomios $m_{f_{V_1}}, \dots, m_{f_{V_n}}$ son coprimos dos a dos, ya que los polinomios p_1, \dots, p_n lo son por hipótesis. Usando esto y la Proposición 5.6.13(ii) aplicada a la descomposición (25), vemos que

$$m_{f_{V_1}} \cdots m_{f_{V_n}} = \text{lcm}\{m_{f_{V_1}}, \dots, m_{f_{V_n}}\} = m_f. \quad (27)$$

Por otro lado, de (26) se sigue que

$$m_f = p_1 \cdots p_n = s_1 \cdots s_n m_{f_{V_1}} \cdots m_{f_{V_n}}$$

y de comparar esto con (27) es claro que el polinomio $s_1 \cdots s_n$ debe ser igual a 1. Esto implica que $m_{f_{V_i}} = p_i$ para cada $i \in \llbracket n \rrbracket$ y, en particular, que V_i es un subespacio no nulo de V , ya que p_i tiene grado positivo. Con esto todas las afirmaciones de la proposición quedan probadas. \square

5.7.2. Usando este resultado sobre descomposiciones primarias, podemos dar una condición necesaria y suficiente para la diagonalizabilidad de un endomorfismo en términos de su polinomio minimal, que es el resultado central de este capítulo:

Teorema. Sea V un espacio vectorial de dimensión finita positiva y sea $f : V \rightarrow V$ un endomorfismo de V . El endomorfismo f es diagonalizable si y solamente si su polinomio minimal m_f se factoriza como producto de polinomios lineales en $\mathbb{k}[X]$ y todas sus raíces son simples.

Demostración. Supongamos primero que f es diagonalizable y sean $\lambda_1, \dots, \lambda_n$ los autovalores de f listados sin repeticiones. De acuerdo a la Proposición 5.3.1, tenemos una descomposición

$$V = E_{\lambda_1}(f) \oplus \cdots \oplus E_{\lambda_n}(f)$$

de V como suma directa de subespacios no nulos y f -invariantes. Es claro que el polinomio minimal de la restricción $f_{E_{\lambda_i}(f)}$ de f al autoespacio $E_{\lambda_i}(f)$ es el polinomio $X - \lambda_i$ y entonces la Proposición 5.6.13(ii) nos dice que el polinomio minimal de f es el mínimo común múltiplo de los n polinomios $X - \lambda_1, \dots, X - \lambda_n$, que es claramente $(X - \lambda_1)\cdots(X - \lambda_n)$. Este polinomio manifiestamente se factoriza en $\mathbb{k}[X]$ como producto de polinomio lineales y tiene sus raíces simples: esto significa que la condición del enunciado es necesaria.

Veamos su suficiencia. Supongamos que el polinomio minimal de f posee en $\mathbb{k}[X]$ una factorización de la forma $m_f = (X - \lambda_1)\cdots(X - \lambda_n)$, con los escalares $\lambda_1, \dots, \lambda_n \in \mathbb{k}$ distintos dos a dos. Como los polinomios $X - \lambda_1, \dots, X - \lambda_n$ son mónicos no constantes y no tiene divisores comunes, la Proposición 5.7.1 nos dice que si para cada $i \in \llbracket n \rrbracket$ ponemos $V_i = \text{Nu}(f - \lambda_i \text{id}_V)$, entonces

$$V = V_1 \oplus \cdots \oplus V_n.$$

Para cada $i \in \llbracket n \rrbracket$ es $V_i = E_{\lambda_i}(f)$, el autoespacio de f correspondiente a λ_i , así que la existencia de esta descomposición, en vista de la Proposición 5.3.1, implica que el endomorfismo f es diagonalizable. \square

5.7.3. Ejemplo. Si $p \in \mathbb{k}[X]$ es un polinomio mónico de grado n , sabemos del Ejemplo 5.6.24 que el polinomio minimal de la matriz compañera $C(p)$ es precisamente p , así que el Teorema 5.7.2 nos dice que $C(p)$ es diagonalizable exactamente cuando p se factoriza como producto de factores mónicos de grado 1 distintos dos a dos, esto es, cuando tiene en \mathbb{k} precisamente n raíces distintas dos a dos. \diamond

5.7.4. Ejemplo. Si $f : V \rightarrow V$ es un proyector de un espacio vectorial V , entonces $f^2 = f$ y, por lo tanto, si ponemos $p = X^2 - X$ tenemos que $p(f) = 0$. Así, si V tiene dimensión finita, entonces el polinomio minimal m_f de f divide a $X(X - 1)$ y, por lo tanto, se factoriza como producto de factores mónicos de grado 1 distintos dos a dos: el Teorema 5.7.2 nos dice entonces que f es diagonalizable. Más aún, como las raíces de p son 0 y 1, los autovalores de f son todos elementos de $\{0, 1\}$. De esta manera obtenemos una vez más la Proposición 2.7.2. \diamond

5.7.5. Ejemplo. Sea V un espacio vectorial sobre \mathbb{C} de dimensión finita y sea $f : V \rightarrow V$ un endomorfismo de V . Decimos que f tiene **orden finito** si existe un entero positivo $k \in \mathbb{N}$ tal que

$f^k = \text{id}_V$. Afirmamos que

si f tiene orden finito, entonces f es diagonalizable.

En efecto, si f tiene orden finito y $k \in \mathbb{N}$ es tal que $f^k - \text{id}_V = 0$, y ponemos $p := X^k - 1 \in \mathbb{k}[X]$, entonces $p(f) = 0$ y, por lo tanto, el polinomio minimal m_f divide a p . Ahora bien, en \mathbb{C} el polinomio p se factoriza como producto de factores lineales coprimos, uno por cada raíz k -ésima de la unidad, así que lo mismo es cierto de m_f : el Teorema 5.7.2 nos dice, entonces, que el endomorfismo f es diagonalizable. \diamond

5.7.6. Una segunda consecuencia de la existencia de descomposiciones primarias es:

Proposición. *Sea V un espacio vectorial de dimensión finita y sea $f : V \rightarrow V$ un endomorfismo de V . Los polinomios minimal y característico de f tienen los mismos divisores irreducibles.*

Observemos que esto completa el resultado del Corolario 5.6.22, que afirma lo mismo pero sólo sobre los divisores de grado 1.

Demostración. El Teorema de Cayley–Hamilton 5.6.20 nos dice que m_f divide a χ_f , así que si un polinomio irreducible divide al primero también divide al segundo.

Supongamos ahora que p_1, \dots, p_n son polinomios mónicos irreducibles distintos dos a dos y que v_1, \dots, v_n son enteros positivos tales que $m_f = p_1^{v_1} \cdots p_n^{v_n}$. Como los polinomios $p_1^{v_1}, \dots, p_n^{v_n}$ son mónicos, no constantes y coprimos dos a dos, la Proposición 5.7.1 nos dice que si para cada $i \in [n]$ ponemos $V_i = \text{Nu}(p_i^{v_i}(f))$, entonces hay una descomposición $V = V_1 \oplus \cdots \oplus V_n$ de V como suma directa de subespacios no nulos y f -invariantes. Para cada $i \in [n]$ podemos considerar la restricciones $f_{V_i} : V_i \rightarrow V_i$ de f a V_i y el Ejemplo 5.4.6(a) nos dice que $\chi_f = \chi_{f_{V_1}} \cdots \chi_{f_{V_n}}$. Si $i \in [n]$, sabemos de la Proposición 5.7.1 que el polinomio minimal de la restricción f_{V_i} es $m_{f_{V_i}} = p_i^{v_i}$ y del Teorema de Cayley–Hamilton que $m_{f_{V_i}}$ divide a $\chi_{f_{V_i}}$: todo esto implica, por supuesto, que p_i divide a χ_f . \square

5.7.7. Terminemos usando el Teorema de descomposición primaria para caracterizar los endomorfismos que admiten vectores cíclicos. Para ello, necesitamos la siguiente observación:

Lema. *Sea V un espacio vectorial, sea $f : V \rightarrow V$ un endomorfismo de V , sea W un subespacio f -invariante de V y sea $f_W : W \rightarrow W$ la restricción de f a W . Si f admite un vector cíclico, entonces f_W también admite uno.*

Demostración. Supongamos que $x \in V$ es un vector cíclico para f . Si W es el subespacio nulo de V es claro que f_W posee un vector cíclico, así que podemos suponer que $W \neq 0$. Consideremos el conjunto

$$I := \{p \in \mathbb{k}[X] : p(f)(x) \in W\}.$$

Se trata claramente de un subespacio de $\mathbb{k}[X]$, y usando la Proposición 5.5.4 y el hecho de que W es f -invariante es inmediato ver que se trata de un ideal de $\mathbb{k}[X]$. Más todavía, $I \neq 0$: si $w \in W \setminus 0$, entonces como x es un vector cíclico para f , existe un polinomio $p \in \mathbb{k}[X]$ tal que $p(f)(x) = w$, de manera que $p \in I$ y, como $w \neq 0$, $p \neq 0$. De acuerdo a la Proposición 5.5.5, hay un polinomio mónico q en I que divide a todos los elementos de I .

Sea $y = q(f)(x)$, que es un elemento de W porque $q \in I$: para probar el lema mostraremos que y es un vector cíclico para f_W . Por supuesto, tenemos que $\langle y \rangle_{f_W} \subseteq W$. Por otro lado, si $w \in W$, entonces como x es cíclico para f , existe $r \in \mathbb{k}[X]$ tal que $r(f)(x) = w$ y, en particular, $r \in I$, así que r es divisible por q : existe $s \in \mathbb{k}[X]$ tal que $r = sq$ y, por lo tanto,

$$w = r(f)(x) = s(f)(q(f)(x)) = s(f)(y) = s(f_W)(y) \in \langle y \rangle_{f_W}.$$

Vemos así que $W \subseteq \langle y \rangle_{f_W}$ y, en definitiva, que vale la igualdad. \square

5.7.8. Podemos dar ya la caracterización de los endomorfismos que admiten subespacios cílicos.

Proposición. *Sea V un espacio vector de dimensión finita y sea $f : V \rightarrow V$ un endomorfismo de V . Las siguientes afirmaciones son equivalentes:*

- (a) *El endomorfismo f admite un vector cíclico.*
- (b) *El polinomio minimal m_f de f y el polinomio característico χ_f de f coinciden.*
- (c) *Si $m_f = p_1^{v_1} \cdots p_r^{v_r}$ es la factorización del polinomio minimal de f como producto de potencias positivas de polinomios monicos irreducibles p_1, \dots, p_r distintos dos a dos, entonces para cada $i \in \llbracket r \rrbracket$ se tiene que $\dim \text{Nu}(p_i(f)) = \deg p_i$.*

Demostración. Sean m_f y χ_f el polinomio minimal y el polinomio característico de f , y sea $m_f = p_1^{v_1} \cdots p_r^{v_r}$ es la factorización del polinomio minimal de f como producto de potencias positivas de polinomios monicos irreducibles p_1, \dots, p_r distintos dos a dos. De acuerdo a la Proposición 5.7.1, si para cada $i \in \llbracket r \rrbracket$ ponemos $V_i := \text{Nu}(p_i^{v_i}(f))$, entonces

$$V = V_1 \oplus \cdots \oplus V_r. \quad (28)$$

Más aún, si $i \in \llbracket r \rrbracket$, el subespacio V_i es f -invariante, no nulo, el polinomio minimal de la restricción $f_{V_i} : V_i \rightarrow V_i$ de f a V_i es precisamente $p_i^{v_i}$, y $\text{Nu}(p_i(f)) \neq 0$.

(a) \Rightarrow (b) Supongamos primero que hay en V un vector cíclico x para f , de manera que $V = \langle x \rangle_f$. El Corolario 5.6.15 nos dice entonces que $m_f = m_{f,x}$, así que

$$\deg m_f = \deg m_{f,x} = \dim \langle x \rangle_f = \dim V = \deg \chi_f.$$

De acuerdo al Teorema de Cayley–Hamilton, el polinomio minimal m_f divide a χ_f y como tienen el mismo grado y son los dos monicos, vemos que $m_f = \chi_f$, de manera que vale (b).

(b) \Rightarrow (a) Supongamos ahora que $m_f = \chi_f$. Si $i \in \llbracket r \rrbracket$, el polinomio minimal de la restricción f_{V_i} de f a V_i es $p_i^{v_i}$, el endomorfismo $p_i^{v_i-1}(f_{V_i})$ de V_i no es nulo y, por lo tanto, su núcleo es un subespacio propio de V_i : existe entonces un vector $x_i \in V_i \setminus \text{Nu}(p_i^{v_i-1}(f_{V_i}))$ y, en particular, el polinomio minimal puntual de x_i para la restricción f_{V_i} es $m_{f_{V_i}, x_i} = p_i^{v_i}$.

Sea $x = x_1 + \dots + x_r$. Observemos que

$$0 = m_{f,x}(f)(x) = m_{f,x}(f)(x_1) + \dots + m_{f,x}(f)(x_r)$$

y como la suma de (28) es directa y los sumandos todos f -invariantes, esto implica que para todo $i \in [r]$ es $m_{f,x}(f_{V_i})(x_i) = m_{f,x}(f)(x_i) = 0$ y, por lo tanto, que $p_i^{v_i} = m_{f_{V_i}, x_i} \mid m_{f,x}$. Como los polinomios $p_1^{v_1}, \dots, p_r^{v_r}$ son dos a dos coprimos, vemos con esto que m_f divide a $m_{f,x}$. Como además $m_{f,x}$ divide a m_f y ambos polinomios son mónicos, tenemos en definitiva que $\chi_f = m_f = m_{f,x}$ y, por lo tanto, que

$$\dim V = \deg \chi_f = \deg m_{f,x} = \dim \langle x \rangle_f.$$

Así, $\langle x \rangle_f = V$ y x es un vector cíclico para f .

(a) \Rightarrow (c) Supongamos que f admite un vector cíclico y sea $i \in [r]$. Como el subespacio $W := \text{Nu}(p_i(f))$ de V es f -invariante y no nulo, el Lema 5.7.7 nos dice que hay en W un vector cíclico x , necesariamente no nulo, para la restricción f_W de f a W . Es $p_i(f_W)(x) = p_i(f)(x) = 0$, así que el polinomio minimal $m_{f_W, x}$ divide a p_i : como $x \neq 0$, el polinomio $m_{f_W, x}$ no es constante, así que —ya que es mónico y p_i irreducible y mónico— tiene que coincidir con p_i . Esto implica que $\dim \text{Nu}(p_i(f)) = \dim \langle x \rangle_{f_i} = \deg p_i$.

(c) \Rightarrow (b) Supongamos ahora que vale la condición (c) del enunciado. Sea $i \in [n]$. Para cada $k \in \mathbb{N}_0$ tenemos una función lineal

$$g_k : x \in \text{Nu}(p_i^{k+1}(f)) \mapsto p_i(f)(x) \in \text{Nu}(p_i^k(f))$$

que tiene por núcleo al subespacio $\text{Nu}(p_i(f))$ de $\text{Nu}(p_i^{k+1}(f))$, así que

$$\dim \text{Nu}(p_i^{k+1}(f)) = \dim \text{Nu}(g_k) + \dim \text{Im}(g_k) \leq \dim \text{Nu}(p_i(f)) + \dim \text{Nu}(p_i^k(f)).$$

De esto se deduce inmediatamente que para todo $k \in \mathbb{N}_0$ es

$$\dim \text{Nu}(p_i^k(f)) \leq k \cdot \dim \text{Nu}(p_i(f))$$

y, en particular y de acuerdo a la hipótesis, que

$$\dim V_i = \dim \text{Nu}(p_i^{v_i}(f)) \leq v_i \cdot \dim \text{Nu}(p_i(f)) = v_i \cdot \deg p_i.$$

Tenemos entonces que

$$\deg \chi_n = \dim V = \dim V_1 + \dots + \dim V_r \leq v_1 \cdot \deg p_1 + \dots + v_r \cdot \deg p_r = \deg m_f$$

y, como m_f divide a χ_f y ambos polinomios son mónicos, que $m_f = \chi_f$. \square

5.7.9. Cuando el polinomio característico del endomorfismo se factoriza como producto de factores lineales —lo que ocurre si, por ejemplo, el cuerpo de base es algebraicamente cerrado— la Proposición 5.7.8 toma una forma particularmente sencilla.

Corolario. Sea V un espacio vectorial de dimensión finita y sea $f : V \rightarrow V$ un endomorfismo de V cuyo polinomio característico se factoriza como producto de factores lineales. Las siguientes afirmaciones son equivalentes:

- (a) El endomorfismo f admite un vector cíclico.
- (b) El polinomio minimal de f y el polinomio característico de f coinciden.
- (c) La multiplicidad geométrica de cada autovalor de f es 1.

Demostración. Esto es consecuencia inmediata de la proposición, ya que la hipótesis sobre el endomorfismo f implica que los factores irreducibles de su polinomio minimal son todos de grado 1. \square

§8. Endomorfismos triangularizables y semisimples

5.8.1. El Teorema 5.7.2 nos dice que un endomorfismo $f : V \rightarrow V$ de un espacio vectorial de dimensión finita es diagonalizable si en la factorización de su polinomio minimal como producto de factores irreducibles mónicos

- todos los factores tienen grado 1 y
- en esa factorización no aparecen factores repetidos.

En esta sección nos proponemos analizar qué sucede cuando el polinomio minimal sólo satisface a una de estas condiciones.

5.8.2. Empezamos con un resultado bien especial:

Lema. Sea $f : V \rightarrow V$ un endomorfismo de un espacio vectorial de dimensión finita. Si existen $\lambda \in \mathbb{k}$ y $v \in \mathbb{N}_0$ tales que el polinomio minimal de f es $m_f = (X - \lambda)^v$, entonces hay una base ordenada \mathcal{B} de V tal que la matriz $[f]_{\mathcal{B}}^{\mathcal{B}}$ es triangular superior.

Demostración. Procedemos haciendo inducción con respecto a la dimensión del espacio V , observando que cuando esta dimensión es nula no hay nada que probar. Sea entonces $n := \dim V$ y supongamos que $n > 0$.

Sean $\lambda \in \mathbb{k}$ y $v \in \mathbb{N}_0$ tales que $m_f = (X - \lambda)^v$. Sean, como siempre, V^* el espacio dual de V y $f^t : V^* \rightarrow V^*$ la función transpuesta de f . Sea ϕ un elemento no nulo de V^* . El polinomio minimal $m_{f^t, \phi}$ tiene grado positivo y divide a m_{f^t} , que coincide con m_f , así que existe $\mu \in \mathbb{N}$ tal que $\mu \leq v$ y $m_{f^t, \phi} = (X - \lambda)^\mu$. Sea $\psi = (f^t - \lambda \text{id}_{V^*})^{\mu-1}$. Como $m_{f^t, \phi}$ no divide a $(X - \lambda)^{\mu-1}$, $\psi \neq 0$. Por otro lado, $(f^t - \lambda \text{id}_V)(\psi) = m_{f^t, \phi}(f)(\phi) = 0$, así que $f^t(\psi) = \lambda\psi$. Vemos de esta manera que ψ es un autovector de f^t de autovalor λ .

Sea ahora $W := \text{Nu}(\psi)$. Si $w \in W$, entonces $\psi(f(w)) = f^t(\psi)(w) = \lambda\psi(w) = 0$, así que $f(w)$ también está en W : vemos de esta forma que W es un subespacio f -invariante de V y que

podemos considerar la restricción $f_W : W \rightarrow W$. Del Corolario 5.6.14 sabemos que el polinomio minimal m_{f_W} divide a $(X - \lambda)^v$, así que es de la forma $(X - \lambda)^{v'}$ para algún entero $v' \in \mathbb{N}_0$. Como la funcional ψ no es nula, $\dim W = \dim V - 1$ y la hipótesis inductiva evidente implica que hay una base ordenada $\mathcal{B}' = (x_1, \dots, x_{n-1})$ tal que a matriz $[f_W]_{\mathcal{B}'}^{\mathcal{B}'}$ es triangular superior. Si x_n es un vector de V tal que $\mathcal{B} = (x_1, \dots, x_n)$ es una base ordenada de V , entonces la matriz $[f]_{\mathcal{B}}^{\mathcal{B}}$ es una matriz de bloques de la forma

$$[f]_{\mathcal{B}}^{\mathcal{B}} = \begin{pmatrix} [f_W]_{\mathcal{B}'}^{\mathcal{B}'} & B \\ 0 & C \end{pmatrix},$$

con $B \in M_{n-1,1}(\mathbb{k})$ y $C \in M_1(\mathbb{k})$. Como es claro que $[f]_{\mathcal{B}}^{\mathcal{B}}$ es triangular superior, esto completa la inducción y, por lo tanto, la prueba del lema. \square

5.8.3. Usando el lema que acabamos de probar y el Teorema de descomposición primaria 5.7.1 podemos obtener el primero de los dos resultados que buscamos:

Teorema. *Sea V un espacio vectorial de dimensión finita y sea $f : V \rightarrow V$ un endomorfismo de V . El polinomio minimal de f se factoriza como producto de factores lineales si y solamente si existe una base ordenada \mathcal{B} de V tal que la matriz $[f]_{\mathcal{B}}^{\mathcal{B}}$ es una matriz triangular superior.*

Decimos en ese caso, por razones obvias, que el endomorfismo f es **triangularizable**. Observemos que la Proposición 5.4.15, que tiene una conclusión similar, se diferencia de ésta en que allí pusimos una hipótesis fuerte sobre el cuerpo de base mientras que aquí la hipótesis importante es sobre el endomorfismo mismo. Por supuesto, si el cuerpo de base es algebraicamente cerrado, todo polinomio de factoriza como producto de factores lineales, y entonces la proposición que estamos por probar implica inmediatamente a la Proposición 5.4.15.

Demostración. Si hay una base \mathcal{B} de V tal que la matriz $[f]_{\mathcal{B}}^{\mathcal{B}}$ es triangular superior, entonces el Ejemplo 5.4.3(b) nos dice que el polinomio característico χ_f , que coincide con el de la matriz $[f]_{\mathcal{B}}^{\mathcal{B}}$, es producto de factores lineales. Como el polinomio minimal de f divide a χ_f , él también es producto de factores lineales.

Supongamos ahora que el polinomio minimal m_f es $(X - \lambda_1)^{v_1} \cdots (X - \lambda_n)^{v_n}$ con $\lambda_1, \dots, \lambda_n \in \mathbb{k}$ escalares dos a dos distintos y $v_1, \dots, v_n \in \mathbb{N}$. De acuerdo a la Proposición 5.7.1, si ponemos $V_i = \text{Nu}((f - \lambda_i \text{id}_V)^{v_i})$ para cada $i \in \llbracket n \rrbracket$, entonces $V = V_1 \oplus \cdots \oplus V_n$, cada uno de los subespacios V_1, \dots, V_n es f -invariante y para cada $i \in \llbracket n \rrbracket$ el polinomio minimal de la restricción $f_{V_i} : V_i \rightarrow V_i$ de f a V_i es $m_{f_{V_i}} = (X - \lambda)^{v_i}$. En vista de esto, podemos deducir del Lema 5.8.2 que hay bases ordenadas $\mathcal{B}_1, \dots, \mathcal{B}_n$ de los subespacios V_1, \dots, V_n tales que las matrices $[f_{V_1}]_{\mathcal{B}_1}^{\mathcal{B}_1}, \dots, [f_{V_n}]_{\mathcal{B}_n}^{\mathcal{B}_n}$ son todas triangulares superiores. Si $\mathcal{B} = \mathcal{B}_1 \sqcup \cdots \sqcup \mathcal{B}_n$ es la base ordenada de V que se obtiene concatenando a las bases ordenadas $\mathcal{B}_1, \dots, \mathcal{B}_n$ en ese orden, entonces la matriz de f con respecto

a \mathcal{B} es la matriz diagonal por bloques

$$[f]_{\mathcal{B}}^{\mathcal{B}} = \begin{pmatrix} [f_{V_1}]_{\mathcal{B}_1}^{\mathcal{B}_1} & & \\ & \ddots & \\ & & [f_{V_1}]_{\mathcal{B}_1}^{\mathcal{B}_1} \end{pmatrix}$$

y es, por lo tanto, triangular superior. \square

5.8.4. Para obtener nuestro segundo resultado, empezamos por una construcción básica de la teoría de cuerpos:

Proposición. *Sea $p \in \mathbb{k}[X]$ un polinomio irreducible.*

- (i) *El conjunto $I := \{pq : q \in \mathbb{k}[X]\}$ es un ideal de $\mathbb{k}[X]$.*
- (ii) *Sea $\mathbb{K} := \mathbb{k}[X]/I$ y para cada $f \in \mathbb{k}[X]$ escribamos $[f]$ a la clase de f en el cociente \mathbb{K} . Hay una única función bilineal $\mu : \mathbb{K} \times \mathbb{K} \rightarrow \mathbb{K}$ tal que*

$$\mu([f], [g]) = [fg].$$

- (iii) *El conjunto \mathbb{K} dotado de la operación de suma + proveniente de su estructura de \mathbb{k} -espacio vectorial y de la operación de producto · dada por la función μ es un cuerpo, y un álgebra sobre \mathbb{k} .*

Demostración. (i) Es inmediato verificar que I es un subespacio de $\mathbb{k}[X]$. Por otro lado, si $x \in I$ y que $y \in \mathbb{k}[X]$, entonces existe $q \in \mathbb{k}[X]$ tal que $x = pq$ y $yx = xy = pqy \in I$. Esto prueba que I es un ideal de $\mathbb{k}[X]$.

(ii) Sea $m : \mathbb{k}[X] \rightarrow \mathbb{k}[X] \rightarrow \mathbb{k}[X]/I$ la función tal que $m(f, g) = [fg]$ para toda elección de f y g en $\mathbb{k}[X]$. Es inmediato verificar que se trata de una función bilineal. Además, si f y g están en $\mathbb{k}[X]$, vale que

$$m(f, g) = 0 \text{ si } f \in I \text{ o } g \in I.$$

En efecto, si por ejemplo $f \in I$, entonces existe $q \in \mathbb{k}[X]$ tal que $f = pq$ y, por lo tanto, $m(f, g) = [pqg] = 0$ en $\mathbb{k}[X]/I$, ya que $pqg \in I$. De acuerdo a la Proposición 2.8.7, existe exactamente una función bilineal $\mu : \mathbb{k}[X]/I \times \mathbb{k}[X]/I \rightarrow \mathbb{k}[X]/I$ tal que $\mu([f], [g]) = m(f, g)$ para toda elección de f y g en $\mathbb{k}[X]$. Esto es exactamente lo que afirma la segunda parte de la proposición.

(iii) Si x e y son dos elementos de \mathbb{K} escribimos $x \cdot y$ en lugar de $\mu(x, y)$. Como \mathbb{K} es un espacio vectorial, su operación de suma + es asociativa, conmutativa, posee un elemento neutro —el cero de \mathbb{K} — y todo elemento de \mathbb{K} posee un opuesto para +. Estas son las condiciones **(K₁)**, **(K₂)**, **(K₃)** y **(K₄)** de la definición 1.1.1.

Sean x, y y z tres elementos de \mathbb{K} , de manera que existen $u, v, w \in \mathbb{k}[X]$ tales que $x = [u]$, $y = [v]$ y $z = [w]$. Es $x \cdot y = \mu(x, y) = \mu([u], [v]) = [uv]$, y entonces

$$(x \cdot y) \cdot z = \mu([uv], [w]) = [(uv)w].$$

De manera similar, vemos que $x \cdot (y \cdot z) = [u(vw)]$. Como la multiplicación de $\mathbb{K}[X]$ es asociativa, tenemos que $(uv)w = u(vw)$ y entonces, por supuesto, que $[(uv)w] = [u(vw)]$. Esto muestra que $(x \cdot y) \cdot z = x \cdot (y \cdot z)$. De manera similar, como el producto en $\mathbb{k}[X]$ es conmutativo, tenemos que

$$x \cdot y = \mu([u], [v]) = [uv] = [vu] = \mu([v], [u]) = y \cdot x.$$

Si 1 es el polinomio constante mónico, entonces

$$x \cdot [1] = \mu([u], [1]) = [u1] = [u] = x$$

y $[1] \cdot x = x$, así que $[1]$ es un elemento neutro para el producto \cdot . Vemos con esto que las condiciones **(K₅)**, **(K₆)** y **(K₇)** se cumplen. Por otro lado, la bilinealidad de la función μ implica inmediatamente que se cumple la condición **(K₉)** de distributividad del producto \cdot sobre la suma $+$. Con todo esto podemos concluir ya que \mathbb{K} es un álgebra sobre \mathbb{k} .

Como el polinomio p no divide al polinomio constante 1, es $1 \notin I$ y, por lo tanto, $[1] \neq 0$ en \mathbb{K} : esto nos dice que se satisface la condición **(K₁₀)**. Para ver que \mathbb{K} también es un cuerpo, nos queda solamente verificar la condición **(K₈)**.

Sea x un elemento no nulo de \mathbb{K} . Existe $f \in \mathbb{K}[X]$ tal que $x = [f]$ y como $x \neq 0$ tenemos que $f \notin I$, esto es, que el polinomio p no divide a f . Como p es irreducible, esto implica que, de hecho, p y f son coprimos y, por lo tanto, que existe polinomios $r, g \in \mathbb{k}[X]$ tales que $rp + gf = 1$. Pero entonces

$$[g] \cdot x = [gf] = [1 - rp] = [1],$$

ya que $rp \in I$, y, en consecuencia, la clase $[g]$ es un inverso para x en \mathbb{K} . □

5.8.5. Decimos que un endomorfismo $f : V \rightarrow V$ de un espacio vectorial V es **semisimple** si todo subespacio f -invariante de V posee un complemento f -invariante, esto es, si cada vez que W es un subespacio f -invariante de V existe otro subespacio f -invariante W' tal que $V = W \oplus W'$.

Lema. *Sea V un espacio vectorial.*

- (i) *Si U, U', W y W' son subespacios de V tales que $V = U \oplus U' = W \oplus W'$ y $U \subseteq W'$, entonces $W = U \oplus U' \cap W$.*
- (ii) *Si $f : V \rightarrow V$ es un endomorfismo semisimple de V y W es un subespacio f -invariante de V , entonces la restricción $f_W : W \rightarrow W$ es un endomorfismo semisimple de W .*

Demostración. (i) Si $w \in W$, entonces existen $u \in U$ y $u' \in U'$ tales que $w = u + u'$ y como $U \subseteq W$ es $W \ni w - u = u' \in U'$, de manera que $u' \in U' \cap W$ y, por lo tanto, $w = u + u' \in U + U' \cap W$. Vemos así que $W \subseteq U + U' \cap W$ y vale de hecho la igualdad porque a inclusión recíproca claramente vale. Así, $W = U + U' \cap W$ y la Proposición 1.9.8 nos dice que la suma es además directa.

(ii) Sea $f : V \rightarrow V$ un endomorfismo semisimple de V , sea W un subespacio f -invariante de V , sea $f_W : W \rightarrow W$ la restricción de f a W y sea U un subespacio f_W -invariante de W . Como

U es f_W -invariante, es un subespacio f -invariante de V y como f es semisimple existe entonces un subespacio f -invariante U' de V tal que $V = U \oplus U'$. Por otro lado, como W es un subespacio f -invariante de V , existe un subespacio f -invariante W' de V tal que $V = W \oplus W'$. Como $U \subseteq W$, la primera parte del lema nos dice que $W = U \oplus U' \cap W$, y es inmediato ver que $U' \cap W$ es f -invariante y, por lo tanto, f_W -invariante. Vemos así que la restricción f_W es semisimple. \square

5.8.6. Podemos ahora probar un caso particular del segundo resultado que buscamos.

Proposición. *Sea $f : V \rightarrow V$ un endomorfismo de un espacio vectorial de dimensión finita. Si el polinomio minimal de f es irreducible, entonces f es semisimple.*

Demostración. Sea p el polinomio minimal de f y supongamos que p es irreducible. Como en esta demostración trabajaremos con dos cuerpos simultáneamente, decoraremos todas las nociones de linealidad con el nombre del cuerpo sobre el que están consideradas.

Sea $\varepsilon_f : \mathbb{k}[X] \rightarrow \text{End}(V)$ el homomorfismo de \mathbb{k} -álgebras dado por la evaluación en f y sea $I := \text{Nu}(\varepsilon_f)$ su núcleo que, como sabemos, es el \mathbb{k} -subespacio $\{pq : q \in \mathbb{k}[X]\}$ de $\mathbb{k}[X]$. De acuerdo a la Proposición 2.8.6, existe exactamente una función \mathbb{k} -lineal $\rho : \mathbb{K} := \mathbb{k}[X]/I \rightarrow \text{End}(V)$ tal que $\rho([q]) = \varepsilon_f(q)$ para todo $q \in \mathbb{k}[X]$. Veamos a \mathbb{K} como una \mathbb{k} -álgebra y como un cuerpo con la estructura descripta en la Proposición 5.8.4. Es inmediato verificar que la función ρ es un homomorfismo de \mathbb{k} -álgebras.

Consideremos la función

$$\cdot : (x, v) \in \mathbb{K} \times V \mapsto \rho(x)(v) \in V.$$

Afirmamos que V , dotado de su operación de suma $+$ y de la multiplicación escalar por elementos de \mathbb{K} dada por este producto, es un espacio vectorial sobre \mathbb{K} . Como V es un espacio vectorial sobre \mathbb{k} , la suma $+$ de V satisface las condiciones (V₁), (V₂), (V₃) y (V₄). Si $x, y \in \mathbb{K}$ y $v \in V$, entonces $\rho(x \cdot y) = \rho(x) \cdot \rho(y)$ porque ρ es un homomorfismo de \mathbb{k} -álgebras, y entonces

$$x \cdot (y \cdot v) = \rho(x)(\rho(y)(v)) = (\rho(x) \cdot \rho(y))(v) = \rho(x \cdot y)(v) = (x \cdot y) \cdot v,$$

así que la condición (V₅) se cumple. Otra vez, como ρ es un homomorfismo de \mathbb{k} -álgebras, tenemos que $\rho(1_{\mathbb{K}}) = \text{id}_V$, y entonces para todo $v \in V$ es

$$1_{\mathbb{K}} \cdot v = \rho(1_{\mathbb{K}})(v) = \text{id}_V(v) = v,$$

y (V₆) también se satisface. Finalmente, las dos condiciones de distributividad de (V₇) valen: si $x, y \in \mathbb{K}$ y $v, w \in W$, entonces

$$(x + y) \cdot v = \rho(x + y)(v) = (\rho(x) + \rho(y))(v) = \rho(x)(v) + \rho(y)(v) = x \cdot v + y \cdot v$$

porque la función ρ es \mathbb{k} -lineal, y

$$x \cdot (v + w) = \rho(x)(v + w) = \rho(x)(v) + \rho(x)(w) = x \cdot v + x \cdot w,$$

ya que la función $\rho(x) : V \rightarrow V$ es \mathbb{k} -lineal.

En este momento el conjunto V es simultáneamente un \mathbb{k} -espacio vectorial y un \mathbb{K} -espacio vectorial. Estas dos estructuras están relacionadas de la siguiente manera:

un subconjunto U de V es un \mathbb{k} -subespacio f -invariante si y solamente si es un \mathbb{K} -subespacio.

Probemos esto. Sea U un subconjunto de V .

- Supongamos primero que U es un \mathbb{k} -subespacio f -invariante de V . En particular, el elemento 0 de V está en U y $v + w \in U$ siempre que $v, w \in U$. Sean, por otro lado, $x \in \mathbb{K}$ y $v \in U$, de manera que existe $q \in \mathbb{k}[X]$ tal que $x = [q]$ y

$$x \cdot v = \rho([q])(v) = \varepsilon_f(q)(v) = q(f)(v) \in U,$$

ya que U es $q(f)$ -invariante. Vemos así que U es un \mathbb{K} -subespacio de V .

- Supongamos ahora que U es un \mathbb{K} -subespacio de V . Otra vez es claro que el elemento 0 de V está en U y que $v + w \in U$ siempre que $v, w \in U$. Si $\lambda \in \mathbb{k}$ y $v \in V$, entonces podemos considerar el polinomio constante $\lambda 1_{\mathbb{k}[X]}$ en $\mathbb{k}[X]$ y calcular que

$$\lambda v = \varepsilon_f(\lambda 1_{\mathbb{k}[X]})(v) = \rho([\lambda 1_{\mathbb{k}[X]}])(v) = [\lambda 1_{\mathbb{k}[X]}] \cdot v \in U.$$

Con esto concluimos que U es un \mathbb{k} -subespacio de V , y nos queda ver que es f -invariante, pero esto es fácil: si $v \in U$, entonces

$$f(v) = \varepsilon_f(X)(v) = \rho([X])(v) = [X] \cdot v \in U,$$

porque U es un \mathbb{K} -subespacio de V .

Con todas estas preparaciones podemos probar la proposición de forma muy sencilla. Sea U un subespacio f -invariante de V . En vista de lo que acabamos de hacer, U es un \mathbb{K} -subespacio de V y, de acuerdo al Corolario 1.10.5, U posee un complemento: existe un \mathbb{K} -subespacio U' de V tal que $V = U \oplus U'$ en tanto \mathbb{K} -espacio vectorial. Ahora bien, como U' es un \mathbb{K} -subespacio de V , es un \mathbb{k} -subespacio f -invariante y tenemos que $V = U \oplus U'$ en tanto \mathbb{k} -espacio vectorial. Esto prueba que f es semisimple. \square

5.8.7. El tercer ingrediente que necesitamos es la siguiente descripción parcial de los subespacios invariantes con respecto a un endomorfismo:

Proposición. *Sea $f : V \rightarrow V$ un endomorfismo de un espacio vectorial de dimensión finita, sea m_f su polinomio minimal, y sean $p_1, \dots, p_r \in \mathbb{k}[X]$ polinomios mónicos y coprimos dos a dos tales que $m_f = p_1 \cdots p_r$. Si para cada $i \in \llbracket r \rrbracket$ ponemos $V_i := \text{Nu}(p_i(f))$, de manera que según la Proposición 5.7.1 es $V = V_1 \oplus \cdots \oplus V_r$, y W es un subespacio f -invariante de V , entonces*

$$W = (W \cap V_1) \oplus \cdots \oplus (W \cap V_r).$$

Demostración. Si para cada $i \in \llbracket r \rrbracket$ ponemos $V_i := \text{Nu}(p_i(f))$, entonces la Proposición 5.7.1 nos dice que

$$V = V_1 \oplus \cdots \oplus V_r \quad (29)$$

y que para cada $i \in \llbracket r \rrbracket$ el subespacio V_i es f -invariante, que el polinomio minimal de la restricción $f_{V_i} : V_i \rightarrow V_i$ de f a V_i es $m_{f_{V_i}} = p_i$, y que si $\pi_i : V \rightarrow V$ es el proyector con imagen V_i asociado a la descomposición (31) de V , entonces hay un polinomio $v_i \in \mathbb{k}[X]$ tal que $\pi_i = v_i(f)$.

Sea W un subespacio f -invariante de V . Afirmamos que

$$W = (W \cap V_1) + \cdots + (W \cap V_r). \quad (30)$$

En efecto, es claro que la suma de la derecha está contenida en W , ya que todos los sumandos lo están. Por otro lado, si $w \in W$, entonces existen $w_1 \in V_1, \dots, w_r \in V_r$ tales que $w = w_1 + \cdots + w_r$ y para cada $i \in \llbracket r \rrbracket$ es

$$V_i \ni w_i = \pi_i(w) = v_i(f)(w) \in W,$$

ya que W es $v_i(f)$ -invariante porque es f -invariante: vemos así que $w_i \in W \cap V_i$ y, en consecuencia, que

$$w = w_1 + \cdots + w_r \in (W \cap V_1) + \cdots + (W \cap V_r).$$

Esto prueba la igualdad (30) y, como la suma de (29) es directa, la Proposición 1.9.8 implica que la suma es de hecho directa. \square

5.8.8. Ya tenemos todas las piezas necesarias para probar nuestro segundo resultado:

Teorema. Sea V un espacio vectorial de dimensión finita y sea $f : V \rightarrow V$ un endomorfismo de V . El endomorfismo f es semisimple si y solamente si su polinomio minimal m_f es libre de cuadrados, esto es, si no es divisible por el cuadrado de ningún polinomio no constante.

Demostración. Supongamos primero que el polinomio minimal m_f de f es libre de cuadrados, de manera que hay una factorización $m_f = p_1 \cdots p_r$ con p_1, \dots, p_r polinomios mónicos irreducibles distintos dos a dos. Si para cada $i \in \llbracket r \rrbracket$ ponemos $V_i := \text{Nu}(p_i(f))$, entonces la Proposición 5.7.1 nos dice que

$$V = V_1 \oplus \cdots \oplus V_r, \quad (31)$$

que para cada $i \in \llbracket r \rrbracket$ el subespacio V_i es f -invariante, que el polinomio minimal de la restricción $f_{V_i} : V_i \rightarrow V_i$ de f a V_i es $m_{f_{V_i}} = p_i$, y que si $\pi_i : V \rightarrow V$ es el proyector con imagen V_i asociado a la descomposición (31) de V , entonces hay un polinomio $v_i \in \mathbb{k}[X]$ tal que $\pi_i = v_i(f)$.

Sea W un subespacio f -invariante de V . De acuerdo a la Proposición 5.8.7,

$$W = (W \cap V_1) \oplus \cdots \oplus (W \cap V_r).$$

Sea ahora $i \in \llbracket r \rrbracket$. El subespacio $W \cap V_i$ de V_i es f_{V_i} -invariante. Como el polinomio minimal de f_{V_i} es irreducible, la Proposición 5.8.6 nos dice que f_{V_i} es semisimple y, por lo tanto, que existe un subespacio W'_i de V_i que es f_{V_i} -invariante y tal que $V_i = (W \cap V_i) \oplus W'_i$. Consideremos el subespacio $W' = W'_1 + \cdots + W'_r$. De acuerdo a la Proposición 1.9.8, la suma es directa, y es fácil ver que se trata de un subespacio f -invariante y que $V = W \oplus W'$. Hemos probado así que el endomorfismo f es semisimple.

Queremos probar ahora que vale la implicación recíproca y lo hacemos por el absurdo. Supongamos que existen endomorfismos semisimples de espacios vectoriales de dimensión finita cuyo polinomio minimal no es libre de cuadrados y sea $f : V \rightarrow V$ uno de ellos elegido de manera que la dimensión de V sea la menor posible. Como m_f no es libre de cuadrados, existen un polinomio irreducible p , un entero $\alpha \geq 2$ y un polinomio q coprimo con p tales que $m_f = p^\alpha q$. De acuerdo a la Proposición 5.7.1, hay una descomposición $V = V_1 \oplus V_2$ de V como suma directa de subespacios f -invariantes tales que los polinomios minimales de las restricciones f_{V_1} y f_{V_2} de f a V_1 y a V_2 son p^α y q , respectivamente. De acuerdo al Lema 5.8.5, el endomorfismo $f_{V_1} : V_1 \rightarrow V_1$ es semisimple: como su polinomio minimal no es libre de cuadrados, la forma en que elegimos a f implica que debe ser $\dim V_1 = \dim V$, es decir, que es $V_2 = 0$ y, por lo tanto, $m_f = p^\alpha$.

Si $v \in V$, entonces el polinomio minimal $m_{f,v}$ divide a m_f , así que existe $\alpha_x \in \mathbb{N}_0$ tal que $m_{f,v} = p^{\alpha_x}$. Como $m_f = \text{lcm}\{m_{f,v} : v \in V\}$, claramente existe $w \in V$ tal que $\alpha_w = \alpha$. Es $\alpha \geq 2$, así que podemos considerar el vector $u := p^{\alpha-1}(f)(w)$. No es nulo, ya que $m_{f,w} = p^\alpha$, y por eso mismo es $p(f)(u) = p^\alpha(f)(w) = 0$: de esto se sigue que $m_{f,u}$ divide a p y, como p es irreducible y u no nulo, que de hecho $m_{f,u} = p$. Sea $U := \langle u \rangle_f$, que es un subespacio f -invariante de V , y que, como f es semisimple, posee un complemento U' en V que es también f -invariante.

Como $w \in V = U \oplus U'$, existen $w_1 \in U$ y $w_2 \in U'$ tales que $w = w_1 + w_2$. Es

$$u = p^{\alpha-1}(f)(w) = p^{\alpha-1}(f)(w_1) + p^{\alpha-1}(f)(w_2) \quad (32)$$

Como $w_1 \in \langle w \rangle_f$ y $m_{f,w} = p$, sabemos que $p(f)(w_1) = 0$ y, como $\alpha \geq 2$, que $p^{\alpha-1}(f)(w_1) = 0$. La igualdad (32) nos dice entonces que $U \ni u = p^{\alpha-1}(f)(w_2) \in U'$, ya que w_2 es f -invariante, y, por lo tanto, que $u \in U \cap U' = 0$: esto es absurdo. \square

5.8.9. En general, si un endomorfismo $f : V \rightarrow V$ de un espacio vectorial V es semisimple, un subespacio f -invariante de V admite muchos complementos en V que son f -invariantes. Muchas veces nos interesa que haya exactamente uno: nuestro resultado final nos dice cuándo eso ocurre.

Proposición. *Sea $f : V \rightarrow V$ un endomorfismo de un espacio vectorial V de dimensión finita.*

- (i) *Todo subespacio f -invariante de V tiene exactamente un complemento f -invariante en V si y solamente si su polinomio minimal m_f es libre de cuadrados y tiene grado igual a $\dim V$.*
- (ii) *Supongamos que las condiciones equivalentes de (i) se satisfacen, sea $m_f = p_1 \cdots p_r$ la factorización de m_f como producto de factores mónicos irreducibles distintos dos a dos y para cada $i \in \llbracket r \rrbracket$ sea $V_i := \text{Nu}(p_i(f))$. Si W es un subespacio f -invariante de V y i_1, \dots, i_t son los*

elementos del conjunto

$$\lambda(W) := \{i \in \llbracket r \rrbracket : W \cap V_i \neq 0\}$$

listados sin repeticiones, entonces

$$W = V_{i_1} \oplus \cdots \oplus V_{i_t}.$$

Si $\mathcal{I}(f)$ es el conjunto de los subespacios f -invariantes de V y $\mathcal{P}(r)$ el de todos los subconjuntos de $\llbracket r \rrbracket$, entonces la función

$$\lambda : W \in \mathcal{I}(f) \mapsto \lambda(W) \in \mathcal{P}(r)$$

es biyectiva y tal que

$$W \subseteq W' \iff \lambda(W) \subseteq \lambda(W'),$$

$$\lambda(W + W') = \lambda(W) \cup \lambda(W'),$$

$$\lambda(W \cap W') = \lambda(W) \cap \lambda(W'),$$

cada vez que W y W' son elementos de $\mathcal{I}(f)$.

Observemos que el grado del polinomio minimal de f es $\dim V$ exactamente cuando ese polinomio coincide con el polinomio característico χ_f .

Demostración. Es suficiente considerar el caso en el que $V \neq 0$ y, en particular, tal que el polinomio minimal de f tiene grado positivo.

Supongamos primero que todo subespacio f -invariante de V posee exactamente un complemento f -invariante. Por supuesto, esto implica que f es semisimple y, de acuerdo al Teorema 5.8.8, que el polinomio minimal m_f de f es libre de cuadrados. Hay entonces polinomios mónicos e irreducibles p_1, \dots, p_r distintos dos a dos y tales que $m_f = p_1 \cdots p_r$. De acuerdo a la Proposición 5.7.1 hay una descomposición $V = V_1 \oplus \cdots \oplus V_r$ de V como suma directa de subespacios f -invariantes tales que para cada $i \in \llbracket r \rrbracket$ la restricción $f_{V_i} : V_i \rightarrow V_i$ tiene polinomio minimal $m_{f_{V_i}} = p_i$.

Observemos que

$$\text{si } x \in V_1 \setminus 0, \text{ entonces } \dim \langle x \rangle_f = \deg(p_1). \quad (33)$$

En efecto, Si $x \in V_1 \setminus 0$, entonces $m_{f,x} = p_1$, ya que $m_{f,V}$ divide a $m_{f_{V_1}}$ y tiene grado positivo, y, en particular, $\dim \langle v \rangle_f = \deg(p_1)$.

Fijemos ahora $v \in V_1 \setminus 0$. Como f es semisimple, el Lema 5.8.5 nos dice que $f_{V_1} : V_1 \rightarrow V_1$ también lo es y existe por lo tanto un subespacio f -invariante U de V_1 tal que $V_1 = \langle v \rangle_f \oplus U$. Supongamos por un momento que $U \neq 0$, de manera que podemos elegir un vector $w \in U \setminus 0$. Como $v + w$ es un elemento nulo de V_1 , de (33) sabemos que $\dim \langle v + w \rangle_f = \deg(p_1)$.

Queremos ver ahora que $V_1 = \langle v + w \rangle_f \oplus U$. Como $v + w \in \langle v + w \rangle_f$ y $w \in U$, tenemos que $v \in \langle v + w \rangle_f + U$ y, como este último subespacio de V es f -invariante, que $\langle v \rangle_f \subseteq \langle v + w \rangle_f + U$.

De esto se deduce inmediatamente que $V = \langle v \rangle_f + U \subseteq \langle v + w \rangle_f + U$. De manera similar, como $v + w \in \langle v \rangle_f + U$, tenemos que $\langle v + w \rangle_f \subseteq \langle v \rangle_f + U$, porque este último espacio es f -invariante, y entonces $\langle v + w \rangle + U \subseteq \langle v \rangle_f + U \subseteq V_1$. Esto prueba que $V_1 = \langle v + w \rangle_f + U$. Ahora bien, como $v + w$ es un elemento no nulo de V_1 , de (33) sabemos que $\dim \langle v + w \rangle_f = \deg(p_1)$ y, por lo tanto,

$$\dim \langle v + w \rangle_f \cap U = \dim V_1 - \dim \langle v + w \rangle_f - \dim U = \dim V_1 - \dim \langle v \rangle_f - \dim U = 0,$$

ya que $V_1 = \langle v_f \rangle \oplus U$: esto significa que la suma $\langle v + w \rangle_f + U$ es directa.

El subespacio $W := U \oplus V_1 \oplus \dots \oplus V_r$ es f -invariante y claramente $V = \langle v \rangle_f \oplus W = \langle v + w \rangle_f \oplus W$. La hipótesis que hicimos sobre f implica entonces que $\langle v \rangle_f = \langle v + w \rangle_f$ y, por lo tanto, que $w \in \langle v \rangle_f$. Esto es absurdo, ya que $0 \neq w \in U$ y $\langle v \rangle_f \cap U = 0$.

Esta contradicción provino de haber supuesto que $U \neq 0$ y prueba, entonces, que $V_1 = \langle v \rangle_f$ y, en particular, que $\dim V_1 = \deg(p_1)$. Por supuesto, el mismo razonamiento de aplica a cada uno de los subespacios V_n, \dots, V_r y, por lo tanto,

$$\dim V = \dim V_1 + \dots + \dim V_r = \deg(p_1) + \dots + \deg(p_r) = \deg(m_f).$$

Esto completa la prueba de la necesidad de la condición del enunciado de la proposición.

Supongamos ahora que $f : V \rightarrow V$ es un endomorfismo semisimple con polinomio minimal de grado $\dim V$. Sabemos que el polinomio minimal m_f es un producto $p_1 \cdots p_r$ de polinomios mónicos e irreducibles p_1, \dots, p_r distintos dos a dos, y que si ponemos $V_i := \text{Nu}(p_i(f))$ para cada $i \in \llbracket r \rrbracket$, entonces V_i es f -invariante y no nulo, la restricción $f_{V_i} : V_i \rightarrow V_i$ tiene polinomio minimal p_i , y que $V = V_1 \oplus \dots \oplus V_r$.

Para cada $i \in \llbracket r \rrbracket$ sea $v_i \in V_i \setminus 0$. El subespacio V_i es f -invariante, así que $\langle v_i \rangle_f \subseteq V_i$. Como $m_{f,v} = p_i$ y, por lo tanto, $\dim \langle v_i \rangle_f = \deg(p_i)$, tenemos que

$$\begin{aligned} \dim V &= \deg(m_f) = \deg(p_1) + \dots + \deg(p_r) = \dim \langle v_1 \rangle_f + \dots + \dim \langle v_r \rangle_f \\ &\leq \dim V_1 + \dots + \dim V_r = \dim V. \end{aligned}$$

Esto implica que $\dim \langle v_1 \rangle_f + \dots + \dim \langle v_r \rangle_f = \dim V_1 + \dots + \dim V_r$ y, como $\langle v_i \rangle_f \subseteq V_i$ para cada $i \in \llbracket r \rrbracket$ y la suma $V_1 \oplus \dots \oplus V_r$ es directa, que, de hecho, $V_i = \langle v_i \rangle_f$ para todo $i \in \llbracket r \rrbracket$.

Sea ahora W un subespacio f -invariante de V . De acuerdo a la Proposición 5.8.7, tenemos que

$$W = (W \cap V_1) \oplus \dots \oplus (W \cap V_r).$$

Si $i \in \llbracket r \rrbracket$ es tal que $W \cap V_i \neq 0$ y v es un vector no nulo de esa intersección, entonces como $\langle v \rangle_f \subseteq W \cap V_i \subseteq V_i$ y $\dim \langle v \rangle_f = \deg(p_i) = \dim V_i$ tenemos que, de hecho, $W \cap V_i = V_i$. Esto nos dice que si ponemos

$$\lambda(W) = \{i \in \llbracket r \rrbracket : W \cap V_i \neq 0\}$$

y i_1, \dots, i_t son los elementos de $\lambda(W)$ listados sin repeticiones, entonces

$$W = V_{i_1} \oplus \dots \oplus V_{i_t}.$$

Sea ahora W' un complemento f -invariante de W en V . Lo que ya hicimos nos dice que si j_1, \dots, j_s son los elementos de $\lambda(W')$ listados sin repeticiones, entonces

$$W' = V_{j_1} \oplus \cdots \oplus V_{j_s}.$$

Como $V = W \oplus W'$, es evidente que $\lambda(W') = [\![r]\!] \setminus \lambda(W)$ y, por lo tanto, W' está únicamente determinado por W .

Con esto hemos probado la parte (i) de la proposición y la primera afirmación de la parte (ii). La verificación del resto de (ii) es inmediata. \square

5.8.10. Corolario. *Sea $f : V \rightarrow V$ un endomorfismo de un espacio vectorial de dimensión finita. Si todo subespacio f -invariante de V tiene exactamente un complemento f -invariante en V , entonces el conjunto $\mathcal{I}(f)$ de los subespacios f -invariantes de V es un reticulado finito y distributivo de subespacios de V .*

Demostración. Ambas afirmaciones se deducen inmediatamente de la segunda parte de la Proposición 5.8.9. \square

Capítulo 6

Formas normales

Supongamos que X es un conjunto y \sim una relación de equivalencia en X , de manera que para cada $x \in X$ tenemos la clase de equivalencia $[x]$ de x con respecto a la relación \sim y una partición $X/\sim := \{[x] : x \in X\}$ del conjunto X . En esta situación, un problema importante es el de decidir, dados dos elementos x e y de X , si $x \sim y$ o no. Una forma de hacer esto consiste en construir una función $v : X \rightarrow X$ de manera tal que se tenga

$$x \sim v(x), \quad x \sim y \iff v(x) = v(y)$$

para todo x e y en X . La primera condición nos dice que $v(x) \in [x]$ cualquier sea x en X , así que $v(x)$ es un *representante* de la clase de x . Por otro lado, la segunda condición nos dice que este representante depende solamente de la clase de x y no realmente de x . Tener una función que satisface estas dos condiciones nos permite reducir el problema de decidir si dos elementos x e y de X están *relacionados* por \sim al de decidir si los dos elementos $v(x)$ y $v(y)$ son *iguales*. Llamamos a una tal función $v : X \rightarrow X$ una **función de normalización** con respecto a la relación \sim , y si $x \in X$, llamamos a $v(x)$ la **forma normal** de x con respecto a v .

En el álgebra lineal aparecen frecuentemente problemas de este tipo: hay distintos conjuntos dotados de relaciones de equivalencia útiles, y se plantea entonces el problema de encontrar formas normales para poder resolver el problema de equivalencia. En este capítulo plantearemos algunos de esos problemas y los resolveremos.

§1. Equivalencia de funciones lineales y de matrices

6.1.1. Sean V y W dos espacios vectoriales. Decimos que dos funciones lineales $f, g : V \rightarrow W$ son *equivalentes* si hay automorfismos $\alpha : V \rightarrow V$ y $\beta : W \rightarrow W$ tales que $f = \beta \circ g \circ \alpha$, y en ese caso escribimos $f \sim g$. Esta relación en el conjunto $\text{hom}(V, W)$ de todos las funciones lineales $V \rightarrow W$ es una relación de equivalencia:

- Si $f \in \text{hom}(V, W)$, entonces por supuesto es $f = \text{id}_W \circ f \circ \text{id}_V$ y, como las funciones $\text{id}_V : V \rightarrow V$ e $\text{id}_W : W \rightarrow W$ son automorfismos de V y de W , respectivamente, se tiene que $f \sim f$.
- Si $f, g \in \text{hom}(V, W)$ son tales que $f \sim g$, entonces existen automorfismos $\alpha : V \rightarrow V$ y $\beta : W \rightarrow W$ tales que $f = \beta \circ g \circ \alpha$ y, por lo tanto, $g = \beta^{-1} \circ f \circ \alpha^{-1}$, de manera que $g \sim f$.
- Si $f, g, h \in \text{hom}(V, W)$ son tales que $f \sim g$ y $g \sim h$ y $\alpha, \alpha' : V \rightarrow V$ y $\beta, \beta' : W \rightarrow W$ son automorfismos tales que $f = \beta \circ g \circ \alpha$ y $g = \beta' \circ h \circ \alpha'$, entonces $f = (\beta \circ \beta') \circ h \circ (\alpha' \circ \alpha)$. Como las composiciones $\alpha' \circ \alpha$ y $\beta \circ \beta'$ son automorfismos de V y de W , respectivamente, vemos así que $f \sim h$.

6.1.2. La relación de equivalencia de funciones lineales tiene una descripción bien concreta en términos de matrices:

Proposición. *Sean V y W dos espacios vectoriales de dimensión finita. Dos funciones lineales $f, g : V \rightarrow W$ son equivalentes si y solamente si existen bases ordenadas $\mathcal{B}_1, \mathcal{B}'_1$ de V y $\mathcal{B}_2, \mathcal{B}'_2$ de W tales que*

$$[f]_{\mathcal{B}_2}^{\mathcal{B}_1} = [g]_{\mathcal{B}_2'}^{\mathcal{B}_1'}. \quad (1)$$

Demostración. Sean $f, g : V \rightarrow W$ funciones lineales y sean n y m las dimensiones de V y de W , respectivamente. Para ver la suficiencia de la condición, supongamos que f y g son equivalentes y sean $\alpha : V \rightarrow V$ y $\beta : W \rightarrow W$ automorfismos tales que $f = \beta \circ g \circ \alpha$. Fijemos además bases ordenadas $\mathcal{B}_1 = (x_1, \dots, x_n)$ y $\mathcal{B}_2 = (y_1, \dots, y_m)$ de V y de W . Como α y β^{-1} son automorfismos de V y de W , respectivamente, $\mathcal{B}'_1 = (\alpha(x_1), \dots, \alpha(x_n))$ y $\mathcal{B}'_2 = (\beta^{-1}(y_1), \dots, \beta^{-1}(y_m))$ también son bases ordenadas de V y de W , y podemos considerar la matriz $[g]_{\mathcal{B}_2'}^{\mathcal{B}_1'} = (a_{i,j}) \in M_{m,n}(\mathbb{k})$. Esto significa que para cada $i \in \llbracket n \rrbracket$ es

$$g(\alpha(x_i)) = a_{1,i}\beta^{-1}(y_1) + \dots + a_{m,1}\beta^{-1}(y_m)$$

y entonces, aplicando la función β a ambos lados de esta igualdad, que

$$f(x_i) = \beta(g(\alpha(x_i))) = a_{1,i}y_1 + \dots + a_{m,1}y_m.$$

Vemos así que $[f]_{\mathcal{B}_2}^{\mathcal{B}_1} = (a_{i,j})$ y, en definitiva, que vale la igualdad (1) del enunciado.

Mostremos ahora la suficiencia de la condición. Supongamos que hay bases ordenadas $\mathcal{B}_1 = (x_1, \dots, x_n)$ y $\mathcal{B}'_1 = (x'_1, \dots, x'_n)$ de V y $\mathcal{B}_2 = (y_1, \dots, y_m)$ y $\mathcal{B}'_2 = (y'_1, \dots, y'_m)$ de W

tales que la igualdad (1) del enunciado se cumple. Hay funciones lineales $\alpha : V \rightarrow V$ y $\beta : W \rightarrow W$ tales que $\alpha(x_i) = x'_i$ para cada $i \in \llbracket n \rrbracket$ y $\beta(y'_j) = y_j$ para cada $j \in \llbracket m \rrbracket$, y estas funciones son isomorfismos. Como claramente $[\alpha]_{\mathcal{B}_1}^{\mathcal{B}_1'} = I_n$ y $[\beta]_{\mathcal{B}_2}^{\mathcal{B}_2'} = I_m$, usando la Proposición 2.6.7 y la hipótesis vemos que

$$[\beta \circ g \circ \alpha]_{\mathcal{B}_2}^{\mathcal{B}_1} = [\beta]_{\mathcal{B}_2}^{\mathcal{B}_2'} \cdot [g]_{\mathcal{B}_2'}^{\mathcal{B}_1'} \cdot [\alpha]_{\mathcal{B}_1}^{\mathcal{B}_1'} = I_m \cdot [g]_{\mathcal{B}_2}^{\mathcal{B}_1'} \cdot I_n = [g]_{\mathcal{B}_2}^{\mathcal{B}_1'} = [f]_{\mathcal{B}_2}^{\mathcal{B}_1}.$$

Así, las funciones lineales $\beta \circ g \circ \alpha$ y f tienen las mismas matrices con respecto a las bases \mathcal{B}_1 y \mathcal{B}_2 , y esto implica que son ellas mismas iguales, de manera que las funciones f y g son equivalentes. \square

6.1.3. La proposición que sigue resuelve de una manera satisfactoria el problema de decidir cuándo dos funciones lineales son equivalentes, ya que lo reduce al cálculo de los rangos de esas funciones.

Proposición. *Sean V y W espacios vectoriales de dimensión finita n y m , respectivamente.*

- (i) *Si $f : V \rightarrow W$ es una función lineal de rango r , entonces existen bases \mathcal{B} de V y \mathcal{B}' de W tales que la matriz $[f]_{\mathcal{B}}^{\mathcal{B}'}$ es la matriz de bloques*

$$\begin{matrix} r & n-r \\ \hline \begin{matrix} I_r & 0 \\ 0 & 0 \end{matrix} \\ m-r \end{matrix}$$

- (ii) *Dos funciones lineales $V \rightarrow W$ son equivalentes si y solamente si tienen el mismo rango*
 (iii) *El número de clases de equivalencia de elementos de $\text{hom}(V, W)$ es $\min\{\dim V, \dim W\} + 1$.*

Demostración. (i) Sea $f : V \rightarrow W$ una función lineal de rango r , de manera que $\dim \text{Im}(f) = r$. Del teorema de la dimensión 2.4.1 sabemos que

$$\dim \text{Nu}(f) = \dim V - \dim \text{Im}(f) = n - r,$$

y entonces existe una base ordenada $\mathcal{B} = (x_1, \dots, x_n)$ de V tal que (x_{r+1}, \dots, x_n) es una base de $\text{Nu}(f)$. Mostremos que $(f(x_1), \dots, f(x_r))$ es una base del subespacio $\text{Im}(f)$ de W :

- Si $y \in \text{Im}(f)$, entonces existe $x \in V$ tal que $y = f(x)$ y, como \mathcal{B} es una base de V , hay escalares $a_1, \dots, a_n \in \mathbb{k}$ tales que $x = a_1x_1 + \dots + a_nx_n$. Aplicando la función f a ambos lados de esta igualdad y recordando que los vectores x_{r+1}, \dots, x_n están en el núcleo de f , vemos que $y = f(x) = a_1f(x_1) + \dots + a_rf(x_r)$. Esto muestra que $\text{Im}(f) = \langle f(x_1), \dots, f(x_r) \rangle$.
- Por otro lado, supongamos que $b_1, \dots, b_r \in \mathbb{k}$ son escalares tales que

$$b_1f(x_1) + \dots + b_rf(x_r) = 0.$$

Esto implica que $f(b_1x_1 + \dots + b_rx_r) = 0$ y, por lo tanto, que el vector $b_1x_1 + \dots + b_rx_r$ está en $\text{Nu}(f)$. En vista de la forma en que elegimos la base \mathcal{B} , entonces, existen escalares $c_{r+1}, \dots, c_n \in \mathbb{k}$ tales que

$$b_1x_1 + \dots + b_rx_r = c_{r+1}x_{r+1} + \dots + c_nx_n$$

y, como el conjunto \mathcal{B} es linealmente independiente, esto implica que $b_1 = \dots = b_r = 0$.

Ahora, como $(f(x_1), \dots, f(x_r))$ es una base de $\text{Im}(f)$, podemos elegir vectores $y_1, \dots, y_{m-r} \in W$ de manera tal que $\mathcal{B}' = (f(x_1), \dots, f(x_r), y_1, \dots, y_{m-r})$ sea una base ordenada de W . Gracias a todas estas elecciones, es inmediato verificar que la matriz de f con respecto a las bases ordenadas \mathcal{B} y \mathcal{B}' es la matriz de bloques que aparece en el enunciado.

(ii) Sean $f, g : V \rightarrow W$ dos funciones lineales. Si f y g tienen el mismo rango r , entonces la primera parte de la proposición nos dice que existen bases ordenadas \mathcal{B}_1 y \mathcal{B}'_1 de V y \mathcal{B}_2 y \mathcal{B}'_2 de W tales que las matrices $[f]_{\mathcal{B}_2}^{\mathcal{B}_1}$ y $[g]_{\mathcal{B}_2}^{\mathcal{B}'_1}$ son iguales a la matriz de bloques que aparece en su enunciado. Se sigue, claro, que estas dos matrices son iguales entre sí y, de acuerdo a la Proposición 6.1.2, que las funciones f y g son equivalentes.

Recíprocamente, si f y g son equivalentes, esa proposición nos dice que existen bases ordenadas \mathcal{B}_1 y \mathcal{B}'_1 de V y \mathcal{B}_2 y \mathcal{B}'_2 de W tales que $[f]_{\mathcal{B}_2}^{\mathcal{B}_1} = [g]_{\mathcal{B}_2}^{\mathcal{B}'_1}$ y, como el rango de f coincide con el de la matriz $[f]_{\mathcal{B}_2}^{\mathcal{B}_1}$ y el de f con el de $[g]_{\mathcal{B}_2}^{\mathcal{B}'_1}$, es claro que f y g tienen el mismo rango.

(iii) Sea $d := \min\{\dim V, \dim W\}$. Si $f : V \rightarrow W$ es una función lineal, entonces

$$0 \leq \dim \text{Im}(f) \leq d,$$

de manera que hay $d+1$ posibles valores para $\dim \text{Im}(f)$. Como dos funciones lineales con el mismo rango son equivalentes, esto nos dice que hay a lo sumo $d+1$ clases de equivalencia en $\text{hom}(V, W)$. Para ver que son exactamente $d+1$ es suficiente con mostrar que para todo $r \in \llbracket 0, d \rrbracket$ existe alguna función $V \rightarrow W$ de rango r . Esto es fácil: si $\mathcal{B} = (x_1, \dots, x_n)$ y $\mathcal{B}' = (y_1, \dots, y_m)$, entonces sabemos que hay una función lineal $f : V \rightarrow W$ tal que

$$f(x_i) = \begin{cases} y_i, & \text{si } i \leq r; \\ 0, & \text{en caso contrario,} \end{cases}$$

y es inmediato que $\dim \text{Im}(f) = r$. □

6.1.4. Como con todo, la noción de equivalencia de funciones lineales se transpone al contexto de las matrices. Si $m, n \in \mathbb{N}$, decimos que dos matrices $A, B \in M_{m,n}(\mathbb{k})$ son **equivalentes** si existen matrices invertibles $P \in GL_m(\mathbb{k})$ y $Q \in GL_n(\mathbb{k})$ tales que $A = PBQ$ y en ese caso escribimos $A \sim B$. Un argumento similar al de 6.1.1 muestra que esto define una relación de equivalencia en el conjunto $M_{m,n}(\mathbb{k})$.

Lema. Sean $m, n \in \mathbb{N}$ y sean $A, B \in M_{m,n}(\mathbb{k})$. Las siguientes afirmaciones son equivalentes:

- (a) Las matrices A y B son equivalentes.
- (b) Las funciones lineales $f_A : x \in \mathbb{k}^n \mapsto Ax \in \mathbb{k}^m$ y $f_B : x \in \mathbb{k}^n \mapsto Bx \in \mathbb{k}^m$ son equivalentes.

Demostración. Si A y B son equivalentes, entonces existen matrices invertibles $P \in GL_m(\mathbb{k})$ y $Q \in GL_n(\mathbb{k})$ tales que $A = PBQ$. Las funciones $\alpha : x \in \mathbb{k}^n \mapsto Qx \in \mathbb{k}^n$ y $\beta : x \in \mathbb{k}^m \mapsto Px \in \mathbb{k}^m$ son entonces automorfismos y es inmediato verificar que $f_A = \beta \circ f_B \circ \alpha$, de manera que f_A es equivalente a f_B .

Recíprocamente, si las funciones lineales f_A y f_B son equivalentes, existen automorfismos $\alpha : \mathbb{k}^n \rightarrow \mathbb{k}^n$ y $\beta : \mathbb{k}^m \rightarrow \mathbb{k}^m$ tales que $f_A = \beta \circ f_B \circ \alpha$ y, si \mathcal{B} y \mathcal{B}' son las bases ordenadas estándares de \mathbb{k}^n y de \mathbb{k}^m , las matrices $Q = [\alpha]_{\mathcal{B}}^{\mathcal{B}}$ y $P = [\beta]_{\mathcal{B}'}^{\mathcal{B}'}$ son inversibles y se tiene que

$$A = [f_A]_{\mathcal{B}'}^{\mathcal{B}} = [\beta \circ f_B \circ \alpha]_{\mathcal{B}'}^{\mathcal{B}} = [\beta]_{\mathcal{B}'}^{\mathcal{B}'} \cdot [f_B]_{\mathcal{B}'}^{\mathcal{B}} \cdot [\alpha]_{\mathcal{B}}^{\mathcal{B}} = PBQ.$$

Esto nos dice que A y B son matrices equivalentes. \square

6.1.5. El resultado correspondiente a la Proposición 6.1.3 para matrices es:

Proposición. Sean $m, n \in \mathbb{N}$.

- (i) Dos matrices de $M_{m,n}(\mathbb{k})$ son equivalentes si y solamente si tienen el mismo rango.
- (ii) El número de clases de equivalencia en $M_{m,n}(\mathbb{k})$ es $\min\{m, n\} + 1$.
- (iii) Toda matriz de $M_{m,n}(\mathbb{k})$ es equivalente a exactamente una de la forma

$$\begin{matrix} & \overbrace{\quad\quad\quad}^{r} \quad \overbrace{\quad\quad\quad}^{n-r} \\ r & \left[\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix} \right] \\ & \overbrace{\quad\quad\quad}^{m-r} \end{matrix} \quad (2)$$

con $r \in [0, \min\{m, n\}]$ y dos matrices de éstas son equivalentes solamente si son iguales.

Demostración. Sean $A, B \in M_{m,n}(\mathbb{k})$ y sean $f_A : x \in \mathbb{k}^n \mapsto Ax \in \mathbb{k}^m$ y $f_B : x \in \mathbb{k}^n \mapsto Bx \in \mathbb{k}^m$. Sabemos del Lema 6.1.4 que las matrices A y B son equivalentes si y solamente si las funciones f_A y f_B lo son, y esto ocurre, según la Proposición 6.1.3(ii), exactamente cuando f_A y f_B tienen el mismo rango. Como A y f_A , por un lado, y B y f_B , por otro, tienen el mismo rango, esto prueba (i). Por otro lado, para ver (ii) basta observar que el rango de una matriz de $M_{m,n}(\mathbb{k})$ es un elemento de $[0, \min\{m, n\}]$ y cada uno los $\min\{m, n\} + 1$ posibles valores ocurre efectivamente como el rango de una matriz de $M_{m,n}(\mathbb{k})$.

Sea, finalmente, $A \in M_{m,n}(\mathbb{k})$ y sea $f_A : x \in \mathbb{k}^n \mapsto Ax \in \mathbb{k}^m$. De la Proposición 6.1.3(i) sabemos que existen bases ordenadas \mathcal{B}_1 y \mathcal{B}'_1 de \mathbb{k}^n y de \mathbb{k}^m tal que la matriz $[f_A]_{\mathcal{B}'_1}^{\mathcal{B}_1}$ es precisamente la matriz (2) del enunciado. Si \mathcal{B}_2 y \mathcal{B}'_2 son las bases ordenadas estándares de \mathbb{k}^n y \mathbb{k}^m , se tiene entonces que

$$A = [f_A]_{\mathcal{B}'_2}^{\mathcal{B}_2} = C(\mathcal{B}'_1, \mathcal{B}'_2) \cdot [f_A]_{\mathcal{B}'_1}^{\mathcal{B}_1} \cdot C(\mathcal{B}_2, \mathcal{B}_1)$$

y, como las matrices de cambio de base $C(\mathcal{B}'_1, \mathcal{B}'_2)$ y $C(\mathcal{B}_2, \mathcal{B}_1)$ pertenecen a $GL_m(\mathbb{k})$ y a $GL_n(\mathbb{k})$, esto muestra que A es equivalente a la matriz $[f_A]_{\mathcal{B}'_1}^{\mathcal{B}_1}$, esto es, a la matriz (2) del enunciado. Para terminar, basta observar que esta última matriz tiene rango r , así que si es equivalente a otra de la misma forma necesariamente tienen que ser exactamente iguales. \square

Operaciones de filas y columnas

6.1.6. Fijemos $n \in \mathbb{N}$.

- Si $i, j \in \llbracket n \rrbracket$, escribimos $E_{i,j}^n$ a la matriz de $M_n(\mathbb{k})$ cuya única entrada no nula es la (i, j) -ésima, que es igual a 1.
- Si $i, j \in \llbracket n \rrbracket$ son tales que $i \neq j$, escribimos $T_{i,j}^n$ a la matriz de permutación de $M_n(\mathbb{k})$ correspondiente a la transposición $(i\ j)$. Claramente, $T_{i,j}^n$ es inversible y, de hecho, $(T_{i,j}^n)^{-1} = T_{i,j}$.
- Si $i, j \in \llbracket n \rrbracket$ son tales que $i \neq j$ y $\lambda \in \mathbb{k}$ es un escalar, ponemos

$$L_{i,j}^n(\lambda) = I_n + \lambda E_{i,j}^n.$$

Se trata de una matriz inversible y es fácil verificar que $(L_{i,j}^n(\lambda))^{-1} = L_{i,j}^n(-\lambda)$.

- Si $i \in \llbracket n \rrbracket$ y $\lambda \in \mathbb{k} \setminus 0$, escribimos

$$D_i^n(\lambda) = I_n + (\lambda - 1)E_{i,i}^n.$$

Otra vez, esta matriz es inversible y $(D_i^n)(\lambda)^{-1} = D_i^n(\lambda^{-1})$.

Llamamos a las matrices

$$\begin{aligned} T_{i,j}^n &\quad \text{con } i, j \in \llbracket n \rrbracket \text{ tales que } i \neq j, \\ L_{i,j}^n(\lambda) &\quad \text{con } i, j \in \llbracket n \rrbracket \text{ y } \lambda \in \mathbb{k} \text{ tales que } i \neq j, \text{ y} \\ D_i^n(\lambda) &\quad \text{con } i \in \llbracket n \rrbracket \text{ y } \lambda \in \mathbb{k} \setminus 0 \end{aligned}$$

las **matrices elementales** de $M_n(\mathbb{k})$. En la Figura 6.1 de la página siguiente están ilustradas algunas. Se sigue de lo anterior que todas las matrices elementales son inversibles y que sus inversas son también matrices elementales. Como consecuencia inmediata de esto, vemos que

todo producto de matrices elementales es inversible y su matriz inversa es también igual a un producto de matrices elementales.

6.1.7. Si $m, n \in \mathbb{N}$ y $A \in M_{m,n}(\mathbb{k})$, entonces calculando directamente es fácil ver que:

- Si $i, j \in \llbracket n \rrbracket$ e $i \neq j$, la matriz $A' = AT_{i,j}^n$ es la que se obtiene de A intercambiando la columna i -ésima y la j -ésima.
- Si $i, j \in \llbracket n \rrbracket$ y $\lambda \in \mathbb{k}$ son tales que $i \neq j$, la matriz $A' = AL_{i,j}^n(\lambda)$ es la que se obtiene de A sumándole a su columna j -ésima el resultado de multiplicar a su columna i -ésima por λ .
- Si $i \in \llbracket n \rrbracket$ y $\lambda \in \mathbb{k} \setminus 0$, la matriz $A' = AD_i^n(\lambda)$ es la que se obtiene de A multiplicando su columna i -ésima por λ .

Decimos en cada uno de estos casos que la matriz A' se obtiene de A haciendo una **operación de columnas**. De manera similar:

- Si $i, j \in \llbracket m \rrbracket$ e $i \neq j$, la matriz $A'' = T_{i,j}^mA$ es la que se obtiene de A intercambiando la fila i -ésima y la j -ésima.
- Si $i, j \in \llbracket m \rrbracket$ y $\lambda \in \mathbb{k}$ son tales que $i \neq j$, la matriz $A'' = L_{i,j}^m(\lambda)A$ es la que se obtiene de A sumándole a su fila i -ésima el resultado de multiplicar a su fila j -ésima por λ .

$$T_{2,4}^6 = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \quad L_{1,5}^6(\lambda) = \begin{pmatrix} 1 & 0 & 0 & 0 & \lambda & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

$$D_3^6(\lambda) = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & \lambda & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

Figura 6.1. Tres matrices elementales. Las entradas pintadas en azul son las únicas en las que estas matrices difieren de la matriz identidad I_6 .

- Si $i \in \llbracket m \rrbracket$ y $\lambda \in \mathbb{k} \setminus 0$, la matriz $A'' = D_i^m(\lambda)A$ es la que se obtiene de A multiplicando su fila i -ésima por λ .

En cada uno de estos casos decimos que la matriz A'' se obtiene de A con la que empezamos haciendo una **operación de filas**.

6.1.8. Una de las razones por las que las matrices elementales son importantes es la siguiente:

Proposición. Sea $n \in \mathbb{N}$. Toda matriz de $\mathrm{GL}_n(\mathbb{k})$ es igual a un producto de matrices elementales.

Demostración. Sea (e_1, \dots, e_n) la base ordenada estándar de \mathbb{k}^n . Mostremos que

existen matrices P_0, \dots, P_n , cada una de ellas igual a un producto de matrices elementales, tales que para cada $i \in \llbracket 0, n \rrbracket$ las primeras i columnas del producto $P_i \cdots P_0 A$ son exactamente los vectores e_1, \dots, e_i , en orden. (3)

Esto probará la proposición ya que entonces tendremos que $P_n \cdots P_0 A = I_n$, ya que la única matriz de $\mathrm{GL}_n(\mathbb{k})$ que tiene sus n columnas iguales a los vectores e_1, \dots, e_n en orden es la matriz identidad I_n y, por lo tanto, $A = P_0^{-1} \cdots P_n^{-1}$ es un producto de matrices elementales.

Construimos las matrices P_0, \dots, P_n inductivamente, empezando con la observación de que podemos tomar simplemente $P_0 = I_n$. Supongamos entonces que $r \in \llbracket 0, n-1 \rrbracket$ y que ya construimos matrices P_0, \dots, P_r , cada una de ellas igual a un producto de matrices elementales, de manera que el producto $B = P_r \cdots P_0 A$ tiene sus primeras r columnas iguales a los vectores e_1, \dots, e_r , en orden.

La matriz B tiene entonces la forma

$$\begin{pmatrix} 1 & b_{1,r+1} & b_{1,r+2} & \cdots & b_{1,n} \\ \ddots & \vdots & \vdots & \ddots & \vdots \\ 1 & b_{r,r+1} & b_{r,r+2} & \cdots & b_{r,n} \\ 0 & \cdots & 0 & b_{r+1,r+1} & b_{r+1,r+2} & \cdots & b_{r+1,n} \\ 0 & \cdots & 0 & b_{r+2,r+1} & b_{r+2,r+2} & \cdots & b_{r+2,n} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & b_{n,r+1} & b_{n,r+2} & \cdots & b_{n,n} \end{pmatrix}$$

Como B es inversible, sus columnas son elementos linealmente independientes de \mathbb{k}^n . Tiene que existir entonces $s \in [r+1, n]$ tal que $b_{s,r+1} \neq 0$: si no fuese ése el caso, la columna $(r+1)$ -ésima de B sería una combinación lineal de las primeras r . Si $s = r+1$, pongamos $C_0 = D_s^n(b_{r+1,r+1}^{-1})$ y si $s \neq r+1$ pongamos $C_0 = D_s^n(b_{s,r+1}^{-1})T_{s,r+1}$. En cualquier caso, tenemos que la matriz

$$B' = C_0 B$$

es de la forma

$$\begin{pmatrix} 1 & b'_{1,r+1} & b'_{1,r+2} & \cdots & b'_{1,n} \\ \ddots & \vdots & \vdots & \ddots & \vdots \\ 1 & b'_{r,r+1} & b'_{r,r+2} & \cdots & b'_{r,n} \\ 0 & \cdots & 0 & 1 & b'_{r+1,r+2} & \cdots & b'_{r+1,n} \\ 0 & \cdots & 0 & b'_{r+2,r+1} & b'_{r+2,r+2} & \cdots & b'_{r+2,n} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & b'_{n,r} & b'_{n,r+2} & \cdots & b'_{n,n} \end{pmatrix}$$

con un 1 en la entrada $(r+1, r+1)$ -ésima. Consideremos ahora la matriz B'' que se obtiene de B reemplazando, para cada $i \in [1, n] \setminus \{r+1\}$, la fila i -ésima por el resultado de restarle a esa fila la fila $(r+1)$ -ésima multiplicada por $b_{i,r+1}$: de acuerdo a nuestras observaciones de 6.1.7, esto significa que

$$B'' = L_{n,r+1}^n(-b_{n,r+1}) \cdots \widehat{L_{n,r+1}^n(-b_{n,r+1})} \cdots L_{r+2,r+1}^n(-b_{r+2,r+1}) B'$$

La matriz B'' es de la forma

$$\begin{pmatrix} 1 & 0 & b''_{1,r+2} & \cdots & b''_{1,n} \\ \ddots & \vdots & \vdots & \ddots & \vdots \\ 1 & 0 & b''_{r,r+2} & \cdots & b''_{r,n} \\ 0 & \cdots & 0 & 1 & b''_{r+1,r+2} & \cdots & b''_{r+1,n} \\ 0 & \cdots & 0 & 0 & b''_{r+2,r+2} & \cdots & b''_{r+2,n} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & 0 & b''_{n,r+2} & \cdots & b''_{n,n} \end{pmatrix}$$

y tiene sus primeras $r + 1$ columnas iguales a los vectores e_1, \dots, e_{r+1} . Esto significa que si ponemos

$$P_{r+1} = \begin{cases} L_{n,r+1}^n(-b_{n,r+1}) \cdots \widehat{L_{n,r+1}^n(-b_{n,r+1})} \cdots L_{r+2,r+1}^n(-b_{r+2,r+1}) D_s^n(b_{s,r+1}^{-1}), & \text{si } s = r + 1; \\ L_{n,r+1}^n(-b_{n,r+1}) \cdots \widehat{L_{n,r+1}^n(-b_{n,r+1})} \cdots L_{r+2,r+1}^n(-b_{r+2,r+1}) D_s^n(b_{s,r+1}^{-1}) T_{s,r+1}, & \text{si } s \neq r + 1; \end{cases}$$

entonces $P_{r+1} \cdots P_0 A$ tiene sus primeras $r + 1$ columnas iguales a e_1, \dots, e_{r+1} . Como manifiestamente P_{r+1} es un producto de matrices elementales, esto prueba la afirmación (3). \square

6.1.9. Una forma de leer la Proposición 6.1.8 y la prueba que dimos de ella es la siguiente: nos dice que toda matriz inversible puede transformarse en la matriz identidad haciendo operaciones de filas. De manera completamente similar, tenemos que toda matriz inversible puede transformarse en la matriz identidad haciendo operaciones de columnas.

6.1.10. Usando la Proposición 6.1.8 podemos dar una nueva descripción de la relación de equivalencia de matrices:

Proposición. Sean $m, n \in \mathbb{N}$ y $A, B \in M_{m,n}(\mathbb{k})$. Las siguientes afirmaciones son equivalentes:

- (a) Las matrices A y B son equivalentes.
- (b) Es posible obtener la matriz A a partir de B haciendo operaciones de filas y columnas.

Demostración. Si A y B son equivalentes, entonces hay matrices inversibles $P \in GL_m(\mathbb{k})$ y $Q \in GL_n(\mathbb{k})$ tales que $A = P B Q$, y de acuerdo a la Proposición 6.1.8 existen $r, s \in \mathbb{N}_0$ y matrices elementales P_1, \dots, P_r en $GL_m(\mathbb{k})$ y $Q_1, \dots, Q_s \in GL_n(\mathbb{k})$ tales que $P = P_1 \cdots P_r$ y $Q = Q_1 \cdots Q_s$, y, por lo tanto, es

$$A = P_1 \cdots P_r B Q_1 \cdots Q_s.$$

Según nuestras observaciones de 6.1.7, esto significa que A puede obtenerse de B haciendo operaciones de filas y columnas.

Recíprocamente, si A puede obtenerse de B haciendo operaciones de filas y columnas, sabemos de 6.1.7 que existen $r, s \in \mathbb{N}_0$ y matrices elementales P_1, \dots, P_r en $GL_m(\mathbb{k})$ y $Q_1, \dots, Q_s \in GL_n(\mathbb{k})$ tales que $A = P_1 \cdots P_r B Q_1 \cdots Q_s$ y entonces, si ponemos $P = P_1 \cdots P_r$ y $Q = Q_1 \cdots Q_s$, tenemos que P y Q son inversibles y que $A = PBQ$, de manera que A y B son equivalentes. \square

§2. Equivalencia a derecha de funciones lineales y de matrices

6.2.1. La Proposición 6.1.10, junto con la Proposición 6.1.5, nos dice que dos matrices de $M_n(\mathbb{k})$ tienen el mismo rango si y solamente si una puede obtenerse de la otra haciendo operaciones de filas y columnas. Tiene sentido preguntarse qué sucede si solamente nos permitimos hacer, por ejemplo, operaciones de columnas. Ese es el problema que consideramos en esta sección.

6.2.2. Decimos que dos matrices A y B de $M_{m,n}(\mathbb{k})$ son *equivalentes a derecha*, y en ese caso escribimos $A \sim_d B$, si existe una matriz inversible $Q \in GL_n(\mathbb{k})$ tal que $A = BQ$. Es inmediato verificar que esto es una relación de equivalencia en el conjunto $M_{m,n}(\mathbb{k})$.

Proposición. Sean $m, n \in \mathbb{N}$. Si A y B son matrices de $M_{m,n}(\mathbb{k})$, entonces las siguientes tres afirmaciones son equivalentes:

- (a) Las matrices A y B son equivalentes a derecha.
- (b) La matriz A puede ser obtenida de la matriz B haciendo operaciones de columnas.
- (c) Los subespacios $\text{col}(A)$ y $\text{col}(B)$ de \mathbb{k}^m generados por las columnas de A y de B , respectivamente, coinciden.

Demostración. (a) \Rightarrow (b) Si $A \sim_d B$, entonces hay una matriz inversible $Q \in GL_n(\mathbb{k})$ tal que $A = BQ$ y, de acuerdo a la Proposición 6.1.8, existen $r \in \mathbb{N}_0$ y matrices elementales Q_1, \dots, Q_r en $GL_n(\mathbb{k})$ tales que $Q = Q_1 \cdots Q_r$: se sigue de esto que $A = BQ_1 \cdots Q_r$ y, según a nuestras observaciones de 6.1.7, que A puede obtenerse de B haciendo operaciones de filas y columnas.

(b) \Rightarrow (a) Recíprocamente, si A puede obtenerse de B haciendo operaciones de columnas, existen $r \in \mathbb{N}_0$ y matrices elementales Q_1, \dots, Q_r en $GL_n(\mathbb{k})$ tales que $A = BQ_1 \cdots Q_r$, y entonces poniendo $Q := Q_1 \cdots Q_r$ sabemos que $Q \in GL_n(\mathbb{k})$ y que $A = BQ$, de manera que $A \sim_d B$.

(a) \Rightarrow (c) Consideremos las funciones lineales $f_A : x \in \mathbb{k}^n \mapsto Ax \in \mathbb{k}^m$ y $f_B : x \in \mathbb{k}^n \mapsto Bx \in \mathbb{k}^m$. Si A y B son equivalentes a derecha, existe una matriz inversible $Q \in GL_n(\mathbb{k})$ tal que $A = BQ$. Si en ese caso ponemos $f_Q : x \in \mathbb{k}^n \mapsto Qx \in \mathbb{k}^n$, entonces f_Q es un automorfismo de \mathbb{k}^n , $f_A = f_B \circ f_Q$ y, en particular, las funciones f_A y f_B tienen la misma imagen. Como la imagen de f_A es el subespacio de \mathbb{k}^m generado por las columnas de A y la de f_B el generado por las columnas de B , vemos que $\text{col}(A) = \text{col}(B)$, como queremos.

(c) \Rightarrow (a) Supongamos ahora que $\text{col}(A) = \text{col}(B)$ y sean r y (x_1, \dots, x_r) la dimensión y una base ordenada de $\text{col}(A)$, respectivamente. Como $\text{col}(A) = \text{Im}(f_A)$, existen vectores $y_1, \dots, y_r \in \mathbb{k}^n$ tales que $f_A(y_i) = x_i$ para cada $i \in [r]$. Por otro lado, como $\dim \text{Nu}(f_A) = n - r$ y podemos elegir una base ordenada (y_{r+1}, \dots, y_n) de $\text{Nu}(f_A)$. Es fácil ver que $\mathcal{B} = (x_1, \dots, x_n)$ es una base de \mathbb{k}^n . De manera similar, como $\text{col}(B) = \text{Im}(f_B)$, hay una base ordenada $\mathcal{B}' = (z_1, \dots, z_n)$ de \mathbb{k}^n tal que $f_B(z_i) = x_i$ para cada $i \in [r]$ y (z_{r+1}, \dots, z_n) es una base de $\text{Nu}(f_B)$. Finalmente, hay una única función lineal $\alpha : \mathbb{k}^n \rightarrow \mathbb{k}^n$ tal que $\alpha(y_i) = z_i$ para todo $i \in [n]$, y se trata de un isomorfismo.

Observemos ahora que si $i \in [\![n]\!]$, entonces

$$(f_B \circ \alpha)(y_i) = f_B(z_i) = \begin{cases} x_i & \text{si } i \leq r; \\ 0 & \text{si } i > r \end{cases} = f_A(y_i),$$

así que $f_B \circ \alpha = f_A$. Si ahora \mathcal{E} y \mathcal{E}' son las bases estándar de \mathbb{k}^n y \mathbb{k}^m , respectivamente, entonces la matriz $Q = [\alpha]_{\mathcal{E}}^{\mathcal{E}'}$ es un elemento de $\mathrm{GL}_n(\mathbb{k})$ porque α es un isomorfismo, y

$$B \cdot Q = [f_B]_{\mathcal{E}'}^{\mathcal{E}} \cdot [\alpha]_{\mathcal{E}}^{\mathcal{E}'} = [f_B \circ \alpha]_{\mathcal{E}'}^{\mathcal{E}'} = [f_A]_{\mathcal{E}'}^{\mathcal{E}'} = A.$$

Esto nos dice que $A \sim_d B$. □

6.2.3. Una consecuencia de esta proposición es que hay tantas clases de equivalencia a derecha en $\mathrm{M}_{m,n}(\mathbb{k})$ como subespacios de \mathbb{k}^m . Más precisamente, tenemos el siguiente corolario, en el que, como en 1.12.13, escribimos $\mathcal{L}(\mathbb{k}^m)$ al conjunto de todos los subespacios de \mathbb{k}^m .

Corolario. *Hay exactamente una función*

$$\Phi : \mathrm{M}_{m,n}(\mathbb{k}) / \sim_d \rightarrow \mathcal{L}(\mathbb{k}^m)$$

tal que $\Phi([A]) = \mathrm{col}(A)$ para toda matriz $A \in \mathrm{M}_{m,n}(\mathbb{k})$, y se trata de una biyección.

Demostración. Que hay una y una sola función con esa propiedad y que es inyectiva es consecuencia inmediata de la Proposición 6.2.2. Veamos que es sobreyectiva. Sea V un subespacio de \mathbb{k}^m , sea r su dimensión y sea (x_1, \dots, x_r) una base ordenada de V : si A es la matriz de $\mathrm{M}_{m,n}(\mathbb{k})$ cuyas primeras r columnas son x_1, \dots, x_r y cuyas restantes $n - r$ columnas son todas nulas claramente tiene $\Phi([A]) = \mathrm{col}(A) = V$. □

6.2.4. Queremos ahora encontrar una función de normalización para la equivalencia a derecha de matrices y para ello necesitamos algunos preparativos.

Para cada $i \in [\![m]\!]$ consideramos la función lineal $\pi_i : (x_1, \dots, x_r)^t \in \mathbb{k} \mapsto x_i \in \mathbb{k}$. Si x es un elemento no nulo de \mathbb{k}^m , escribimos

$$\ell(x) := \min\{i \in [\![m]\!] : \pi_i(x) \neq 0\}$$

y decimos que x es **mónico** si $\pi_{\ell(x)}(x) = 1$, esto es, si su primera componente no nula es igual a 1. Por otro lado, si V es un subespacio no nulo de \mathbb{k}^m ponemos

$$\ell(V) := \min\{\ell(x) : x \in V \setminus 0\}, \quad \tau(V) := \{x \in V : \ell(x) > \ell(V)\}.$$

Lema. *Si V es un subespacio no nulo de \mathbb{k}^m , entonces el subconjunto $\tau(V)$ es un subespacio de V , existe un vector mónico $x_V \in V$ tal que $\ell(x_V) = \ell(V)$ y $V = \langle x_V \rangle \oplus \tau(V)$.*

Demostración. Sea V un subespacio no nulo de \mathbb{k}^m y sea r su dimensión. En vista de la definición de $\ell(V)$, existe un vector $x \in V \setminus 0$ tal que $\ell(x) = \ell(V)$ y entonces el vector $x_V := \pi_{\ell(x)}(x)^{-1}x$ es mónico y tal que $\ell(x_V) = \ell(V)$.

Por otro lado, la restricción $\phi := \pi_{\ell(V)}|_V : V \rightarrow \mathbb{k}$ es lineal y, como $\phi(x_V) = 1$, no nula: esto implica inmediatamente que el núcleo $\text{Nu}(\phi)$ tiene dimensión $r - 1$ y que $V = \langle x_V \rangle \oplus \text{Nu}(\phi)$. Si $y \in \tau(V)$, entonces $\ell(y) > \ell(V)$ y, por lo tanto, $\phi(y) = \pi_{\ell(V)}(y) = 0$ por la forma en que definimos $\ell(y)$: vemos así que $y \in \text{Nu}(\phi)$. Recíprocamente, si $y \in \text{Nu}(\phi)$ entonces por un lado es $\ell(y) \geq \ell(V)$ y, por otro, $\pi_{\ell(V)}(y) = \phi(y) = 0$, así que $\ell(y) > \ell(V)$, de manera que $y \in \tau(V')$. Esto muestra que $\text{Nu}(\phi) = \tau(V)$. \square

6.2.5. Sea V un subespacio de \mathbb{k}^m y sea r su dimensión. Una base ordenada $\mathcal{B} = (x_1, \dots, x_r)$ es *reducida* si

- $\ell(x_1) < \ell(x_2) < \dots < \ell(x_r)$,
- todos sus vectores de \mathcal{B} son mónicos, y
- cada vez que i y j son elementos distintos de $\llbracket r \rrbracket$ se tiene que $\pi_{\ell(x_j)}(x_i) = 0$.

Proposición. *Cada subespacio de \mathbb{k}^m posee exactamente una base reducida.*

Demostración. Sea V un subespacio de \mathbb{k}^m y sea r su dimensión. Probaremos la proposición haciendo inducción con respecto a r . Si $r = 0$ claramente ni hay nada que probar, así que supongamos que la dimensión r es positiva. Sean $x_V \in V$ como en el Lema 6.2.4, de manera que x_V es mónico, $\ell(x_V) = \ell(V)$ y $V = \langle x_V \rangle \oplus \tau(V)$. Como la dimensión de $\tau(V)$ es $r - 1$, la hipótesis inductiva nos dice que hay una base ordenada reducida (x_2, \dots, x_r) de $\tau(V)$. Consideremos los vectores

$$y := \pi_{\ell(x_2)}(x) \cdot x_2 + \dots + \pi_{\ell(x_r)}(x) \cdot x_r, \quad x_1 := x_V - y,$$

y mostremos que $\mathcal{B} = (x_1, x_2, \dots, x_r)$ es una base ordenada reducida de V .

- Como $V = \langle x_V \rangle \oplus \tau(V)$, sabemos que (x_V, x_2, \dots, x_r) es una base de V y, por lo tanto, (x_1, x_2, \dots, x_r) también lo es: en efecto, genera a $\tau(V)$ y a x_V , ya que $x_V = x_1 + y \in x_1 + \tau(V)$, y tiene cardinal r .
- Los vectores x_2, \dots, x_r son mónicos por construcción porque la base (x_2, \dots, x_r) de $\tau(V)$ es reducida. Por otro lado, como $y \in \tau(V)$, es $\ell(y) > \ell(V)$, así que $\pi_{\ell(V)}(y) = 0$ y por lo tanto $\pi_{\ell(V)}(x_1) = \pi_{\ell(V)}(x_V) - \pi_{\ell(V)}(y) = 1$, ya que x_V es mónico y $\ell(x_V) = \ell(V)$. Como $\ell(x_1) \geq \ell(V)$ por la definición de $\ell(V)$, esto muestra que $\ell(x_1) = \ell(V)$ y que x_1 es mónico.
- Como (x_2, \dots, x_r) es una base ordenada reducida de $\tau(V)$ y $\ell(x_1) = \ell(V)$, tenemos que

$$\ell(x_1) = \ell(V) < \ell(x_2) < \dots < \ell(x_r).$$

- Sean ahora i y j dos elementos distintos de $\llbracket r \rrbracket$. Si $i = 1$, entonces tenemos que

$$\pi_{\ell(x_1)}(x_j) = \pi_{\ell(V)}(x_j) = 0$$

porque $j > 1$ y, por lo tanto, $x_j \in \tau(V)$. Si $i > 1$ y $j > 1$, entonces $\pi_{\ell(x_i)}(x_j) = 0$ porque

(x_2, \dots, x_r) es una base reducida de $\tau(V)$. Finalmente, si $i > 1$ y $j = 1$, entonces

$$\begin{aligned}\pi_{\ell(x_i)}(x_1) &= \pi_{\ell(x_i)}(x_V) - \pi_{\ell(x_i)}(y) \\ &= \pi_{\ell(x_i)}(x_V) - (\pi_{\ell(x_2)}(x) \cdot \pi_{\ell(x_i)}(x_2) + \dots + \pi_{\ell(x_r)}(x) \cdot \pi_{\ell(x_i)}(x_r)) \\ &= 0,\end{aligned}$$

ya que $\pi_{\ell(x_i)}(x_k) = \delta_{i,k}$ si $i, j \in \llbracket 2, r \rrbracket$. En definitiva, en cualquier caso tenemos que $\pi_{\ell(x_1)}(x_j) = 0$.

Para ver la unicidad, supongamos ahora que $\mathcal{B}' = (x'_1, \dots, x'_r)$ es otra base ordenada reducida de V . Como $\ell(V) \leq \ell(x'_1) < \ell(x'_i)$ para cada $i \in \llbracket 2, r \rrbracket$, los vectores x'_2, \dots, x'_r están en $\tau(V)$ y, como son linealmente independientes y $r - 1$ en número, (x'_2, \dots, x'_r) es, de hecho, una base ordenada de $\tau(V)$. Más aún, como \mathcal{B}' es una base ordenada reducida de V , es inmediato que (x'_2, \dots, x'_r) es una base ordenada reducida de $\tau(V)$: la hipótesis inductiva, entonces, nos dice que $x'_i = x_i$ para cada $i \in \llbracket 2, r \rrbracket$. Por otro lado, como \mathcal{B}' es una base de V , hay escalares $\alpha_1, \dots, \alpha_r \in \mathbb{k}$ tales que $x_1 = \alpha_1 x'_1 + \dots + \alpha_r x'_r$. Si $j \in \llbracket 2, r \rrbracket$, entonces

$$0 = \pi_{\ell(x_j)}(x_1) = \pi_{\ell(x'_j)}(\alpha_1 x'_1 + \dots + \alpha_r x'_r) = \alpha_j,$$

porque la base ordenada \mathcal{B}' es reducida y entonces

$$1 = \pi_{\ell(x_1)}(x_1) = \pi_{\ell(V)}(\alpha_1 x'_1) = \alpha_1 \cdot \pi_{\ell(V)}(x'_1). \quad (4)$$

Esto nos dice que $\pi_{\ell(V)}(x'_1) \neq 0$ y, por lo tanto, que $\ell(x'_1) = 1$: de la igualdad (4) vemos que $\alpha_1 = 1$ y, en definitiva, que $x_1 = x'_1$. Así, las bases ordenadas \mathcal{B} y \mathcal{B}' de V coinciden. \square

6.2.6. Si V es un subespacio de \mathbb{k}^m de dimensión r y (x_1, \dots, x_r) es la base ordenada reducida de V , llamamos a la r -upla estrictamente creciente de enteros positivos $(\ell(x_1), \dots, \ell(x_r))$ el **tipo** de V . Este invariante tiene la siguiente descripción alternativa.

Proposición. Sea (e_1, \dots, e_m) la base ordenada estándar de \mathbb{k}^m y consideremos en \mathbb{k}^m los subespacios $F_0 := V$ y $F_i := \langle e_i, \dots, e_n \rangle$ para cada $i \in \llbracket m \rrbracket$. Si V es un subespacio de \mathbb{k}^m de dimensión r y tipo (t_1, \dots, t_r) , entonces

$$\{t_1, \dots, t_r\} = \{i \in \llbracket n \rrbracket : V \cap F_{i-1} \neq V \cap F_i\}.$$

Los subespacios V_0, \dots, V_m del enunciado se organizan en una cadena estrictamente decreciente

$$\mathbb{k}^m = F_0 \supsetneq F_1 \supsetneq F_2 \supsetneq \dots \supsetneq F_m \supsetneq 0.$$

Si los intersecamos con V obtenemos otra cadena decreciente de subespacios \mathbb{k}^m ,

$$V = V \cap F_0 \supsetneq V \cap F_1 \supsetneq V \cap F_2 \supsetneq \dots \supsetneq V \cap F_m \supsetneq 0$$

pero esta no es en general estrictamente decreciente: la proposición nos dice que el tipo del subespacio V es el conjunto de puntos en esta cadena donde la inclusión es estricta.

	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8
$\ell(x_1) = 3$								
	1							
$\ell(x_2) = 6$		1						
			1					
$\ell(x_3) = 9$				1				
					1			
$\ell(x_4) = 13$					1			
						1		
$\ell(x_5) = 15$							1	
								1

Figura 6.2. Una matriz de $M_{17,8}(\mathbb{k})$ reducida por columnas de rango 5 y tipo $(3, 6, 9, 13, 15)$. Las entradas pintadas de color son todas nulas: las amarillas se anulan porque el rango de la matriz es 5, las azules por la primera condición de la definición 6.2.5, y las rojas por la tercera condición. Los unos garantizan que los vectores x_1, \dots, x_5 sean mónicos. Finalmente, las entradas rayadas son escalares arbitrarios.

6.2.7. Una matriz $A \in M_{m,n}(\mathbb{k})$ cuyas columnas son los vectores $x_1, \dots, x_n \in \mathbb{k}^m$ es **reducida por columnas** si existe $r \in [0, n]$ tal que (x_1, \dots, x_r) es una base ordenada reducida del subespacio de \mathbb{k}^n que genera y $x_i = 0$ para cada $i \in [r+1, n]$.

Proposición. Si $A \in M_{m,n}(\mathbb{k})$, existe exactamente una matriz $\rho(A) \in M_{m,n}(\mathbb{k})$ reducida por columnas y tal que $A \sim_d \rho(A)$.

Esto nos dice que cada clase de equivalencia a derecha en $M_{m,n}(\mathbb{k})$ contiene exactamente una matriz reducida por columnas y que entonces la función $\rho : A \in M_{m,n}(\mathbb{k}) \rightarrow \rho(A) \in M_{m,n}(\mathbb{k})$ es una función de normalización para la relación de equivalencia a derecha.

Demostración. **HACER.**

□

§3. Conjugación de endomorfismos y de matrices cuadradas

6.3.1. Cuando trabajamos con endomorfismos de un espacio vectorial, es de interés una noción diferente de equivalencia: si V es un espacio vectorial y $f, g : V \rightarrow V$ son endomorfismos de V , decimos que f y g son **conjugados** si existe un automorfismo $\alpha : V \rightarrow V$ tal que $f = \alpha^{-1} \circ g \circ \alpha$, y en ese caso escribimos $f \approx g$. Esta es una relación de equivalencia en el conjunto $\text{End}(V)$:

- Si $f \in \text{End}(V)$, entonces $f = \text{id}_V^{-1} \circ f \circ \text{id}_V$, así que $f \approx f$.
- Si $f, g \in \text{End}(V)$ son tales que $f \approx g$, entonces existe un automorfismo $\alpha : V \rightarrow V$ tal que $f = \alpha^{-1} \circ g \circ \alpha$ y se sigue de esto que $g = (\alpha^{-1})^{-1} \circ f \circ \alpha^{-1}$ y, como α^{-1} es un automorfismo de V , que $g \approx f$.
- Si $f, g, h \in \text{End}(V)$ son tales que $f \approx g$ y $g \approx h$, existen automorfismos $\alpha, \beta : V \rightarrow V$ tales que $f = \alpha^{-1} \circ g \circ \alpha$ y $g = \beta^{-1} \circ h \circ \beta$, y entonces

$$f = \alpha^{-1} \circ \beta^{-1} \circ h \circ \beta \circ \alpha = (\beta \circ \alpha)^{-1} \circ h \circ (\beta \circ \alpha).$$

Como la composición $\beta \circ \alpha$ es un automorfismo de V , se sigue de esto que $f \approx h$.

De manera similar, si $n \in \mathbb{N}$, decimos que dos matrices A y B de $M_n(\mathbb{k})$ son **conjugadas**, y escribimos $A \approx B$, si existe una matriz inversible $P \in GL_n(\mathbb{k})$ tal que $A = P^{-1}BP$. Como para los endomorfismos, la relación de conjugación es una relación de equivalencia sobre $\text{End}(V)$.

Más aún la conjugación de matrices y de endomorfismos está estrechamente relacionadas:

Lema. Sea $n \in \mathbb{N}$ y sean $A, B \in M_n(\mathbb{k})$. Las siguientes afirmaciones son equivalentes:

- (a) Las matrices A y B son conjugadas.
- (b) Los endomorfismos $f_A : x \in \mathbb{k}^n \mapsto Ax \in \mathbb{k}^n$ y $f_B : x \in \mathbb{k}^n \mapsto Bx \in \mathbb{k}^n$ de \mathbb{k}^n son conjugados.

Demostración. Si A y B son conjugadas, existe una matriz inversible $P \in GL_n(\mathbb{k})$ tal que $A = P^{-1}BP$ y entonces la función $f_P : x \in \mathbb{k}^n \mapsto Px \in \mathbb{k}^n$ es un automorfismo y $f_A = f_P^{-1} \circ f_B \circ f_P$, de manera que los isomorfismos f_A y f_B son conjugados.

Recíprocamente, si f_A y f_B son conjugados, de manera que existe un automorfismo $\alpha : V \rightarrow V$ tal que $f_A = \alpha^{-1} \circ f_B \circ \alpha$, y \mathcal{B} es la base ordenada estándar de \mathbb{k}^n , entonces la matriz $[\alpha]_{\mathcal{B}}^{\mathcal{B}}$ es inversible, es

$$A = [f_A]_{\mathcal{B}}^{\mathcal{B}} = [\alpha^{-1} \circ f_B \circ \alpha]_{\mathcal{B}}^{\mathcal{B}} = [\alpha^{-1}]_{\mathcal{B}}^{\mathcal{B}} \cdot [f_B]_{\mathcal{B}}^{\mathcal{B}} \cdot [\alpha]_{\mathcal{B}}^{\mathcal{B}} = ([\alpha]_{\mathcal{B}}^{\mathcal{B}})^{-1} \cdot B \cdot [\alpha]_{\mathcal{B}}^{\mathcal{B}}$$

y, por lo tanto, que las matrices A y B son conjugadas. \square

6.3.2. Como en el caso de la equivalencia de funciones lineales, la relación de conjugación entre endomorfismos tiene una interpretación concreta en términos de matrices muy sencilla, análoga a la de la Proposición 6.1.2:

Proposición. Sea V un espacio vectorial de dimensión finita. Dos endomorfismos $f, g : V \rightarrow V$ son conjugados si y solo si existen bases ordenadas \mathcal{B} y \mathcal{B}' de V tales que $[f]_{\mathcal{B}}^{\mathcal{B}'} = [g]_{\mathcal{B}'}$.

Demostración. Sean $f, g : V \rightarrow V$ dos endomorfismos de V y supongamos primero que f y g son conjugados, de manera que existe un automorfismo $\alpha : V \rightarrow V$ tal que $f = \alpha^{-1} \circ g \circ \alpha$. Sea $\mathcal{B} = (x_1, \dots, x_n)$ una base ordenada de V . Como α es un automorfismo, sabemos que $\mathcal{B}' = (\alpha(x_1), \dots, \alpha(x_n))$ es otra base ordenada de V . Sea $[g]_{\mathcal{B}'}^{\mathcal{B}'} = (a_{i,j}) \in M_n(\mathbb{k})$ la matriz de g con respecto a la base \mathcal{B}' , de manera que para cada $j \in [n]$ tenemos que

$$g(\alpha(x_j)) = a_{1,j}\alpha(x_1) + \dots + a_{n,j}\alpha(x_n).$$

Si ahora $j \in [k]$, entonces

$$f(x_j) = (\alpha^{-1} \circ g \circ \alpha)(x_j) = \alpha^{-1}(g(\alpha(x_j))) = \alpha^{-1}(a_{1,j}\alpha(x_1) + \dots + a_{n,j}\alpha(x_n)) = a_{1,j}x_1 + \dots + a_{n,j}x_n.$$

Esto nos dice que $[f]_{\mathcal{B}}^{\mathcal{B}} = (a_{i,j})$ y, en definitiva, que las matrices $[f]_{\mathcal{B}}^{\mathcal{B}}$ y $[g]_{\mathcal{B}'}^{\mathcal{B}'}$ coinciden: la condición del enunciado es, por lo tanto, necesaria.

Para probar su suficiencia, supongamos ahora que hay bases ordenadas $\mathcal{B} = (x_1, \dots, x_n)$ y $\mathcal{B}' = (y_1, \dots, y_n)$ de V tales que $[f]_{\mathcal{B}}^{\mathcal{B}} = [g]_{\mathcal{B}'}^{\mathcal{B}'}$. Hay exactamente una función lineal $\alpha : V \rightarrow V$ tal que $\alpha(x_i) = y_i$ para todo $i \in [n]$, se trata de un automorfismo de V , y es $[\alpha]_{\mathcal{B}'}^{\mathcal{B}} = I_n$ y $[\alpha^{-1}]_{\mathcal{B}}^{\mathcal{B}'} = I_n$. Se sigue de esto, que

$$[\alpha^{-1} \circ g \circ \alpha]_{\mathcal{B}}^{\mathcal{B}} = [\alpha^{-1}]_{\mathcal{B}}^{\mathcal{B}'} \cdot [g]_{\mathcal{B}'}^{\mathcal{B}'} \cdot [\alpha]_{\mathcal{B}'}^{\mathcal{B}} = I_n \cdot [f]_{\mathcal{B}}^{\mathcal{B}} \cdot I_n = [f]_{\mathcal{B}}^{\mathcal{B}}$$

y, por lo tanto, que $f = \alpha^{-1} \circ g \circ \alpha$, ya que los endomorfismos que están a ambos lados de la igualdad tienen la misma matriz con respecto a la base ordenada \mathcal{B} . \square

6.3.3. Es importante notar que la relación de conjugación es estrictamente más restrictiva que la de equivalente. Dos endomorfismos $f, g : V \rightarrow V$ de un espacio vectorial V son equivalentes si son conjugados, pero no vale la implicación recíproca. Por ejemplo, las matrices $A = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$ y $B = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$ son equivalentes, porque tienen claramente el mismo rango, pero no son conjugadas. En efecto, sabemos de la Proposición 5.4.4 que si lo fueran ambas tendrían el mismo polinomio característico, pero $\chi_A = X^2$ mientras que $\chi_B = X(X - 1)$.

6.3.4. La descripción de las clases de equivalencia para la relación de conjugación en $\text{End}(V)$ y en $M_n(\mathbb{k})$ es considerablemente más difícil que para la de equivalencia y ocupará las próximas secciones: nuestro objetivo es obtener un resultado como los de las Proposiciones 6.1.3 y 6.1.5.

§4. Endomorfismos descomponibles e indescomponibles

6.4.1. Decimos que un endomorfismo $f : V \rightarrow V$ de un espacio vectorial V es **descomponible** si existen subespacios no nulos y f -invariantes W_1 y W_2 de V tales que $V = W_1 \oplus W_2$, y que es **indescomponible** si no es descomponible y $V \neq 0$. Observemos que, de acuerdo a estas definiciones, el endomorfismo nulo del espacio vectorial nulo $0 : 0 \rightarrow 0$ no es ni descomponible ni indescomponible.

6.4.2. Ejemplos.

- (a) Un endomorfismo $f : V \rightarrow V$ diagonalizable de un espacio vectorial V con $n := \dim V \geq 2$ es descomponible. En efecto, si $\mathcal{B} = (x_1, \dots, x_n)$ es una base de V de autovectores de f , entonces los subespacios $V_1 := \langle x_1 \rangle$ y $V_2 := \langle x_2, \dots, x_n \rangle$ son f -invariantes, no nulos y tales que $V = V_1 \oplus V_2$.
- (b) El endomorfismo $f : (x, y)^t \in \mathbb{R}^2 \mapsto (-y, x)^t \in \mathbb{R}^2$ del espacio vectorial real \mathbb{R}^2 es indescomponible. En efecto, supongamos que V_1 y V_2 son dos subespacios f -invariantes de V tales que $V = V_1 \oplus V_2$, y supongamos que $V_1 \neq 0$. Hay entonces un vector no nulo $v = (a, b)^t$ en V_1 : el vector $f(v)$ pertenece entonces a V_1 , porque V_1 es f -invariante, y es linealmente independiente de v , así que $\dim V_1 \geq \dim \langle v, f(v) \rangle = 2$ y, por lo tanto, $\dim V_1 = \dim V - \dim V_1 \leq 0$, de manera que V_2 es necesariamente el subespacio nulo de V . \diamond

6.4.3. La descomponibilidad tiene una interpretación sencilla en términos de matrices:

Proposición. *Un endomorfismo $f : V \rightarrow V$ de un espacio vectorial V de dimensión finita es descomponible si y solamente si existe una base \mathcal{B} de V tal que la matriz de f con respecto a \mathcal{B} es diagonal por bloques de manera no trivial.*

Aquí decimos que la matriz es diagonal por bloques de manera no trivial si los bloques que aparecen a lo largo de la diagonal tienen todos tamaño positivo.

Demostración. Sea $f : V \rightarrow V$ un endomorfismo descomponible de un espacio vectorial de dimensión finita y sea $V = V_1 \oplus V_2$ una descomposición en suma directa de V como suma de subespacios no nulos y f -invariantes. Sean $\mathcal{B}_1 = (x_1, \dots, x_r)$ y $\mathcal{B}_2 = (y_1, \dots, y_s)$ bases ordenadas de V_1 y V_2 , respectivamente. Sabemos que $\mathcal{B} = (x_1, \dots, x_r, y_1, \dots, y_s)$ es una base ordenada de V y es inmediato ver que la matriz de f con respecto a \mathcal{B} es

$$[f]_{\mathcal{B}}^{\mathcal{B}} = \begin{pmatrix} [f_{V_1}]_{\mathcal{B}_1}^{\mathcal{B}_1} & 0 \\ 0 & [f_{V_2}]_{\mathcal{B}_2}^{\mathcal{B}_2} \end{pmatrix},$$

una matriz diagonal por bloques de tamaños n y m , ambos positivos.

Sea ahora $f : V \rightarrow V$ un endomorfismo de un espacio vectorial de dimensión finita n tal que existe una base ordenada $\mathcal{B} = (z_1, \dots, z_n)$ de V tal que la matriz $[f]_{\mathcal{B}}^{\mathcal{B}}$ es diagonal por bloques de manera no trivial,

$$[f]_{\mathcal{B}}^{\mathcal{B}} = \begin{pmatrix} A_1 & 0 \\ 0 & A_1 \end{pmatrix}$$

con $A_1 \in M_r(\mathbb{k})$, $A_2 \in M_s(\mathbb{k})$, $r > 0$ y $s > 0$. Por supuesto tenemos entonces que $n = r + s$ y que los subespacios $V_1 = \langle z_1, \dots, z_r \rangle$ y $V_2 = \langle z_{r+1}, \dots, z_n \rangle$ son no nulos, f -invariantes y tales que $V = V_1 \oplus V_2$: esto nos dice que el endomorfismo f es descomponible. \square

6.4.4. El siguiente resultado nos dice que siempre podemos descomponer un endomorfismo de un espacio vectorial de dimensión finita en “componentes indescomponibles”, de manera completamente similar a como un entero puede escribirse como producto de números primos.

Proposición. Sea V un espacio vectorial de dimensión finita y sea $f : V \rightarrow V$ un endomorfismo. Existen $r \in \mathbb{N}_0$ y subespacios f -invariantes V_1, \dots, V_r de V tales que

- para cada $i \in \llbracket r \rrbracket$ la restricción $f_{V_i} : V_i \rightarrow V_i$ es indescomponible, y
- $V = V_1 \oplus \dots \oplus V_r$.

Demostración. Si V es el espacio nulo claramente podemos elegir $r = 0$. Supongamos entonces que $V \neq 0$. En ese caso existen descomposiciones $V = W_1 \oplus \dots \oplus W_r$ de V como suma directa de subespacios no nulos f -invariantes: por ejemplo, podemos tomar $r = 1$ y $W_1 = V$. Más aún, el número r de sumandos en una tal descomposición es a lo sumo igual a $\dim V$, ya que cada uno de ellos tiene dimensión al menos igual a 1. Podemos entonces elegir una descomposición como esa en la que el número de sumandos es máximo.

Ahora bien, si existiese $i \in \llbracket r \rrbracket$ tal que la restricción f_{V_i} fuera descomponible, habría subespacios no nulos y f -invariantes U_1 y U_2 de V_i tales que $V_i = U_1 \oplus U_2$ y, en consecuencia,

$$V = V_1 \oplus \dots \oplus V_{i-1} \oplus U_1 \oplus U_2 \oplus V_{i+1} \oplus \dots \oplus V_r$$

sería una descomposición de V como suma directa de subespacios no nulos y f -invariantes con $r + 1$ sumandos, lo que es imposible. Esta significa que cada una de las restricciones f_{V_1}, \dots, f_{V_r} es indescomponible y prueba la proposición. \square

6.4.5. En general la descomposición de un espacio vectorial V en componentes indescomponibles con respecto a un endomorfismo $f : V \rightarrow V$ que nos da la Proposición 6.4.4 no es única. Por ejemplo, si $f = \text{id}_V$ es la función identidad de V y $\dim V > 1$, entonces cualquiera de las varias descomposiciones de V como suma directa de subespacios de dimensión 1 satisface esas condiciones.

6.4.6. Una de las razones por las que los endomorfismos indescomponibles son importantes es que tienen polinomios mínimos de una forma especial. Este punto será central varias veces en lo que sigue.

Proposición. Sea V un espacio vectorial de dimensión finita y positiva. Si $f : V \rightarrow V$ es un endomorfismo indescomponible de V , entonces existen un polinomio mónico e irreducible $p \in \mathbb{k}[X]$ y un entero positivo v tales que el polinomio minimal de f es $m_f = p^v$.

Demostración. Sabemos que existen un entero positivo r , polinomios monicos, irreducibles y coprimos dos a dos p_1, \dots, p_r y enteros positivos v_1, \dots, v_r tales que $m_f = p_1^{v_1} \cdots p_r^{v_r}$, y la Proposición 5.7.1 nos dice que los subespacios $V_i = \text{Nu}((p_i^{v_i})(f))$ con $i \in \llbracket n \rrbracket$ son no nulos, f -invariantes y tales que $V = V_1 \oplus \dots \oplus V_n$. No puede ser que n sea mayor que 1, porque f es indescomponible, y esto prueba la proposición. \square

§5. Endomorfismos nilpotentes

6.5.1. Si V es un espacio vectorial, un endomorfismo $f : V \rightarrow V$ de V es **nilpotente** si existe $k \in \mathbb{N}_0$ tal que $f^k = 0$ y en ese caso el número $r := \min\{i \in \mathbb{N}_0 : f^i = 0\}$ es el **índice de nilpotencia** de f . Si V es el espacio nulo, entonces f es siempre nilpotente y su índice es 0. Si en cambio V no es nulo, el índice de nilpotencia de un endomorfismo nilpotente de V es un entero positivo.

Una observación sencilla y útil que usaremos siempre sin mencionarla explícitamente es:

Lema. Si $f : V \rightarrow V$ es un endomorfismo nilpotente de índice de nilpotencia r , entonces para todo $i \in \mathbb{N}_0$ se tiene que $f^i = 0$ si y solamente si $i \geq r$.

Demostración. Sea $f : V \rightarrow V$ un endomorfismo nilpotente de V , sea r su índice de nilpotencia y sea $i \in \mathbb{N}_0$. Si $f^i = 0$, entonces la definición del índice de nilpotencia implica inmediatamente que $i \geq r$. Recíprocamente, si $i \geq r$ entonces el entero $i - r$ es no negativo y podemos calcular que $f^i = f^{i-r} \circ f^r = 0$. \square

6.5.2. Si $n \in \mathbb{N}$, una matriz $A \in M_n(\mathbb{k})$ es **nilpotente** si existe $k \in \mathbb{N}_0$ tal que $A^k = 0$ y en ese caso el número $\min\{i \in \mathbb{N}_0 : A^i = 0\}$ es el **índice de nilpotencia** de A . Como es de esperar, la nilpotencia de una matriz $M_n(\mathbb{k})$ es equivalente a la del endomorfismo de \mathbb{k}^n que determina:

Lema. Sea $n \in \mathbb{N}$, sea $A \in M_n(\mathbb{k})$ y sea $f_A : x \in \mathbb{k}^n \mapsto Ax \in \mathbb{k}^n$. La matriz A es nilpotente si y solamente si el endomorfismo f_A de \mathbb{k}^n es nilpotente, y cuando ese es el caso los índices de nilpotencia de A y de f_A coinciden.

Demostración. Las dos afirmaciones del lema son consecuencia inmediata que los conjuntos $\{i \in \mathbb{N}_0 : A^i = 0\}$ y $\{i \in \mathbb{N}_0 : f_A^i = 0\}$ coinciden. \square

6.5.3. No es difícil dar algunas caracterizaciones alternativas útiles de la nilpotencia.

Proposición. Sea V un espacio vectorial de dimensión finita n y sea $f : V \rightarrow V$ un endomorfismo de V . Las siguientes afirmaciones son equivalentes:

- (a) f es nilpotente.
- (b) $f^n = 0$.
- (c) Existe $k \in \mathbb{N}_0$ tal que el polinomio minimal de f es $m_f = X^k$.
- (d) El polinomio característico de f es $\chi_f = X^n$.

Si se cumplen, el entero k que aparece en (c) es el índice de nilpotencia de f y es a lo sumo igual a n .

Demostración. Si V es el espacio nulo todas las afirmaciones del enunciado se cumplen trivialmente, así que podemos suponer que $V \neq 0$.

(a) \Rightarrow (b) Supongamos que el endomorfismo f es nilpotente y sea r su índice de nilpotencia, que es un entero positivo porque $V \neq 0$.

Para cada $i \in [0, r]$ sea $V_i := \text{Nu}(f^i)$. Afirmamos que para cada $i \in [0, r-1]$ se tiene que

$$V_i \subseteq V_{i+1}, \quad V_i = V_{i+1} \implies V_{i+1} = V_{i+2}. \quad (5)$$

En efecto, si $x \in V_i$, entonces $f^i(x) = 0$ y, por lo tanto, $f^{i+1}(x) = f(f^i(x)) = f(0) = 0$, de manera que $x \in V_{i+1}$: esto prueba la que $V_i \subseteq V_{i+1}$. Supongamos, por otro lado, que $V_i = V_{i+1}$ y sea $x \in V_{i+2}$. Es entonces $f^{i+1}(f(x)) = f^{i+2}(x) = 0$ y, en consecuencia, $f(x) \in V_{i+1} = V_i$ por la hipótesis: se sigue de esto que $f^{i+1}(x) = f^i(f(x)) = 0$, esto es, que $x \in V_{i+1}$.

Observemos que $V_0 = \text{Nu}(f^0) = \text{Nu}(\text{id}_V) = 0$ y $V_r = \text{Nu}(f^r) = V$, ya que $f^r = 0$. De la primera de las afirmaciones de (5) deducimos que tenemos una cadena creciente de subespacios de V ,

$$V_0 \subseteq V_1 \subseteq V_2 \subseteq \dots \subseteq V_r = V. \quad (6)$$

De la implicación de (5) se deduce inmediatamente que si existe $i \in [0, r-1]$ tal que $V_i = V_{i+1}$, entonces de hecho $V_j = V_i$ para todo $j \in [i, r]$ y, en particular, que $\text{Nu}(f^i) = V_i = V_r = V$, de manera que $f^i = 0$: esto es absurdo, porque $i < r$ y r es el índice de nilpotencia de f . Vemos así que todas las inclusiones de (6) son estrictas y, por lo tanto, que $\dim V_i - \dim V_{i-1} \geq 1$ cualquiera sea $i \in [r]$ y que

$$n = \dim V_r - \dim V_0 = \sum_{i=1}^r (\dim V_i - \dim V_{i-1}) \geq r.$$

El entero $n - r$ es en consecuencia no negativo y podemos calcular ahora que $f^n = f^{n-r} \circ f^r = 0$.

(b) \Rightarrow (c) Si ponemos $p = X^n \in \mathbb{k}[X]$, entonces la hipótesis es que $p(f) = 0$ y la propiedad característica del polinomio minimal m_f implica que m_f divide a X^n y, claro, que existe $k \in [0, n]$ tal que $m_f = X^k$.

(c) \Rightarrow (d) De la Proposición 5.7.6 sabemos que el polinomio minimal m_f y el polinomio característico χ_f de f tienen los mismos divisores irreducibles. Si $m_f = X^k$ para algún $k \in \mathbb{N}_0$, entonces el único divisor irreducible de m_f es X , así que χ_f tiene que ser de la forma X^l para algún $l \in \mathbb{N}_0$. Como el grado de χ_f es igual a la dimensión de V , debe ser $\chi_f = X^n$.

(d) \Rightarrow (a) Si el polinomio característico de f es $\chi_f = X^n$, entonces el Teorema de Cayley-Hamilton 5.6.20 nos dice que $f^n = \chi_f(f) = 0$: el endomorfismo f es por lo tanto nilpotente.

Esto completa la prueba de la equivalencia de las cuatro afirmaciones del enunciado. Si suponemos que valen, de manera que existe $k \in \mathbb{N}_0$ tal que $m_f = X^k$, entonces $f^k = m_f(f) = 0$ y $f^{k-1} \neq 0$ porque f no anula al polinomio X^{k-1} , ya que éste tiene grado menor que el de m_f . Esto significa que k es el índice de nilpotencia de f . \square

6.5.4. Ejemplos.

- (a) Si $f : V \rightarrow V$ es un endomorfismo nilpotente de un espacio vectorial V y W es un subespacio f -invariante de V , entonces la restricción $f_W : W \rightarrow W$ es nilpotente y su índice de nilpotencia no es mayor que el de f . En efecto, si el índice de nilpotencia de f es k , entonces $(f_W)^k = (f^k)_W = 0$.

- (b) Si $n \in \mathbb{N}$ y $A \in M_n(\mathbb{k})$ es una matriz triangular inferior o superior, entonces A es nilpotente si y solamente si todas las entradas de su diagonal son nulas. En efecto, si $A = (a_{i,j})$, sabemos que el polinomio característico de A es $\chi_A = (X - a_{1,1})\cdots(X - a_{n,n})$ y A es nilpotente, de acuerdo a la Proposición 6.5.3, exactamente cuando este polinomio es igual a X^n .
- (c) Un endomorfismo $f : V \rightarrow V$ diagonalizable es nilpotente si y solamente si es nulo. En efecto, como es diagonalizable existe una base \mathcal{B} de V cuyos elementos son autovectores de f . Para cada $x \in \mathcal{B}$ sea λ_x el autovalor de f correspondiente a x , de manera que $f(x) = \lambda_x x$. Supongamos que f es nilpotente y sea $k \in \mathbb{N}_0$ tal que $f^k = 0$. Si $x \in \mathcal{B}$, entonces $0 = f^k(x) = \lambda_x^k x$, de manera que $\lambda_x = 0$ y, en particular, $f(x) = \lambda_x x = 0$. Vemos así que f se anula en cada uno de los elementos de la base \mathcal{B} y que es, por lo tanto, el endomorfismo nulo de V . Recíprocamente, es claro que si f es nulo, entonces es nilpotente.
- (d) Sea $n \in \mathbb{N}_0$ y consideremos el endomorfismo $D : p \in \mathbb{R}[X]_{\leq n} \mapsto p' \in \mathbb{R}[X]_{\leq n}$ del espacio vectorial $\mathbb{R}[X]_{\leq n}$ de polinomios con coeficientes reales de grado a lo sumo n dado por la derivación. Se trata de un endomorfismo nilpotente. En efecto, sabemos que $D^{n+1}(p) = 0$ si $p \in \mathbb{R}[X]_{\leq n}$. Esto nos dice además que el índice de nilpotencia de D es a lo sumo $n+1$ y de hecho es igual: es $D^n(X^n) = n! \neq 0$, así que D^n no es el endomorfismo nulo.

De manera similar, es fácil ver que el endomorfismo

$$L : p \in \mathbb{R}[X]_{\leq n} \mapsto p(X) - p(X-1) \in \mathbb{R}[X]_{\leq n}$$

de «diferencias finitas» es nilpotente de índice de nilpotencia igual a $n+1$.

- (e) Sean ahora p un número primo y n un entero no negativo, y consideremos la función lineal $D : p \in \mathbb{F}_p[X]_{\leq n} \mapsto p' \in \mathbb{F}_p[X]_{\leq n}$ dada por derivación formal de polinomios con coeficientes en el cuerpo \mathbb{F}_p . Se trata del único endomorfismo de $\mathbb{F}_p[X]_{\leq n}$ tal que $D(X^i) = iX^{i-1}$ para todo $i \in \llbracket 0, n \rrbracket$. Una inducción evidente muestra que para cada $i \in \llbracket 0, n \rrbracket$ vale que

$$D^p(X^i) = i(i-1)\cdots(i-p+1)X^{i-p}$$

y —como uno de los p enteros sucesivos $i, i-1, \dots, i-p+1$ es divisible por p y, por lo tanto, igual a 0 en \mathbb{F}_p — esto implica que $D^p = 0$. Así, D es un endomorfismo nilpotente de $\mathbb{F}_p[X]_{\leq n}$ y su índice de nilpotencia es $\min\{n, p\}$. \diamond

6.5.5. El resultado más importante sobre endomorfismos nilpotentes es que cuando son indecomponibles admiten un vector cíclico:

Proposición. *Sea V un espacio vectorial de dimensión finita n . Si $f : V \rightarrow V$ es un endomorfismo nilpotente e indecomponible, entonces existe $x \in V$ tal que $V = \langle x \rangle_f$, el índice de nilpotencia de f es igual a la dimensión de V y existe una base ordenada \mathcal{B} de V tal que la matriz de f con respecto*

a \mathcal{B} es

$$N_n = \begin{pmatrix} 0 & & & \\ 1 & 0 & & \\ & 1 & 0 & \\ & & \ddots & \ddots \\ & & & 1 & 0 \end{pmatrix}$$

Llamamos a la matriz N_n que aparece aquí un **bloque de Jordan nilpotente de tamaño n** .

Demostración. Sea $f : V \rightarrow V$ un endomorfismo nilpotente e indecomponible y sea r su índice de nilpotencia, de manera que el polinomio minimal de f es $m_f = X^r$. De acuerdo al Corolario 5.6.18, existe un vector $x \in V$ tal que $m_{f,x} = m_f$, ese vector no es nulo ya que el polinomio $m_{f,x}$ tiene grado positivo, y la Proposición 5.6.11 nos dice que $\mathcal{B} = (x, f(x), \dots, f^{r-1}(x))$ es una base ordenada del subespacio f -cíclico $\langle x \rangle_f$ generado por x . En particular, hay una función lineal $\phi : \langle x \rangle_f \rightarrow \mathbb{k}$ tal que

$$\phi(f^i(x)) = \begin{cases} 1 & \text{si } i = r - 1; \\ 0 & \text{si } 0 \leq i < r - 1. \end{cases}$$

Más aún, como $\langle x \rangle_f$ es un subespacio de V , existe una función lineal $\Phi : V \rightarrow \mathbb{k}$ que extiende a ϕ , esto es, tal que $\Phi(v) = \phi(v)$ siempre que $v \in \langle x \rangle_f$.

Sea ahora $\pi : V \rightarrow V$ la función lineal tal que para cada $v \in V$ es

$$\phi(v) = \sum_{i=0}^{r-1} \Phi(f^i(v)) \cdot f^{r-1-i}(x).$$

El núcleo $W := \text{Nu}(\pi)$ es un subespacio f -invariante de V . En efecto, si $v \in W$, entonces

$$\pi(f(v)) = \sum_{i=0}^{r-1} \Phi(f^i(f(v))) \cdot f^{r-1-i}(x) = \sum_{i=0}^{r-1} \Phi(f^{i+1}(v)) \cdot f^{r-1-i}(x)$$

y cambiando el índice de la suma por $j = i + 1$ vemos que esto es

$$= \sum_{j=1}^r \Phi(f^j(v)) \cdot f^{r-j}(x)$$

que, a su vez, como $f^r = 0$, es

$$\begin{aligned} &= \sum_{j=0}^{r-1} \Phi(f^j(v)) \cdot f^{r-j}(x) = f \left(\sum_{j=0}^{r-1} \Phi(f^j(v)) \cdot f^{r-1-j}(x) \right) \\ &= f(\pi(v)) = 0, \end{aligned}$$

de manera que $f(w) \in W$.

Por otro lado, afirmamos que

$$\pi(\nu) = \nu \text{ siempre que } \nu \in \langle x \rangle_f. \quad (7)$$

Para verlo, es suficiente mostrar que vale esa igualdad cuando $\nu = f^j(x)$ para algún $j \in \llbracket 0, r-1 \rrbracket$, ya que esos vectores generan al subespacio $\langle x \rangle_f$, entonces podemos calcular que

$$\pi(f^j(x)) = \sum_{i=0}^{r-1} \Phi(f^i(f^j(x))) \cdot f^{r-1-i}(x) = \sum_{i=0}^{r-1} \Phi(f^{i+j}(x)) \cdot f^{r-1-i}(x).$$

En esta última suma la función Φ aparece evaluada en elementos de $\langle x \rangle_f$, donde coincide con ϕ , así que tenemos que

$$\pi(f^j(x)) = \sum_{i=0}^{r-1} \underbrace{\phi(f^{i+j}(x))}_{\cdot} \cdot f^{r-1-i}(x).$$

Por la forma en que definimos la función ϕ , la expresión marcada aquí se anula si $i + j$ es distinto de $r - 1$, y vale 1 en caso contrario, así que el único término posiblemente no nulo de la suma es el que corresponde a $i = r - 1 - j$, y entonces

$$\pi(f^j(x)) = f^{r-1-(r-1-j)}(x) = f^j(x).$$

Probemos ahora que $V = \langle x \rangle_f \oplus W$. Sea $\nu \in V$. Como $\pi(\nu) \in \langle x \rangle_f$, de (7) sabemos que $\pi(\pi(\nu)) = \pi(\nu)$ y, por lo tanto que

$$\pi(\nu - \pi(\nu)) = \pi(\nu) - \pi(\pi(\nu)) = 0,$$

de manera que $\nu - \pi(\nu) \in W$. Tenemos entonces que

$$\nu = \pi(\nu) + (\nu - \pi(\nu)) \in \langle x \rangle_f + W,$$

lo que muestra que $V = \langle x \rangle_f + W$. Por otro lado, si $\nu \in \langle x \rangle_f \cap W$, entonces por un lado $\pi(\nu) = \nu$ porque $\nu \in \langle x \rangle_f$ y, por otro, $\pi(\nu) = 0$ porque $\nu \in W$: por supuesto, esto nos dice que $\nu = 0$ y, en definitiva, que $\langle x \rangle_f \cap W = 0$.

Ahora bien, como $V = \langle x \rangle_f \oplus W$ y los subespacios $\langle x \rangle_f$ y W son f -invariantes y el primero no es nulo —ya que contiene al vector x , que no es cero— que el endomorfismo f sea indecomponible implica inmediatamente que $W = 0$, esto es, que $V = \langle x \rangle_f$. Tenemos entonces que $\mathcal{B} = (x, f(x), \dots, f^{r-1}(x))$ es una base ordenada de V y es trivial ahora verificar que la matriz $[f]_{\mathcal{B}}$ es la que aparece en el enunciado de la proposición. \square

6.5.6. Si $f : V \rightarrow V$ es un endomorfismo de un espacio vectorial de dimensión finita e $i \in \mathbb{N}_0$, definimos el número

$$\delta_i(f) := \dim \text{Nu}(f^i).$$

Proposición. Sea V un espacio vectorial de dimensión finita.

- (i) Si $f, g : V \rightarrow V$ son endomorfismos conjugados, entonces para todo $i \in \mathbb{N}_0$ es $\delta_i(f) = \delta_i(g)$.
- (ii) Si $f : V \rightarrow V$ es un endomorfismo y V_1, \dots, V_r son subespacios f -invariantes de V tales que $V = V_1 \oplus \dots \oplus V_r$, entonces para todo $i \in \mathbb{N}_0$ se tiene que $\delta_i(f) = \delta_i(f_{V_1}) + \dots + \delta_i(f_{V_r})$.

Demostración. (i) Sean $f, g : V \rightarrow V$ dos endomorfismos conjugados y sea $\alpha : V \rightarrow V$ un automorfismo tal que $f = \alpha^{-1} \circ g \circ \alpha$. De esto se sigue que $f^i = \alpha^{-1} \circ g^i \circ \alpha$ para todo $i \in \mathbb{N}_0$. En efecto, esta igualdad es evidente si $i = 0$, es la hipótesis si $i = 1$, y si vale para $i \in \mathbb{N}_0$ entonces

$$f^{i+1} = f^i \circ f = \alpha^{-1} \circ g^i \circ \alpha \circ (\alpha^{-1} \circ g \circ \alpha) = \alpha^{-1} \circ g^i \circ g \circ \alpha = \alpha^{-1} \circ g^{i+1} \circ \alpha.$$

Es $\delta_0(f) = \delta_0(\text{id}_V) = \delta_0(g)$. Por otro lado, si $x \in \text{Nu}(f)$, tenemos que

$$g(\alpha(x)) = \alpha((\alpha^{-1} \circ g \circ \alpha)(x)) = \alpha(f(x)) = 0$$

y entonces $\alpha(x) \in \text{Nu}(g)$. Como consecuencia de esto, podemos considerar la función

$$\beta : x \in \text{Nu}(f) \mapsto \alpha(x) \in \text{Nu}(g).$$

De manera similar, tenemos una función

$$\gamma : y \in \text{Nu}(g) \mapsto \alpha^{-1}(y) \in \text{Nu}(f)$$

y es inmediato verificar que β y γ son isomorfismos inversos. En particular, vale que

$$\delta_1(f) = \dim \text{Nu}(f) = \dim \text{Nu}(g) = \delta_1(g).$$

Finalmente, si $i \geq 2$, entonces vimos arriba que f^i y g^i son conjugados y por lo que ya probamos para δ_1 nos dice que

$$\delta_i(f) = \delta_1(f^i) = \delta_1(g^i) = \delta_i(g).$$

(ii) Sea $f : V \rightarrow V$ un endomorfismo de V , sean V_1, \dots, V_r subespacios f -invariantes de V tales que

$$V = V_1 \oplus \dots \oplus V_r \tag{8}$$

y sea $i \in \mathbb{N}_0$. Para ver que $\delta_i(f) = \delta_i(f_{V_1}) + \dots + \delta_i(f_{V_r})$, es decir, que

$$\dim \text{Nu}(f^i) = \dim \text{Nu}(f_{V_1}^i) + \dots + \dim \text{Nu}(f_{V_r}^i)$$

es suficiente con mostrar que

$$\text{Nu}(f^i) = \text{Nu}(f_{V_1}^i) \oplus \dots \oplus \text{Nu}(f_{V_r}^i). \tag{9}$$

Supongamos que $x \in \text{Nu}(f^i)$. Como tenemos la descomposición (8), existen $x_1 \in V_1, \dots, x_r \in V_r$ tales que $x = x_1 + \dots + x_r$ y, por lo tanto,

$$0 = f^i(x) = f^i(x_1) + \dots + f^i(x_r) = f_{V_1}^i(x_1) + \dots + f_{V_r}^i(x_r).$$

Como la suma (8) es directa y $f_{V_j}^i(x_j) \in V_j$ para cada $j \in \llbracket r \rrbracket$, se sigue de esto que para cada $j \in \llbracket r \rrbracket$ es $f_{V_j}(x_j) = 0$, esto es, que $x_j \in \text{Nu}(f_{V_j}^i)$. Esto muestra que

$$\text{Nu}(f^i) = \text{Nu}(f_{V_1}^i) + \cdots + \text{Nu}(f_{V_r}^i).$$

Esta igualdad, el hecho evidente de que $\text{Nu}(f_{V_j}^i) \subseteq V_j$ para cada $j \in \llbracket r \rrbracket$, y la descomposición directa (8) implican, de acuerdo a la Proposición 1.9.8, que tenemos la descomposición (9) que queremos. \square

6.5.7. La razón por la que las funciones δ_i nos interesan es que vale el siguiente resultado:

Proposición. *Sea V un espacio vectorial de dimensión finita n . Si $f : V \rightarrow V$ un endomorfismo nilpotente e indescomponible de V , entonces para todo $i \in \mathbb{N}$ se tiene que*

$$-\delta_{i+1}(f) + 2\delta_i(f) - \delta_{i-1}(f) = \begin{cases} 1, & \text{si } i = n; \\ 0, & \text{en caso contrario.} \end{cases}$$

Demostración. Sea $f : V \rightarrow V$ un endomorfismo nilpotente e indescomponible de V . De la Proposición 6.5.5 sabemos que existe una base ordenada $\mathcal{B} = (x_1, \dots, x_n)$ de V tal que

$$[f]_{\mathcal{B}}^{\mathcal{B}} = \begin{pmatrix} 0 & & & & \\ 1 & 0 & & & \\ & 1 & 0 & & \\ & & \ddots & \ddots & \\ & & & 1 & 0 \end{pmatrix}$$

Calculando explícitamente las potencias de esta matriz vemos que para cada $i \geq 0$ es

$$\delta_i(f) = \begin{cases} i, & \text{si } i \leq n; \\ n, & \text{si } i > n; \end{cases}$$

y entonces

$$\begin{aligned} -\delta_{i+1}(f) + 2\delta_i(f) - \delta_{i-1}(f) &= \begin{cases} -(i+1) + 2i + (i-1), & \text{si } i \leq n-1; \\ -n + 2i - (i-1), & \text{si } i = n; \\ -n + 2n - (i-1), & \text{si } i = n+1; \\ -n + 2n - n, & \text{si } i > n+1; \end{cases} \\ &= \begin{cases} 1, & \text{si } i = n; \\ 0, & \text{en caso contrario.} \end{cases} \end{aligned}$$

Esto es lo que afirma el enunciado. \square

6.5.8. Ya tenemos todo lo que necesitamos para probar la existencia de formas normales de Jordan para endomorfismos nilpotentes.

Proposición. Sea V un espacio vectorial de dimensión finita y positiva n , y sea $f : V \rightarrow V$ un endomorfismo nilpotente de V .

- (i) Existe una única secuencia (n_1, n_2, \dots, n_r) de enteros positivos de longitud r finita y positiva que es débilmente decreciente con $n_1 + \dots + n_r = n$ y tal que hay una base ordenada \mathcal{B} de V con respecto a la cual la matriz de f es diagonal por bloques, de la forma

$$[f]_{\mathcal{B}}^{\mathcal{B}} = \begin{pmatrix} N_{n_1} & & & \\ & N_{n_2} & & \\ & & \ddots & \\ & & & N_{n_r} \end{pmatrix},$$

con cada N_{n_i} un bloque de Jordan nilpotente de tamaño n_i .

- (ii) Para cada $s \in \mathbb{N}$ el número de bloques de Jordan nilpotentes de tamaño s que aparecen en esta matriz es

$$b_s(0) := \#\{i \in \llbracket r \rrbracket : n_i = s\} = -\delta_{s+1}(f) + 2\delta_s(f) - \delta_{s-1}(f).$$

La cantidad total de bloques es igual a $\delta_1(f)$.

Llamamos a la secuencia (n_1, \dots, n_r) el **tipo** del endomorfismo f .

Demostración. De acuerdo a la Proposición 6.4.4, existen $r \in \mathbb{N}$ y subespacios f -invariantes V_1, \dots, V_r de V tales que $V = V_1 \oplus \dots \oplus V_r$ y para cada $i \in \llbracket r \rrbracket$ la restricción $f_{V_i} : V_i \rightarrow V_i$ de f a V_i es indescomponible. Sin pérdida de generalidad, podemos suponer que además que la secuencia $(n_1, \dots, n_r) := (\dim V_1, \dots, \dim V_r)$ es débilmente decreciente, ya que si no es ése el caso podemos reindexar los subespacios. Como V no es nulo, el número r es positivo, y como los subespacios V_1, \dots, V_r son indescomponibles, los números n_1, \dots, n_r también son positivos.

Como $f^n = 0$, es claro que $(f_{V_i})^n = 0$ para cada $i \in \llbracket n \rrbracket$, así que los endomorfismos f_{V_1}, \dots, f_{V_r} son nilpotentes. Son además indescomponibles, así que la Proposición 6.5.5 nos dice que para cada $i \in \llbracket r \rrbracket$ hay un vector $x_i \in V_i$ con $V_i = \langle x_i \rangle_f$, que $\mathcal{B}_i = (x_i, f(x_i), \dots, f^{n_i-1}(x_i))$ es una base ordenada de V_i , y que

$$[f_{V_i}]_{\mathcal{B}_i}^{\mathcal{B}_i} = N_{n_i},$$

un bloque de Jordan nilpotente de tamaño n_i . Si $\mathcal{B} = \mathcal{B}_1 \cup \dots \cup \mathcal{B}_r$ es la base ordenada de V que se obtiene concatenando, en orden, las bases $\mathcal{B}_1, \dots, \mathcal{B}_r$, es claro que la matriz $[f]_{\mathcal{B}}^{\mathcal{B}}$ es la matriz descripta en el enunciado. Esto prueba la existencia que se afirma en la parte (a) de la proposición.

De acuerdo a la segunda afirmación de la Proposición 6.5.6, tenemos que

$$\delta_i(f) = \delta_i(f_{V_1}) + \dots + \delta_i(f_{V_r})$$

para todo $i \in \mathbb{N}_0$ y entonces, si $s \in \mathbb{N}$ se tiene que

$$-\delta_{s+1}(f) + 2\delta_s(f) - \delta_{s-1}(f) = \sum_{i=1}^r (-\delta_{s+1}(f_{V_i}) + 2\delta_s(f_{V_i}) - \delta_{s-1}(f_{V_i})).$$

n	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	...
p_n	1	2	3	5	7	11	15	22	30	42	56	77	101	135	176	...

Cuadro 6.1. El número p_n de particiones de un entero positivo n .

Esto, de acuerdo a la Proposición 6.5.7, es igual al numero $\#\{i \in [\![r]\!] : \dim V_i = s\}$: en efecto, el término de la suma correspondiente a $i \in [\![r]\!]$ es igual a 1 si $\dim V_i = s$ y es igual a 0 en caso contrario, así que la suma cuenta cuantos de los subespacios V_i tienen dimensión s . Esto prueba la segunda afirmación de la proposición. Por otro lado, tenemos que

$$\delta_1(f) = \delta_1(f_{V_1}) + \cdots + \delta_1(f_{V_r}) = 1 + \cdots + 1 = r,$$

que es el número de bloques de la matriz del enunciado.

Finalmente, observemos que en la secuencia (n_1, \dots, n_r) la cantidad de veces que aparece cada número $s \in \mathbb{N}$ es precisamente $-\delta_{s+1}(f) + 2\delta_s(f) - \delta_{s-1}(f)$, que depende de f y de nada más: esto prueba la afirmación de unicidad de la parte (i) de la proposición. \square

6.5.9. Si $f : V \rightarrow V$ es un endomorfismo nilpotente de un espacio vectorial de dimensión finita y positiva n , entonces el tipo de f es una secuencia finita y débilmente decreciente (n_1, \dots, n_r) de enteros positivos tal que $n_1 + \cdots + n_r = n$. Decimos que una tal secuencia es una **partición** de n . Por ejemplo, las particiones de 7 son las siguientes quince secuencias:

$$(1, 1, 1, 1, 1, 1, 1), \quad (2, 1, 1, 1, 1, 1), \quad (2, 2, 1, 1, 1), \quad (2, 2, 2, 1), \quad (3, 1, 1, 1, 1), \\ (3, 2, 1, 1), \quad (3, 2, 2), \quad (3, 3, 1), \quad (4, 1, 1, 1), \quad (4, 2, 1), \\ (4, 3), \quad (5, 1, 1), \quad (5, 2), \quad (6, 1), \quad (7).$$

En el Cuadro 6.1 está tabulado el número de particiones de los primeros 15 enteros positivos. Ese número crece muy rápidamente y, por ejemplo, hay

$$24\,061\,467\,864\,032\,622\,473\,692\,149\,727\,991 \sim 24 \cdot 10^{30}$$

particiones de 1000. *Godfrey Harold Hardy* y *Srinivasa Ramanujan* probaron en 1918 que el número de particiones de n es aproximadamente igual a

$$\frac{1}{4n\sqrt{3}} \exp\left(\pi\sqrt{\frac{2n}{3}}\right)$$

cuando n es grande y, de hecho, se conocen aproximaciones mucho mejores y hasta fórmulas exactas para ese número. El estudio de las particiones de enteros es parte de la llamada *teoría aditiva de números* y es una bella, importante y profunda parte de la matemática. Una buena introducción al tema es el libro [And98] de *George E. Andrews*.

6.5.10. La razón por las que nos interesa el tipo de un endomorfismo nilpotente es que nos permite resolver el problema de decidir si dos endomorfismos son conjugados o no:

Proposición. *Sea V un espacio vectorial de dimensión finita y positiva. Dos endomorfismos nilpotentes de V son conjugados si y solamente si tienen el mismo tipo.*

Demostración. Sean f y g dos endomorfismos nilpotentes de V . Si los dos tienen tipo (n_1, \dots, n_r) , entonces hay bases ordenadas \mathcal{B} y \mathcal{B}' de V tales que las matrices $[f]_{\mathcal{B}}^{\mathcal{B}}$ y $[g]_{\mathcal{B}}^{\mathcal{B}'}$ son ambas iguales a

$$\begin{pmatrix} N_{n_1} & & & \\ & N_{n_2} & & \\ & & \ddots & \\ & & & N_{n_r} \end{pmatrix}$$

y esto implica, según la Proposición 6.3.2, que f y g son endomorfismos conjugados.

Recíprocamente, si f y g son conjugados, entonces la Proposición 6.5.6(i) nos dice que $\delta_s(f) = \delta_s(g)$ para cada $s \in \mathbb{N}$ y, de acuerdo a la Proposición 6.5.8(ii), de esto se sigue que f y g tienen el mismo tipo. \square

6.5.11. Una consecuencia inmediata de los resultados que obtuvimos en esta sección es:

Corolario. *Si V es un espacio vectorial de dimensión finita y positiva n , entonces*

- (i) *el número de clases de conjugación de elementos nilpotentes de $\text{End}(V)$ es finito e igual a p_n , el número de particiones de n , y*
- (ii) *todos los endomorfismos nilpotentes e indescomponibles de V son conjugados, de manera que son los elementos de exactamente una clase de conjugación de $\text{End}(V)$.*

Demostración. La primera parte es consecuencia de la Proposición 6.5.10, ya que, por un lado, el tipo de un endomorfismo nilpotente de V es, por definición, una partición de n y, por otro, es claro que toda partición de n es el tipo de algún endomorfismo nilpotente de V .

El tipo de un endomorfismo indescomponible de V tiene una única componente —esto se sigue de la Proposición 6.5.5— así que tiene que ser necesariamente igual a (n) . Esto implica que, de acuerdo a la primera parte, que todos los endomorfismos nilpotentes de V son conjugados, como afirma la segunda parte de la proposición. \square

Un algoritmo para buscar la base que realiza la forma de Jordan

6.5.12. Hay muchos algoritmos para encontrar la forma normal de Jordan de un endomorfismo nilpotente. Queremos describir uno de ellos. Antes, sin embargo, consideraremos un caso especial en el que todo es más sencillo, que encontramos frecuentemente en la práctica, y que sirve de modelo para el caso general.

6.5.13. Si $f : V \rightarrow V$ es un endomorfismo de un espacio vectorial V , llamamos *f -cadena* a toda secuencia ordenada (x_1, \dots, x_n) de vectores no nulos de V tal que $f(x_i) = x_{i+1}$ si $i \in \llbracket n-1 \rrbracket$ y

$f(x_n) = 0$. Muchas veces ilustramos una tal f -cadena con un dibujo de la forma

$$x_1 \xrightarrow{f} x_2 \xrightarrow{f} x_3 \xrightarrow{f} \cdots \cdots \cdots \xrightarrow{f} x_{n-1} \xrightarrow{f} x_n$$

Proposición. Sea V un espacio vectorial de dimensión finita n y sea $f : V \rightarrow V$ un endomorfismo nilpotente de V .

- (i) El endomorfismo f es indecomponible si y solamente si su núcleo $\text{Nu}(f)$ tiene dimensión 1. Si ese es el caso, entonces
- (ii) si (x_1, \dots, x_m) es una f -cadena de longitud m y $m < n$, entonces existe $x \in V$ tal que (x, x_1, \dots, x_m) es una f -cadena de longitud $m + 1$, y
- (iii) toda f -cadena $\mathcal{B} = (x_1, \dots, x_n)$ de longitud n es una base ordenada de V y la matriz $[f]_{\mathcal{B}}^{\mathcal{B}}$ es la forma normal de Jordan de f .

Demostración. De acuerdo a la Proposición 6.5.8, la cantidad de bloques que hay en la forma normal de Jordan de f es $\delta_1(f) = \dim \text{Nu}(f)$, y este es un número positivo. Se sigue de esto que si $\dim \text{Nu}(f) > 1$, entonces f es descomponible. Recíprocamente, si f es descomponible y V_1 y V_2 son dos subespacios no nulos f -invariantes de V tales que $V = V_1 \oplus V_2$, entonces $\text{Nu}(f_{V_1})$ y $\text{Nu}(f_{V_2})$ son subespacios no nulos de V_1 y de V_2 , respectivamente, así que $\text{Nu}(f) = \text{Nu}(f_{V_1}) \oplus \text{Nu}(f_{V_2})$ tiene dimensión al menos 2. Esto prueba la parte (i) de la proposición.

Supongamos ahora que f es indecomponible. De acuerdo a la Proposición 6.5.5, el polinomio minimal de f es X^n . Como $f^n = 0$ y $f^{n-1} \neq 0$, tenemos que $\text{Im}(f) \subseteq \text{Nu}(f^{n-1}) \subsetneq V$ y, por lo tanto, que

$$\dim \text{dim } \text{Im}(f) \leq \dim \text{Nu}(f^{n-1}) < \dim V = n.$$

Ahora bien, es $\dim \text{Im}(f) = \dim V - \dim \text{Nu}(f) = n - 1$, así que esto nos dice que $\text{Im}(f) = \text{Nu}(f^{n-1})$.

Probemos ahora (ii). Si (x_1, \dots, x_m) una f -cadena en V de longitud m tal que $m < n$, entonces claramente $f^m(x_1) = 0$, así que $x_1 \in \text{Nu}(f^m) \subseteq \text{Nu}(f^{n-1}) = \text{Im}(f)$ y, en consecuencia, existe $x \in V$ tal que $x_1 = f(x)$, así que (x, x_1, \dots, x_m) es una f -cadena de longitud $m + 1$.

Supongamos finalmente que $\mathcal{B} = (x_1, \dots, x_n)$ es una f -cadena de longitud n . Para probar (iii) es suficiente que probemos que \mathcal{B} es linealmente independiente, ya que en ese caso es una base ordenada de V porque tiene n elementos y es inmediato que $[f]_{\mathcal{B}}^{\mathcal{B}}$ es un bloque de Jordan nilpotente de tamaño n . Supongamos entonces que $a_1, \dots, a_n \in \mathbb{k}$ escalares tales que $a_1x_1 + \dots + a_nx_n = 0$ y supongamos que no son todos nulos, de manera que podemos considerar el entero $k = \min\{i \in \llbracket n \rrbracket : a_i \neq 0\}$. Tenemos entonces que

$$0 = f^{n-k}(a_1x_1 + \dots + a_nx_n) = a_1f^{n-k}(x_1) + \dots + a_nf^{n-k}(x_n). \quad (10)$$

Si $i \in \llbracket n \rrbracket$ es menor que k , entonces $a_i = 0$, y si es mayor que k , entonces $f^{n-k}(x_i) = f^n(x_{i-k}) = 0$. Esto nos dice que en la suma de (10) hay un sólo sumando posiblemente no nulo, el k -ésimo, y, por lo tanto, que $a_kf^{n-k}(x_k) = a_kx_n = 0$. Como $x_n \neq 0$, esto es absurdo. \square

6.5.14. Ejemplo. Consideremos la matriz

$$A = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 1 & 0 \end{pmatrix} \in M_5(\mathbb{k})$$

y sea $f_A : x \in \mathbb{k}^5 \mapsto Ax \in \mathbb{k}^5$. Como A es estrictamente triangular inferior, es nilpotente. Por otro lado, su rango es 4, ya que el menor $A^{(5,1)}$ es claramente inversible, así que su núcleo tiene dimensión 1 y, por lo tanto, es indescomponible. Queremos una base ordenada \mathcal{B} tal que la matriz $[f]_{\mathcal{B}}$ sea N_5 . De acuerdo a la Proposición 6.5.13, para esto es suficiente encontrar un vector no nulo x_5 del núcleo de f_A y luego sucesivamente vectores x_4, x_3, x_2 y x_1 tales que $Ax_4 = x_5$, $Ax_3 = x_4$, $Ax_2 = x_3$ y $Ax_1 = x_2$ — que esto es posible es lo que afirma la segunda parte de esa proposición — y poner $\mathcal{B} = (x_1, \dots, x_5)$. Es inmediato hacer esto: podemos elegir

$$\begin{aligned} x_1 &= (1, -3, 3, -1, 0)^t, & x_2 &= (0, 1, -2, 1, 0)^t, & x_3 &= (0, 0, 1, -1, 0)^t, \\ x_4 &= (0, 0, 0, 1, 0)^t, & x_5 &= (0, 0, 0, 0, 1)^t. \end{aligned}$$

La matriz

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

que tiene a esos vectores por columnas es inversible y $C^{-1}AC = N_5$.

Hay un patrón más o menos claro en el resultado de este cálculo. En efecto, se puede mostrar fácilmente que si $n \in \mathbb{N}$ y consideramos las matrices $A = (a_{i,j})$ y $C = (c_{i,j})$ de $M_n(\mathbb{k})$ con

$$a_{i,j} = \begin{cases} 1 & \text{si } i > j; \\ 0 & \text{en caso contrario.} \end{cases} \quad c_{i,j} = (-1)^{i+j} \binom{n-j}{i-j}$$

para cada $i, j \in \llbracket n \rrbracket$, entonces $C^{-1}AC = N_n$. \diamond

6.5.15. Extendamos ahora la idea de este ejemplo al caso de un endomorfismo nilpotente no necesariamente indescomponible. Si $f : V \rightarrow V$ es un endomorfismo nilpotente de un espacio vectorial V de dimensión finita, para cada $i \in \mathbb{N}$ consideramos el subespacio

$$N_i(f) := \text{Nu}(f) \cap \text{Im}(f^{i-1}).$$

Lema. Sea $f : V \rightarrow V$ un endomorfismo nilpotente de un espacio vectorial V de dimensión finita y ea r su índice de nilpotencia.

(i) Hay una cadena decreciente

$$N_1(f) \supseteq N_2(f) \supseteq \cdots \supseteq N_r(f) \supseteq N_{r+1}(f) = 0$$

de subespacios de V que termina en el subespacio nulo.

(ii) Para cada $i \in \mathbb{N}$ es

$$\dim N_i(f) = \dim \text{Nu}(f^i) - \dim \text{Nu}(f^{i-1}). \quad (11)$$

(iii) Si $i \in \mathbb{N}$ y x_1, \dots, x_n son vectores de V tales que $\text{Nu}(f^i) = \langle x_1, \dots, x_n \rangle$, entonces

$$N_i(f) = \langle f^{i-1}(x_1), \dots, f^{i-1}(x_n) \rangle.$$

En general, las inclusiones de la primera parte de este lema no son estrictas y el conjunto generador de $N_i(f)$ que nos da la tercera parte no es linealmente independiente.

Demostración. De acuerdo a la definición, $N_{r+1}(f) \subseteq \text{Im}(f^r) = 0$, porque $f^r = 0$. Por otro lado, para cada $i \in \llbracket r \rrbracket$ se tiene claramente que $\text{Nu}(f^{i-1}) \supseteq \text{Nu}(f^i)$, así que

$$N_i(f) = \text{Nu}(f) \cap \text{Nu}(f^{i-1}) \supseteq \text{Nu}(f) \cap \text{Nu}(f^i) = N_{i+1}(f).$$

Esto prueba la primera parte del lema. Para probar la segunda, fijemos $i \in \mathbb{N}$. Si $x \in \text{Im}(f^{i-1})$, entonces $f(x) \in \text{Im}(f^i)$, así que tenemos una función lineal

$$\phi : x \in \text{Im}(f^{i-1}) \mapsto f(x) \in \text{Im}(f^i).$$

Esta función es sobreyectiva: si $y \in \text{Nu}(f^i)$, entonces existe $z \in V$ tal que $y = f^i(z)$ y, por lo tanto, $x := f^{i-1}(z) \in \text{Im}(f^{i-1})$ y $\phi(x) = f(f^{i-1}(z)) = y$. Por otro lado, es evidente que $\text{Nu}(\phi) = \text{Nu}(f) \cap \text{Im}(f^{i-1})$, así que el Teorema 2.4.1 nos dice que

$$\dim N_i(f) = \dim \text{Nu}(f) \cap \text{Im}(f^{i-1}) = \dim \text{Im}(f^{i-1}) - \dim \text{Im}(f^i).$$

Como ese mismo teorema nos dice además que

$$\begin{aligned} \dim \text{Im}(f^{i-1}) - \dim \text{Im}(f^i) &= (\dim V - \dim \text{Im}(f^{i-1})) - (\dim V - \dim \text{Im}(f^i)) \\ &= \dim \text{Nu}(f^i) - \dim \text{Nu}(f^{i-1}), \end{aligned}$$

vemos que la igualdad (11) del enunciado vale.

Veamos finalmente la parte (iii) del lema. Sea $i \in \mathbb{N}$ y sean x_1, \dots, x_n vectores de V tales que $\text{Nu}(f^i) = \langle x_1, \dots, x_n \rangle$. En particular, tenemos que $f(f^{i-1}(x_j)) = f^i(x_j) = 0$ para cada $j \in \llbracket n \rrbracket$ y, por lo tanto, que

$$\langle f^{i-1}(x_1), \dots, f^{i-1}(x_n) \rangle \subseteq \text{Nu}(f) \cap \text{Im}(f^{i-1}) = N_i(f).$$

Recíprocamente, si y es un elemento de $N_i(f)$, entonces y está en $\text{Im}(f^{i-1})$, así que existe $x \in V$ tal que $y = f^{i-1}(x)$, e $y \in \text{Nu}(f)$, de manera que $0 = f(y) = f^i(x)$ y, en consecuencia, $x \in \text{Nu}(f^i)$.

Existen entonces escalares $a_1, \dots, a_n \in \mathbb{k}$ tales que $x = a_1x_1 + \dots + a_nx_n$ y, por lo tanto,

$$y = f^{i-1}(a) = a_1f(x_1) + \dots + a_nf(x_n) \in \langle f^{i-1}(x_1), \dots, f^{i-1}(x_n) \rangle.$$

Concluimos con esto que $N_i(f) = \langle f^{i-1}(x_1), \dots, f^{i-1}(x_n) \rangle$, como queremos. \square

6.5.16. Antes de describir en general el algoritmo que tenemos en mente, demos un ejemplo de cómo funciona. Supongamos que tenemos un espacio vectorial V de dimensión 17 y un endomorfismo nilpotente $f : V \rightarrow V$ tal que las dimensiones de los núcleos de las potencias sucesivas son como en la siguiente tabla:

i	0	1	2	3	4
$\dim \text{Nu}(f^i)$	0	6	11	15	17

El índice de nilpotencia de f es entonces 4 y usando la segunda parte del lema que acabamos de probar podemos calcular las dimensiones de los subespacios $N_i(f)$:

i	1	2	3	4
$\dim N_i(f)$	6	5	4	2

Ahora hacemos sucesivamente los siguientes pasos:

- Sea $\mathcal{B}_4 = \{x_1, \dots, x_{17}\}$ una base cualquiera de $\text{Nu}(f^4)$. De acuerdo a la tercera parte del lema, es $N_4(f) = \langle f^3(x_1), \dots, f^3(x_{17}) \rangle$: como sabemos que $N_4(f)$ tiene dimensión 2, podemos elegir dos vectores u_1, u_2 en el conjunto \mathcal{B}_4 de manera que

$$\{f^3(u_1), f^3(u_2)\} \text{ sea una base de } N_4(f).$$

- Sea ahora $\mathcal{B}_3 = \{y_1, \dots, y_{15}\}$ una base de $\text{Nu}(f^3)$, de manera que, otra vez por el lema, $N_3(f) = \langle f^2(y_1), \dots, f^2(y_{15}) \rangle$. Como $N_4(f) \subseteq N_3(f)$, sabemos que $f^3(u_1)$ y $f^3(u_2)$ están en $N_3(f)$, así que, de hecho, tenemos que

$$N_3(f) = \langle f^3(u_1), f^3(u_2), f^2(y_1), \dots, f^2(y_{15}) \rangle.$$

El subespacio $N_3(f)$ tiene dimensión 4, así que entre los 17 generadores aquí listados podemos quedarnos con 4 que sean linealmente independientes — más aún, como los dos primeros de esos 17 generadores son linealmente independientes, podemos incluir a esos dos entre los 4. Esto nos dice que hay dos vectores v_1, v_2 en el conjunto \mathcal{B}_3 tales que

$$\{f^3(u_1), f^3(u_2), f^2(v_1), f^2(v_2)\} \text{ es una base de } N_3(f).$$

- Sea $\mathcal{B}_2 = \{z_1, \dots, z_{11}\}$ una base de $\text{Nu}(f^2)$. Es $N_2(f) = \langle f(z_1), \dots, f(z_{11}) \rangle$ y

$$\{f^3(u_1), f^3(u_2), f^2(v_1), f^2(v_2)\} \subseteq N_3(f) \subseteq N_2(f),$$

así que también

$$N_2(f) = \langle f^3(u_1), f^3(u_2), f^2(v_1), f^2(v_2), f(z_1), \dots, f(z_{11}) \rangle.$$

Como $\dim N_2(f) = 5$ y los primeros cuatro de los vectores aquí listados son linealmente independientes, podemos elegir un vector w_1 en \mathcal{B}_2 de manera que

$\{f^3(u_1), f^3(u_2), f^2(v_1), f^2(v_2), f(w_1)\}$ es una base de $N_2(f)$.

- Finalmente, sea $\mathcal{B}_1 = \{t_1, \dots, t_6\}$ una base de $\text{Nu}(f)$. Como antes, el lema nos dice que $N_1(f) = \langle t_1, \dots, z_6 \rangle$ y, como $N_2(f) \subseteq N_1(f)$, tenemos que

$$N_1(f) = \langle f^3(u_1), f^3(u_2), f^2(v_1), f^2(v_2), f(w_1), t_1, \dots, t_6 \rangle.$$

La dimensión de $N_1(f)$ es 6 y los primeros 5 de estos 11 vectores son linealmente independientes, así que existe $s \in \mathcal{B}_1$ tal que

$\{f^3(u_1), f^3(u_2), f^2(v_1), f^2(v_2), f(w_1), s\}$ es una base de $N_1(f)$.

Podemos organizar los vectores que fuimos encontrando a lo largo de esta construcción en el siguiente diagrama:

u_1	u_2					
$f(u_1)$	$f(u_2)$	v_1	v_2			
$f^2(u_1)$	$f^2(u_2)$	$f(v_1)$	$f(v_2)$	w_1		
$f^3(u_1)$	$f^3(u_2)$	$f^2(v_1)$	$f^2(v_2)$	$f(w_1)$	s	

(12)

Los 6 vectores de la última fila generan libremente a $N_1(f) = \text{Nu}(f)$, así que f se anula en cada uno de ellos. Por otro lado, la imagen por f de cualquiera de los vectores del diagrama que no están en la última fila de este es el vector que está inmediatamente abajo de él.

Los vectores del diagrama son 17: mostremos que son linealmente independientes. Consideraremos una combinación lineal de ellos que es igual a 0: podemos representarla gráficamente con un diagrama de la misma forma que el anterior, pero en el que en cada casillero ponemos el coeficiente que acompaña al vector correspondiente en la combinación lineal:

$a_{1,0}$	$a_{2,0}$					
$a_{1,1}$	$a_{2,1}$	$a_{3,0}$	$a_{4,0}$			
$a_{1,2}$	$a_{2,2}$	$a_{3,1}$	$a_{4,1}$	$a_{5,0}$		
$a_{1,3}$	$a_{2,3}$	$a_{3,2}$	$a_{4,2}$	$a_{5,1}$	$a_{6,0}$	

= 0. (13)

Si aplicamos la función f^3 a ambos lados de esta igualdad, de acuerdo a como dijimos que

funciona f , obtenemos la igualdad

$$\begin{array}{|c|c|c|c|} \hline & & & \\ \hline & & & \\ \hline & & & \\ \hline \end{array} = 0.$$

$a_{1,0} \quad a_{2,0}$

Como los vectores de la última fila son linealmente independientes, esto implica que $a_{1,0} = a_{2,0} = 0$.

Si ahora aplicamos la función f^2 a ambos lados de la igualdad (13) y usamos lo que ya sabemos, vemos que

$$\begin{array}{|c|c|c|c|} \hline & & & \\ \hline & & & \\ \hline & & & \\ \hline \end{array} = 0$$

$a_{1,1} \quad a_{2,1} \quad a_{3,0} \quad a_{4,0}$

y, por lo tanto, que $a_{1,1} = a_{2,1} = a_{3,0} = a_{4,0} = 0$. Usando todo esto y aplicando f en la igualdad (13) ahora vemos que

$$\begin{array}{|c|c|c|c|} \hline & & & \\ \hline & & & \\ \hline & & & \\ \hline \end{array} = 0,$$

$a_{1,2} \quad a_{2,2} \quad a_{3,1} \quad a_{4,1} \quad a_{5,0}$

así que $a_{1,1} = a_{2,1} = a_{3,0} = a_{4,0} = 0$ y, finalmente, con toda la información que ya tenemos la igualdad (13) queda en la forma

$$\begin{array}{|c|c|c|c|} \hline & & & \\ \hline & & & \\ \hline & & & \\ \hline \end{array} = 0,$$

$a_{1,3} \quad a_{2,3} \quad a_{3,2} \quad a_{4,2} \quad a_{5,1} \quad a_{6,0}$

y $a_{1,3} = a_{2,3} = a_{3,2} = a_{4,2} = a_{5,1} = a_{6,0} = 0$. Hemos probado que todos los coeficientes de la combinación lineal eran nulos, así que los vectores de (12) son linealmente independientes.

Por supuesto, esto implica que

$$\mathcal{B} = (u_1, f(u_1), f^2(u_1), f^3(u_1), u_2, f(u_2), f^2(u_2), f^3(u_2), v_1, f(v_1), f^2(v_1), v_2, f(v_2), f^2(v_2), w_2, f(w_1), s)$$

es una base ordenada de V y es inmediato verificar que, de acuerdo a como describimos la acción de f , la matriz de f con respecto a esta base es

$$[f]_{\mathcal{B}}^{\mathcal{B}} = \begin{pmatrix} 0 & & & & & \\ 1 & 0 & & & & \\ 1 & 0 & & & & \\ 1 & 0 & & & & \\ & & 0 & & & \\ & & 1 & 0 & & \\ & & 1 & 0 & & \\ & & 1 & 0 & & \\ & & & 0 & & \\ & & & 1 & 0 & \\ & & & 1 & 0 & \\ & & & 0 & 1 & 0 \\ & & & & 0 & 1 & 0 \\ & & & & & 0 & 1 & 0 \\ & & & & & & 0 & 1 & 0 \\ & & & & & & & 0 & \\ & & & & & & & & 0 \end{pmatrix}$$

Claramente esta es la forma normal de Jordan de f y el tipo de f es $(4, 4, 3, 3, 2, 1)$.

6.5.17. El algoritmo para hacer esto en general es el de la Figura 6.3 de la página siguiente. Empieza con un endomorfismo $f : V \rightarrow V$ de un espacio vectorial V de dimensión finita y termina con la lista J conteniendo una base ordenada de V con respecto a la cual la matriz de f es la forma normal de Jordan de f .

Para ver que este algoritmo es correcto, ante que nada, tenemos que verificar que que en cada iteración del bucle que empieza en la línea 7 es efectivamente posible elegir el subconjunto S_i , y para esto es suficiente mostrar que para cada $i \in \llbracket r \rrbracket$

si al empezar la iteración i -ésima del bucle que empieza en la linea 7 el conjunto \mathcal{B} es una base de $N_{r+2-i}(f)$ y sus elementos son

$$f^{r-j}(x), \quad j \in \llbracket i-1 \rrbracket, \quad x \in S_{r+1-j}, \tag{14}$$

entonces es posible elegir el conjunto S_{r+1-i} en la línea 9, y al terminar esa iteración, el conjunto \mathcal{B} es una base de $N_{r+1-i}(f)$ con elementos

$$f^{r-j}(x), \quad j \in \llbracket i \rrbracket, \quad x \in S_{r+1-j}. \tag{15}$$

En efecto, lo que queremos sigue de esto por inducción, ya que al empezar la primera iteración \mathcal{B} es vacío y es, por lo tanto, una base de $N_{r+2-1}(f)$, que es el subespacio nulo de V . Para hacerlo, supongamos que $i \in \llbracket r \rrbracket$ y que al empezar la iteración i -ésima del bucle el conjunto \mathcal{B} es una base de $N_{r+2-i}(f)$ cuyos elementos son los listados en (14), en particular, que tiene $n_{r+2-i} - n_{r+1-i}$

```

1    $r \leftarrow$  índice de nilpotencia de  $f$ 
2    $n_0 \leftarrow 0$ 
3   para cada  $i$  de 1 a  $r$ 
4   |  $n_i \leftarrow \dim \text{Nu}(f^i)$ 
5   fin
6    $\mathcal{B} \leftarrow \emptyset$ 
7   para cada  $i$  de 1 a  $r$ 
8   | Elegir una base  $\mathcal{B}_{r+1-i}$  de  $\text{Nu}(f^{r+1-i})$ 
9   | Elegir un subconjunto  $S_{r+1-i}$  de  $\mathcal{B}_{r+1-i}$  tal que  $\mathcal{B} \cup \{f^{r-i}(x) : x \in S_{r+1-i}\}$  sea
    | linealmente independiente y tenga  $n_{r+1-i} - n_{r-i}$  elementos
10  |  $\mathcal{B} \leftarrow \mathcal{B} \cup \{f^{r-i}(x) : x \in S_{r+1-i}\}$ 
11  fin
12   $J \leftarrow []$ , la lista vacía
13  para cada  $i$  de 1 a  $r$ 
14  | para cada  $x$  de  $S_{r+1-i}$ 
15  | |  $J \leftarrow J + [x, f(x), \dots, f^{r-i}(x)]$ 
16  | fin
17  fin

```

Figura 6.3. Un algoritmo para determinar el tipo de un endomorfismo nilpotente $f : V \rightarrow V$ de um espacio vectorial de dimensión finita y una base de V con respecto a la que la matriz de f es la forma normal de Jordan de f . Al terminar el procedimiento, la lista τ contiene el tipo de f y la lista J una base ordenada con la propiedad deseada.

elementos. Como $N_{r+1-i}(f)$ está generado por el conjunto $X := \{f^{r-i}(x) : x \in \mathcal{B}_{r+1-i}\}$, tiene dimensión $n_{r+1-i} - n_{r-i}$ y contiene a \mathcal{B} , es posible elegir en X un subconjunto Y de manera que $Y \cup \mathcal{B}$ sea una base de $N_{r+1-i}(f)$. El cardinal de este conjunto Y es

$$(n_{r+1-i} - n_{r-i}) - (n_{r+2-i} - n_{r+1-i}) = -n_{r-i} + 2n_{r+1-i} - n_{r+2-i},$$

que es el número b_{r+1-i} de bloques de tamaño $r+1-i$ en la forma normal de Jordan de f . Más aún, en vista de la definición de X , hay un subconjunto S_{r+1-i} de \mathcal{B}_{r+1-i} de b elementos tal que $Y = \{f^{r-i}(x) : x \in S_{r+1-i}\}$. Al terminar la iteración, entonces, \mathcal{B} es una base de $N_{r+1-i}(f)$ y claramente sus elementos son los listados en (15).

Al llegar a la línea 12, construimos subconjuntos S_1, \dots, S_r de V tales que los elementos

$$f^{j-1}(x), \quad j \in \llbracket r \rrbracket, \quad x \in S_j, \tag{16}$$

que son $b_1 + \dots + b_r$, son linealmente independientes y generan a $N_1(f) = \text{Nu}(f)$. Queremos

mostrar ahora que los vectores

$$f^l(x), \quad j \in \llbracket r \rrbracket, \quad x \in S_j, \quad l \in \llbracket 0, j-1 \rrbracket, \quad (17)$$

que son precisamente los que contiene la lista J al terminar el algoritmo, son linealmente independientes. Como son

$$1 \cdot b_1 + 2 \cdot b_2 + \cdots + r \cdot b_r = \dim V$$

en número, esto probará que J es una base ordenada de V .

Supongamos entonces que para cada $j \in \llbracket r \rrbracket$, cada $x \in S_j$ y cada $l \in \llbracket 0, j-1 \rrbracket$ tenemos un escalar $a_{x,l} \in \mathbb{k}$, de manera tal que

$$\sum_{j=1}^r \sum_{x \in S_j} \sum_{l=0}^{j-1} a_{x,l} f^l(x) = 0 \quad (18)$$

y, para llegar a una contradicción, supongamos que no son todos esos escalares nulos. Podemos entonces considerar el entero

$$k := \max\{j - l : j \in \llbracket r \rrbracket, l \in \llbracket 0, j-1 \rrbracket, \text{ existe } x \in S_j \text{ con } a_{x,l} \neq 0\}.$$

Observemos que $k \in \llbracket r \rrbracket$ y que

$$\text{existen } j_0 \in \llbracket k, r \rrbracket \text{ y } x_0 \in S_{j_0} \text{ tales que } a_{x_0,j_0-k} \neq 0. \quad (19)$$

Como $k-1 \geq 0$, tiene sentido considerar la potencia f^{k-1} y aplicársela a los dos lados de la igualdad (18). Vemos así que

$$\sum_{j=1}^r \sum_{x \in S_j} \sum_{l=0}^{j-1} a_{x,l} f^{l+k-1}(x) = 0. \quad (20)$$

Ahora bien, si $j \in \llbracket r \rrbracket$ y $x \in S_j$, entonces tenemos que

- si $l \in \llbracket 0, j-1 \rrbracket$ es tal que $l < j-k$, entonces $k < j-l$ y, por la forma en que elegimos al entero k , es $a_{x,l} = 0$, mientras que
- si $l \in \llbracket 0, j-1 \rrbracket$ es tal que $l > j-k$, entonces $l+k-1 > j-1$ y, por lo tanto, como $f^{j-1}(x) \in \text{Nu}(f)$, es $f^{l+k-1}(x) = 0$.

Esto nos dice que la igualdad (20) se puede reescribir en la forma

$$\sum_{j=k}^r \sum_{x \in S_j} a_{x,j-k} f^{j-1}(x) = 0.$$

A la izquierda tenemos una combinación lineal de los vectores de listados en (16), que generan libremente el núcleo $\text{Nu}(f)$ y, en consecuencia, todos los coeficientes que aparecen en ella son nulos: esto es,

$$j \in \llbracket k, r \rrbracket, \quad x \in S_j \implies a_{x,j-k} = 0.$$

Esto contradice a (19) y esta contradicción, como dijimos, prueba que los vectores listados en (17) generan libremente al espacio V .

Si para cada $j \in [r]$ escribimos son $x_{j,1}, \dots, x_{j,b_j}$ a los elementos de S_j , entonces J al terminar el algoritmo contiene los vectores del diagrama de la Figura 6.4 de la página siguiente listados por filas. Como f se anula en cada uno de los vectores que aparecen a final de cada una de las filas del diagrama, es claro que la matriz de f con respecto a la base ordenada J de V es diagonal por bloques, con bloques de Jordan nilpotentes a lo largo de la diagonal y es, por lo tanto, la forma normal de Jordan de f . Esto prueba que nuestro algoritmo efectivamente encuentra una base de V que realiza la forma normal de Jordan de f .

Es importante observar que nuestra demostración de la corrección del algoritmo no prueba por sí sola que existe una base que realiza la forma normal de Jordan: usamos en el argumento que dimos que ya sabemos que la forma normal de Jordan existe y que podemos expresar el número de bloques de Jordan de cada tamaño en esa forma normal en términos de las dimensiones de los núcleos de las potencias de f como en la Proposición 6.5.8(ii).

S_r	$x_{r,1}$	$f(x_{r,1})$	\dots	$f^{r-3}(x_{r,1})$	$f^{r-2}(x_{r,1})$	$f^{r-1}(x_{r,1})$
	$x_{r,2}$	$f(x_{r,2})$	\dots	$f^{r-3}(x_{r,1})$	$f^{r-2}(x_{r,1})$	$f^{r-1}(x_{r,2})$
	\vdots	\vdots		\vdots	\vdots	\vdots
	x_{r,b_r}	$f(x_{r,b_r})$	\dots	$f^{r-3}(x_{r,1})$	$f^{r-2}(x_{r,1})$	$f^{r-1}(x_{r,b_r})$
S_{r-1}	$x_{r-1,1}$	$f(x_{r-1,2})$	\dots	$f^{r-3}(x_{r-1,1})$	$f^{r-2}(x_{r-1,1})$	
	$x_{r-1,2}$	$f(x_{r-1,2})$	\dots	$f^{r-3}(x_{r-1,1})$	$f^{r-2}(x_{r-1,1})$	
	\vdots	\vdots		\vdots	\vdots	
	$x_{r-1,b_{r-1}}$	$f(x_{r-1,b_{r-1}})$	\dots	$f^{r-3}(x_{r-1,1})$	$f^{r-2}(x_{r-1,b_{r-1}})$	
	\vdots	\vdots		\vdots	\vdots	
	\vdots	\vdots		\vdots	\vdots	
	\vdots	\vdots		\vdots	\vdots	
S_2	$x_{2,1}$	$f(x_{2,1})$				
	$x_{2,2}$	$f(x_{2,2})$				
	\vdots	\vdots				
	x_{2,b_2}	$f(x_{2,b_2})$				
S_1	$x_{1,1}$					
	$x_{1,2}$					
	\vdots					
	x_{1,b_1}					

Figura 6.4. La base ordenada J . La imagen por f de cada uno de estos vectores es el que está a su derecha, salvo por aquellos que están al final de cada fila, sobre los que f se anula .

§6. La forma normal de Jordan

6.6.1. Proposición. *Sea V un espacio vectorial de dimensión finita y positiva n . Si $f : V \rightarrow V$ es un endomorfismo indescomponible de V y existe un autovalor λ de f en \mathbb{k} , entonces existe una base ordenada \mathcal{B} de V tal que la matriz de f con respecto a \mathcal{B} es*

$$J_n(\lambda) = \begin{pmatrix} \lambda & & & \\ 1 & \lambda & & \\ & 1 & \lambda & \\ & & \ddots & \ddots \\ & & & 1 & \lambda \end{pmatrix}$$

Si $m \in \mathbb{N}$ y $\rho \in \mathbb{k}$, entonces

$$-\delta_{m-1}(f - \rho \text{id}_V) + 2\delta_m(f - \rho \text{id}_V) - \delta_{m+1}(f - \rho \text{id}_V) = \begin{cases} 1 & \text{si } (m, \rho) = (n, \lambda); \\ 0 & \text{en caso contrario.} \end{cases}$$

La matriz $J_n(\lambda)$ es un **bloque de Jordan de autovalor λ de tamaño n** . Observemos que cuando $\lambda = 0$ la matriz $J_n(\lambda)$ coincide con la matriz N_n de la sección anterior.

Demostración. Sea $f : V \rightarrow V$ un endomorfismo indescomponible de V y sea $\lambda \in \mathbb{k}$ un autovalor de f . Sea $m_f \in \mathbb{k}[X]$ el polinomio minimal de f . De acuerdo a la Proposición 6.4.6, hay un polinomio mónico e irreducible $p \in \mathbb{k}[X]$ y un entero positivo v tal que $m_f = p^v$.

Como λ es una raíz de m_f , claramente tiene que ser también una raíz de p y, entonces, necesariamente debe ser $p = X - \lambda$, ya que p es irreducible. Consideremos el endomorfismo $g = f - \lambda \text{id}_V$ de V . Es $g^v = m_f(f) = 0$, así que el polinomio minimal de g divide a X^v y, en particular, g es nilpotente. Por otro lado, si g fuese descomponible, habría subespacios no nulos y g -invariantes W_1 y W_2 en V tales que $V = W_1 \oplus W_2$: esto es imposible, ya que W_1 y W_2 son también f -invariantes y f es indescomponible por hipótesis.

Vemos así que la Proposición 6.5.5 se aplica a g y que, por lo tanto, existe una base ordenada \mathcal{B} de V tal que $[g]_{\mathcal{B}}^{\mathcal{B}} = N_n$, un bloque de Jordan nilpotente de tamaño n . Se sigue de esto, claro, que

$$[f]_{\mathcal{B}}^{\mathcal{B}} = [g + \lambda \text{id}_V]_{\mathcal{B}}^{\mathcal{B}} = [g]_{\mathcal{B}}^{\mathcal{B}} + \lambda I_n = J_n(\lambda),$$

la matriz descripta en el enunciado.

Sean ahora $m \in \mathbb{N}$ y $\rho \in \mathbb{k}$. Si $\rho \neq \lambda$, entonces $f - \rho \text{id}_V$ es un automorfismo de V , porque ρ no es un autovalor de f , y entonces $\delta_t(f - \rho \text{id}_V) = 0$ para todo $t \in \mathbb{N}$. Si en cambio $\rho = \lambda$, entonces

$$-\delta_{m-1}(f - \rho \text{id}_V) + 2\delta_m(f - \rho \text{id}_V) - \delta_{m+1}(f - \rho \text{id}_V) = -\delta_{m-1}(g) + 2\delta_m(g) - \delta_{m+1}(g).$$

Como g es un endomorfismo nilpotente e indescomponible, la Proposición 6.5.7 nos dice que el lado derecho de esta igualdad vale 1 si $m = n$ y 0 en cualquier otro caso. Esto prueba la segunda afirmación de la proposición. \square

6.6.2. Estamos en posición de probar el resultado central de este capítulo, el llamado *Teorema de la forma normal de Jordan*, por Camille Jordan (1838–1922, Francia). Jordan enunció y probó un resultado similar pero distinto, en el contexto de las funciones lineales sobre cuerpos finitos, en su libro *Traité des substitutions et des équations algébriques* [Jor70] de 1870. Es de notar que Jordan no habla nunca de matrices, ni de espacios vectoriales ni de funciones lineales en ese texto —esos conceptos no habían sido inventados aún cuando Jordan escribía— sino de *substitutiones lineales*, esto es, de cambios lineales de coordenadas.

Teorema. Sea V un espacio vectorial de dimensión finita y positiva n , sea $f : V \rightarrow V$ un endomorfismo de V y supongamos que el polinomio minimal de f se factoriza en $\mathbb{k}[X]$ como producto de factores lineales. Existen un entero positivo r y r secuencias ordenadas

$$(\lambda_1, n_{1,1}, \dots, n_{1,m_1}), \quad (\lambda_2, n_{2,2}, \dots, n_{2,m_2}), \quad \dots, \quad (\lambda_r, n_{r,r}, \dots, n_{r,m_r}) \quad (21)$$

tales que

- $\lambda_1, \dots, \lambda_r \in \mathbb{k}$ son escalares distintos dos a dos,
- para cada $i \in [\![r]\!]$ es $m_i \in \mathbb{N}$ y $n_{i,1}, \dots, n_{i,m_i}$ es una secuencia débilmente decreciente de enteros positivos,
- $\sum_{i=1}^r \sum_{j=1}^{r_i} n_{i,j} = n$,
- existe una base ordenada \mathcal{B} de V tal que la matriz de f con respecto a \mathcal{B} es la matriz diagonal por bloques

$$\left(\begin{array}{cccccc} J_{n_{1,1}}(\lambda_1) & & & & & \\ & \ddots & & & & \\ & & J_{n_{1,m_1}}(\lambda_1) & & & \\ & & & \ddots & & \\ & & & & J_{n_{1,1}}(\lambda_1) & \\ & & & & & \ddots \\ & & & & & & J_{n_{1,m_1}}(\lambda_1) \\ & & & & & & & \ddots \\ & & & & & & & & J_{n_{r,1}}(\lambda_r) \\ & & & & & & & & & \ddots \\ & & & & & & & & & & J_{n_{r,m_r}}(\lambda_r) \end{array} \right). \quad (22)$$

Más aún:

- (i) Los escalares $\lambda_1, \dots, \lambda_r$ son los autovalores de f listados sin repeticiones.
- (ii) Para cada $i \in [\![r]\!]$ el autovalor λ_i de f tiene
 - multiplicidad geométrica m_i , que coincide con la cantidad de bloques de Jordan de autovalor λ_i que aparecen en la matriz (22), y
 - multiplicidad algebraica $n_{1,1} + \dots + n_{1,m_i}$.

Para cada $s \in \mathbb{N}$ la cantidad de bloques de la forma $J_s(\lambda_i)$ que aparecen en la matriz (22) es

$$b_s(\lambda_i) := \#\{j \in [\![m_i]\!]: n_j = s\} = -\delta_{s+1}(f - \lambda_i \text{id}_V) + 2\delta_s(f - \lambda_i \text{id}_V) - \delta_{i-1}(f - \lambda_i \text{id}_V).$$

El conjunto de las r secuencias de está unívocamente determinado por f .

Llamamos a la matriz (22) la **forma normal de Jordan** del endomorfismo f y al conjunto que tiene como elementos las r secuencias de (21) el **tipo** de f . El teorema nos dice que la forma normal de Jordan de f y el tipo de f están unívocamente determinados por f , a menos de una permutación de los bloques de la matriz.

Demostración. De acuerdo a la Proposición 6.4.4, hay un entero $s \in \mathbb{N}$ y subespacios f -invariantes V_1, \dots, V_s de V tales que $V = V_1 \oplus \dots \oplus V_s$ y para cada $i \in [\![s]\!]$ la restricción $f_{V_i} : V_i \rightarrow V_i$ de f a V_i es indecomponible. Más aún, si $i \in [\![s]\!]$, como el subespacio V_i es f -invariante, el Corolario 5.6.14 nos dice que el polinomio minimal $m_{f_{V_i}}$ de la restricción f_{V_i} divide al polinomio minimal m_f de f : como m_f se factoriza como producto de factores lineales, lo mismo es cierto de $m_{f_{V_i}}$, y, como V_i no es un espacio nulo, esto implica que el polinomio $m_{f_{V_i}}$ tiene una raíz μ_i en \mathbb{k} y, en consecuencia, que μ_i es un autovalor de f_{V_i} . La Proposición 6.6.1 nos dice, entonces, que hay bases ordenadas $\mathcal{B}_1, \dots, \mathcal{B}_s$ de los subespacios V_1, \dots, V_s tales que

$$[f_{V_i}]_{\mathcal{B}_i}^{\mathcal{B}_i} = J_{d_i}(\mu_i)$$

para cada $i \in [\![s]\!]$, con $d_i := \dim V_i$.

Sean $\lambda_1, \dots, \lambda_r$ los elementos del conjunto $\{\mu_1, \dots, \mu_s\}$ listados sin repeticiones, y consideremos los s pares

$$\binom{d_1}{\mu_1}, \dots, \binom{d_s}{\mu_s}. \tag{23}$$

Para cada $i \in [\![r]\!]$ sea m_i el número de estos pares que tienen a λ_i en su segunda componente, y sean $n_{i,1}, \dots, n_{i,m_i}$ las primeras componentes de esos m_i pares ordenadas en orden decreciente. Claramente los pares

$$\binom{n_{1,1}}{\lambda_1}, \dots, \binom{n_{1,m_1}}{\lambda_1}, \quad \binom{n_{2,1}}{\lambda_2}, \dots, \binom{n_{2,m_1}}{\lambda_2}, \quad \dots, \dots, \quad \binom{n_{r,1}}{\lambda_r}, \dots, \binom{n_{r,m_r}}{\lambda_r} \tag{24}$$

son exactamente los mismos que los listados en (23). En otras palabras, hay una biyección $\pi : [\![s]\!] \rightarrow [\![s]\!]$ tal que para cada $i \in [\![s]\!]$ el par i -ésimo de (24) coincide con el par $\pi(i)$ -ésimo de (23). Como conclusión de esto vemos que, a menos de reindexar los pares de (23), podemos suponer las listas (23) y (24) son las mismas.

Sea ahora \mathcal{B} la base ordenada de V que se obtiene concatenando las bases $\mathcal{B}_1, \dots, \mathcal{B}_s$ en ese orden. Como $V = V_1 \oplus \dots \oplus V_s$ y cada uno de los subespacios de esta descomposición es

f -invariante y tiene como base el subconjunto \mathcal{B}_i de \mathcal{B} , tenemos que

$$[f]_{\mathcal{B}}^{\mathcal{B}} = \begin{pmatrix} [f|_{V_1}]_{\mathcal{B}_1}^{\mathcal{B}_1} & & \\ & \ddots & \\ & & [f|_{V_s}]_{\mathcal{B}_s}^{\mathcal{B}_s} \end{pmatrix} = \begin{pmatrix} J_{d_1}(\mu_1) & & \\ & \ddots & \\ & & J_{d_s}(\mu_s) \end{pmatrix}.$$

Gracias a la forma en que ordenamos los pares en la lista (24), esta matriz tiene la forma de la matriz (22) del enunciado del teorema. Es inmediato ahora que las r secuencias

$$(\lambda_1, n_{1,1}, \dots, n_{1,m_1}), \quad (\lambda_2, n_{2,2}, \dots, n_{2,m_2}), \quad \dots, \quad (\lambda_r, n_{r,r}, \dots, n_{r,m_r})$$

tienen las cuatro propiedades descriptas en la primera parte del enunciado del teorema.

La matriz $[f]_{\mathcal{B}}^{\mathcal{B}}$ es triangular inferior, así que los autovalores de f son los escalares que aparecen a lo largo de su diagonal y la multiplicidad algebraica de cada uno de ellos es el número de veces que aparece en esa diagonal. Así, los autovalores de f son los escalares $\lambda_1, \dots, \lambda_r$, y para cada $i \in \llbracket r \rrbracket$ la multiplicidad algebraica de λ_i como autovalor de f es $n_{i,1} + \dots + n_{i,m_i}$.

Sea $\rho \in \mathbb{k}$ y sea $t \in \mathbb{N}$. Para cada $i \in \llbracket s \rrbracket$ el subespacio V_i es $(f - \rho \text{id}_V)$ -invariante y la restricción $(f - \rho \text{id}_V)|_{V_i}$ claramente coincide con $f_{V_i} - \rho \text{id}_{V_i}$, así que la Proposición 6.5.6(ii) nos dice que

$$\delta_t(f - \rho \text{id}_V) = \delta_t(f_{V_1} - \rho \text{id}_{V_1}) + \dots + \delta_t(f_{V_r} - \rho \text{id}_{V_r})$$

y, por lo tanto, que

$$\begin{aligned} & -\delta_{t-1}(f - \rho \text{id}_V) + 2\delta_t(f - \rho \text{id}_V) - \delta_{t+1}(f - \rho \text{id}_V) \\ &= \sum_{i=1}^s \underbrace{[-\delta_{t-1}(f_{V_i} - \rho \text{id}_{V_i}) + 2\delta_t(f_{V_i} - \rho \text{id}_{V_i}) - \delta_{t+1}(f_{V_i} - \rho \text{id}_{V_i})]}_{}. \end{aligned}$$

De acuerdo a la Proposición 6.6.1, la expresión marcada es igual a 1 si $(\rho, t) = (\mu_i, d_i)$, y a 0 en caso contrario, y entonces la suma es igual a la cantidad de índices $i \in \llbracket s \rrbracket$ tal que $[f|_{V_i}]_{\mathcal{B}_i}^{\mathcal{B}_i}$ es la matriz $J_t(\rho)$, es decir, el número de bloques de Jordan de tamaño t y autovalor ρ en la matriz (22).

Finalmente, si $\rho \in \mathbb{k}$, la multiplicidad geométrica de ρ como autovalor de f es $\delta_1(f - \rho \text{id}_V)$, y de acuerdo a las observaciones que hicimos arriba esto es

$$\delta_1(f - \rho \text{id}_V) = \delta_1(f_{V_1} - \rho \text{id}_{V_1}) + \dots + \delta_1(f_{V_s} - \rho \text{id}_{V_s})$$

y para cada $i \in \llbracket s \rrbracket$ el número $\delta_1(f_{V_i} - \rho \text{id}_{V_i})$ es 1 si $\rho = \mu_i$ y 0 en caso contrario: esto nos dice que $\delta_1(f - \rho \text{id}_V)$ es el número de bloques de Jordan en la matriz (22) de autovalor ρ . Así, para cada $i \in \llbracket r \rrbracket$ la multiplicidad geométrica de λ_i como autovalor de f es precisamente el entero m_i . \square

6.6.3. La prueba que da Jordan de este teorema en [Jor70] es completamente distinta de la que presentamos nosotros. En la tesis [Breo6] de Frédéric Brechenmacher puede encontrarse una discusión detallada de la demostración original. La importancia de este teorema se manifiesta, entre otras formas, en la cantidad de pruebas radicalmente distintas que se han dado de él a lo largo del tiempo: por citar sólo algunas, podemos mencionar las de [Bru87], [Cat62], [Fil71], [FS83], [GW80], [GG96], [Roi99] y [Väl86].

6.6.4. La forma normal de Jordan nos permite resolver el problema que nos habíamos planteado en la Sección 6.3 de decidir si dos endomorfismos son conjugados, bajo una hipótesis razonable sobre el cuerpo con el que estamos trabajando:

Proposición. *Supongamos que \mathbb{k} es un cuerpo algebraicamente cerrado. Dos endomorfismos de un espacio vectorial de dimensión finita y positiva son conjugados si y solamente si tienen el mismo tipo.*

Demostración. Sean $f, g : V \rightarrow V$ dos endomorfismos de un espacio vectorial de dimensión finita. Como el cuerpo \mathbb{k} es algebraicamente cerrado, los polinomios minimales de f y de g se factorizan como producto de factores lineales y entonces el Teorema 6.6.2 nos dice que hay bases ordenadas \mathcal{B} y \mathcal{B}' de V tales que las matrices $[f]_{\mathcal{B}}$ y $[g]_{\mathcal{B}'}$ son las formas normales de Jordan de f y de g , respectivamente. Si f y g tienen el mismo tipo, entonces las formas normales de Jordan de f y g coinciden, de manera que $[f]_{\mathcal{B}} = [g]_{\mathcal{B}'}$: la Proposición 6.3.2 nos dice entonces que los endomorfismos f y g son conjugados.

Recíprocamente, si los endomorfismos f y g son conjugados, entonces para todo $\lambda \in \mathbb{k}$ los endomorfismos $f - \lambda \text{id}_V$ y $g - \lambda \text{id}_V$ son conjugados y, de acuerdo a la Proposición 6.5.6(i), tenemos que $\delta_s(f - \lambda \text{id}_V) = \delta_s(g - \lambda \text{id}_V)$ para todo $s \in \mathbb{N}_0$. Como consecuencia de esto y de la última parte del Teorema 6.6.2, vemos que las formas normales de Jordan de f y de g tienen la misma cantidad de bloques de Jordan de cada autovalor y cada tamaño y, por lo tanto, que los tipos de f y de g coinciden. La proposición queda así probada. \square

6.6.5. Si V es un espacio vectorial y X un conjunto, decimos que una función $\Phi : \text{End}(V) \rightarrow X$ es un **invariante** si cada vez que f y g son dos elementos de $\text{End}(V)$ que son conjugados se tiene que $\Phi(f) = \Phi(g)$. Ya conocemos varios ejemplos de invariantes: el determinante, la traza, el polinomio característico, el polinomio minimal.

Una consecuencia inmediata de la Proposición 6.6.4 es que si $\Phi : \text{End}(V) \rightarrow X$ es un invariante, entonces para todo endomorfismo $f : V \rightarrow V$ cuyo polinomio minimal se factoriza como producto de factores lineales el valor $\Phi(f)$ depende solamente del tipo de f .

Proposición. *Sea V un espacio vectorial de dimensión finita y positiva n , sea $f : V \rightarrow V$ un endomorfismo de V y supongamos que el polinomio minimal de f se factoriza en $\mathbb{k}[X]$ como producto de factores lineales. Si*

$$\{(\lambda_1, n_{1,1}, \dots, n_{1,m_1}), (\lambda_2, n_{2,2}, \dots, n_{2,m_2}), \dots, (\lambda_r, n_{r,r}, \dots, n_{r,m_r})\}$$

es el tipo de f y para cada $i \in [r]$ ponemos $n_i = n_{i,1} + \dots + n_{i,m_i}$ y $v_i = \max\{n_{i,1}, \dots, n_{i,m_i}\}$, entonces

- *el polinomio característico de f es $\chi_f = (X - \lambda_1)^{n_1} \cdots (X - \lambda_r)^{n_r}$,*
- *el polinomio minimal de f es $m_f = (X - \lambda_1)^{v_1} \cdots (X - \lambda_r)^{v_r}$,*
- *el determinante y la traza de f son, respectivamente,*

$$\det(f) = \lambda_1^{n_1} \cdots \lambda_r^{n_r}, \quad \text{tr}(f) = n_1 \lambda_1 + \cdots + n_r \lambda_r,$$

Demostración. El polinomio característico, el polinomio minimal, el determinante y la traza de f pueden calcularse a partir de la matriz $[f]_{\mathcal{B}}^{\mathcal{B}}$ con respecto a cualquier base ordenada \mathcal{B} de V , y si elegimos a \mathcal{B} de manera que esa matriz sea la forma normal de Jordan de f , entonces las afirmaciones de la proposición son evidentes. \square

§7. Algunas ejemplos y aplicaciones

La forma normal de Jordan de la matriz compañera de un polinomio

6.7.1. Sea $p \in \mathbb{k}[X]$ un polinomio mónico que se factoriza como producto de factores lineales, de manera que hay escalares $\lambda_1, \dots, \lambda_r \in \mathbb{k}$ distintos dos a dos y enteros positivos $n_1, \dots, n_r \in \mathbb{N}$ tales que $p = (X - \lambda_1)^{m_1} \cdots (X - \lambda_r)^{m_r}$, y sea $A = C(p)^t$ la matriz transpuesta de la matriz compañera de p . El polinomio característico de A es p y en el Ejemplo 5.6.24 vimos que el polinomio minimal de A también es p . De acuerdo a la Proposición 6.6.5, esto nos dice que para cada $i \in \llbracket r \rrbracket$ hay al menos un bloque de Jordan $J_{m_i}(\lambda_i)$ de tamaño m_i y autovalor λ_i en la forma normal de Jordan de A . Como $m_1 + \cdots + m_r = \text{gr}(p) = n$, no puede haber más que esos bloques. Concluimos de esta manera que la forma norma de Jordan de la matriz A es

$$J = \begin{pmatrix} J_{m_1}(\lambda_1) & & \\ & \ddots & \\ & & J_{m_r}(\lambda_r) \end{pmatrix}.$$

Queremos encontrar explícitamente una matriz inversible C tal que $C^{-1}AC$ sea J .

6.7.2. Supongamos que $p = a_n X^n + a_{n-1} X^{n-1} + \cdots + a_1 X + a_0$, de manera que, en particular, $a_n = 1$, y sea λ una de las raíces de p . Como λ es un autovalor de A , hay un vector $x = (x_1, \dots, x_n)^t$ que es un autovector de A de autovector λ y entonces

$$\begin{pmatrix} \lambda x_1 \\ \lambda x_2 \\ \vdots \\ \lambda x_{n-2} \\ \lambda x_{n-1} \\ \lambda x_n \end{pmatrix} = \begin{pmatrix} 0 & 1 & \cdots & 0 & 0 & 0 \\ 0 & 0 & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & & \vdots & \vdots & \\ 0 & 0 & \cdots & 0 & 1 & 0 \\ 0 & 0 & \cdots & 0 & 0 & 1 \\ -a_0 & -a_1 & \cdots & -a_{n-3} & -a_{n-2} & -a_{n-1} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_{n-2} \\ x_{n-1} \\ x_n \end{pmatrix} = \begin{pmatrix} x_2 \\ x_3 \\ \vdots \\ x_{n-1} \\ x_n \\ -\sum_{i=0}^{n-1} a_i x_{i+1} \end{pmatrix},$$

de manera que, en particular, $x_i = \lambda x_{i-1}$ para cada $i \in \llbracket 2, n \rrbracket$. Tiene que ser $x_1 \neq 0$, porque si no el vector x sería nulo y no lo es, así que a menos de multiplicar por un escalar podemos suponer que $x_1 = 1$ y, por lo tanto, concluir que el vector x es

$$x_{\lambda}^{(0)} := (1, \lambda, \lambda^2, \dots, \lambda^{n-1})^t.$$

Nos preguntamos ahora si hay un vector $y = (y_1, \dots, y_n)^t \in \mathbb{k}$ tal que $x_\lambda^{(0)} = (A - \lambda I_n)y$. Si y es una solución a esta ecuación, entonces $y - y_1 x_\lambda^{(0)}$ es otra que además tiene su primera componente nula: podemos suponer entonces, sin pérdida de generalidad, que $y_1 = 0$. La ecuación que queremos resolver es, en forma matricial, la siguiente:

$$\begin{pmatrix} 1 \\ \lambda \\ \vdots \\ \lambda^{n-3} \\ \lambda^{n-2} \\ \lambda^{n-1} \end{pmatrix} = \begin{pmatrix} -\lambda & 1 & \cdots & 0 & 0 & 0 \\ 0 & -\lambda & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & & \vdots & \vdots & \\ 0 & 0 & \cdots & -\lambda & 1 & 0 \\ 0 & 0 & \cdots & 0 & -\lambda & 1 \\ -a_0 & -a_1 & \cdots & -a_{n-3} & -a_{n-2} & -a_{n-1} - \lambda \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_{n-2} \\ y_{n-1} \\ y_n \end{pmatrix} = \begin{pmatrix} y_2 - \lambda y_1 \\ y_3 - \lambda y_2 \\ \vdots \\ y_{n-1} - \lambda y_{n-2} \\ y_n - \lambda y_{n-1} \\ -\sum_{i=0}^{n-1} a_i y_{i+1} - \lambda y_n \end{pmatrix}.$$

Mirando las primeras $n-1$ coordenadas de estos vectores, vemos que tiene que ser $y_i = \lambda^{i-2} + \lambda y_{i-1}$ para cada $i \in \llbracket 2, n \rrbracket$ y, como $y_1 = 0$, que, de hecho, $y_i = (i-1)\lambda^{i-2}$ para cada $i \in \llbracket n \rrbracket$. Por otro lado, mirando la última coordenada vemos que tiene que ser

$$\lambda^{n-1} = -\sum_{i=0}^{n-1} a_i y_{i+1} - \lambda y_n = -\sum_{i=1}^{n-1} a_i i \lambda^{i-1} - (n-1)\lambda^{n-1}$$

o, equivalentemente,

$$p'(\lambda) = n\lambda^{n-1} + \sum_{i=1}^{n-1} a_i i \lambda^{i-1} = 0.$$

La conclusión de esto es que hay un vector y en \mathbb{k}^n tal que $x_\lambda^{(0)} = (A - \lambda I_n)y$ si y solamente si $p'(\lambda) = 0$ y, como sabemos, esto ocurre si y solamente si la multiplicidad de λ como raíz de m es al menos 2. Cuando ese es el caso, un vector y que satisface la ecuación es

$$x_\lambda^{(1)} = (0, 1, 2\lambda, 3\lambda^2, \dots, (n-1)\lambda^{n-2})^t.$$

Esto es el principio de un patrón general:

Lema. Si λ es una raíz de p de multiplicidad m , entonces los vectores

$$x_\lambda^{(k)} := \left(\underbrace{0, \dots, 0}_{k \text{ copias}}, 1, \dots, \binom{i-1}{k} \lambda^{i-1-k}, \dots, \binom{n-1}{k} \lambda^{n-1-k} \right), \quad k \in \llbracket 0, m-1 \rrbracket$$

forman una $(A - \lambda I_n)$ -cadena de longitud m ,

$$x_\lambda^{(m-1)} \xrightarrow{A - \lambda I_n} x_\lambda^{(m-2)} \xrightarrow{A - \lambda I_n} x_\lambda^{(m-3)} \quad \dots \quad \dots \xrightarrow{A - \lambda I_n} x_\lambda^{(1)} \xrightarrow{A - \lambda I_n} x_\lambda^{(0)}$$

Demostración. Ya sabemos que $(A - \lambda I_n)x_\lambda^{(0)} = 0$, así que es suficiente que fijemos $k \in \llbracket m-1 \rrbracket$ y veamos que

$$(A - \lambda I_n)x_\lambda^{(k)} = x_\lambda^{(k-1)}.$$

Sea $i \in \llbracket n-1 \rrbracket$. La i -ésima componente del lado izquierdo de esta igualdad es

$$-\lambda(x_\lambda^{(k)})_i + (x_\lambda^{(k)})_{i+1} \quad (25)$$

Si $i < k$, entonces esto es 0, porque las componentes i -ésima e $(i+1)$ -ésimas de $x_\lambda^{(k)}$ son nulas. Si $i = k$, entonces el valor de (25) es 1, porque la k -ésima coordenada de $x_\lambda^{(k)}$ es nula mientras que la $(k+1)$ -ésima es 1. Finalmente, si $k < i < n$, entonces (25) es

$$-\lambda \binom{i-1}{k} \lambda^{i-1-k} + \binom{i}{k} \lambda^{i-k} = \binom{i-1}{k-1} \lambda^{i-1-(k-1)}.$$

Vemos así que para todo $i \in \llbracket n-1 \rrbracket$ las coordenadas i -ésimas de $(A - \lambda I_n)x_\lambda^{(k)}$ y de $x_\lambda^{(k-1)}$ coinciden.

Por otro lado, que las componentes n -ésimas de esos dos vectores sean iguales es equivalente a que sea

$$\begin{aligned} 0 &= (x_\lambda^{(k-1)})_n + \sum_{i=0}^{n-1} a_i (x_\lambda^{(k)})_{i+1} + \lambda (x_\lambda^{(k)})_n = \binom{n-1}{k-1} \lambda^{n-k} + \sum_{i=k}^{n-1} a_i \binom{i}{k} \lambda^{i-k} + \binom{n-1}{k} \lambda^{n-k} \\ &= \sum_{i=k}^n a_i \binom{i}{k} \lambda^{i-k}. \end{aligned} \quad (26)$$

Ahora bien, como λ es una raíz de multiplicidad m de p , hay un polinomio $q \in \mathbb{k}[X]$ tal que $p = (X - \lambda)^m q$ y entonces

$$X^m q(\lambda + X) = p(\lambda + X) = \sum_{i=0}^n a_i (\lambda + X)^i = \sum_{i=0}^n \sum_{j=0}^i a_i \binom{i}{j} \lambda^{i-j} X^j = \sum_{j=0}^n \left(\sum_{i=j}^n a_i \binom{i}{j} \lambda^{i-j} \right) X^j$$

y como X^m divide al primer miembro de esta cadena de igualdades, también divide al último y, en consecuencia, tenemos que

$$\sum_{i=j}^n a_i \binom{i}{j} \lambda^{i-j} = 0 \quad \text{para cada } j \in \llbracket 0, m-1 \rrbracket.$$

Vemos así que la suma (26) se anula, como queremos: esto completa la prueba. \square

6.7.3. Si para cada $j \in \llbracket 0, n \rrbracket$ ponemos

$$\Delta_j(p)(\lambda) := \sum_{i=j}^n a_i \binom{i}{j} \lambda^{i-j},$$

en el final de la prueba de este lema mostramos que

$$p(\lambda + X) = \sum_{j=0}^n \Delta_j(p)(\lambda) X^j. \quad (27)$$

Es inmediato verificar que para cada $j \in \llbracket n \rrbracket$ es $j! \cdot \Delta_j(p)(\lambda) = p^{(j)}(\lambda)$, el valor de la derivada j -ésima de p en λ , así que si la característica del cuerpo \mathbb{k} es 0 (o, más generalmente, si esa característica es mayor que n), de manera que $n!$ es inversible en \mathbb{k} , se tiene que

$$\Delta_j(p)(\lambda) = \frac{p^{(j)}(\lambda)}{j!} \quad (28)$$

para cada $j \in \llbracket 0, n \rrbracket$ y, por lo tanto, la igualdad (27) puede reescribirse en la forma

$$p(\lambda + X) = \sum_{j=0}^n \frac{p^{(j)}(\lambda)}{j!} X^j,$$

que es el desarrollo de Taylor de p alrededor de λ . Para un cuerpo cualquiera, sin embargo, el miembro derecho de la igualdad (28) simplemente no tiene sentido. Llamamos al escalar $\Delta_j(p)(\lambda)$ la j -ésima **derivada de Dieudonné** del polinomio p en λ , por *Jean Dieudonné*, que las estudió en [Die57]. Estas derivadas tienen propiedades formales muy similares a las derivadas usuales y, cuando trabajamos con cuerpos de característica positiva, mejores. Por ejemplo, es fácil ver que

un escalar λ es raíz de un polinomio $f \in \mathbb{k}[X]$ de multiplicidad al menos m si y solamente si $\Delta_j(f)(\lambda) = 0$ para cada $j \in \llbracket 0, m-1 \rrbracket$,

y esto es cierto cualquiera sea el cuerpo \mathbb{k} , mientras que la afirmación más conocida

un escalar λ es raíz de un polinomio $f \in \mathbb{k}[X]$ de multiplicidad al menos m si y solamente si $f^{(j)}(\lambda) = 0$ para cada $j \in \llbracket 0, m-1 \rrbracket$

vale solamente si el cuerpo \mathbb{k} tiene característica nula.

6.7.4. El Lema 6.7.2 nos permite calcular inmediatamente una matriz C tal que $C^{-1}AC = J$: la que tiene por columnas a los vectores

$$x_{\lambda_1}^{(m_1-1)}, \dots, x_{\lambda_1}^{(0)}, x_{\lambda_2}^{(m_2-1)}, \dots, x_{\lambda_2}^{(0)}, \dots, \dots, x_{\lambda_r}^{(m_r-1)}, \dots, x_{\lambda_r}^{(0)}.$$

Por ejemplo, si $p = (X - \alpha)^2(X - \beta)^3$, de manera que $r = 2$, $\lambda_1 = \alpha$, $\lambda_2 = \beta$, $m_1 = 2$ y $m_2 = 3$, esta matriz es

$$\begin{pmatrix} 0 & 1 & 0 & 0 & 1 \\ 1 & \alpha & 0 & 1 & \beta \\ 2\alpha & \alpha^2 & 1 & 2\beta & \beta^2 \\ 3\alpha^2 & \alpha^3 & 3\beta & 3\beta^2 & \beta^3 \\ 4\alpha^3 & \alpha^4 & 6\beta^2 & 4\beta^3 & \beta^4 \end{pmatrix}$$

Esta matriz es una generalización natural de la matriz de Vandermonde de 4.8.6 y como la de esta su determinante es importante:

6.7.5. Proposición. *Supongamos que el cuerpo \mathbb{k} tiene característica 0. Sean $r \in \mathbb{N}$, sean $\lambda_1, \dots, \lambda_r \in \mathbb{k}$ escalares y sean $m_1, \dots, m_r \in \mathbb{N}$. El determinante de la matriz cuadrada C de tamaño $n = m_1 + \dots + m_r$*

que tiene por columnas a los vectores

$$x_{\lambda_1}^{(m_1-1)}, \dots, x_{\lambda_1}^{(0)}, x_{\lambda_2}^{(m_2-1)}, \dots, x_{\lambda_2}^{(0)}, \dots, \dots, x_{\lambda_r}^{(m_r-1)}, \dots, x_{\lambda_r}^{(0)}.$$

es

$$\prod_{1 \leq i < j \leq r} (\lambda_j - \lambda_i)^{m_j m_j}.$$

Este resultado es cierto para cualquier cuerpo, pero la demostración que daremos requiere que tenga característica nula.

Demostración. Escribamos $m := m_1$, para cada elección de r enteros $i_1, \dots, i_r \in \mathbb{N}_0$ sea $C(i_1, \dots, i_r)$ la matriz de $\mathbf{M}_n(\mathbb{k}[X])$ cuyas columnas son

$$x_X^{(i_1)}, \dots, x_X^{(i_r)}, x_{\lambda_2}^{(m_2-1)}, \dots, x_{\lambda_2}^{(0)}, \dots, \dots, x_{\lambda_r}^{(m_r-1)}, \dots, x_{\lambda_r}^{(0)},$$

y sea $c(i_1, \dots, i_r)$ su determinante, que es un elemento de $\mathbb{k}[X]$. La matriz C del enunciado es $C(m-1, \dots, 0)(\lambda_1)$ y el determinante que queremos calcular es el valor del polinomio

$$p := c(m-1, \dots, 0)$$

en λ_1 . Necesitamos hacer la siguiente observación sobre estos polinomios:

si $i_1, \dots, i_r \in \mathbb{N}_0$, entonces la derivada del polinomio $c(i_1, \dots, i_r)$ es una combinación lineal de los polinomios $c(i_1, \dots, i_{t-1}, i_t + 1, i_{t+1}, \dots, i_r)$ con $t \in \llbracket r \rrbracket$.

si $i_1, \dots, i_r \in \mathbb{N}_0$ y $l \in \mathbb{N}_0$, entonces la derivada l -ésima del polinomio $c(i_1, \dots, i_r)$ es una combinación lineal de los polinomios $c(i_1 + j_1, \dots, i_r + j_r)$ con $j_1, \dots, j_r \in \mathbb{N}_0$ tales que $j_1 + \dots + j_r = l$.

□

Potencias matriciales

6.7.6. El problema que queremos resolver es el de calcular de la manera más explícita posible las potencias de una matriz de $\mathbf{M}_n(\mathbb{k})$. La primera observación que podemos hacer es muy sencilla:

Lema. Sea $n \in \mathbb{N}$ y sean A, B y C tres matrices de $\mathbf{M}_n(\mathbb{k})$. Si C es inversible y $A = CBC^{-1}$, entonces para todo $k \in \mathbb{N}_0$ es $A^k = CB^k C^{-1}$.

Demostración. Supongamos que C es inversible y que $A = CBC^{-1}$ y probemos que $A^k = CB^k C^{-1}$ para todo $k \in \mathbb{N}_0$ haciendo inducción con respecto a k . Es $A^0 = I_n = CC^{-1} = CIC^{-1} = CB^0 C^{-1}$. Por otro lado, si $k \in \mathbb{N}_0$ y $A^k = CB^k C^{-1}$, entonces $A^{k+1} = A^k \cdot A = CB^k C^{-1} \cdot CBC^{-1} = CB^{k+1} C^{-1}$. □

6.7.7. Si tenemos una matriz $A \in M_n(\mathbb{k})$ y queremos calcular las potencias de A , entonces el lema nos dice que si elegimos una matriz inversible cualquiera $C \in GL_n(\mathbb{k})$ y ponemos $B := C^{-1}AC$, tenemos que $A^k = CB^kC^{-1}$ para todo $k \in \mathbb{N}_0$. Si elegimos la matriz C de manera tal que la matriz B tenga potencias que sepamos calcular, entonces podremos calcular las potencias de A . Haremos esto eligiendo C de manera que B sea la forma normal de Jordan de A .

Para que esto funcione, tenemos que poder calcular las potencias de esta última y para eso empezamos por el siguiente cálculo:

Lema. Sea $\lambda \in \mathbb{k}$ y sea $n \in \mathbb{N}$. Para todo $k \in \mathbb{k}$ es

$$J_n(\lambda)^k = \begin{pmatrix} \lambda^k & & & & & \\ \binom{k}{1}\lambda^{k-1} & \lambda^k & & & & \\ \vdots & \binom{k}{1}\lambda^{k-1} & \ddots & & & \\ \vdots & \vdots & \ddots & \ddots & & \\ \binom{k}{n-3}\lambda^{k-n+3} & \vdots & & \ddots & & \\ \binom{k}{n-2}\lambda^{k-n+2} & \binom{k}{n-3}\lambda^{k-n+3} & \dots & \dots & \binom{k}{1}\lambda^{k-1} & \lambda^k \\ \binom{k}{n-1}\lambda^{k-n+1} & \binom{k}{n-2}\lambda^{k-n+2} & \binom{k}{n-3}\lambda^{k-n+3} & \dots & \dots & \binom{k}{1}\lambda^{k-1} & \lambda^k \end{pmatrix}.$$

Explícitamente, si $i, j \in \llbracket n \rrbracket$, entonces la componente (i, j) -ésima de esta matriz es

$$\begin{cases} \binom{k}{i-j}\lambda^{k-i+j} & \text{si } i \geq j; \\ 0 & \text{en caso contrario.} \end{cases}$$

Demostración. Sea $k \in \mathbb{N}_0$ y sea N_n el bloque de Jordan nilpotente de tamaño n , como en la Proposición 6.5.5. Es

$$J_n(\lambda) = \lambda I_n + N_n$$

y las matrices λI_n y N_n commutan, así que la fórmula de Newton nos dice que

$$J_n(\lambda)^k = (\lambda I_n + N_n)^k = \sum_{l=0}^k \binom{k}{l} (\lambda I_n)^{k-l} N_n^l = \sum_{l=0}^k \binom{k}{l} \lambda^{k-l} N_n^l. \quad (29)$$

Si para cada $l \in \llbracket 0, n-1 \rrbracket$ escribimos $N_n^l = (n_{i,j}^{(l)})$, entonces para cada elección de i y j en $\llbracket n \rrbracket$ tenemos que

$$n_{i,j}^{(l)} = \begin{cases} 1 & \text{si } i - j = l \\ 0 & \text{en caso contrario.} \end{cases}$$

Usando esto y la expresión (29) vemos que si $i, j \in \llbracket n \rrbracket$ entonces la componente (i, j) -ésima de la

matriz $J_n(\lambda)^k$ es

$$\sum_{l=0}^k \binom{k}{l} \lambda^{k-l} n_{i,j}^{(l)} = \begin{cases} \binom{k}{i-j} \lambda^{k-i+j} & \text{si } 0 \leq i - j \leq k; \\ 0 & \text{en caso contrario.} \end{cases}$$

Esto es lo que afirma el lema. \square

6.7.8. Sea $A \in M_n(\mathbb{k})$ una matriz cuyo polinomio minimal se factoriza como producto de factores de grado 1 en $\mathbb{k}[X]$. De acuerdo al Teorema ??, hay una matriz inversible C tal que $J := C^{-1}AC$ es la forma normal de Jordan de A . Hay entonces un entero $r \in \mathbb{N}$, escalares $\lambda_1, \dots, \lambda_r \in \mathbb{k}$, y enteros $n_1, \dots, n_r \in \mathbb{N}$ tales que

$$J = \begin{pmatrix} J_{n_1}(\lambda_1) & & \\ & \ddots & \\ & & J_{n_r}(\lambda_r) \end{pmatrix}$$

y para cada $k \in \mathbb{N}_0$ tenemos que

$$J^k = \begin{pmatrix} J_{n_1}(\lambda_1)^k & & \\ & \ddots & \\ & & J_{n_r}(\lambda_r)^k \end{pmatrix}.$$

Como consecuencia de esto y del Lema 6.7.6

$$A^k = CJ^kC^{-1} = C \begin{pmatrix} J_{n_1}(\lambda_1)^k & & \\ & \ddots & \\ & & J_{n_r}(\lambda_r)^k \end{pmatrix} C^{-1}$$

y el Lema 6.7.7 nos da una expresión explícita para cada una de las potencias de bloques de Jordan que aparecen en esta matriz. Esto nos da la expresión que queríamos para las potencias de A .

6.7.9. Ejemplo. Consideremos la matriz

$$A := \begin{pmatrix} 4 & -6 & -1 & 4 & 4 \\ 1 & -3 & -1 & 5 & 5 \\ 4 & -4 & -1 & -2 & -1 \\ -1 & 1 & 1 & -3 & 0 \\ 1 & -1 & -1 & 6 & 3 \end{pmatrix} \in M_5(\mathbb{C}).$$

Su polinomio característico es

$$\chi_A = X^5 - 15X^3 - 10X^2 + 60X + 72 = (X + 2)^3(X - 3)^2,$$

así que los autovalores de A son -2 y 3 . Calculando, determinamos los rangos de las matrices $(A + 2I_5)^i$ y $(A - 3I_5)^i$ y usando eso las dimensiones de sus núcleos:

i	0	1	2	3	4
$\dim \text{Nu}(A - 3I_5)^i$	0	1	2	2	2
$\dim \text{Nu}(A + 2I_5)^i$	0	1	2	3	3

De esto vemos inmediatamente que la forma normal de Jordan de A es la matriz diagonal de bloques

$$J = \begin{pmatrix} J_2(3) & 0 \\ 0 & J_3(-2) \end{pmatrix}$$

Como hay un sólo bloque para cada autovalor, es fácil calcular una matriz C tal que $C^{-1}AC = J$: buscamos vectores x_1 y x_5 que generen los núcleos de $A - 3I_5$ y de $A + 2I_5$, respectivamente, y luego vectores x_1, x_3, x_4 que formen cadenas

$$x_1 \xrightarrow{A+2I_5} x_2 \quad x_3 \xrightarrow{A+2I_5} x_4 \xrightarrow{A+2I_5} x_5$$

Por ejemplo, podemos elegir como x_1, x_2, x_3, x_4 y x_5 a las columnas de la matriz

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

Sabemos entonces que C es inversible y que $C^{-1}AC = J$, como queremos. Por otro lado, para cada $k \in \mathbb{N}_0$ el Lema 6.7.7 nos dice que

$$J^k = \begin{pmatrix} J_2(3)^k & 0 \\ 0 & J_3(-2)^k \end{pmatrix} = \begin{pmatrix} 3^k & 0 & 0 & 0 & 0 \\ k3^{k-1} & 3^k & 0 & 0 & 0 \\ 0 & 0 & (-2)^k & 0 & 0 \\ 0 & 0 & k(-2)^{k-1} & (-2)^k & 0 \\ 0 & 0 & \frac{1}{2}k(k-1)(-2)^{k-2} & k(-2)^{k-1} & (-2)^k \end{pmatrix}$$

y usando esto (y una computadora...) podemos ver que la matriz $A^k = CJ^kC^{-1}$ es

$$\begin{pmatrix} 3^k - k(k-5)(-2)^{k-3} & (k^2 - 5k - 8)(-2)^{k-3} - 3^k & k(k-5)(-2)^{k-3} & (8+k-k^2)(-2)^{k-3} + (3-k)3^{k-1} & (3-k)3^{k-1} - (-2)^k \\ (5k - k^2)(-2)^{k-3} & (k^2 - 5k - 8)(-2)^{k-3} & (k^2 - 5k)(-2)^{k-3} & (8-k-k^2)(-2)^{k-3} + 3^k & 3^k - (-2)^k \\ (2-k)(-2)^{k-1} + 3^k & (k-2)(-2)^{k-1} - 3^k & (k-2)(-2)^{k-1} & -k3^{k-1} - k(-2)^{k-1} & -k3^{k-1} \\ -k(-2)^{k-1} & k(-2)^{k-1} & k(-2)^{k-1} & -(k+2)(-2)^{k-1} & 0 \\ k(-2)^{k-1} & -k(-2)^{k-1} & -k(-2)^{k-1} & (k+2)(-2)^{k-1} + 3^k & 3^k \end{pmatrix}.$$

◇

Capítulo 7

Espacios con producto interno

§1. Espacios con producto interno

7.1.1. En todo este capítulo escribiremos \mathbb{k} para referirnos o bien al cuerpo \mathbb{R} de los números reales o bien al cuerpo \mathbb{C} de los números complejos. En cualquiera de los dos casos, si $\lambda \in \mathbb{k}$, escribimos $\bar{\lambda}$ al número conjugado de λ ; si $\mathbb{k} = \mathbb{R}$, entonces por supuesto es $\bar{\lambda} = \lambda$ para todo $\lambda \in \mathbb{k}$.

7.1.2. Si V es un espacio vectorial sobre \mathbb{k} , un *producto interno* sobre V es una función

$$\langle -, - \rangle : V \times V \rightarrow \mathbb{k}$$

tal que para cada $x, x', y \in V$ y cada $\lambda \in \mathbb{k}$ se tiene que

$$(\textbf{PI}_1) \quad \langle x + x', y \rangle = \langle x, y \rangle + \langle x', y \rangle,$$

$$(\textbf{PI}_2) \quad \langle \lambda x, y \rangle = \lambda \langle x, y \rangle,$$

$$(\textbf{PI}_3) \quad \langle x, y \rangle = \overline{\langle y, x \rangle},$$

$$(\textbf{PI}_4) \quad \langle x, x \rangle > 0 \text{ si } x \neq 0.$$

Notemos que si $x \in V$, entonces la tercera condición implica que $\langle x, x \rangle = \overline{\langle x, x \rangle}$ y, por lo tanto, el escalar $\langle x, x \rangle$ es un número real cualquiera sea \mathbb{k} : vemos así, en particular, que la última condición de esta definición tiene sentido.

Las condiciones **(PI₁)** y **(PI₂)** nos dicen que un producto interno es una función lineal de su primera variable, esto es, que para cada $v \in V$ la función

$$u \in V \mapsto \langle u, v \rangle \in \mathbb{k}$$

es lineal. Si $\mathbb{k} = \mathbb{R}$, la condición **(PI₃)** implica inmediatamente entonces que $\langle -, - \rangle$ es también lineal en su segunda variable y por lo tanto en ese caso se trata, de hecho, de una función bilineal. Si en cambio $\mathbb{k} = \mathbb{C}$, un producto interno es una función *semilineal* de su segundo argumento: esto significa que para cada $u, v, v' \in V$ y cada $\lambda \in \mathbb{k}$ es $\langle u, v + v' \rangle = \langle u, v \rangle + \langle u, v' \rangle$ y $\langle u, \lambda v \rangle = \bar{\lambda} \langle u, v \rangle$.

En cualquier caso, se sigue de la linealidad en la primera variable que

$$\langle 0, 0 \rangle = \langle 0 \cdot 0, 0 \rangle = 0 \cdot \langle 0, 0 \rangle = 0$$

y esto, junto con (PI₄), implica que, de hecho, para cada $x \in V$ vale que

$$\langle x, x \rangle = 0 \iff x = 0.$$

7.1.3. Un *espacio vectorial con producto interno* es un par ordenado $(V, \langle -, - \rangle)$ en el que V es un espacio vectorial sobre \mathbb{k} y $\langle -, - \rangle$ es un producto interno sobre V . Salvo en situaciones excepcionales, escribiremos simplemente V en lugar de $(V, \langle -, - \rangle)$ y diremos que V es un espacio con producto interno, dejando implícita la notación para el producto interno.

7.1.4. Ejemplos.

- (a) Si $n \in \mathbb{N}$, definimos una función $\langle -, - \rangle : \mathbb{k}^n \times \mathbb{k}^n \rightarrow \mathbb{k}$ poniendo, para cada par de vectores $x = (x_1, \dots, x_n)$, $y = (y_1, \dots, y_n) \in \mathbb{k}^n$,

$$\langle x, y \rangle = x_1 \bar{y}_1 + \dots + x_n \bar{y}_n.$$

Es fácil verificar que se trata de un producto interno sobre \mathbb{k}^n , al que llamamos el *producto interno estándar* de \mathbb{k}^n . Salvo indicación en contrario, cada vez que consideremos a \mathbb{k}^n como un espacio vectorial con producto interno será con respecto a este producto.

Más generalmente, si $\omega = (\omega_1, \dots, \omega_n)$ es un elemento de \mathbb{R}^n con componentes estrictamente positivas, la función $\langle -, - \rangle_\omega : \mathbb{k}^n \times \mathbb{k}^n \rightarrow \mathbb{k}$ tal que

$$\langle x, y \rangle_\omega = \omega_1 x_1 \bar{y}_1 + \dots + \omega_n x_n \bar{y}_n$$

cada vez que $x = (x_1, \dots, x_n)$ e $y = (y_1, \dots, y_n)$ son elementos de \mathbb{k}^n es un producto interno sobre \mathbb{k}^n , al que llamamos el *producto interno estándar de peso ω* sobre \mathbb{k}^n .

- (b) Escribamos \mathbb{k}^∞ al espacio vectorial de las sucesiones $x = (x_i)_{i \geq 1}$ de elementos de \mathbb{k} cuyas entradas son «casi todas nulas», esto es, aquellas para las que existe $n \geq 1$ tal que $x_i = 0$ para todo $i \geq n$. Hay una función $\langle -, - \rangle : \mathbb{k}^\infty \times \mathbb{k}^\infty \rightarrow \mathbb{k}$ tal que para cada par de sucesiones $x = (x_i)_{i \geq 1}$ e $y = (y_i)_{i \geq 1}$ de \mathbb{k}^∞ es

$$\langle x, y \rangle = \sum_{i=1}^{\infty} x_i \bar{y}_i.$$

Notemos que la suma de la derecha tiene sentido, ya que tiene un número finito de términos no nulos. Esta función $\langle -, - \rangle$ es un producto interno sobre \mathbb{k}^∞ .

- (c) Sea $\mathbb{k}[X]$ el espacio vectorial de los polinomios con coeficientes en \mathbb{k} y sea $I = [a, b]$ un intervalo cerrado y acotado de \mathbb{R} con $a < b$. Hay un producto interno $\langle -, - \rangle : \mathbb{k}[X] \times \mathbb{k}[X] \rightarrow \mathbb{k}$ tal que para cada $p, q \in \mathbb{k}[X]$ se tiene que

$$\langle p, q \rangle = \int_a^b p(t) \overline{q(t)} dt.$$

La verificación de las condiciones **(PI₁)**, **(PI₂)** y **(PI₃)** de la definición 7.1.2 es inmediata. Veamos la de **(PI₄)**. Sea $p \in \mathbb{k}[X]$ un polinomio. Es claro que

$$\langle p, p \rangle = \int_a^b |p(t)|^2 dt$$

es un número no negativo, ya que el integrando es real y no negativo. Más aún, si $\langle p, p \rangle = 0$, entonces es nula la integral y, como el integrando es una función continua y no negativa en todo el intervalo I , esto sólo es posible si ese integrando es allí idénticamente nulo: tiene que ser entonces $p(t) = 0$ para cada $t \in I$ y en consecuencia, ya que p es un polinomio y el intervalo I un conjunto infinito, $p = 0$. Esto prueba **(PI₄)**

Más generalmente, si $\omega : I \rightarrow \mathbb{R}$ es una función continua, no negativa y no idénticamente nula, es fácil ver que hay un producto interno $\langle -, - \rangle_\omega : \mathbb{k}[X] \times \mathbb{k}[X] \rightarrow \mathbb{k}$ sobre $\mathbb{k}[X]$ tal que para cada $p, q \in \mathbb{k}[X]$ se tiene que

$$\langle p, q \rangle_\omega = \int_a^b \omega(t) p(t) \overline{q(t)} dt.$$

- (d) Sea $I = [a, b]$ un intervalo cerrado y acotado de \mathbb{R} y sea $C(I)$ el espacio vectorial de las funciones $I \rightarrow \mathbb{k}$ que son continuas. Hay una función $\langle -, - \rangle : C(I) \times C(I) \rightarrow \mathbb{k}$ tal que cada vez que f y g son elementos de $C(I)$ se tiene que

$$\langle f, g \rangle = \int_a^b f(t) \overline{g(t)} dx.$$

En efecto, si f y g están en (I) la función $t \in I \mapsto f(t) \overline{g(t)} \in \mathbb{k}$ es continua y, por lo tanto, está definida su integral sobre I . Es fácil ver que $\langle -, - \rangle$ es un producto interno sobre $C(I)$.

- (e) Sea $C(\mathbb{R})$ el espacio vectorial de todas las funciones continuas $\mathbb{R} \rightarrow \mathbb{k}$ y sea $V \subseteq C_{\mathbb{k}}(\mathbb{R})$ un subespacio tal que

cualquiera sea $f \in V$ la integral de Riemann impropia $\int_{-\infty}^{\infty} |f(x)|^2 dx$ es finita.

Por ejemplo, podemos tomar como V al espacio de todas las funciones $f \in C(\mathbb{R})$ para las que existe $M > 0$ tal que $f(x) = 0$ si $|x| > M$.

Mostremos que

si f y g están en V , entonces la integral impropia $\int_{-\infty}^{\infty} f(x) \overline{g(x)} dx$ converge absolutamente. (1)

Sean para ello $f, g \in V$, sean $a = (\int_{-\infty}^{\infty} |f(x)|^2 dx)^{1/2}$ y $b = (\int_{-\infty}^{\infty} |g(x)|^2 dx)^{1/2}$, y fijemos $R > 0$. Si $t \in \mathbb{R}$, entonces

$$\begin{aligned} 0 &\leq \int_{-R}^R (|f(x)| - t|g(x)|)^2 dx \\ &= \int_{-R}^R |f(x)|^2 dx - 2t \int_{-R}^R |f(x)g(x)| dx + t^2 \int_{-R}^R |g(x)|^2 dx. \end{aligned}$$

El último miembro de esta desigualdad es un polinomio en t con coeficientes reales que no toma valores negativos. Su discriminante tiene que ser, por lo tanto, no positivo, y entonces

$$\left(\int_{-R}^R |f(x)g(x)| dx \right)^2 \leq \int_{-R}^R t^2 |f(x)| dx \cdot \int_{-R}^R t^2 |g(x)| dx \leq a^2 b^2.$$

Vemos así que para todo $R > 0$ es $\int_{-R}^R |f(x)g(x)| dx \leq ab$ y se sigue de esto que la integral impropia $\int_{-\infty}^{\infty} |f(x)g(x)| dx$ converge y, por lo tanto, que la integral impropia $\int_{-\infty}^{\infty} f(x)g(x) dx$ converge absolutamente, como queríamos ver.

Como consecuencias de (1) es claro que hay una función $\langle -, - \rangle : V \times V \rightarrow \mathbb{R}$ tal que

$$\langle f, g \rangle = \int_{-\infty}^{\infty} f(x) \overline{g(x)} dx$$

para cada $f, g \in V$. Esta función es un producto interno sobre V . \diamond

7.1.5. Proposición. *Sea V un espacio vectorial con producto interno. Si $x, x' \in V$, entonces*

- (i) $\langle x, y \rangle = 0$ para todo $x \in V$ si y solamente si $x = 0$, y
- (ii) $\langle x, y \rangle = \langle x', y \rangle$ para todo $y \in V$ si y solamente si $x = x'$.

Demostración. (i) Para ver la necesidad de la condición, notemos que si x la satisface se tiene en particular que $\langle x, x \rangle = 0$, así que $x = 0$. La suficiencia es inmediata.

(ii) Esto es consecuencia inmediata de la parte (i), ya que $\langle x, y \rangle - \langle x', y \rangle = \langle x - x', y \rangle$ cualquiera sea $y \in V$. \square

7.1.6. Podemos restringir el producto interno de un espacio vectorial con producto interno a cualquiera de sus subespacios:

Proposición. *Sea V un espacio vectorial con producto interno y sea W un subespacio de V . La función $\langle -, - \rangle_W : W \times W \rightarrow \mathbb{k}$ que se obtiene restringiendo el producto interno $\langle -, - \rangle : V \times V \rightarrow \mathbb{k}$ de V a $W \times W$ es un producto interno sobre W .*

Siempre que consideremos a un subespacio de un espacio vectorial con producto interno lo dotaremos con el producto interno construido como en esta proposición.

Demostración. La función $\langle -, - \rangle_W$ satisface cada una de las condiciones de la definición 7.1.2 simplemente porque $\langle -, - \rangle$ lo hace. \square

7.1.7. Proposición. *Sean $(V, \langle -, - \rangle_V)$ y $(W, \langle -, - \rangle_W)$ dos espacios con producto interno. Si $V \boxplus W$ es la suma directa externa de V y W , entonces la función*

$$\langle -, - \rangle : (V \boxplus W) \times (V \boxplus W) \rightarrow \mathbb{k}$$

tal que para cada de vectores (x, y) y (x', y') de $V \boxplus W$ es

$$\langle (x, y), (x', y') \rangle = \langle x, x' \rangle + \langle y, y' \rangle$$

es un producto interno sobre $V \boxplus W$.

Cuando lo dotamos de este producto interno, llamamos a $V \boxplus W$ la *suma directa ortogonal* de los espacios vectoriales con producto interno V y W .

Demostración. **HACER.**

□

§2. Normas y métricas

7.2.1. Si V es un espacio vectorial sobre \mathbb{k} , una *norma* sobre V es una función $\|-\| : V \rightarrow \mathbb{R}_{\geq 0}$ tal que para cada $x, y \in V$ y cada $\lambda \in \mathbb{k}$ se satisfacen las condiciones

$$(N_1) \quad \|x\| = 0 \text{ si y solamente si } x = 0;$$

$$(N_2) \quad \|\lambda x\| = |\lambda| \|x\|; \text{ y}$$

$$(N_3) \quad \|x + y\| \leq \|x\| + \|y\|.$$

Un *espacio vectorial normado* es un par ordenado $(V, \|-\|)$ en el que V es un espacio vectorial y $\|-\|$ es una norma sobre V . La desigualdad (N_3) es llamada la *desigualdad triangular*.

7.2.2. Proposición. *Sea V un espacio vectorial con producto interno. La función*

$$\|-\| : x \in V \mapsto \langle x, x \rangle^{1/2} \in \mathbb{R}.$$

es una norma sobre V y para cada $x, y \in V$ vale la desigualdad

$$|\langle x, y \rangle| \leq \|x\| \|y\|. \tag{2}$$

Más aún, vale aquí la igualdad si y solamente si el conjunto $\{x, y\}$ es linealmente dependiente.

De ahora en adelante consideraremos a todo espacio vectorial con producto interno implícitamente dotado de la norma que nos provee esta proposición, y la llamaremos la *norma asociada* a su producto interno. La desigualdad (2) que aparece en esta proposición es la *desigualdad de Cauchy-Bunyakovsky-Schwartz*, por Augustin-Louis Cauchy (1789–1857, Francia), Viktor Bunyakovsky (1804–1889, Rusia) y Hermann Schwarz (1843–1921, Alemania). El primero de estos autores probó la desigualdad en el caso del espacio \mathbb{k}^n con su producto interno estándar, el segundo para espacios vectoriales de funciones con producto interno dado por una integral, como en el Ejemplo 7.1.4(d), y el tercero probó esencialmente el caso general.

Demostración. Tenemos que verificar las tres condiciones de la definición 7.2.1. Las dos primeras son inmediatas. Para probar la tercera, mostremos antes la desigualdad (2). Sean $x, y \in V$; si $y = 0$,

la desigualdad es inmediata, así que podemos suponer que no es ése el caso. Para cada $\lambda \in \mathbb{k}$ es

$$0 \leq \|x - \lambda y\|^2 = \langle x - \lambda y, x - \lambda y \rangle = \langle x, x \rangle - \lambda \langle y, x \rangle - \bar{\lambda} (\langle x, y \rangle - \lambda \langle y, y \rangle).$$

Si tomamos $\lambda = \langle x, y \rangle / \langle y, y \rangle$, la expresión entre paréntesis se anula, así que tenemos que

$$0 \leq \langle x, x \rangle - \frac{\langle x, y \rangle \langle y, x \rangle}{\langle y, y \rangle} = \|x\|^2 - \frac{|\langle x, y \rangle|^2}{\|y^2\|},$$

que es equivalente a (2). Más aún, si vale la igualdad, tenemos que $\|x - \lambda y\|^2 = 0$, así que $x = \lambda y$.

Probemos ahora, usando la desigualdad de Cauchy–Bunyakovsky–Schwartz, que la tercera condición de la definición 7.2.1 también se satisface. Si $x, y \in V$, es

$$\begin{aligned} \|x + y\|^2 &= \langle x + y, x + y \rangle = \langle x, x \rangle + \langle x, y \rangle + \langle y, x \rangle + \langle y, y \rangle \\ &= \|x\|^2 + 2 \operatorname{Re} \langle x, y \rangle + \|y\|^2 \leq \|x\|^2 + 2|\langle x, y \rangle| + \|y\|^2 \\ &\leq \|x\|^2 + 2\|x\|\|y\| + \|y\|^2 = (\|x\| + \|y\|)^2 \end{aligned}$$

y claramente esto implica que vale la desigualdad triangular. \square

7.2.3. Corolario. Sea V un espacio vectorial con producto interno.

(i) (Ley del paralelogramo) Si $x, y \in V$, vale que

$$\|x + y\|^2 + \|x - y\|^2 = 2\|x\|^2 + 2\|y\|^2.$$

(ii) Si $\mathbb{k} = \mathbb{R}$, para cada $x, y \in V$ se tiene que

$$\langle x, y \rangle = \frac{1}{4}\|x + y\|^2 - \frac{1}{4}\|x - y\|^2.$$

Si $\mathbb{k} = \mathbb{C}$, en cambio, es

$$\langle x, y \rangle = \frac{1}{4}\|x + y\|^2 - \frac{1}{4}\|x - y\|^2 + \frac{i}{4}\|x + iy\|^2 - \frac{i}{4}\|x - iy\|^2.$$

La ley del paralelogramo tiene una interpretación geométrica muy directa, que puede verse representada gráficamente en la Figura 7.1. Por otro lado, la segunda parte de este corolario nos dice que en un espacio con producto interno el producto interno queda determinado por la norma correspondiente.

Demostración. Para ver la primera parte, sean x e y en V y calculamos:

$$\begin{aligned} \|x + y\|^2 + \|x - y\|^2 &= \langle x + y, x + y \rangle + \langle x - y, x - y \rangle \\ &= (\langle x, x \rangle + \langle x, y \rangle + \langle y, x \rangle + \langle y, y \rangle) + (\langle x, x \rangle - \langle x, y \rangle - \langle y, x \rangle + \langle y, y \rangle) \\ &= 2\langle x, x \rangle + 2\langle y, y \rangle \\ &= 2\|x\|^2 + 2\|y\|^2. \end{aligned}$$

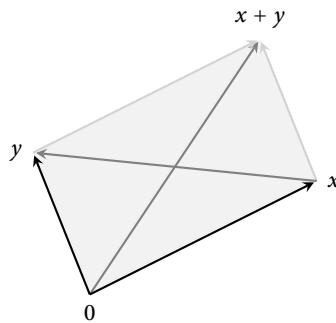


Figura 7.1. La igualdad de la ley del paralelogramo dice que la suma de los cuadrados de las longitudes de las diagonales de un paralelogramo es igual a la suma de los cuadrados de longitudes de los lados.

Para la segunda parte, suponemos que $\mathbb{k} = \mathbb{R}$ y vemos que

$$\begin{aligned}\|x + y\|^2 - \|x - y\|^2 &= \langle x + y, x + y \rangle - \langle x - y, x - y \rangle \\ &= (\langle x, x \rangle + \langle x, y \rangle + \langle y, x \rangle + \langle y, y \rangle) - (\langle x, x \rangle - \langle x, y \rangle - \langle y, x \rangle + \langle y, y \rangle) \\ &= 4\langle x, y \rangle.\end{aligned}$$

La igualdad que se afirma en la tercera parte de la proposición es consecuencia de un cálculo directo similar a este último, que omitimos. \square

7.2.4. La primera parte del Corolario 7.2.3 nos dice que la Ley del Paralelogramo es una condición necesaria para que la norma de un espacio normado provenga de un producto interno. El siguiente teorema de Pascual Jordan (1902–1980, Alemania) y John von Neumann (1903–1957, Hungría) publicado en [JvN35] afirma que también es una condición suficiente:

Proposición. *Sea V un espacio vectorial normado. Existe un producto interno sobre V cuya norma asociada es la de V si y solamente si satisface la ley del paralelogramo, esto es, si para todo par de vectores x, y de V se tiene que*

$$\|x + y\|^2 + \|x - y\|^2 = 2\|x\|^2 + 2\|y\|^2.$$

Demostración. Bastará que probemos que la condición es suficiente. Esta condición implica que

$$\|x + x' + y\|^2 = \|(x + y) + x'\|^2 = 2\|x + y\|^2 + 2\|x'\|^2 - \|x + y - x'\|^2$$

y que

$$\|x + x' - y\|^2 = \|x + (x' - y)\|^2 = 2\|x\|^2 + 2\|x' - y\|^2 - \|x - x' + y\|^2.$$

Restando miembro a miembro estas dos igualdades vemos que

$$\|x + x' + y\|^2 - \|x + x' - y\|^2 = 2\|x + y\|^2 + 2\|x'\|^2 - 2\|x\|^2 - 2\|x' - y\|^2$$

e intercambiando aquí los roles de x y de x' concluimos que también

$$\|x + x' + y\|^2 - \|x + x' - y\|^2 = 2\|x' + y\|^2 + 2\|x\|^2 - 2\|x'\|^2 - 2\|x - y\|^2.$$

Sumando ahora miembro a miembro estas dos igualdades llegamos a que

$$\|x + x' + y\|^2 - \|x + x' - y\|^2 = \|x + y\|^2 - \|x - y\|^2 + \|x' + y\|^2 - \|x' - y\|^2 \quad (3)$$

Consideremos primero el caso en que $\mathbb{k} = \mathbb{R}$. Sea $\langle -, - \rangle : V \times V \rightarrow \mathbb{R}$ la función tal que

$$\langle x, y \rangle = \frac{1}{4}\|x + y\|^2 - \frac{1}{4}\|x - y\|^2$$

para cada x y cada $y \in V$, y verifiquemos que se trata de un producto interno.

- Si $x, x', y \in V$, entonces vale que

$$\langle x + x', y \rangle = \langle x, y \rangle + \langle x', y \rangle,$$

ya que esta igualdad es, de acuerdo a la definición de la función $\langle -, - \rangle$, equivalente a (3).

Esto significa que la condición **(PI₁)** de la definición 7.1.2 se satisface.

- Si x e y son elementos de V se tiene que

$$\langle x, y \rangle = \frac{1}{4}\|x + y\|^2 - \frac{1}{4}\|x - y\|^2 = \frac{1}{4}\|y + x\|^2 - \frac{1}{4}\|y - x\|^2 = \langle y, x \rangle$$

y

$$\langle x, x \rangle = \frac{1}{4}\|x + x\|^2 - \frac{1}{4}\|x - x\|^2 = \|x\|^2 \geq 0, \quad (4)$$

así que las condiciones **(PI₃)** y **(PI₄)** también se cumplen.

- Para cada $r \in \mathbb{Q}$ sea $P(r)$ la afirmación

$$\text{para todo } x, y \in V \text{ se tiene que } \langle rx, y \rangle = r\langle x, y \rangle.$$

Queremos mostrar que $P(r)$ vale para todo $r \in \mathbb{Q}$, y lo hacemos en varias etapas:

- Observemos primero que si r es un número racional cualquiera y vale $P(r)$, entonces para cada $x, y \in V$ se tiene que

$$\langle (r+1)x, y \rangle = \langle rx + x, y \rangle = \langle rx, y \rangle + \langle x, y \rangle = r\langle x, y \rangle + \langle x, y \rangle = (r+1)\langle x, y \rangle,$$

de manera que también vale $P(r+1)$. Como claramente vale $P(1)$, esto implica que, de hecho, vale $P(r)$ para todo $r \in \mathbb{N}$.

- Vale $P(r)$ para todo $r \in \mathbb{Q}$ positivo. En efecto, si $r = \frac{p}{q}$ con $p, q \in \mathbb{N}$, para cada $x, y \in V$ se tiene, usando $P(p)$ y $P(q)$, que

$$p\langle x, y \rangle = \langle px, y \rangle = \left\langle q\left(\frac{p}{q}x\right), y \right\rangle = q\left\langle \frac{p}{q}x, y \right\rangle,$$

$$\text{así que } \left\langle \frac{p}{q}x, y \right\rangle = \frac{p}{q}\langle x, y \rangle.$$

– Vale $P(0)$, ya que

$$\langle 0x, y \rangle = \frac{1}{4} \|0x + y\|^2 - \frac{1}{4} \|0x - y\|^2 = \frac{1}{4} \|y\|^2 - \frac{1}{4} \|y\|^2 = 0 = 0\langle x, y \rangle,$$

y si vale $P(r)$ para algún $r \in \mathbb{Q}$ entonces también vale $P(-r)$: para todo x e y de V se tiene en ese caso que

$$0 = \langle 0, y \rangle = \langle rx - rx, y \rangle = \langle rx + (-r)x, y \rangle = \langle rx, y \rangle + \langle -rx, y \rangle,$$

por lo que $\langle -rx, y \rangle = -\langle rx, y \rangle = -r\langle x, y \rangle$. Junto con el paso anterior, esto nos permite concluir que vale $P(r)$ para todo $r \in \mathbb{Q}$, como queríamos.

- Sean otra vez x e y dos elementos de V y mostremos que

$$|\langle x, y \rangle| \leq \|x\| \|y\|. \quad (5)$$

Esto es evidente si $y = 0$, así que podemos suponer que no es ése el caso. En vista de lo que probamos en los dos pasos anteriores, sabemos que cada vez que $r \in \mathbb{Q}$ es

$$0 \leq \|x - ry\|^2 = \langle x - ry, x - ry \rangle = \langle x, x \rangle - 2r\langle x, y \rangle + r^2\langle y, y \rangle.$$

Esto nos dice que el polinomio $\|y\|^2 X^2 - 2\langle x, y \rangle X + \|x\|^2 \in \mathbb{R}[X]$ no toma valores negativos sobre \mathbb{Q} : como es una función continua, no toma entonces valores negativos en ningún punto de \mathbb{R} y, por lo tanto, su discriminante es no negativo, esto es, $\langle x, y \rangle^2 - \|x\|^2 \|y\|^2 \leq 0$. La desigualdad (5) es consecuencia inmediata de esto.

- Si $x, y \in V$, entonces la función $\zeta : t \in \mathbb{R} \mapsto \langle tx, y \rangle \in \mathbb{R}$ es continua. Para verlo basta observar que, de acuerdo a lo que sabemos ya de la función $\langle -, - \rangle$, se tiene que

$$|\zeta(s) - \zeta(t)| = |\langle sx, y \rangle - \langle tx, y \rangle| = |((s-t)x, y)| \leq \|(s-t)x\| \|y\| \leq |s-t| \|x\| \|y\|.$$

- Sean x e y dos elementos de V . Sabemos, gracias a los dos pasos anteriores, que la función $t \in \mathbb{R} \mapsto \langle tx, y \rangle - t\langle x, y \rangle \in \mathbb{R}$ es continua y que se anula sobre \mathbb{Q} . Esto implica, por supuesto, que es idénticamente nula y, por lo tanto, que para todo $t \in \mathbb{R}$ vale que $\langle tx, y \rangle = t\langle x, y \rangle$. La función $\langle -, - \rangle$ satisface entonces la condición (PI₂).

Concluimos de esta forma que la función $\langle -, - \rangle$ es un producto interno sobre V . De acuerdo a la igualdad (4), la norma asociada a este producto interno coincide con la norma de V : esto completa la prueba de la proposición en el caso que estamos considerando en el que $\mathbb{k} = \mathbb{R}$.

HACER: El caso complejo. □

7.2.5. Ejemplo. La función $\|-| : (x, y) \in \mathbb{k}^2 \mapsto |x| + |y| \in \mathbb{R}_{\geq 0}$ es una norma sobre \mathbb{k}^2 . Si $e_1 = (1, 0)$ y $e_2 = (0, 1)$ son los dos vectores de la base estándar de \mathbb{k}^2 , entonces

$$\|e_1 + e_2\|^2 + \|e_1 - e_2\|^2 = 8 \neq 4 = \|e_1\|^2 + 2\|e_2\|^2.$$

Vemos así que esta norma $\|-|$ no satisface la ley de paralelogramo y, en consecuencia, que no existe ningún producto interno sobre V que la tenga como norma asociada. ◇

7.2.6. Si V es un espacio vectorial, una función $d : V \times V \rightarrow \mathbb{R}_{\geq 0}$ es una **métrica** sobre V si para cada $x, y, z \in V$ se tiene que

- (M₁) $d(x, y) = d(y, x);$
- (M₂) $d(x, y) = 0$ si y solamente si $x = y$; y
- (M₃) $d(x, z) \leq d(x, y) + d(y, z).$

Decimos que esa métrica es **invariante** si

$$(M_4) \quad d(x, y) = d(x + z, y + z)$$

y que es **homogénea** si para cada $\lambda \in \mathbb{k}$ es

$$(M_5) \quad d(\lambda x, \lambda y) = |\lambda| d(x, y).$$

7.2.7. Proposición. Sea V un espacio vectorial con producto interno y sea $\| - \| : V \rightarrow \mathbb{R}$ la norma asociada. La función

$$d : (v, w) \in V \times V \mapsto \|v - w\| \in \mathbb{R}_{\geq 0}$$

es una métrica invariante y homogénea sobre V .

Demostración. Esto es una consecuencia inmediata de la definición de la función d y de las propiedades de la norma $\| - \|$. \square

§3. Ortogonalidad

7.3.1. Sea V un espacio con producto interno. Si x e y son dos vectores de V , decimos que x e y son **ortogonales** si $\langle u, v \rangle = 0$, y en ese caso escribimos $u \perp v$. Es inmediato que esto define una relación \perp entre los elementos de V que es simétrica. Además, tenemos que

Lema. Sea V un espacio vectorial con producto interno.

- (i) Todo vector de V es ortogonal con 0.
- (ii) Si $x \in V$ es tal que $x \perp y$ para todo $y \in V$, entonces $x = 0$.

Demostración. La primera afirmación es inmediata, y la segunda es consecuencia de la primera parte de la Proposición 7.1.5. \square

7.3.2. Intuitivamente, dos vectores son ortogonales si el ángulo que forman es recto. Esto se ve reflejado en que vale el teorema de Pitágoras:

Proposición. Sea V un espacio con producto interno. Si $x, y \in V$, entonces

$$(i) \text{ si } x \perp y, \text{ entonces } \|x + y\|^2 = \|x\|^2 + \|y\|^2.$$

Más generalmente, vale que

$$(ii) \ \|x + y\|^2 = \|x\|^2 + \|y\|^2 + 2 \operatorname{Re}\langle x, y \rangle.$$

Demostración. La primera afirmación es consecuencia inmediata de la segunda, así que basta probar esta última, que a su vez sigue de un cálculo directo: si $x, y \in V$, la definición de la norma implica que

$$\begin{aligned} \|x + y\|^2 &= \langle x + y, x + y \rangle \\ &= \langle x, x \rangle + \langle x, y \rangle + \langle y, x \rangle + \langle y, y \rangle \\ &= \|x\|^2 + \|y\|^2 + \langle x, y \rangle + \overline{\langle x, y \rangle} \\ &= \|x\|^2 + \|y\|^2 + 2 \operatorname{Re}\langle x, y \rangle. \end{aligned}$$

□

7.3.3. Motivados por la Proposición 7.3.2, hacemos la siguiente definición: si V es un espacio vectorial con producto interno y x e y son dos vectores de V , el **ángulo** entre x e y es el único número real $\theta(x, y) \in [0, \pi]$ tal que

$$\cos \theta(x, y) = \frac{\operatorname{Re}\langle x, y \rangle}{\|x\| \|y\|}. \quad (6)$$

Esta definición tiene sentido: de la desigualdad de Cauchy-Bunyakovsky-Schwartz sabemos que

$$|\operatorname{Re}\langle x, y \rangle| \leq |\langle x, y \rangle| \leq \|x\| \|y\|,$$

así que el cociente que aparece a la derecha en la definición (6) es un elemento del intervalo $[-1, 1]$ y, por lo tanto, es el coseno de exactamente un número θ en el intervalo $[0, \pi]$. Es inmediato que dos vectores son ortogonales si y solamente si el ángulo entre ellos es $\pi/2$.

Usando esta definición, podemos reinterpretar la segunda parte de la Proposición 7.3.2 como el *teorema del coseno* de la geometría euclídea:

Proposición. Sea V un espacio vectorial con producto interno. Si x e y son dos vectores de V y $\theta(x, y)$ es el ángulo entre x e y , entonces

$$\|x - y\|^2 = \|x\|^2 + \|y\|^2 - 2 \cos \theta(x, y).$$

Demostración. Basta reemplazar y por $-y$ en la igualdad de la Proposición 7.3.2(ii) y usar la definición de $\theta(x, y)$. □

7.3.4. Si V es un espacio con producto interno y A es un subconjunto de V , entonces decimos que

- A es **ortogonal** si para cada par $x, y \in A$ de elementos distintos se tiene que $x \perp y$, y que
- A es **ortonormal** si es ortogonal y además $\|x\| = 1$ para cada $x \in A$.

7.3.5. Hay una relación muy útil entre la ortogonalidad y la independencia lineal:

Proposición. Sea V un espacio con producto interno y sea A un subconjunto de V .

- (i) Si A es ortogonal y $0 \notin A$, entonces A es linealmente independiente.
- (ii) Si A es ortonormal y si $x = \lambda_1 x_1 + \dots + \lambda_n x_n$, con $x_1, \dots, x_n \in A$ distintos dos a dos y $\lambda_1, \dots, \lambda_n \in \mathbb{k}$, entonces $\lambda_j = \langle x, x_j \rangle$ para todo $j \in [\![n]\!]$.
- (iii) Si A es ortonormal, entonces para cada $x \in \langle A \rangle$ el conjunto $A_x = \{y \in A : \langle x, y \rangle \neq 0\}$ es finito, y vale que $x = \sum_{y \in A_x} \langle x, y \rangle y$.

Demostración. (i) Sean $n \geq 1$, sean $x_1, \dots, x_n \in A$ distintos dos a dos y sean $\lambda_1, \dots, \lambda_n \in \mathbb{k}$ tales que $\sum_{i=1}^n \lambda_i x_i = 0$. Si $j \in [\![n]\!]$, entonces

$$0 = \langle 0, x_j \rangle = \left\langle \sum_{i=1}^n \lambda_i x_i, x_j \right\rangle = \sum_{i=1}^n \lambda_i \langle x_i, x_j \rangle.$$

Como A es ortogonal, en esta última suma el único término que puede ser no nulo es aquél en el que i es igual a j , así que tenemos que $\lambda_j \langle x_j, x_j \rangle = 0$. Como $0 \notin A$, es $\langle x_j, x_j \rangle \neq 0$ y entonces $\lambda_j = 0$. Esto vale para cada $j \in [\![n]\!]$, así que podemos concluir que A es linealmente independiente.

(ii) Si $j \in [\![n]\!]$, entonces

$$\langle x, x_j \rangle = \left\langle \sum_{i=1}^n \lambda_i x_i, x_j \right\rangle = \sum_{i=1}^n \lambda_i \langle x_i, x_j \rangle.$$

Como A es ortonormal, el único término de esta suma que es posiblemente no nulo es el que tiene i igual a j , que es igual a $\lambda_j \langle x_j, x_j \rangle = \lambda_j$. Esto prueba lo que queremos.

(iii) Sea $x \in \langle A \rangle$, de manera que existen $n \geq 0$, $x_1, \dots, x_n \in A$ y $\lambda_1, \dots, \lambda_n \in \mathbb{k}$ tales que $x = \sum_{i=1}^n \lambda_i x_i$, y supongamos, sin pérdida de generalidad, que $\lambda_i \neq 0$ para cada $i \in [\![n]\!]$.

Si $y \in A$, entonces $\langle x, y \rangle = \sum_{i=1}^n \lambda_i \langle x_i, y \rangle$. Como A es ortonormal, esta suma tiene todos sus términos nulos si $y \notin \{x_1, \dots, x_n\}$. Así, vemos que A_x es un subconjunto de $\{x_1, \dots, x_n\}$ y, en particular, que es finito. Para terminar, es suficiente con mostrar que $\lambda_i = \langle x, x_i \rangle$ para cada $i \in [\![n]\!]$: esto es consecuencia inmediata de la afirmación (ii). \square

7.3.6. Corolario. Sea V un espacio vectorial con producto interno y de dimensión finita n , y supongamos que $\mathcal{B} = \{x_1, \dots, x_n\}$ es una base ortonormal de V . Para cada $x \in V$ es

$$x = \sum_{i=1}^n \langle x, x_i \rangle x_i.$$

Demostración. Esto es un caso particular de la tercera parte de la Proposición 7.3.5(iii). \square

7.3.7. Es fácil construir conjuntos ortonormales:

Proposición. Sea V un espacio vectorial con producto interno. Si $n \geq 1$ y x_1, \dots, x_n son n vectores de V tal que el conjunto $\{x_1, \dots, x_n\}$ tiene n elementos y es linealmente independiente, entonces hay vectores y_1, \dots, y_n en V tales que

- el conjunto $\{y_1, \dots, y_n\}$ es ortonormal,
- para todo $i \in \llbracket n \rrbracket$ se tiene que $\langle x_1, \dots, x_i \rangle = \langle y_1, \dots, y_i \rangle$, y
- para todo $i \in \llbracket n \rrbracket$ es $\langle x_i, y_i \rangle > 0$.

Estas condiciones determinan únicamente a los vectores y_1, \dots, y_n .

Demostración. Probemos la existencia haciendo inducción con respecto a n .

Supongamos primero que $n = 1$. Como $\{x_1\}$ es linealmente independiente, es $x_1 \neq 0$ y entonces $\|x_1\| \neq 0$: podemos considerar entonces el vector $y_1 = x_1 / \|x_1\|$. Es inmediato que $\|y_1\| = 1$, de manera que $\{y_1\}$ es un conjunto ortonormal, que $\langle x_1 \rangle = \langle y_1 \rangle$ y que $\langle x_1, y_1 \rangle > 0$. La afirmación de la proposición es por lo tanto cierta en este caso.

Supongamos ahora que $n > 1$. Como $\{x_1, \dots, x_{n-1}\}$ es linealmente independiente y tiene $n - 1$ elementos, podemos suponer inductivamente que hay un conjunto ortonormal $\{y_1, \dots, y_{n-1}\}$ tal que para cada $i \in \llbracket n-1 \rrbracket$ es $\langle x_1, \dots, x_i \rangle = \langle y_1, \dots, y_i \rangle$. Sea

$$\tilde{y}_n = x_n - \sum_{i=1}^{n-1} \langle x_n, y_i \rangle y_i.$$

Si fuese $\tilde{y}_n = 0$, tendríamos que

$$x_n = \sum_{i=1}^{n-1} \langle x_n, y_i \rangle y_i \in \langle y_1, \dots, y_{n-1} \rangle = \langle x_1, \dots, x_{n-1} \rangle,$$

lo que es imposible, ya que $\{v_1, \dots, v_n\}$ es linealmente independiente. Podemos entonces considerar el vector $y_n = \tilde{y}_n / \|\tilde{y}_n\|$, que tiene norma unitaria. Si $j \in \llbracket n-1 \rrbracket$, entonces

$$\begin{aligned} \|\tilde{y}_n\| \langle y_n, y_j \rangle &= \left\langle x_n - \sum_{i=1}^{n-1} \langle x_n, y_i \rangle y_i, y_j \right\rangle = \langle x_n, y_j \rangle - \sum_{i=1}^{n-1} \langle x_n, y_i \rangle \langle y_i, y_j \rangle \\ &= \langle x_n, y_j \rangle - \langle x_n, y_j \rangle = 0, \end{aligned}$$

y esto, junto con el hecho de que $\{y_1, \dots, y_{n-1}\}$ es un conjunto ortonormal, implica que el conjunto $\{y_1, \dots, y_{n-1}, y_n\}$ es ortonormal. Como es claro que

$$\langle x_1, \dots, x_n \rangle = \langle y_1, \dots, y_n \rangle,$$

esto completa la inducción.

HACER: Unidad. □

7.3.8. La demostración que dimos de la Proposición 7.3.7 nos da un procedimiento efectivo para obtener conjuntos ortonormales, llamado **procedimiento de ortogonalización de Gram–Schmidt**, por Jørgen Pedersen Gram (1850–1916, Dinamarca) y Erhard Schmidt (1876–1959, Alemania).

Empezando con n vectores x_1, \dots, x_n linealmente independientes calculamos, en orden,

$$\begin{aligned}\tilde{y}_1 &= x_1, & y_1 &= \tilde{y}_1 / \|\tilde{y}_1\|, \\ \tilde{y}_2 &= x_2 - \langle x_2, y_1 \rangle y_1, & y_2 &= \tilde{y}_2 / \|\tilde{y}_2\|, \\ \tilde{y}_3 &= x_3 - \langle x_3, y_1 \rangle y_1 - \langle x_3, y_2 \rangle y_2, & y_3 &= \tilde{y}_3 / \|\tilde{y}_3\|, \\ &\dots, & &\dots, \\ \tilde{y}_n &= x_n - \langle x_n, y_1 \rangle y_1 - \langle x_n, y_2 \rangle y_2 - \dots - \langle x_n, y_{n-1} \rangle y_{n-1}, & y_n &= \tilde{y}_n / \|\tilde{y}_n\|.\end{aligned}$$

Al terminar, el conjunto $\{y_1, \dots, y_n\}$ es ortonormal y genera el mismo espacio que $\{x_1, \dots, x_n\}$.

7.3.9. Ejemplos.

- (a) Consideremos al espacio vectorial \mathbb{K}^3 dotado de su producto interno estándar y los tres vectores

$$x_1 = (1, 1, 0), \quad x_2 = (2, 2, 2), \quad x_3 = (0, 1, 1).$$

Siguiendo el procedimiento de ortogonalización descripto arriba, encontramos:

$$\begin{aligned}y_1 &= x_1 / \|x_1\| = \left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0\right), \\ \tilde{y}_2 &= x_2 - \langle x_2, y_1 \rangle y_1 = (0, 0, 2), \\ y_2 &= \tilde{y}_2 / \|\tilde{y}_2\| = (0, 0, 1), \\ \tilde{y}_3 &= x_3 - \langle x_3, y_1 \rangle y_1 - \langle x_3, y_2 \rangle y_2 = \left(-\frac{1}{2}, \frac{1}{2}, 0\right), \\ y_3 &= \tilde{y}_3 / \|\tilde{y}_3\| = \left(-\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0\right).\end{aligned}$$

El conjunto $\{\left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0\right), (0, 0, 1), \left(-\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0\right)\}$ es entonces ortonormal y genera el mismo subespacio que $\{x_1, x_2, x_3\}$.

- (b) Sea ahora $\mathbb{R}[X]$ el espacio vectorial real de los polinomios con coeficientes en \mathbb{R} y consideremos sobre él el producto interno $\langle -, - \rangle : \mathbb{R}[X] \times \mathbb{R}[X] \rightarrow \mathbb{R}$ tal que

$$\langle p, q \rangle = \frac{2}{\pi} \int_{-1}^1 f(t)g(t)\sqrt{1-t^2} dt$$

para cada $p, q \in \mathbb{R}[X]$; observemos que esto es un caso particular de la construcción hecha en el Ejemplo 7.1.4(c) con $I = [-1, 1]$ y $\omega : t \in I \mapsto \frac{2}{\pi}\sqrt{1-t^2} \in \mathbb{R}$. Un cálculo directo muestra que si realizamos el procedimiento de ortogonalización de Gram–Schmidt empezando con los polinomios

$$1, \quad X, \quad X^2, \quad X^3, \quad X^4, \quad \dots$$

obtenemos los polinomios

$$1, \quad 2X, \quad 4X^2 - 1, \quad 8X^3 - 4X, \quad 16X^4 - 12X^2 + 1, \quad 32X^5 - 32X^3 + 6X, \quad \dots$$

Estos polinomios son precisamente los polinomios de Chebyshev que encontramos en el Ejemplo 5.4.12(b) del Capítulo 5. Probemos esto.

Para cada $n \in \mathbb{N}_0$ sea U_n el n -ésimo polinomio de Chebyshev, de manera que

$$U_0 = 1, \quad U_1 = 2X \tag{7}$$

y

$$U_n = 2XU_{n-1} - U_{n-2} \text{ si } n \geq 2. \tag{8}$$

Si $n, m \in \mathbb{N}_0$, entonces

$$\langle U_n, U_m \rangle = \frac{2}{\pi} \int_{-1}^1 U_n(t) U_m(t) \sqrt{1-t^2} dt. \tag{9}$$

Como vimos en el Ejemplo 5.4.12(b), vale que si $\theta \in (0, \pi)$ es

$$U_n(\cos \theta) = \frac{\sin((n+1)\theta)}{\sin \theta}.$$

y entonces, poniendo $t = \cos \theta$ en la integral (9), vemos que

$$\begin{aligned} \langle U_n, U_m \rangle &= \frac{2}{\pi} \int_0^\pi U_n(\cos \theta) U_m(\cos \theta) \sin^2 \theta d\theta \\ &= \frac{2}{\pi} \int_0^\pi \sin((n+1)\theta) \cdot \sin((m+1)\theta) d\theta \\ &= \frac{1}{\pi} \int_0^\pi \cos((n-m)\theta) d\theta - \frac{1}{\pi} \int_0^\pi \cos((n+m+2)\theta) d\theta \end{aligned} \tag{10}$$

Ahora bien, es fácil ver que para cada $k \in \mathbb{Z}$ se tiene que

$$\int_0^\pi \cos k\theta d\theta = \begin{cases} \pi, & \text{si } k = 0; \\ 0, & \text{en caso contrario}; \end{cases}$$

y usando esto en (10) concluimos inmediatamente que

$$\langle U_n, U_m \rangle = \begin{cases} 1, & \text{si } n = m; \\ 0, & \text{en caso contrario}. \end{cases}$$

Esto prueba que el conjunto $\{U_n : n \in \mathbb{N}_0\}$ es ortonormal.

Por otro lado, una demostración inductiva a partir de (7) y (8) muestra inmediatamente que para todo $n \in \mathbb{N}_0$ el grado de U_n es n y que el coeficiente de X^n en U_n es 2^n . **HACER:**
Terminar.

◇

7.3.10. La consecuencia más importante de la Proposición 7.3.7 es la siguiente:

Corolario. *Un espacio vectorial con producto interno y de dimensión finita tiene una base ortonormal.*

Demostración. Si V es un espacio vectorial con producto interno y de dimensión finita y \mathcal{B} es una base de V , la proposición nos da un conjunto ortonormal \mathcal{B}' con la misma cantidad de elementos que \mathcal{B} y que, entre otras cosas, tiene la propiedad de que $\langle \mathcal{B}' \rangle = \langle \mathcal{B} \rangle = V$. Esto implica que \mathcal{B}' es una base de V . Como \mathcal{B}' es ortonormal, esto prueba el corolario. □

7.3.11. Vimos como un producto interno nos permite «medir» la distancia entre dos elementos de un espacio vectorial. Extendemos ahora esa idea para poder hablar de la distancia de un punto a un conjunto. Si V es un espacio vectorial con producto interno, S un subconjunto no vacío de V y $x \in V$, la *distancia de x a S* es el número

$$d(x, S) = \inf\{d(x, y) : y \in S\}.$$

Observemos que esta definición tiene sentido, ya que el subconjunto $\{d(x, y) : y \in S\}$ de \mathbb{R} no es vacío y está acotado inferiormente por 0, así que posee un ínfimo.

7.3.12. Proposición. *Sea V un espacio vectorial con producto interno, sea S un subespacio de dimensión finita de V y sea $x \in V$. Si $\mathcal{B} = \{y_1, \dots, y_n\}$ es una base ortonormal de S y*

$$x_S = \sum_{j=1}^n \langle x, y_j \rangle y_j,$$

entonces

- (i) $x - x_S \perp S$,
- (ii) $d(x, S) = d(x, x_S)$, y
- (iii) $d(x, x_S) < d(x, y)$ para todo $y \in S \setminus \{x_S\}$.

Esta proposición nos dice que el vector x_S es un punto de S tal que $d(x, S) = d(x, x_S)$ y que es, más aún, el único con esa propiedad.

Demuestração. (i) Sea $y \in S$. Como \mathcal{B} es una base ortonormal de S , es $y = \sum_{i=1}^n \langle y, y_i \rangle y_i$ y entonces

$$\begin{aligned} \langle x - x_S, y \rangle &= \left\langle x - \sum_{j=1}^n \langle x, y_j \rangle y_j, \sum_{i=1}^n \langle y, y_i \rangle y_i \right\rangle \\ &= \sum_{i=1}^n \overline{\langle y, y_i \rangle} \langle x, y_i \rangle - \sum_{j=1}^n \sum_{i=1}^n \langle x, y_j \rangle \overline{\langle y, y_i \rangle} \langle y_j, y_i \rangle \\ &= \sum_{i=1}^n \overline{\langle y, y_i \rangle} \langle x, y_i \rangle - \sum_{j=1}^n \langle x, y_j \rangle \overline{\langle y, y_j \rangle} = 0. \end{aligned}$$

Vemos así que $x - x_S \perp S$.

(ii) Como $x_S \in S$, es claro que $d(x, S) \leq d(x, x_S)$. Si, por otro lado, es $y \in S$, tenemos que

$$\begin{aligned} d(x, y)^2 &= \|x - y\|^2 \\ &= \|(x - x_S) + (x_S - y)\|^2 \end{aligned}$$

y, como $x - x_S \perp x_S - y$ porque $x_S - y \in S$, esto es

$$\begin{aligned} &= \|x - x_S\|^2 + \|x_S - y\|^2 \\ &\geq \|x - x_S\|^2, \end{aligned}$$

de manera que $d(x, y) \geq d(x, x_S)$. Vemos así que $d(x, S) \geq d(x, x_S)$.

(iii) Si $y \in S \setminus \{x_S\}$, entonces $d(x_S, y) > 0$ y, como recién,

$$d(x, y)^2 = d(x, x_S)^2 + d(x_S, y)^2 > d(x, x_S)^2,$$

así que $d(x, y) > d(x, x_S)$. □

§4. Complementos ortogonales

7.4.1. Si V es un espacio vectorial con producto interno y S es un subconjunto arbitrario de V , el **complemento ortogonal** de S es el subconjunto

$$S^\perp = \{x \in V : x \perp s \text{ para todo } s \in S\}.$$

7.4.2. Proposición. *Sea V un espacio vectorial con producto interno.*

- (i) *Si S es un subconjunto de V , entonces S^\perp es un subespacio de V .*
- (ii) *Se tiene que $0^\perp = V$ y $V^\perp = 0$.*
- (iii) *Si S y T son subconjuntos de V y $S \subseteq T$, entonces $T^\perp \subseteq S^\perp$.*
- (iv) *Si S es un subconjunto de V , entonces $\langle S \rangle^\perp = S^\perp$.*
- (v) *Si S y T son subespacios de V , entonces $(S + T)^\perp = S^\perp \cap T^\perp$.*

Demostración. (i) Sean $x, y \in S^\perp$ y $a, b \in \mathbb{k}$. Para cada $s \in S$, es

$$\langle ax + by, s \rangle = a\langle x, s \rangle + b\langle y, s \rangle = 0,$$

porque ambos términos se anulan. Esto nos dice que $ax + by \in S^\perp$. Como S^\perp no es vacío, ya que contiene a 0, se trata de un subespacio de V .

(ii) Todo vector es ortogonal a 0, así que $V \subseteq 0^\perp$. La primera igualdad es entonces inmediata. Por otro lado, si $x \in V^\perp$, entonces $\langle x, y \rangle = 0$ para todo $y \in V$ y la Proposición 7.1.5 nos dice que $x = 0$. Como por supuesto $0 \in V^\perp$, esto prueba la segunda igualdad del enunciado.

(iii) Supongamos que S y T son subconjuntos de V y que $S \subseteq T$. Si $y \in T^\perp$, entonces para cada $s \in S$ se tiene que $y \perp s$, ya que $s \in T$ y, en consecuencia, $x \in S^\perp$. Así, es $T^\perp \subseteq S^\perp$, como queremos.

(iv) Como $S \subseteq \langle S \rangle$, la parte (iii) implica que $\langle S \rangle^\perp \subseteq S^\perp$. Sea, por otro lado, $x \in S^\perp$ y sea $y \in \langle S \rangle$. Existen entonces $n \geq 0$, $s_1, \dots, s_n \in S$ y $\lambda_1, \dots, \lambda_n \in \mathbb{k}$ tales que $y = \sum_{i=1}^n \lambda_i s_i$ y, en consecuencia,

$$\langle x, y \rangle = \left\langle x, \sum_{i=1}^n \lambda_i s_i \right\rangle = \sum_{i=1}^n \bar{\lambda}_i \langle x, s_i \rangle = 0$$

porque $x \in S^\perp$. Esto muestra que $S^\perp \subseteq \langle S \rangle^\perp$.

(v) Sean ahora S y T dos subespacios de V . Como S y T están contenidos en $S + T$, la parte (ii) implica que $(S + T)^\perp \subseteq S^\perp$ y $(S + T)^\perp \subseteq T^\perp$, así que $(S + T)^\perp \subseteq S^\perp \cap T^\perp$. Por otro lado, si $x \in S^\perp \cap T^\perp$ e $y \in S + T$, de manera que existen $s \in S$ y $t \in T$ tales que $y = s + t$, entonces

$$\langle x, y \rangle = \langle x, s + t \rangle = \langle x, s \rangle + \langle x, t \rangle = 0.$$

Esto nos dice que $x \in (S + T)^\perp$ y, en definitiva, que $S^\perp \cap T^\perp \subseteq (S + T)^\perp$. \square

7.4.3. Si S es un subconjunto de un espacio vectorial con producto interno, escribimos $S^{\perp\perp} = (S^\perp)^\perp$ y $S^{\perp\perp\perp} = (S^{\perp\perp})^\perp$. Observemos que $S^{\perp\perp\perp}$ coincide con $(S^\perp)^\perp$.

Proposición. *Sea V un espacio vectorial con producto interno. Si S es un subconjunto de V , entonces*

- (i) $S \cap S^\perp \subseteq 0$,
- (ii) $S \subseteq S^{\perp\perp}$,
- (iii) $S^\perp = S^{\perp\perp\perp}$.

Demostración. (i) Si $x \in S \cap S^\perp$, entonces $\langle x, x \rangle = 0$, ya que $x \in S$ y $x \in S^\perp$, y por lo tanto $x = 0$.

(ii) Si $s \in S$, para cada $t \in S^\perp$ es $s \perp t$. Esto dice que $s \in (S^\perp)^\perp = S^{\perp\perp}$.

(iii) Acabamos de probar que $S \subseteq S^{\perp\perp}$, así que usando la Proposición 7.4.2(iii) vemos que $S^{\perp\perp\perp} = (S^{\perp\perp})^\perp \subseteq S^\perp$. Por otro lado, según la parte (ii), $S^\perp \subseteq (S^\perp)^\perp = S^{\perp\perp\perp}$. \square

7.4.4. En la situación de la Proposición 7.4.3(ii), la inclusión es en general estricta. Vale la igualdad, sin embargo, en un caso importante:

Proposición. *Sea V un espacio vectorial con producto interno. Si S es un subespacio de V de dimensión finita, entonces $V = S \oplus S^\perp$ y $S = S^{\perp\perp}$.*

Demostración. Sea $x \in V$. De acuerdo a la Proposición 7.3.12, existe $x_S \in S$ tal que $x - x_S \perp S$, esto es, tal que $x - x_S \in S^\perp$. Entonces $x = x_S + (x - x_S) \in S + S^\perp$ y vemos que $V = S + S^\perp$. Como sabemos que $S \cap S^\perp = 0$, esto implica que $V = S \oplus S^\perp$.

Para probar que $S = S^{\perp\perp}$ basta mostrar, en vista de la Proposición 7.4.3(ii), que $S^{\perp\perp} \subseteq S$. Sea entonces $x \in S^{\perp\perp}$. Como $V = S \oplus S^\perp$, existen $s \in S$ y $t \in S^\perp$ tales que $x = s + t$. Como $x \in S^{\perp\perp}$, es

$$0 = \langle x, t \rangle = \langle s + t, t \rangle = \langle s, t \rangle + \langle t, t \rangle = \langle t, t \rangle,$$

así que $t = 0$ y, por lo tanto, $x = s + t = s \in S$. \square

7.4.5. Corolario. *Sea V un espacio vectorial con producto interno y de dimensión finita. Si $S \subseteq V$ es un subespacio, entonces*

$$\dim V = \dim S + \dim S^\perp.$$

Demostración. Esto es consecuencia de que, de acuerdo a la proposición, $V = S \oplus S^\perp$. \square

7.4.6. Ejemplo. Demos un ejemplo de una situación en la que la inclusión de la Proposición 7.4.3(ii) es propia. Sea ℓ_2 el conjunto de todas las sucesiones $(x_i)_{i \geq 1}$ de elementos de \mathbb{k} tales que la serie $\sum_{i=1}^{\infty} |x_i|^2$ converge a un número finito. Esto es un subconjunto del espacio vectorial complejo de todas las sucesiones $(x_i)_{i \geq 1}$ y, de hecho, se trata de un subespacio: es inmediato que si $\xi \in \ell_2$ y $\lambda \in \mathbb{C}$ se tiene que $\lambda\xi \in \ell_2$, así que bastará que mostremos que ℓ_2 es cerrado para la suma.

Sean $\xi = (x_i)_{i \geq 1}$ y $\zeta = (y_i)_{i \geq 1}$ dos elementos de ℓ_2 y consideremos los números $a = \sum_{i=1}^{\infty} |x_i|^2$ y $b = \sum_{i=1}^{\infty} |y_i|^2$. Si $N \in \mathbb{N}$, entonces los vectores $x = (x_1, \dots, x_N)$ e $y = (y_1, \dots, y_N)$ pertenecen a \mathbb{k}^N y la desigualdad triangular de la norma asociada al producto interno estándar de \mathbb{k}^N implica que

$$\begin{aligned} \sum_{i=1}^N |x_i + y_i|^2 &= \|x + y\|^2 \leq (\|x\| + \|y\|)^2 = \left(\sqrt{\sum_{i=1}^N |x_i|^2} + \sqrt{\sum_{i=1}^N |y_i|^2} \right)^2 \\ &\leq \left(\sqrt{\sum_{i=1}^{\infty} |x_i|^2} + \sqrt{\sum_{i=1}^{\infty} |y_i|^2} \right)^2 = (\sqrt{a} + \sqrt{b})^2. \end{aligned}$$

Esto nos dice que las sumas parciales de la serie $\sum_{i=1}^{\infty} |x_i + y_i|^2$, que tiene términos no negativos, están acotadas y, por lo tanto, que esta serie converge: vemos así que $\xi + \zeta$ es un elemento de ℓ_2 , como queríamos.

Mostremos ahora que

$$\text{si } \xi = (x_i)_{i \geq 1}, \zeta = (y_i)_{i \geq 1} \in \ell_2, \text{ entonces la serie } \sum_{i=1}^{\infty} x_i \bar{y}_i \text{ converge absolutamente.} \quad (11)$$

Sean para ello $\xi = (x_i)_{i \geq 1}$ y $\zeta = (y_i)_{i \geq 1}$ dos elementos de ℓ_2 . De manera similar a lo que hicimos recién, consideremos un entero $N \in \mathbb{N}$ y los vectores $x = (|x_1|, \dots, |x_N|)$ e $y = (|y_1|, \dots, |y_N|)$ de \mathbb{k}^N . La desigualdad de Cauchy–Bunyakovsky–Schwartz para el producto interno estándar de \mathbb{k}^N nos dice que

$$\sum_{i=1}^N |x_i \bar{y}_i| = |\langle x, y \rangle_{\mathbb{k}^N}| \leq \|x\|_{\mathbb{k}^N} \cdot \|y\|_{\mathbb{k}^N} = \left(\sum_{i=1}^N |x_i|^2 \right)^{1/2} \left(\sum_{i=1}^N |y_i|^2 \right)^{1/2} \leq \left(\sum_{i=1}^{\infty} |x_i|^2 \right)^{1/2} \left(\sum_{i=1}^{\infty} |y_i|^2 \right)^{1/2}$$

y por lo tanto las sumas parciales de la serie $\sum_{i=1}^{\infty} |x_i \bar{y}_i|$ están acotadas. Esto implica que esa serie converge y que la serie $\sum_{i=1}^{\infty} x_i \bar{y}_i$ converge absolutamente, como afirma (11).

Ahora, en vista de (11), hay una función $\langle -, - \rangle : \ell_2 \times \ell_2 \rightarrow \mathbb{k}$ tal que

$$\langle \xi, \zeta \rangle = \sum_{i=1}^{\infty} x_i \bar{y}_i$$

cada vez que $\xi = (x_i)_{i \geq 1}$ y $\zeta = (y_i)_{i \geq 1}$ están en ℓ_2 , y es fácil verificar que se trata de un producto interno sobre ℓ_2 .

Sea S el subconjunto de ℓ_2 de las sucesiones $\xi = (x_i)_{i \geq 1}$ que tienen un número finito de componentes no nulas; se trata, de hecho, de un subespacio. Afirmamos que $S^\perp = 0$. En efecto,

supongamos que $\zeta = (y_i)_{i \geq 1}$ es un elemento de S^\perp : para cada $j \in \mathbb{N}$, la sucesión $e = (e_i)_{i \geq 1}$ tal que $e_j = 1$ y $e_i = 0$ si $i \neq j$ es claramente un elemento de S y se tiene que $y_i = \langle \zeta, e \rangle = 0$.

Por supuesto, de esto se sigue que $S^{\perp\perp} = \ell_2$ y, en consecuencia, que $S \subsetneq S^{\perp\perp}$. \diamond

§5. Proyectores ortogonales

7.5.1. Si V es un espacio vectorial, decimos que un endomorfismo $p : V \rightarrow V$ de V es un *proyector* si $p^2 = p$.

Lema. *Sea V un espacio vectorial. Si $p : V \rightarrow V$ es un proyector, entonces*

- (i) *un vector $x \in V$ está en la imagen de p si y solamente si $x = p(x)$, y*
- (ii) *hay una descomposición $V = \text{Im}(p) \oplus \text{Nu}(p)$.*

Demostración. Sea $p : V \rightarrow V$ un proyector.

- (i) Si x está en la imagen de p , de manera que existe $y \in V$ tal que $p(y) = x$, entonces $p(x) = p^2(x) = p(y) = x$. Recíprocamente, es evidente que si $p(x) = x$ entonces x está en $\text{Im}(p)$.
- (ii) Si $x \in V$, entonces $p(x - p(x)) = p(x) - p^2(x) = 0$, así que $x - p(x) \in \text{Nu}(p)$ y

$$x = p(x) + (x - p(x)) \in \text{Im}(p) + \text{Nu}(p).$$

Vemos así que $V = \text{Im}(p) + \text{Nu}(p)$. Esta suma es directa: si $x \in \text{Im}(p) \cap \text{Nu}(p)$, entonces como $x \in \text{Im}(p)$ la primera parte nos dice que $x = p(x)$ y, como $x \in \text{Nu}(p)$, esto implica que $x = 0$. \square

7.5.2. Proposición. *Sea V un espacio vectorial y sea S un subespacio de V .*

- (i) *Si T es un complemento de S en V , entonces hay un único proyector $p : V \rightarrow V$ tal que $\text{Im}(p) = S$ y $\text{Nu}(p) = T$.*
- (ii) *Hay un proyector $p : V \rightarrow V$ que tiene a S por imagen si y solamente si S posee un complemento T en V .*

Demostración. (i) Supongamos que T es un complemento de S en V , de manera que $V = S \oplus T$. Hay una función $p : V \rightarrow V$ tal que si $x \in V$ y $s \in S$ y $t \in T$ son tales que $x = s + t$ entonces $p(x) = s$: esto es consecuencia de que la suma es directa, con lo que esos vectores s y t están únicamente determinados por x . Mostremos que p es un proyector de V que satisface las condiciones del enunciado:

- Sean x e y vectores de V y $a, b \in \mathbb{k}$. Si $s_1, s_2 \in S$ y $t_1, t_2 \in T$ son tales que $x = s_1 + t_1$ e $y = s_2 + t_2$, entonces $p(x) = s_1$, $p(y) = s_2$ y, como

$$ax + by = (as_1 + bs_2) + (at_1 + bt_2)$$

con $as_1 + bs_2 \in S$ y $at_1 + bt_s \in T$,

$$p(ax + by) = as_1 + bs_2 = ap(x) + bp(y).$$

Esto nos dice que p es una función lineal.

- Si $x \in V$ y los vectores $s \in S$ y $t \in T$ son tales que $x = s + t$, entonces $p(x) = s \in S$: esto nos dice que $\text{Im}(p) \subseteq S$. Por otro lado, si $s \in S$, entonces es claro que $p(s) = s$ y, por lo tanto, que $S \subseteq \text{Im}(p)$. Vemos así que $\text{Im}(p) = S$ y que $p^2 = p$, de manera que p es un proyector.
- Por otro lado, si $x \in V$ es un elemento de $\text{Nu}(p)$ y $s \in S$ y $t \in T$ son tales que $x = s + t$, entonces $0 = p(x) = s$: esto implica que $x = t \in T$ y, en consecuencia, que $\text{Nu}(p) \subseteq T$. Recíprocamente, si $t \in T$ entonces de la definición de p se siguen inmediatamente que $p(t) = 0$. En definitiva, tenemos que $\text{Nu}(p) = T$.

Supongamos ahora que $q : V \rightarrow V$ es otro proyector tal que $\text{Im}(q) = S$ y $\text{Nu}(q) = T$. Sea $x \in V$ y sean $s \in S$ y $t \in T$ tales que $x = s + t$. Como $s \in \text{Im}(q)$, de la Proposición 7.5.1(i) sabemos que $s = q(s)$; por otro lado, como $t \in \text{Nu}(q)$, tenemos que $q(t) = 0$. Se sigue de esto, entonces, que $q(x) = q(s) + q(t) = s = p(s)$. Así, es $q = p$ y esto prueba la unicidad que afirma la proposición.

(ii) Si hay un proyector $p : V \rightarrow V$ que tiene a S por imagen, entonces de la Proposición 7.5.1(ii) se sigue que el núcleo $\text{Nu}(p)$ es un complemento para S . Recíprocamente, la parte (i) que acabamos de probar nos dice que si S posee un complemento, entonces existe un proyector $p : V \rightarrow V$ que tiene a S por imagen. \square

7.5.3. Sea V un espacio vectorial con producto interno. Si $p : V \rightarrow V$ es un proyector tal que $\text{Im}(p) \perp \text{Nu}(p)$, entonces decimos que p es un **proyector ortogonal**.

Proposición. *Sea V un espacio vectorial con producto interno.*

- (i) *Si $p : V \rightarrow V$ es un proyector ortogonal, entonces $\text{Im}(p)^\perp = \text{Nu}(p)$ y $\text{Nu}(p)^\perp = \text{Im}(p)$.*
- (ii) *Si S es un subespacio de V y existe un proyector ortogonal $p : V \rightarrow V$ cuya imagen es S , entonces $S^{\perp\perp} = S$ y $V = S \oplus S^\perp$.*

Demostración. (i) Como p es un proyector ortogonal, tenemos que $\text{Im}(p) \perp \text{Nu}(p)$ y entonces $\text{Im}(p) \subseteq \text{Nu}(p)^\perp$ y $\text{Nu}(p) \subseteq \text{Im}(p)^\perp$. Veamos que valen las igualdades.

Como p es un proyector, sabemos que $V = \text{Im}(p) \oplus \text{Nu}(p)$. Sea $x \in \text{Nu}(p)^\perp$ y sean $y \in \text{Im}(p)$ y $z \in \text{Nu}(p)$ tales que $x = y + z$. Entonces

$$0 = \langle x, z \rangle = \langle y + z, z \rangle = \langle y, z \rangle + \langle z, z \rangle$$

y el primer sumando se anula porque $\text{Im}(p) \perp \text{Nu}(p)$. Así, es $\langle z, z \rangle = 0$, de manera que $z = 0$ y $x = y \in \text{Im}(p)$. Concluimos de esta forma que $\text{Nu}(p)^\perp \subseteq \text{Im}(p)$. De manera enteramente similar puede verse que $\text{Im}(p)^\perp \subseteq \text{Nu}(p)$.

(ii) Si $p : V \rightarrow V$ es un proyector ortogonal con $S = \text{Im}(p)$, la primera parte de la proposición nos dice que $S^\perp = \text{Im}(p)^\perp = \text{Nu}(p)$ y que entonces $S^{\perp\perp} = \text{Nu}(p)^\perp = \text{Im}(p) = S$. Por otro lado, como p es un proyector tenemos que $V = \text{Im}(p) \oplus \text{Nu}(p)$, y todo esto prueba lo que queremos. \square

7.5.4. En vista de la Proposición 7.5.3(ii) y el Ejemplo 7.4.6, no todo subespacio de un espacio vectorial con producto interno es la imagen de un proyector ortogonal. La siguiente proposición nos dice que eso puede ocurrir solamente si el subespacio tiene dimensión infinita:

Proposición. *Sea V un espacio vectorial con producto interno. Si S es un subespacio de V de dimensión finita, entonces $V = S \oplus S^\perp$ y existe exactamente un proyector ortogonal $p : V \rightarrow V$ tal que $\text{Im}(p) = S$. Más aún, vale que $\text{Nu}(p) = S^\perp$ y, si $\mathcal{B} = \{x_1, \dots, x_n\}$ es una base ortonormal de S con n elementos, que para cada $x \in V$ se tiene que*

$$p(x) = \sum_{i=1}^n \langle x, x_i \rangle x_i.$$

Demostración. Sea S un subespacio de V de dimensión finita n y sea $\mathcal{B} = \{x_1, \dots, x_n\}$ una base de S . La función $p : V \rightarrow V$ tal que $p(x) = \sum_{i=1}^n \langle x, x_i \rangle x_i$ para todo $x \in V$ es claramente linear y, de acuerdo a la Proposición 7.3.12, se tiene que $x - p(x) \in S^\perp$ para todo $x \in V$.

Si $x \in \text{Nu}(p)$, entonces $x = x - p(x) \in S^\perp$ y, reciprocamente, si $x \in S^\perp$, de manera que en particular $\langle x, x_i \rangle = 0$ para cada $i \in \llbracket n \rrbracket$, entonces $p(x) = \sum_{i=1}^n \langle x, x_i \rangle x_i = 0$. Así, es $\text{Nu}(p) = S^\perp$. De la definición de p es claro que $\text{Im}(p) \subseteq S$. Por otro lado, si $x \in S$, entonces el Corolario 7.3.6 nos dice que $x = \sum_{i=1}^n \langle x, x_i \rangle x_i$ porque \mathcal{B} es una base ortonormal de S y, como consecuencia de esto, que $x = p(x) \in \text{Im}(p)$. Luego $\text{Im}(p) = S$ y $p^2 = p$, como queríamos. \square

7.5.5. Hay una relación muy estrecha entre los proyectores ortogonales de un espacio vectorial con producto interno y sus subespacios. Un ejemplo de esto es el siguiente resultado:

Proposición. *Sea V un espacio vectorial con producto interno y sean $p, q : V \rightarrow V$ dos proyectores ortogonales.*

- (i) *Es $p \circ q = 0$ si y solamente si $\text{Im}(p) \perp \text{Im}(q)$.*
- (ii) *Es $p \circ q = p$ si y solamente si $q \circ p = p$ y esto ocurre si y solamente si $\text{Im}(p) \subseteq \text{Im}(q)$.*
- (iii) *Es $p \circ q = p \circ p$ si y solamente $p \circ q$ es un proyector ortogonal, y en ese caso se tiene que $\text{Im}(p \circ q) = \text{Im}(p) \cap \text{Im}(q)$.*

Demostración. Observemos primero que

$$\text{si } p : V \rightarrow V \text{ es un proyector ortogonal y } x, y \in V, \text{ entonces } \langle p(x), y \rangle = \langle x, p(y) \rangle. \quad (12)$$

En efecto, si $x = s + t$ e $y = s' + t'$ con $s, s' \in \text{Im}(p)$ y $t, t' \in \text{Nu}(p)$, entonces

$$\langle p(x), y \rangle = \langle s, s' + t' \rangle = \langle s, s' \rangle = \langle s + t, s' \rangle = \langle x, p(y) \rangle.$$

(i) Supongamos que $p \circ q = 0$. Si $x \in \text{Im}(p)$ e $y \in \text{Im}(q)$, entonces $q(y) = y$ y

$$p(y) = p(q(y)) = 0,$$

de manera que $y \in \text{Nu}(p) = \text{Im}(p)^\perp$: esto implica que $x \perp y$ y muestra que $\text{Im}(p) \perp \text{Im}(q)$.

Para probar la implicación recíproca, supongamos que $\text{Im}(p) \perp \text{Im}(q)$ y sea $x \in V$. Como $q(x) \in \text{Im}(q) \subseteq \text{Im}(p)^\perp = \text{Nu}(p)$, tenemos que $p(q(x)) = 0$. Vemos así que $p \circ q = 0$.

(ii) Probemos la equivalencia de las tres condiciones del enunciado.

- Supongamos que $p \circ q = p$. Si $x \in V$, entonces usando (12) vemos que para todo $y \in V$ vale que

$$\langle q(p(x)), y \rangle = \langle p(x), q(y) \rangle = \langle x, p(q(y)) \rangle = \langle x, p(y) \rangle = \langle p(x), y \rangle$$

y, por lo tanto, que $q(p(x)) = p(x)$: esto nos dice que $q \circ p = p$.

- De manera similar, si $q \circ p = p$, entonces para cada $x \in V$ se tiene que

$$\langle p(q(x)), y \rangle = \langle q(x), p(y) \rangle = \langle x, q(p(y)) \rangle = \langle x, p(y) \rangle = \langle p(x), y \rangle,$$

así que $p(q(x)) = p(x)$ y, en definitiva, $p \circ q = p$.

- Si $q \circ p = p$ entonces claramente $\text{Im}(p) \subseteq \text{Im}(q)$.
- Recíprocamente, si $\text{Im}(p) \subseteq \text{Im}(q)$, entonces para cada $x \in V$ es $p(x) \in \text{Im}(q)$, así que $q(p(x)) = p(x)$: esto significa quer $q \circ p = p$.

(iii) Supongamos primero que $p \circ q = q \circ p$. Tenemos entonces que

$$(p \circ q)^2 = p \circ q \circ p \circ p = p \circ p \circ q \circ q = p \circ q,$$

así que $r = p \circ q$ es un proyector. Como $\text{Im}(r) = \text{Im}(p \circ q) \subseteq \text{Im}(p)$ e $\text{Im}(r) = \text{Im}(q \circ p) \subseteq \text{Im}(q)$, es claro que $\text{Im}(r) \subseteq \text{Im}(p) \cap \text{Im}(q)$. Recíprocamente, si $x \in \text{Im}(p) \cap \text{Im}(q)$, entonces sabemos que $p(x) = x$ y que $q(x) = x$, así que $x = p(x) = p(q(x)) = r(x) \in \text{Im}(r)$. Esto prueba que $\text{Im}(r) = \text{Im}(p) \cap \text{Im}(q)$.

Supongamos ahora que la composición $r = p \circ q$ es un proyector ortogonal. Si $x \in V$, entonces para cada $y \in V$ vale que

$$\langle q(p(x)), y \rangle = \langle p(x), q(y) \rangle = \langle x, p(q(y)) \rangle = \langle x, r(y) \rangle = \langle r(x), y \rangle = \langle p(q(x)), y \rangle$$

y esto nos dice que $q \circ p = p \circ q$. Esto muestra que la condición del enunciado es suficiente. \square

§6. El teorema de representación de Riesz

7.6.1. Lema. Sea V un espacio vectorial con producto interno y para cada $y \in V$ sea la función

$$\phi_y : x \in V \mapsto \langle x, y \rangle \in \mathbb{k}.$$

- (i) Para cada $y \in V$, la función ϕ_y es lineal, de manera que $\phi_y \in V^*$.
- (ii) La función $\Phi : y \in V \mapsto \phi_y \in V^*$ es sesquilineal, de manera que para cada $y, y' \in V$ y cada

$a \in \mathbb{k}$ se tiene que

$$\phi_{y+y'} = \phi_y + \phi_{y'}, \quad \phi_{day} = \bar{a}\phi_y,$$

e inyectiva.

Demostración. (i) Sea $y \in V$. Si $x, x' \in V$ y $a, b \in \mathbb{k}$, entonces

$$\phi_y(ax + bx') = \langle ax + bx', y \rangle = a\langle x, y \rangle + b\langle x', y \rangle = a\phi_y(x) + b\phi_y(x'),$$

así que la función ϕ_y es lineal.

(ii) Sean $y, y' \in V$. Para cada $y \in V$, es

$$\phi_{y+y'}(x) = \langle x, y + y' \rangle = \langle x, y \rangle + \langle x, y' \rangle = \phi_y(x) + \phi_{y'}(x) = (\phi_y + \phi_{y'})(x),$$

así que $\phi_{y+y'} = \phi_y + \phi_{y'}$. De manera similar, si $y \in V$ y $a \in \mathbb{k}$, para cada $x \in V$ se tiene que

$$\phi_{ay}(x) = \langle y, ay \rangle = \bar{a}\langle x, y \rangle = \bar{a}\phi_y(x) = (\bar{a}\phi_y)(x),$$

de manera que $\phi_{ay} = \bar{a}\phi_y$. Esto prueba que la función Φ es sesquilineal.

Finalmente, si $y, y' \in V$ son tales que $\Phi(y) = \Phi(y')$, entonces para cada $x \in V$ se tiene que

$$\langle x, y - y' \rangle = \langle x, y \rangle - \langle x, y' \rangle = \Phi(y)(x) - \Phi(y')(x) = 0,$$

y entonces $y - y' = 0$, esto es, $y = y'$. Vemos así que la función Φ es inyectiva. \square

7.6.2. El siguiente resultado es el *teorema de representación de Riesz*, por Frigyes Riesz (1880–1956, Hungría), quien probó un resultado más general para espacios de Hilbert:

Teorema. *Sea V un espacio vectorial con producto interno y de dimensión finita. Si $f \in V^*$, entonces existe un único vector $x_f \in V$ tal que $\phi_{x_f} = f$, esto es, tal que para todo $x \in V$ se tiene que*

$$f(x) = \langle x, x_f \rangle. \tag{13}$$

Demostración. Sea n la dimensión de V , sea $\mathcal{B} = \{x_1, \dots, x_n\}$ una base ortonormal de V y sea

$$x_f = \sum_{i=1}^n \overline{f(x_i)} x_i.$$

Si $x \in V$, entonces $x = \sum_{i=1}^n \langle x, x_i \rangle x_i$ y

$$f(x) = f\left(\sum_{i=1}^n \langle x, x_i \rangle x_i\right) = \sum_{i=1}^n \langle x, x_i \rangle f(x_i) = \left\langle x, \sum_{i=1}^n \overline{f(x_i)} x_i \right\rangle = \langle x, x_f \rangle.$$

Esto nos dice que x_f satisface la condición (13).

Por otro lado, si $x'_f \in V$ es otro vector que la satisface, entonces para todo $x \in V$ se tiene que

$$\langle x, x_f - x'_f \rangle = \langle x, x_f \rangle - \langle x, x'_f \rangle = f(x) - f(x) = 0.$$

Esto es solo posible si $x_f = x'_f$. □

7.6.3. Corolario. La función $\Phi : V \rightarrow V^*$ del Lema 7.6.1(ii) es una biyección.

Demostración. En efecto, el Teorema 7.6.2 nos dice que es sobreyectiva y ese lema nos dice que es inyectiva. □

7.6.4. Ejemplo. En general, si V es un espacio vectorial con producto interno y $f : V \rightarrow \mathbb{k}$ es lineal, no existe un vector $y \in V$ tal que $f(x) = \langle x, y \rangle$ para todo $x \in V$. Sea, por ejemplo, \mathbb{k}^∞ el espacio vectorial de las sucesiones $(x_i)_{i \geq 1}$ de elementos de \mathbb{k} con casi todas las componentes nulas, dotado del producto interno tal que si $x = (x_i)_{i \geq 1}$ e $y = (y_i)_{i \geq 1}$ es

$$\langle x, y \rangle = \sum_{i \geq 1} x_i \bar{y}_i.$$

Notemos que esto tiene sentido, ya que cualesquiera sean x e y en V la suma es finita. Para cada $n \geq 1$ sea e_n el elemento de V que tiene todas sus componentes nulas salvo la n -ésima, que es igual a 1, y consideremos la función lineal $f : V \rightarrow \mathbb{k}$ tal que $f(x) = \sum_{i \geq 1} x_i$ si $x = (x_i)_{i \geq 1} \in V$. Si hubiera un vector $y = (y_i)_{i \geq 1} \in V$ tal que $f(x) = \langle x, y \rangle$ para todo $x \in V$, tendríamos en particular que para todo $n \geq 1$ sería $\bar{y}_n = \langle e_n, y \rangle = f(e_n) = 1$, lo que es absurdo. ◇

§7. Funciones adjuntas

7.7.1. Sean V y W espacios vectoriales con producto interno y sea $f : V \rightarrow W$ una función lineal. Decimos que una función lineal $g : W \rightarrow V$ es **adjunta** a f si para todo $x \in V$ y todo $y \in W$ se tiene que

$$\langle f(x), y \rangle_W = \langle x, g(y) \rangle_V.$$

Lema. Sean V y W espacios vectoriales con producto interno. Una función lineal $f : V \rightarrow W$ posee a lo sumo una función lineal adjunta.

En vista de este lema, si una función lineal $f : V \rightarrow W$ tiene una función adjunta, tiene una sola: podemos en ese caso denotarla sin ambigüedad f^* .

Demostración. Supongamos que $g, g' : W \rightarrow V$ son funciones lineales adjuntas a f . Si $y \in W$, entonces para todo $x \in V$ se tiene que

$$\langle x, g(y) - g'(y) \rangle_V = \langle x, g(y) \rangle_V - \langle x, g'(y) \rangle_V = \langle f(x), y \rangle_W - \langle f(x), y \rangle_W = 0,$$

así que $g(y) = g'(y)$. Esto nos dice que $g = g'$. □

7.7.2. Ejemplos.

- (a) Si V es un espacio vectorial con producto interno y $\lambda \in \mathbb{k}$ es un escalar, entonces la función lineal $\lambda \text{id}_V : V \rightarrow V$ posee una adjunta, que es $\bar{\lambda} \text{id}_V : V \rightarrow V$.
- (b) Sea $V = C_c^\infty(\mathbb{R})$ el espacio vectorial de todas las funciones infinitamente diferenciables $\mathbb{R} \rightarrow \mathbb{k}$ que tienen soporte compacto, dotado del producto interno $\langle -, - \rangle$ tal que

$$\langle f, g \rangle = \int_{-\infty}^{\infty} f(x) \overline{g(x)} dx$$

para cada $f, g \in V$; notemos que esta integral tiene sentido, porque el soporte de $f\bar{g}$ es compacto y la función es allí continua. Si $f \in V$, entonces también $f' \in V$, así que podemos considerar la función lineal

$$D : f \in V \mapsto f' \in V.$$

Si $f, g \in V$, entonces la fórmula de integración por partes nos dice que

$$\langle Df, g \rangle = \int_{-\infty}^{\infty} f'(x) \overline{g(x)} dx = - \int_{-\infty}^{\infty} f(x) \overline{g'(x)} dx = \langle f, -Dg \rangle,$$

y esto significa que D posee una función adjunta y que $D^* = -D$. \diamond

7.7.3. Ejemplo. En general, un endomorfismo de un espacio vectorial con producto interno no tiene una función adjunta. Para dar un ejemplo, sea $V = \mathbb{k}^\infty$ el espacio vectorial de las sucesiones $(x_i)_{i \geq 1}$ de elementos de \mathbb{k} con casi todas las componentes nulas, con el producto interno tal que si $x = (x_i)_{i \geq 1}$ e $y = (y_i)_{i \geq 1}$ es

$$\langle x, y \rangle = \sum_{i \geq 1} x_i \overline{y_i}.$$

Notemos que esto tiene sentido, ya que cualesquiera sean x e y en V la suma es finita. Para cada $n \in \mathbb{N}$ escribimos e_n al elemento de V que tiene todas sus componentes nulas salvo la n -ésima, que es igual a 1, y consideramos la función lineal $L : V \rightarrow \mathbb{k}$ tal que si $x = (x_i)_{i \geq 1}$ es un elemento de V entonces

$$L(x) = \sum_{i \geq 1} x_i.$$

Supongamos que L posee una función adjunta $L^* : \mathbb{k} \rightarrow V$, de manera que, en particular, si escribimos $L^*(1) = (u_i)_{i \geq 1}$, es

$$1 = \langle L(e_n), 1 \rangle = \langle e_n, L^*(1) \rangle = \overline{u_n}$$

para cada $i \geq 1$: esto es absurdo, porque nos dice que *todas* las componentes de $L^*(1) \in \mathbb{k}^\infty$ son no nulas. Este ejemplo está en evidente relación con el Ejemplo 7.6.4 visto arriba. \diamond

7.7.4. Proposición. Sean V y W dos espacios vectoriales con producto interno.

- (i) Si $f, g : V \rightarrow W$ son funciones lineales que poseen funciones adjuntas, entonces para cada $a, b \in \mathbb{k}$ la función lineal $af + bg : V \rightarrow W$ posee función adjunta y, de hecho, se tiene que

$$(af + bg)^* = \bar{a} f^* + \bar{b} g^*.$$

- (ii) Si $f : V \rightarrow W$ es una función lineal que posee una función adjunta, entonces f^* también posee una función adjunta y

$$(f^*)^* = f.$$

Demostración. (i) Si $x \in V$ e $y \in W$, entonces

$$\begin{aligned} \langle (af + bg)(x), y \rangle &= \langle af(x) + bg(x), y \rangle = a\langle f(x), y \rangle + b\langle g(x), y \rangle \\ &= a\langle x, f^*(y) \rangle + b\langle x, g^*(y) \rangle = \langle x, \bar{a} f^*(y) + \bar{b} g^*(y) \rangle \\ &= \langle x, (\bar{a} f^* + \bar{b} g^*)(y) \rangle. \end{aligned}$$

Esto significa que la función lineal $\bar{a} f^* + \bar{b} g^* : V \rightarrow W$ es adjunta a $af + bg$.

- (ii) Si $x \in V$ e $y \in W$, tenemos que

$$\langle f^*(x), y \rangle = \overline{\langle y, f^*(x) \rangle} = \overline{\langle f(y), x \rangle} = \langle x, f(y) \rangle,$$

y esto nos dice que $f : V \rightarrow W$ es adjunta a $f^* : W \rightarrow V$, esto es, que $(f^*)^* = f$. \square

7.7.5. Proposición.

- (i) Si V es un espacio con producto interno, entonces la función identidad $\text{id}_V : V \rightarrow V$ posee función adjunta y

$$(\text{id}_V)^* = \text{id}_V.$$

- (ii) Si V, W y U son espacios vectoriales con producto interno y $f : V \rightarrow W$ y $g : W \rightarrow U$ son funciones lineales que poseen funciones adjuntas, entonces la composición $g \circ f : V \rightarrow U$ posee función adjunta y

$$(g \circ f)^* = f^* \circ g^*.$$

Demostración. Para ver la primera parte, es suficiente notar que si x e y son elementos de V se tiene que

$$\langle \text{id}_V(x), y \rangle = \langle x, y \rangle = \langle x, \text{id}_V(y) \rangle,$$

así que $(\text{id}_V)^* = \text{id}_V$. Por otro lado, si $f : V \rightarrow W$ y $g : W \rightarrow U$ son funciones lineales que poseen funciones adjuntas, entonces para cada $x \in V$ y cada $y \in U$ se tiene que

$$\langle (g \circ f)(x), y \rangle = \langle g(f(x)), y \rangle = \langle f(x), g^*(y) \rangle = \langle x, f^*(g^*(y)) \rangle = \langle x, (f^* \circ g^*)(y) \rangle.$$

Esto significa que $g \circ f$ tiene como función adjunta a $f^* \circ g^*$. \square

7.7.6. Teorema. *Sean V y W espacios vectoriales con producto interno. Si V tiene dimensión finita, entonces Toda función lineal $f : V \rightarrow W$ posee una función adjunta $f^* : W \rightarrow V$.*

Demostración. Si $y \in V$, la función

$$\psi_y : x \in V \mapsto \langle f(x), y \rangle_W \in \mathbb{k}$$

es lineal, así que el Teorema 7.6.2 nos dice que existe un único un vector $f^*(y) \in V$ con la propiedad de que para todo $x \in V$ es $\psi_y(x) = \langle x, f^*(y) \rangle_V$, esto es, para todo $y \in V$ es

$$\langle f(x), y \rangle_V = \langle x, f^*(y) \rangle_W.$$

De esta forma tenemos definida una función $f^* : W \rightarrow V$ que satisface la identidad deseada. Para terminar la prueba del teorema resta únicamente mostrar que f^* es lineal.

Sean $y, y' \in W$. Para todo $x \in V$ se tiene que

$$\begin{aligned} \langle x, f^*(y + y') \rangle_V &= \langle f(x), y + y' \rangle_W = \langle f(x), y \rangle_W + \langle f(x), y' \rangle_W \\ &= \langle x, f^*(y) \rangle_W + \langle x, f^*(y') \rangle_W = \langle x, f^*(y) + f^*(y') \rangle_W, \end{aligned}$$

así que $f^*(y + y') = f^*(y) + f^*(y')$. Por otro lado, si $y \in W$ y $\lambda \in \mathbb{k}$, para cada $x \in W$ tenemos que

$$\begin{aligned} \langle x, f^*(\lambda y) \rangle_V &= \langle f(x), \lambda y \rangle_W = \langle \bar{\lambda} f(x), y \rangle_W = \langle f(\bar{\lambda} x), y \rangle_W = \langle \bar{\lambda} x, f^*(y) \rangle_V \\ &= \langle x, \lambda f^*(y) \rangle_V, \end{aligned}$$

de manera que $f^*(\lambda w) = \lambda f^*(w)$. Concluimos que f^* es lineal, como queríamos. \square

7.7.7. Proposición. *Sean V y W espacios vectoriales con producto interno de dimensiones finita y positivas m y n y sean \mathcal{B} y \mathcal{B}' bases ordenadas ortonormales de V y de W , respectivamente. Si $f : V \rightarrow W$ es una función lineal, entonces la matriz de la función adjunta $f^* : W \rightarrow V$ con respecto a las bases \mathcal{B}' y \mathcal{B} es*

$$[f^*]_{\mathcal{B}}^{\mathcal{B}'} = \overline{([f]_{\mathcal{B}'}^{\mathcal{B}})^t}.$$

Demostración. Supongamos que las bases de V y de W son $\mathcal{B} = (x_1, \dots, x_m)$ y $\mathcal{B}' = (y_1, \dots, y_n)$, y que $[f]_{\mathcal{B}'}^{\mathcal{B}} = (a_{i,j}) \in M_{n,m}(\mathbb{k})$. Si $i \in [m]$ y $j \in [n]$, es

$$\overline{\langle f^*(y_j), x_i \rangle} = \langle x_i, f^*(y_j) \rangle = \langle f(x_i), y_j \rangle = \left\langle \sum_{k=1}^m a_{k,i} y_k, y_j \right\rangle = \sum_{k=1}^m a_{k,i} \langle y_k, y_j \rangle = a_{j,i}$$

porque la base \mathcal{B}' es ortonormal y entonces $\langle f^*(y_j), x_i \rangle = \overline{a_{j,i}}$. De acuerdo al Corolario 7.3.6, se sigue de esto que

$$f^*(y_j) = \sum_{i=1}^m \overline{a_{j,i}} x_i$$

y entonces la matriz de f^* con respecto a las bases \mathcal{B}' y \mathcal{B} es $(\overline{a_{j,i}})_{i,j}$, esto es, la matriz que se obtiene de $[f]_{\mathcal{B}'}$ transponiendo y conjugando, como afirma la proposición. \square

7.7.8. Ejemplo. Es importante observar que la igualdad que afirma la Proposición 7.7.7 no es cierta, en general, si las bases consideradas no son ortonormales. Demos un ejemplo de esto.

Consideremos el espacio vectorial \mathbb{k}^2 dotado de su producto interno estándar y la matriz $A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \in M_2(\mathbb{k})$, y sea $f : x \in \mathbb{k}^2 \mapsto Ax \in \mathbb{k}^2$. Si $\mathcal{B} = (e_1, e_2)$ es la base ordenada estándar de \mathbb{k}^2 , entonces sabemos que $[f]_{\mathcal{B}} = A$ y, como \mathcal{B} es de hecho una base ortonormal, que $[f^*]_{\mathcal{B}} = A^t$, la matriz transpuesta de A . En cambio, si ponemos $e'_2 = e_1 + e_2$, entonces $\mathcal{B}' = (e_1, e'_2)$ también es una base ordenada de \mathbb{k}^2 , se tiene que $[f]_{\mathcal{B}'} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ mientras que $[f^*]_{\mathcal{B}'} = \begin{pmatrix} -1 & 0 \\ 1 & 0 \end{pmatrix}$. \diamond

7.7.9. Proposición. Sean V y W espacios vectoriales con producto interno y sea $f : V \rightarrow W$ una función lineal que posee adjunta. Entonces

- (i) $\text{Nu}(f^*) = \text{Im}(f)^\perp$,
- (ii) $\text{Im}(f^*) \subseteq \text{Nu}(f)^\perp$, y
- (iii) $\text{Im}(f^*)^\perp \subseteq \text{Nu}(f^*)$.

Si V tiene dimensión finita, entonces se tiene además que

- (iv) $\text{Im}(f^*) = \text{Nu}(f)^\perp$.

Demostración. (i) Sea $x \in \text{Nu}(f^*)$. Si $y \in \text{Im}(f)$, de manera que existe $z \in V$ tal que $f(z) = y$, entonces $\langle y, x \rangle = \langle f(z), x \rangle = \langle z, f^*(x) \rangle = 0$. Esto nos dice que $x \in \text{Im}(f)^\perp$. Recíprocamente, si $x \in \text{Im}(f)^\perp$, se tiene que para todo $y \in V$ es $\langle y, f^*(x) \rangle = \langle f(y), x \rangle = 0$, así que $f^*(x) = 0$, esto es, $x \in \text{Nu}(f^*)$.

(ii) Sea $x \in \text{Im}(f^*)$, de manera que existe $y \in V$ tal que $x = f^*(y)$. Para cada $z \in \text{Nu}(f)$ es $\langle z, x \rangle = \langle z, f^*(y) \rangle = \langle f(z), y \rangle = 0$, así que $x \in \text{Nu}(f)^\perp$: hemos probado que $\text{Im}(f^*) \subseteq \text{Nu}(f)^\perp$.

(iii) Sea $x \in \text{Im}(f^*)^\perp$. Si $y \in V$, entonces $\langle f(x), y \rangle = \langle x, f^*(y) \rangle = 0$, así que $f(x) = 0$. Vemos de esta forma que $\text{Im}(f^*)^\perp \subseteq \text{Nu}(f^*)$.

(iv) Si V tiene dimensión finita, usando la Proposición 7.4.2(iii) y la Proposición 7.4.4 y las partes (ii) y (iii) de esta proposición, tenemos que $\text{Nu}(f^*)^\perp \subseteq \text{Im}(f^*)^{\perp\perp} = \text{Im}(f^*)$. \square

7.7.10. Ejemplo. La igualdad de la parte (iv) de la Proposición 7.7.9 en general no vale. **HACER.** \diamond

§8. Funciones lineales autoadjuntas

7.8.1. Sea V un espacio vectorial con producto interno. Si $f : V \rightarrow V$ es una función lineal que posee adjunta f^* , decimos que f es **autoadjunta** si $f^* = f$. En otras palabras, la función lineal f es autoadjunta si y solamente si para cada $x, y \in V$ si tiene que

$$\langle f(x), y \rangle = \langle x, f(y) \rangle.$$

7.8.2. **Proposición.** Sea V un espacio vectorial con producto interno y de dimensión finita y sea \mathcal{B} una base ortonormal de V . Una función lineal $f : V \rightarrow V$ es autoadjunta si y solamente si

$$[f]_{\mathcal{B}}^{\mathcal{B}} = \overline{([f]_{\mathcal{B}}^{\mathcal{B}})^t}.$$

En particular,

- (i) si $\mathbb{k} = \mathbb{R}$, f es autoadjunta si y solamente si la matriz $[f]_{\mathcal{B}}^{\mathcal{B}}$ es simétrica, y
- (ii) si $\mathbb{k} = \mathbb{C}$, f es autoadjunta si y solamente si la matriz $[f]_{\mathcal{B}}^{\mathcal{B}}$ es hermitiana.

Demostración. Esto es consecuencia inmediata de la Proposición 7.7.7 y de que dos transformaciones lineales $V \rightarrow V$ son iguales si y solamente si sus matrices con respecto a la base \mathcal{B} coinciden. \square

7.8.3. **Proposición.** Sea V un espacio vectorial con producto interno y sea $p : V \rightarrow V$ un proyector. Entonces p es ortogonal si y solamente si es autoadjunto.

Demostración. Supongamos primero que el proyector p es ortogonal. Sean $x, y \in V$. Como $V = \text{Im}(p) \oplus \text{Nu}(p)$, existen $x', y' \in \text{Im}(p)$ y $x'', y'' \in \text{Nu}(p)$ tales que $x = x' + x''$ e $y = y' + y''$. Más aún, $p(x) = x'$, $p(y) = y'$ y, como por hipótesis es $\text{Im}(p) \perp \text{Nu}(p)$, es $\langle x', y'' \rangle = \langle x'', y' \rangle = 0$. Usando todo esto, vemos que

$$\begin{aligned}\langle p(x), y \rangle &= \langle x', y' + y'' \rangle = \langle x', y' \rangle + \langle x', y'' \rangle = \langle x', y' \rangle = \langle x', y' \rangle + \langle x'', y' \rangle \\ &= \langle x' + x'', y' \rangle = \langle x, p(y) \rangle.\end{aligned}$$

Esto nos dice que p posee adjunta y que, de hecho, $p^* = p$.

Supongamos ahora que p es autoadjunto. Sean $x \in \text{Im}(p)$ e $y \in \text{Nu}(p)$. Como p es autoadjunto,

$$\langle x, y \rangle = \langle p(x), y \rangle = \langle x, p(y) \rangle = 0.$$

Esto nos dice que $x \perp y$ y, en definitiva, que $\text{Im}(p) \perp \text{Nu}(p)$. \square

7.8.4. **Proposición.** Sea V un espacio vectorial con producto interno y sea $f : V \rightarrow V$ una función lineal autoadjunta.

- (i) Todo autovalor de f es real.
- (ii) Si $x, y \in V$ son autovectores de f correspondientes a autovalores distintos, entonces $x \perp y$.

Demostración. (i) Sea $\lambda \in \mathbb{k}$ un autovalor y sea $x \in V$ un autovector de f de autovalor λ . Entonces

$$\lambda \langle x, x \rangle = \langle \lambda x, x \rangle = \langle f(x), x \rangle = \langle x, f(x) \rangle = \langle x, \lambda x \rangle = \bar{\lambda} \langle x, x \rangle$$

y, como $\langle x, x \rangle \neq 0$, esto implica que $\lambda = \bar{\lambda}$, esto es, que $\lambda \in \mathbb{R}$.

(ii) Supongamos que $x, y \in V$ son autovectores de autovalores $\lambda, \mu \in \mathbb{k}$, respectivamente, y que $\lambda \neq \mu$. De la parte (i) sabemos que $\lambda, \mu \in \mathbb{R}$ y entonces

$$\lambda \langle x, y \rangle = \langle \lambda x, y \rangle = \langle f(x), y \rangle = \langle x, f(y) \rangle = \langle x, \mu y \rangle = \bar{\mu} \langle x, y \rangle = \mu \langle x, y \rangle.$$

Vemos así que $(\lambda - \mu) \langle v, w \rangle = 0$ y, como $\lambda \neq \mu$, que $\langle v, w \rangle = 0$. \square

7.8.5. Proposición. *Sea V un espacio vectorial con producto interno y de dimensión finita. Una función lineal $f : V \rightarrow V$ que es autoadjunta posee un autovalor.*

Demostración. Si $\mathbb{k} = \mathbb{C}$, ya sabemos que todo endomorfismo de un espacio vectorial de dimensión finita tiene un autovalor, ya que el cuerpo \mathbb{C} es algebraicamente cerrado. Nos queda entonces solamente considerar el caso en que $\mathbb{k} = \mathbb{R}$.

Sea $n = \dim V$, sea \mathcal{B} una base ortonormal de V y sea $A = [f]_{\mathcal{B}}^{\mathcal{B}} \in M_n(\mathbb{R})$. Como f es autoadjunta, la matriz A es simétrica. Consideremos el endomorfismo $g : x \in \mathbb{C}^n \mapsto Ax \in \mathbb{C}^n$ del espacio vectorial complejo \mathbb{C}^n y sea \mathcal{B}' la base ordenada estándar de \mathbb{C}^n . Si dotamos a \mathbb{C}^n de su producto interno estándar, la base \mathcal{B}' es ortonormal, así que, como $[g]_{\mathcal{B}'}^{\mathcal{B}'} = A$ es una matriz hermitiana (ya que es simétrica y real), la función g es autoadjunta y todos sus autovalores son reales. Esto implica que el polinomio característico χ_g de g tiene todas sus raíces reales. Como este polinomio coincide con el polinomio característico χ_A de A y éste último con el polinomio característico χ_f de f , vemos que χ_f tiene sus raíces en \mathbb{R} : esto nos dice que f posee algún autovalor. \square

7.8.6. Podemos ahora probar el resultado más importante sobre las funciones lineales autoadjuntas: son diagonalizables.

Proposición. *Sea V un espacio vectorial con producto interno y de dimensión finita. Si $f : V \rightarrow V$ es una función lineal autoadjunta, existe una base ortonormal \mathcal{B} cuyos elementos son autovectores de f y, en particular, la matriz $[f]_{\mathcal{B}}^{\mathcal{B}}$ es diagonal.*

Demostración. Hacemos inducción en $\dim V$. Es claro que cuando $\dim V = 0$ no hay nada que probar, así que supongamos que $\dim V > 0$.

Sea $f : V \rightarrow V$ una función lineal autoadjunta. De acuerdo a la proposición anterior, existe un autovalor $\lambda \in \mathbb{k}$ y entonces existe $x_1 \in V \setminus \{0\}$ tal que $f(x_1) = \lambda x_1$. Sin pérdida de generalidad, podemos suponer que $\|x_1\| = 1$; si ése no fuese el caso, podríamos simplemente reemplazar a x_1 por el vector $x_1/\|x_1\|$.

Sea $W = \langle x_1 \rangle^\perp$. Si $x \in W$, entonces

$$\langle f(x), x_1 \rangle = \langle x, f(x_1) \rangle = \langle x, \lambda x_1 \rangle = \lambda \langle x, x_1 \rangle = 0,$$

de manera que $f(x) \in W$: esto nos dice que el subespacio W es f -invariante. Dotemos a W del producto interno que se obtiene restringiendo el de V y sea $f_W : W \rightarrow W$ la restricción de f a W . Es inmediato ver que f_W es entonces un endomorfismo autoadjunto de W y, como $\dim W = \dim V - 1$, inductivamente podemos suponer que hay una base ordenada ortonormal (x_2, \dots, x_n) de W cuyos elementos son autovectores de f_W . Por supuesto, se sigue inmediatamente de esto que $\mathcal{B} = (x_1, \dots, x_n)$ es una base ordenada ortonormal de V cuyos elementos son autovectores de f , lo que prueba la proposición. \square

7.8.7. Corolario. *Sea V un espacio vectorial con producto interno y de dimensión finita y sea $f : V \rightarrow V$ una función lineal. La función f es autoadjunta si y solamente si existe una base ordenada ortonormal \mathcal{B} de V tal que la matriz $[f]_{\mathcal{B}}$ es diagonal y real.*

Demostración. La necesidad de la condición es consecuencia de la Proposición 7.8.6 y de la Proposición 7.8.4(i). La suficiente, por su parte, sigue inmediatamente de la Proposición 7.7.7. \square

§9. Funciones lineales normales

7.9.1. Sea V un espacio vectorial con producto interno. Una función lineal $f : V \rightarrow V$ es **normal** si posee adjunta f^* y

$$f^* f = f f^*.$$

Es claro que si $f : V \rightarrow V$ es autoadjunta, de manera que $f^* = f$, entonces f es normal: en ese caso es $f^* f = f^2 = f f^*$. La implicación recíproca es falsa.

7.9.2. Ejemplo. Consideremos a \mathbb{R}^2 dotado de su producto interno usual y sea \mathcal{B} la base ordenada estándar de \mathbb{R}^2 . Sea $\alpha \in \mathbb{R}$ y sea $f : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ la función lineal tal que

$$[f]_{\mathcal{B}} = \begin{pmatrix} \cos \alpha & \sin \alpha \\ -\sin \alpha & \cos \alpha \end{pmatrix}$$

Como \mathcal{B} es una base ortonormal, sabemos que

$$[f^*]_{\mathcal{B}} = \overline{([f]_{\mathcal{B}})^t} = \begin{pmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{pmatrix},$$

y un cálculo inmediato muestra que $[f^*]_{\mathcal{B}} [f]_{\mathcal{B}} = I_2 = [f]_{\mathcal{B}} [f^*]_{\mathcal{B}}$, la matriz identidad, de manera que $f^* f = f f^* = \text{id}_{\mathbb{R}^2}$. Así, f es normal cualquiera sea $\alpha \in \mathbb{R}$. Es claro, por otro lado, que $f^* \neq f$ si el número α no es un múltiplo entero de π . \diamond

7.9.3. Proposición. *Sea V un espacio vectorial con producto interno. Sea $f : V \rightarrow V$ una función lineal normal y sean $x \in V$ y $\lambda \in \mathbb{k}$. Las siguientes afirmaciones son equivalentes:*

- (a) x es un autovector para f de autovalor λ .
- (b) x es un autovector para f^* de autovalor $\bar{\lambda}$.

Demostración. Sea $g = f - \lambda \text{id}_V$. Sabemos que g posee adjunta y que $g^* = f^* - \bar{\lambda} \text{id}_V$ y, como f es normal, es inmediato verificar que g es normal. Usando esto, tenemos que

$$\|g(x)\|^2 = \langle g(x), g(x) \rangle = \langle x, g^*(g(x)) \rangle = \langle x, g(g^*(x)) \rangle = \langle g^*(x), g^*(x) \rangle = \|g^*(x)\|^2.$$

Esto nos dice que $g(v) = 0$ si y solo si $g^*(v) = 0$, es decir, que $f(v) = \lambda v$ si y solo si $f^*(v) = \bar{\lambda} v$, que es precisamente lo que afirma la proposición. \square

7.9.4. Teorema. *Sea V un espacio vectorial complejo con producto interno y de dimensión finita. Si $f : V \rightarrow V$ es una función lineal normal, entonces existe una base ortonormal de V cuyos elementos son autovectores de f .*

Demostración. Hagamos inducción con respecto a la dimensión de V , notando que si $V = 0$ no hay nada que probar. Como \mathbb{C} es algebraicamente cerrado, sabemos que existen $\lambda \in \mathbb{C}$ y $x \in V$ tales que $\|x\| = 1$ y $f(x) = \lambda x$. De la Proposición 7.9.3, entonces, vale también que $f^*(x) = \bar{\lambda} x$.

Sea $W = \langle x \rangle^\perp$. Si $y \in W$, entonces $y \perp x$ y tenemos que

$$\langle f(y), x \rangle = \langle y, f^*(x) \rangle = \langle y, \bar{\lambda} x \rangle = \bar{\lambda} \langle y, x \rangle = 0$$

y

$$\langle f^*(y), x \rangle = \langle y, f(x) \rangle = \langle y, \lambda x \rangle = \bar{\lambda} \langle y, x \rangle = 0.$$

Vemos así que el subespacio W es f -y f^* -invariante y que, en particular, podemos considerar las restricciones $f_W, (f^*)_W : W \rightarrow W$. Es inmediato verificar que f_W posee una adjunta y que, de hecho, $(f_W)^* = (f^*)_W$. Más aún, la función f_W es normal.

Como $\dim W = \dim V - 1$, podemos suponer que el teorema es cierto para f_W y, entonces, que existe una base ortonormal \mathcal{B}' de W formada por autovectores de f_W . Es claro entonces que $\mathcal{B} = \{x\} \cup \mathcal{B}'$ es una base ortonormal de V formada por autovectores de f . \square

7.9.5. Corolario. *Sea V un espacio vectorial complejo con producto interno y de dimensión finita y sea $f : V \rightarrow V$ una transformación lineal. Las siguientes afirmaciones son equivalentes:*

- (a) La función f es normal.
- (b) Existe una base ortonormal \mathcal{B} de V formada por autovectores de f .

Demostración. La implicación $(a) \Rightarrow (b)$ es el contenido del Teorema 7.9.4. Veamos la recíproca. Si \mathcal{B} es una base ortonormal formada por autovectores de f , entonces la matriz $[f]_{\mathcal{B}}^{\mathcal{B}}$ es diagonal y lo mismo es cierto de la matriz $[f^*]_{\mathcal{B}}^{\mathcal{B}} = \overline{([f]_{\mathcal{B}}^{\mathcal{B}})^t}$. Esto implica que estas dos matrices $[f]_{\mathcal{B}}^{\mathcal{B}}$ y $[f^*]_{\mathcal{B}}^{\mathcal{B}}$ comutan y, por lo tanto, que f y f^* comutan: en otras palabras, f es normal. \square

7.9.6. El siguiente resultado es conocido como el *Teorema espectral para transformaciones normales*:

Proposición. Sea V un espacio vectorial complejo con producto interno y de dimensión finita y sea $f : V \rightarrow V$ una transformación lineal normal. Sean $\lambda_1, \dots, \lambda_n \in \mathbb{C}$ los autovalores distintos de f y para cada $i \in \llbracket n \rrbracket$ sea $V_i = \text{Nu}(f - \lambda_i \text{id}_V)$ y $p_i : V \rightarrow V$ el proyector ortogonal de imagen V_i . Entonces:

- (i) Si $i, j \in \llbracket n \rrbracket$ son distintos, entonces $V_i \perp V_j$ y $p_i p_j = 0$.
- (ii) Hay una descomposición en suma directa $V = V_1 \oplus \dots \oplus V_n$.
- (iii) Se tiene que $\text{id}_V = p_1 + \dots + p_n$ y $f = \lambda_1 p_1 + \dots + \lambda_n p_n$.

Demostración. Si $i, j \in \llbracket n \rrbracket$ son distintos y $x \in V_i \setminus 0$ e $y \in V_j \setminus 0$, entonces x e y son autovectores de f de autovalores λ_i y λ_j y la Proposición 7.9.3 nos dice que $f^*(y) = \bar{\lambda}_j y$ y que, en consecuencia,

$$\lambda_i \langle x, y \rangle = \langle \lambda_i x, y \rangle = \langle f(x), y \rangle = \langle x, f^*(y) \rangle = \langle x, \lambda_j y \rangle = \lambda_j \langle x, y \rangle. \quad (14)$$

Como $\lambda_i \neq \lambda_j$, esto implica que $\langle x, y \rangle = 0$. Concluimos de esta forma que $V_i \perp V_j$ y, usando la Proposición 7.5.5, que $p_i p_j = 0$. Más aún, de acuerdo al Teorema 7.9.4, hay una base de autovectores de V así que, de hecho, es $V = \bigoplus_{i=1}^n V_i$.

Si $x \in V$, entonces existen $x_1 \in V_1, \dots, x_n \in V_n$ tales que $x = x_1 + \dots + x_n$. Si $i, j \in \llbracket n \rrbracket$, se tiene que

$$p_i(x_j) = \begin{cases} x_i, & \text{si } i = j; \\ 0, & \text{en caso contrario.} \end{cases}$$

En efecto, en el primer caso esto es así porque $x_j \in V_i$ y p_i es el proyector ortogonal con imagen V_i , en el segundo porque $x_j \in V_j \subseteq V_i^\perp = \text{Nu}((\lambda_i)p_i)$. Como consecuencia de esto se tiene que

$$(p_1 + \dots + p_n)(x) = \sum_{i=1}^n \sum_{j=1}^n p_i(x_j) = \sum_{i=1}^n p_i(x_i) = x$$

y

$$\begin{aligned} (\lambda_1 p_1 + \dots + \lambda_n p_n)(x) &= \sum_{i=1}^n \sum_{j=1}^n \lambda_i p_i(x_j) = \sum_{i=1}^n \lambda_i p_i(x_i) = \sum_{i=1}^n \lambda_i x_i \\ &= \sum_{i=1}^n f(x_i) = f\left(\sum_{i=1}^n x_i\right) = f(x) \end{aligned}$$

y esto implica que $p_1 + \dots + p_n = \text{id}_V$ y que $\lambda_1 p_1 + \dots + \lambda_n p_n = f$. \square

7.9.7. Lema. Sea V un espacio vectorial, sean $\lambda_1, \dots, \lambda_n \in \mathbb{k}$ escalares y sean $p_1, \dots, p_n : V \rightarrow V$ proyectores ortogonales tales que $\sum_{i=1}^n p_i = \text{id}_V$ y $p_i p_j = 0$ siempre que $i, j \in \llbracket n \rrbracket$ son distintos. Si $f = \sum_{i=1}^n \lambda_i p_i$, entonces para cada $h \in \mathbb{k}[X]$ es

$$h(f) = \sum_{i=1}^n h(\lambda_i) p_i.$$

Demostración. Por linealidad, es suficiente probar la igualdad del enunciado para el caso particular en que $h = X^r$ es un monomio y hacemos esto procediendo por inducción en r .

Cuando $r = 0$, de manera que $h = 1$, la igualdad (14) vale porque por hipótesis es $\text{id}_V = \sum_{i=1}^n p_i$. Por otro lado, si $p = X^{r+1}$, entonces

$$h(f) = f^{r+1} = f^r f = \left(\sum_{i=1}^n \lambda_i^r p_i \right) \left(\sum_{j=1}^n \lambda_j p_j \right) = \sum_{i=1}^n \sum_{j=1}^n \lambda_i^r \lambda_j p_i p_j$$

y, como $p_i p_j = 0$ si $i \neq j$, esto es

$$= \sum_{i=1}^n \lambda_i^{r+1} p_i = \sum_{i=1}^n h(\lambda_i) p_i.$$

Esto completa la inducción. \square

7.9.8. Proposición. *Sea V un espacio vectorial complejo con producto interno y de dimensión finita y sea $f : V \rightarrow V$ una función lineal. Las siguientes afirmaciones son equivalentes:*

- (a) *La función f es normal.*
- (b) *Existe un polinomio $h \in \mathbb{C}[X]$ tal que $f^* = p(f)$.*

Demostración. (b) \Rightarrow (a) Si existe $h \in \mathbb{C}[X]$ tal que $f^* = h(f)$, entonces

$$f^* \circ f = h(f) \circ f = (hX)(f) = (Xh)(f) = f \circ h(f) = f \circ f^*,$$

así que f es normal.

(a) \Rightarrow (b) De acuerdo al Teorema 7.9.6, hay escalares $\lambda_1, \dots, \lambda_n \in \mathbb{C}$ distintos dos a dos y proyectores ortogonales $p_1, \dots, p_n : V \rightarrow V$ tales que $p_i p_j = 0$ cada vez que $i, j \in \llbracket n \rrbracket$ son distintos. Consideraremos el polinomio

$$h(X) = \sum_{i=1}^n \bar{\lambda}_i \prod_{\substack{1 \leq j \leq n \\ j \neq i}} \frac{X - \lambda_j}{\lambda_i - \lambda_j} \in \mathbb{C}[X].$$

Es inmediato verificar que $h(\lambda_i) = \bar{\lambda}_i$ para cada $i \in \llbracket n \rrbracket$ y entonces, de acuerdo al Lema 7.9.7, se tiene que

$$h(f) = \sum_{i=1}^n h(\lambda_i) p_i = \sum_{i=1}^n \bar{\lambda}_i p_i.$$

Por otro lado, como los proyectores p_i son ortogonales y, entonces, autoadjuntos, tenemos que

$$f^* = \left(\sum_{i=1}^n \lambda_i p_i \right)^* = \sum_{i=1}^n \bar{\lambda}_i p_i,$$

Así, vemos que $h(f) = f^*$. \square

7.9.9. Ejemplo. En el Teorema 7.9.4 la hipótesis de que el espacio vectorial sea complejo es importante. En efecto, la función lineal $f : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ construida en el Ejemplo 7.9.2 es normal cualquiera sea $\alpha \in \mathbb{R}$ pero no tiene ningún autovector si α no es un múltiplo entero de π . \diamond

§10. Funciones unitarias y ortogonales

7.10.1. Si V es un espacio vectorial con producto interno y $f : V \rightarrow V$ es un endomorfismo de V tal que

$$\langle f(x), f(y) \rangle = \langle x, y \rangle$$

para todo $x, y \in V$, decimos que f es **unitario** cuando el cuerpo \mathbb{k} es \mathbb{C} y que es **ortogonal** cuando el cuerpo \mathbb{k} es \mathbb{R} .

7.10.2. Los endomorfismos unitarias u ortogonales de un espacio vectorial con producto interno son los endomorfismos de éste que preservan el producto interno. El siguiente resultado muestra que es suficiente para ello con que preserven la norma asociada a ese producto interno.

Proposición. Sea V un espacio vectorial con producto interno y sea $f : V \rightarrow V$ un endomorfismo de V . Las siguientes afirmaciones son equivalentes:

- (a) La función f es unitaria si $\mathbb{k} = \mathbb{C}$ u ortogonal si $\mathbb{k} = \mathbb{R}$.
- (b) Para todo $x \in V$ se tiene que $\|f(x)\| = \|x\|$.

Demostración. (a) \Rightarrow (b) Si f es unitario u ortogonal, según el caso, y $x \in V$, entonces

$$\|f(x)\| = \sqrt{\langle f(x), f(x) \rangle} = \sqrt{\langle x, x \rangle} = \|x\|.$$

(b) \Rightarrow (a) Sean x e y dos elementos de V . Si $\mathbb{k} = \mathbb{R}$, en vista del Corolario 7.2.3, tenemos que

$$\langle f(x), f(y) \rangle = \frac{1}{4} \|f(x) + f(y)\|^2 - \frac{1}{4} \|f(x) - f(y)\|^2 = \frac{1}{4} \|f(x+y)\|^2 - \frac{1}{4} \|f(x-y)\|^2$$

y, de acuerdo a la hipótesis, esto es

$$= \frac{1}{4} \|x+y\|^2 - \frac{1}{4} \|x-y\|^2 = \langle x, y \rangle,$$

de manera que f es ortogonal. De manera similar, si $\mathbb{k} = \mathbb{C}$ ese mismo corolario nos dice que

$$\begin{aligned} & \langle f(x), f(y) \rangle \\ &= \frac{1}{4} \|f(x) + f(y)\|^2 - \frac{1}{4} \|f(x) - f(y)\|^2 \frac{i}{4} \|f(x) + if(y)\|^2 - \frac{i}{4} \|f(x) - if(y)\|^2 \\ &= \frac{1}{4} \|f(x+y)\|^2 - \frac{1}{4} \|f(x-y)\|^2 \frac{i}{4} \|f(x+iy)\|^2 - \frac{i}{4} \|f(x-iy)\|^2 \\ &= \frac{1}{4} \|x+y\|^2 - \frac{1}{4} \|x-y\|^2 \frac{i}{4} \|x+iy\|^2 - \frac{i}{4} \|x-iy\|^2 \\ &= \langle x, y \rangle, \end{aligned}$$

por lo que f es en este caso unitario. \square

7.10.3. Proposición. Sea V un espacio vectorial con producto interno y sea $f : V \rightarrow V$ un endomorfismo de V que posee un endomorfismo adjunto. Las siguientes afirmaciones son equivalentes:

- (a) El endomorfismo f es unitario si $\mathbb{k} = \mathbb{C}$ u ortogonal si $\mathbb{k} = \mathbb{R}$.
- (b) Se tiene que $f^* \circ f = \text{id}_V = f \circ f^*$, de manera que f es inversible y tiene a su adjunto f^* como inverso.

Demostración. (a) \Rightarrow (b) Si f es unitario u ortogonal y $x \in V$, para cada $y \in V$ se tiene que

$$\langle (f^* \circ f)(x), y \rangle = \langle f^*(f(x)), y \rangle = \langle f(x), f(y) \rangle = \langle x, y \rangle,$$

así que $f^* \circ f = \text{id}_V$. Un argumento idéntico muestra que $f \circ f^* = \text{id}_V$.

(b) \Rightarrow (a) Supongamos que $f^* \circ f = \text{id}_V$. Si $x, y \in V$, entonces

$$\langle f(x), f(y) \rangle = \langle f^*(f(x)), y \rangle = \langle x, y \rangle,$$

de manera que f es unitario u ortogonal, según el caso. \square

7.10.4. Proposición. Sea V un espacio vectorial con producto interno y dimensión finita y sea $f : V \rightarrow V$ una función lineal. Las siguientes afirmaciones son equivalentes:

- (a) La función f es unitaria si $\mathbb{k} = \mathbb{C}$ u ortogonal si $\mathbb{k} = \mathbb{R}$.
- (b) Existe una base ortonormal \mathcal{B} de V tal que $f(\mathcal{B})$ es una base ortonormal de V .
- (c) Para cada base ortonormal \mathcal{B} de V se tiene que $f(\mathcal{B})$ es una base ortonormal de V .

Demostración. **HACER.** \square

§11. Matrices ortonormales y unitarias

7.11.1. Sea $n \in \mathbb{N}$ y consideremos a los espacios vectoriales \mathbb{R}^n y \mathbb{C}^n dotados con sus productos internos estándares.

- Decimos que una matriz $A \in M_n(\mathbb{R})$ es **ortogonal** si cada vez que x e y son vectores de \mathbb{R}^n , se tiene que $\langle Ax, Ay \rangle = \langle x, y \rangle$. Escribimos $O_n(\mathbb{R})$ al conjunto de las matrices ortogonales de $M_n(\mathbb{R})$.
- De manera similar, decimos que una matriz $A \in M_n(\mathbb{C})$ es **unitaria** si para cada $x, y \in \mathbb{C}^n$ vale que $\langle Ax, Ax \rangle = \langle x, y \rangle$. Escribimos $U_n(\mathbb{C})$ al conjunto de las matrices unitarias de $M_n(\mathbb{C})$ ■

7.11.2. Proposición. Sea $n \in \mathbb{N}$ y dotemos a \mathbb{R}^n y a \mathbb{C}^n de sus productos internos estándares.

- (i) Si $A \in M_n(\mathbb{R})$, entonces las siguientes afirmaciones son equivalentes:

- (a) La matriz A es ortogonal.
 (b) La función lineal $f_A : x \in \mathbb{R}^n \mapsto Ax \in \mathbb{R}^n$ es ortogonal.
 (c) Se tiene que $AA^t = I_n = A^tA$.
 (d) El conjunto de las columnas de A es una base ortonormal de \mathbb{R}^n .
- (ii) Si $A \in M_n(\mathbb{C})$, entonces las siguientes afirmaciones son equivalentes:
- (a) La matriz A es unitaria
 (b) La función lineal $f_A : x \in \mathbb{C}^n \mapsto Ax \in \mathbb{C}^n$ es unitaria.
 (c) Se tiene que $AA^* = I_n = A^*A$.
 (d) El conjunto de las columnas de A es una base ortonormal de \mathbb{C}^n .

Demostración. **HACER.**

□

7.11.3. Proposición. Sea $n \in \mathbb{N}$.

- (i) Si $A \in M_n(\mathbb{R})$ es una matriz simétrica, entonces existe una matriz ortogonal $C \in O_n(\mathbb{R})$ tal que la matriz $C^t AC$ es diagonal.
 (ii) Si $A \in M_n(\mathbb{C})$ es una matriz hermitiana, entonces existe una matriz unitaria $C \in U_n(\mathbb{C})$ tal que la matriz $C^* AC$ es diagonal.

Demostración. **HACER.**

□

7.11.4. La siguiente proposición —una consecuencia directa del procedimiento de ortogonalización de Gram–Schmidt— nos dice que toda matriz cuadrada con entradas en \mathbb{k} tiene una factorización como producto de una matriz ortogonal o unitaria y de una matriz triangular superior. Esta factorización se conoce usualmente con el nombre de **factorización QR** y es de mucho interés en las aplicaciones: por ejemplo, su uso es frecuente al resolver problemas de cuadrados mínimos en estadística y es la base del llamado *algoritmo QR* para calcular los autovalores de una matriz real.

Proposición. Sea $n \in \mathbb{N}$ y sea $A \in M_n(\mathbb{k})$.

- (i) Existen una matriz Q ortogonal si $\mathbb{k} = \mathbb{R}$ y unitaria si $\mathbb{k} = \mathbb{C}$, y una matriz triangular superior $R \in M_n(\mathbb{k})$ tales que $A = QR$.
 (ii) Si A es inversible, entonces podemos elegir estas matrices Q y R de manera que las entradas que aparecen a lo largo de la diagonal en la matriz R son números positivos, y si hacemos esto tanto Q como R están únicamente determinadas por A .

Demostración. **HACER.**

□

§12. Una familia importante de ejemplos

7.12.1. Si X es un conjunto, un **grafo** con conjunto de vértices X es una relación simétrica y anti-reflexiva sobre X , esto es, un subconjunto Γ de $X \times X$ tal que

- cada vez que x e y son elementos de X tales que $(x, y) \in \Gamma$, se tiene que $(y, x) \in \Gamma$, y
- para todo $x \in X$ vale que $(x, x) \notin \Gamma$.

Una forma eficiente de presentar un grafo Γ sobre X es dibujándolo: hacemos un punto \circ por cada elemento de X , decorándolo si es necesario con su nombre, y cada vez que x e y son vértices de Γ tales que $(x, y) \in \Gamma$, conectamos los puntos correspondientes a x y a y en el dibujo con un segmento $\circ — \circ$. Por ejemplo, representamos al grafo Γ sobre el conjunto $X = [5]$ tal que

$$\Gamma = \{(1, 2), (2, 1), (2, 3), (3, 2), (3, 1), (1, 3), (1, 4), (4, 1), (4, 5), (5, 4)\}$$

por el dibujo



Es importante observar que las posiciones relativas de los puntos en el dibujo no es importante: lo único significativo en uno de estos dibujos es qué vértices están conectados entre sí. Así, los siguientes dibujos representan el mismo grafo que el anterior:



7.12.2. Si Γ es un grafo con conjunto de vértices X , consideremos el espacio vectorial real \mathbb{R}^X de todas las funciones $X \rightarrow \mathbb{R}$ y, si $f \in \mathbb{R}^X$ e $i \in X$, escribamos f_i en lugar de $f(i)$. Definimos una función

$$\langle -, - \rangle_\Gamma : \mathbb{R}^X \times \mathbb{R}^X \rightarrow \mathbb{R}^X$$

poniendo

$$\langle f, g \rangle_\Gamma = 2 \sum_{i=1}^n f_i g_i - \sum_{(i,j) \in \Gamma} f_i g_j$$

para cada par de funciones $f, g \in \mathbb{R}^X$. Así, si Γ es el grafo con conjunto de vértices $[5]$ representado

en la figura (15), entonces

$$\begin{aligned}\langle f, g \rangle_{\Gamma} &= 2f_1g_1 + 2f_2g_2 + 2f_3g_3 + 2f_4g_4 + 2f_5g_5 \\ &\quad - f_1g_2 - f_2g_1 - f_2g_3 - f_3g_2 - f_3g_1 - f_1g_3 - f_1g_4 - f_4g_1 - f_4g_5 - f_5g_4\end{aligned}$$

para cada $f, g \in \mathbb{R}^{[5]}$.

Digamos que un grafo Γ es **bueno** si la función $\langle -, - \rangle_{\Gamma}$ es un producto interno. Es inmediato verificar que esta función satisface las condiciones **(PI₁)**, **(PI₂)** y **(PI₃)** de la definición 7.1.2 cualquiera sea Γ . Nuestro objetivo en esta sección es describir cuáles son precisamente los grafos buenos, esto es, los que hacen que la condición **(PI₄)** se satisfaga.

7.12.3. Una primera observación que podemos hacer es que que un grafo sea bueno depende solamente de la forma en que están conectados sus vértices y no de las etiquetas que éstos tienen. Para precisar esta afirmación necesitamos la siguiente definición.

Si Γ y Γ' son dos grafos con conjuntos de vértices X y X' , respectivamente, decimos que Γ y Γ' son **isomorfos** si existe una función biyectiva $\phi : X \rightarrow X'$ tal que para cada par de vértices $i, j \in X$ se tiene que

$$(i, j) \in \Gamma \iff (\phi(i), \phi(j)) \in X',$$

y en ese caso decimos que esa función ϕ es un **isomorfismo de grafos** y vale que la función

$$(i, j) \in \Gamma \mapsto (\phi(i), \phi(j)) \in \Gamma'$$

es una biyección. La idea detrás de esta definición es que dos grafos son isomorfos exactamente cuando ambos pueden ser representados por el mismo dibujo, una vez que eliminamos las etiquetas que identifican a los vértices.

Proposición. Sean Γ y Γ' grafos con conjuntos de vértices X y X' , respectivamente. Si Γ es bueno y los grafos Γ y Γ' son isomorfos, entonces Γ' también es bueno.

Demostración. Supongamos que Γ es bueno y que hay un isomorfismo $\phi : X \rightarrow X'$ entre los grafos Γ y Γ' . Consideremos la función $\Phi : f \in \mathbb{R}^{X'} \mapsto f \circ \phi \in \mathbb{R}^X$, de manera que para cada $f \in \mathbb{R}^{X'}$ y cada $i \in X$ se tiene $\Phi(f)_i = f_{\phi(i)}$. Esta función es lineal, como puede verse inmediatamente, y se trata, de hecho, de un isomorfismo: si $\psi : X' \rightarrow X$ es la biyección inversa a ϕ , entonces la función lineal $\Psi : g \in \mathbb{R}^X \mapsto g \circ \psi \in \mathbb{R}^{X'}$ es inversa a Φ .

Si $f, g \in \mathbb{R}^{X'}$, entonces se tiene que

$$\begin{aligned}\langle \Phi(f), \Phi(g) \rangle_{\Gamma} &= 2 \sum_{i \in X} \Phi(f)_i \Phi(g)_i - \sum_{(i,j) \in \Gamma} \Phi(f)_i \Phi(g)_j \\ &= 2 \sum_{i \in X} f_{\phi(i)} g_{\phi(i)} - \sum_{(i,j) \in \Gamma} f_{\phi(i)} g_{\phi(j)}\end{aligned}$$

y, como las funciones ϕ y $(i, j) \in \Gamma \mapsto (\phi(i), \phi(j)) \in \Gamma'$ son biyecciones, esto es

$$= 2 \sum_{i \in X'} f_i g_i - \sum'_{(i,j) \in \Gamma} f_i g_j = \langle f, g \rangle_{\Gamma'}.$$

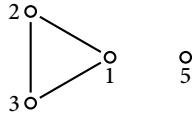
En particular, si $f \in \mathbb{R}^{X'}$ entonces

$$\langle f, f \rangle_{\Gamma'} = \langle \Phi(f), \Phi(f) \rangle_{\Gamma} \geq 0,$$

porque $\langle -, - \rangle_{\Gamma}$ es un producto interno y, más aún, $\langle f, f \rangle_{\Gamma'} = 0$ si y solamente si $\langle \Phi(f), \Phi(f) \rangle_{\Gamma} = 0$, y esto ocurre si y solamente si $\Phi(f) = 0$, esto es, si $f = 0$. Esto nos dice que la función $\langle -, - \rangle_{\Gamma'}$ satisface la condición **(PI4)** y que, por lo tanto, el grafo Γ' es bueno. \square

7.12.4. La segunda observación importante que tenemos que hacer es que si a un grafo bueno le sacamos vértices entonces lo que queda sigue siendo bueno. Formalicemos esta afirmación.

Si Γ y Γ' son grafos con conjuntos de vértices X e Y , respectivamente, decimos que Γ' es un **subgrafo** de Γ si $Y \subseteq X$ y $\Gamma' = \Gamma \cap (Y \times Y)$. Notemos que en ese caso el grafo Γ' queda completamente determinado por Γ y su conjunto de vértices Y . Por ejemplo, el subgrafo del grafo de la figura (15) de la página 369 que tiene como conjunto de vértices a $\{1, 2, 3, 5\}$ es el siguiente grafo no conexo:



Proposición. *Sea Γ un grafo y sea Γ' un subgrafo de Γ . Si Γ es bueno, entonces Γ' también lo es.*

Demostración. Supongamos que X es el conjunto de vértices de Γ , que $Y \subseteq X$ es el conjunto de vértices de Γ' y que Γ es bueno. Hay una función lineal $\zeta : \mathbb{R}^Y \rightarrow \mathbb{R}^X$ tal que para cada $f \in \mathbb{R}^Y$ y cada $i \in X$ se tiene que

$$\zeta(f)_i = \begin{cases} f_i, & \text{si } i \in Y; \\ 0, & \text{si } i \in X \setminus Y. \end{cases}$$

En otras palabras, para cada $f \in \mathbb{R}^Y$ la función $\zeta(f) : X \rightarrow \mathbb{R}$ se obtiene de f extendiendo por cero fuera del conjunto Y .

Si $f, g \in \mathbb{R}^Y$, entonces

$$\langle \zeta(f), \zeta(g) \rangle_{\Gamma} = 2 \sum_{i \in X} \zeta(f)_i \zeta(g)_i - \sum_{(i,j) \in \Gamma} \zeta(f)_i \zeta(g)_j$$

y, como $\zeta(f)_i = 0$ si $i \in X \setminus Y$ y $\Gamma' = \Gamma \cap Y \times Y$, esto es

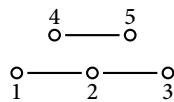
$$\begin{aligned} &= 2 \sum_{i \in Y} \zeta(f)_i \zeta(g)_i - \sum_{(i,j) \in \Gamma'} \zeta(f)_i \zeta(g)_j \\ &= 2 \sum_{i \in Y} f_i g_i - \sum_{(i,j) \in \Gamma'} f_i g_j \\ &= \langle f, g \rangle_{\Gamma'} \end{aligned}$$

En particular, si $f \in \mathbb{R}^Y$, entonces

$$\langle f, f \rangle_{\Gamma'} = \langle \zeta(f), \zeta(f) \rangle_{\Gamma} \geq 0$$

y $\langle f, f \rangle_{\Gamma'} = 0$ si y solamente si $\langle \zeta(f), \zeta(f) \rangle_{\Gamma} = 0$, lo que ocurre exactamente cuando $\zeta(f) = 0$. Como ζ es una función inyectiva, esto implica que $\langle -, - \rangle_{\Gamma'}$ satisface la condición **(PI₄)**: vemos así que el grafo Γ' es bueno, como queríamos. \square

7.12.5. Decimos que un grafo Γ con conjunto de vértices X es **conexo** si cada vez que i y j son vértices de X existe $k \in \mathbb{N}_0$ y una sucesión de vértices $i_0, i_1, \dots, i_k \in X$ tales que $i_0 = i$, $i_k = j$ y $(i_{t-1}, i_t) \in \Gamma$ para cada $t \in \llbracket k \rrbracket$. El grafo de la figura (15) de la página 369 es conexo, mientras que el grafo



con conjunto de vértices $\llbracket 5 \rrbracket$ claramente no lo es.

7.12.6. Estamos en posición de enunciar y probar el resultado principal de esta sección:

Teorema. *Un grafo conexo Γ es bueno si y solamente si Γ es isomorfo a uno de los grafos A_n con $n \geq 1$, D_n con $n \geq 4$, E_6 , E_7 o E_8 ilustrados en la Figura 7.2 de la página 373.*

Los grafos de esa figura son los llamados **diagramas de Dynkin simplemente enlazados**, por Eugene Dynkin (1924–2014, Rusia), y aparecen en una gran cantidad de contextos distintos. El nombre de cada uno de estos grafos tiene como subíndice el número de vértices que tiene.

Demostración. Empezamos observando que los grafos \tilde{A}_n con $n \geq 2$, \tilde{D}_n con $n \geq 4$, \tilde{E}_6 , \tilde{E}_7 y \tilde{E}_8 de las Figuras 7.3 y 7.4 de las páginas 375 y 376 no son buenos. \square

Acompaña a cada uno de los grafos una descomposición de la función $\langle -, - \rangle_{\Gamma}$ como suma de cuadrados, que puede verificarse por medio de un cálculo directo. Por ejemplo, la Figura 7.3 nos informa que si $\Gamma = \tilde{D}_n$, entonces para cada $x = (x_1, \dots, x_{n+1}) \in \mathbb{R}^{n+1}$ vale que

$$4\langle x, x \rangle_{\Gamma} = (2x_1 - x_3)^2 + (2x_2 - x_3)^2 + 2 \sum_{i=3}^{n-2} (x_i - x_{i+1})^2 + (x_{n-1} - 2x_n)^2 + (x_{n-1} - 2x_{n+1})^2.$$

$$\text{A}_n, \quad n \geq 1$$

$$\text{E}_6$$

$$\text{D}_n, \quad n \geq 4$$

$$\text{E}_7$$

$$\text{E}_8$$

Figura 7.2. Los diagramas de Dynkin simplemente enlazados.

Es claro entonces que $\langle x, x \rangle_{\Gamma} \geq 0$ para todo $x \in \mathbb{R}^n$ y que $\langle x, x \rangle_{\Gamma} = 0$ si y solamente si se satisfacen las igualdades

$$2x_1 - x_3 = 0, \quad 2x_2 - x_3 = 0, \quad x_{n-1} - 2x_n = 0, \quad x_{n-1} - 2x_{n+1} = 0$$

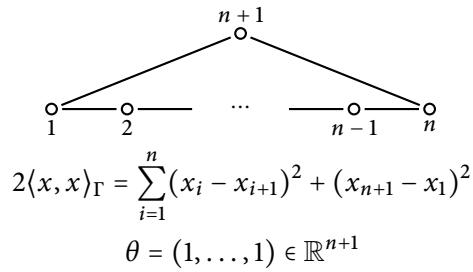
y

$$x_i - x_{i+1} = 0 \text{ para cada } i \in [3, n-2].$$

Es fácil resolver estas ecuaciones: se satisfacen si y solamente si x es un múltiplo escalar del vector $\theta = (1, 1, 2, \dots, 2, 1, 1)$ mencionado en la figura.

Exactamente lo mismo ocurre con cada uno de estos grafos —omitimos las verificación de esto, que es completamente rutinaria.

$\widetilde{\mathbf{A}}_n, \quad n \geq 2$



$\widetilde{\mathbf{D}}_n, \quad n \geq 4$

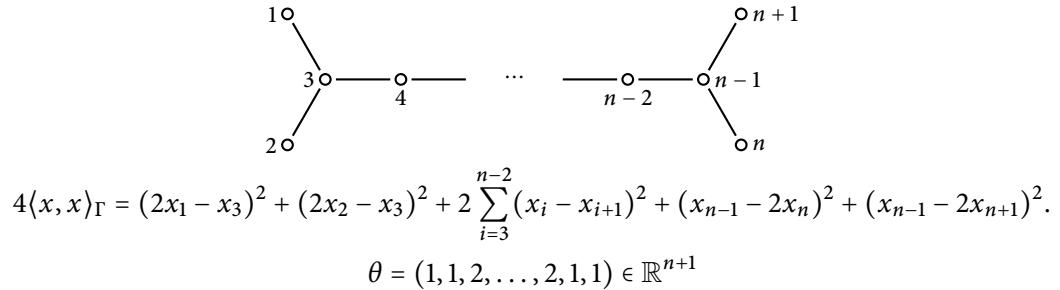
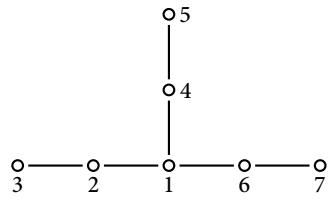


Figura 7.3. Los diagramas euclídeos $\widetilde{\mathbf{A}}_n$ y $\widetilde{\mathbf{D}}_n$.

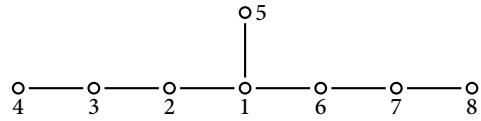
\widetilde{E}_6



$$36\langle x, x \rangle_{\Gamma} = (6x_3 - 3x_2)^2 + (6x_7 - 3x_6)^2 + (6x_5 - 3x_4)^2 + 3(3x_2 - 2x_1)^2 + 3(3x_6 - 2x_1)^2 + 2(3x_4 - 2x_1)^2$$

$$\theta = (3, 2, 1, 2, 1, 2, 1) \in \mathbb{R}^7$$

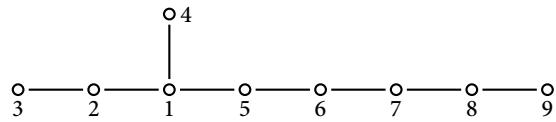
\widetilde{E}_7



$$24\langle x, x \rangle_{\Gamma} = 6(2x_4 - x_3)^2 + 6(2x_8 - x_7)^2 + 2(3x_3 - 2x_2)^2 + 2(3x_7 - 2x_6)^2 + (4x_2 - 3x_1)^2 + (4x_6 - 3x_1)^2 + 6(2x_5 - x_1)^2$$

$$\theta = (4, 3, 2, 1, 2, 3, 2, 1) \in \mathbb{R}^8$$

\widetilde{E}_8



$$120\langle x, x \rangle_{\Gamma} = 30(2x_9 - x_8)^2 + 10(3x_8 - 2x_7)^2 + 5(4x_7 - 3x_6)^2 + 3(5x_6 - 4x_5)^2 + 30(2x_3 - x_2)^2 + 2(6x_5 - 5x_1)^2 + 10(3x_2 - 2x_1)^2 + 30(2x_4 - x_1)^2$$

$$\theta = (6, 4, 2, 3, 5, 4, 3, 2, 1) \in \mathbb{R}^9$$

Figura 7.4. Los diagramas euclídeos \widetilde{E}_6 , \widetilde{E}_7 y \widetilde{E}_8 .

Referencias

- [AM57] Abraham Adrian Albert y Benjamin Muckenhoupt, *On matrices of trace zero*, Michigan Math. J. **4** (1957), 1–3. MR83961 ↑
- [And98] George E. Andrews, *The theory of partitions*, Cambridge Mathematical Library, Cambridge University Press, Cambridge, 1998. Reprint of the 1976 original. MR1634067 ↑305
- [Bla84] Andreas Blass, *Existence of bases implies the axiom of choice*, Axiomatic set theory (Boulder, Colo., 1983), Contemp. Math., vol. 31, Amer. Math. Soc., Providence, RI, 1984, pp. 31–33, DOI 10.1090/conm/031/763890. MR763890 ↑27
- [Bre06] Frédéric Brechenmacher, *Histoire du théorème de Jordan de la décomposition matricielle (1870–1930)*, Tesis de doctorado, École des Hautes Études en Sciences sociales, París, Francia, 2006. ↑321
- [Bru87] Richard A. Brualdi, *The Jordan canonical form: an old proof*, Amer. Math. Monthly **94** (1987), no. 3, 257–267, DOI 10.2307/2323392. MR883292 ↑321
- [Cat62] S. Cater, *An Elementary Development of the Jordan Canonical Form*, Amer. Math. Monthly **69** (1962), no. 5, 391–393, DOI 10.2307/2312130. MR1531677 ↑321
- [Cay58] Arthur Cayley, *A Memoir on the Theory of Matrices*, Phil. Trans. R. Soc. of London **148** (1858), 17–37, disponible en <http://www.jstor.org/stable/108649>. ↑259
- [CL55] Earl A. Coddington y Norman Levinson, *Theory of ordinary differential equations*, McGraw-Hill Book Company, Inc., New York-Toronto-London, 1955. MR0069338 ↑223
- [Dedoo] Richard Dedekind, *Ueber die von drei Moduln erzeugte Dualgruppe*, Math. Ann. **53** (1900), no. 3, 371–403, DOI 10.1007/BF01448979 (German). MR1511094 ↑73

- [Die57] Jean Dieudonné, *Le calcul différentiel dans les corps de caractéristique $p > 0$* , Proceedings of the International Congress of Mathematicians, Amsterdam, 1954, Vol. 1, Erven P. Noordhoff N.V., Groningen; North-Holland Publishing Co., Amsterdam, 1957, pp. 240–252 (French). MR0095218 ↑326
- [FS83] Roger Fletcher y Danny C. Sorensen, *An algorithmic derivation of the Jordan canonical form*, Amer. Math. Monthly **90** (1983), no. 1, 12–16, DOI 10.2307/2975686. MR691009 ↑321
- [Fil71] Aleksei Fedorovich Filippov, *A short proof of the theorem on reduction of a matrix to Jordan form.*, Vestnik Moskov. Univ. Ser. I Mat. Meh. **26** (1971), no. 2, 18–19. MR0279113 ↑321
- [GW80] Anatoly M. Galperin y Zeev Waksman, *An elementary approach to Jordan theory*, Amer. Math. Monthly **87** (1980), no. 9, 728–732. MR602830 ↑321
- [GG96] Israel Gohberg y Seymour Goldberg, *A simple proof of the Jordan decomposition theorem for matrices*, Amer. Math. Monthly **103** (1996), no. 2, 157–159, DOI 10.2307/2975110. MR1375060 ↑321
- [Ham53] William Rowan Hamilton, *Lectures on Quaternions*, Hodges and Smith, Dublin, Ireland, 1853. Containing a Systematic Statement of a New Mathematical Method; of which the Principles Were Communicated in 1843 to the Royal Irish Academy; and which Has Since Formed the Subject of Successive Courses of Lectures, Delivered in 1848 and Subsequent Years, in the Halls of Trinity College, Dublin: with Numerous Illustrative Diagrams, and with Some Geometrical and Physical Applications. ↑259
- [Gin10] Harry Gindi, *Slick proof?: A vector space has the same dimension as its dual if and only if it is finite dimensional*, MathOverflow, 2010, <https://mathoverflow.net/q/13322>. (version: 2010-01-29). ↑141
- [Jor70] Camille Jordan, *Traité des substitutions et des équations algébriques*, Gauthier-Villars, Paris, Francia, 1870. ↑319, 321
- [JvN35] Pascual Jordan y John von Neumann, *On inner products in linear, metric spaces*, Ann. of Math. (2) **36** (1935), no. 3, 719–723. MR1503247 ↑337
- [KM75] D. Kleitman y G. Markowsky, *On Dedekind's problem: the number of isotone Boolean functions. II*, Trans. Amer. Math. Soc. **213** (1975), 373–390, DOI 10.2307/1998052. MR382107 ↑73
- [MB56] Romolo Musti y Ettore Buttafuoco, *Sui subreticolati distributivi dei reticolati modulari*, Boll. Un. Mat. Ital. (3) **11** (1956), 584–587 (Italian). MR0083976 ↑72

- [Noe26] Emmy Noether, *Abstrakter Aufbau der Idealtheorie in algebraischen Zahl und Funktionenkörpern*, Math. Ann. **96** (1926), 26–61. ↑119
- [OEIS] OEIS Foundation Inc., *The On-Line Encyclopedia of Integer Sequences* (2020), <http://oeis.org/>. ↑73
- [Pas40] Blaise Pascal, *Essay pour les coniques*, Niedersächsische Landesbibliothek (1640), disponible en <http://digitale-sammlungen.gwlb.de/resolve?id=00066632>. ↑147
- [PPo5] Alexander Polishchuk y Leonid Positselski, *Quadratic algebras*, University Lecture Series, vol. 37, American Mathematical Society, Providence, RI, 2005. MR2177131 ↑72
- [Roi99] Moshe Roitman, *A short proof of the Jordan decomposition theorem*, Linear and Multilinear Algebra **46** (1999), no. 3, 245–247, DOI [10.1080/03081089908818616](https://doi.org/10.1080/03081089908818616). MR1708591 ↑321
- [Sho37] Kenjiro Shoda, *Einige Sätze über Matrizen*, Jpn. J. Math. **13** (1937), no. 3, 361–365. MR3223061 ↑106
- [Väl86] Hannu Väliaho, *An elementary approach to the Jordan form of a matrix*, Amer. Math. Monthly **93** (1986), no. 9, 711–714, DOI [10.2307/2322285](https://doi.org/10.2307/2322285). MR863972 ↑321
- [vdW30] Bartel Leendert van der Waerden, *Moderne Algebra. Bd. I. Unter Benutzung von Vorlesungen von E. Artin und E. Noether*, Die Grundlehren der mathematischen Wissenschaften in Einzeldarstellungen mit besonderer Berücksichtigung der Anwendungsgebiete, vol. 23, J. Springer, Berlin, 1930. ↑119
- [Wie91] Doug Wiedemann, *A computation of the eighth Dedekind number*, Order **8** (1991), no. 1, 5–6, DOI [10.1007/BF00385808](https://doi.org/10.1007/BF00385808). MR1129608 ↑73
- [Yca13] Bernard Ycart, *A case of mathematical eponymy: the Vandermonde determinant*, Rev. Histoire Math. **19** (2013), no. 1, 43–77, disponible en <http://arxiv.org/abs/1204.4716>. MR3155603 ↑199
- [vYz93] Jan van Yzeren, *A simple proof of Pascal’s hexagon theorem*, Amer. Math. Monthly **100** (1993), no. 10, 930–931, DOI [10.2307/2324214](https://doi.org/10.2307/2324214). MR1252929 ↑147

Índice alfabético

A

Albert, Abraham, 106
Álgebra, 93
Andrews, George E., 305
Autoespacio
 de un endomorfismo, 215
 de una matriz, 217
Autovalor
 de un endomorfismo, 215
 de una matriz, 217
Autovector
 de un endomorfismo, 215
 de una matriz, 217

B

Base, 23
 dual, 138
 ordenada, 46
 reducida, 290
Bloque de Jordan, 318
 nilpotente, 300
Buttafuoco, Ettore, 72

C

Cadena, 306
Característica de un cuerpo, 5
Cayley, Arthur, 259
Chebyshev, Pafnuti, 236

Codimensión de un subespacio, 44
Combinación lineal, 13
Complejo, 124
Complemento de un subespacio, 42
Componente primaria, 261
Congruencia módulo un subespacio, 112
Conjugación, 293
Comutador de dos matrices, 106
Conúcleo, 118
Coordenadas
 de Plücker, 213
 de un vector, 46
Cramer, Daniel, 194
Cuerpo, 1
 de funciones racionales, 4
 de Galois, 4
 primo, 6

D

Dedekind, Richard, 72
Defecto de una terna de subespacios, 56
Dependencia lineal, 18
Descomposición primaria, 261
Determinante
 de un endomorfismo, 180
 de una matriz, 177
Dieudonné, Jean, 326

- Dimensión de un espacio, 29
- E**
- Ecuación de Klein, 214
 - Endomorfismo, 75
 - de orden finito, 264
 - descomponible, 294
 - diagonalizable, 224
 - idempotente, 106
 - indescomponible, 294
 - nilpotente, 297
 - semisimple, 271
 - triangularizable, 269
 - Epimorfismo, 82
 - Equivalencia
 - a derecha, 288
 - Espacio vectorial, 6
 - cociente, 112
 - complejo, 7
 - de dimensión finita, 29
 - dual, 137
 - finitamente generado, 24
 - nulo, 8
 - real, 7
 - Espectro
 - de un endomorfismo, 215
 - de una matriz, 217
- F**
- Forma normal de Jordan, 320
 - Fórmula
 - de Laplace, 191
 - de Leibniz, 188
 - Función
 - alternante, 164
 - anti-simétrica, 166
 - bilineal, 91, 163
 - de tipo Fredholm, 133
 - lineal, 75
 - multilineal, 163
- G**
- Generador mónico, 246
 - Grado
 - de una función multilineal, 163
- H**
- Hamilton, William Rowan, 259
 - Hardy, Godfrey Harold, 305
 - Hermite, Charles, 239
 - Homomorfismo
 - de álgebras, 94
 - de conexión, 129
 - de espacios vectoriales, 75
- I**
- Ideal de un álgebra, 94
 - Identidad de Jacobi, 106
 - Imagen
 - de un homomorfismo, 81
 - de un subespacio por un homomorfismo, 81
 - Independencia lineal, 18
 - Índice
 - de Fredholm, 133
 - de nilpotencia, 297
 - Isomorfismo
 - de álgebras, 94
 - de espacios vectoriales, 83
- J**
- Jacobi, Carl Gustav Jacob, 106
- K**
- Klein, Felix, 214
 - Kleitman, D., 73
- L**
- Laplace, Pierre-Simon, 191
 - Leibniz, Gottfried Wilhelm, 188

- M**
- Markowsky, G.*, 73
- Matriz
- compañera, 198
 - de cambio de base, 47
 - de cofactores, 193
 - de permutación, 183
 - de una función lineal, 95
 - de Vandermonde, 199
 - diagonalizable, 226
 - elemental, 284
 - nilpotente, 297
 - reducida por columnas, 292
 - triangular inferior, 196
 - triangular superior, 196
- Menor de una matriz, 190
- Mínimo común múltiplo, 255
- Monomorfismo, 82
- Muckenhoupt, Benjamin*, 106
- Multiplicidad
- algebraica, 233
 - geométrica, 215
- Musti, Romolo*, 72
- N**
- Noether, Emmy*, 119
- Núcleo
- de un homomorfismo, 81
 - de un homomorfismo de álgebras, 94
- Nulidad de una función lineal, 88
- Número de Dedekind, 72
- O**
- Operación
- de columnas, 284
 - de filas, 285
- P**
- Partición de un entero, 305
- Pascal, Blaise*, 147
- Permutación, 181
- identidad, 181
- impar, 186
- par, 186
- Plücker, Julius*, 213
- Polinomio
- de Chebyshev, 236
 - de Hermite, 239
- Polinomio característico
- de un endomorfismo, 231
 - de una matriz, 228
- Polinomio minimal
- de un endomorfismo, 247
 - de un vector, 254
- Polishchuk, Alexander*, 72
- Positselski, Leonid*, 72
- Preimagen
- de un subespacio, 81
- Proyección canónica a un cociente, 112
- Proyector, 106
- R**
- Ramanujan, Srinivasa*, 305
- Rango
- de una función lineal, 88
- Regla de Cramer, 194
- Relación de dependencia lineal, 18
- Relaciones de Plücker, 214
- Restricción de un homomorfismo, 134
- Reticulado de subespacios, 69
- distributivo, 71
 - generado por un conjunto, 70
- S**
- Sarrus, Pierre Frédéric*, 178
- Shoda, Kenjiro*, 106
- Signo de una permutación, 185
- Sistema completo de proyectores ortogonales, 111
- Subespacio
- cíclico, 253

f-invariante, 134, 251
Subespacio vectorial, 10
 de codimensión finita, 44
 de codimensión infinita, 44
 generado por un conjunto, 14
 impropio, 10
 no trivial, 10
 nulo, 10
 triviales, 10
Subespacios
 independientes, 37
Subespacios comparables, 65
Sucesión exacta, 124
 corta, 124
Suma de subespacios, 34
Suma directa de subespacios, 37

T

Terna distributiva de subespacios, 58
Tipo
 de un endomorfismo, 320
 de un endomorfismo nilpotente, 291, 304
Transposición, 182
Traza
 de un endomorfismo, 103
 de una matriz, 101

V

van der Waerden, Bartel Leendert, 119
Vandermonde, Alexandre-Théophile, 199
Vector cíclico, 256

W

Wiedemann, Doug, 73

