

Un curso de Álgebra Lineal II

Camilo Sanabria

Universidad de los Andes
Departamento de Matemáticas
Bogotá - Colombia.

Índice general

1. Espacios vectoriales y transformaciones lineales	1
1.1. Espacios vectoriales	1
1.2. Base y dimensión	6
1.3. Transformaciones lineales	11
1.4. Matrices y vectores de coordenadas	18
1.5. Sumas directas y proyecciones	27
1.6. Espacios cociente	34
2. Estructura de las transformaciones lineales	39
2.1. Descomposición por suma directa	39
2.2. Espacios estables, cíclicos y propios	40
2.3. Polinomios de operadores	44
2.4. El polinomio característico y el teorema de Cayley-Hamilton	45
2.5. Descomposición en espacios estables y el lema de los núcleos	48
2.6. Operadores nilpotentes y forma normal de Jordan	55
2.7. Descomposición diagonalizable-nilpotente	65
3. Espacio dual	71
3.1. Funcionales lineales	71
3.2. Transformación dual	75
4. Espacios euclídeos	81
4.1. Producto interno	81
4.2. Operador adjunto	88
4.3. Operadores ortogonales	93
5. Espacios unitarios	97
5.1. Producto hermítico	97
5.2. Operador adjunto	101
5.3. Operadores unitarios	107
5.4. Estructura compleja	108

6. Espacios simplécticos	115
6.1. Forma simpléctica	115
6.2. Subespacios isotrópicos y bases de Darboux	118
6.3. Operadores adjuntos	125
7. Álgebra Multilineal	133
7.1. Tensores	133
7.2. Tensores alternantes	136
7.3. (l, k) -Tensores	144
7.4. Convenciones en notación de tensores	148
A. Cuerpos	151

Índice de figuras

2.1. Subespacios cíclicos colapsando	61
3.1. Transformación dual	75
5.1. Estructura compleja	112
6.1. Descomposición lagrangiana	121

Capítulo 1

Espacios vectoriales y transformaciones lineales

Sea \mathbb{K} un cuerpo (ver la definición A.1). A los elementos de \mathbb{K} los llamaremos *escalares*.

1.1. Espacios vectoriales

Definición 1.1. Un *espacio vectorial sobre \mathbb{K}* es un conjunto V junto con dos operaciones binarias

$$\begin{array}{ccc} + : V \times V & \longrightarrow & V \\ (v, w) & \longmapsto & v + w \end{array} \qquad \begin{array}{ccc} \cdot : \mathbb{K} \times V & \longrightarrow & V \\ (c, v) & \longmapsto & cv, \end{array}$$

que llamamos respectivamente *suma* y *producto por escalar* (o *adición* y *multiplicación por escalar*), y que contiene un elemento $O \in V$, que llamamos *el origen*, los cuales satisfacen las siguientes propiedades.

- (i) *La estructura $(V, +, O)$ es un grupo abeliano:* para todo $u, v, w \in V$ se tiene

$$v + w = w + v, \quad u + (v + w) = (u + v) + w, \quad v + O = v,$$

y para todo $v \in V$ existe $-v \in V$ que satisface la igualdad $-v + v = O$.

- (ii) *El producto por escalar es unitario y asociativo:* para todo $a, b \in \mathbb{K}$ y todo $v \in V$ se tiene

$$1v = v, \quad a(bv) = (ab)v.$$

- (iii) *El producto por escalar se distribuye sobre la suma:* para todo $a, b \in \mathbb{K}$ y todo $v, w \in V$ se tiene

$$a(v + w) = av + aw, \quad (a + b)v = av + bv.$$

Un *vector* es un elemento de un espacio vectorial.

Ejemplo 1.2. Los siguientes espacios vectoriales sobre \mathbb{K} son ejemplos imprescindibles.

- (i) *Espacio cero-dimensional*: El conjunto $\{O\}$ junto con las únicas operaciones posibles.
- (ii) *Espacio uni-dimensional*: El conjunto \mathbb{K} junto con las operaciones del cuerpo y $O = 0$.
- (iii) *Espacio n -dimensional de coordenadas*: El conjunto \mathbb{K}^n formado por el producto cartesiano de n copias del cuerpo K , junto con las operaciones

$$\begin{aligned}(a_1, \dots, a_n) + (b_1, \dots, b_n) &= (a_1 + b_1, \dots, a_n + b_n) \\ c(a_1, \dots, a_n) &= (ca_1, \dots, ca_n)\end{aligned}$$

y $O = (0, \dots, 0)$.

- (iv) *Espacio de funciones con valor en \mathbb{K}* : Dado un conjunto X , el conjunto \mathbb{K}^X de funciones $X \rightarrow \mathbb{K}$, junto con las operaciones

$$\begin{aligned}f + g : x &\mapsto f(x) + g(x) \\ cf : x &\mapsto cf(x)\end{aligned}$$

y el origen es la función $O : x \mapsto 0$.

- (v) *Espacio de funciones con valor en \mathbb{K} , con casi todos los valores iguales a 0*: Dado un conjunto X , el conjunto \mathbb{K}_0^S de funciones $X \rightarrow \mathbb{K}$ para las que todos los valores, salvo para un número finito de elementos en X , son 0, junto con las operaciones y el origen definidos para \mathbb{K}^X .
- (vi) *Espacio de polinomios con coeficientes en \mathbb{K}* : El conjunto $\mathbb{K}[t]$ de polinomios en la variable t con coeficientes en \mathbb{K} junto con las operaciones de suma y producto por escalar usuales y con el origen O dado por el polinomio constante 0 (ver la definición A.11).

Proposición 1.3. Sea V un espacio vectorial sobre \mathbb{K}

- (i) Ley de cancelación: Dados $u, v, w \in V$, la igualdad $u + v = w + v$ implica $u = w$.
- (ii) Unicidad del origen: Si $o \in V$ es tal que $v + o = v$ para algún $v \in V$ entonces $v = 0$.
- (iii) Unicidad del opuesto: Dado $v \in V$, si $w \in V$ es tal que $v + w = O$, entonces $w = -v$.

Dem.

- (i) A partir de la igualdad $u + v = w + v$, si sumamos $-v$ a ambos lados obtenemos $u = w$.
- (ii) Se sigue de la ley cancelativa aplicada a $v + o = v = v + O$.
- (iii) Se sigue de la ley cancelativa aplicada a $v + w = O = v + (-v)$.

Proposición 1.4. *Sea V es un espacio vectorial sobre \mathbb{K} .*

- (i) *Para todo $c \in \mathbb{K}$ y $v \in V$ tenemos $cO = 0v = O$.*
- (ii) *Para todo $v \in V$ tenemos $(-1)v = -v$.*
- (iii) *Si tenemos $cv = O$, entonces $c = 0$ ó $v = O$.*

Dem.

- (i) Tenemos $cO + O = cO = c(O + O) = cO + cO$, luego, por la ley de cancelación, $cO = O$. Igualmente, tenemos $0v + O = 0v = (0 + 0)v = 0v + 0v$, luego $0v = O$.
- (ii) Por la unicidad del opuesto, basta ver que $v + (-1)v = O$, lo cual se sigue de las igualdades $v + (-1)v = 1v + (-1)v = (1 + (-1))v = 0v$.
- (iii) Suponga que $cv = O$ con $c \neq 0$, entonces tenemos $O = c^{-1}O = c^{-1}(cv) = (c^{-1}c)v = 1v = v$.

□

Definición 1.5. Sea V un espacio vectorial sobre \mathbb{K} y $U \subseteq V$. Decimos que U es un *subespacio* de V si U , con las operaciones heredadas de V , es un espacio vectorial.

Notación 1.6. Usaremos los símbolos \leq , $<$, \geq y $>$ para representar respectivamente *subespacio de*, *subespacio propio de*, *superespacio de* y *superespacio propio de*.

Proposición 1.7 (Caracterización de subespacios). *Sea V un espacio vectorial sobre \mathbb{K} . Si U es un subconjunto de V , entonces $U \leq V$ si y solo si U satisface las siguientes propiedades.*

- (i) $O \in U$.
- (ii) El conjunto U es cerrado bajo suma: *para todo $v, w \in U$, se tiene $v + w \in U$.*
- (iii) El conjunto U es cerrado bajo multiplicación por escalar: *para todo $c \in \mathbb{K}$ y $v \in U$, se tiene $cv \in U$.*

Dem. Suponga primero $U \leq V$. Luego U contiene un neutro respecto a la suma y, como la operación de suma es la de V , este es O . Así tenemos $O \in U$. Por otro lado, como U es un espacio vectorial con las operaciones de V , la suma de dos elementos en U está en U y producto por escalar de un elemento en U

también está en U .

Recíprocamente, si U contiene al origen y es cerrado bajo suma, la restricción de la suma a $U \times U$ da una operación $+: U \times U \rightarrow U$ que cumple con el axioma (i) de la definición 1.1. Similarmente sucede con la restricción del producto por escalar a $\mathbb{K} \times U$ y el axioma (ii). Finalmente, el axioma (iii) de distributividad se hereda de V . \square

Ejemplo 1.8. Las soluciones a sistemas lineales homogéneos son subespacios. De hecho, tome $a_{i1}, \dots, a_{in} \in \mathbb{K}$, para $i \in \{1, \dots, m\}$, con $m \leq n$. El conjunto U de soluciones $(x_1, \dots, x_n) \in \mathbb{K}^n$ al sistema

$$\begin{cases} a_{11}x_1 + \dots + a_{1n}x_n &= 0 \\ &\vdots \\ a_{m1}x_1 + \dots + a_{mn}x_n &= 0 \end{cases}$$

contiene al origen, es cerrado mediante suma y producto por escalar, y es así un subespacio.

Definición 1.9. Sea V un espacio vectorial sobre \mathbb{K} y sean $v_1, \dots, v_n \in V$. Una *combinación lineal* de v_1, \dots, v_n es un vector $v \in V$ tal que

$$v = c_1v_1 + \dots + c_nv_n$$

con $c_1, \dots, c_n \in \mathbb{K}$. A los elementos c_1, \dots, c_n los llamamos los *coeficientes* de la combinación lineal.

Proposición 1.10. Sea V un espacio vectorial sobre \mathbb{K} . Dado un subconjunto no vacío C de V , el conjunto formado por todas las combinaciones lineales de elementos en C es un subespacio de V .

Dem. Sea U el conjunto $\left\{ \sum_{i=1}^n c_i v_i \mid c_1, \dots, c_n \in \mathbb{K}, v_1, \dots, v_n \in C \right\}$. Como C no es vacío, para cualquier $c \in C$ se tiene $0c = O$ y así U contiene al origen. Para todo $v, w \in U$, existen $a_1, \dots, a_n, b_1, \dots, b_m \in \mathbb{K}$ y $v_1, \dots, v_n, w_1, \dots, w_m \in S$ tales que $v = \sum_{i=1}^n a_i v_i$ y $w = \sum_{j=1}^m b_j w_j$, y así

$$v + w = a_1v_1 + \dots + a_nv_n + b_1w_1 + \dots + b_mw_m.$$

Luego, $v + w$ es una combinación lineal de elementos de C y $v + w$ pertenece a U . Similarmente, dado $c \in \mathbb{K}$, tenemos

$$cv = ca_1v_1 + \dots + ca_nv_n,$$

y así cv está en U . Por lo tanto, U es cerrado bajo suma y multiplicación por escalar, luego, por la propiedad 1.7, U es subespacio de V . \square

Definición 1.11. Sea V un espacio vectorial sobre \mathbb{K} y sea C un subconjunto de V no vacío. Al subespacio formado por todas las combinaciones lineales de elementos de C lo llamamos el *espacio generado por C* y lo denotamos por $\langle C \rangle$. Por convención, definimos $\langle \emptyset \rangle = \{O\}$.

Proposición 1.12. Sea V un espacio vectorial sobre \mathbb{K} .

- (i) Si $\{U_i\}_{i \in I}$ es una familia de subespacios de V , entonces la intersección $\bigcap_{i \in I} U_i$ es un subespacio de V .
- (ii) Si C es un subconjunto de V , entonces $\langle C \rangle$ es el mínimo subespacio vectorial en V que contiene a C . Es decir, si W es un subespacio de V que contiene a C , entonces $\langle C \rangle$ es un subespacio de W .

Dem.

- (i) Sea U el conjunto $\bigcap_{i \in I} U_i$. Para todo $i \in I$, el origen O está en U_i y así $O \in U$. Sea $v, w \in U$ y $a \in \mathbb{K}$. Para todo $i \in I$ tenemos $v, w \in U_i$, luego $v + w$ y cv pertenecen a todo U_i y así también están en U .
- (ii) Sea W un subespacio de V que contiene a C . Como W es cerrado bajo suma y bajo multiplicación por escalar, dados $v_1, \dots, v_n \in C$ y $c_1, \dots, c_n \in \mathbb{K}$, la combinación lineal $\sum_i c_i v_i$ está en W . Es decir que toda combinación lineal de elementos de C está en W . Luego, tenemos $\langle C \rangle \subseteq W$. Pero como $\langle C \rangle$ es un espacio vectorial, obtenemos $\langle C \rangle \leq W$. \square

Ejercicios

- Verifique que los conjuntos en el ejemplo 1.2 con sus respectivas operaciones son espacios vectoriales.
- Verifique que el conjunto de funciones continuas de \mathbb{R} en \mathbb{R} es un espacio vectorial sobre \mathbb{R} al igual que el conjunto de funciones de \mathbb{R} en \mathbb{R} diferenciables en todas partes.
- Verifique que el conjunto de funciones racionales $\mathbb{C}(z_1, \dots, z_n)$ es un espacio vectorial sobre \mathbb{C} .
- Explique por qué los conjuntos \mathbb{N} y \mathbb{Z} no son espacios vectoriales.
- ¿Es el conjunto $\{(x, y) \in \mathbb{R}^2 | y > 0\}$ un subespacio de \mathbb{R}^2 ?
- ¿Es el conjunto $\{(x, y, z) \in \mathbb{K}^3 | x + y + z = 0\}$ un subespacio de \mathbb{K}^3 ?
- ¿Es el conjunto $\{(x, -x, 2x) \in \mathbb{K}^3 | x \in \mathbb{R}\}$ un subespacio de \mathbb{K}^3 ?
- Demuestre que el conjunto $\{(x, y, z) \in \mathbb{R}^3 | 2x + y - z = 0, x - y = 0\}$ es un subespacio de \mathbb{Q}^3 y encuentre un conjunto de vectores que los genere.
- Demuestre que el conjunto $\{(x, y, z) \in \mathbb{R}^3 | 2x + y - z = 0\}$ es un subespacio de \mathbb{Q}^3 y encuentre un conjunto de vectores que lo genere.
- Describa geoméricamente el subconjunto $\langle (1, 0, 1), (0, 1, 0) \rangle$ de \mathbb{R}^3 .
- Considere los siguientes subespacio de \mathbb{Q}^3 , $U_1 = \langle (1, 1, 0), (0, 1, 1) \rangle$ y $U_2 = \langle (1, 0, 1), (0, 1, 0) \rangle$. Describa los elementos (x, y, z) de $U_1 \cap U_2$ como soluciones de un sistema y encuentre un conjunto generador.

1.2. Base y dimensión

Dado un conjunto de índices I y un conjunto C , una *familia de elementos de C indexada por I* , o simplemente una *familia de C* , es una función $I \rightarrow C$ que denotaremos por $(x_i)_{i \in I}$ cuando la imagen de i es x_i . Cuando no haya lugar a confusión, regularmente identificaremos esta familia con el subconjunto $\{x_i\}_{i \in I}$ de C .

Sea V un espacio vectorial sobre \mathbb{K} .

Notación 1.13. Sea $C = (v_i)_{i \in I}$ una familia de vectores de V indexados por un conjunto I y $(c_i)_{i \in I}$ una familia de escalares indexados por I , tales que $c_i = 0$ para todo índice $i \in I$, salvo para una subcolección finita, es decir que $(c_i)_{i \in I} \in \mathbb{K}_0^I$ (ver el ejemplo 1.2(v)). A la combinación lineal de elementos de C con coeficientes c_i la denotaremos por $\sum_{i \in I} c_i v_i$ y a c_i lo llamamos el i -ésimo coeficiente. Note que estas sumas son finitas pues los coeficientes que no son 0 son finitos. Cuando I es un conjunto finito de n elementos, lo usual es tomar $I = \{1, \dots, n\}$.

Definición 1.14. Sea $\mathcal{B} = (v_i)_{i \in I}$ una familia de vectores en V . Decimos que \mathcal{B} es una *base* de V si para cada $v \in V$ existe una única combinación lineal de elementos de C igual a v . Si $v = \sum_{i \in I} c_i v_i$, al coeficiente c_i lo llamamos la *coordenada i de v en la base \mathcal{B}* . Por convención, la base del espacio cero-dimensional es \emptyset .

Lema 1.15. Suponga que V no es el espacio cero-dimensional y sea $\mathcal{B} = (v_i)_{i \in I}$ una base. Si $c_i, i \in I$, son escalares para los cuales $0 = \sum_{i \in I} c_i v_i$, entonces $c_i = 0$ para todo $i \in I$.

Dem. La combinación lineal $\sum_{i \in I} 0v_i$ es la única combinación lineal de elementos en \mathcal{B} igual al origen, entonces $c_i = 0$ para todo i . \square

Definición 1.16. Decimos que V tiene dimensión finita si tiene una base finita. De lo contrario decimos que V tiene dimensión infinita.

Teorema 1.17 (Teorema de la dimensión). *Si V tiene dimensión finita, el número de elementos de la base es independiente de la base escogida.*

Dem. La afirmación para el caso $V = \{0\}$, se sigue del hecho de que su única base es de cero elementos. Tome V diferente del espacio cero-dimensional y suponga por contradicción que V tiene dos bases finitas (v_1, \dots, v_n) y (w_1, \dots, w_m) de diferentes tamaños, es decir $n \neq m$. Sin pérdida de generalidad, podemos asumir que $m > n$.

Para $j \in \{1, \dots, m\}$, tome escalares a_{1j}, \dots, a_{nj} tales que

$$w_j = \sum_{i=1}^n a_{ij} v_i.$$

Así, dado $v \in V$ con $v = x_1 w_1 + \dots + x_m w_m$ y $x_1, \dots, x_m \in \mathbb{K}$, tendríamos

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

Por el lema anterior, se cumple $v = O$ si y solo si

$$\sum_{j=1}^m a_{ij}x_j = 0,$$

para $i \in \{1, \dots, n\}$. Estas n igualdades forman un sistema homogéneo subdeterminado, es decir con más variables que ecuaciones. En particular, este sistema tiene al menos una solución no trivial. Si (x_1, \dots, x_m) es una de ellas, entonces obtenemos el origen como una combinación lineal de los elementos de la base $\{w_1, \dots, w_m\}$ con coeficientes no todos iguales a cero, lo cual contradice el lema. Por lo tanto tenemos $m = n$.

Cuando asumimos por contradicción que V admite una base infinita $(w_i)_{i \in I}$, y tomamos $v \in V$, entonces existe una combinación lineal $v = x_1w_1 + \dots + x_nw_n$ de vectores $w_1, \dots, w_n \in \{w_i\}_{i \in I}$ y obtenemos una contradicción similarmente a como la acabamos de obtener. \square

Definición 1.18. Suponga que V tiene dimensión finita. Al número de elementos en una base de V lo llamamos dimensión de V y lo denotamos por $\dim_{\mathbb{K}}(V)$, ó $\dim(V)$. Si V tiene dimensión infinita, escribimos $\dim(V) = \infty$, y usamos la convención de que $n < \infty$ para todo entero n .

Definición 1.19. Sea C un subconjunto de V . Decimos que C es *linealmente dependiente* si algún elemento de C es combinación lineal de los otros, es decir si existen $v_0, v_1, \dots, v_n \in C$, dos a dos distintos, tales que

$$v_0 = c_1v_1 + \dots + c_nv_n$$

para algunos escalares c_1, \dots, c_n . Si C no es linealmente dependiente, decimos que S es *linealmente independiente*. Por convención diremos que \emptyset es linealmente independiente.

Proposición 1.20 (Caracterización de independencia lineal). *Un conjunto C de vectores es linealmente independiente si y solo si la única combinación lineal de vectores en C igual al origen es la que todos sus coeficientes son iguales a cero.*

Dem. Establecemos la contrapositiva: C es linealmente dependiente si y solo si existen $v_1, \dots, v_n \in C$ y $c_1, \dots, c_n \in \mathbb{K}$, con $c_j \neq 0$ para algún $j \in \{1, \dots, n\}$, tales que

$$O = c_1v_1 + \dots + c_nv_n.$$

Lo cual se sigue al notar que esta igualdad equivale a

$$v_j = \sum_{i=1, i \neq j}^n (-c_i/c_j)v_i$$

cuando c_j es invertible, es decir cuando $c_j \neq 0$ para algún $j \in \{1, \dots, n\}$. \square

Teorema 1.21 (Caracterización de las bases). *Sea $\mathcal{B} = (v_i)_{i \in I}$ una familia de vectores de V indexada por I , con $I \neq \emptyset$. Entonces \mathcal{B} es una base de V si y solo si satisface las siguientes dos propiedades.*

- (i) *El conjunto \mathcal{B} es linealmente independiente.*
- (ii) *El espacio V es generado por \mathcal{B} .*

Dem. Es suficiente demostrar que si tenemos $\langle \mathcal{B} \rangle = V$ entonces \mathcal{B} es una base si y solo si \mathcal{B} es linealmente independiente. Con esto en mente, suponga primero que \mathcal{B} es base, luego por el lema 1.15, \mathcal{B} es linealmente independiente. Recíprocamente, suponga que \mathcal{B} es linealmente independiente y asuma por contradicción que \mathcal{B} no es base. Luego, existen dos combinaciones lineales con coeficientes distintos, ambas iguales a algún $v \in V$. La diferencia de estas dos combinaciones lineales daría una combinación lineal igual a O con no todos los coeficientes nulos, lo cual contradice la propiedad 1.20. \square

Ejemplo 1.22. Por la propiedad 1.21, se puede verificar que, para cada uno de los siguientes espacios V , el respectivo conjunto \mathcal{B} es una base.

- (i) En $V = K^n$, para $i \in \{1, \dots, n\}$, defina $e_i \in V$ como el elemento con ceros en todas las entradas salvo en la i -ésima, donde tiene 1, es decir

$$e_1 = (1, 0, \dots, 0), \quad e_2 = (0, 1, \dots, 0), \quad e_n = (0, 0, \dots, 1).$$

Tome $\mathcal{B} = (e_1, \dots, e_n)$.

- (ii) Sea I un conjunto y $V = (K^I)_0$ (ver Ejemplo 1.2.5). Para $i \in I$, defina la función $\delta_i : I \rightarrow K$ por $\delta_i(i) = 1$ y $\delta_i(j) = 0$ para $j \neq i$. Tome $\mathcal{B} = (\delta_i)_{i \in I}$.
- (iii) Para $V = K[t]$, sea $\mathcal{B} = (1, t, t^2, \dots) = (t^n)_{n \in \mathbb{N}}$.

Notación 1.23. Para K^n , la *base canónica* es la base (e_1, \dots, e_n) definida en el ejemplo 1.22.(i), que denotaremos por \mathcal{C}_n , o simplemente \mathcal{C} .

Lema 1.24. *Sea C un subconjunto de V linealmente independiente. Si $v \in V$ es tal que $C \cup \{v\}$ es linealmente dependiente, entonces v pertenece a $\langle C \rangle$.*

Dem. Como $C \cup \{v\}$ es linealmente dependiente, por la proposición 1.20 existen escalares c_1, \dots, c_n, c_{n+1} , no todos iguales a cero, y $v_1, \dots, v_n \in C$, dos a dos distintos, tales que

$$O = a_1 v_1 + \dots + a_n v_n + a_{n+1} v.$$

Note que a_{n+1} es diferente de 0, o de lo contrario tendríamos una combinación lineal de v_1, \dots, v_n igual a O , con no todos los coeficientes iguales a cero, lo que contradiría la independencia lineal de C . Obtenemos

$$v = \sum_{i=1}^n (-a_i/a_{n+1}) v_i \in \langle v_1, \dots, v_n \rangle,$$

y por lo tanto v pertenece a $\langle v_1, \dots, v_n \rangle$ que está contenido en $\langle C \rangle$. \square

Proposición 1.25. *Suponga que V tiene dimensión finita distinta de cero y sea C un subconjunto no vacío finito de V . Si C_0 es un subconjunto linealmente independiente de C , entonces existe $C' \subseteq C$ maximal entre los subconjuntos linealmente independientes de C que contienen a C_0 . Es decir, dado un subconjunto C_1 de C que contiene a C' , si C_1 es linealmente independiente, entonces $C_1 = C'$, y si C_1 contiene propiamente a C' , entonces C_1 es linealmente dependiente. Más aún se tiene $\langle C' \rangle = \langle C \rangle$.*

Dem. Sean v_1, \dots, v_n los elementos de C , enumerados de forma tal que $\{v_1, \dots, v_m\}$ sea C_0 . Definimos $m = 0$ cuando $C_0 = \emptyset$. Tenemos $m \leq n$. Iterativamente, para $i \in \{1, \dots, n - m\}$, definimos

$$C_i = \begin{cases} C_{i-1} & \text{si } v_{m+i} \in \langle C_{i-1} \rangle \\ C_{i-1} \cup \{v_{m+i}\} & \text{si } v_{m+i} \notin \langle C_{i-1} \rangle \end{cases}.$$

Por el lema anterior, la independencia lineal de C_{i-1} implica la de C_i . Tome $C' = C_{n-m}$. En particular, C' es linealmente independiente. Por construcción, tenemos $C_0 \subseteq C_1 \subseteq \dots \subseteq C_{n-m} = C'$. Para ver la maximalidad de C' , tomemos $j \in \{1, \dots, n\}$. Si v_j no pertenece a C , esto quiere decir que v_j está en $\langle C_{j-1} \rangle$, el cual es un subconjunto de $\langle C' \rangle$, y así $C' \cup \{v_j\}$ es linealmente dependiente. Luego $C' \subset C$ es maximal respecto a las propiedades de contener a C_0 y ser linealmente independiente. El mismo argumento demuestra la contención de C en $\langle C' \rangle$, y la proposición 1.12.2. implica que $\langle C \rangle \leq \langle C' \rangle$. Para la contención opuesta, note que $C' \subseteq C$, y por lo tanto $\langle C' \rangle \leq \langle C \rangle$, y así $\langle C' \rangle = \langle C \rangle$. \square

Teorema 1.26 (Extensión a una base). *Suponga que V tiene dimensión finita. Si \mathcal{B}_0 es una familia finita de V linealmente independiente, entonces existe una base \mathcal{B} de V que contiene a \mathcal{B}_0 .*

Dem. Como V tiene dimensión finita, existe una base finita \mathcal{B}_1 de V . Denote C al conjunto $\mathcal{B}_0 \cup \mathcal{B}_1$. Por la propiedad anterior, existe un subconjunto \mathcal{B} de C linealmente independiente maximal que contiene a \mathcal{B}_0 . Tenemos $\langle \mathcal{B} \rangle = \langle C \rangle \geq \langle \mathcal{B}_1 \rangle = V$. Así \mathcal{B} forma una familia linealmente independiente que genera a V , y luego, por la proposición 1.21, \mathcal{B} forma una base. \square

Lema 1.27. *Si V tiene dimensión infinita, entonces para todo entero $n \geq 1$, existe un subconjunto C de V linealmente independiente con n elementos.*

Dem. Hacemos inducción en n . Como V tiene dimensión infinita, tenemos $V \neq \{O\}$, y así, para cualquier $v \in V$ distinto de O , el conjunto $\{v\}$ es linealmente independiente y tiene 1 elemento. Esto establece el caso base $n = 1$. Para el paso inductivo, suponga que tenemos un subconjunto linealmente independiente $\{v_1, \dots, v_n\}$ en V . Como V tiene dimensión infinita, (v_1, \dots, v_n) no es una base de V , luego tampoco lo genera y por ende existe $v_{n+1} \in V$ fuera de $\langle v_1, \dots, v_n \rangle$. Por lo tanto, el lema 1.24 implica que el conjunto $\{v_1, \dots, v_n, v_{n+1}\}$ es linealmente independiente. \square

Teorema 1.28 (Monotonía de la dimensión). *Si U es un subespacio de V , entonces tenemos $\dim U \leq \dim V$. Más aún, si V tiene dimensión finita, entonces $\dim U = \dim V$ si y solo si $U = V$.*

Dem. Suponga primero que $\dim(U) = \infty$. Asuma por contradicción que V tiene dimensión finita. Por el lema anterior, existe una familia \mathcal{B}_0 de U finita y linealmente independiente con $\dim(V) + 1$ elementos. Por el teorema 1.26 de extensión a una base, existe una base \mathcal{B}_1 de V que extiende a \mathcal{B}_0 . Luego, V tiene una base con más elementos que su dimensión, contradiciendo el teorema 1.17. Por lo tanto, tenemos $\dim(V) = \infty$, y así $\dim(U) \leq \dim(V)$.

Suponga ahora que U tiene dimensión finita. Si V tiene dimensión infinita, obtenemos $\dim(U) \leq \dim(V)$. Ahora asumamos que V tiene dimensión finita. Tome una base \mathcal{B}_0 de U , la cual, nuevamente por el teorema 1.26, podemos extender a una base \mathcal{B}_1 de V . Tenemos entonces $|\mathcal{B}_0| \leq |\mathcal{B}_1|$, es decir $\dim(U) \leq \dim(V)$. Finalmente, suponga que V tiene dimensión finita. Evidentemente, si $U = V$, entonces $\dim U = \dim V$. Recíprocamente, si $\dim(U) = \dim(V)$, entonces toda base \mathcal{B}_0 de U es también una base de V . De hecho, ya vimos que una base de U puede ser extendida a una base de V , pero si esta extensión contiene más elementos, la dimensión de V sería mayor a la de U , contradiciendo la hipótesis $\dim(U) = \dim(V)$. Obtenemos así $U = \langle \mathcal{B}_0 \rangle = V$. \square

Observación 1.29 (Existencias de bases y dimensión infinita). El teorema 1.26 de extensión a una base se puede usar para demostrar que, partiendo de un conjunto vacío, todo espacio vectorial de dimensión finita tiene una base. La cuestión para dimensión infinita toca la fibra de los fundamentos de la matemática y, para demostrar la existencia de una base, requiere admitir el axioma de elección. Siendo más precisos, usamos una versión equivalente a este.

Lema de Zorn. Si (P, \preceq) es un conjunto con un orden parcial en el que toda cadena admite una cota superior, entonces P contiene al menos un elemento maximal.

A partir de este axioma, suponga que V tiene dimensión infinita y tome un familia \mathcal{B}_0 de V linealmente independiente. Sea P el conjunto de familias de V linealmente independientes que contienen a \mathcal{B}_0 , el cual ordenamos por contención. Dada una cadena en P , la unión de todos sus elementos también está en P y es una cota superior de ella. Por el lema de Zorn, P contiene un maximal \mathcal{B}_1 . Usando un argumento similar al del lema 1.24, se demuestra que \mathcal{B}_1 es una base de V . De hecho, si existe $v \in V$ fuera de $\langle \mathcal{B}_1 \rangle$, con el conjunto $\mathcal{B}_1 \cup \{v\}$ formamos una familia linealmente independiente que contiene estrictamente a \mathcal{B}_1 y a \mathcal{B}_0 , contradiciendo la maximalidad de \mathcal{B}_1 en P .

De esta forma todo espacio vectorial tiene una base y aún más, en cualquier espacio vectorial, todo conjunto linealmente independiente se puede extender a una.

Similarmente podemos extender la proposición 1.25 para concluir que si dentro de un subconjunto C de V , tomamos un subconjunto C_0 linealmente independiente, entonces existe un subconjunto C' de C linealmente independiente maximal que contiene a C_0 , y para un tal C' se tiene $\langle C \rangle = \langle C' \rangle$. De hecho, tomamos, usando Lema de Zorn, un maximal C' en la colección, ordenada por inclusión, de subconjunto linealmente independientes de C que contienen a C_0 . El lema 1.24 implica la igualdad $\langle C \rangle = \langle C' \rangle$.

Ejercicios

1. Determine si el conjunto $\{(1, 2), (1, 3), (1, -1)\}$ es linealmente dependiente en \mathbb{R}^2 , y en tal caso exprese un vector como combinación lineal de los otros.
2. Determine si el conjunto $\{(1, 2, 3), (2, 3, 4), (3, 4, 5)\}$ es linealmente dependiente en \mathbb{Q}^3 , y en tal caso exprese un vector como combinación lineal de los otros.
3. Determine si el conjunto $\{(1, 0, 0, 1), (1, 0, 1, 0), (0, 1, 1, 0), (0, 1, 0, 1)\}$ es linealmente dependiente en \mathbb{K}^4 , y en tal caso exprese un vector como combinación lineal de los otros.
4. Encuentre una base del subespacio $\{(x, y, z) \in \mathbb{R}^3 \mid 2x + y - z = 0, x - y = 0\}$ de \mathbb{R}^3 .
5. Encuentre una base del subespacio $\{(x, y, z) \in \mathbb{R}^3 \mid 2x + y - z = 0\}$ de \mathbb{Q}^3 .
6. Para cada uno de las siguientes conjuntos de vectores encuentre un subconjunto linealmente independiente maximal.
 - a) $\{(1, 0, -2, 1), (0, -1, 1, 1), (-1, 2, 1, 0), (2, 1, -4, 4)\} \subseteq \mathbb{Q}^4$
 - b) $\{(1, 0, -2, 1), (0, -1, 1, 1), (-1, 2, 1, 0), (2, 1, 2, -2)\} \subseteq \mathbb{Q}^4$
 - c) $\{(1, -1, 5, -8, 6), (-1, 1, -5, 5, -3), (1, 0, 3, -3, 5), (2, 3, 4, -1, 1), (0, 1, 0, -1, 2)\} \subseteq \mathbb{Q}^5$
 - d) $\{(1, -1, 5, -8, 6), (-1, 1, -5, 5, -3), (1, 0, 3, -3, 5), (2, 1, 4, 8, 0)\} \subseteq \mathbb{Q}^5$
 - e) $\{(1, 1, 0, 0, 0, 0), (0, 0, 1, 1, 0, 0), (0, 0, 0, 0, 1, 1), (1, 0, 1, 0, 1, 0), (0, 1, 0, 1, 0, 1)\} \subseteq \mathbb{Q}^6$

1.3. Transformaciones lineales

Sean U , V y W espacios vectoriales sobre \mathbb{K} .

Definición 1.30. Sea $f : V \rightarrow W$ una función. Decimos que f es una *transformación lineal* (o un *morfismo de espacios vectoriales*) si satisface las siguientes propiedades.

- (i) *Preserva sumas:* Para todo $v_1, v_2 \in V$ se tiene $f(v_1 + v_2) = f(v_1) + f(v_2)$.
- (ii) *Preserva productos por escalar:* Para todo $c \in \mathbb{K}$ y todo $v \in V$ se tiene $f(cv) = cf(v)$.

Al conjunto de transformaciones lineales de V en W lo denotamos por $\text{Hom}_{\mathbb{K}}(V, W)$. A las transformaciones lineales de un espacio en él mismo las llamamos *operadores* (o *endomorfismos de espacios vectoriales*). Al conjunto $\text{Hom}_{\mathbb{K}}(V, V)$ de operadores de V lo denotamos por $\text{End}_{\mathbb{K}}(V)$.

Ejemplo 1.31. (i) *Transformación lineal cero:* La función $\underline{O} : V \rightarrow W$ definida por $\underline{O}(v) = O$ para todo v es una transformación lineal.

(ii) *Operador identidad:* La función $\text{id}_V : V \rightarrow V$ definida por $\text{id}_V(v) = v$ para todo v es un operador.

Proposición 1.32. Para toda $f \in \text{Hom}_{\mathbb{K}}(V, W)$ se tiene

$$(i) \quad f(O) = O,$$

$$(ii) \quad f(-v) = -f(v) \text{ para todo } v \in V, \text{ y}$$

$$(iii) \quad f(c_1v_1 + \cdots + c_nv_n) = c_1f(v_1) + \cdots + c_nf(v_n) \text{ para todo } c_1, \dots, c_n \in K \text{ y todo } v_1, \dots, v_n \in V.$$

Dem.

$$(i) \quad \text{Tenemos } f(O) = f(0O) = 0f(O) = O.$$

(ii) Dado $v \in V$, se tiene $f(v) + f(-v) = f(v + (-v)) = f(O) = O$, y así, por la unicidad del opuesto, $f(-v) = -f(v)$.

(iii) Usaremos inducción en n , siendo el caso base, $n = 2$, cierto por el axioma (ii) en la definición 1.30 de transformación lineal. Ahora, si asumimos que la propiedad es cierta para n , entonces para todo $c_1, \dots, c_{n+1} \in K$ y todo $v_1, \dots, v_{n+1} \in V$ tenemos

$$\begin{aligned} f(c_1v_1 + \cdots + c_nv_n + c_{n+1}v_{n+1}) &= f(c_1v_1 + \cdots + c_nv_n) + f(c_{n+1}v_{n+1}) \\ &= c_1f(v_1) + \cdots + c_nf(v_n) + c_{n+1}f(v_{n+1}). \end{aligned}$$

Por lo tanto, es cierto para $n + 1$ y la propiedad se sigue por inducción. □

Observación 1.33. Dados $f, g \in \text{Hom}_{\mathbb{K}}(V, W)$ y $c \in K$ definimos las transformaciones lineales $f + g$ y cg en $\text{Hom}_K(V, W)$ por $(f + g)(v) = f(v) + g(v)$ y $(cg)(v) = cg(v)$ para todo $v \in V$. El conjunto $\text{Hom}_{\mathbb{K}}(V, W)$ junto con estas operaciones y el origen dado por \underline{O} es un espacio vectorial sobre \mathbb{K} .

Proposición 1.34. Si $f \in \text{Hom}_{\mathbb{K}}(V, W)$ y $g \in \text{Hom}_{\mathbb{K}}(W, U)$, entonces $(g \circ f) \in \text{Hom}_{\mathbb{K}}(V, U)$.

Dem. Para todo $u, v \in V$ tenemos

$$g \circ f(u + v) = g(f(u) + f(v)) = g \circ f(u) + g \circ f(v).$$

Para todo $v \in V$ y todo $c \in \mathbb{K}$, tenemos

$$g \circ f(cv) = g(cf(v)) = cg \circ f(v).$$

□

Proposición 1.35 (Rigidez de las transformaciones lineales). Sean $v_1, \dots, v_n \in V$ tales que $\langle v_1, \dots, v_n \rangle = V$. Si w_1, \dots, w_n son elementos en W , entonces tenemos las siguientes dos propiedades.

- (i) Existe a lo sumo una transformación lineal $f \in \text{Hom}_{\mathbb{K}}(V, W)$ tal que $f(v_i) = w_i$, para $i \in \{1, \dots, n\}$.
- (ii) Si $\{v_1, \dots, v_n\}$ es linealmente independiente, entonces existe una transformación $f \in \text{Hom}_{\mathbb{K}}(V, W)$ tal que $f(v_i) = w_i$, para $i \in \{1, \dots, n\}$.

Dem.

- (i) Considere dos transformaciones lineales $f, g \in \text{Hom}_{\mathbb{K}}(V, W)$ tales que $f(v_i) = w_i = g(v_i)$, para $i \in \{1, \dots, n\}$. Dado $v \in V$, existen $c_1, \dots, c_n \in \mathbb{K}$ para los cuales $v = c_1 v_1 + \dots + c_n v_n$ y así

$$f(v) = f\left(\sum_{i=1}^n c_i v_i\right) = \sum_{i=1}^n c_i f(v_i) = \sum_{i=1}^n c_i w_i = \sum_{i=1}^n c_i g(v_i) = g(v).$$

Por lo tanto $f = g$.

- (ii) Si $\{v_1, \dots, v_n\}$ es linealmente independiente, por la proposición 1.21, se sigue que (v_1, \dots, v_n) es una base de V . Dado $v \in V$, existe un único elemento $(c_1, \dots, c_n) \in \mathbb{K}^n$ para el que se tiene $v = c_1 v_1 + \dots + c_n v_n$. Defina la función $f : V \rightarrow W$ por

$$f(v) = c_1 w_1 + \dots + c_n w_n.$$

Veamos que f es una transformación lineal. Dados $u, v \in V$, para (a_1, \dots, a_n) y (b_1, \dots, b_n) en \mathbb{K}^n tales que $u = a_1 v_1 + \dots + a_n v_n$ y $v = b_1 v_1 + \dots + b_n v_n$, obtenemos $u + v = (a_1 + b_1)v_1 + \dots + (a_n + b_n)v_n$ y

$$\begin{aligned} f(u + v) &= (a_1 + b_1)w_1 + \dots + (a_n + b_n)w_n \\ &= a_1 w_1 + \dots + a_n w_n + b_1 w_1 + \dots + b_n w_n \\ &= f(u) + f(v). \end{aligned}$$

Luego, f respecta sumas. Dado $v \in V$, si se tiene $v = c_1 v_1 + \dots + c_n v_n$ con $(c_1, \dots, c_n) \in \mathbb{K}^n$, entonces, para todo $c \in \mathbb{K}$, $cv = (ca_1)v_1 + \dots + (ca_n)v_n$ y

$$\begin{aligned} f(cv) &= (ca_1)w_1 + \dots + (ca_n)w_n \\ &= c(a_1 w_1 + \dots + a_n w_n) \\ &= cf(v). \end{aligned}$$

Por lo tanto, f también respecta productos por escalar y se sigue que f es una transformación lineal.

□

Proposición 1.36. Sea $f \in \text{Hom}_{\mathbb{K}}(V, W)$. Si f es biyectiva, entonces f^{-1} es una transformación lineal.

Dem. Sean $w_1, w_2 \in W$. Como f es sobreyectiva, tenemos $f(v_1) = w_1$ y $f(v_2) = w_2$ para algunos $v_1, v_2 \in V$ y, así, $f(v_1 + v_2) = f(v_1) + f(v_2) = w_1 + w_2$. Se sigue que

$$f^{-1}(w_1 + w_2) = f^{-1}(f(v_1 + v_2)) = v_1 + v_2,$$

y

$$f^{-1}(w_1) + f^{-1}(w_2) = f^{-1}(f(v_1)) + f^{-1}(f(v_2)) = v_1 + v_2.$$

Luego, para todo $w_1, w_2 \in W$ tenemos $f^{-1}(w_1 + w_2) = f^{-1}(w_1) + f^{-1}(w_2)$. Sea $w \in W$. Dado $v \in V$ tal que $f(v) = w$, vemos que $f(cv) = cf(v) = cw$ para todo $c \in \mathbb{K}$. Por ende,

$$f^{-1}(cw) = f^{-1}(f(cv)) = cv,$$

y

$$cf^{-1}(w) = cf^{-1}(f(v)) = cv.$$

Luego, para todo $w \in W$ y todo $c \in \mathbb{K}$ se tiene que $f^{-1}(cw) = cf^{-1}(w)$. Por lo tanto, f^{-1} respeta sumas y productos por escalar, y por consiguiente f^{-1} es una transformación lineal.

Definición 1.37. Sea $f \in \text{Hom}_{\mathbb{K}}(V, W)$. Si f es una biyección, entonces decimos que f es un *isomorfismo*. Si existe un isomorfismo $f \in \text{Hom}_{\mathbb{K}}(V, W)$ decimos que V y W son *isomorfos* y lo denotamos por $V \simeq_{\mathbb{K}} W$.

Ejemplo 1.38. Suponga que V es unidimensional. Dado $\lambda \in \mathbb{K}$, la función $f : V \rightarrow V$ definida por $f(v) = \lambda v$ para todo v es una transformación lineal. Recíprocamente, si $f : V \rightarrow V$ es una transformación lineal, existe $\lambda \in \mathbb{K}$ tal que $f(v) = \lambda v$ para todo v . De hecho, si $v_0 \neq 0$ entonces $V = \langle v_0 \rangle$ así $f(v_0) = \lambda v_0$ para algún $\lambda \in \mathbb{K}$. Ahora, dado cualquier $v \in V$ existe $c \in \mathbb{K}$ tal que $cv_0 = v$, y tenemos que

$$f(v) = cf(f(v_0)) = c\lambda v_0 = \lambda cv_0 = \lambda v.$$

Entonces, si $\Phi : \mathbb{K} \rightarrow \text{End}_{\mathbb{K}}(V)$ es la función definida por $\Phi(\lambda) = m_{\lambda}$, con m_{λ} definida por $m_{\lambda}(v) = \lambda v$ para todo $v \in V$, vemos que Φ es un isomorfismo y $\text{End}_{\mathbb{K}}(V) \simeq_{\mathbb{K}} \mathbb{K}$.

Definición 1.39. Sea $f \in \text{End}_{\mathbb{K}}(V)$. Si f es una biyección, entonces decimos que f es un *automorfismo*. Al conjunto de automorfismos de V lo denotamos por $\text{GL}_{\mathbb{K}}(V)$.

Teorema 1.40. Si que V y W tienen dimensión finita, entonces $V \simeq_{\mathbb{K}} W$ si y solo si $\dim(V) = \dim(W)$.

Dem. Primero establecemos la necesidad. Suponga que $V \simeq_{\mathbb{K}} W$, luego existe un isomorfismo $f \in \text{Hom}_{\mathbb{K}}(V, W)$. Tome (v_1, \dots, v_n) una base de V , donde $n = \dim(V)$. Para $i \in \{1, \dots, n\}$, defina $w_i = f(v_i)$. Veamos que (w_1, \dots, w_n)

es base de W , y así obtener que $\dim(W) = n = \dim(V)$. Si $c_1, \dots, c_n \in \mathbb{K}$ son tales que $c_1 w_1 + \dots + c_n w_n = O$, entonces

$$f(c_1 v_1 + \dots + c_n v_n) = c_1 f(v_1) + \dots + c_n f(v_n) = c_1 w_1 + \dots + c_n w_n = O.$$

Como f es inyectiva y $f(O) = O$, se sigue que $c_1 v_1 + \dots + c_n v_n = O$, pero al ser $\{v_1, \dots, v_n\}$ linealmente independiente, entonces $c_1 = \dots = c_n = 0$. De esta forma, el conjunto $\{w_1, \dots, w_n\}$ es linealmente independiente. Veamos ahora que es un conjunto generador. Sea $w \in W$. Como f es sobreyectiva, existe $v \in V$ tal que $f(v) = w$. Como $\{v_1, \dots, v_n\}$ genera V , existen $c_1, \dots, c_n \in \mathbb{K}$ tales que $v = c_1 v_1 + \dots + c_n v_n$. Luego

$$w = f(v) = f(c_1 v_1 + \dots + c_n v_n) = c_1 f(v_1) + \dots + c_n f(v_n) = c_1 w_1 + \dots + c_n w_n.$$

De donde $W = \langle w_1, \dots, w_n \rangle$ y así $\{w_1, \dots, w_n\}$ es una base de W . Con esto completamos la prueba de la necesidad.

Ahora establecemos la suficiencia. Suponga que $\dim(V) = \dim(W) = n$ y tome dos bases $\{v_1, \dots, v_n\}$ y $\{w_1, \dots, w_n\}$ respectivamente de V y W . Por la proposición 1.35.(ii), existe $f \in \text{Hom}(V, W)$ tal que $f(v_i) = w_i$, para $i \in \{1, \dots, n\}$. La prueba estará completa cuando establezcamos que f es biyectiva. Sean $u, v \in V$ tales que $f(u) = f(v)$. Existen (a_1, \dots, a_n) y (b_1, \dots, b_n) en \mathbb{K}^n tales que $u = a_1 v_1 + \dots + a_n v_n$ y $v = b_1 v_1 + \dots + b_n v_n$. Como $f(u - v) = f(u) - f(v) = O$ y

$$u - v = (a_1 - b_1)v_1 + \dots + (a_n - b_n)v_n,$$

entonces

$$O = f(u - v) = (a_1 - b_1)w_1 + \dots + (a_n - b_n)w_n.$$

La independencia lineal de $\{w_1, \dots, w_n\}$ implica que, para $i \in \{1, \dots, n\}$, $a_i - b_i = 0$, o equivalentemente, $a_i = b_i$. De esta forma tenemos que $u = v$ y se sigue que f es inyectiva. Sea $w \in W$ y sea $(c_1, \dots, c_n) \in \mathbb{K}^n$ tal que $w = c_1 w_1 + \dots + c_n w_n$. Al tomar $v = c_1 v_1 + \dots + c_n v_n$, se obtiene

$$f(v) = f(c_1 v_1 + \dots + c_n v_n) = c_1 f(v_1) + \dots + c_n f(v_n) = c_1 w_1 + \dots + c_n w_n = w.$$

Por lo cual, f es sobreyectiva. Al ser inyectiva y sobreyectiva, f es biyectiva. \square

Definición 1.41. Sea $f \in \text{Hom}_{\mathbb{K}}(V, W)$.

- (i) El *núcleo* (o el *kernel*) de f es el conjunto $\{v \in V \mid f(v) = O\}$ y lo denotamos por $\ker(f)$.
- (ii) La *imagen* de f es el conjunto $\{w \in W \mid \exists v \in V : f(v) = w\}$ y la denotamos por $\text{im}(f)$.

Proposición 1.42. Sea $f \in \text{Hom}_{\mathbb{K}}(V, W)$, entonces $\ker(f)$ e $\text{im}(f)$ son respectivamente subespacios de V y W .

Dem. Para la demostración, mostramos que los dos conjuntos contienen al origen y son cerrados mediante suma y producto por escalar. Como $f(O) = O$, entonces $O \in \ker(f)$ y $O \in \operatorname{im}(f)$. Para todo $v_1, v_2 \in \ker(f)$, tenemos $f(v_1 + v_2) = f(v_1) + f(v_2) = O + O = O$ y así $u + v \in \ker(f)$. Para todo $v \in \ker(f)$ y todo $c \in \mathbb{K}$, $f(cv) = cf(v) = cO = O$ y así $cv \in \ker(f)$. Luego, de la propiedad 1.7, se sigue que $\ker(f)$ es un subespacio de V . Para todo $w_1, w_2 \in \operatorname{im}(f)$ existen $v_1, v_2 \in V$ tales que $f(v_1) = w_1$ y $f(v_2) = w_2$. Luego $f(v_1 + v_2) = f(v_1) + f(v_2) = w_1 + w_2$ y así $w_1 + w_2 \in \operatorname{im}(f)$. Para todo $w \in \operatorname{im}(f)$ existe $v \in V$ tal que $f(v) = w$. Luego, para todo $c \in \mathbb{K}$, $f(cv) = cf(v) = cw$ y así $cv \in \operatorname{im}(f)$. Luego, de la propiedad 1.7, se sigue que $\operatorname{im}(f)$ es un subespacio de V . \square

Proposición 1.43. Sea $f \in \operatorname{Hom}_{\mathbb{K}}(V, W)$, entonces f es inyectiva si y solo si $\ker(f) = \{O\}$.

Dem. Suponga primero que f es inyectiva, entonces, como $f(O) = O$, tenemos $\ker(f) = \{O\}$. Suponga ahora que $\ker(f) = \{O\}$. Si $u, v \in V$ son tales que $f(u) = f(v)$, entonces

$$f(u - v) = f(u) - f(v) = O,$$

De donde $u - v \in \ker(f)$, luego $u - v = O$, o equivalentemente $u = v$, y así, f es inyecta. \square

Definición 1.44. Suponga que V tiene dimensión finita. Sea $f \in \operatorname{Hom}_{\mathbb{K}}(V, W)$.

- (i) La *nulidad* de f , que denotamos $\nu(f)$ es la dimensión de $\ker(f)$.
- (ii) El *rango* de f , que denotamos $\rho(f)$ es la dimensión de $\operatorname{im}(f)$.

Teorema 1.45 (Teorema del rango). Suponga que V tiene dimensión finita. Para todo $f \in \operatorname{Hom}_{\mathbb{K}}(V, W)$, se tiene

$$\nu(f) + \rho(f) = \dim(V)$$

Dem. Note que para el caso $\rho(f) = 0$ el teorema se sigue inmediatamente. Asumamos que $\rho(f) > 0$. Como $\ker(f) \leq V$, por la monotonía de la dimensión tenemos $\nu(f) = \dim(\ker(f)) \leq \dim(V)$. Sean $n = \nu(f)$ y $n + m = \dim(V)$. Sea $\mathcal{B}_0 = (v_1, \dots, v_n) \subseteq V$ una base de $\ker(f)$ (si $n = 0$ tomamos $\mathcal{B}_0 = \emptyset$). Extendemos \mathcal{B}_0 a una base $\mathcal{B} = (v_1, \dots, v_n, v_{n+1}, \dots, v_{n+m})$ de V .

Para $i \in \{1, \dots, m\}$, defina $w_i = f(v_{n+i})$. Basta demostrar que $\{w_1, \dots, w_m\}$ es una base de $\operatorname{im}(f)$, pues en tal caso tendríamos que $m = \dim(\operatorname{im}(f)) = \rho(f) = \dim(V) - \nu(f)$. Para establecerlo usaremos la proposición 1.21.

Veamos primero que $\{w_1, \dots, w_m\}$ es linealmente independiente. Suponga que $a_1, \dots, a_m \in \mathbb{K}$ son tales que

$$a_1 w_1 + \dots + a_m w_m = O.$$

Luego, si $v = a_1 v_{n+1} + \dots + a_m v_{n+m}$, entonces

$$f(v) = a_1 f(v_{n+1}) + \dots + a_m f(v_{n+m}) = a_1 w_1 + \dots + a_m w_m = O,$$

y así $v \in \ker(f)$. Pero como $\langle v_1, \dots, v_n \rangle = \ker(f)$, entonces existe $(b_1, \dots, b_n) \in \mathbb{K}^n$ tal que

$$v = b_1 v_1 + \dots + b_n v_n$$

es decir que $b_1 v_1 + \dots + b_n v_n = v = a_1 v_{n+1} + \dots + a_m v_{n+m}$ y así

$$O = (-b_1)v_1 + \dots + (-b_n)v_n + a_1 v_{n+1} + \dots + a_m v_{n+m}.$$

Por ende, como $\{v_1, \dots, v_n, v_{n+1}, \dots, v_{n+m}\}$ es linealmente independiente, la igualdad anterior implica que $a_1 = \dots = a_m = 0$ y la independencia lineal de $\{w_1, \dots, w_m\}$ se sigue ahora de la proposición 1.20.

Veamos que $\{w_1, \dots, w_m\}$ genera a $\text{im}(f)$ y con eso completamos la prueba. Como $w_1, \dots, w_m \in \text{im}(f)$ entonces $\langle w_1, \dots, w_m \rangle \subseteq \text{im}(f)$. Basta entonces establecer la otra inclusión. Dado $w \in \text{im}(f)$, existe $v \in V$ tal que $f(v) = w$. Sean $c_1, \dots, c_n, c_{n+1}, \dots, c_{n+m} \in \mathbb{K}$ tales que

$$v = c_1 v_1 + \dots + c_n v_n + c_{n+1} v_{n+1} + \dots + c_{n+m} v_{n+m},$$

de forma que

$$\begin{aligned} w = f(v) &= f(\underbrace{c_1 v_1 + \dots + c_n v_n}_{\in \ker(f)} + c_{n+1} v_{n+1} + \dots + c_{n+m} v_{n+m}) \\ &= c_{n+1} w_1 + \dots + c_{n+m} w_m, \end{aligned}$$

y así $w \in \langle w_1, \dots, w_m \rangle$. De donde $\text{im}(f) \subseteq \langle w_1, \dots, w_m \rangle$. \square

Corolario 1.46. *Suponga que V tiene dimensión finita y sea $f \in \text{Hom}_{\mathbb{K}}(V, W)$. Entonces las siguientes dos propiedades son equivalentes:*

- (i) $\nu(f) = 0$
- (ii) $\rho(f) = \dim(V)$

Dem. Se sigue inmediatamente del teorema del rango.

Ejercicios

- Demuestre que la función $f : \mathbb{K}^2 \rightarrow \mathbb{K}^3$ dada por $f(x, y) = (x + 2y, 2x - y, x + y)$ es una transformación lineal. Asuma que $\text{char}(K) \neq 2$.
- ¿Existe alguna transformación lineal f de \mathbb{Q}^2 en \mathbb{Q}^3 tal que $f(1, -2) = (1, 1, 1)$ y $f(-2, 4) = (1, -1, -1)$?
- ¿Cuántas transformaciones lineales f de \mathbb{Q}^2 en \mathbb{Q}^3 existen tales que $f(1, 1) = (3, 2, 4)$ y $f(1, 2) = (1, 1, 1)$? Calcule $f(0, 1)$ y $f(1, 0)$ para estas transformaciones.
- Sea $f : V \rightarrow W$ una transformación lineal, con $\dim_{\mathbb{K}}(V) = \dim_{\mathbb{K}}(W)$. Demuestre que f es inyectiva si y solo si es sobreyectiva.

5. Considere las funciones $d : \mathbb{R}[t] \rightarrow \mathbb{R}[t]$ e $i : \mathbb{R}[t] \rightarrow \mathbb{R}[t]$ dadas por:

$$d(P(t)) = \frac{d}{dt}P(t) \quad i(P(t)) = \int_0^t P(x)dx.$$

Demuestre que d e i son transformaciones lineales de espacios vectoriales sobre \mathbb{R} . Demuestre además que d es sobreyectiva pero no inyectiva, y que i es inyectiva pero no sobreyectiva.

6. Para cada una de las siguientes transformaciones lineales, encuentre una base del núcleo y del rango.

- a) $f \in \text{Hom}_{\mathbb{Q}}(\mathbb{Q}^3, \mathbb{Q}^3)$ dada por $f(x, y, z) = (3x - 3y - z, -y - 4z, 5x - 7y + 2z)$.
- b) $f \in \text{Hom}_{\mathbb{Q}}(\mathbb{Q}^4, \mathbb{Q}^4)$ dada por $f(x, y, z, w) = (x - y + z + w, -x - y - z, 2x - y - 2z - w, 2x - y - z)$.
- c) $f \in \text{Hom}_{\mathbb{Q}}(\mathbb{Q}^2, \mathbb{Q}^3)$ dada por $f(x, y) = (y, x, -x)$.
- d) $f \in \text{Hom}_{\mathbb{Q}}(\mathbb{Q}^2, \mathbb{Q}^3)$ dada por $f(x, y) = \frac{1}{2}(x + y, x + y, -x - y)$.
- e) $f \in \text{Hom}_{\mathbb{Q}}(\mathbb{Q}^2, \mathbb{Q}^4)$ dada por $f(x, y) = \frac{1}{3}(2y, 2y, 3x - y, -3x - y)$.
- f) $f \in \text{Hom}_{\mathbb{Q}}(\mathbb{Q}^3, \mathbb{Q}^4)$ dada por $f(x, y, z) = \frac{1}{3}(2x + y + z, 2x + y + z, -x + y - 2z, -x - 2y + z)$.

1.4. Matrices y vectores de coordenadas

Sea \mathbb{K} un cuerpo. Denotaremos por m, n, r, s cuatro enteros estrictamente positivos y por $[m]$ y $[n]$ los conjuntos $\{1, \dots, m\}$ y $\{1, \dots, n\}$.

Definición 1.47. Una *matriz* $m \times n$ sobre \mathbb{K} (o una *matriz* $m \times n$ con entradas en \mathbb{K}) es una familia A de escalares indexada por $[m] \times [n]$, es decir $A \in \mathbb{K}^{[m] \times [n]}$. Si $a_{ij} = A(i, j)$, para $(i, j) \in [m] \times [n]$, denotaremos $A = (a_{ij})$ ó

$$A = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix}.$$

y llamaremos a a_{ij} la entrada ij de A . La fila i de A es la matrix $1 \times n$

$$[a_{i1} \dots a_{in}]$$

y la columna j de A es la matrix $m \times 1$,

$$\begin{bmatrix} a_{1j} \\ \vdots \\ a_{mj} \end{bmatrix}.$$

El conjunto $M_{m \times n}(\mathbb{K})$ de matrices $m \times n$ sobre \mathbb{K} es el espacio vectorial $\mathbb{K}^{[m] \times [n]}$ (ver el ejemplo 1.2(iv)).

Definición 1.48. Un *vector de n coordenadas sobre \mathbb{K}* (o un *vector de n coordenadas con entradas en \mathbb{K}*) es una matriz $n \times 1$. Si $\bar{x} \in M_{n \times 1}(\mathbb{K})$ es tal que $\bar{x}(i, 1) = x_i$, para $i \in \{1, \dots, n\}$, entonces denotaremos $\bar{x} = (x_i)$ ó

$$\bar{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$$

y llamamos a x_i la coordenada i de \bar{x} . Para simplificar, denotaremos $\bar{x}(i, 1)$ por $\bar{x}(i)$.

Definición 1.49. Sean A y B respectivamente matrices $m \times n$ y $n \times r$. Definimos el *producto de A y B* , que denotamos por AB , como la matriz $m \times r$ cuya entrada ij es

$$\sum_{k=1}^n a_{ik} b_{kj} = a_{i1} b_{1j} + \dots + a_{in} b_{nj}.$$

Proposición 1.50. La multiplicación matricial satisface las siguientes propiedades.

- (i) *Commutatividad con escalares:* Para $c \in \mathbb{K}$, $A \in M_{m \times n}(\mathbb{K})$ y $B \in M_{n \times r}(\mathbb{K})$ cualesquiera se tiene $A(cB) = cAB = (cA)B$.
- (ii) *Distributividad:* Para $A \in M_{m \times n}(\mathbb{K})$ y $B, C \in M_{n \times r}(\mathbb{K})$ cualesquiera se tiene $A(B + C) = AB + AC$.
- (iii) *Asociatividad:* Para $A \in M_{m \times n}(\mathbb{K})$, $B \in M_{n \times r}(\mathbb{K})$ y $C \in M_{r \times s}(\mathbb{K})$ cualesquiera se tiene $A(BC) = (AB)C$.

Dem. Sean $A = (a_{ij})$, $B = (b_{jk})$ y $C = (c_{kl})$.

- (i) La entrada ij de $A(cB)$ es

$$\sum_{k=1}^n a_{ik} c b_{kj} = c \sum_{k=1}^n a_{ik} b_{kj} = \sum_{k=1}^n c a_{ik} b_{kj},$$

y es igual a la entrada ij de cAB y de $(cA)B$.

- (ii) La entrada ij de $A(B + C)$ es

$$\sum_{k=1}^n a_{ik} (b_{kj} + c_{kj}) = \sum_{k=1}^n a_{ik} b_{kj} + \sum_{k=1}^n a_{ik} c_{kj},$$

y es igual a la entrada ij de $AB + AC$.

(iii) La entrada ij de $A(BC)$ es

$$\sum_{l=1}^n a_{il} \left(\sum_{k=1}^r b_{lk} c_{kj} \right) = \sum_{k=1}^r \sum_{l=1}^n a_{il} b_{lk} c_{kj} = \sum_{k=1}^r \left(\sum_{l=1}^n a_{il} b_{lk} \right) c_{kj},$$

y es igual a la entrada ij de $(AB)C$.

□

Definición 1.51. La *matriz identidad* $n \times n$ es la matriz $I_n \in M_{n \times n}(\mathbb{K})$ cuya entrada ij es 1 cuando $i = j$ y 0 cuando $i \neq j$. En particular tenemos

$$I_n = \begin{bmatrix} 1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 1 \end{bmatrix}$$

e $I_n = (\delta_{ij})$ donde definimos $\delta_{ij} = 1$ cuando $i = j$ y $\delta_{ij} = 0$ cuando $i \neq j$. La función $\delta \in \mathbb{K}^{[n] \times [n]}$ definida por $\delta(i, j) = \delta_{ij}$ se llama *la función delta de Kronecker*.

Proposición 1.52. Para toda matriz $A \in M_{m \times n}(\mathbb{K})$, se tiene $I_m A = A = A I_n$.

Dem. Si a_{ij} es la entrada ij de A , entonces la entrada ij de $I_m A$ es $\sum_{k=1}^m \delta_{ik} a_{kj} = a_{ij}$ y la de $A I_n$ es $\sum_{k=1}^n a_{ik} \delta_{kj} = a_{ij}$. □

Definición 1.53. Sea $A \in M_{n \times n}(\mathbb{K})$. Decimos que A es una *matriz invertible* si existe una matriz $A^{-1} \in M_{n \times n}(\mathbb{K})$ para la cual se tiene $AA^{-1} = I_n = A^{-1}A$. En tal caso, llamamos a A^{-1} *la matriz inversa de A* y decimos que A es *invertible*.

Proposición 1.54 (Unicidad de la inversa). Sea $A \in M_{n \times n}(\mathbb{K})$ una matriz invertible. Si $B \in M_{n \times n}$ es tal que $BA = I_n$ ó $AB = I_n$ entonces $B = A^{-1}$.

Dem. Si BA es igual a I_n , entonces tenemos $B = BI_n = B(AA^{-1}) = (BA)A^{-1} = I_n A^{-1} = A^{-1}$. Similarmente se establece $B = A^{-1}$ cuando AB es igual a I_n . □

Matrices de transformaciones

Sean V , W y U espacios vectoriales sobre \mathbb{K} de dimensión finita, n , m y r sus respectivas dimensiones. Fijamos una base $\mathcal{B}_V = (v_1, \dots, v_n)$, $\mathcal{B}_W = (w_1, \dots, w_m)$ y $\mathcal{B}_U = (u_1, \dots, u_r)$ de cada uno de estos tres espacios.

Definición 1.55. Dado $v \in V$, sean $c_1, \dots, c_n \in \mathbb{K}$ tales que

$$c_1 v_1 + \cdots + c_n v_n.$$

El *vector de coordenadas de v en la base \mathcal{B}_V* es el vector de n coordenadas

$$\begin{bmatrix} v \end{bmatrix}^{\mathcal{B}_V} = \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix}$$

Proposición 1.56. *La función*

$$\begin{aligned} \left[\bullet \right]^{\mathcal{B}_V} : V &\longrightarrow \mathbb{K}^{[n]} \\ v &\longmapsto \left[v \right]^{\mathcal{B}_V} \end{aligned}$$

es un isomorfismo.

Dem. Tenemos $[v]^{\mathcal{B}_V} = O$ si y solo si $v = O$, así la proposición se sigue del corolario 1.46 del teorema del rango.

Definición 1.57. Dada $f \in \text{Hom}_{\mathbb{K}}(V, W)$, para $i \in \{1, \dots, m\}$ y $j \in \{1, \dots, n\}$, definimos $a_{ij} \in \mathbb{K}$ tal que

$$f(v_j) = a_{1j}w_1 + \dots + a_{mj}w_m.$$

La matriz de f para las bases \mathcal{B}_V y \mathcal{B}_W es la matrix $m \times n$ cuya entrada ij es a_{ij} , la cual denotamos $\left[f \right]_{\mathcal{B}_V}^{\mathcal{B}_W}$.

Observación 1.58. (i) La columna j de $\left[f \right]_{\mathcal{B}_V}^{\mathcal{B}_W}$ es el vector de coordenadas

$\left[f(v_j) \right]^{\mathcal{B}_W}$, en particular tenemos

$$\left[f \right]_{\mathcal{B}_V}^{\mathcal{B}_W} = \left[\left[f(v_1) \right]^{\mathcal{B}_W} \mid \dots \mid \left[f(v_n) \right]^{\mathcal{B}_W} \right]$$

$$\text{y } \left[\text{id}_V \right]_{\mathcal{B}_V}^{\mathcal{B}_V} = I_n.$$

(ii) Cuando V y W son los espacios \mathbb{K}^n y \mathbb{K}^m matriz para las bases canónicas \mathcal{C}_n y \mathcal{C}_m la llamamos la *matriz canónica*.

Proposición 1.59. *Dado $f \in \text{Hom}_{\mathbb{K}}(V, W)$, para todo $v \in V$, se tiene*

$$\left[f(v) \right]^{\mathcal{B}_W} = \left[f \right]_{\mathcal{B}_V}^{\mathcal{B}_W} \left[v \right]^{\mathcal{B}_V}.$$

En particular, si $A = \left[f \right]_{\mathcal{B}_V}^{\mathcal{B}_W}$, $\bar{x} = \left[v \right]^{\mathcal{B}_V}$ e $\bar{y} = \left[f(v) \right]^{\mathcal{B}_W}$ entonces

$$\bar{y} = A\bar{x}.$$

Dem. Sean $\left[f \right]_{\mathcal{B}_V}^{\mathcal{B}_W} = (a_{ij})$ y $\left[v \right]^{\mathcal{B}_V} = (c_i)$. La coordenada i del vector de coor-

denadas $\begin{bmatrix} f \end{bmatrix}_{\mathcal{B}_V}^{\mathcal{B}_W} \begin{bmatrix} v \end{bmatrix}_{\mathcal{B}_V}^{\mathcal{B}_V}$ es $\sum_j a_{ij} c_j$, y así, de las igualdades

$$\begin{aligned} f(v) &= f(c_1 v_1 + \cdots + c_n v_n) = c_1 f(v_1) + \cdots + c_n f(v_n) \\ &= c_1 \sum_{i=1}^m a_{i1} w_i + \cdots + c_n \sum_{i=1}^m a_{in} w_i \\ &= \sum_{j=1}^n \sum_{i=1}^m a_{ij} c_j w_i = \sum_{i=1}^m \left(\sum_{j=1}^n a_{ij} c_j \right) w_i, \end{aligned}$$

se sigue que la coordenada i de $f(v)$ en la base \mathcal{B}_W es igual a la coordenada i del vector de coordenadas $\begin{bmatrix} f \end{bmatrix}_{\mathcal{B}_V}^{\mathcal{B}_W} \begin{bmatrix} v \end{bmatrix}_{\mathcal{B}_V}^{\mathcal{B}_V}$. \square

Proposición 1.60. Dado $f \in \text{Hom}_{\mathbb{K}}(V, W)$ y $g \in \text{Hom}_{\mathbb{K}}(W, U)$, se tiene

$$\begin{bmatrix} g \circ f \end{bmatrix}_{\mathcal{B}_V}^{\mathcal{B}_U} = \begin{bmatrix} g \end{bmatrix}_{\mathcal{B}_W}^{\mathcal{B}_U} \begin{bmatrix} f \end{bmatrix}_{\mathcal{B}_V}^{\mathcal{B}_W}.$$

En particular, si $A = \begin{bmatrix} f \end{bmatrix}_{\mathcal{B}_V}^{\mathcal{B}_W}$, $B = \begin{bmatrix} g \end{bmatrix}_{\mathcal{B}_W}^{\mathcal{B}_U}$ y $C = \begin{bmatrix} g \circ f \end{bmatrix}_{\mathcal{B}_V}^{\mathcal{B}_U}$ entonces

$$C = BA.$$

Dem. Sea $\begin{bmatrix} f \end{bmatrix}_{\mathcal{B}_V}^{\mathcal{B}_W} = (a_{ij})$ y $\begin{bmatrix} g \end{bmatrix}_{\mathcal{B}_W}^{\mathcal{B}_U} = (b_{ij})$. La entrada ij de $\begin{bmatrix} g \end{bmatrix}_{\mathcal{B}_W}^{\mathcal{B}_U} \begin{bmatrix} f \end{bmatrix}_{\mathcal{B}_V}^{\mathcal{B}_W}$ es $\sum_k b_{ik} a_{kj}$, y así, de las igualdades

$$\begin{aligned} g \circ f(v_j) &= g(a_{1j} w_1 + \cdots + a_{mj} w_m) = a_{1j} g(w_1) + \cdots + a_{mj} g(w_m) \\ &= a_{1j} \sum_{i=1}^r b_{i1} u_i + \cdots + a_{mj} \sum_{i=1}^r b_{im} u_i \\ &= \sum_{k=1}^r \sum_{i=1}^r b_{ik} a_{kj} u_i = \sum_{i=1}^r \left(\sum_{k=1}^r b_{ik} a_{kj} \right) u_i, \end{aligned}$$

se sigue que la coordenada i de $g \circ f(v_j)$ en la base \mathcal{B}_U es igual a la entrada ij de la matriz $\begin{bmatrix} g \end{bmatrix}_{\mathcal{B}_W}^{\mathcal{B}_U} \begin{bmatrix} f \end{bmatrix}_{\mathcal{B}_V}^{\mathcal{B}_W}$. \square

Proposición 1.61. El mapa $\begin{bmatrix} \bullet \end{bmatrix}_{\mathcal{B}_V}^{\mathcal{B}_W}$ definido por

$$\begin{aligned} \begin{bmatrix} \bullet \end{bmatrix}_{\mathcal{B}_V}^{\mathcal{B}_W} : \text{Hom}_{\mathbb{K}}(V, W) &\longrightarrow M_{m \times n}(\mathbb{K}) \\ f &\longmapsto \begin{bmatrix} f \end{bmatrix}_{\mathcal{B}_V}^{\mathcal{B}_W} \end{aligned}$$

es un isomorfismo de espacios vectoriales sobre \mathbb{K} . Más aún, f es un isomorfismo si y solo si $A = \begin{bmatrix} f \end{bmatrix}_{\mathcal{B}_V}^{\mathcal{B}_W}$ es invertible, y en tal caso se tiene $\begin{bmatrix} f^{-1} \end{bmatrix}_{\mathcal{B}_V}^{\mathcal{B}_W} = A^{-1} = \left(\begin{bmatrix} f \end{bmatrix}_{\mathcal{B}_V}^{\mathcal{B}_W} \right)^{-1}$.

Dem. Sean $f, g \in \text{Hom}_{\mathbb{K}}(V, W)$. Como $(f + g)(v_j) = f(v_j) + g(v_j)$ y $(cf)(v_j) = cf(v_j)$ para todo $j \in \{1, \dots, n\}$ y todo $c \in \mathbb{K}$, entonces $\left[\bullet \right]_{\mathcal{B}_V}^{\mathcal{B}_W}$ es una transformación lineal. Para establecer que es una biyección basta notar que, por la proposición 1.35, dada una matriz $A \in M_{m \times n}(\mathbb{K})$, con $A = (a_{ij})$, existe una única transformación lineal $f \in \text{Hom}_{\mathbb{K}}(V, W)$ tal que $f(v_j) = \sum_i a_{ij} w_i$.

Sea $\left[f \right]_{\mathcal{B}_V}^{\mathcal{B}_W} = A$. Si f es biyectiva y B es la matriz $\left[f^{-1} \right]_{\mathcal{B}_W}^{\mathcal{B}_V}$, entonces se tiene

$$BA = \left[f^{-1} \right]_{\mathcal{B}_W}^{\mathcal{B}_V} \left[f \right]_{\mathcal{B}_V}^{\mathcal{B}_W} = \left[f^{-1} \circ f \right]_{\mathcal{B}_V}^{\mathcal{B}_V} = \left[\text{id}_V \right]_{\mathcal{B}_V}^{\mathcal{B}_V} = I_n,$$

luego A es invertible y $B = A^{-1}$. Si A es invertible, entonces por la proposición 1.61 existe una única transformación lineal $g \in \text{Hom}_{\mathbb{K}}(W, V)$ tal que A^{-1} es la matriz $\left[g \right]_{\mathcal{B}_W}^{\mathcal{B}_V}$. Las igualdades

$$\left[g \circ f \right]_{\mathcal{B}_V}^{\mathcal{B}_V} = \left[g \right]_{\mathcal{B}_W}^{\mathcal{B}_V} \left[f \right]_{\mathcal{B}_V}^{\mathcal{B}_W} = A^{-1}A = I_n = \left[\text{id}_V \right]_{\mathcal{B}_V}^{\mathcal{B}_V},$$

implican $g \circ f = \text{id}_V$. Similarmente, tenemos $f \circ g = \text{id}_W$. Así, f es una biyección y g es la inversa f^{-1} . \square

Definición 1.62. Dadas dos bases $\mathcal{B} = (v_1, \dots, v_n)$ y $\mathcal{B}' = (v'_1, \dots, v'_n)$ de V . La matriz de cambio de coordenadas de la base \mathcal{B} a la base \mathcal{B}' es la matriz $\left[\text{id}_V \right]_{\mathcal{B}}^{\mathcal{B}'}$.

Observación 1.63. La columna j de $\left[\text{id}_V \right]_{\mathcal{B}}^{\mathcal{B}'}$ es el vector de coordenadas $\left[v_j \right]^{\mathcal{B}'}$, en particular tenemos

$$\left[\text{id}_V \right]_{\mathcal{B}}^{\mathcal{B}'} = \left[\left[v_1 \right]^{\mathcal{B}'} \mid \dots \mid \left[v_n \right]^{\mathcal{B}'} \right].$$

Proposición 1.64. Dadas dos bases $\mathcal{B} = (v_1, \dots, v_n)$ y $\mathcal{B}' = (v'_1, \dots, v'_n)$ de V .

- (i) Para todo $v \in V$ se tiene $\left[v \right]^{\mathcal{B}'} = \left[\text{id}_V \right]_{\mathcal{B}}^{\mathcal{B}'} \left[v \right]^{\mathcal{B}}$.
- (ii) La matriz de cambio de coordenadas de una base a la otra y la matriz de cambio inverso son una la inversa de la otra, es decir que se tiene la igualdad

$$\left[\text{id}_V \right]_{\mathcal{B}'}^{\mathcal{B}} = \left(\left[\text{id}_V \right]_{\mathcal{B}}^{\mathcal{B}'} \right)^{-1}.$$

Dem.

- (i) Se sigue de las igualdades $\left[v \right]^{\mathcal{B}'} = \left[\text{id}_V(v) \right]^{\mathcal{B}'} = \left[\text{id}_V \right]_{\mathcal{B}}^{\mathcal{B}'} \left[v \right]^{\mathcal{B}}$.

(ii) Se sigue de la proposición 1.61 aplicada a id_V y de la igualdad $\text{id}_V^{-1} = \text{id}_V$.

□

Proposición 1.65. Sean \mathcal{B}_V , y \mathcal{B}'_V bases de V , y \mathcal{B}_W , y \mathcal{B}'_W bases de W . Para toda $f \in \text{Hom}_{\mathbb{K}}(V, W)$, se tiene

$$\left[f \right]_{\mathcal{B}'_V}^{\mathcal{B}'_W} = \left[\text{id}_W \right]_{\mathcal{B}_W}^{\mathcal{B}'_W} \left[f \right]_{\mathcal{B}_V}^{\mathcal{B}_W} \left[\text{id}_V \right]_{\mathcal{B}'_V}^{\mathcal{B}_V}.$$

En particular, si $A = \left[f \right]_{\mathcal{B}_V}^{\mathcal{B}_W}$, $B = \left[f \right]_{\mathcal{B}'_V}^{\mathcal{B}'_W}$, $C_1 = \left[\text{id}_V \right]_{\mathcal{B}'_V}^{\mathcal{B}_V}$ y $C_2 = \left[\text{id}_W \right]_{\mathcal{B}'_W}^{\mathcal{B}_W}$ entonces

$$B = C_2^{-1} A C_1.$$

Dem. La propiedad se sigue de la igualdad $f = \text{id}_W \circ f \circ \text{id}_V$ y de la proposición 1.60. □

Observación 1.66. Sean \mathcal{B} y \mathcal{B}' bases de V y sea $f \in \text{End}_{\mathbb{K}}(V)$. Para

$$A = \left[f \right]_{\mathcal{B}}^{\mathcal{B}}, \quad B = \left[f \right]_{\mathcal{B}'}^{\mathcal{B}'}, \quad \text{y} \quad C = \left[\text{id}_V \right]_{\mathcal{B}'}^{\mathcal{B}}$$

tenemos

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

Ejemplo 1.67. Suponga que $\text{char}(\mathbb{K}) \neq 2$, de forma que $-1 \neq 1$ en \mathbb{K} . Sea $f \in \text{Hom}_{\mathbb{K}}(\mathbb{K}^2, \mathbb{K}^2)$ el operador definido por

$$f(x, y) = (y, x).$$

Para la base canónica \mathcal{C} se tiene

$$\left[f \right]_{\mathcal{C}}^{\mathcal{C}} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

y para $\mathcal{B} = ((1, 1), (1, -1))$ se tiene

$$\left[\text{id}_{\mathbb{K}^2} \right]_{\mathcal{B}}^{\mathcal{C}} = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix},$$

luego

$$\begin{aligned} \left[f \right]_{\mathcal{B}}^{\mathcal{B}} &= \left[\text{id}_{\mathbb{K}^2} \right]_{\mathcal{C}}^{\mathcal{B}} \left[f \right]_{\mathcal{C}}^{\mathcal{C}} \left[\text{id}_{\mathbb{K}^2} \right]_{\mathcal{B}}^{\mathcal{C}} \\ &= \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}^{-1} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \\ &= \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}. \end{aligned}$$

Observación 1.68 (Vector de coordenadas en dimensión infinita). Dada una base \mathcal{B} de V , con $\mathcal{B} = (v_i)_{i \in I}$, para cada $v \in V$ existe una única combinación lineal de \mathcal{B} igual a v . Para $i \in I$, sean $c_i \in \mathbb{K}$ los coeficientes de esta combinación lineal, es decir tenemos $\sum_{i \in I} c_i v_i = v$. La unicidad de esta combinación lineal nos permite identificar cada elemento en v con el elemento $\begin{bmatrix} v \end{bmatrix}^{\mathcal{B}} \in (\mathbb{K}^I)_0$ definido por

$$\begin{aligned} \begin{bmatrix} v \end{bmatrix}^{\mathcal{B}} : I &\longrightarrow \mathbb{K} \\ i &\longrightarrow c_i. \end{aligned}$$

Observación 1.69 (Dimensión infinita e isomorfismo). La caracterización de los espacios lineales, salvo isomorfismos, por su dimensión se puede reescribir así: si \mathcal{B}_V y \mathcal{B}_W son respectivamente bases V y W , con $\mathcal{B}_V = (v_j)_{j \in J}$ y $\mathcal{B}_W = (w_i)_{i \in I}$, entonces $V \simeq_{\mathbb{K}} W$ si y solo si existe una biyección $\phi : J \rightarrow I$. Empecemos por establecer la suficiencia. Note que los mapas (ver notación en Observación 1.68)

$$\begin{aligned} V &\longrightarrow (\mathbb{K}^J)_0 & y & & W &\longrightarrow (\mathbb{K}^I)_0 \\ v &\longmapsto \begin{bmatrix} v \end{bmatrix}^{\mathcal{B}_V} & & & w &\longmapsto \begin{bmatrix} w \end{bmatrix}^{\mathcal{B}_W}, \end{aligned}$$

son isomorfismos. Ahora, dada una biyección $\phi : J \rightarrow I$ entre los índices de las bases, la transformación lineal $\Phi \in \text{Hom}_{\mathbb{K}}((\mathbb{K}^J)_0, (\mathbb{K}^I)_0)$ definida en la base $\{\delta_j\}_{j \in J}$ de $(\mathbb{K}^J)_0$ (ver el ejemplo 1.22.(iii)) por $\Phi(\delta_j) = \delta_{\phi(j)}$ es un isomorfismo. Tenemos así $V \simeq_{\mathbb{K}} (\mathbb{K}^J)_0 \simeq_{\mathbb{K}} (\mathbb{K}^I)_0 \simeq_{\mathbb{K}} W$.

Para establecer la necesidad necesitamos demostrar que si V y W son isomorfos, entonces existe una biyección entre J e I . Para esto basta demostrar que dos bases cualesquiera de V están en correspondencia biyectiva, pues si $f \in \text{Hom}_{\mathbb{K}}(W, V)$ es un isomorfismo, entonces la imagen $f(\mathcal{B}_W)$ es una base de V y es un conjunto en biyección con J , y luego, si existe una biyección entre \mathcal{B}_V y $f(\mathcal{B}_W)$, entonces hay una biyección entre J e I . Supongamos entonces que V tiene dimensión infinita y sean \mathcal{B}_V y \mathcal{B}'_V respectivamente las bases $(v_j)_{j \in J}$ y $(v_{j'})_{j' \in J'}$ de V . Para cada $j \in J$ denote por J'_j al conjunto de índices $j' \in J'$ para los cuales v_j tiene coordenada diferente de cero en la base \mathcal{B}'_V , es decir tenemos $J'_j = \left\{ j' \in J' \mid \begin{bmatrix} v_j \end{bmatrix}^{\mathcal{B}'_V}_{j'} \neq 0 \right\}$. En particular J'_j es la mínima colección de índices $j' \in J'$ con la propiedad $v_j \in \langle v'_{j'} \rangle_{j' \in J'_j}$. Tenemos que, para $j \in J$, cada J'_j es finito. Veamos que $\cup_{j \in J} J'_j = J'$. De hecho, en caso contrario, si existe $j' \in J' \setminus \cup_{j \in J} J'_j$ y $\{v_{j_1}, \dots, v_{j_n}\} \subseteq \mathcal{B}_V$ es tal que $v_{j'}$ está en $\langle v_{j_1}, \dots, v_{j_n} \rangle$, entonces $v_{j'}$ está en $\langle v_{k'} \mid k' \in \bigcup_{k=1}^n J'_{j_k} \rangle$ que es un subespacio de $\langle v_{k'} \mid k' \in J' \setminus \{j'\} \rangle$, lo cual, por el lema 1.24, violaría la independencia lineal de $\mathcal{B}'_V = \{v_{k'}\}_{k' \in J'}$. Defina \mathcal{J}' como la unión disyunta de todos los J'_j , para $j \in J$, es decir $\mathcal{J}' = \coprod_{j \in J} J'_j$. Como \mathcal{J}' es una unión de conjuntos finitos y disyuntos indexada por J y J es infinito, entonces J y \mathcal{J}' son conjuntos biyectivos. Como tenemos $\cup_{j \in J} J'_j = J'$, en \mathcal{J}' podemos inyectar a J' , y, así también, en J . Simétricamente podemos

inyectar J en J' . Luego, por el teorema de Schroeder-Bernstein, J y J' son biyectivos.

Observación 1.70 (Dimensión arbitraria y descomposición del dominio). El teorema 1.45 del rango se demostró descomponiendo una base del dominio de la transformación lineal. Este resultado lo podemos generalizar a espacios de dimensión infinita de la siguiente manera. Sea $f \in \text{Hom}_{\mathbb{K}}(V, W)$, entonces existe una base \mathcal{B} de V tal que \mathcal{B} es igual a la unión de dos conjuntos disyuntos \mathcal{B}_0 y \mathcal{B}_1 , donde \mathcal{B}_0 es una base de $\ker(f)$ y $f(\mathcal{B}_1)$ es una base de $\text{im}(f)$. De hecho, basta tomar una base \mathcal{B}_0 de $\ker(f)$ y extenderla a una de V (ver Observación 1.29). El resto de detalles son similares a los de la demostración del teorema.

Observación 1.71 (Dimensión infinita y transformaciones lineales). La proposición 1.35 de rigidez de las transformaciones lineales también se puede generalizar a espacios de dimensión infinita. Sea C un conjunto generador de V . Dada una función $f_0 : C \rightarrow W$, existe a los sumo una transformación lineal $f \in \text{Hom}_{\mathbb{K}}(V, W)$ para la cual se tiene $f(v) = f_0(v)$ para todo $v \in C$. Si además C es linealmente independiente, entonces una tal transformación lineal f existe. La demostración es fundamentalmente la misma que en el caso de base finita. Note que un caso particular de esta observación ya se usó en Observación 1.69.

Ejercicios

1. Sea $f \in \text{Hom}_{\mathbb{Q}}(\mathbb{Q}^3, \mathbb{Q}^3)$ tal que $f(1, 0, 0) = (0, 1, 2)$, $f(0, 1, 0) = (1, 2, 3)$ y $f(0, 0, 1) = (2, 0, -1)$. Encuentre la matriz canónica de f y verifique que $f(x, y, z) = (y + 2z, x + 2y, 2x + 3y - z)$.

2. Obtenga la matrix canónica de $f \in \text{Hom}_{\mathbb{Q}}(\mathbb{Q}^3, \mathbb{Q}^3)$ donde

$$f(x, y, z) = (3x - 3y - z, -y - 4z, 5x - 7y + 2z).$$

3. Obtenga la matrix canónica de $f \in \text{Hom}_{\mathbb{Q}}(\mathbb{Q}^4, \mathbb{Q}^4)$ donde

$$f(x, y, z, w) = (x - y + z + w, -x - y - z, 2x - y - 2z - w, 2x - y - z)$$

4. Obtenga la matrix canónica de cada una de las transformaciones en el ejercicio 1.3.6.

5. Sean $f, g \in \text{Hom}_{\mathbb{Q}}(\mathbb{Q}^3, \mathbb{Q}^3)$ definidas por

$$f(x, y, z) = (y + 2z, x + 2y, 2x + 3y - z)$$

y

$$g(x, y, z) = (3x - 3y - z, -y - 4z, 5x - 7y + 2z).$$

Encuentre la matrix canónica de $f \circ g$ y $g \circ f$.

6. Sea $V = \mathbb{R}[t]_{\leq 2}$ el espacio de polinomios de grado menor o igual a 2 con coeficientes reales y \mathcal{B} la base $(1, t, t^2)$. Encuentre la matrix para la base \mathcal{B} de $f \in \text{End}_{\mathbb{R}}(V)$ con $f(P(t)) = P(t) + \frac{d}{dt}P(t)$. Determine si f es invertible y en tal caso encuentre la matrix de f^{-1} para \mathcal{B} .

1.5. Sumas directas y proyecciones

Sea V un espacio vectorial sobre \mathbb{K} .

Definición 1.72. Dados $V_1, \dots, V_n \leq V$, definimos su *suma* como el conjunto $\{v_1 + \dots + v_n \in V \mid v_i \in V_i, i = 1, \dots, n\}$ que denotamos por $V_1 + \dots + V_n$ ó por $\sum_{i=1}^n V_i$.

Proposición 1.73. Si V_1, \dots, V_n son subespacios de V , entonces $V_1 + \dots + V_n$ es un subespacio de V .

Dem. Usamos la proposición 1.7. Primero, note que $V_1 + \dots + V_n$ contiene al origen. Tome $v, v' \in V_1 + \dots + V_n$ y $c \in \mathbb{K}$. Para $i \in \{1, \dots, n\}$, sean $v_i, v'_i \in V_i$ tales que se tiene $v = v_1 + \dots + v_n$ y $v' = v'_1 + \dots + v'_n$. Obtenemos así $v + v' = (v_1 + v'_1) + \dots + (v_n + v'_n)$, y por ende, como $v_i + v'_i$ pertenece a V_i , para $i \in \{1, \dots, n\}$, entonces $v + v'$ pertenece a $V_1 + \dots + V_n$. Igualmente, como tenemos $cv_i \in V_i$, para $i \in \{1, \dots, n\}$, de la igualdad $cv = cv_1 + \dots + cv_n$ vemos que v está en $V_1 + \dots + V_n$. \square

Teorema 1.74. Si V_1 y V_2 son subespacios de V de dimensión finita, entonces $V_1 \cap V_2$ y $V_1 + V_2$ también lo son. Más aún se tiene

$$\dim(V_1 + V_2) = \dim(V_1) + \dim(V_2) - \dim(V_1 \cap V_2),$$

o equivalentemente, $\dim(V_1) + \dim(V_2) = \dim(V_1 + V_2) + \dim(V_1 \cap V_2)$.

Dem. Como V_1 tiene dimensión finita, y $V_1 \cap V_2$ es un subespacio de V_1 , entonces, por el teorema 1.28 de monotonía de la dimensión, $V_1 \cap V_2$ también tiene dimensión finita. Sean $n_1 = \dim(V_1)$, $n_2 = \dim(V_2)$, $p = \dim(V_1 \cap V_2)$ y $\{v_1, \dots, v_p\}$ una base de $V_1 \cap V_2$. Extendemos esta base de $V_1 \cap V_2$ a una base \mathcal{B}_1 de V_1 constituida por los vectores $v_1, \dots, v_p, v'_{p+1}, \dots, v'_{n_1}$, y a una base \mathcal{B}_2 de V_2 constituida por los vectores $v_1, \dots, v_p, v''_{p+1}, \dots, v''_{n_2}$. Así, el conjunto $\{v_1, \dots, v_p, v'_{p+1}, \dots, v'_{n_1}, v''_{p+1}, \dots, v''_{n_2}\}$ genera a $V_1 + V_2$, luego este subespacio tiene dimensión finita. Sea

$$\mathcal{B} = (v_1, \dots, v_p, v'_{p+1}, \dots, v'_{n_1}, v''_{p+1}, \dots, v''_{n_2}).$$

El teorema se sigue si demostramos que \mathcal{B} es una base, para esto nos hace falta demostrar que es linealmente independiente y lo haremos usando la proposición 1.20. Suponga que $a_1, \dots, a_p, a'_{p+1}, \dots, a'_{n_1}, a''_{p+1}, \dots, a''_{n_2}$ son tales que se tiene

$$O = \underbrace{\sum_{i=1}^p a_i v_i + \sum_{i=p+1}^{n_1} a'_i v'_i}_v + \underbrace{\sum_{i=p+1}^{n_2} a''_i v''_i}_{-v}.$$

Luego, para $v = \sum_{i=1}^p a_i v_i + \sum_{i=p+1}^{n_1} a'_i v'_i$ tenemos que v está en V_1 y también $v = -\sum_{i=p+1}^{n_2} a''_i v''_i$ también está en V_2 , y por ende v está en $V_1 \cap V_2$. Sean $b_1, \dots, b_p \in \mathbb{K}$ para los cuales se tiene

$$v = b_1 v_1 + \dots + b_p v_p,$$

entonces obtenemos

$$O = v - v = \sum_{i=1}^p (a_i - b_i)v_i + \sum_{i=p+1}^{n_1} a'_i v'_i.$$

Por la independencia lineal de \mathcal{B}_1 tenemos $a'_{p+1} = \dots = a'_{n_1} = 0$ y

$$O = \sum_{i=1}^p a_i v_i + \sum_{i=p+1}^{n_2} a''_i v''_i.$$

Por la independencia lineal de \mathcal{B}_2 , tenemos $a_1 = \dots = a_p = a''_{p+1} = \dots = a''_{n_2} = 0$. \square

Observación 1.75. Note que si i, s, n_1, n_2 son tales que $i \leq n_1 \leq n_2 \leq s$ y s es menor que la dimensión de V , entonces existen $V_1, V_2 \leq V$ con $\dim(V_1) = n_1$, $\dim(V_2) = n_2$, $\dim(V_1 \cap V_2) = i$ y $\dim(V_1 + V_2) = s$, siempre que

$$n_1 + n_2 = s + i.$$

De hecho si $\{v_1, \dots, v_s\}$ es una colección de s vectores linealmente independientes en V basta tomar $V_1 = \langle v_1, \dots, v_{n_1} \rangle$ y $V_2 = \langle v_1, \dots, v_i, v_{n_1+1}, \dots, v_s \rangle$. Más aún, si V'_1 y V'_2 son subespacios de V para los que se tiene $\dim(V'_1) = n_1$, $\dim(V'_2) = n_2$, $\dim(V'_1 \cap V'_2) = i$ y $\dim(V'_1 + V'_2) = s$, entonces existe un automorfismo $f \in \text{End}_{\mathbb{K}}(V)$ tal que $f(V_1) = V'_1$ y $f(V_2) = V'_2$.

Definición 1.76. Sean $V_1, V_2 \leq V$. Decimos que V_1 y V_2 están en *posición general* si $\dim(V_1 + V_2)$ es tan grande y $\dim(V_1 \cap V_2)$ es tan pequeño como lo es posible.

Ejemplo 1.77. Dos subespacios bidimensional de un espacio tridimensional están en posición general si su intersección es un espacio unidimensional. Dos subespacio cuatridimensional de un espacio sexadimensional están en posición general si su intersección es un espacio bidimensional. Dos subespacios tridimensionales en un espacio septadimensional están en posición general si su intersección es trivial.

Definición 1.78. Sean $V_1, \dots, V_r \leq V$, decimos que V es la *suma directa* de V_1, \dots, V_r , lo cual denotamos por

$$V = V_1 \oplus \dots \oplus V_r = \bigoplus_{i=1}^r V_i$$

si para cada $v \in V$ existe un único $(v_1, \dots, v_r) \in V_1 \times \dots \times V_r$ que cumple

$$v = v_1 + \dots + v_r.$$

Proposición 1.79. Sean $V_1, \dots, V_r \leq V$. Se tiene $V = \bigoplus_{i=1}^r V_i$ si y solo si V_1, \dots, V_r satisfacen las siguientes dos propiedades.

(i) La suma $\sum_{i=1}^r V_i$ es igual a V .

(ii) Para todo $i \in \{1, \dots, r\}$ se tiene $V_i \cap \sum_{j \neq i} V_j = \{O\}$.

Dem. Suponga primero que tenemos $V = \bigoplus_{i=1}^r V_i$, luego por definición tenemos $V = \sum_{i=1}^r V_i$. Por otro lado, sea $i \in \{1, \dots, r\}$ y tome $v \in V_i \cap \sum_{j \neq i} V_j$. Así, existe $(v_1, \dots, v_r) \in V_1 \times \dots \times V_r$ para el cual se cumple $v = -v_i = \sum_{j \neq i} v_j$ y por lo cual se tiene $O = v_1 + \dots + v_r$. Pero por otro lado, para $(O, \dots, O) \in V_1 \times \dots \times V_r$ se tiene $O = O + \dots + O$, luego, por unicidad de esta descomposición se siguen las igualdades $v_1 = \dots = v_r = O$, $v = O$ y $V_i \cap \sum_{j \neq i} V_j = \{O\}$.

Recíprocamente, suponga que $\sum_{i=1}^r V_i$ es igual a V y que para todo $i \in \{1, \dots, r\}$ se tiene $V_i \cap \sum_{j \neq i} V_j = \{O\}$. Sea $v \in V$. Entonces existe $(v_1, \dots, v_r) \in V_1 \times \dots \times V_r$ para el cual se tiene $v = v_1 + \dots + v_r$. Veamos que esta descomposición es única. De hecho, si $(v'_1, \dots, v'_r) \in V_1 \times \dots \times V_r$ es tal que se tiene $v = v'_1 + \dots + v'_r$, dado $i \in \{1, \dots, r\}$, se obtiene

$$\underbrace{v_i - v'_i}_{\in V_i} = \underbrace{\sum_{j \neq i} (v'_j - v_j)}_{\in \sum_{j \neq i} V_j}.$$

Luego tenemos $v_i - v'_i \in V_i \cap \sum_{j \neq i} V_j = \{O\}$, es decir $v_i - v'_i = O$ y así $v_i = v'_i$. \square

Proposición 1.80. Sean $V_1, \dots, V_r \leq V$ tales que $V = \sum_{i=1}^r V_i$ y suponga que V tiene dimensión finita. Entonces la siguientes propiedades son equivalentes.

(i) Para todo $i \in \{1, \dots, r\}$ se tiene $V_i \cap \sum_{j \neq i} V_j = \{O\}$.

(ii) Se tiene $\sum_{i=1}^n \dim(V_i) = \dim(V)$.

Dem. Suponga primero que tenemos $V_i \cap \sum_{j \neq i} V_j = \{O\}$, para todo $i \in \{1, \dots, r\}$. Por el teorema 1.74 se obtiene

$$\dim(V) = \dim(V_1) + \dim\left(\sum_{j>1} V_j\right).$$

La inclusión $V_2 \cap \sum_{j>2} V_j \subseteq V_2 \cap \sum_{j \neq 2} V_j$ implica la igualdad $(V_2 \cap \sum_{j>2} V_j) = \{O\}$. Por el teorema 1.74 se obtiene

$$\dim\left(\sum_{j>1} V_j\right) = \dim(V_2) + \dim\left(\sum_{j>2} V_j\right).$$

Inductivamente, obtenemos

$$\begin{aligned}
 \dim(V) &= \dim(V_1) + \dim\left(\sum_{j>1} V_j\right) \\
 &= \dim(V_1) + \dim(V_2) + \dim\left(\sum_{j>2} V_j\right) \\
 &\vdots \\
 &= \dim(V_1) + \cdots + \dim(V_r)
 \end{aligned}$$

Suponga ahora que se tiene $\sum_{i=1}^r \dim(V_i) = \dim(V)$. Por el teorema 1.74 se tienen las desigualdades

$$\begin{aligned}
 \dim(V) &\leq \dim(V_1) + \dim\left(\sum_{j>1} V_j\right) \\
 &\leq \dim(V_1) + \dim(V_2) + \dim\left(\sum_{j>2} V_j\right) \\
 &\vdots \\
 &\leq \sum_{i=1}^r \dim(V_i).
 \end{aligned}$$

Pero se tiene $\sum_{i=1}^r \dim(V_i) = \dim(V)$, luego estas desigualdades son igualdades y en particular obtenemos $\dim(V) = \dim(V_1) + \dim\left(\sum_{j>1} V_j\right)$. De donde, por el mismo teorema 1.74, se tiene $\dim(V_1 \cap \sum_{j \neq 1} V_j) = 0$, es decir $V_1 \cap \sum_{j \neq 1} V_j = \{0\}$. Reordenando los subespacios V_i , $i = 1, \dots, n$, obtenemos $V_i \cap \sum_{j \neq i} V_j = \{0\}$, para todo $i \in \{1, \dots, r\}$. \square

Observación 1.81. De la proposición anterior se sigue que si $\mathcal{B} = (v_1, \dots, v_n)$ es una base de V y C_1, \dots, C_r forman una partición de \mathcal{B} , entonces $V = V_1 \oplus \cdots \oplus V_r$ donde $V_i = \langle C_i \rangle$, $i \in \{1, \dots, r\}$. Y reciprocamente, si $V = V_1 \oplus \cdots \oplus V_r$ y \mathcal{B}_i es una base de V_i para $i \in \{1, \dots, r\}$, la concatenación de estas bases $\mathcal{B} = (\mathcal{B}_1 | \cdots | \mathcal{B}_r)$ es una base de V .

Definición 1.82. Sea $p \in \text{End}_{\mathbb{K}}(V)$. Decimos que p es una *proyección* si $p \circ p = p$.

Observación 1.83. Si $p \in \text{End}_{\mathbb{K}}(V)$ es una proyección y V_0 es la imagen de p , entonces $p(v_0) = v_0$ para todo $v_0 \in V_0$. De hecho si v_0 pertenece a V_0 , existe $v \in V$ que satisface $p(v) = v_0$, luego obtenemos $p(v_0) = p \circ p(v) = p(v) = v_0$.

Observación 1.84. Suponga que tenemos $V = V_1 \oplus V_2$, y defina los operadores $p_1, p_2 \in \text{End}_{\mathbb{K}}(V)$ por $p_1(v) = v_1$ y $p_2(v) = v_2$ si se tiene $v = v_1 + v_2$ con $(v_1, v_2) \in V_1 \times V_2$. Note que p_1 y p_2 son proyecciones que cumplen $p_1 \circ p_2 = p_2 \circ p_1 = \underline{O}$ y $p_1 + p_2 = \text{id}_V$. Similarmente, si tenemos $V = \bigoplus_{i=1}^n V_i$, podemos definir n proyecciones p_1, \dots, p_n que satisfacen $p_i(V) = V_i$, para cada $i \in \{1, \dots, n\}$, $\sum_{i=1}^n p_i = \text{id}_V$, y $p_i \circ p_j = \underline{O}$ si $i \neq j$. Esto nos sugiere otra forma de caracterizar sumas directas, como se hace en el siguiente teorema.

Teorema 1.85. Sean $p_1, \dots, p_n \in \text{End}_{\mathbb{K}}(V)$ proyecciones y V_1, \dots, V_n sus respectivas imágenes. Si $\sum_{i=1}^n p_i = \text{id}_V$ y $p_i \circ p_j = 0$ para $i \neq j$, entonces $V = \bigoplus_{i=1}^n V_i$.

Dem. Usamos la propiedad 1.79 para establecer este teorema. Veamos que tenemos $V = \sum_{i=1}^n V_i$. Dado $v \in V$, tomamos, para $i \in \{1, \dots, n\}$, $v_i \in V_i$ con $v_i = p_i(v)$ y obtenemos

$$v = \text{id}_V(v) = \sum_{i=1}^n p_i(v) = \sum_{i=1}^n v_i.$$

Basta ahora establecer que $V_i \cap \sum_{j \neq i} V_j = \{O\}$ para todo $i \in \{1, \dots, n\}$. Para ello, fijemos un índice i , tomemos $v \in V_i \cap \sum_{j \neq i} V_j$ y veamos que $v = O$. Como v pertenece a $\sum_{j \neq i} V_j$, tenemos $v = \sum_{j \neq i} v_j$ para algunos $v_j \in V_j$, con $j \neq i$. En particular, para cada $j \neq i$, existe $v'_j \in V$ que satisface $v_j = p_j(v'_j)$, y así obtenemos $v = \sum_{j \neq i} p_j(v_j)$. Por otro lado, como v pertenece a V_i , existe $v'_i \in V$ que satisface $v = p_i(v'_i)$. Por ende, se siguen las igualdades

$$v = p_i(v'_i) = p_i \circ p_i(v'_i) = p_i(v) = p_i \left(\sum_{j \neq i} p_j(v_j) \right) = \sum_{j \neq i} p_i \circ p_j(v_j) = O.$$

□

Definición 1.86. Suponga que $V = V_1 \oplus \dots \oplus V_r$. Para $i \in \{1, \dots, r\}$, sean $f_{V_i} \in \text{Hom}_{\mathbb{K}}(V_i, V_i)$. La suma directa $f = f_{V_1} \oplus \dots \oplus f_{V_r}$ es el operador en $\text{End}_{\mathbb{K}}(V)$ definido por

$$f(v) = f_{V_1}(v_1) + \dots + f_{V_r}(v_r)$$

donde $(v_1, \dots, v_r) \in V_1 \times \dots \times V_r$ es tal que $v = v_1 + \dots + v_r$.

Observación 1.87. Suponga que $V = V_1 \oplus \dots \oplus V_r$. Si para $i \in \{1, \dots, r\}$, se tiene una base \mathcal{B}_i de V_i , y $\mathcal{B} = (\mathcal{B}_1 | \dots | \mathcal{B}_r)$ es la base de V que se obtiene al concatenar estas bases, cuando $f = f_{V_1} \oplus \dots \oplus f_{V_r}$ con $f_{V_i} \in \text{Hom}_{\mathbb{K}}(V_i, V_i)$, entonces la matriz de f para la base \mathcal{B}

$$A = \left[\begin{array}{c|c|c} A_1 & \cdots & 0 \\ \hline \vdots & \ddots & \vdots \\ \hline 0 & \cdots & A_r \end{array} \right]$$

es una matriz diagonal por bloques donde cada bloque cooresponde a la matriz $A_i = [f_{V_i}]_{\mathcal{B}_i}^{\mathcal{B}_i}$.

Observación 1.88. Para terminar esta sección, vamos a definir dos generalizaciones de la suma directa, que son la suma directa externa y el producto directo. En estas definiciones combinamos una colección, no necesariamente finita, de espacios para obtener un nuevo espacio. La suma directa externa y el producto directo de una misma colección finita de espacios producen espacio isomorfos.

Definición 1.89. Sea I una colección de índices y $(V_i)_{i \in I}$ una familia de espacios vectoriales sobre \mathbb{K} . Definimos los siguientes espacios.

(i) El *producto directo* de $\{V_i\}_{i \in I}$ es el espacio

$$\prod_{i \in I} V_i = \left\{ \phi : I \rightarrow \prod_{i \in I} V_i \mid \phi(i) \in V_i \right\},$$

el cual es un espacio vectorial sobre \mathbb{K} bajo las operaciones

$$(\phi + \psi)(i) = \phi(i) + \psi(i) \quad (c\phi)(i) = c\psi(i)$$

para todo $\phi, \psi \in \prod_{i \in I} V_i$ y $c \in \mathbb{K}$. Para $j \in I$, definimos la j -ésima inclusión como la transformación lineal

$$\begin{aligned} \iota_j : V_j &\longrightarrow \prod_{i \in I} V_i \\ v_j &\longmapsto \iota_j(v_j) : \begin{cases} j &\mapsto v_j \\ i &\mapsto 0 \text{ si } i \neq j \end{cases} \end{aligned}$$

(ii) La *suma directa externa* de $\{V_i\}_{i \in I}$ es el espacio

$$\bigoplus_{i \in I} V_i = \left\{ \phi \in \prod_{i \in I} V_i \mid \phi(i) \neq 0 \text{ únicamente para finitos índices } i \in I \right\},$$

el cual es un subespacio de $\prod_{i \in I} V_i$. Para $j \in I$, definimos la j -ésima proyección como la transformación lineal

$$\begin{aligned} p_j : \bigoplus_{i \in I} V_i &\longrightarrow V_j \\ \phi &\longmapsto \phi(j) \end{aligned}$$

Observación 1.90. Note que si $(V_i)_{i \in I}$ es una familia de espacios vectoriales sobre \mathbb{K} indexada por los índices $i \in I$, y nos son dadas $f_i \in \text{End}_{\mathbb{K}}(V, V_i)$ y $g_i \in \text{Hom}_{\mathbb{K}}(V_i, V)$, para todo $i \in I$, podemos definir las transformaciones lineales

$$\begin{aligned} f : V &\longrightarrow \prod_{i \in I} V_i & g : \bigoplus_{i \in I} V_i &\longrightarrow V \\ v &\longmapsto f(v) : i \mapsto f_i(v) & \phi &\longmapsto \sum_{i \in I} g_i(\phi(i)). \end{aligned}$$

Note que la suma $\sum_{i \in I} g_i(\phi(i))$ es finita pues $\phi(i) = 0$ para todos los $i \in I$ salvo un número finito de índices.

Observación 1.91. Si $\mathcal{B} = (v_i)_{i \in I}$ es una base de V , entonces se obtienen los isomorfismos

$$V \simeq_{\mathbb{K}} (\mathbb{K}^I)_0 \simeq_{\mathbb{K}} \bigoplus_{i \in I} \mathbb{K}.$$

En general, se tiene

$$\mathbb{K}^I \simeq_{\mathbb{K}} \prod_{i \in I} \mathbb{K}.$$

Definición 1.92. Suponga que $(V_i)_{i \in I}$ y $(W_i)_{i \in I}$ son familias de espacios vectoriales sobre \mathbb{K} indexadas por I . Si para cada $i \in I$ tenemos un $f_i \in \text{Hom}_{\mathbb{K}}(V_i, W_i)$, definimos las siguientes funciones.

(i) El *producto externo* de $(f_i)_{i \in I}$ por

$$\begin{aligned} \prod_{i \in I} f_i : \prod_{i \in I} V_i &\longrightarrow \prod_{i \in I} W_i \\ \phi &\longmapsto \left(\prod_{i \in I} f_i \right) (\phi) : i \mapsto f_i(\phi(i)). \end{aligned}$$

(ii) La *suma directa externa* de $(f_i)_{i \in I}$ por

$$\begin{aligned} \bigoplus_{i \in I} f_i : \bigoplus_{i \in I} V_i &\longrightarrow \bigoplus_{i \in I} W_i \\ \phi &\longmapsto \left(\prod_{i \in I} f_i \right) (\phi) : i \mapsto f_i(\phi(i)), \end{aligned}$$

la cual es la función inducida por $\prod_{i \in I} f_i$ entre los subespacios $\bigoplus_{i \in I} V_i$ y $\bigoplus_{i \in I} W_i$.

Ejercicios

1. Suponga que $V_1, V_2, V_3 \leq V$ son tales que $\dim(V_1 + V_2 + V_3) < \infty$. Demuestre que

$$\begin{aligned} \dim(V_1 + V_2 + V_3) &= \dim(V_1) + \dim(V_2) + \dim(V_3) \\ &\quad - \dim(V_1 \cap V_2) - \dim(V_1 \cap V_3) - \dim(V_2 \cap V_3) \\ &\quad + \dim(V_1 \cap V_2 \cap V_3) \end{aligned}$$

2. Suponga que $V_1, \dots, V_r \leq V$ son tales que $\dim(V_1 + \dots + V_r) < \infty$. Demuestre que

$$\begin{aligned} \dim(V_1 + \dots + V_r) &= \sum_i^r \dim(V_i) - \sum_{i < j} \dim(V_i \cap V_j) + \sum_{i < j < k} \dim(V_i \cap V_j \cap V_k) \\ &\quad - \dots + (-1)^{r-1} \dim(V_1 \cap \dots \cap V_r) \end{aligned}$$

3. Demuestre la afirmación hecha en la observación 1.81.
4. Demuestre la afirmación hecha en la observación 1.87.
5. Sea $p_1 \in \text{End}_{\mathbb{K}}(V)$ una proyección. Demuestre que $p_2 = \text{id}_V - p_1$ es una proyección y que $V = V_1 \oplus V_2$ donde $V_1 = \text{im}(p_1)$ y $V_2 = \text{im}(p_2)$. Pruebe además que $V_1 = \ker(p_2)$ y $V_2 = \ker(p_1)$.
6. En \mathbb{Q}^3 , tome $v_1 = (1, 0, 1)$, $v_2 = (0, 1, 1)$, $v_3 = (1, 1, 0)$ y $\mathcal{B} = (v_1, v_2, v_3)$. Sea $V_1 = \langle v_1, v_2 \rangle$ y $V_2 = \langle v_3 \rangle$, y p_1 la proyección con imagen V_1 y núcleo V_2 , y p_2 la proyección con imagen V_2 y núcleo V_1 .
 - a) Encuentre las matrices de p_1 y p_2 para la base \mathcal{B} .
 - b) Encuentre las matrices canónicas de p_1 y p_2 .
7. En \mathbb{Q}^3 , tome $v_1 = (2, 1, 1)$, $v_2 = (1, 2, 1)$, $v_3 = (1, 1, 2)$ y $\mathcal{B} = (v_1, v_2, v_3)$. Sea $V_1 = \langle v_1 \rangle$ y $V_2 = \langle v_2, v_3 \rangle$, y p_1 la proyección con imagen V_1 y núcleo V_2 , y p_2 la proyección con imagen V_2 y núcleo V_1 . Encuentre las matrices canónicas de p_1 y p_2 .
8. Suponga que $V = V_1 \oplus \cdots \oplus V_r$ y que p_1, \dots, p_r son las proyecciones sobre cada uno de los sumandos de acuerdo a esta descomposición de V como suma directa. Tomamos un operador $f \in \text{End}_{\mathbb{K}}(V)$ y una base \mathcal{B} de V .
 - a) Demuestre que, cuando $\dim(V) < \infty$, se tiene $A = A_1 + \cdots + A_r$, donde A es la matriz de f para \mathcal{B} , y para $i \in \{1, \dots, r\}$, $A_i = AP_i$ donde P_i es la matriz de p_i para \mathcal{B} .
 - b) Suponga además que $f \in \text{End}_{\mathbb{K}}(V)$ es tal que $f = f_{V_1} \oplus \cdots \oplus f_{V_r}$ con $f_{V_i} \in \text{End}_{\mathbb{K}}(V_i)$. Demuestre que se tiene $AP_i = P_i A$ para $i \in \{1, \dots, r\}$.
9. Demuestre que las j -ésimas inclusiones ι_j y proyecciones p_j en la definición 1.89 son transformaciones lineales.
10. Demuestre que las funciones f y g definidas en la observación 1.90 son únicas con la propiedad $f \circ \iota_j = f_j$ y $p_j \circ g = g_j$ para todo $j \in I$.
11. Demuestre que las funciones $\prod_{i \in I} f_i$ y $\bigoplus_{i \in I} f_i$ definidas en la definición 1.92 son transformaciones lineales.

1.6. Espacios cociente

Sean V, W espacios vectoriales sobre \mathbb{K} .

Definición 1.93. Dados $V_0 \leq V$ y $v \in V$, la *translación de V_0 por v* es el conjunto

$$v + V_0 = \{v' \in V \mid v' = v + v_0, v_0 \in V_0\}.$$

Observación 1.94. Tenemos $v + V_0 = v' + V_0$ si y solo si $v - v' \in V_0$. De hecho, si $v + V_0 = v' + V_0$, como $v \in v + V_0 = v' + V_0$, entonces existe $v_0 \in V_0$ tal que $v = v' + v_0$, es decir $v - v' = v_0 \in V_0$; recíprocamente, si $v_0 = v - v' \in V_0$, cualquier $w \in v + V_0$ es de la forma $w = v + w_0$ para algún $w_0 \in V_0$, en particular $w = v' + (v_0 + w_0) \in v' + V_0$, y cualquier $w' \in v' + V_0$ es de la forma $w' = v' + w'_0$ para algún $w'_0 \in V_0$, en particular $w' = v + (w'_0 - v_0) \in v + V_0$.

Definición 1.95. Sea $V_0 \leq V$, el *espacio cociente* V módulo V_0 es el conjunto de traslaciones de V_0 :

$$V/V_0 = \{v + V_0 \mid v \in V\}$$

Proposición 1.96. Para todo $V_0 \leq V$ se tienen las siguientes propiedades.

- (i) Si $v, v', w, w' \in V$ son tales que $v + V_0 = v' + V_0$ y $w + V_0 = w' + V_0$ entonces $(v + w) + V_0 = (v' + w') + V_0$
- (ii) Si $v, v' \in V$ son tales que $v + V_0 = v' + V_0$, entonces $cv + V_0 = cv' + V_0$ para todo $c \in \mathbb{C}$.

Dem. Se tiene que $v + V_0 = v' + V_0$ y $w + V_0 = w' + V_0$ si y solo si $v - v' \in V_0$ y $w - w' \in V_0$. En tal caso se sigue que $(v + w) - (v' + w') = (v - v') + (w - w') \in V_0$, y por ende $(v + w) + V_0 = (v' + w') + V_0$. Similarmente, se tiene $cv - cv' = c(v - v') \in V_0$, y por ello $cv + V_0 = cv' + V_0$. \square

Proposición 1.97. Sea $V_0 \leq V$. El espacio cociente V/V_0 es un espacio vectorial sobre \mathbb{K} bajo las operaciones

$$(v + V_0) + (w + V_0) = (v + w) + V_0 \quad c(v + V_0) = cv + V_0,$$

y con origen $O + V_0 = V_0$. El mapa

$$\begin{aligned} \pi_{V_0} : V &\longrightarrow V/V_0 \\ v &\longmapsto v + V_0 \end{aligned}$$

es una transformación lineal sobreyectiva con $\ker(\pi_{V_0}) = V_0$

Dem. La proposición anterior garantiza que tales operaciones están bien definidas. Las propiedades de la definición 1.1 se heredan de las de V . La misma proposición implica la linealidad de π_{V_0} . Por definición de V/V_0 , π_{V_0} es sobreyectiva. Por último, se tiene $v \in \ker(\pi_{V_0})$ si y solo si $\pi_{V_0}(v) = V_0$, es decir si y solo si $v + V_0 = V_0$, que equivale a $v \in V_0$. \square

Proposición 1.98. Sea $V_0 \leq V$ y suponga que V tiene dimensión finita, entonces

$$\dim(V/V_0) = \dim(V) - \dim(V_0)$$

Dem. Se sigue inmediatamente del teorema 1.45 del rango y de la proposición anterior.

Proposición 1.99. *Suponga que V tiene dimensión finita y sea $V_0 \leq V$. Si $\mathcal{B}_0 = (v_1, \dots, v_{n_0})$ es una base de V_0 y la extendemos a una base $\mathcal{B} = (v_1, \dots, v_{n_0}, v_{n_0+1}, \dots, v_{n_0+n_1})$ de V , entonces $\mathcal{B}_1 = (v_{n_0+1} + V_0, \dots, v_{n_0+n_1} + V_0)$ es una base de V/V_0 .*

Dem. Veamos que \mathcal{B}_1 es linealmente independiente. Si $c_1(v_{n_0+1} + V_0) + \dots + c_{n_1}(v_{n_0+n_1} + V_0) = O + V_0$ entonces $c_1 v_{n_0+1} + \dots + c_{n_1} v_{n_0+n_1}$ está en V_0 , luego existen $a_1, \dots, a_{n_0} \in \mathbb{K}$ tales que $c_1 v_{n_0+1} + \dots + c_{n_1} v_{n_0+n_1} = a_1 v_1 + \dots + a_{n_0} v_{n_0}$, y así

$$-a_1 v_1 - \dots - a_{n_0} v_{n_0} + c_1 v_{n_0+1} + \dots + c_{n_1} v_{n_0+n_1} = O.$$

La independencia lineal de los elementos en \mathcal{B} implica que $c_1 = \dots = c_{n_0} = 0$. Veamos que \mathcal{B}_1 genera a V/V_0 . De hecho, dado $v + V_0 \in V/V_0$ con

$$v = a_1 v_1 + \dots + a_{n_0} v_{n_0} + c_1 v_{n_0+1} + \dots + c_{n_1} v_{n_0+n_1},$$

si $v' = c_1 v_{n_0+1} + \dots + c_{n_1} v_{n_0+n_1}$, entonces $v - v' = a_1 v_1 + \dots + a_{n_0} v_{n_0}$ está en V_0 , luego $v + V_0 = v' + V_0$, y así $v + V_0 = c_1(v_{n_0+1} + V_0) + \dots + c_{n_1}(v_{n_0+n_1} + V_0)$. \square

Teorema 1.100. *Sea $f \in \text{Hom}_{\mathbb{K}}(V, W)$. Si $V_0 = \ker(f)$, entonces existe una única transformación lineal $f_0 \in \text{Hom}_{\mathbb{K}}(V/V_0, W)$ tal que $f = f_0 \circ \pi_{V_0}$. La transformación f_0 es inyectiva, y, si f es además sobreyectiva, entonces f_0 es un isomorfismo.*

Dem. Note que $f(v) = f(v')$ si y solo si $v - v' \in V_0$, es decir si y solo si $v + V_0 = v' + V_0$. Defina entonces

$$\begin{aligned} f_0 : V/V_0 &\longrightarrow W \\ v + V_0 &\longmapsto f(v). \end{aligned}$$

Así, f_0 es lineal pues f lo es, y además es inyectiva pues $f(v) = f(v')$ si y solo si $v + V_0 = v' + V_0$. Por contrucción $f = f_0 \circ \pi_{V_0}$. Ahora si f es sobreyectiva, entonces f_0 es biyectiva y, así, es un isomorfismo. \square

Teorema 1.101. *Sea $f \in \text{End}_{\mathbb{K}}(V)$ y $V_0 \leq V$ tal que $f(V_0) \leq V_0$, entonces la función*

$$\begin{aligned} f_{/V_0} : V/V_0 &\longrightarrow V/V_0 \\ v + V_0 &\longmapsto f(v) + V_0 \end{aligned}$$

define un operador en $\text{End}_{\mathbb{K}}(V/V_0)$

Dem. Veamos que $f_{/V_0}$ está bien definida. Si $v, v' \in V$ son tales que $v + V_0 = v' + V_0$, es decir $v - v' \in V_0$, entonces $f(v) - f(v') = f(v - v') \in V_0$, así $f(v) + V_0 = f(v') + V_0$. Ahora, la linealidad de $f_{/V_0}$ se hereda de la de f .

Observación 1.102. *Suponga que V tiene dimensión finita y que $V_0 \leq V$ es tal que $f(V_0) \leq V_0$ para algún operador $f \in \text{End}_{\mathbb{K}}(V)$. Considere bases $\mathcal{B}_0, \mathcal{B}$ y*

\mathcal{B}_1 de V_0 , V y V/V_0 como en la proposición 1.99. Se tiene entonces que la matriz de f para \mathcal{B} es de la forma

$$[f]_{\mathcal{B}} = \left[\begin{array}{c|c} A & B \\ \hline O & C \end{array} \right]$$

donde A , B y C son respectivamente matrices $n_0 \times n_0$, $n_0 \times n_1$ y $n_1 \times n_1$; y la matriz de $f_{/V_0}$ para \mathcal{B}_1 es

$$[f_{/V_0}]_{\mathcal{B}_1} = C.$$

Ejercicios

1. Considere el operador f de \mathbb{Q}^3 dado por $f(x, y, z) = (-2x + 4z, -x + 3y, -3x + 2y + 3z)$. Verifique que para $V_0 = \langle (1, 1, 1) \rangle$ se tiene $f(V_0) = V_0$. Encuentre una base de V/V_0 y la matriz de $f_{/V_0}$ para esta base.
2. Considere el operador f de \mathbb{Q}^4 dado por $f(x, y, z, w) = (5x + 5z + w, -3x - 3y, -3x + y, x - y + w)$. Verifique que para $V_0 = \langle e_1, e_4 \rangle$ se tiene $f(V_0) = V_0$. Encuentre una base de V/V_0 y la matriz de $f_{/V_0}$ para esta base.
3. Dada $f \in \text{Hom}_{\mathbb{K}}(V, W)$ definimos el co-núcleo de f por $\text{cok}(f) = W/\text{im}(f)$. Demuestre que f es sobreyectiva si y solo si $\text{cok}(f) = \{O + \text{im}(f)\}$.
4. Suponga que $f \in \text{Hom}_{\mathbb{K}}(V, W)$ es una transformación lineal donde W es un espacio finito-dimensional. Demuestre que $\rho(f) + \dim(\text{cok}(f)) = \dim(W)$. Compare esta igualdad con el teorema 1.45 del rango.
5. Sea $f \in \text{Hom}_{\mathbb{K}}(V, W)$ y suponga que $\dim(W) < \infty$. Demuestre que las siguientes propiedades son equivalentes.

- (i) $\dim(\text{cok}(f)) = 0$
- (ii) $\dim(\text{im}(f)) = \dim(W)$

Compare esta equivalencia con el corolario 1.46 del teorema 1.45 del rango.

Capítulo 2

Estructura de las transformaciones lineales

Sea \mathbb{K} un cuerpo.

2.1. Descomposición por suma directa

Sean V y W espacios vectoriales sobre \mathbb{K} .

Definición 2.1. Sean V_1, \dots, V_r subespacios de V . Cuando tenemos $V = V_1 \oplus \dots \oplus V_r$ llamamos a esta igualdad una *descomposición en suma directa*, o simplemente una *descomposición*, de V .

Teorema 2.2. Dada $f \in \text{Hom}_{\mathbb{K}}(V, W)$, existen $V_0, V_1 \leq V$ y $W_1, W_2 \leq W$ tales que $V = V_0 \oplus V_1$ y $W = W_1 \oplus W_2$ con $\ker(f) = V_0$ e $\text{im}(f) = W_1$. En particular f induce un isomorfismo entre V_1 y W_1 .

Dem. Sea \mathcal{B}_0 una base de $V_0 = \ker(f)$, la cual extendemos a una base $(\mathcal{B}_0 | \mathcal{B}_1)$ de V . Defina $V_1 = \langle \mathcal{B}_1 \rangle$. Así, por la observación 1.81 se tiene $V = V_0 \oplus V_1$. Por otro lado, si $v, v' \in V_1$ son tales que $f(v) = f(v')$, entonces $v - v' \in \ker(f) = V_0$, luego $v - v' \in V_0 \cap V_1 = \{0\}$, luego $v = v'$. Es decir la restricción de f a V_1 es inyectiva.

Sea $\mathcal{B}'_1 = f(\mathcal{B}_1)$. Como la restricción de f a V_1 es inyectiva, es decir que $f(v) = 0$ con $v \in V_1$ si y solo si $v = 0$, entonces \mathcal{B}'_1 es linealmente independiente. Defina $W_1 = \langle \mathcal{B}'_1 \rangle$, de forma que \mathcal{B}'_1 es una base de W_1 y $W_1 = f(V_1)$. Por construcción $\text{im}(f) = W_1$; pues, dado $w \in \text{im}(f)$, existe $v \in V$ tal que $w = f(v)$, si $v = v_0 + v_1$ con $(v_0, v_1) \in V_0 \times V_1$, se tiene $w = f(v) = f(v_0) + f(v_1) = f(v_1)$. Finalmente, extienda \mathcal{B}'_1 a una base $(\mathcal{B}'_1 | \mathcal{B}'_2)$ de W . Si $W_2 = \langle \mathcal{B}'_2 \rangle$ se tiene $W = W_1 \oplus W_2$. Como f es inyectiva en V_1 y $f(V_1) = W_1$, f induce un isomorfismo entre V_1 y W_1 . \square

Observación 2.3. Suponga que en el teorema 2.2 los espacios V y W tienen dimensión finita, con $n = \dim(V)$, $m = \dim(W)$ y $r = \dim(\text{im}(f))$. Si, como en

la demostración, $\mathcal{B}_0, \mathcal{B}_1$ son bases de V_0 y V_1 , y \mathcal{B}'_1 y \mathcal{B}'_2 son bases de W_1 y W_2 . Entonces, para las bases $\mathcal{B} = (\mathcal{B}_1 | \mathcal{B}_0)$ y $\mathcal{B}' = (\mathcal{B}'_1 | \mathcal{B}'_2)$ de V y W , en la matriz de f

$$A = \left[f \right]_{\mathcal{B}}^{\mathcal{B}'} = (a_{ij}),$$

se tiene $a_{ii} = 1$ si $0 \leq i \leq r$ y $a_{ij} = 0$ si $i \neq j$, o si $r < i$ e $i = j$. Es decir

$$A = \left[\begin{array}{c|c} I_r & 0 \\ \hline 0 & 0 \end{array} \right]$$

donde I_r denota la matriz identidad $r \times r$ y 0 denota los orígenes de $M_{r \times (n-r)}(\mathbb{K})$, $M_{(m-r) \times r}(\mathbb{K})$ y $M_{(m-r) \times (n-r)}(\mathbb{K})$.

Ejercicios

1. Para cada una de las siguientes transformaciones lineales f encuentre bases para el dominio y el rango respecto a las cuales la matriz de f tiene un matriz identidad en la esquina superior izquierda y ceros en el resto.

- (i) $f \in \text{Hom}_{\mathbb{Q}}(\mathbb{Q}^2, \mathbb{Q}^3)$ donde

$$f(x, y) = (y, x, -x).$$

- (ii) $f \in \text{Hom}_{\mathbb{Q}}(\mathbb{Q}^2, \mathbb{Q}^3)$ donde

$$f(x, y) = \frac{1}{2}(x + y, x + y, -x - y).$$

- (iii) $f \in \text{Hom}_{\mathbb{Q}}(\mathbb{Q}^2, \mathbb{Q}^4)$ donde

$$f(x, y) = \frac{1}{3}(2y, 2y, 3x - y, -3x - y).$$

- (iv) $f \in \text{Hom}_{\mathbb{Q}}(\mathbb{Q}^3, \mathbb{Q}^4)$ donde

$$f(x, y) = \frac{1}{3}(2x + y + z, 2x + y + z, -x + y - 2z, -x - 2y + z).$$

2. Suponga que $f \in \text{Hom}_{\mathbb{K}}(V, W)$ es una transformación lineal entre espacios finito-dimensionales y sean $V_0, V_1 \leq V$ y $W_1, W_2 \leq W$ como en el teorema 2.2. Demuestre que $\dim(\text{cok}(f)) = \dim(W_2)$ (El espacio co-núcleo $\text{cok}(f)$ está definido en el ejercicio 1.6.3).

2.2. Espacios estables, cíclicos y propios

Sea V un espacio vectorial sobre \mathbb{K} y f un operador de V , e.d. $f \in \text{End}_{\mathbb{K}}(V)$.

Observación 2.4. Sean $A = (a_{ij})_{i,j=1}^n$ y $C = (c_{ij})_{i,j=1}^n$ matrices $n \times n$. Se tiene

$$\operatorname{tr}(AC) = \sum_{i=1}^n \sum_{j=1}^n a_{ij} c_{ji} = \sum_{j=1}^n \sum_{i=1}^n c_{ij} a_{ji} = \sum_{i=1}^n \sum_{j=1}^n c_{ij} a_{ji} = \operatorname{tr}(CA).$$

Ahora, si V tiene dimensión finita igual a n , dadas dos bases \mathcal{B} y \mathcal{B}' de V , tenemos dos matrices $n \times n$ de f , $A = [f]_{\mathcal{B}}^{\mathcal{B}}$ y $B = [f]_{\mathcal{B}'}^{\mathcal{B}'}$. Entonces, si además $C = [\operatorname{id}_V]_{\mathcal{B}'}^{\mathcal{B}}$ es la matriz de cambio de coordenadas de \mathcal{B} a \mathcal{B}' , se tiene

$$B = C^{-1}AC,$$

y

$$\begin{aligned} \operatorname{tr}(B) &= \operatorname{tr}(C^{-1}AC) = \operatorname{tr}(ACC^{-1}) \\ &= \operatorname{tr}(A) \\ \det(B) &= \det(C^{-1}AC) = \det(C)^{-1} \det(A) \det(C) \\ &= \det(A) \end{aligned}$$

Es decir que la traza y el determinante de la matriz de un operador lineal, tomando la misma base para el dominio y el rango, es independiente de la base escogida.

Definición 2.5. Suponga que V tiene dimensión finita. Definimos el *determinante* y la *traza* de f respectivamente por

$$\det(f) = \det \left([f]_{\mathcal{B}}^{\mathcal{B}} \right) \quad \operatorname{tr}(f) = \operatorname{tr} \left([f]_{\mathcal{B}}^{\mathcal{B}} \right).$$

donde \mathcal{B} es cualquier base de V .

Definición 2.6. Sea $V_0 \leq V$. Decimos que V_0 es *estable bajo f* (ó *f -estable*) si $f(V_0) \leq V_0$. En tal caso, a la restricción de f a V_0 la denotamos f_{V_0} , esto es $f_{V_0} \in \operatorname{Hom}_{\mathbb{K}}(V_0, V_0)$ es el operador definido por:

$$\begin{aligned} f_{V_0} : V_0 &\longrightarrow V_0 \\ v_0 &\longmapsto f(v_0) \end{aligned}$$

Observación 2.7. De acuerdo con la proposición 1.101, cuando $V_0 \leq V$ es estable bajo f , el operador f induce un operador $f_{/V_0}$ de V/V_0 .

Observación 2.8. Suponga que V es finito-dimensional. Dado $v \in V$ no nulo, existe un entero r maximal tal que $\{v, f(v), \dots, f^{r-1}(v)\}$ es linealmente independiente. Luego, por la propiedad 1.24, existen $a_0, \dots, a_{r-1} \in \mathbb{K}$ tales que

$$f^r(v) = a_{r-1}f^{r-1}(v) + \dots + a_1f(v) + a_0v.$$

De esta forma, $V_0 = \langle v, f(v), \dots, f^{r-1}(v) \rangle$ es f -estable. Más aún, si \mathcal{B} es una base de V que extiende la base $\mathcal{B}_0 = (v, f(v), \dots, f^{r-1}(v))$ de V_0 , entonces la matrix de f para \mathcal{B} es de la forma

$$[f]_{\mathcal{B}} = \left[\begin{array}{c|c} A & B \\ \hline O & C \end{array} \right]$$

donde A es la matrix compañera

$$A = \left[\begin{array}{ccccc} 0 & 0 & \cdots & 0 & a_0 \\ 1 & 0 & \cdots & 0 & a_1 \\ 0 & 1 & \cdots & 0 & a_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & a_{r-1} \end{array} \right] = [f_{V_0}]_{\mathcal{B}_0}.$$

Definición 2.9. Si $V_0 \leq V$ es tal que $V_0 = \langle f^i(v) \mid i \in \mathbb{Z}_{\geq 0} \rangle$ para algún vector $v \in V$, decimos que V_0 es un subespacio f -cíclico (o *cíclico bajo f*). Si además $\mathcal{B}_0 = (v, f(v), \dots, f^{r-1}(v))$ es una base de V_0 , decimos que v es un vector f -cíclico (o *cíclico bajo f*) de V_0 .

Observación 2.10. Si $V_0 \leq V$ es un subespacio f -cíclico finito-dimensional, entonces V_0 es f -estable. Si además $v \in V$ es tal que $\mathcal{B}_0 = (v, f(v), \dots, f^{r-1}(v))$ es una base de V_0 , e.d. v es un vector f -cíclico de V_0 , entonces la matrix de f_{V_0} para \mathcal{B}_0 es una matrix compañera.

Definición 2.11. Un escalar $\lambda \in \mathbb{K}$ es un *valor propio* de f si $f(v) = \lambda v$ para algún vector no nulo $v \in V$, y en tal caso decimos que v es un *vector propio* de f con valor propio λ . Si λ es un valor propio de f , al subespacio de vectores $v \in V$ tales que $f(v) = \lambda v$ lo llamamos el *espacio propio de f asociado a λ* .

Observación 2.12. Si $V_0 \leq V$ es el espacio propio asociado al valor propio λ_0 , entonces V_0 es un subespacio f -estable tal que $f_{V_0} = \lambda_0 \text{id}_{V_0}$.

Definición 2.13. Sea $A = (a_{ij})$ una matrix $n \times n$. Decimos que A es *diagonal* si $a_{ij} = 0$ siempre que $i \neq j$. Si V tiene dimensión finita, decimos que el operador f es *diagonalizable* si $A = [f]_{\mathcal{B}}^{\mathcal{B}}$ es diagonal para alguna base \mathcal{B} de V .

Teorema 2.14. Si V tiene dimensión finita, el operador f es diagonalizable si y solo si existe una descomposición $V = V_1 \oplus \cdots \oplus V_r$ en subespacios propios. En tal caso, si V_i es el espacio propio asociado a λ_i , para $i \in \{1, \dots, r\}$, entonces $f = \lambda_1 \text{id}_{V_1} \oplus \cdots \oplus \lambda_r \text{id}_{V_r}$ y $f = \lambda_1 p_1 + \cdots + \lambda_r p_r$ donde p_1, \dots, p_r son las proyecciones de acuerdo a la descomposición $V = V_1 \oplus \cdots \oplus V_r$ (ver la observación 1.84).

Dem. Suponga primero que f es diagonalizable y sea $\mathcal{B} = (v_1, \dots, v_n)$ una base de V tal que $A = (a_{ij}) = [f]_{\mathcal{B}}^{\mathcal{B}}$ es diagonal. Para cada $i \in \{1, \dots, n\}$, defina $V_i = \langle v_i \rangle$, de forma que $V = V_1 \oplus \cdots \oplus V_n$ (ver la observación 1.81). Dado

$j \in \{1, \dots, n\}$, se tiene, por definición de A , $f(v_j) = a_{1j}v_1 + \dots + a_{nj}v_n$. Pero, como A es diagonal, entonces $a_{ij} = 0$ si $i \neq j$ y así $f(v_j) = a_{jj}v_j$. En particular, cada V_j es f -estable, $f_{V_j} = \lambda_j \text{id}_{V_j}$ con $\lambda_j = a_{jj}$ y tenemos que $V = V_1 \oplus \dots \oplus V_n$ es una descomposición en subespacios f -estables. Finalmente, los espacios V_j con un mismo valor λ_j se puede sumar para obtener un espacio propio.

Suponga ahora que tenemos una descomposición $V = V_1 \oplus \dots \oplus V_r$ en espacios propios asociados a los valores propios $\lambda_1, \dots, \lambda_r \in \mathbb{K}$. Para cada $i \in \{1, \dots, r\}$, tomamos una base \mathcal{B}_i de V_i . En particular, si $v \in V_i$, entonces $f(v) = \lambda_i v$. La concatenación $\mathcal{B} = (\mathcal{B}_1 | \dots | \mathcal{B}_r)$ es una base de V . Sea $\mathcal{B} = (v_1, \dots, v_n)$ y $A = (a_{ij}) = \begin{bmatrix} f \end{bmatrix}_{\mathcal{B}}^{\mathcal{B}}$, es decir que $f(v_j) = a_{1j}v_1 + \dots + a_{nj}v_n$. Pero si $v_j \in V_i$, como $f(v_j) = \lambda_i v_j$, entonces $a_{ij} = 0$ si $i \neq j$ y $a_{jj} = \lambda_i$, y por ende A es diagonal.

Finalmente, si $V = V_1 \oplus \dots \oplus V_r$ es una descomposición en espacios propios, entonces para $v = v_1 + \dots + v_r$ con $(v_1, \dots, v_r) \in V_1 \times \dots \times V_r$, se tiene $f(v) = \lambda_1 v_1 + \dots + \lambda_r v_r = \lambda_1 p_1(v) + \dots + \lambda_r p_r(v)$ y así $f = \lambda_1 \text{id}_{V_1} \oplus \dots \oplus \lambda_r \text{id}_{V_r}$ (ver definición 1.86) y $f = \lambda_1 p_1 + \dots + \lambda_r p_r$. \square

Ejercicios

- Demuestre que el operador de \mathbb{Q}^3 dado por $f(x, y, z) = (-2x, -2x + 2z, -2x + 2y)$, y cuyos valores propios son -2 y 2 , es diagonalizable.
- Considere el operador de \mathbb{Q}^3 dado por $f(x, y, z) = (-x + y - z, -2x + 2z, -x + 3y - z)$, cuyos valores propios son -2 y 2 .
 - Demuestre que f no es diagonalizable.
 - Demuestre que $v = (1, 1, 0)$ es un vector f -cíclico para $V_0 = \ker(f + 2 \text{id}_{\mathbb{Q}^3})$.
 - Encuentre una base \mathcal{B} de \mathbb{Q}^3 que extienda a la base $\mathcal{B}_0 = (v, f(v))$ de V_0 y las matrices de f_{V_0} y de f para estas bases.
- Demuestre que si $\lambda \in \mathbb{K}$ es un valor propio de $f \in \text{End}_{\mathbb{K}}(V)$, entonces el espacio propio de f asociado a λ es $\ker(f - \lambda \text{id}_V)$.
- Demuestre que $\lambda \in \mathbb{K}$ es un valor propio de $f \in \text{End}_{\mathbb{K}}(V)$ si y solo si $\det(f - \lambda \text{id}_V) = 0$.
- Demuestre que $f \in \text{End}_{\mathbb{K}}(V)$ es diagonalizable si y solo si el polinomio $P(\lambda) = \det(f - \lambda \text{id}_V)$ se factoriza como $P(\lambda) = (\lambda - \lambda_1)^{m_1} \dots (\lambda - \lambda_r)^{m_r}$ con $\lambda_1, \dots, \lambda_r \in \mathbb{K}$ y V se descompone como $V = V_1 \oplus \dots \oplus V_r$ donde $V_i = \ker(f - \lambda_i \text{id}_V)$, para $i \in \{1, \dots, r\}$.

2.3. Polinomios de operadores, polinomio anulador y polinomio minimal

Definición 2.15. Dado $P(t) \in \mathbb{K}[t]$, con $P(t) = a_n t^n + a_{n-1} t^{n-1} + \cdots + a_1 t + a_0$, definimos el operador $P(f) \in \text{End}_{\mathbb{K}}(V)$ por

$$P(f) = a_n f^n + a_{n-1} f^{n-1} + \cdots + a_1 f + a_0 \text{id}_V,$$

donde

$$f^k = \underbrace{f \circ \cdots \circ f}_{k\text{-veces}}$$

para cualquier entero positivo k , y $f^0 = \text{id}_V$. Similarmente, para toda matriz $A \in M_{n \times n}(\mathbb{K})$ se define la matriz $P(A) \in M_{n \times n}(\mathbb{K})$ por

$$P(A) = a_n A^n + a_{n-1} A^{n-1} + \cdots + a_1 A + a_0 I_n.$$

Decimos que $P(t)$ es un *polinomio anulador* de f (ó de A) si $P(f) = \underline{0}$ (ó $P(A) = O$).

Observación 2.16. (i) Si V tiene dimensión finita, para toda base \mathcal{B} de V y todo $P(t) \in \mathbb{K}[t]$ se tiene

$$[P(f)]_{\mathcal{B}}^{\mathcal{B}} = P([f]_{\mathcal{B}}^{\mathcal{B}}).$$

(ii) Dados $P_1(t), P_2(t) \in \mathbb{K}[t]$, si $P(t) = P_1(t)P_2(t)$, entonces $P(f) = P_1(f) \circ P_2(f) = P_2(f) \circ P_1(f)$, es decir que los operadores $P_1(f)$ y $P_2(f)$ conmutan, pues para los monomios at^m y bt^n se tiene $(af^m) \circ (bf^n) = abf^{m+n} = (bf^n) \circ (af^m)$ y de esto se sigue la igualdad para polinomios más generales.

Ejemplo 2.17. Considere el operador $f \in \text{End}_{\mathbb{K}}(\mathbb{K}^2)$ definido por $f(x, y) = (x + y, y)$ y el polinomio $P(t) = (t - 1)^2$. Como $(f - \text{id}_V)(x, y) = (y, 0)$, entonces $P(f) = \underline{0}$ y $P(t)$ es un polinomio anulador de f .

Definición 2.18. El *polinomio minimal* de f es el polinomio mónico anulador de f de menor grado.

Observación 2.19. El teorema de Cayley-Hamilton (teorema 2.28 abajo) explica que todo operador en un espacio finito dimensional admite un polinomio anulador.

Proposición 2.20. Suponga que $P_0(t) \in \mathbb{K}[t]$ es el polinomio minimal de f . Si $P(t) \in \mathbb{K}[t]$ es un polinomio anulador de f , entonces $P_0(t)$ divide a $P(t)$.

Dem. Sean $Q(t)$ y $R(t)$ el cociente y el residuo en la división de $P(t)$ por $P_0(t)$, luego $\underline{0} = P(f) = Q(f) \circ P_0(f) + R(f) = R(f)$. Por ende, $R(t)$ es un polinomio anulador de f , y como es de menor grado que el polinomio minimal, entonces $R(t) = 0$, es decir que $P_0(t)$ divide a $P(t)$. \square

Proposición 2.21. *Suponga que $P_0(t) \in \mathbb{K}[t]$ es el polinomio minimal de f . Si $W \leq V$ es f -estable, entonces los polinomios minimales de f_W y $f_{/W}$ dividen a $P_0(t)$.*

Dem. Como $P_0(f)(W) = \{O\}$, entonces $P_0(f_W) = \underline{O}$, es decir que $P_0(t)$ es un anulador de f_W y, por la proposición anterior, el polinomio minimal de f_W lo divide. Similarmente, como $P_0(f)(V) = \{O\} \leq W$, entonces $P_0(f_{/W}) = \underline{O}$, es decir que $P_0(t)$ es un anulador de $f_{/W}$ y, por la proposición anterior, el polinomio minimal de $f_{/W}$ lo divide.

Ejercicios

1. Complete la demostración de la observación 2.16.(ii) en la que se afirma que, dados dos polinomios $P_1(t), P_2(t) \in \mathbb{K}[t]$ y un operador $f \in \text{End}_{\mathbb{K}}(V)$, los dos operadores $P_1(f)$ y $P_2(f)$ conmutan.
2. Considere el operador de \mathbb{Q}^3 del ejercicio 2.2.2 dado por $f(x, y, z) = (-x + y - z, -2x + 2z, -x + 3y - z)$, junto con el subespacio f -cíclico $W = \langle v, f(v) \rangle$ donde $v = (0, 0, 1)$. Demuestre que los polinomios minimales de f , f_W y $f_{/W}$ son respectivamente $t^3 + 2t^2 + 8t$, $t^2 + 2t + 8$ y t .

2.4. El polinomio característico y el teorema de Cayley-Hamilton

El teorema de Cayley-Hamilton (teorema 2.28 abajo) nos permitirá generalizar la descomposición del teorema 2.14 con espacios estables más generales que los espacios propios. Esto se hace con el lema de los núcleos (proposición 2.35 abajo).

Como en la sección anterior, V denotará un espacio sobre \mathbb{K} y f un operador de V .

Observación 2.22. Del mismo modo en que definimos una matriz sobre los escalares como una familia de escalares indexados por dos índices, podemos definir matrices sobre $\mathbb{K}[t]$. A este conjunto lo denotaremos $M_{m \times n}(\mathbb{K}[t])$. Podemos sumar y multiplicar matrices con entradas polinomiales cuando las dimensiones de las matrices lo permiten. Además, podemos también hablar del determinante y de la traza de matrices cuadradas con entradas en $\mathbb{K}[t]$, los cuales serán polinomios en $\mathbb{K}[t]$. De hecho, toda matriz con entradas en $\mathbb{K}[t]$ puede ser vista como un caso especial de una matriz con entradas en el cuerpo $\mathbb{K}(t)$.

Observación 2.23. Sea n un entero positivo y A una matriz en $M_{n \times n}(\mathbb{K})$. Dada cualquier $C \in M_{n \times n}(\mathbb{K})$, invertible, tenemos

$$\det(tI_n - A) = \det\left(C^{-1}(tI_n - A)C\right) = \det(tI_n - C^{-1}AC)$$

donde $tI_n - A$ y $tI_n - C^{-1}AC$ son matrices en $M_{n \times n}(\mathbb{K}[t])$. Esta observación nos permite formular la siguiente definición.

Definición 2.24. Suponga que V tiene dimensión finita n . Dado $f \in \text{End}_{\mathbb{K}}(V)$, definimos el *polinomio caraterístico* de f por

$$P_f(t) = \det(tI_n - A) \in \mathbb{K}[t]$$

donde A es la matriz de f para cualquier base \mathcal{B} de V . Similarmente, dada una matriz A de tamaño $n \times n$, definimos su polinomio característico por

$$P_A(t) = \det(tI_n - A) \in \mathbb{K}[t]$$

Observación 2.25. Si V tiene dimensión finita, entonces el grado del polinomio característico de f coincide con la dimensión de V .

Teorema 2.26. Suponga que V tiene dimensión finita y sea λ un escalar. Entonces, λ es un valor propio de f si y solo si $P_f(\lambda) = 0$.

Dem. Sea \mathcal{B} una base de V . El escalar λ es un valor propio de f si y solo si existe $v \in V$, con $v \neq 0$, tal que $f(v) = \lambda v$, o, equivalentemente, tal que $(\lambda \text{id}_V - f)(v) = 0$. Es decir $\lambda \in \mathbb{K}$ es un valor propio de f si y solo si $\lambda \text{id}_V - f$ no es inyectiva, lo que equivale a

$$0 = \det(\lambda \text{id}_V - f) = \det \left(\lambda I_n - [f]_{\mathcal{B}}^{\mathcal{B}} \right) = P_f(\lambda).$$

□

Observación 2.27. El polinomio característico de la matriz compañera (ver observación 2.8)

$$A = \begin{bmatrix} 0 & 0 & \cdots & 0 & a_0 \\ 1 & 0 & \cdots & 0 & a_1 \\ 0 & 1 & \cdots & 0 & a_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & a_{r-1} \end{bmatrix}$$

es $P_A(t) = t^r - a_{r-1}t^{r-1} - \cdots - a_1t - a_0$.

Teorema 2.28 (Cayley-Hamilton). Si V tiene dimensión finita, entonces $P_f(f) = 0$.

Dem. Sea $v \in V$ no nulo. Como en la observación 2.8, existe un entero r maximal tal que $\{v, f(v), \dots, f^{r-1}(v)\}$ es linealmente independiente, y escalares a_0, \dots, a_{r-1} tales que

$$f^r(v) = a_{r-1}f^{r-1}(v) + \cdots + a_1f(v) + a_0v.$$

El subespacio $V_0 = \langle v, f(v), \dots, f^{r-1}(v) \rangle$ es f -estable y si \mathcal{B} es una base de V que extiende la base $\mathcal{B}_0 = (v, f(v), \dots, f^{r-1}(v))$ de V_0 , entonces la matrix de f para \mathcal{B} es de la forma

$$[f]_{\mathcal{B}}^{\mathcal{B}} = \left[\begin{array}{c|c} A & B \\ \hline O & C \end{array} \right]$$

donde A es una matriz compañera. Ahora, $P_f(t) = P_A(t)P_C(t)$, así, por las observaciones 2.27 (i) y (ii),

$$\begin{aligned}
 P_f(f)(v) &= P_A(f) \circ P_C(f)(v) \\
 &= P_C(f) \circ P_A(f)(v) \\
 &= P_C(f)(f^r - a_{r-1}f^{r-1} - \cdots - a_1f - a_0 \text{id}_V)(v) \\
 &= P_C(f)(f^r(v) - a_{r-1}f^{r-1}(v) - \cdots - a_1f(v) + a_0v) \\
 &= P_C(f)(O) = O.
 \end{aligned}$$

Por ende, para todo $v \in V$, se tiene $P_f(f)(v) = O$, es decir que $P_f(f) = O$. \square

Observación 2.29 (Otra demostración del teorema de Cayley-Hamilton). Antes de mostrar la otra demostración arrancamos con una serie de observaciones.

- (i) Dada una matriz $C = (c_{ij}(t))$ en $M_{m \times n}(\mathbb{K}[t])$ definimos la transformación lineal

$$C_f : \underbrace{V \times \cdots \times V}_{n-\text{veces}} \longrightarrow \underbrace{V \times \cdots \times V}_{m-\text{veces}}$$

por

$$C_f(v_1, \dots, v_n) = \left(\sum_{j=1}^n c_{1j}(f)(v_j), \dots, \sum_{j=1}^n c_{mj}(f)(v_j) \right).$$

- (ii) Note que si nos son dadas $C_1 \in M_{m \times n}(\mathbb{K}[t])$ y $C_2 \in M_{l \times m}(\mathbb{K}[t])$, entonces

$$(C_1 C_2)_f = C_{1f} \circ C_{2f}.$$

- (iii) Dada $B = (b_{ij}(t))$ en $M_{n \times n}(\mathbb{K}[t])$, denotamos por \tilde{B} su matriz de cofactores, es decir la matriz $n \times n$ con entradas en $\mathbb{K}[t]$ cuya ij -ésima entrada es

$$\tilde{b}_{ij}(t) = (-1)^{i+j} \det(B_{ij})$$

donde B_{ij} es la matriz $(n-1) \times (n-1)$ que se obtiene a partir de B eliminando la i -ésima fila y la j -ésima columna. De tal forma que

$$B\tilde{B}^\top = \begin{bmatrix} \det(B) & 0 & \cdots & 0 \\ 0 & \det(B) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \det(B) \end{bmatrix} = (B\tilde{B}^\top)^\top = \tilde{B}B^\top$$

donde \tilde{B}^\top denota la transpuesta de \tilde{B} , es decir la matriz $n \times n$ cuya ij -ésima entrada es la entrada ji -ésima de \tilde{B} .

Para hacer la demostración, consideramos una base $\mathcal{B} = (v_1, \dots, v_n)$ de V y la matriz $A = (a_{ij})$ de f para \mathcal{B} , de forma que

$$f(v_j) = a_{1j}v_1 + \dots + a_{nj}v_n = \sum_{i=1}^n a_{ij}v_i.$$

Al tomar para B la matriz $B = tI_n - A$ en $M_{n \times n}(\mathbb{K}[t])$ obtenemos $\det(\tilde{B}B^\top) = P_f(t)I_n$, pues $\det(B) = P_f(t)$. Ahora, por un lado se tiene

$$\begin{aligned} (B^\top)_f(v_1, \dots, v_n) &= \left(f(v_1) - \left(\sum_{j=1}^n a_{j1}v_j \right), \dots, f(v_n) - \left(\sum_{j=1}^n a_{jn}v_j \right) \right) \\ &= (0, \dots, 0); \end{aligned}$$

y, por otro,

$$\begin{aligned} (P_f(f)(v_1), \dots, P_f(f)(v_n)) &= (P_f(t)I_n)_f(v_1, \dots, v_n) \\ &= (\tilde{B}B^\top)_f(v_1, \dots, v_n) \\ &= \tilde{B}_f \circ (B^\top)_f(v_1, \dots, v_n) \\ &= \tilde{B}_f(0, \dots, 0) \\ &= (0, \dots, 0). \end{aligned}$$

Luego, todos los vectores de la base \mathcal{B} se anulan con $P_f(f)$, y así, por el teorema 1.35 de rigidez de las transformaciones lineales se sigue $P_f(f) = \underline{0}$.

Ejercicios

1. Verifique la fórmula en la observación 2.27 para polinomio característico de una matriz compañera.
2. Verifique el teorema 2.28 de Cayley-Hamilton para el operador de \mathbb{Q}^3 del ejercicio 2.2.1 dado por $f(x, y, z) = (-2x, -2x + 2z, -2x + 2y)$.
3. Verifique el teorema 2.28 de Cayley-Hamilton para el operador de \mathbb{Q}^3 del ejercicio 2.2.2 dado por $f(x, y, z) = (-x + y - z, -2x + 2z, -x + 3y - z)$.

2.5. Descomposición en espacios estables y el lema de los núcleos

Sea V un espacio vectorial sobre \mathbb{K} y f un operador de V .

Proposición 2.30. *Para todo $P(t) \in \mathbb{K}[t]$, el subespacio $V_0 = \ker(P(f))$ es estable bajo f .*

Dem. Dado $v \in V_0$, por la observación 2.16.(ii), se tiene

$$P(f)(f(v)) = P(f) \circ f(v) = f \circ P(f)(v) = f(O) = O.$$

Es decir que $f(v) \in \ker(P(f)) = V_0$. \square

Proposición 2.31. *Suponga que $P(t) \in \mathbb{K}[t]$ es tal que*

$$(i) \ P(f) = 0,$$

$$(ii) \ P(t) = P_1(t)P_2(t) \text{ con } (P_1(t), P_2(t)) = 1 \text{ y } P_1(t), P_2(t) \in \mathbb{K}[t].$$

Entonces, $V = V_1 \oplus V_2$ y $f = f_{V_1} \oplus f_{V_2}$, donde $V_1 = \ker(P_1(f))$ y $V_2 = \ker(P_2(f))$. Más aún, existen polinomios $\Pi_1(t), \Pi_2(t) \in \mathbb{K}[t]$, tales que $p_1 = \Pi_1(f)$ y $p_2 = \Pi_2(f)$ son las proyecciones en V_1 y V_2 relativas a la descomposición $V = V_1 \oplus V_2$ (ver la observación 1.84).

Dem. Como $(P_1(t), P_2(t)) = 1$, entonces existen $Q_1, Q_2 \in \mathbb{K}[t]$ tales que $Q_1(t)P_1(t) + P_2(t)Q_2(t) = 1$, luego

$$Q_1(f) \circ P_1(f) + P_2(f) \circ Q_2(f) = \text{id}_V,$$

y en particular, para todo $v \in V$ se tiene

$$\begin{aligned} v &= \underbrace{Q_1(f) \circ P_1(f)(v)}_{v_2} + \underbrace{P_2(f) \circ Q_2(f)(v)}_{v_1} \\ &= v_2 + v_1. \end{aligned}$$

Ahora

$$P_2(f)(v_2) = P_2(f) \circ Q_1(f) \circ P_1(f)(v) = Q_1(f) \circ P_1(f) \circ P_2(f)(v) = Q_1(f) \circ P(f)(v) = 0$$

y

$$P_1(f)(v_1) = P_1(f) \circ Q_2(f) \circ P_2(f)(v) = Q_2(f) \circ P_1(f) \circ P_2(f)(v) = Q_2(f) \circ P(f)(v) = 0,$$

luego $v_2 \in V_2$ y $v_1 \in V_1$. Así, $V = V_1 + V_2$. Pero, si asumimos que $v \in V_1 \cap V_2$, es decir que $P_1(f)(v) = O = P_2(f)(v)$, entonces en la descomposición $v = v_2 + v_1$, se tiene $v_2 = O = v_1$, luego $v = O$. Por ende, $V_1 \cap V_2 = \{O\}$ y $V = V_1 \oplus V_2$. Por la proposición anterior V_1 y V_2 son f -estables, luego $f = f_{V_1} \oplus f_{V_2}$.

Finalmente, si $\Pi_1(t) = Q_2(t)P_2(t)$ y $\Pi_2(t) = Q_1(t)P_1(t)$, tenemos $\Pi_2(t) + \Pi_1(t) = 1$ y así $\Pi_2(f) + \Pi_1(f) = \text{id}_V$, es decir $p_1 + p_2 = \text{id}_V$, para $p_1 = \Pi_1(f)$ y $p_2 = \Pi_2(f)$. Ahora, como

$$\Pi_1(t)\Pi_2(t) = Q_2(t)P_2(t)Q_1(t)P_1(t) = Q_2(t)Q_1(t)P(t),$$

entonces $\Pi_1(f) \circ \Pi_2(f) = \underline{O}$, es decir $p_1 \circ p_2 = \underline{O}$, y como

$$\Pi_2(t) = \Pi_2(t)(\Pi_2(t) + \Pi_1(t)) = (\Pi_2(t))^2 + \Pi_2(t)\Pi_1(t),$$

entonces $\Pi_2(f) = (\Pi_2(f))^2$, es decir $p_2^2 = p_2$. Similarmente, $p_1^2 = p_1$. Por el teorema 1.85, basta ver que $\text{im}(p_1) = V_1$ y $\text{im}(p_2) = V_2$. Pero, si $v_1 \in V_1 = \ker(P_1(f))$, entonces $v_1 = p_1(v_1) + p_2(v_1) = p_1(v_1) + Q_1(f) \circ P_1(f)(v_1) = p_1(v_1)$, luego $v_1 \in \text{im}(p_1)$; y si $v_1 \in \text{im}(p_1)$, entonces para algún $v \in V$ tenemos $v_1 = p_1(v)$, así $P_1(f)(v_1) = P_1(f) \circ p_1(v) = P_1(f) \circ Q_2(f) \circ P_2(f) = Q_2(f) \circ P(f)(v) = O$, luego $v_1 \in \ker(P_1(f))$. Similarmente se establece $\text{im}(p_2) = V_2$. \square

Ejemplo 2.32. Sea $p \in \text{End}_{\mathbb{K}}(V)$ una proyección, es decir que $p^2 = p$. Luego, si $P(t) = t^2 - t$, entonces $P(p) = p^2 - p = \underline{0}$ y $P(t) = P_1(t)P_2(t)$ con $P_1(t) = t - 1$ y $P_2(t) = t$. Note que $(P_1(t), P_2(t)) = 1$ y

$$-P_1(t) + P_2(t) = 1.$$

Así, por la demostración de la propiedad anterior obtenemos que, para

$$V_1 = \ker(P_1(p)) = \ker(p - \text{id}_V)$$

$$V_2 = \ker(P_2(p)) = \ker(p),$$

$V = V_1 \oplus V_2$ y si

$$\Pi_1(t) = P_2(t) = t$$

$$\Pi_2(t) = -P_1(t) = 1 - t,$$

entonces $p_1 = \Pi_2(p) = p$ y $p_2 = \Pi_1(p) = \text{id}_V - p$ son proyecciones sobre V_1 y V_2 tales que $p_1 + p_2 = \text{id}_V$.

Ejemplo 2.33. Suponga que $\text{char}(\mathbb{K}) \neq 2$, de forma que $-1 \neq 1$ y sea $f \in \text{End}_{\mathbb{K}}(\mathbb{K}^2)$ el operador

$$f(x, y) = (y, x).$$

La matriz canónica de f es

$$[f]_C^C = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

y así su polinomio característico es $P_f(t) = t^2 - 1$. Note que $P_f(t) = P_1(t)P_2(t)$ con $P_1(t) = t - 1$ y $P_2(t) = t + 1$. Por el teorema de Cayley-Hamilton se tiene $P_f(f) = 0$, y por la proposición, $\mathbb{K}^2 = V_1 \oplus V_2$ donde $V_1 = \ker(f - \text{id}_{\mathbb{K}^2})$ y $V_2 = \ker(f + \text{id}_{\mathbb{K}^2})$. Como

$$-\frac{1}{2}P_1(t) + \frac{1}{2}P_2(t) = 1,$$

entonces

$$p_1 = \frac{1}{2}(f + \text{id}_{\mathbb{K}^2}) \quad \text{y} \quad p_2 = -\frac{1}{2}(f - \text{id}_{\mathbb{K}^2})$$

son las proyecciones sobre V_1 y V_2 . Explícitamente:

$$p_1(x, y) = \frac{1}{2}(x + y, x + y) \quad \text{y} \quad p_2(x, y) = \frac{1}{2}(x - y, y - x).$$

Observación 2.34. Bajo las hipótesis de la proposición, para $i \in \{1, 2\}$, tenemos $P_i(f_{V_i}) = \underline{0}$, pues $V_i = \ker(P_i(f))$. Así, podemos a su vez aplicar el resultado a f_{V_i} dada una factorización de $P_i(t)$ en factores primos relativos. Inductivamente, obtenemos el siguiente resultado.

Proposición 2.35 (Lema de los núcleos). *Suponga que $P(t) \in \mathbb{K}[t]$ es tal que*

$$(i) \ P(f) = 0,$$

$$(ii) \ P(t) = P_1(t) \cdots P_r(t) \text{ con } P_1(t), \dots, P_r(t) \in \mathbb{K}[t] \text{ y } (P_i(t), P_j(t)) = 1 \text{ si } i \neq j.$$

Entonces, $V = V_1 \oplus \cdots \oplus V_r$ y $f = f_{V_1} \oplus \cdots \oplus f_{V_r}$, donde $V_i = \ker(P_i(f))$ para $i \in \{1, \dots, r\}$. Más aún, existen polinomios $\Pi_1(t), \dots, \Pi_r(t) \in \mathbb{K}[t]$, tales que $p_i = \Pi_i(f)$ es la proyección sobre V_i relativa a la descomposición $V = V_1 \oplus \cdots \oplus V_r$ (ver la observación 1.84).

Dem. Aunque la demostración se sigue por inducción de la proposición anterior, se puede hacer directamente de forma similar. Para $i = 1, \dots, r$ tomamos $R_i(t) = \prod_{j \neq i} P_j(t)$, y de esta forma se obtiene $(R_1(t), \dots, R_r(t)) = 1$. Así que existen $Q_1(t), \dots, Q_n(t) \in \mathbb{K}[t]$ para los cuales $Q_1(t)R_1(t) + \cdots + Q_n(t)R_n(t) = 1$. Si para $i \in \{1, \dots, r\}$ definimos $\Pi_i(t) = Q_i(t)R_i(t)$ y $p_i = \Pi_i(f)$, entonces tenemos $p_1 + \cdots + p_r = \text{id}_V$, $p_i \circ p_j = \underline{0}$ cuando $i \neq j$, y $p_i^2 = p_i$. El resultado se sigue del teorema 1.85 si demostramos que $\text{im}(p_i) = V_i$. Para ello, dado $v_i \in V_i = \ker(P_i(f))$ se tiene

$$v_i = p_1(v_i) + \cdots + p_r(v_i) = p_i(v_i) + \sum_{j \neq i} Q_j(f) \circ R_j(f)(v_i),$$

pero $P_i(t)$ es factor de los $R_j(t)$, así $R_j(f)(v_i) = 0$, luego $v_i = p_i(v_i)$ y $v_i \in \text{im}(p_i)$. Recíprocamente, dado $v_i \in \text{im}(p_i)$, tenemos $v_i = p_i(v)$ para algún $v \in V$, entonces

$$P_i(f)(v_i) = P_i(f) \circ Q_i(f) \circ R_i(f)(v) = Q_i(f) \circ P(f)(v) = 0$$

y así $v_i \in \ker(P_i(f))$. □

Ejemplo 2.36. Sea $f \in \text{End}_{\mathbb{Q}}(\mathbb{Q}^4)$ el operador

$$f(x, y, z, w) = (x - y + w, -x - z + 2w, 2x - y - z - w, 2x - y)$$

La matriz canónica de f es

$$[f]_C^C = \begin{bmatrix} 1 & -1 & 0 & 1 \\ -1 & 0 & -1 & 2 \\ 2 & -1 & -1 & -1 \\ 2 & -1 & 0 & 0 \end{bmatrix}$$

y así $P_f(t) = P_1(t)P_2(t)P_3(t)$ donde $P_1(t) = (t+1)$, $P_2(t) = (t-1)$ y $P_3(t) = (t^2-2)$. Luego, de la proposición anterior y el teorema de Cayley-Hamilton, si para $i = 1, 2, 3$ definimos $V_i = \ker(P_i(f))$, cada uno de estos espacios es estable bajo f y obtenemos la descomposición:

$$\mathbb{Q}^4 = V_1 \oplus V_2 \oplus V_3.$$

Si usamos la misma notación de la demostración, tenemos $R_1(t) = P_2(t)P_3(t) = (t-1)(t^2-2)$, $R_2(t) = P_1(t)P_3(t) = (t+1)(t^2-2)$, $R_3(t) = (t-1)(t+1)$, y como

$$\frac{1}{2}R_1(t) - \frac{1}{2}R_2(t) + R_3(t) = 1,$$

para

$$\begin{aligned}\Pi_1(t) &= \frac{1}{2}R_1(t) = \frac{(t-1)(t^2-2)}{2}, \\ \Pi_2(t) &= -\frac{1}{2}R_2(t) = -\frac{(t+1)(t^2-2)}{2}, \text{ y} \\ \Pi_3(t) &= R_3(t) = (t-1)(t+1),\end{aligned}$$

los operadores $p_i = \Pi_i(f)$, para $i = 1, 2, 3$, son las respectivas proyecciones sobre V_i de acuerdo a nuestra descomposición de \mathbb{Q}^4 . Las representaciones matriciales en la base canónica de estas proyecciones son:

$$\begin{aligned}\begin{bmatrix} p_1 \end{bmatrix}_c^c &= \begin{bmatrix} 1 & 0 & 0 & -1 \\ 2 & 0 & 0 & -2 \\ 1 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 \end{bmatrix} \\ \begin{bmatrix} p_2 \end{bmatrix}_c^c &= \begin{bmatrix} -3 & 2 & -1 & 2 \\ -3 & 2 & -1 & 2 \\ 0 & 0 & 0 & 0 \\ -3 & 2 & -1 & 2 \end{bmatrix} \\ \begin{bmatrix} p_3 \end{bmatrix}_c^c &= \begin{bmatrix} 3 & -2 & 1 & -1 \\ 1 & -1 & 1 & 0 \\ -1 & 0 & 1 & 1 \\ 3 & -2 & 1 & -1 \end{bmatrix},\end{aligned}$$

así pues $V_1 = \text{im}(p_1) = \langle (1, 2, 1, 0) \rangle$, $V_2 = \text{im}(p_2) = \langle (1, 1, 0, 1) \rangle$ y $V_3 = \text{im}(p_3) = \langle (1, 1, 1, 1), (1, 0, -1, 1) \rangle$. Para $\mathcal{B} = ((1, 2, 1, 0), (1, 1, 0, 1), (1, 1, 1, 1), (1, 0, -1, 1))$, la matriz de f es

$$\begin{bmatrix} f \end{bmatrix}_{\mathcal{B}}^{\mathcal{B}} = \left[\begin{array}{c|c|c|c} -1 & 0 & 0 & 0 \\ \hline 0 & 1 & 0 & 0 \\ \hline 0 & 0 & 0 & 2 \\ \hline 0 & 0 & 1 & 0 \end{array} \right]$$

la cual es una matriz diagonal por bloques, donde cada bloque describe la restricción de f a cada uno de los subespacios estables de la descomposición.

Observación 2.37. En el ejemplo anterior, podemos observar que la dimensión de los subespacios $V_i = \ker(P_i(f))$ corresponde al grado de los factores $P_i(t)$ del polinomio característico $P_f(t)$. Este es el resultado que pasamos a establecer.

Lema 2.38. *Si V tiene dimensión finita, entonces todo factor irreducible del polinomio característico de f es un factor del polinomio minimal de f .*

Dem. Denotemos el polinomio minimal de f por $Q(t)$. Sea $P(t)$ un factor irreducible del polinomio característico $P_f(t)$. Por inducción en la dimensión $n = \dim_{\mathbb{K}}(V)$, estableceremos que $P(t)$ divide a $Q(t)$; siendo el caso $n = 1$ inmediato, pues en él $P(t) = P_f(t) = Q(t)$. Como en la demostración del teorema

2.28 de Cayley-Hamilton, tomamos un $v \in V$ no nulo y así existe un entero r maximal tal que $\{v, f(v), \dots, f^{r-1}(v)\}$ es linealmente independiente. El subespacio $W = \langle v, f(v), \dots, f^{r-1}(v) \rangle$ es f -estable y si \mathcal{B} es una base de V que extiende la base \mathcal{B}_0 de W , entonces la matrix de f para \mathcal{B} es de la forma

$$[f]_{\mathcal{B}}^{\mathcal{B}} = \left[\begin{array}{c|c} A & B \\ \hline O & C \end{array} \right]$$

donde A es una matrix de f_W y C corresponde una matrix de $f_{V/W}$. En particular, se tiene $P_f(t) = P_A(t)P_C(t)$ con $P_A(t) = P_{f_W}(t)$ y $P_C(t) = P_{f_{V/W}}(t)$. Ahora, como $P(t)$ es un polinomio irreducible que divide a $P_f(t)$, entonces, por el lema A.33, $P(t)$ divide a unos de los factores $P_{f_W}(t)$ ó $P_{f_{V/W}}(t)$, que son ambos polinomios característicos. Luego, como W y V/W tienen dimensión estrictamente menor que V , por hipótesis de inducción se sigue que $P(t)$ divide al polinomio minimal de f_W ó de $f_{V/W}$ y como estos, por la proposición 2.21, dividen a $Q(t)$, se sigue que $P(t)$ divide a $Q(t)$. \square

Teorema 2.39. *Si V tiene dimensión finita y*

$$P_f(t) = P_1(t)^{m_1} \cdots P_r(t)^{m_r}$$

es una factorización del polinomio característico donde los $P_i(t)$ son polinomios irreducibles dos a dos distintos, entonces el operador f se descompone como

$$f = f_{V_1} \oplus \cdots \oplus f_{V_r}$$

donde, para $i \in \{1, \dots, r\}$, los espacios f -estables $V_i = \ker(P_i(f)^{m_i})$ tienen dimensión $\dim(V_i) = m_i \deg(P_i(t))$ y los polinomios característicos de las restricciones f_{V_i} son $P_{f_{V_i}}(t) = P_i(t)^{m_i}$. Además, si \mathcal{B}_i es una base de V_i , entonces la concatenación $\mathcal{B} = (\mathcal{B}_1 | \cdots | \mathcal{B}_r)$ es una base de V para la cual la matriz de f es la matriz diagonal por bloques

$$[f]_{\mathcal{B}}^{\mathcal{B}} = \left[\begin{array}{c|c|c} A_1 & \cdots & 0 \\ \hline \vdots & \ddots & \vdots \\ \hline 0 & \cdots & A_r \end{array} \right]$$

donde cada bloque A_i es la matriz de f_{V_i} para \mathcal{B}_i .

Dem. Del teorema 2.28 de Cayley-Hamilton y la proposición 2.35 se sigue la descomposición $f = f_{V_1} \oplus \cdots \oplus f_{V_r}$ donde, para $i \in \{1, \dots, r\}$, los espacios $V_i = \ker(P_i(f)^{m_i})$ son f -estables. Por ello, si \mathcal{B}_i es una base de V_i , entonces la concatenación $\mathcal{B} = (\mathcal{B}_1 | \cdots | \mathcal{B}_r)$ es una base de V para la cual la matriz de f es una matriz diagonal por bloques donde cada bloque A_i es la matriz de f_{V_i} para \mathcal{B}_i . Falta ver que $\dim(V_i) = \deg(P_i(t)^{m_i})$ y el polinomio característico de las restricciones f_{V_i} es $P_{f_{V_i}} = P_i(f)^{m_i}$.

De hecho, si $n_i = \dim(V_i)$, al calcular el polinomio característico $P_f(t)$ con la matriz de f para \mathcal{B} obtenemos

$$P_f(t) = \det(tI_{n_1} - A_1) \cdots \det(tI_{n_r} - A_r) = P_{f_{V_1}}(t) \cdots P_{f_{V_r}}(t).$$

Luego, $P_1(t)^{m_1} \cdots P_r(t)^{m_r} = P_{f_{V_1}}(t) \cdots P_{f_{V_r}}(t)$. Suponga por contradicción que $P_j(t)$ divide a $P_{f_{V_i}}(t)$ con $i \neq j$, entonces, por el lema anterior, $P_j(t)$ divide al polinomio minimal de f_{V_i} . Pero, $P_i(t)^{m_i}$ es un polinomio anulador de f_{V_i} , entonces el polinomio minimal de f_{V_i} lo divide y así $P_j(t)$ divide a $P_i(t)^{m_i}$, lo cual contradice el teorema A.33 de factorización única. Así, $P_{f_{V_i}}(t) = P_i(t)^{m_i}$ y se sigue que $\dim(V_i) = m_i \deg(P_i(t))$ (ver la observación 2.25). \square

Ejercicios

Considere los siguientes operadores:

(i) $f \in \text{Hom}_{\mathbb{Q}}(\mathbb{Q}^3, \mathbb{Q}^3)$ donde

$$f(x, y, z) = (4x, 3x + y - 3z, 3x - 3y + z),$$

(ii) $f \in \text{Hom}_{\mathbb{Q}}(\mathbb{Q}^3, \mathbb{Q}^3)$ donde

$$f(x, y, z) = \frac{1}{2}(5x - y + z, 5x - y - 3z, 4x - 4y)$$

(iii) $f \in \text{Hom}_{\mathbb{Q}}(\mathbb{Q}^4, \mathbb{Q}^4)$ donde

$$f(x, y, z, w) = \frac{1}{3}(11x - y - z + 2w, 14x + 2y - 7z - 4w, 13x - 5y + z - 5w, 13x - 8y - 5z + 4w)$$

(iv) $f \in \text{Hom}_{\mathbb{Q}}(\mathbb{Q}^4, \mathbb{Q}^4)$ donde

$$f(x, y, z, w) = \frac{1}{3}(-2x + y + 4z - 2w, 3x - 6y + 6z, 2x - 4y + 5z + 2w, -6y + 6z + 3w),$$

(v) $f \in \text{Hom}_{\mathbb{R}}(\mathbb{R}^4, \mathbb{R}^4)$ donde

$$f(x, y, z, w) = \frac{1}{3}(-2x + y + 4z - 2w, 3x - 6y + 6z, 2x - 4y + 5z + 2w, -6y + 6z + 3w),$$

(vi) $f \in \text{Hom}_{\mathbb{Q}}(\mathbb{Q}^4, \mathbb{Q}^4)$ donde

$$f(x, y, z, w) = \frac{1}{3}(9x - 3y - 6w, 3x - 3y + 6z, 8x - 10y + 5z - w, 10x - 8y + z + w),$$

(vii) $f \in \text{Hom}_{\mathbb{C}}(\mathbb{C}^4, \mathbb{C}^4)$ donde

$$f(x, y, z, w) = \frac{1}{3}(9x - 3y - 6w, 3x - 3y + 6z, 8x - 10y + 5z - w, 10x - 8y + z + w).$$

Para cada operador:

(a) encuentre la representación matricial de f en la base canónica;

- (b) encuentre la factorización del polinomio característico $P_f(t)$ como producto de potencias de polinomios irreducibles de $K[t]$:

$$P_f(t) = \prod_{i=1}^n P_i(t)^{r_i}$$

donde cada $P_i(t) \in \mathbb{K}[t]$ es irreducible;

- (c) sean $V_i = \ker(P_i(f)^{r_i})$, $i = 1, \dots, n$, y p_1, \dots, p_n las proyecciones sobre V_1, \dots, V_n respecto a la descomposición $V = \bigoplus_{i=1}^n V_i$, encuentre polinomios $\Pi_1(t), \dots, \Pi_n(t) \in K[t]$ tales que $\Pi_i(f) = p_i$, $i = 1, \dots, n$;
- (d) encuentre una base de V respecto a la cual la representación matricial de f es diagonal por bloques de tamaño $\dim(V_i)$, $i = 1, \dots, n$, y cada bloque corresponde a una representación matricial de $f|_{V_i}$.

2.6. Operadores nilpotentes y forma normal de Jordan

Sea V un espacio vectorial sobre \mathbb{K} y f un operador de V .

Suponga que V tiene dimensión finita y que $P_f(t)$ se descompone en factores lineales, es decir

$$P_f(t) = (t - \lambda_1)^{m_1} (t - \lambda_2)^{m_2} \dots (t - \lambda_r)^{m_r}, \quad \lambda_1, \lambda_2, \dots, \lambda_r \in \mathbb{K}.$$

con $\lambda_i \neq \lambda_j$ si $i \neq j$. De esta forma, si $V_i = \ker((f - \lambda_i \text{id}_V)^{m_i})$, para $i = 1, \dots, r$, tenemos

$$V = V_1 \oplus V_2 \oplus \dots \oplus V_r.$$

En particular, si $g_i \in \text{Hom}_{\mathbb{K}}(V_i, V_i)$ es la restricción de $f - \lambda_i \text{id}_V$ a V_i , tenemos que para todo $v_i \in V_i$

$$g_i^{m_i}(v_i) = (f - \lambda_i \text{id}_V)^{m_i}(v_i) = 0$$

es decir que $g_i^{m_i} = \underline{0}$, y se tiene

$$f = (\lambda_1 \text{id}_{V_1} + g_1) \oplus \dots \oplus (\lambda_r \text{id}_{V_r} + g_r).$$

Esto justifica el estudio de operadores que se anulan a partir de una potencia.

Definición 2.40. Decimos que f es *nilpotente* si existe $r \in \mathbb{Z}_{>0}$ tal que $f^r = \underline{0}$, y al mínimo entre estos lo llamamos el *índice de nilpotencia* de f . Similarmente, una matriz $n \times n$ es *nilpotencia* si alguna potencia se anula, y a la mínima potencia que la anula la llamamos el *índice de nilpotencia*.

Proposición 2.41. Suponga que f es nilpotente de índice r , y que $V \neq \{O\}$, entonces tenemos una cadena de contenencias estrictas

$$\{O\} < \ker(f) < \ker(f^2) < \dots < \ker(f^r) = V.$$

En particular, si V tiene dimensión finita, entonces $r \leq \dim(V)$.

Dem. Note primero que para todo $j \in \mathbb{Z}_{>0}$, si $v \in V$ es tal que $f^j(v) = 0$, entonces $f^{j+1}(v) = 0$, luego $\ker(f^j) \leq \ker(f^{j+1})$.

Si $r = 1$, no hay nada que demostrar pues $f = \underline{O}$, y así la cadena corresponde a $\{O\} < V$. Ahora suponga que $r > 1$, luego $f^{r-1} \neq \underline{O}$, y así existe $v \in V$ tal que $f^{r-1}(v) \neq O$. Note que para $j = 1, \dots, r-1$

$$\begin{aligned} f^{j-1}(f^{r-j}(v)) &= f^{r-1}(v) \neq O, \text{ y} \\ f^j(f^{r-j}(v)) &= f^r(v) = O, \end{aligned}$$

luego se tiene $f^{r-j}(v) \in \ker(f^j) \setminus \ker(f^{j-1})$, y obtenemos la contención estricta $\ker(f^{j-1}) < \ker(f^j)$.

Suponga ahora que V tiene dimensión finita y denote $n_j = \dim(\ker(f^j))$. Entonces,

$$0 < n_1 < n_2 < \dots < n_r = \dim(V)$$

es una cadena de $r+1$ enteros estrictamente creciente que arranca en 0, luego $1 \leq n_1, 2 \leq n_2, \dots, r \leq n_r = \dim(V)$. \square

Ejemplo 2.42. Sea $f \in \text{Hom}_{\mathbb{K}}(\mathbb{K}^4, \mathbb{K}^4)$ el operador nilpotente de índice 4 definido por

$$f(x, y, z, w) = (y, z, w, 0).$$

Así,

$$\begin{aligned} f^2(x, y, z, w) &= (z, w, 0, 0), \\ f^3(x, y, z, w) &= (w, 0, 0, 0), \\ f^4(x, y, z, w) &= (0, 0, 0, 0) \end{aligned}$$

y

$$n_1 = 1, \quad n_2 = 2, \quad n_3 = 3, \quad n_4 = 4.$$

La matriz canónica de f es

$$[f]_{\mathcal{C}}^{\mathcal{C}} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

y su polinomio característico es $P_f(t) = t^4$.

Ejemplo 2.43. Sea $f \in \text{Hom}_{\mathbb{K}}(\mathbb{K}^4, \mathbb{K}^4)$ el operador nilpotente de índice 3 definido por

$$f(x, y, z, w) = (y, z, 0, 0).$$

Así,

$$\begin{aligned} f^2(x, y, z, w) &= (z, 0, 0, 0), \\ f^3(x, y, z, w) &= (0, 0, 0, 0) \end{aligned}$$

y

$$n_1 = 2, \quad n_2 = 3, \quad n_3 = 4.$$

La matriz canónica de f es

$$[f]_{\mathcal{C}}^{\mathcal{C}} \left[\begin{array}{ccc|c} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 \end{array} \right]$$

y su polinomio característico es $P_f(t) = t^4$.

Ejemplo 2.44. Sea $f \in \text{Hom}_{\mathbb{K}}(\mathbb{K}^4, \mathbb{K}^4)$ el operador nilpotente de índice 2 definido por

$$f(x, y, z, w) = (y, 0, w, 0).$$

Así,

$$f^2(x, y, z, w) = (0, 0, 0, 0)$$

y

$$n_1 = 2, \quad n_2 = 4$$

La matriz canónica de f es

$$[f]_{\mathcal{C}}^{\mathcal{C}} \left[\begin{array}{cc|cc} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{array} \right]$$

y el polinomio característico es $P_f(t) = t^4$.

Ejemplo 2.45. Sea $f \in \text{Hom}_{\mathbb{K}}(\mathbb{K}^4, \mathbb{K}^4)$ el operador nilpotente de índice 2 definido por

$$f(x, y, z, w) = (y, 0, 0, 0).$$

Así,

$$f^2(x, y, z, w) = (0, 0, 0, 0)$$

y

$$n_1 = 3, \quad n_2 = 4.$$

La matriz canónica de f es

$$[f]_{\mathcal{C}}^{\mathcal{C}} \left[\begin{array}{cc|cc} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 \end{array} \right]$$

y el polinomio característico es $P_f(t) = t^4$.

Definición 2.46. Suponga que f es nilpotente. Dado $v \in V$, al mínimo $k \in \mathbb{Z}_{>0}$ tal que $f^k(v) = O$ lo llamamos el *índice* de v (bajo f).

Observación 2.47. Suponga que f es nilpotente de índice r y denotemos K_j al núcleo $\ker(f^j)$, para $j \in \{0, 1, \dots, r\}$. En particular se tiene $K_0 = \{O\}$ y $K_r = V$. Para $j = 2, \dots, r$, la propiedad 2.41 indica que $K_{j-1} < K_j$, luego podemos tomar subespacios $K'_j \leq K_j$ transversales a K_{j-1} , para así descomponer los núcleos K_j como

$$K_j = K_{j-1} \oplus K'_j.$$

Ahora, note que el índice de todo $v \in K'_j$, salvo el origen, es j y que el índice de $f(v)$ es $j-1$. Luego, podemos tomar los espacios K'_j de forma tal que cumplan

$$f(K'_j) \leq K'_{j-1}.$$

La manera de hacer esto es inductivamente: se toma primero K'_r y una vez definido K'_j , se toma K'_{j-1} que contenga $f(K'_j)$. Tenemos entonces

$$\begin{aligned} V &= K_r \\ &= K_{r-1} \oplus K'_r \\ &\vdots \\ &= K_1 \oplus K'_2 \oplus \dots \oplus K'_r \end{aligned}$$

y

$$K'_r \xrightarrow{f} K'_{r-1} \xrightarrow{f} \dots \xrightarrow{f} K'_2 \xrightarrow{f} K_1 \xrightarrow{f} \{O\}$$

Proposición 2.48. Suponga que f es nilpotente y que $v \in V$ tiene índice k . Entonces, $C = \{v, f(v), \dots, f^{k-1}(v)\}$ es linealmente independiente.

Dem. Suponga que $c_0, c_1, \dots, c_{k-1} \in K$ son tales que

$$c_0 v + c_1 f(v) + \dots + c_{k-1} f^{k-1}(v) = O.$$

Aplicando f^{k-1} a esta igualdad obtenemos $c_0 f^{k-1}(v) = O$. Pero $f^{k-1}(v) \neq O$, luego $c_0 = 0$. Inductivamente, si hemos establecido que $c_0 = c_1 = \dots = c_{i-1} = 0$, al aplicar f^{k-i-1} obtenemos $c_i f^{k-1}(v) = O$ y así $c_i = 0$. Por ende, obtenemos $c_0 = c_1 = \dots = c_{k-1} = 0$. \square

Observación 2.49. Suponga que V tiene dimensión finita y f es nilpotente de índice $r = \dim(V)$. Si $v \in V$ es tal que $v \notin \ker(f^{r-1})$, entonces v tiene índice r . Luego, si denotamos $v_i = f^{r-i}(v)$ para $i = 1, \dots, r$, por la propiedad anterior se obtiene que V es cíclico bajo f ,

$$\mathcal{B} = (v_1, \dots, v_r) = (f^{r-1}(v), \dots, f(v), v)$$

es una base de V , y

$$[f]_{\mathcal{B}}^{\mathcal{B}} = \begin{bmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \\ 0 & 0 & 0 & \dots & 0 \end{bmatrix}.$$

Una matriz $n \times n$ como la anterior, con ceros en todas las entradas salvo en las entradas inmediatamente encima de la diagonal en donde hay unos, se denotará por $J_n(0)$.

Definición 2.50. Suponga que V tiene dimensión finita y f es nilpotente de índice $r = \dim(V)$. Una *base de Jordan de V (relativa a f)* es una base de la forma

$$\mathcal{B}_J = (f^{r-1}(v), \dots, f(v), v),$$

o equivalentemente, una base \mathcal{B}_J para la cual la matriz de f es $J_n(0)$.

Observación 2.51. En caso de que f sea nilpotente de índice estrictamente inferior a $\dim(V)$, el espacio V no es cíclico, pero, como veremos en el siguiente teorema, se puede descomponer en subespacios cíclicos. Para su demostración establecemos un lema.

Lema 2.52. *Suponga que f es nilpotente de grado $r > 1$. Fijamos un índice $j \in \{2, \dots, r\}$, y consideramos los núcleos $K_j = \ker(f^j)$ y $K_{j-1} = \ker(f^{j-1})$. Si $K'_j \leq K_j$ es tal que $K_j = K_{j-1} \oplus K'_j$, entonces, para toda colección $v_1, \dots, v_{n'_j} \in K'_j$ linealmente independiente, se tiene que $f(v_1), \dots, f(v_{n'_j})$ es linealmente independiente.*

Dem. En la observación 2.47 vimos que por construcción todos los vectores en K'_j , a excepción de O , tienen índice bajo f igual a j . Suponga ahora que c_1, \dots, c_m son escalares tales que $c_1 f(v_1) + \dots + c_{n'_j} f(v_{n'_j}) = O$, entonces $f(c_1 v_1 + \dots + c_{n'_j} v_{n'_j}) = O$, luego el vector $c_1 v_1 + \dots + c_{n'_j} v_{n'_j}$ tiene índice 1 y así, $c_1 v_1 + \dots + c_{n'_j} v_{n'_j} = O$. Por ende, se sigue que $c_1 = \dots = c_{n'_j} = 0$. \square

Teorema 2.53. *Suponga que V tiene dimensión finita y f es nilpotente de grado $r > 1$. Si $m = \dim(\ker(f))$, entonces V se descompone en m subespacios f -cíclicos.*

Dem. Para $i \in \{0, 1, \dots, r\}$, denotamos $K_j = \ker(f^j)$. Como en la observación 2.47, para $j \in \{2, \dots, n\}$, tomamos subespacios $K'_j \leq K_j$ tales que $f(K'_j) \leq K'_{j-1}$ y así,

$$V = K_1 \oplus K'_2 \oplus \dots \oplus K'_r$$

y

$$K'_r \xrightarrow{f} K'_{r-1} \xrightarrow{f} \dots \xrightarrow{f} K'_2 \xrightarrow{f} K_1 \xrightarrow{f} \{O\}.$$

Para identificar los m subespacios cíclicos, construiremos una base de V arrancando por una base \mathcal{B}_r de K'_r que una vez extendida a una base \mathcal{B}_j de $K'_j \oplus \dots \oplus K'_{r-1} \oplus K'_r$ la extenderemos a una de $K'_{j-1} \oplus K'_j \oplus \dots \oplus K'_r$ que contenga $f(\mathcal{B}_j)$; lo cual se puede hacer gracias al lema 2.52 anterior.

Para $j = 2, \dots, r$, denotaremos $n_j = \dim(K_j)$ y $n'_j = \dim(K'_j)$, y $n_1 = n'_1 = m = \dim(K_1)$ de forma que

$$n_j = n'_j + n_{j-1}$$

y

$$\begin{aligned}
 \dim(V) &= n_r \\
 &= n'_r + n_{r-1} \\
 &\vdots \\
 &= n'_r + n'_{r-1} + \cdots + n_2 \\
 &= n'_r + n'_{r-1} + \cdots + n'_2 + n_1 \\
 &= n'_r + n'_{r-1} + \cdots + n'_2 + n'_1 \\
 &= n'_r + n'_{r-1} + \cdots + n'_2 + m
 \end{aligned}$$

Sea $\mathcal{B}_r = (v_{r,1}, \dots, v_{r,n'_r})$ una base de K'_r . Para $i \in \{1, \dots, n'_r\}$, denotamos $v_{r-1,i} = f(v_{r,i})$. Entonces, por el lema 2.52, $f(\mathcal{B}_r) = (v_{r-1,1}, \dots, v_{r-1,n'_r})$ forman un conjunto de vectores en K'_{r-1} linealmente independiente.

Sea $\mathcal{B}_{r-1} = (v_{r-1,1}, \dots, v_{r-1,n'_{r-1}})$ un base de K'_{r-1} que extiende a $f(\mathcal{B}_r)$. Para $i \in \{1, \dots, n'_{r-1}\}$, denotamos $v_{r-2,i} = f(v_{r-1,i})$. Entonces, por el lema 2.52, $f(\mathcal{B}_{r-1}) = (v_{r-2,1}, \dots, v_{r-2,n'_{r-1}})$ forman un conjunto de vectores en K'_{r-2} linealmente independiente.

Iterativamente, obtenemos bases $\mathcal{B}_r, \mathcal{B}_{r-1}, \dots, \mathcal{B}_2$ y \mathcal{B}_1 respectivamente de $K'_r, K'_{r-1}, \dots, K'_2$ y K_1 tales que $f(\mathcal{B}_{j+1}) \subseteq \mathcal{B}_j$ para $i \in \{1, \dots, r-1\}$. Además, la concatenación

$$\mathcal{B} = (\mathcal{B}_1 | \mathcal{B}_2 | \dots | \mathcal{B}_r)$$

es una base de V .

Note que en la notación de los vectores que componen a \mathcal{B} , el subíndice j de $v_{j,i}$ determina su índice bajo f y el espacio K'_j al que pertenece, y el subíndice i indica que es el i -ésimo vector entre los n'_j de la base \mathcal{B}_j de K'_j . Definimos (ver la figura 2.1)

$$\begin{aligned}
 V_1 &= \langle v_{j,1} \in \mathcal{B} \mid 1 \leq n'_j \rangle \\
 V_2 &= \langle v_{j,2} \in \mathcal{B} \mid 2 \leq n'_j \rangle \\
 &\vdots \\
 V_m &= \langle v_{j,m} \in \mathcal{B} \mid m \leq n'_j \rangle
 \end{aligned}$$

de esta forma, por construcción, obtenemos la descomposición

$$V = V_1 \oplus V_2 \oplus \cdots \oplus V_m$$

en subespacios f -cíclicos. □

Observación 2.54.

Si nos apegamos a la notación de la demostración, el lema 2.52 implica que $\dim(K'_j) \leq \dim(K'_{j-1})$, o equivalentemente $n'_j \leq n'_{j-1}$, para $j \in \{2, \dots, r\}$. En particular, el j -ésimo piso en la figura 2.1 tiene menos vectores que el piso $j-1$.

Aunque el diagrama en la figura 2.1 es algo complejo porque busca representar la situación genérica, en la práctica es bastante más concreto y sencillo. De hecho, lo usual es simplificarlo representando cada vector $v_{j,i}$ por un punto en la coordenada (i, j) en el plano cartesiano. Note que, fijado un j , hay $n'_j = n_j - n_{j-1}$ puntos en el nivel j .

Ejemplo 2.55. Para el operador $f(x, y, z, w) = (y, z, w, 0)$ del ejemplo 2.42 se tiene

$$m = n'_1 = 1, \quad n'_2 = 1, \quad n'_3 = 1, \quad n'_4 = 1$$

y así, el diagrama es



Ejemplo 2.56. Para el operador $f(x, y, z, w) = (y, z, 0, 0)$ del ejemplo 2.43 se tiene

$$m = n'_1 = 2, \quad n'_2 = 1, \quad n'_3 = 1$$

y así, el diagrama es

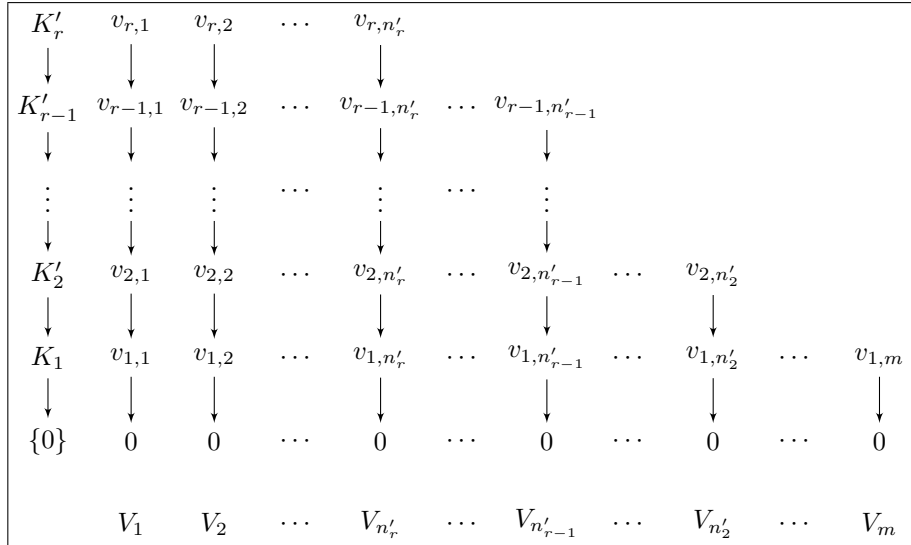


Figura 2.1: Subespacios cíclicos colapsando



Ejemplo 2.57. Para el operador $f(x, y, z, w) = (y, 0, w, 0)$ del ejemplo 2.44 se tiene

$$m = n'_1 = 2, \quad n'_2 = 2$$

y así, el diagrama es



Ejemplo 2.58. Para el operador $f(x, y, z, w) = (y, 0, 0, 0)$ del ejemplo 2.45 se tiene

$$m = n'_1 = 3, \quad n'_2 = 1$$

y así, el diagrama es



Observación 2.59. Bajo la hipótesis y notación del teorema 2.53 obtenemos que para cada V_i , $i = 1, \dots, m$, tenemos una base de Jordan $\mathcal{B}_{J,i}$ de V_i relativa a f_{V_i} . De esta forma la unión de ella forma una base \mathcal{B} de V . La representación matricial de f en la base \mathcal{B} es una matriz diagonal por bloques:

$$[f]_{\mathcal{B}}^{\mathcal{B}} = \left[\begin{array}{c|c|c} J_{r_1}(0) & \cdots & 0 \\ \hline \vdots & \ddots & \vdots \\ \hline 0 & \cdots & J_{r_m}(0) \end{array} \right]$$

donde cada $J_{r_i}(0)$, con $r_i = \dim(V_i)$, es una matriz como en la observación 2.49.

Observación 2.60. Como corolario de la prueba del teorema y por la observación 2.59 anterior tenemos que cuando V tiene dimensión finita y f es nilpotente de índice r , la información suministrada por las cantidades $m = n'_1, n'_2, \dots, n'_r$ determinan unívocamente la transformación f , salvo cambio de coordenadas. De hecho dadas dos transformaciones con igual información, para cada una podemos encontrar una base de V que arrojan la misma representación matricial. Específicamente, $m = n'_1$ indica el número de bloques de Jordan y n'_j el número de bloques de Jordan de tamaño mayor o igual a j .

Definición 2.61. Se le llama *matriz en bloque de Jordan* a una matriz cuadrada $n \times n$ de la forma

$$J_n(\lambda) = \begin{bmatrix} \lambda & 1 & 0 & \cdots & 0 \\ 0 & \lambda & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ 0 & 0 & 0 & \cdots & \lambda \end{bmatrix}.$$

En particular, se tiene $J_n(\lambda) = J_n(0) + \lambda I_n$.

Lema 2.62. Suponga que V tiene dimensión finita igual a n y que $P(t) = (t - \lambda)^r \in \mathbb{K}[t]$ es un polinomio anulador de f . Entonces, existe una base \mathcal{B} de V para la cual la matriz de f es una matriz diagonal por bloques de Jordan:

$$[f]_{\mathcal{B}} = \left[\begin{array}{c|c|c} J_{r_1}(\lambda) & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & J_{r_m}(\lambda) \end{array} \right]$$

con $r_1 + r_2 + \dots + r_m = n$.

Dem. Como $P(f) = (f - \lambda \text{id}_V)^r = 0$, entonces el operador $g = f - \lambda \text{id}_V$ es nilpotente. Por el teorema 2.53 V se descompone en m espacios g -cíclicos, y por la observación 2.59, existe una base \mathcal{B} para la cual la matriz de g es diagonal por bloques de Jordan $J_{r_1}(0), \dots, J_{r_m}(0)$. Luego,

$$\begin{aligned} [f]_{\mathcal{B}} &= [g + \lambda \text{id}_V]_{\mathcal{B}} = [g]_{\mathcal{B}} + [\lambda \text{id}_V]_{\mathcal{B}} \\ &= \left[\begin{array}{c|c|c} J_{r_1}(0) & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & J_{r_m}(0) \end{array} \right] + \lambda I_m \\ &= \left[\begin{array}{c|c|c} J_{r_1}(\lambda) & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & J_{r_m}(\lambda) \end{array} \right] \end{aligned}$$

□

Teorema 2.63 (Teorema de Jordan). Suponga que V tiene dimensión finita y que

$$P_f(t) = (t - \lambda_1)^{m_1} (t - \lambda_2)^{m_2} \dots (t - \lambda_r)^{m_r}, \quad \lambda_1, \lambda_2, \dots, \lambda_r \in \mathbb{K}.$$

Entonces, existe una base \mathcal{B} de V tal que la matriz de f es una matriz diagonal por bloques de Jordan.

Dem. Sin pérdida de generalidad podemos asumir que $\lambda_i \neq \lambda_j$ si $i \neq j$. Así,

$$((t - \lambda_i)^{m_i}, (t - \lambda_j)^{m_j}) = 1$$

si $i \neq j$. Por el teorema de Cayley-Hamilton $P_f(f) = 0$, luego por el lema de los núcleos (proposición 2.35), V se descompone en

$$V = V_1 \oplus \dots \oplus V_r$$

donde cada $V_i = \ker((f - \lambda_i \text{id}_V)^{m_i})$, para $i \in \{1, \dots, r\}$, es f -estable. En particular, el polinomio $P_i = (t - \lambda_i)^{m_i}$ es un anulador de $f|_{V_i}$. Por lo tanto, el lema anterior implica que existe una base \mathcal{B}_i de V_i para la cual $[f_i]_{\mathcal{B}_i}^{\mathcal{B}_i}$ es una matriz diagonal por bloques de Jordan. Finalmente, para la concatenación de bases $\mathcal{B} = (\mathcal{B}_1 | \dots | \mathcal{B}_r)$, la matriz de f tiene la forma buscada. \square

Definición 2.64. Generalizamos la definición 2.50 de base de Jordan. Si V tiene dimensión finita, decimos que una base de V es una *base de Jordan (relativa a f)* si la matriz para aquella base es diagonal en bloques de Jordan, y a esta matriz la llamamos *forma normal de Jordan de f* .

Ejercicios

Considere los siguientes operadores:

(i) $f \in \text{Hom}_{\mathbb{Q}}(\mathbb{Q}^4, \mathbb{Q}^4)$ donde

$$f(x, y, z, w) = (3x - 3y + 3w, 2x - y - z + 2w, 2x - y - z + 2w, 2x - 4y + 2z + 2w),$$

(ii) $f \in \text{Hom}_{\mathbb{Q}}(\mathbb{Q}^4, \mathbb{Q}^4)$ donde

$$f(x, y, z, w) = (-x + 2y - z + 2w, x + y - 2z + w, -x + 2y - z + 2w, -2x + y + z + w),$$

(iii) $f \in \text{Hom}_{\mathbb{Q}}(\mathbb{Q}^5, \mathbb{Q}^5)$ donde

$$f(x_1, x_2, x_3, x_4, x_5) = (4x_1 - 4x_2 - 4x_3 + 8x_5, -4x_3 + 8x_5, -x_1 - x_2 - x_3 - x_4 + 7x_5, \\ -x_1 - 5x_2 - x_3 + 3x_4 + 7x_5, 2x_1 - 2x_2 - 2x_3 + 2x_4 + 6x_5),$$

(iv) $f \in \text{Hom}_{\mathbb{Q}}(\mathbb{Q}^5, \mathbb{Q}^5)$ donde

$$f(x_1, x_2, x_3, x_4, x_5) = (x_1 - 3x_2 - 3x_3 + x_4 + 9x_5, -4x_3 + 8x_5, -4x_1 + 8x_5, \\ -4x_1 - 4x_2 + 4x_4 + 8x_5, -x_1 - x_2 - x_3 + 3x_4 + 7x_5).$$

Para cada operador:

- (a) encuentre la representación matricial de f en la base canónica,
- (b) encuentre una base de Jordan para f ,
- (c) encuentre la representación matricial de f respecto a la base de Jordan encontrada.

2.7. Descomposición diagonalizable-nilpotente

Sea V un espacio vectorial sobre \mathbb{K} de dimensión finita y f un operador de V .

Suponga que el polinomio característico de f se factoriza en factores lineales

$$P_f(t) = (t - \lambda_1)^{m_1} (t - \lambda_2)^{m_2} \dots (t - \lambda_r)^{m_r}, \quad \lambda_1, \lambda_2, \dots, \lambda_r \in \mathbb{K}.$$

Por el teorema 2.63 de Jordan, existe una base \mathcal{B} de Jordan de V relativa a f . En particular, si la forma normal de Jordan de f tiene s bloques, cada uno de tamaño d_i , con los primeros bloques con λ_1 en la diagonal, los siguientes λ_2 , y así hasta los últimos con λ_r , entonces

$$\begin{aligned} A = [f]_{\mathcal{B}}^{\mathcal{B}} &= \left[\begin{array}{c|c|c} J_{d_1}(\lambda_1) & \cdots & 0 \\ \hline \vdots & \ddots & \vdots \\ \hline 0 & \cdots & J_{d_s}(\lambda_r) \end{array} \right] \\ &= \underbrace{\left[\begin{array}{c|c|c} \lambda_1 I_{d_1}(0) & \cdots & 0 \\ \hline \vdots & \ddots & \vdots \\ \hline 0 & \cdots & \lambda_r I_{d_s}(0) \end{array} \right]}_D + \underbrace{\left[\begin{array}{c|c|c} J_{d_1}(0) & \cdots & 0 \\ \hline \vdots & \ddots & \vdots \\ \hline 0 & \cdots & J_{d_s}(0) \end{array} \right]}_N, \end{aligned}$$

así, $A = D + N$ donde D es una matriz diagonal y N una nilpotente.

Teorema 2.65 (Descomposición de Jordan-Chevalley). *Suponga que V tiene dimensión finita y que*

$$P_f(t) = (t - \lambda_1)^{m_1} (t - \lambda_2)^{m_2} \dots (t - \lambda_r)^{m_r}, \quad \lambda_1, \lambda_2, \dots, \lambda_r \in \mathbb{K}.$$

Entonces, existen operadores $f_D, f_N \in \text{End}_{\mathbb{K}}(V)$, tales que

1. f_D es diagonalizable y f_N es nilpotente;
2. $f_D + f_N = f$;
3. $f_D \circ f_N = f_N \circ f_D$.

Más aún, esta descomposición es única respecto a estas tres propiedades. Además existen polinomios $P_D(t), P_N(t) \in \mathbb{K}[t]$ tales que $f_N = P_N(f)$ y $f_D = P_D(f)$.

Dem. (Existencia) Defina $P_i(t) = (t - \lambda_i)^{m_i}$, $i = 1, \dots, n$. Por Propiedad 2.35, existen polinomios $\Pi_1(t), \dots, \Pi_n(t) \in \mathbb{K}[t]$ tales que los operadores $\Pi_i(f) = p_i$, para $i \in \{1, \dots, n\}$, son las proyecciones sobre los subespacios $V_i = \ker((f - \lambda_i \text{id}_V)^{m_i})$ respecto a la descomposición

$$V = V_1 \oplus \dots \oplus V_r.$$

Defina $P_D(t) = \lambda_1 \Pi_1(t) + \dots + \lambda_n \Pi_n(t)$, y $f_D = P_D(f)$. De esta forma, si $v_i \in V_i$,

$$f_D(v_i) = \lambda_1 p_1(v_i) + \dots + \lambda_n p_n(v_i) = \lambda_i v_i,$$

y así f_D es diagonalizable por el teorema 2.14. Defina $P_N(t) = t - P_D(t)$ y $f_N = P_N(f) = f - f_D$, luego se tiene $f_D + f_N = f$. Dado $v_i \in V_i$

$$f_N(v_i) = f(v_i) - f_D(v_i) = f(v_i) - \lambda_i(v_i) = (f - \lambda_i \text{id}_V)(v_i),$$

luego la restricción de f_N a V_i es nilpotente con índice $\leq m_i$. Entonces, f_N es nilpotente de grado $\leq \max\{m_1, \dots, m_n\}$. Finalmente, como dos operadores definidos por polinomios de un mismo operador conmutan se tiene

$$f_D \circ f_N = P_D(f) \circ P_N(f) = P_N(f) \circ P_D(f) = f_N \circ f_D.$$

Para la demostración de la unicidad de esta descomposición necesitamos los siguientes lemas. \square

Lema 2.66. *Si $\lambda_1, \dots, \lambda_r \in \mathbb{K}$ son valores propios de f dos a dos distintos, y $v_1, \dots, v_r \in V$ son vectores propios de f , con $f(v_i) = \lambda_i v_i$ para $i = 1, \dots, r$, entonces $\{v_1, \dots, v_r\}$ es un conjunto linealmente independiente.*

Dem. Por inducción en r , siendo el caso base $r = 1$ inmediato, pues $\{v_1\}$ es linealmente independiente si $v_1 \neq 0$, lo cual se cumple pues v_1 es vector propio. Para el paso inductivo, tomamos c_1, \dots, c_r son tales que $c_1 v_1 + \dots + c_r v_r = 0$, y por contradicción asumimos que cada $c_i \neq 0$ (pues de lo contrario, por hipótesis de inducción, si algún c_i es 0 el resto también lo son). Entonces,

$$0 = (f - \lambda_n \text{id}_V)(c_1 v_1 + \dots + c_r v_r) = c_1(\lambda_1 - \lambda_r)v_1 + \dots + c_{r-1}(\lambda_{r-1} - \lambda_r)v_{r-1};$$

y así, por la hipótesis de inducción, se sigue que $c_i(\lambda_i - \lambda_n) = 0$ para $i = 1, \dots, r-1$. Pero como asumimos que $c_i \neq 0$ y se tiene que $\lambda_i - \lambda_r \neq 0$ para $i \in \{1, \dots, r-1\}$, entonces $c_i(\lambda_i - \lambda_n) \neq 0$, lo cual es una contradicción. \square

Lema 2.67. *Si f es diagonalizable, se siguen las siguientes propiedades.*

1. *Si $W \neq \{0\}$ es f -estable, entonces la restricción f_W también es diagonalizable.*
2. *Si $g \in \text{End}_{\mathbb{K}}(V)$ es diagonalizable y $f \circ g = g \circ f$, entonces existe una descomposición $V = V_1 \oplus \dots \oplus V_s$ para la cual $f_{V_i} = \lambda_i \text{id}_{V_i}$ y $g_{V_i} = \mu_i \text{id}_{V_i}$ con $\lambda_i, \mu_i \in \mathbb{K}$ para cada $i \in \{1, \dots, s\}$. En particular, existe una base de V para la cual las matrices f y g son simultáneamente diagonales. Más aún, si v es un vector propio simultáneamente de f y g , entonces v es un vector propio de $af + bg$ para todo $a, b \in \mathbb{K}$.*

Dem. Por el teorema 2.14, $f = \lambda_1 \text{id}_{V_1} \oplus \dots \oplus \lambda_r \text{id}_{V_r}$ donde $\lambda_1, \dots, \lambda_r \in \mathbb{K}$ son los valores propios de f y $V_1, \dots, V_r \leq V$ son los respectivos espacios propios.

1. Para $i \in \{1, \dots, r\}$, sea $W_i = W \cap V_i$. De esta forma, $f_W = \lambda_1 \text{id}_{W_1} \oplus \dots \oplus \lambda_r \text{id}_{W_r}$, y así, por el teorema 2.14, f_W es diagonalizable.
2. Para $i \in \{1, \dots, r\}$, si $v \in V_i$, entonces

$$f(g(v)) = g(f(v)) = \lambda_i g(v),$$

y así V_i es invariante bajo g . Por la primera parte del lema, sigue que la restricción g_{V_i} es diagonalizable. Luego, si tomamos para cada V_i una base \mathcal{B}_i para la cual la matriz de g_{V_i} es diagonal, por el lema anterior, para la concatenación $\mathcal{B} = (\mathcal{B}_1 | \cdots | \mathcal{B}_r)$ se obtiene una base para la cual las matrices de f y g son simultáneamente diagonales. Finalmente, si para $v \in V$ se tiene $f(v) = \lambda v$ y $g(v) = \mu v$, entonces $(af + bg)(v) = (a\lambda + b\mu)v$. \square

Lema 2.68. *Suponga que f es nilpotente y que g es otro operador de V nilpotente que conmuta con f . Entonces, $af + bg$ es nilpotente para todo $a, b \in \mathbb{K}$.*

Dem. Sean r y s respectivamente los índices de nilpotencia de f y g . Como f y g conmutan entonces, para todo entero positivo m , se tiene

$$(af + bg)^n = \sum_{i=1}^m \binom{m}{i} a^i b^{m-i} f^i \circ g^{m-i}.$$

En particular, para $m = r + s$ se tiene $(af + bg)^m = \underline{0}$.

Dem. (Unicidad en el teorema 2.65) Si $f'_D, f'_N \in \text{End}_{\mathbb{K}}(V)$ conmutan y son respectivamente diagonalizable y nilpotente tales que $f = f'_D + f'_N$, entonces

$$f \circ f'_D = (f'_D + f'_N) \circ f'_D = f'_D \circ f'_D + f'_N \circ f'_D = f'_D \circ f'_D + f'_D \circ f'_N = f'_D \circ f,$$

es decir que f y f'_D conmutan. Por lo cual, $P_D(f) = f_D$ y f'_D también lo hacen. Similarmente, f_N y f'_N conmutan.

Como f_D y f'_D son diagonalizables y conmutan, por los lemas anteriores, V admite una base de vectores simultáneamente propios, que son todos vectores propios de $f_D - f'_D$. Pero $f_D - f'_D = f'_N - f_N$, y, como f'_N y f_N igualmente conmutan, entonces por los lemas $f'_N - f_N$ es también nilpotente. Así, $f_D - f'_D$ es diagonalizable y, a su vez, nilpotente, luego el único valor propio del operador diagonalizable $f_D - f'_D$ es 0, y por el teorema 2.14 $f_D - f'_D = \underline{0}$, $\underline{0} = f_D - f'_D = f'_N - f_N$ y entonces, $f'_D = f_D$ y $f'_N = f_N$. \square

Ejercicios

1. Considere los siguientes operadores de los ejercicios de la sección 2.6:

- (i) $f \in \text{Hom}_{\mathbb{Q}}(\mathbb{Q}^4, \mathbb{Q}^4)$ donde

$$f(x, y, z, w) = (3x - 3y + 3w, 2x - y - z + 2w, 2x - y - z + 2w, 2x - 4y + 2z + 2w),$$

- (ii) $f \in \text{Hom}_{\mathbb{Q}}(\mathbb{Q}^4, \mathbb{Q}^4)$ donde

$$f(x, y, z, w) = (-x + 2y - z + 2w, x + y - 2z + w, -x + 2y - z + 2w, -2x + y + z + w),$$

- (iii) $f \in \text{Hom}_{\mathbb{Q}}(\mathbb{Q}^5, \mathbb{Q}^5)$ donde

$$f(x_1, x_2, x_3, x_4, x_5) = (4x_1 - 4x_2 - 4x_3 + 8x_5, -4x_3 + 8x_5, -x_1 - x_2 - x_3 - x_4 + 7x_5, \\ -x_1 - 5x_2 - x_3 + 3x_4 + 7x_5, 2x_1 - 2x_2 - 2x_3 + 2x_4 + 6x_5),$$

(iv) $f \in \text{Hom}_{\mathbb{Q}}(\mathbb{Q}^5, \mathbb{Q}^5)$ donde

$$f(x_1, x_2, x_3, x_4, x_5) = (x_1 - 3x_2 - 3x_3 + x_4 + 9x_5, -4x_3 + 8x_5, -4x_1 + 8x_5, \\ -4x_1 - 4x_2 + 4x_4 + 8x_5, -x_1 - x_2 - x_3 + 3x_4 + 7x_5).$$

Encuentre para cada operador:

- (a) polinomios $P_D(t)$ y $P_N(t)$, tales que $f = f_D + f_N$ donde $f_D = P_D(f)$ es diagonalizable y $f_N = P_N(f)$ es nilpotente,
 - (b) las matrices canónica de f_D y f_N .
2. Decimos que un operador f es *simple* si los únicos espacios invariantes bajo f son $\{0\}$ y V . Decimos que f es *semi-simple* si para todo $V_1 \leq V$ invariante bajo f existe $V_2 \leq V$ invariante bajo f tal que $V = V_1 \oplus V_2$.
- (a) Suponga que f es semi-simple y sea $V_1 \leq V$ un subespacio invariante bajo f . Demuestre que la restricción de f a V_1 , $f_1 \in \text{Hom}_K(V_1, V_1)$, es semi-simple.
 - (b) Suponga que f es semi-simple, demuestre que existe una descomposición

$$V = V_1 \oplus \dots \oplus V_r$$

tal que, para $i = 1, \dots, r$, V_i es invariante bajo f y la restricción de este a V_i , $f_i \in \text{Hom}_K(V_i, V_i)$, es simple. (*Ayuda:* Use inducción fuerte en $n = \dim_K(V)$, es decir asuma que el resultado es cierto para todo operador en un espacio de dimensión estrictamente menor que n y use el punto anterior).

- (c) Demuestre que el polinomio minimal de un operador semi-simple es un producto de polinomios irreducibles ninguno de ellos repetido (es decir, elevados únicamente a la primera potencia; es decir, sin cuadrados que lo dividan) usando los siguientes pasos:
 - (i) Usando la descomposición del punto anterior, demuestre que el polinomio minimal de f_i divide al polinomio minimal de f . (*Ayuda:* divida el polinomio minimal de f por el de f_i y evalúe el residuo en f_i ; demuestre que es cero.)
 - (ii) Demuestre que el mínimo común múltiplo de los polinomios minimales de f_1, \dots, f_r es el polinomio minimal de f . (*Ayuda:* Sea $P(t)$ este mínimo común múltiplo, dado $v \in V$, tome $v_i \in V_i$, $i = 1, \dots, r$, tales que $v = v_1 + \dots + v_r$, demuestre que $P(f)(v) = P(f)(v_1) + \dots + P(f)(v_r) = 0 + \dots + 0 = 0$.)
 - (iii) Concluya. (*Ayuda:* El polinomio minimal de un operador simple es irreducible.)
3. Sea $\lambda \in \mathbb{K}$ y $f \in \text{Hom}_K(\mathbb{K}^2, \mathbb{K}^2)$ definido por $f(x, y) = (\lambda x + y, \lambda y)$. Verifique que f no es semi-simple. Encuentre el polinomio minimal de f .

4. Sea $\lambda \in K$ y $f \in \text{Hom}_K(\mathbb{K}^2, \mathbb{K}^2)$ definido por $f(x, y) = (\lambda x, \lambda y)$. Encuentre el polinomio minimal de f . Verifique que f es semi-simple.
5. Considere los operadores $g, h \in \text{Hom}_{\mathbb{R}}(\mathbb{R}^4, \mathbb{R}^4)$ definidos por

$$\begin{aligned} g(x, y, z, w) &= \left(\frac{\sqrt{3}}{2}x + \frac{1}{2}w, \frac{\sqrt{3}}{2}y + \frac{1}{2}z, -\frac{1}{2}y + \frac{\sqrt{3}}{2}z, -\frac{1}{2}x + \frac{\sqrt{3}}{2}w \right) \\ h(x, y, z, w) &= \left(\frac{\sqrt{3}}{2}x + \frac{1}{2}z, \frac{\sqrt{3}}{2}y + z - \frac{1}{2}w, \frac{1}{2}x - y + \frac{\sqrt{3}}{2}z, -\frac{1}{2}y + \frac{\sqrt{3}}{2}w \right), \end{aligned}$$

y los operadores $g_{\mathbb{C}}, h_{\mathbb{C}} \in \text{Hom}_{\mathbb{C}}(\mathbb{C}^4, \mathbb{C}^4)$ dadas por las mismas fórmulas.

- (a) Verifique que g y h no tienen valores propios y que los valores propios de $g_{\mathbb{C}}$ y $h_{\mathbb{C}}$ son $\lambda = \frac{1}{2}(\sqrt{3} + i)$ y $\bar{\lambda} = \frac{1}{2}(\sqrt{3} - i)$.
- (b) Verifique que si $u + iv$, con $u, v \in \mathbb{R}^4$, es un vector propio de $g_{\mathbb{C}}$ (resp. de $h_{\mathbb{C}}$) asociado a λ , entonces $u - iv$ es un vector propio de $g_{\mathbb{C}}$ (resp. de $h_{\mathbb{C}}$) asociado a $\bar{\lambda}$.
- (c) Encuentre bases de Jordan para $g_{\mathbb{C}}$ y $h_{\mathbb{C}}$ de la forma $\mathcal{B}_{\mathbb{C}} = (u_1 + iv_1, u_2 + iv_2, u_1 - iv_1, u_2 - iv_2)$.
- (d) Usando la notación de los vectores en el literal anterior, encuentre las matrices de g y h para las bases $\mathcal{B} = (u_1, -v_1, u_2, -v_2)$. (*Ayuda:* Por bloques 2×2 estas se deben ver

$$\begin{bmatrix} g \end{bmatrix}_{\mathcal{B}}^{\mathcal{B}} = \begin{bmatrix} R & 0 \\ 0 & R \end{bmatrix} \quad \text{y} \quad \begin{bmatrix} h \end{bmatrix}_{\mathcal{B}}^{\mathcal{B}} = \begin{bmatrix} R & I \\ 0 & R \end{bmatrix},$$

donde R es una matriz de rotación e I es la matriz identidad.)

- (e) Encuentre los polinomios minimales de g y h .
- (f) Identifique entre los operadores g y h cual es semi-simple.

Capítulo 3

Espacio dual

Sea K un cuerpo y V, W espacios vectoriales sobre K .

Notación 3.1. Dada una colección de índices I , definimos para cada $i, j \in I$ el símbolo *delta de Kronecker*:

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

3.1. Funcionales lineales

Definición 3.2. El *espacio dual de V* es el espacio vectorial $V^* = \text{Hom}_K(V, K)$, es decir la colección de transformaciones lineales

$$\lambda : V \longrightarrow K.$$

A los elementos $\lambda \in V^*$ los llamamos *funcionales lineales*.

Proposición 3.3. Si V tiene dimensión finita $\dim(V) = \dim(V^*)$.

Dem. Sea $\{v_1, \dots, v_n\}$ una base de V , donde $n = \dim(V)$. Por Proposición 1.35.2, $\lambda \in V^*$ está unívocamente por los valores $\lambda(v_1), \dots, \lambda(v_n)$. Defina $\lambda_1, \dots, \lambda_n \in V^*$ por

$$\lambda_i(v_j) = \delta_{ij}.$$

Veamos que $\{\lambda_1, \dots, \lambda_n\}$ es una base de V^* probando que es linealmente independiente y que genera a V^* ; y así obtenemos $\dim(V^*) = n$. Para la independencia lineal, tome $a_1, \dots, a_n \in K$ tales que

$$\sum_{i=1}^n a_i \lambda_i = 0.$$

De forma que, para $j = 1, \dots, n$,

$$0 = \left(\sum_{i=1}^n a_i \lambda_i \right) (v_j) = \sum_{i=1}^n a_i \lambda_i(v_j) = \sum_{i=1}^j a_i \delta_{ij} = a_j.$$

Para ver que $V^* = \langle \lambda_1, \dots, \lambda_n \rangle$, dado $\lambda \in V^*$, defina $a_i = \lambda(v_i)$ y sea

$$\mu = \sum_{i=1}^n a_i \lambda_i.$$

De forma que, para $j = 1, \dots, n$,

$$\mu(v_j) = \left(\sum_{i=1}^n a_i \lambda_i \right) (v_j) = \sum_{i=1}^j a_i \delta_{ij} = a_j = \lambda(v_j),$$

luego $\mu = \lambda$. □

Definición 3.4. Suponga que V tiene dimensión finita y sea $\mathcal{B} = \{v_1, \dots, v_n\}$ una base de V , donde $n = \dim(V)$. A la base $\mathcal{B}^* = \{\lambda_1, \dots, \lambda_n\}$ de V^* donde

$$\lambda_i(v_j) = \delta_{ij}.$$

la llamamos *base dual* de \mathcal{B} .

Observación 3.5. Si V tiene dimensión finita, $\mathcal{B} = \{v_1, \dots, v_n\}$ es una base de V y $\mathcal{B}^* = \{\lambda_1, \dots, \lambda_n\}$ es la base dual, entonces para todo $v \in V$

$$v = \sum_{i=1}^n \lambda_i(v) v_i.$$

De hecho si $a_1, \dots, a_n \in K$ son tales que $a_1 v_1 + \dots + a_n v_n = v$,

$$\lambda_i(v) = \lambda_i \left(\sum_{j=1}^n a_j v_j \right) = \sum_{j=1}^n a_j \delta_{ij} = a_i.$$

Es decir, λ_i arroja la coordenada en v_i .

Observación 3.6. Si V tiene dimensión infinita y $\{v_i\}_{i \in I}$ es una base de V , igualmente podemos definir la colección $\{\lambda_i\}_{i \in I} \subseteq V^*$ por

$$\lambda_i(v_j) = \delta_{ij}.$$

e igualmente tenemos que para todo $v \in V$

$$v = \sum_{i \in I} \lambda_i(v) v_i.$$

Note que $\lambda_i(v) = 0$ para todo $i \in I$ salvo para una subcolección finita de índices. La diferencia con el caso en dimensión infinita es que $\{\lambda_i\}_{i \in I}$ no es una base de V^* , pues en tal caso el funcional lineal λ definido por $\lambda(v_i) = 1$, para todo $i \in I$, no es una combinación lineal de $\{\lambda_i\}_{i \in I}$.

Definición 3.7. Sean $S \subseteq V$ y $L \subseteq V^*$, definimos:

1. el *anulador* de S por

$$S^0 = \{\lambda \in V^* \mid \lambda(v) = 0, \forall v \in S\};$$

2. el *cero* de L por

$$L_0 = \{v \in V \mid \lambda(v) = 0, \forall \lambda \in L\}.$$

Propiedad 3.8. Sean $S \subseteq V$ y $L \subseteq V^*$. Tenemos:

1. $S^0 \leq V^*$ y $L_0 \leq V$;
 2. si $S_1, S_2 \subseteq V$ y $L_1, L_2 \subseteq V^*$ son tales que

$$S_1 \subseteq S_2 \quad y \quad L_1 \subseteq L_2,$$

entonces

$$S_2^0 \leq S_1^0 \quad y \quad (L_2)_0 \leq (L_1)_0;$$

3. $\langle S^0 \rangle_0 = \text{Sp}(S)$;
 4. si $V_1, V_2 \leq V$ y $V_1^*, V_2^* \leq V^*$, entonces

$$(V_1 + V_2)^0 = V_1^0 \cap V_2^0, \quad y \quad (V_1^* + V_2^*)_0 = (V_1^*)_0 \cap (V_2^*)_0;$$

5. si V tiene *dimensión finita*,

$$\dim(\langle S \rangle) + \dim(S^0) = \dim(V), \quad y \quad \dim(L_0) + \dim(\langle L \rangle) = \dim(V^*).$$

Dem.

1. Si $\lambda_1, \lambda_2 \in S^0$ y $c \in K$, entonces para todo $v \in S$

$$(\lambda_1 + \lambda_2)(v) = \lambda_1(v) + \lambda_2(v) = 0, \quad y \quad (c\lambda_1)(v) = c\lambda_1(v) = 0,$$

es decir $\lambda_1 + \lambda_2 \in S^0$ y $c\lambda_1 \in S^0$. Luego S_0 es un subespacio de V^* .
 Similarmente L_0 es un subespacio de V .

2. Sea $\lambda \in S_2^0$, luego, si $v \in S_1$, como $v \in S_2$, entonces $\lambda(v) = 0$; en particular $\lambda \in S_1^0$. Similarmente, sea $v \in (L_2)_0$, luego, si $\lambda \in L_1$, como $\lambda \in L_2$, entonces $\lambda(v) = 0$; en particular $v \in (L_1)_0$.

3. Sea $v \in \langle S \rangle$, entonces existen $v_1, \dots, v_m \in S$ y $a_1, \dots, a_m \in K$ tales que

$$v = a_1v_1 + \dots + a_mv_m$$

así, si $\lambda \in S^0$, $\lambda(v) = a_1\lambda(v_1) + \dots + a_m\lambda(v_m) = 0$. Luego $\langle S \rangle \leq (S^0)_0$.
 Tome ahora un subconjunto $S' \subseteq S$ linealmente independiente tal que

$\langle S' \rangle = \langle S \rangle$, el cual extendemos a una base $\mathcal{B} = \{v_i\}_{i \in I}$ de V . Defina, para cada $i \in I$, $\lambda_i \in V^*$ por

$$\lambda_i(v_j) = \delta_{ij}$$

para todo $j \in I$. De esta forma $\lambda_i \in S^0$ si y solo si $v_i \notin S'$. Sea $J \subset I$ la subcolección de índices definida por $j \in J$ si $v_j \in S'$. Entonces $L = \{\lambda_i\}_{i \in I \setminus J} \subseteq S^0$ y $(S^0)_0 \leq L_0$. Ahora si $v \in V$, como

$$v = \sum_{i \in J} \lambda_i(v)v_i + \sum_{i \in I \setminus J} \lambda_i(v)v_i$$

entonces $v \in L_0$ si y solo si $v \in \langle S' \rangle$, es decir $L_0 = \langle S' \rangle$. Luego $(S^0)_0 \leq \langle S' \rangle = \langle S \rangle$.

4. Suponga que $\lambda \in (V_1 + V_2)^0$, luego, si $v \in V_i$, con $i = 1$ ó $i = 2$, entonces $v \in V_1 + V_2$ y $\lambda(v) = 0$, en particular $\lambda \in V_1^0 \cap V_2^0$. Recíprocamente, si $\lambda \in V_1^0 \cap V_2^0$ y $v \in V_1 + V_2$, con $v = v_1 + v_2$, $v_1 \in V_1$ y $v_2 \in V_2$, $\lambda(v) = \lambda(v_1) + \lambda(v_2) = 0$, en particular $\lambda \in (V_1 + V_2)^0$.
 Similarmente, suponga que $v \in (V_1^* + V_2^*)_0$, luego, si $\lambda \in V_i^*$, con $i = 1$ ó $i = 2$, entonces $\lambda \in V_1^* + V_2^*$ y $\lambda(v) = 0$, en particular $v \in (V_1^*)_0 \cap (V_2^*)_0$. Recíprocamente, si $v \in (V_1^*)_0 \cap (V_2^*)_0$ y $\lambda \in V_1^* + V_2^*$, con $\lambda = \lambda_1 + \lambda_2$, $\lambda_1 \in V_1^*$ y $\lambda_2 \in V_2^*$, $\lambda(v) = \lambda_1(v) + \lambda_2(v) = 0$, en particular $v \in (V_1^* + V_2^*)_0$.
5. Tome un subconjunto $\{v_1, \dots, v_k\} = S' \subseteq S$ linealmente independiente tal que $\langle S' \rangle = \langle S \rangle$, el cual extendemos a una base $\mathcal{B} = \{v_1, \dots, v_n\}$ de V . Sea $\mathcal{B}^* = \{\lambda_1, \dots, \lambda_n\}$ la base dual a \mathcal{B} . Defina $f \in \text{Hom}_K(V, V)$ por

$$f(v) = \sum_{i=1}^k \lambda_i(v)v_i.$$

Por construcción, $\text{im}(f) = \langle S' \rangle = \langle S \rangle$ y $\ker(f) = \langle \lambda_{k+1}, \dots, \lambda_n \rangle_0$. Pero $\langle \lambda_{k+1}, \dots, \lambda_n \rangle = S^0$, luego $\dim(\langle S \rangle) + \dim(S^0) = n$.

Similarmente, tome un subconjunto $\{\lambda_1, \dots, \lambda_k\} = L' \subseteq L$ linealmente independiente tal que $\langle L' \rangle = \langle L \rangle$, el cual extendemos a una base $\mathcal{B}^* = \{\lambda_1, \dots, \lambda_n\}$ de V . Sea $\{v_1, \dots, v_n\}$ una base de V . Defina $f \in \text{Hom}_K(V, V)$ por

$$f(v) = \sum_{i=1}^k \lambda_i(v)v_i.$$

Por construcción, $\ker(f) = L'_0 = L_0$ y $\text{im}(f) = \langle v_1, \dots, v_k \rangle$. Luego $\dim \text{im}(f) = \dim(\langle L \rangle)$ y $\dim(L_0) + \dim(\langle L \rangle) = n$. \square

Proposición 3.9. Sea $V_1 \leq V$ entonces $(V/V_1)^* = V_1^0$.

Dem. Tome $\pi_{V_1} : V \rightarrow V/V_1$ con $\pi_{V_1}(v) = v + V_1$; y, defina la transformación lineal $f : (V/V_1)^* \rightarrow V^*$ por $f(\lambda) = \lambda \circ \pi_{V_1}$. Veamos que f es un isomorfismo. Primero es inyectiva pues si $\lambda \circ \pi_{V_1} = 0$, entonces, para todo $v + V_1 \in V/V_1$, $\lambda(v + V_1) = \lambda \circ \pi_{V_1}(v) = 0$. Es decir $\lambda = 0$. Por otro lado f es sobreyectiva,

pues dado $\mu \in V_1^0$, si $v - v' \in V_1$ entonces $\mu(v) - \mu(v') = \mu(v - v') = 0$, luego la función $\lambda : V/V_1 \rightarrow K$ tal que $\lambda(v + V_1) = \mu(v)$ es un funcional lineal tal que $f(\lambda) = \mu$. \square

Teorema 3.10. *Existe una transformación lineal canónica inyectiva*

$$\begin{aligned}\widehat{\bullet} : V &\longrightarrow (V^*)^* \\ v &\longmapsto \widehat{v} : \lambda \mapsto \lambda(v).\end{aligned}$$

Si V tiene dimensión finita, $\widehat{\bullet}$ es un isomorfismo.

Dem. Por definición $\widehat{\bullet}$ es lineal, pues

$$\widehat{v_1 + v_2}(\lambda) = \lambda(v_1 + v_2) = \lambda(v_1) + \lambda(v_2) = (\widehat{v_1} + \widehat{v_2})(\lambda).$$

Ahora sea $\{v_i\}_{i \in I}$ una base de V , y $\{\lambda_i\}_{i \in I} \subseteq V^*$ la colección tal que $\lambda_i(v_j) = \delta_{ij}$. Como para todo $v \in V$, $v = \sum_{i \in I} \lambda_i(v)v_i$, si $\widehat{v} = 0$ entonces $\lambda_i(v) = 0$ para todo i y así $v = 0$. Luego $\widehat{\bullet}$ es inyectiva.

Si V tiene dimensión finita, $\dim(V) = \dim(V^*) = \dim((V^*)^*)$, entonces $\widehat{\bullet}$ es un isomorfismo. \square

3.2. Transformación dual

Definición 3.11. Sea $f \in \text{Hom}_K(V, W)$. Definimos la *transformación dual*, $f^* \in \text{Hom}_K(W^*, V^*)$, por

$$f^*(\lambda) = \lambda \circ f$$

para todo $\lambda \in W^*$.

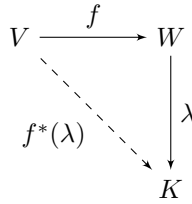


Figura 3.1: Transformación dual

Observación 3.12. La linealidad del mapa f^* se sigue de las siguientes igualdades, validas para todo $f \in \text{Hom}_K(V, W)$, $\lambda_1, \lambda_2 \in W^*$, $c \in K$:

$$\begin{aligned}(\lambda_1 + \lambda_2) \circ f &= \lambda_1 \circ f + \lambda_2 \circ f \\ (c\lambda_1) \circ f &= c(\lambda_1 \circ f)\end{aligned}$$

Propiedad 3.13. Sean U un espacio vectorial sobre K y $f \in \text{Hom}_K(V, W)$ y $g \in \text{Hom}_K(W, U)$, entonces

$$(g \circ f)^* = f^* \circ g^*$$

Dem. Para $\lambda \in U^*$, tenemos

$$(g \circ f)^* \lambda = \lambda \circ (g \circ f) = (\lambda \circ g) \circ f = g^*(\lambda) \circ f = f^* \circ g^*(\lambda)$$

□

Propiedad 3.14. Sea $f \in \text{Hom}_K(V, W)$, entonces

1. Si f es sobreyectiva, f^* es inyectiva; y,
2. Si f es inyectiva, f^* es sobreyectiva.

Dem.

1. Sea $\lambda \in W^*$ tal que $f^*(\lambda) = 0$, entonces, dado $w \in W$, como f es sobreyectiva, existe $v \in V$ tal que $w = f(v)$, así

$$\lambda(w) = \lambda(f(v)) = f^*(\lambda)(v) = 0$$

luego $\lambda = 0$ y f^* es inyectiva.

2. Sea $W_1, W_2 \leq W$ tales que $W = W_1 \oplus W_2$ y $W_1 = f(V)$. Tome $\mu \in V$, y defina $\lambda : W \rightarrow K$ por

$$\lambda(w) = \mu(v_1)$$

donde $w = w_1 + w_2$, $w_1 \in W_1$ y $w_2 \in W_2$ y $f(v_1) = w_1$. Como f es inyectiva, v_1 es único, y la función λ está bien definida. Como la descomposición de $w = w_1 + w_2$ es lineal y μ y f son lineales, entonces λ lineal, es decir $\lambda \in W^*$. Por construcción $\mu = f^*\lambda$ pues

$$f^*(\lambda)(v_1) = \lambda(f(v_1)) = \lambda(w_1) = \mu(v_1),$$

luego f^* es sobreyectiva. □

Propiedad 3.15. Sean $V_1, V_2 \leq V$ tales que $V = V_1 \oplus V_2$; y, $\pi_1 : V \rightarrow V_1$ y $\pi_2 : V \rightarrow V_2$ respectivamente las proyecciones sobre V_1 y V_2 dadas por la descomposición $V = V_1 \oplus V_2$. Entonces

$$V^* = \pi_1^*(V_1^*) \oplus \pi_2^*(V_2^*)$$

Dem. Dado $\lambda \in V^*$, defina $\lambda_1 \in V_1^*$ y $\lambda_2 \in V_2^*$ por

$$\lambda_1(v_1) = \lambda(v_1) \quad \lambda_2(v_2) = \lambda(v_2).$$

De tal forma que si $v = v_1 + v_2 \in V$ con $v_1 = \pi_1(v)$ y $v_2 = \pi_2(v)$, entonces

$$\begin{aligned} \lambda(v) &= \lambda(v_1) + \lambda(v_2) \\ &= \lambda_1(\pi_1(v)) + \lambda_2(\pi_2(v)) \\ &= (\pi_1^*(\lambda_1) + \pi_2^*(\lambda_2))(v) \end{aligned}$$

luego $V^* = \pi_1^*(V_1^*) + \pi_2^*(V_2^*)$. Ahora, si $\lambda \in \pi_1^*(V_1^*) \cap \pi_2^*(V_2^*)$, existen $\lambda_1 \in V_1^*$ y $\lambda_2 \in V_2^*$ tales que $\lambda = \pi_1^*(\lambda_1) = \pi_2^*(\lambda_2)$, de esta forma, para todo $v = v_1 + v_2 \in V$ con $v_1 = \pi_1(v)$ y $v_2 = \pi_2(v)$

$$\lambda(v) = \lambda(v_1) + \lambda(v_2) = \lambda_2(\pi_2(v_1)) + \lambda_1(\pi_1(v_2)) = \lambda_2(0) + \lambda_1(0) = 0.$$

Luego $\pi_1^*(V_1^*) \cap \pi_2^*(V_2^*) = \{0\}$. □

Teorema 3.16. *El mapa*

$$\begin{aligned} \bullet^* : \text{Hom}_K(V, W) &\longrightarrow \text{Hom}_K(W^*, V^*) \\ f &\longmapsto f^* \end{aligned}$$

es una transformación lineal inyectiva. Si W tiene dimensión finita entonces es un isomorfismo.

Dem. Sean $f, g \in \text{Hom}_K(V, W)$, y $c \in K$. Dado $\lambda \in W^*$ tenemos:

$$\begin{aligned} (f + g)^*(\lambda) &= \lambda \circ (f + g) \\ &= \lambda \circ f + \lambda \circ g \\ &= f^*(\lambda) + g^*(\lambda) \\ &= (f^* + g^*)(\lambda) \\ (cf)^*(\lambda) &= \lambda \circ (cf) \\ &= c(\lambda \circ f) \\ &= cf^*(\lambda). \end{aligned}$$

Es decir $(f + g)^* = f^* + g^*$ y $(cf)^* = cf^*$, y así \bullet^* es lineal.

Ahora suponga que $f^* = 0$, es decir $f^*(\lambda) = 0$ para todo $\lambda \in W^*$. Tomamos una base $\{w_i\}_{i \in I}$ de W y definimos $\{\lambda_i\}_{i \in I} \subseteq W^*$ por

$$\delta_i(v_j) = \delta_{ij}$$

Así, como en Observación 3.6, tenemos para todo $v \in V$

$$f(v) = \sum_{i \in I} \lambda_i(f(v)) w_i = \sum_{i \in I} f^*(\lambda_i)(v) w_i = \sum_{i \in I} 0 w_i = 0.$$

Es decir $f = 0$ y así \bullet^* es inyectiva.

Si además asumimos que W tiene dimensión finita, entonces $\{\lambda_i\}_{i \in I}$ es la base de W^* , dual de $\{w_i\}_{i \in I}$. En particular $\phi \in \text{Hom}^*(W^*, V^*)$ está unívocamente determinado por la imagen $\{\phi(\lambda_i)\}_{i \in I} \subseteq V^*$. Defina $f \in \text{Hom}_K(V, W)$ por la suma finita

$$f(v) = \sum_{j \in I} [\phi(\lambda_j)(v)] w_j,$$

de forma tal que para todo $i \in I$, $v \in V$,

$$\begin{aligned} f^*(\lambda_i)(v) &= \lambda_i(f(v)) \\ &= \lambda_i \left(\sum_{j \in I} [\phi(\lambda_j)(v)] w_j \right) \\ &= \sum_{j \in I} [\phi(\lambda_j)(v)] \lambda_i(w_j) \\ &= \sum_{j \in I} [\phi(\lambda_j)(v)] \delta_{ij} = \phi(\lambda_i)(v). \end{aligned}$$

Es decir $f^*(\lambda_i) = \phi(\lambda_i)$, para todo $i \in I$. Luego $f^* = \phi$, de donde \bullet^* es también sobreyectiva, así es un isomorfismo. \square

Observación 3.17. Suponga que W tiene dimensión finita, sea $\mathcal{B}_W = \{w_1, \dots, w_m\}$ una base de W y $\lambda \in W^* = \text{Hom}_K(W, K)$. Si tomamos la base $\{1\}$ de K , entonces la representación matricial de λ respecto a las base \mathcal{B}_W y $\{1\}$ es

$$[\lambda]_{\mathcal{B}_W}^{\{1\}} = [\lambda(w_1) \cdots \lambda(w_m)].$$

Ahora, si \mathcal{B}_V es una base de V y $f \in \text{Hom}_K(V, W)$, tenemos

$$[f^*(\lambda)]_{\mathcal{B}_V}^{\{1\}} = [\lambda \circ f]_{\mathcal{B}_V}^{\{1\}} = [\lambda]_S^{\{1\}} [f]_{\mathcal{B}_W}^{\mathcal{B}_V}.$$

Por otro lado, podemos tomar las coordenadas de λ en la base $\mathcal{B}_W^* = \{\lambda_1, \dots, \lambda_m\}$ de W^* dual de \mathcal{B}_W :

$$[\lambda]_{\mathcal{B}_W}^{\mathcal{B}_W^*} = \begin{bmatrix} \lambda(w_1) \\ \vdots \\ \lambda(w_m) \end{bmatrix}$$

Si además asumimos que V tiene también dimensión finita y tomamos la base \mathcal{B}_V^* de V^* dual de T , entonces

$$[f^*(\lambda)]_{\mathcal{B}_V}^{\mathcal{B}_V^*} = [f^*]_{\mathcal{B}_W}^{\mathcal{B}_V^*} [\lambda]_{\mathcal{B}_W}^{\mathcal{B}_W^*}.$$

La pregunta inmediata es: ¿Cuál es la relación entre $[f]_{\mathcal{B}_V}^{\mathcal{B}_W}$ y $[f^*]_{\mathcal{B}_W}^{\mathcal{B}_V^*}$?

Definición 3.18. Sean I, J conjuntos y $A \in M_{I \times J}(K)$, definimos la *matriz traspuesta* de A por $A^\top \in M_{J \times I}(K)$ tal que

$$A^\top(j, i) = A(i, j)$$

para todo $(j, i) \in J \times I$. Es decir el valor en (j, i) de A^\top es el valor en (i, j) de A . Similarmente si $m, n \in \mathbb{Z}_{>0}$, y $A \in M_{m \times n}(K)$, definimos su traspuesta por $A^\top \in M_{n \times m}(K)$ tal que

$$A^\top(j, i) = A(i, j)$$

Sea $A \in M_{I \times I}(K)$, o $A \in M_{n \times n}(K)$, decimos que A es *simétrica* si $A^\top = A$.

Teorema 3.19. Suponga que V y W tienen dimensión finita, $n = \dim(V) > 0$ y $m = \dim(W) > 0$, y $f \in \text{Hom}_K(V, W)$. Sean $\mathcal{B}_W = \{w_1, \dots, w_m\}$ y $\mathcal{B}_V = \{v_1, \dots, v_n\}$ respectivamente bases de W y V ; y, sean $\mathcal{B}_W^* = \{\lambda_1, \dots, \lambda_m\}$ y $\mathcal{B}_V^* = \{\mu_1, \dots, \mu_n\}$ respectivamente las bases de W^* y V^* duales de S y T . Sea $A \in M_{m \times n}(K)$ la representación matricial de f respecto a las bases \mathcal{B}_V y \mathcal{B}_W , entonces $A^\top \in M_{n \times m}(K)$ es la representación matricial de f^* respecto a las bases \mathcal{B}_V^* y \mathcal{B}_W^* . Es decir,

$$[f^*]_{\mathcal{B}_W}^{\mathcal{B}_V^*} = \left([f]_{\mathcal{B}_V}^{\mathcal{B}_W} \right)^\top$$

Dem. Si $a_{ij} \in K$ es la ij -ésima entrada de $A = \left[f \right]_{\mathcal{B}_V}^{\mathcal{B}_W}$, entonces

$$a_{ij} = \lambda_i(f(v_j)) = f^*(\lambda_i)(v_j);$$

y,

$$f^*(\lambda_i) = \sum_{l=1}^n [f^*(\lambda_i)(v_l)] \mu_l,$$

luego la ji -ésima entrada de $\left[f^* \right]_{\mathcal{B}_W}^{\mathcal{B}_V^*}$ es $f^*(\lambda_i)(v_j) = a_{ij}$. □

Capítulo 4

Espacios euclídeos

Sea V un espacio vectorial sobre \mathbb{R} .

4.1. Producto interno

Definición 4.1. Un *producto interno* en V es una función

$$\begin{aligned}\langle \bullet, \bullet \rangle : V \times V &\longrightarrow \mathbb{R} \\ (v_1, v_2) &\longmapsto \langle v_1; v_2 \rangle\end{aligned}$$

tal que:

1. *es bilineal*: para todo $v, v_1, v_2 \in V$ y $c \in \mathbb{R}$

$$\begin{aligned}\langle v_1 + v_2; v \rangle &= \langle v_1; v \rangle + \langle v_2; v \rangle \\ \langle cv_1; v_2 \rangle &= c\langle v_1; v_2 \rangle \\ \langle v; v_1 + v_2 \rangle &= \langle v; v_1 \rangle + \langle v; v_2 \rangle \\ \langle v_1; cv_2 \rangle &= c\langle v_1; v_2 \rangle;\end{aligned}$$

2. *es simétrica*: para todo $v_1, v_2 \in V$

$$\langle v_2; v_1 \rangle = \langle v_1; v_2 \rangle;$$

3. *es definitivamente positiva* para todo $v \in V$, $v \neq 0$,

$$\langle v; v \rangle > 0.$$

Un *espacio euclídeo* es un espacio vectorial sobre \mathbb{R} provisto de un producto interno.

Observación 4.2. Se sigue que $\langle v; v \rangle = 0$ si y solo si $v = 0$.

Ejemplo 4.3. 1. Sobre $V = \mathbb{R}^n$,

$$\langle (x_1, \dots, x_n); (y_1, \dots, y_n) \rangle = \sum_{i=1}^n x_i y_i.$$

2. Sobre $V = M_{n \times n}(\mathbb{R})$,

$$\langle A; B \rangle = \text{tr}(B^T A).$$

3. Sea $[a, b] \subseteq \mathbb{R}$ un intervalo cerrado. Sobre $V = C^0[a, b]$, el conjunto de funciones continuas $[a, b] \rightarrow \mathbb{R}$,

$$\langle f; g \rangle = \int_a^b f(x)g(x)dx.$$

Definición 4.4. Dado un espacio euclídeo V , definimos la norma de $v \in V$ por $\|v\| = \sqrt{\langle v; v \rangle}$.

Propiedad 4.5. Sea V un espacio euclídeo, entonces:

1. $\|cv\| = |c|\|v\|$, para todo $c \in \mathbb{R}$ y $v \in V$;
2. Desigualdad de Cauchy-Schwarz: $|\langle v_1; v_2 \rangle| \leq \|v_1\|\|v_2\|$, para todo $v_1, v_2 \in V$, más aún se tiene $|\langle v_1; v_2 \rangle| = \|v_1\|\|v_2\|$ únicamente cuando $\{v_1, v_2\}$ es linealmente dependiente; y,
3. Desigualdad triangular: $\|v_1 + v_2\| \leq \|v_1\| + \|v_2\|$, para todo $v_1, v_2 \in V$, más aún se tiene $\|v_1 + v_2\| = \|v_1\| + \|v_2\|$ si y solo si $av_1 = bv_2$ con $a, b \geq 0$.

Dem.

1. Dados $c \in \mathbb{R}$ y $v \in V$

$$\|cv\| = \sqrt{\langle cv; cv \rangle} = \sqrt{c^2 \langle v; v \rangle} = |c| \sqrt{\langle v; v \rangle} = |c|\|v\|.$$

2. Si $v_1 = 0$ o $v_2 = 0$ tenemos $0 = |\langle v_1, v_2 \rangle| = \|v_1\|\|v_2\|$. En el caso general, para todo $a, b \in \mathbb{R}$,

$$\begin{aligned} 0 \leq \|av_1 - bv_2\|^2 &= \langle av_1 - bv_2, av_1 - bv_2 \rangle \\ &= a^2 \langle v_1, v_1 \rangle - ab \langle v_2, v_1 \rangle - ab \langle v_1, v_2 \rangle + b^2 \langle v_2, v_2 \rangle \\ &= a^2 \|v_1\|^2 - 2ab \langle v_1, v_2 \rangle + b^2 \|v_2\|^2. \end{aligned}$$

Si suponemos ahora que $v_2 \neq 0$, es decir $\|v_2\|^2 > 0$, y $a \neq 0$, dividiendo por a^2 , obtenemos así una expresión cuadrática en b/a , positiva o nula para todo valor, luego su discriminante satisface

$$4\langle v_1, v_2 \rangle^2 - 4\|v_1\|^2\|v_2\|^2 \leq 0$$

Lo cual implica la desigualdad deseada. Igualmente, esta expresión cuadrática se anula, es decir $av_1 - bv_2 = 0$ para algún $a, b \in \mathbb{R}$, si y solo si su discriminante es cero, es decir si y solo si $\langle v_1, v_2 \rangle^2 = \|v_1\|^2\|v_2\|^2$.

3. Tomando $a = 1$ y $b = -1$ en la expresión arriba, obtenemos

$$\|v_1 + v_2\|^2 = \|v_1\|^2 + 2\langle v_1; v_2 \rangle + \|v_2\|^2;$$

y la desigualdad de Cauchy-Schwarz implica

$$\|v_1 + v_2\|^2 \leq \|v_1\|^2 + 2\|v_1\|\|v_2\| + \|v_2\|^2 = (\|v_1\| + \|v_2\|)^2$$

que es equivalente a la desigualdad afirmada. Se obtiene una igualdad en la desigualdad triangular si y solo si $\langle v_1, v_2 \rangle = \|v_1\|\|v_2\|$, lo cual equivale a $av_1 - bv_2 = 0$ donde $a = \|v_2\|$ y $b = \|v_1\|$. \square

Observación 4.6. En vista de la desigualdad de Cauchy-Schwarz, es usual definir el ángulo entre v_1 y v_2 por $\theta = \arccos(\langle v_1; v_2 \rangle / (\|v_1\|\|v_2\|))$, siempre que $v_1 \neq 0$ y $v_2 \neq 0$. La dirección, o el signo, del ángulo no se puede definir intrínsecamente a partir del producto interno, hace falta definir una orientación. Con esta definición del ángulo entre dos elementos distintos del origen en un espacio euclídeo, se empata con la definición clásica, según la cual

$$\langle v_1; v_2 \rangle = \|v_1\|\|v_2\| \cos \theta.$$

cuando $v_1 \neq 0$ y $v_2 \neq 0$. El caso particular de mayor interés es naturalmente cuando existe una relación de ortogonalidad entre v_1 y v_2 .

Definición 4.7. Sea V un espacio euclídeo y $S \subseteq V$. Decimos que S es *ortogonal* si $\langle v_1; v_2 \rangle = 0$ para todo $v_1, v_2 \in S$ tales que $v_1 \neq v_2$. Si además $\|v\| = 1$ para todo $v \in S$, decimos que S es *ortonormal*.

Observación 4.8. Note que $S = \{v_i\}_{i \in I}$ es ortonormal si y solo si

$$\langle v_i; v_j \rangle = \delta_{ij}.$$

para todo $i, j \in I$.

Propiedad 4.9. Sea V un espacio euclídeo. Si $S \subset V$ es ortogonal, y $0 \notin S$, entonces S es linealmente independiente.

Dem. Suponga que $v_1, \dots, v_n \in S$ y $a_1, \dots, a_n \in S$ son tales que

$$a_1 v_1 + \dots + a_n v_n = 0.$$

Entonces, para $i = 1, \dots, n$,

$$0 = \langle 0; v_i \rangle = \langle a_1 v_1 + \dots + a_n v_n; v_i \rangle = a_1 \langle v_1; v_i \rangle + \dots + a_n \langle v_n; v_i \rangle = a_i \|v_i\|^2,$$

pero $v_i \neq 0$ pues $0 \notin S$, luego $\|v_i\|^2 \neq 0$, y así $a_i = 0$. \square

Teorema 4.10 (Ortogonalización de Gram-Schmidt). Sea V un espacio euclídeo. Suponga que V tiene dimensión finita, entonces V tiene una base ortonormal. Más aún, si $\{v_1, \dots, v_n\}$ es una base de V , existe una base ortonormal $\{u_1, \dots, u_n\}$ de V tal que para $k = 1, \dots, n$

$$\langle v_1, \dots, v_k \rangle = \langle u_1, \dots, u_k \rangle.$$

Dem. Sea $\{v_1, \dots, v_n\}$ una base de V , definimos recursivamente $\{v'_1, \dots, v'_n\}$ por

$$\begin{aligned} v'_1 &= v_1 \\ v'_{k+1} &= v_{k+1} - \sum_{i=1}^k \frac{\langle v_{k+1}; v'_i \rangle}{\|v'_i\|^2} v'_i. \end{aligned}$$

Veamos que $\{v'_1, \dots, v'_n\}$ es ortogonal. Para esto, basta establecer por inducción que, si $1 \leq k < n$, entonces $\langle v'_{k+1}; v'_j \rangle = 0$ para $1 \leq j \leq k$. Para el caso base, $k = 1 = j$ y

$$\begin{aligned} \langle v'_2; v'_1 \rangle &= \langle v_2; v'_1 \rangle - \frac{\langle v_2; v'_1 \rangle}{\|v'_1\|^2} \langle v'_1; v'_1 \rangle \\ &= \langle v_2; v'_1 \rangle - \langle v_2; v'_1 \rangle \\ &= 0. \end{aligned}$$

Para el paso inductivo, asumimos que $\langle v'_i; v'_j \rangle = 0 = \langle v'_j; v'_i \rangle$, siempre que $1 \leq i < j \leq k$, de tal forma que si $1 \leq j \leq k$,

$$\begin{aligned} \langle v'_{k+1}; v'_j \rangle &= \langle v_{k+1}; v'_j \rangle - \sum_{i=1}^k \frac{\langle v_{k+1}; v'_i \rangle}{\|v'_i\|^2} \langle v'_i; v'_j \rangle \\ &= \langle v_{k+1}; v'_j \rangle - \frac{\langle v_{k+1}; v'_j \rangle}{\|v'_j\|^2} \langle v'_j; v'_j \rangle \\ &= \langle v_{k+1}; v'_j \rangle - \langle v_{k+1}; v'_j \rangle \\ &= 0 \end{aligned}$$

Además recursivamente vemos que

$$\begin{aligned} \langle v_1 \rangle &= \langle v'_1 \rangle \\ \langle v_1, \dots, v_{k+1} \rangle &= \langle v'_1, \dots, v'_{k+1} \rangle \end{aligned}$$

pues $v_{k+1} - v'_{k+1} \in \langle v'_1, \dots, v'_k \rangle = \langle v_1, \dots, v_k \rangle$. Note que, como $\{v_1, \dots, v_n\}$ es linealmente independiente, entonces $v_{k+1} \notin \langle v_1, \dots, v_k \rangle = \langle v'_1, \dots, v'_k \rangle$, luego $v'_{k+1} \notin \langle v'_1, \dots, v'_k \rangle$ y así, para $i = 1, \dots, n$, $v'_i \neq 0$. De donde, como $\{v'_1, \dots, v'_n\}$ es ortogonal y no contiene al origen, es un conjunto linealmente independiente con el mismo número de elementos que la dimensión de V , entonces es una base de V .

Finalmente, para obtener la base ortonormal basta tomar, para $i = 1, \dots, n$,

$$u_i = \frac{1}{\|v'_i\|} v'_i.$$

Tenemos para $k = 1, \dots, n$

$$\langle v_1, \dots, v_k \rangle = \langle v'_1, \dots, v'_k \rangle = \langle u_1, \dots, u_k \rangle.$$

□

Propiedad 4.11. Sea V un espacio euclídeo. Suponga que V tiene dimensión finita y $\{u_1, \dots, u_n\}$ es una base ortonormal de V , entonces para todo $v \in V$

$$v = \sum_{i=1}^n \langle v; u_i \rangle u_i.$$

En particular, si $v_1, v_2 \in V$ son tales que

$$v_1 = \sum_{i=1}^n x_i u_i \quad v_2 = \sum_{i=1}^n y_i u_i,$$

entonces

$$\langle v_1; v_2 \rangle = \sum_{i=1}^n x_i y_i.$$

Dem. Como $\{u_1, \dots, u_n\}$ es base, existen $a_1, \dots, a_n \in K$ tal que $v = \sum_{i=1}^n a_i u_i$. De esta forma, para $j = 1, \dots, n$,

$$\langle v; u_j \rangle = \sum_{i=1}^n a_i \langle u_i; u_j \rangle = \sum_{i=1}^n a_i \delta_{ij} = a_j$$

(ver Observación 4.8). Finalmente,

$$\langle v_1; v_2 \rangle = \left\langle \sum_{i=1}^n x_i u_i; \sum_{j=1}^n y_j u_j \right\rangle = \sum_{i=1}^n \sum_{j=1}^n x_i y_j \delta_{ij} = \sum_{i=1}^n x_i y_i.$$

□

Observación 4.12. Si $\mathcal{B} = \{u_1, \dots, u_n\}$ es una base ortonormal de V , la propiedad anterior implica que para todo $v_1, v_2 \in V$ tenemos

$$\langle v_1; v_2 \rangle = \left([v_2]^{\mathcal{B}} \right)^{\top} [v_1]^{\mathcal{B}}$$

Definición 4.13. Sean V un espacio euclídeo y $S \subseteq V$, el *conjunto ortogonal* a S está definido por

$$S^{\perp} = \{v \in V \mid \langle v; u \rangle = 0, \text{ para todo } u \in S\}$$

Propiedad 4.14. Sean V un espacio euclídeo y $S \subseteq V$, entonces $S^{\perp} \leq V$.

Dem. Como $\langle 0, v \rangle = 0$ para todo $v \in V$, $0 \in S^{\perp}$. Tome ahora $v_1, v_2, v \in S^{\perp}$ y $a \in \mathbb{R}$, entonces para todo $u \in S$

$$\begin{aligned} \langle v_1 + v_2; u \rangle &= \langle v_1; u \rangle + \langle v_2; u \rangle = 0 \\ \langle av; u \rangle &= a \langle v; u \rangle = 0 \end{aligned}$$

luego $v_1 + v_2 \in S^{\perp}$ y $av \in S^{\perp}$; así, S^{\perp} es un subespacio de V .

□

Teorema 4.15. *Sea V un espacio euclídeo. Suponga que V tiene dimensión finita y sea $U \leq V$, entonces*

$$V = U \oplus U^\perp$$

Dem. Sean $n = \dim(V)$, $m = \dim(U)$ y $\{v_1, \dots, v_n\}$ una base de V tal que $\{v_1, \dots, v_m\}$ es una base de U . Sea $\{u_1, \dots, u_n\}$ la base obtenida mediante ortogonalización a partir de $\{v_1, \dots, v_n\}$. En particular $U = \langle u_1, \dots, u_m \rangle$ y $\langle u_{m+1}, \dots, u_n \rangle \leq U^\perp$. Tome $v \in U^\perp$, tenemos

$$v = \sum_{i=1}^n \langle v; u_i \rangle u_i = \sum_{i=m+1}^n \langle v; u_i \rangle u_i,$$

luego $v \in \langle u_{m+1}, \dots, u_n \rangle$. Entonces $U^\perp = \langle u_{m+1}, \dots, u_n \rangle$ y $V = U \oplus U^\perp$. \square

Definición 4.16. Sea V un espacio euclídeo de dimensión finita y $U \leq V$. Llamamos a U^\perp el *complemento ortogonal de U* . A la proyección

$$p_U^\perp : V \longrightarrow V$$

sobre U , definida por la descomposición $V = U \oplus U^\perp$ la llamamos *proyección ortogonal sobre U* .

Propiedad 4.17. *Sea V un espacio euclídeo de dimensión finita, $\dim(V) = n$, y $\langle v_1, \dots, v_m \rangle = U \leq V$. Si \mathcal{B} es una base ortonormal de V y $\{v_1, \dots, v_m\}$ es linealmente independiente entonces*

$$\left[p_U^\perp \right]_{\mathcal{B}}^{\mathcal{B}} = A(A^\top A)^{-1} A^\top$$

donde $A \in M_{n \times m}(\mathbb{R})$ es la matrix cuya j -ésima columna es $\left[v_j \right]_{\mathcal{B}}^{\mathcal{B}}$, $j = 1, \dots, m$.

Dem. Para todo $v \in V$ tenemos $p_U^\perp(v) \in U$ luego existe un único $\bar{c}_v \in M_{m \times 1}(\mathbb{R})$ tal que

$$\left[p_U^\perp(v) \right]_{\mathcal{B}}^{\mathcal{B}} = A \bar{c}_v$$

Si $P = \left[p_U^\perp \right]_{\mathcal{B}}^{\mathcal{B}}$, y $\bar{x} = \left[v \right]_{\mathcal{B}}^{\mathcal{B}}$ entonces

$$P\bar{x} = A\bar{c}_v.$$

Ahora como $v - p_U^\perp(v) \in U^\perp$ entonces $\langle v - p_U^\perp(v); v_j \rangle = 0$ para $j = 1, \dots, m$ luego (ver Observación 4.12)

$$\left(\left[v_j \right]_{\mathcal{B}}^{\mathcal{B}} \right)^\top (\bar{x} - A\bar{c}_v) = 0$$

y

$$\begin{aligned} 0 &= A^\top (\bar{x} - A\bar{c}_v) \\ &= A^\top \bar{x} - A^\top A \bar{c}_v \end{aligned}$$

Veamos que $A^T A \in M_{m \times m}(\mathbb{R})$ es invertible. De hecho, si $\{u_1, \dots, u_m\}$ es una base ortonormal de U y $c_{ij} \in \mathbb{R}$, $i, j = 1, \dots, m$, son tales que

$$u_j = c_{1j}v_1 + \dots + c_{mj}v_m, \quad j = 1, \dots, m$$

y $C = (c_{ij})$ entonces $C \in M_{m \times m}(\mathbb{R})$ es invertible y la j -ésima columna de AC es $\begin{bmatrix} u_j \end{bmatrix}^B$. Así

$$I_m = (AC)^T AC = C^T A^T AC$$

y $A^T A = (CC^T)^{-1}$, luego $A^T A$ es invertible. Tenemos entonces

$$\bar{c}_v = (A^T A)^{-1} A \bar{x}$$

$$A \bar{c}_v = A(A^T A)^{-1} A \bar{x}$$

$$P \bar{x} = A(A^T A)^{-1} A \bar{x}$$

y se sigue $P = A(A^T A)^{-1} A^T$. □

Propiedad 4.18. Sea V un espacio euclídeo. Suponga que V tiene dimensión finita y sea $U \leq V$. Tenemos para todo $v_1, v_2 \in V$

$$\langle p_U^\perp(v_1); v_2 \rangle = \langle v_1; p_U^\perp(v_2) \rangle.$$

Si $\{u_1, \dots, u_m\}$ es una base ortonormal de U entonces para todo $v \in V$

$$p_U^\perp(v) = \sum_{i=1}^m \langle v; u_i \rangle u_i.$$

Dem. Sea $v'_1, v'_2 \in U^\perp$ tales que

$$v_1 = p_U^\perp(v_1) + v'_1 \quad v_2 = p_U^\perp(v_2) + v'_2.$$

Entonces

$$\begin{aligned} \langle p_U^\perp(v_1); v_2 \rangle &= \langle p_U^\perp(v_1); p_U^\perp(v_2) \rangle + \langle p_U^\perp(v_1); v'_2 \rangle = \langle p_U^\perp(v_1); p_U^\perp(v_2) \rangle \\ \langle v_1; p_U^\perp(v_2) \rangle &= \langle p_U^\perp(v_1); p_U^\perp(v_2) \rangle + \langle v'_1; p_U^\perp(v_2) \rangle = \langle p_U^\perp(v_1); p_U^\perp(v_2) \rangle. \end{aligned}$$

Finalmente, si completamos la base ortonormal $\{u_1, \dots, u_m\}$ de U a una base ortonormal $\{u_1, \dots, u_n\}$ de V ,

$$v = \underbrace{\sum_{i=1}^m \langle v; u_i \rangle u_i}_{\in U} + \underbrace{\sum_{i=m+1}^n \langle v; u_i \rangle u_i}_{\in U^\perp}.$$

□

Observación 4.19. Los operadores sobre un espacio euclídeo que, como la proyección ortogonal, pasan de un lado al otro del producto interno tienen varias propiedades, la más importante de ellas es que son diagonalizables, el cual es el contenido del Teorema Espectral. Antes de establecer este resultado, necesitamos elaborar la teoría de los operadores adjuntos.

4.2. Operador adjunto

Sea V un espacio euclídeo y $f \in \text{Hom}_{\mathbb{R}}(V, V)$ un operador.

Definición 4.20. Sea $g \in \text{Hom}_{\mathbb{R}}(V, V)$, decimos que g es un *operador adjunto de f* si para todo $v_1, v_2 \in V$

$$\langle g(v_1); v_2 \rangle = \langle v_1; f(v_2) \rangle.$$

Decimos que f es *auto-adjunto* si f es un operador adjunto de f .

Observación 4.21. Note que si g es adjunto de f , entonces f es adjunto de g . De hecho

$$\langle f(v_1); v_2 \rangle = \langle v_2; f(v_1) \rangle = \langle g(v_2); v_1 \rangle = \langle v_1; g(v_2) \rangle$$

Proposición 4.22. Suponga que V tiene dimensión finita, entonces existe un único operador $g \in \text{Hom}_{\mathbb{R}}(V, V)$ adjunto de f . Más aún, si $\mathcal{B} = \{u_1, \dots, u_n\}$ es una base ortonormal de V , entonces

$$[g]_{\mathcal{B}}^{\mathcal{B}} = \left([f]_{\mathcal{B}}^{\mathcal{B}} \right)^{\top}$$

Dem. Defina el operador $g \in \text{Hom}_{\mathbb{R}}(V, V)$ por la imagen de la base \mathcal{B} :

$$g(u_j) = \sum_{i=1}^n \langle u_j; f(u_i) \rangle u_i.$$

De esta forma

$$\langle g(u_j); u_i \rangle = \langle u_j; f(u_i) \rangle$$

y por bilinearidad del producto interno, g es adjunto de f . Por otro lado si, $h \in \text{Hom}_{\mathbb{R}}(V, V)$ es adjunto de f , por Propiedad 4.11,

$$\begin{aligned} h(u_j) &= \sum_{i=1}^n \langle h(u_j); u_i \rangle u_i \\ &= \sum_{i=1}^n \langle u_j; f(u_i) \rangle u_i \\ &= g(u_j), \end{aligned}$$

luego $h = g$.

Ahora, para ver que la representación matricial de g respecto a \mathcal{B} es la traspuesta

de la de f basta observar que

$$\begin{aligned} \left[g \right]_{\mathcal{B},(j,i)}^{\mathcal{B}} &= \left[g(u_i) \right]_j^{\mathcal{B}} \\ &= \langle g(u_i); u_j \rangle \\ &= \langle u_i; f(u_j) \rangle \\ &= \langle f(u_j); u_i \rangle \\ &= \left[f(u_j) \right]_i^{\mathcal{B}} \\ &= \left[f \right]_{\mathcal{B},(i,j)}^{\mathcal{B}} \end{aligned}$$

□

Notación 4.23. Si V tiene dimensión finita, a la adjunta de f la denotaremos por f^* .

Observación 4.24. En particular $(f^*)^* = f$. Note que la notación de la adjunta es la misma que la notación para transformación dual. Mientras que la adjunta sigue siendo un operador de V , el dual de un operador es un operador en V^* . Confundir las dos notaciones tiene su fundamento en la siguiente propiedad.

Propiedad 4.25. *El mapa*

$$\begin{aligned} \iota : V &\longrightarrow V^* \\ v &\longmapsto \iota_v = \langle \bullet; v \rangle : v' \mapsto \langle v'; v \rangle \end{aligned}$$

es una transformación lineal inyectiva, la cual es un isomorfismo si V tiene dimensión finita.

Dem. Por la linealidad en el segundo factor del producto interno, ι es una transformación lineal. Suponga ahora que $\iota_v = 0$, en particular $0 = \iota_v(v) = \langle v, v \rangle = \|v\|^2$, luego $v = 0$ y así ι es inyectiva. Finalmente como $\dim(V) = \dim(V^*)$ cuando V tiene dimensión finita, entonces ι en este caso también es sobreyectiva, y es un isomorfismo.

Proposición 4.26. *Suponga que V tiene dimensión finita y sea*

$$\begin{aligned} \hat{\bullet} : V &\longrightarrow (V^*)^* \\ v &\longmapsto \hat{v} : \lambda \mapsto \lambda(v). \end{aligned}$$

el isomorfismo canónico. Entonces para todo $v \in V$

$$\iota^*(\hat{v}) = \iota(v).$$

Dem. Para todo $v' \in V$

$$\begin{aligned} \iota^*(\hat{v})(v') &= \hat{v}(\iota(v')) \\ &= \iota(v')(v) \\ &= \langle v; v' \rangle \\ &= \langle v'; v \rangle \\ &= \iota(v)(v'). \end{aligned}$$

□

Observación 4.27. Note que si V tiene dimensión finita, para todo $v' \in V$,

$$\begin{aligned} f^*(\iota_v)(v') &= \iota_v(f(v')) \\ &= \langle f(v'); v \rangle \\ &= \langle v'; f^*(v) \rangle \\ &= \iota_{f^*(v)}(v') \end{aligned}$$

luego $f^*(\iota_v) = \iota_{f^*(v)}$, es decir

$$f^* \circ \iota = \iota \circ f^*,$$

lo cual justifica la confusión entre las dos notaciones de dual y adjunto (en la última igualdad, f^* a la izquierda es el dual de f , mientras que a la derecha es el adjunto); pues, a través del isomorfismo ι , ambos conceptos coinciden.

Observación 4.28. Suponga que V tiene dimensión finita y sean $\mathcal{B} = \{v_1, \dots, v_m\}$ una base de V y $\mathcal{B}^* = \{\lambda_1, \dots, \lambda_m\}$ la base de V^* dual de \mathcal{B} . Tomamos la imagen de \mathcal{B} mediante el isomorfismo canónico $V \mapsto (V^*)^*$, la cual es la base $\widehat{\mathcal{B}} = \{\widehat{v}_1, \dots, \widehat{v}_m\}$ de $(V^*)^*$ dual de \mathcal{B}^* . La proposición anterior implica que si tomamos las representaciones matriciales en $M_{m \times m}(\mathbb{R})$

$$A = [\iota]_{\mathcal{B}}^{\mathcal{B}^*}, \text{ y } B = [\iota^*(\widehat{\bullet})]_{\mathcal{B}}^{\mathcal{B}^*} = [\iota^*]_{\widehat{\mathcal{B}}}^{\mathcal{B}^*} [\widehat{\bullet}]_{\mathcal{B}}^{\widehat{\mathcal{B}}} = [\iota^*]_{\widehat{\mathcal{B}}}^{\mathcal{B}^*},$$

entonces $B = A$, pero por otro lado $B = A^\top$, luego $A^\top = A$. Es decir, la representación matricial de ι respecto a una base y su dual es simétrica.

Observación 4.29. Note que si V tiene dimensión finita, $f^* \circ f$ es auto-adjunta, de hecho para todo $v_1, v_2 \in V$

$$\langle f^* \circ f(v_1); v_2 \rangle = \langle f(v_1); f(v_2) \rangle = \langle v_1; f^* \circ f(v_2) \rangle$$

Proposición 4.30. Si V tiene dimensión finita, las siguientes dos propiedades son equivalentes:

1. f es auto-adjunta; y,
2. la representación matricial de f respecto a toda base ortonormal es simétrica.

Dem. Proposición 4.22 implica que si f es auto-adjunta, su representación matricial respecto a una base ortogonal es simétrica. Para obtener el converso, tomamos una base ortonormal de V , $\mathcal{B} = \{u_1, \dots, u_n\}$ y asumimos que $[f]_{\mathcal{B}}^{\mathcal{B}}$ es simétrica, es decir para todo $i, j \in \{1, \dots, n\}$

$$\langle f(u_j); u_i \rangle = [f]_{\mathcal{B},(i,j)}^{\mathcal{B}} = [f]_{\mathcal{B},(j,i)}^{\mathcal{B}} = \langle f(u_i); u_j \rangle,$$

luego

$$\langle f(u_j); u_i \rangle = \langle f(u_i); u_j \rangle = \langle u_j; f(u_i) \rangle$$

lo cual, por bilinearidad del producto interno, implica que f es auto-adjunta. \square

Observación 4.31. Nos disponemos ahora a estudiar la descomposición de Jordan-Chevalley de los operadores auto-adjuntos. El ingrediente fundamental será establecer que las partes diagonalizable y nilpotente son en este caso también auto-adjuntos.

Lema 4.32. *Suponga que f es auto-adjunto, entonces*

1. *Para todo $P(t) \in \mathbb{R}[t]$, $P(f)$ es auto-adjunto;*
2. *si f es nilpotente, $f = 0$; y,*
3. *si $v_1, v_2 \in V$ son vectores propios asociados a valores propios distintos, entonces $\langle v_1, v_2 \rangle = 0$.*

Dem.

1. Note primero que para todo $v_1, v_2 \in V$, recursivamente establecemos que para $k \in \mathbb{Z}_{\geq 0}$,

$$\langle f^i(v_1); v_2 \rangle = \langle v_1; f^i(v_2) \rangle.$$

Ahora, si $P(t) = \sum_{k=0}^n a_k t^k$ entonces para todo $v_1, v_2 \in V$

$$\begin{aligned} \langle P(f)(v_1); v_2 \rangle &= \left\langle \sum_{k=0}^n a_k f^k(v_1); v_2 \right\rangle \\ &= \sum_{k=0}^n a_k \langle f^k(v_1); v_2 \rangle \\ &= \sum_{k=0}^n a_k \langle v_1; f^k(v_2) \rangle \\ &= \left\langle v_1; \sum_{k=0}^n a_k f^k(v_2) \right\rangle \\ &= \langle v_1; P(f)(v_2) \rangle. \end{aligned}$$

2. Sea $r \in \mathbb{Z}_{>0}$ el grado de nilpotencia de f . Asuma por contradicción que $r > 1$, luego $r \geq 2$, y existe $v \in V$ tal que $f^{r-1}(v) \neq 0$; pero en tal caso

$$\|f^{r-1}(v)\|^2 = \langle f^{r-1}(v); f^{r-1}(v) \rangle = \langle f^r(v); f^{r-2}(v) \rangle = \langle 0; f^{r-2}(v) \rangle = 0$$

luego $f^{r-1}(v) = 0$, lo cual contradice la elección de v . Luego $r = 1$ y así $f = 0$.

3. Sean $\lambda_1, \lambda_2 \in \mathbb{R}$, $\lambda_1 - \lambda_2 \neq 0$, tales que $f(v_1) = \lambda_1 v_1$ y $f(v_2) = \lambda_2 v_2$. Así

$$\lambda_1 \langle v_1; v_2 \rangle = \langle \lambda_1 v_1; v_2 \rangle = \langle f(v_1); v_2 \rangle = \langle v_1; f(v_2) \rangle = \langle v_1; \lambda_2 v_2 \rangle = \lambda_2 \langle v_1; v_2 \rangle,$$

luego

$$(\lambda_1 - \lambda_2) \langle v_1; v_2 \rangle = 0,$$

pero como $\lambda_1 - \lambda_2 \neq 0$, $\langle v_1; v_2 \rangle = 0$. \square

Teorema 4.33 (Teorema Espectral). *Suponga que V tiene dimensión finita, f es auto-adjunta y que*

$$P_f(t) = (t - \lambda_1)^{m_1} (t - \lambda_2)^{m_2} \dots (t - \lambda_r)^{m_r}, \quad \lambda_1, \lambda_2, \dots, \lambda_r \in \mathbb{R}.$$

entonces existe una base ortonormal $\mathcal{B} = \{u_1, \dots, u_n\}$ de V tal que $\left[f\right]_{\mathcal{B}}^{\mathcal{B}}$ es diagonal.

Dem. Por Teorema 2.65 existen $P_D(t), P_N(t) \in \mathbb{R}[t]$, tales que si $f_D = P_D(f)$ y $f_N = P_N(f)$ entonces $f = f_D + f_N$ es la descomposición de Jordan-Chevalley, es decir f_D es diagonalizable y f_N nilpotente y estas conmutan. Ahora, por el lema, f_N es auto-adjunta y así, como es nilpotente, $f_N = 0$. Luego $f = f_D$ es diagonalizable. Para $i = 1, \dots, r$, denote V_i el espacio generado por los vectores propios de f asociados a λ_i , es decir

$$V_i = \{v \in V \mid f(v) = \lambda_i v\},$$

de forma que, como f es diagonalizable,

$$V = V_1 \oplus \dots \oplus V_r.$$

Por el lema también sabemos que si $v_i \in V_i$ y $v_j \in V_j$, $i \neq j$, tenemos $\langle v_i; v_j \rangle = 0$, luego si $\mathcal{B}_1, \dots, \mathcal{B}_r$ son respectivamente bases ortonormales de V_1, \dots, V_r ,

$$\mathcal{B} = \mathcal{B}_1 \cup \dots \cup \mathcal{B}_r$$

es una base ortonormal de V formada por vectores propios de f , en particular, $\left[f\right]_{\mathcal{B}}^{\mathcal{B}}$ es diagonal. \square

Observación 4.34. Más adelante, cuando estudiemos la versión compleja del teorema espectral, veremos que el polinomio característico de un operador auto-adjunto sobre un espacio euclídeo de dimensión finita siempre se puede factorizar en factores lineales en $\mathbb{R}[t]$, lo cual a su vez implicará que estos operadores son siempre diagonalizables mediante una base ortonormal. Estudiemos ahora con más detalle este tipo bases.

4.3. Operadores ortogonales

Sea V un espacio euclídeo y $f \in \text{Hom}_{\mathbb{R}}(V, V)$ un operador.

Definición 4.35. Decimos que f es un *operador ortogonal* si para todo $v_1, v_2 \in V$

$$\langle f(v_1); f(v_2) \rangle = \langle v_1; v_2 \rangle.$$

Observación 4.36. Tenemos

$$\begin{aligned} \|v_1 + v_2\|^2 &= \langle v_1 + v_2; v_1 + v_2 \rangle \\ &= \langle v_1; v_1 \rangle + \langle v_2; v_1 \rangle + \langle v_1; v_2 \rangle + \langle v_2; v_2 \rangle \\ &= \|v_1\|^2 + 2\langle v_1; v_2 \rangle + \|v_2\|^2, \end{aligned}$$

de forma que el producto interno se puede expresar en términos de la norma:

$$\langle v_1; v_2 \rangle = \frac{\|v_1 + v_2\|^2 - (\|v_1\|^2 + \|v_2\|^2)}{2}.$$

De esto podemos concluir que f es ortogonal si y solo f preserva la norma, es decir $\|f(v)\| = \|v\|$ para todo $v \in V$.

Proposición 4.37. Si V tiene dimensión finita, las siguientes propiedades son equivalentes

1. f es ortogonal;
2. f preserva la norma;
3. $f^* \circ f = \text{id}_V$;
4. la imagen por f de una base ortonormal es una base ortonormal.

Dem. La observación muestra la equivalencia entre las dos primeras propiedades. Veamos ahora la equivalencia entre la primera y la tercera. Suponga primero que f es ortogonal y sea $\mathcal{B} = \{u_1, \dots, u_n\}$ una base ortonormal de V . Para todo $v \in V$

$$\begin{aligned} f^* \circ f(v) &= \sum_{i=1}^n \langle f^* \circ f(v); u_i \rangle u_i \\ &= \sum_{i=1}^n \langle f(v); f(u_i) \rangle u_i \\ &= \sum_{i=1}^n \langle v; u_i \rangle u_i \\ &= v. \end{aligned}$$

Suponga ahora que $f^* \circ f = \text{id}_V$, entonces para todo $v_1, v_2 \in V$

$$\langle f(v_1); f(v_2) \rangle = \langle f^* \circ f(v_1); v_2 \rangle = \langle v_1; v_2 \rangle.$$

Finalmente establecemos la equivalencia entre la primera propiedad y la cuarta. Si f es ortogonal y $\mathcal{B} = \{u_1, \dots, u_n\}$ entonces

$$\langle f(u_i); f(u_j) \rangle = \langle u_i; u_j \rangle = \delta_{ij}$$

para todo $i, j \in \{1, \dots, n\}$, así $f(\mathcal{B})$ es una base ortonormal (ver Observación 4.8). Recíprocamente, asuma que f envía una base ortonormal $\mathcal{B} = \{u_1, \dots, u_n\}$, a la base ortonormal $f(\mathcal{B}) = \{f(u_1), \dots, f(u_n)\}$. Luego, para todo $v_1, v_2 \in V$, si $x_1, \dots, x_n, y_1, \dots, y_n \in \mathbb{R}$ son tales que

$$v_1 = \sum_{i=1}^n x_i u_i \quad v_2 = \sum_{i=1}^n y_i u_i,$$

entonces

$$f(v_1) = \sum_{i=1}^n x_i f(u_i) \quad v_2 = \sum_{i=1}^n y_i f(u_i)$$

y, por Propiedad 4.11,

$$\langle f(v_1); f(v_2) \rangle = \sum_{i=1}^n x_i y_i = \langle v_1; v_2 \rangle.$$

□

Observación 4.38. Note que implícitamente la tercera propiedad está diciendo que los operadores ortogonales sobre espacios euclídeos de dimensión finita son invertibles, pues su inversa es su adjunta; y la cuarta que esta inversa es también ortogonal, pues también envía una base ortogonal en una base ortogonal.

Definición 4.39. Decimos que $A \in M_{n \times n}(\mathbb{R})$ es una *matriz ortogonal* si

$$A^\top A = I_n$$

donde I_n denota la matriz con unos en la diagonal y ceros en el resto de entradas.

Proposición 4.40. Si V tiene dimensión finita y $\mathcal{B} = \{u_1, \dots, u_n\}$ es una base ortonormal de V , f es ortogonal si y solo si $\begin{bmatrix} f \end{bmatrix}_{\mathcal{B}}^{\mathcal{B}}$ es ortogonal.

Dem. La proposición se sigue del hecho que si $A = \begin{bmatrix} f \end{bmatrix}_{\mathcal{B}}^{\mathcal{B}}$, entonces $A^\top = \begin{bmatrix} f^* \end{bmatrix}_{\mathcal{B}}^{\mathcal{B}}$ y

$$A^\top A = \begin{bmatrix} f^* \end{bmatrix}_{\mathcal{B}}^{\mathcal{B}} \begin{bmatrix} f \end{bmatrix}_{\mathcal{B}}^{\mathcal{B}} = \begin{bmatrix} f^* \circ f \end{bmatrix}_{\mathcal{B}}^{\mathcal{B}}.$$

Entonces $A^\top A = I_n$ si y solo si $f^* \circ f = \text{id}_V$. □

Teorema 4.41. Si V tiene dimensión finita igual a n , la colección de operadores ortogonales de V está en correspondencia biyectiva con la colección de n -tuplas $(v_1, \dots, v_n) \in V \times \dots \times V$ tales que $\{v_1, \dots, v_n\}$ es una base ortonormal de V .

Dem. Sea $\mathcal{B} = \{u_1, \dots, u_n\}$ una base ortonormal de V . Dado un operador ortogonal $g \in \text{Hom}_{\mathbb{R}}(V, V)$, le asociamos la n -tupla

$$(g(u_1), \dots, g(u_n)).$$

Dada la n -tupla $(v_1, \dots, v_n) \in V \times \dots \times V$ tal que $\{v_1, \dots, v_n\}$ es una base ortonormal de V , le asociamos el operador definido sobre la base \mathcal{B} por

$$u_1 \mapsto v_1, \dots, u_n \mapsto v_n.$$

Las equivalencias en Proposición 4.37 implican que $(g(u_1), \dots, g(u_n))$ es una n -tupla cuyas componentes forman una base ortonormal de V y que el operador tal que $u_1 \mapsto v_1, \dots, u_n \mapsto v_n$ es ortogonal. Las dos asociaciones son una la inversa de la otra. \square

Observación 4.42. Note que la correspondencia descrita en la prueba del teorema depende de la escogencia de la base $\mathcal{B} = \{u_1, \dots, u_n\}$ y del ordenamiento de los elementos que la conforman. Vale la pena subrayar el hecho que los operadores ortogonales se pueden componer entre si y obtener un tercer operador ortogonal, y también invertir y obtener otro operador ortogonal. Conjuntos con estas propiedades se les llama grupos. Una vez se fija una base, junto con un ordenamiento de sus elementos, la correspondencia del teorema respeta la estructura de grupo.

Formalmente, si G denota la colección de operadores ortogonales de V , X la de n -tuplas cuyas componentes formán bases ortogonales y

$$\begin{aligned} \Phi_T : G &\longrightarrow X \\ g &\longmapsto (g(u_1), \dots, g(u_n)) \end{aligned}$$

es la correspondencia biyectiva definida por T (junto con un ordenamiento), entonces

$$\begin{aligned} \Phi_T(\text{id}_V) &= (u_1, \dots, u_n) \\ \Phi_T(g \circ h) &= g(\Phi_T(h)) \end{aligned}$$

para todo $g, h \in G$, donde definimos $g(v_1, \dots, v_n) = (g(v_1), \dots, g(v_n))$ para todo $(v_1, \dots, v_n) \in X$. Más general

$$\begin{aligned} \Phi : X \times G &\longmapsto X \times X \\ (x, g) &\longmapsto (x, gx) \end{aligned}$$

es una biyección, en la cual, si fijamos el primer factor en $x = (u_1, \dots, u_n)$, obtenemos en el segundo la biyección ϕ_T .

Capítulo 5

Espacios unitarios

Sea V un espacio vectorial sobre \mathbb{C} .

Notación 5.1. Dado un escalar $c \in \mathbb{C}$, denotamos por \bar{c} a su conjugado el cual está definido por

$$\bar{c} = a - bi, \text{ si } c = a + bi, \ a, b \in \mathbb{R},$$

por $|c| = \sqrt{a^2 + b^2} = \sqrt{c\bar{c}}$ a su norma, por $\operatorname{Re}(c) = a = (c + \bar{c})/2$ su parte real y por $\operatorname{Im}(c) = b = (c - \bar{c})/2i$ su parte imaginaria.

Observación 5.2. Bastantes elementos elaborados en este capítulo se establecen por argumentos similares a los expuestos durante el desarrollo de la teoría de espacios euclídeos. En tales casos, dejaremos la verificación de los detalles al lector.

5.1. Producto hermítico

Definición 5.3. Un *producto hermítico* en V es una función

$$\begin{aligned} \langle \bullet, \bullet \rangle : V \times V &\longrightarrow \mathbb{C} \\ (v_1, v_2) &\longmapsto \langle v_1; v_2 \rangle \end{aligned}$$

tal que:

1. *es sesquilineal*: para todo $v, v_1, v_2 \in V$ y $c \in \mathbb{C}$

$$\begin{aligned} \langle v_1 + v_2; v \rangle &= \langle v_1; v \rangle + \langle v_2; v \rangle \\ \langle cv_1; v_2 \rangle &= c\langle v_1; v_2 \rangle \\ \langle v; v_1 + v_2 \rangle &= \langle v; v_1 \rangle + \langle v; v_2 \rangle \end{aligned}$$

2. *es hermítica*: para todo $v_1, v_2 \in V$

$$\langle v_2; v_1 \rangle = \overline{\langle v_1; v_2 \rangle};$$

3. *es definitivamente positiva* para todo $v \in V$, $v \neq 0$,

$$\langle v; v \rangle > 0.$$

Un *espacio unitario* es un espacio vectorial sobre \mathbb{C} provisto de un producto hermítico.

Observación 5.4. Se sigue que $\langle v; v \rangle = 0$ si y solo si $v = 0$ y que para todo $v, v_1, v_2 \in V$ y $c \in \mathbb{C}$

$$\langle v_1; cv_2 \rangle = \bar{c} \langle v_1; v_2 \rangle$$

Ejemplo 5.5. 1. Sobre $V = \mathbb{C}^n$,

$$\langle (z_1, \dots, z_n); (w_1, \dots, w_n) \rangle = \sum_{i=1}^n z_i \bar{w}_i.$$

2. Sobre $V = M_{n \times n}(\mathbb{C})$,

$$\langle A; B \rangle = \text{tr}(B^* A).$$

donde $B^* = \bar{B}^T$ es la matriz cuyas entradas son las conjugadas de las entradas de la matriz traspuesta de B .

3. Sea $[a, b] \subseteq \mathbb{R}$ un intervalo cerrado. Sobre $V = C_{\mathbb{C}}^0[a, b]$, el conjunto de funciones continuas $[a, b] \rightarrow \mathbb{C}$,

$$\langle f; g \rangle = \int_a^b f(x) \overline{g(x)} dx.$$

Definición 5.6. Dado un espacio hermítico V , definimos la norma de $v \in V$ por $\|v\| = \sqrt{\langle v; v \rangle}$.

Propiedad 5.7. Sea V un espacio unitario, entonces:

1. $\|cv\| = |c| \|v\|$, para todo $c \in \mathbb{C}$ y $v \in V$;
2. Desigualdad de Cauchy-Schwarz: $|\langle v_1; v_2 \rangle| \leq \|v_1\| \|v_2\|$, para todo $v_1, v_2 \in V$, más aún se tiene $\langle v_1; v_2 \rangle = \|v_1\| \|v_2\|$ únicamente cuando $\{v_1, v_2\}$ es linealmente dependiente; y,
3. Desigualdad triangular: $\|v_1 + v_2\| \leq \|v_1\| + \|v_2\|$, para todo $v_1, v_2 \in V$, más aún se tiene $\|v_1 + v_2\| \leq \|v_1\| + \|v_2\|$ si y solo si $av_1 = bv_2$ con $a, b \geq 0$.

Dem.

1. Dados $c \in \mathbb{C}$ y $v \in V$

$$\|cv\| = \sqrt{\langle cv; cv \rangle} = \sqrt{c\bar{c} \langle v; v \rangle} = |c| \sqrt{\langle v; v \rangle} = |c| \|v\|.$$

2. Si $v_1 = 0$ o $v_2 = 0$, se tiene $0 = \langle v_1; v_2 \rangle = \|v_1\| \|v_2\|$ y se sigue la desigualdad. En general, para todo $a, b \in \mathbb{C}$,

$$\begin{aligned} 0 \leq \|av_1 - bv_2\|^2 &= \langle av_1 - bv_2, av_1 - bv_2 \rangle \\ &= a\bar{a}\langle v_1; v_1 \rangle - b\bar{a}\langle v_2; v_1 \rangle - a\bar{b}\langle v_1; v_2 \rangle + b\bar{b}\langle v_2; v_2 \rangle \\ &= |a|^2\|v_1\|^2 - \left(\bar{a}b\overline{\langle v_1; v_2 \rangle} + a\bar{b}\langle v_1; v_2 \rangle \right) + |b|^2\|v_2\|^2. \end{aligned}$$

En particular si $a = \|v_2\|^2$ y $b = \langle v_1; v_2 \rangle$, obtenemos

$$0 \leq \|v_2\|^4\|v_1\|^2 - \|v_2\|^2|\langle v_1, v_2 \rangle|^2,$$

lo cual, si suponemos que $v_2 \neq 0$ (e.d. $|v_2|^2 \neq 0$), implica la desigualdad deseada. Ahora bien, remontando las igualdades, $|\langle v_1, v_2 \rangle| = \|v_1\| \|v_2\|$ equivale a $\|av_1 - bv_2\|^2 = 0$, donde $a = \|v_2\|^2$ y $b = \langle v_1; v_2 \rangle$.

3. Tomando $a = 1$ y $b = -1$, obtenemos

$$\begin{aligned} \|v_1 + v_2\|^2 &= \|v_1\|^2 + \left(\overline{\langle v_1; v_2 \rangle} + \langle v_1; v_2 \rangle \right) + \|v_2\|^2 \\ &= \|v_1\|^2 + 2\operatorname{Re}(\langle v_1; v_2 \rangle) + \|v_2\|^2 \\ &\leq \|v_1\|^2 + 2|\langle v_1; v_2 \rangle| + \|v_2\|^2 \end{aligned}$$

La desigualdad de Cauchy-Schwarz implica

$$\|v_1 + v_2\|^2 \leq \|v_1\|^2 + 2\|v_1\|\|v_2\| + \|v_2\|^2 = (\|v_1\| + \|v_2\|)^2$$

que es equivalente a la desigualdad afirmada. La igualdad se obtiene si y solo si $\operatorname{Re}(\langle v_1, v_2 \rangle) = \langle v_1, v_2 \rangle = \|v_1\|\|v_2\|$, lo cual equivale a $av_1 - bv_2 = 0$ donde $a = \|v_2\|$ y $b = \|v_1\|$ pues

$$\|av_1 - bv_2\|^2 = |a|^2\|v_1\|^2 - 2\operatorname{Re}(a\bar{b}\langle v_1; v_2 \rangle) + |b|^2\|v_2\|^2.$$

□

Definición 5.8. Sea V un espacio unitario y $S \subseteq V$. Decimos que S es *ortogonal* si $\langle v_1, v_2 \rangle = 0$ para todo $v_1, v_2 \in S$ tales que $v_1 \neq v_2$. Si además $\|v\| = 1$ para todo $v \in S$, decimos que S es *ortonormal*.

Observación 5.9. Note que $S = \{v_i\}_{i \in I}$ es ortonormal si y solo si

$$\langle v_i; v_j \rangle = \delta_{ij}.$$

para todo $i, j \in I$.

Propiedad 5.10. Sea V un espacio unitario. Si $S \subset V$ es ortogonal, y $0 \notin S$, entonces S es linealmente independiente.

Dem. El argumento es similar al de la demostración de Propiedad 4.9.

□

Teorema 5.11 (Ortogonalización de Gram-Schmidt). *Sea V un espacio unitario. Suponga que V tiene dimensión finita, entonces V tiene una base ortonormal. Más aún, si $\{v_1, \dots, v_n\}$ es una base de V , existe una base ortonormal $\{u_1, \dots, u_n\}$ de V tal que para $k = 1, \dots, n$*

$$\langle v_1, \dots, v_k \rangle = \langle u_1, \dots, u_k \rangle.$$

Dem. El argumento es similar al de la demostración de Teorema 4.10. \square

Propiedad 5.12. *Sea V un espacio unitario. Suponga que V tiene dimensión finita y $\{u_1, \dots, u_n\}$ es una base ortonormal de V , entonces para todo $v \in V$*

$$v = \sum_{i=1}^n \langle v; u_i \rangle u_i.$$

En particular, si $v_1, v_2 \in V$ son tales que

$$v_1 = \sum_{i=1}^n z_i u_i \quad v_2 = \sum_{i=1}^n w_i u_i,$$

entonces

$$\langle v_1; v_2 \rangle = \sum_{i=1}^n z_i \overline{w_i}.$$

Dem. Como $\{u_1, \dots, u_n\}$ es base, existen $a_1, \dots, a_n \in K$ tal que $v = \sum_{i=1}^n a_i u_i$. De esta forma, para $j = 1, \dots, n$,

$$\langle v; u_j \rangle = \sum_{i=1}^n a_i \langle u_i, u_j \rangle = \sum_{i=1}^n a_i \delta_{ij} = a_j$$

(ver Observación 5.9). Finalmente,

$$\langle v_1; v_2 \rangle = \left\langle \sum_{i=1}^n z_i u_i, \sum_{j=1}^n w_j u_j \right\rangle = \sum_{i=1}^n \sum_{j=1}^n z_i \overline{w_j} \delta_{ij} = \sum_{i=1}^n z_i \overline{w_i}.$$

\square

Definición 5.13. Sean V un espacio unitario y $S \subseteq V$, el *conjunto ortogonal* a S está definido por

$$S^\perp = \{v \in V \mid \langle v, u \rangle = 0, \text{ para todo } u \in S\}$$

Propiedad 5.14. *Sean V un espacio unitario y $S \subseteq V$, entonces $S^\perp \leq V$.*

Dem. El argumento es similar al de la demostración de Propiedad 4.14. \square

Teorema 5.15. *Sea V un espacio unitario. Suponga que V tiene dimensión finita y sea $U \leq V$, entonces*

$$V = U \oplus U^\perp$$

Dem. El argumento es similar al de la demostración de Teorema 4.15. \square

Definición 5.16. Sea V un espacio unitario. Suponga que V tiene dimensión finita y sea $U \leq V$. Llamamos a U^\perp el *complemento ortogonal de U* . A la proyección

$$p_U^\perp : V \longrightarrow V$$

sobre U , definida por la descomposición $V = U \oplus U^\perp$ la llamamos *proyección ortogonal sobre U* .

Propiedad 5.17. Sea V un espacio unitario. Suponga que V tiene dimensión finita y sea $U \leq V$. Tenemos para todo $v_1, v_2 \in V$

$$\langle p_U^\perp(v_1); v_2 \rangle = \langle v_1; p_U^\perp(v_2) \rangle$$

Si $\{u_1, \dots, u_m\}$ es una base ortonormal de U entonces para todo $v \in V$

$$p_U^\perp(v) = \sum_{i=1}^m \langle v; u_i \rangle u_i.$$

Dem. El argumento es similar al de la demostración de Propiedad 4.18. \square

5.2. Operador adjunto

Sea V un espacio unitario y $f \in \text{Hom}_{\mathbb{C}}(V, V)$ un operador.

Definición 5.18. Sea $g \in \text{Hom}_{\mathbb{C}}(V, V)$, decimos que g es un *operador adjunto de f* si para todo $v_1, v_2 \in V$

$$\langle g(v_1); v_2 \rangle = \langle v_1; f(v_2) \rangle.$$

Decimos que f es *auto-adjunto* si f es un operador adjunto de f .

Observación 5.19. Note que si g es adjunto de f , entonces f es adjunto de g . De hecho

$$\langle f(v_1); v_2 \rangle = \overline{\langle v_2; f(v_1) \rangle} = \overline{\langle g(v_2); v_1 \rangle} = \langle v_1; g(v_2) \rangle$$

Definición 5.20. Sean I, J conjuntos y $A \in M_{I \times J}(\mathbb{C})$, definimos la *matriz adjunta* de A por $A^* \in M_{J \times I}(\mathbb{C})$ tal que

$$A^*(j, i) = \overline{A(i, j)}$$

para todo $(j, i) \in J \times I$. Es decir el valor en (j, i) de A^* es el conjugado del valor en (i, j) de A . Similarmente si $m, n \in \mathbb{Z}_{>0}$, y $A \in M_{m \times n}(\mathbb{C})$, definimos su adjunta por $A^* \in M_{n \times m}(\mathbb{C})$ tal que

$$A^*(j, i) = \overline{A(i, j)}$$

Sea $A \in M_{I \times I}(K)$, o $A \in M_{n \times n}(K)$, decimos que A es *hermítica* si $A^* = A$.

Proposición 5.21. *Suponga que V tiene dimensión finita, entonces existe un único operador $g \in \text{Hom}_{\mathbb{C}}(V, V)$ adjunto de f . Más aún, si $\mathcal{B} = \{u_1, \dots, u_n\}$ es una base ortonormal de V , entonces*

$$[g]_{\mathcal{B}}^{\mathcal{B}} = \left([f]_{\mathcal{B}}^{\mathcal{B}} \right)^*$$

Dem. Defina el operador $g \in \text{Hom}_{\mathbb{C}}(V, V)$ por la imagen de la base \mathcal{B} :

$$g(u_j) = \sum_{i=1}^n \langle u_j; f(u_i) \rangle u_i.$$

De esta forma

$$\langle g(u_j); u_i \rangle = \langle u_j; f(u_i) \rangle$$

y por Propiedad 5.12

$$\begin{aligned} \langle g(v_1); v_2 \rangle &= \sum_{i=1}^n \langle g(v_1); u_i \rangle \overline{\langle v_2; u_i \rangle} \\ &= \sum_{i,j=1}^n \langle v_1; u_j \rangle \langle g(u_j); u_i \rangle \overline{\langle v_2; u_i \rangle} \\ &= \sum_{i,j=1}^n \langle v_1; u_j \rangle \langle u_j; f(u_i) \rangle \overline{\langle v_2; u_i \rangle} \\ &= \sum_{j=1}^n \langle v_1; u_j \rangle \langle u_j; f(v_2) \rangle \\ &= \sum_{j=1}^n \langle v_1; u_j \rangle \overline{\langle f(v_2); u_j \rangle} \\ &= \langle v_1; f(v_2) \rangle \end{aligned}$$

Por otro lado si, $h \in \text{Hom}_{\mathbb{C}}(V, V)$ es adjunto de f , por Propiedad 5.12,

$$\begin{aligned} h(u_j) &= \sum_{i=1}^n \langle h(u_j); u_i \rangle u_i \\ &= \sum_{i=1}^n \langle u_j; f(u_i) \rangle u_i \\ &= g(u_j), \end{aligned}$$

luego $h = g$.

Ahora, para ver que la representación matricial de g respecto a \mathcal{B} es la adjunta

de la de f basta observar que

$$\begin{aligned}
 \left[g \right]_{\mathcal{B},(j,i)}^{\mathcal{B}} &= \left[g(u_i) \right]_j^{\mathcal{B}} \\
 &= \langle g(u_i); u_j \rangle \\
 &= \langle u_i; f(u_j) \rangle \\
 &= \overline{\langle f(u_j); u_i \rangle} \\
 &= \overline{\left[f(u_j) \right]_i^{\mathcal{B}}} \\
 &= \overline{\left[f \right]_{\mathcal{B},(i,j)}^T}
 \end{aligned}$$

□

Notación 5.22. Si V tiene dimensión finita, a la adjunta de f la denotaremos por f^* .

Propiedad 5.23. *El mapa*

$$\begin{aligned}
 h : V &\longrightarrow V^* \\
 v &\longmapsto h_v = \langle \bullet; v \rangle : v' \mapsto \langle v'; v \rangle
 \end{aligned}$$

es semilineal, es decir para todo $v_1, v_2, v \in V$ y $c \in \mathbb{C}$

$$\begin{aligned}
 h(v_1 + v_2) &= h(v_1) + h(v_2) \\
 h(cv) &= \bar{c}h(v),
 \end{aligned}$$

e inyectivo. Si además V tiene dimensión finita, h es biyectivo.

Dem. Para todo $v' \in V$, dados $v_1, v_2, v \in V$ y $c \in \mathbb{C}$

$$\begin{aligned}
 h(v_1 + v_2)(v') &= \langle v'; v_1 + v_2 \rangle \\
 &= \langle v'; v_1 \rangle + \langle v'; v_2 \rangle \\
 &= h(v_1)(v') + h(v_2)(v') \\
 &= (h(v_1) + h(v_2))(v'), \text{ y} \\
 h(cv)(v') &= \langle v'; cv \rangle \\
 &= \bar{c}\langle v'; v \rangle \\
 &= \bar{c}h(v)(v')
 \end{aligned}$$

Suponga ahora que $v_1, v_2 \in V$ son tales $h(v_1) = h(v_2)$, es decir $h(v_1 - v_2) = 0$, en particular $0 = h(v_1 - v_2)(v_1 - v_2) = \langle v_1 - v_2; v_1 - v_2 \rangle = \|v_1 - v_2\|^2$, luego $v_1 - v_2 = 0$ y así $v_1 = v_2$, es decir h es inyectiva. Finalmente suponga que V tiene dimensión finita y sea $\mathcal{B} = \{u_1, \dots, u_n\}$ una base de V , entonces $h(\mathcal{B}) = \{h(u_1), \dots, h(u_n)\}$ es la base de V^* dual de V , pues

$$h(u_j)(u_i) = \langle u_i; u_j \rangle = \delta_{ij}.$$

En particular dado $\lambda \in V^*$, si $a_1, \dots, a_n \in \mathbb{C}$ son tales que $\lambda = \sum_{i=1}^n a_i h(u_i)$ entonces $h(\sum_{i=1}^n \bar{a}_i u_i) = \lambda$, luego h es también sobreyectiva y así biyectiva. □

Observación 5.24. Note que si V tiene dimensión finita, para todo $v' \in V$,

$$\begin{aligned} f^*(h_v)(v') &= h_v(f(v')) \\ &= \langle f(v'); v \rangle \\ &= \langle v'; f^*(v) \rangle \\ &= h_{f^*(v)}(v') \end{aligned}$$

luego $f^*(h_v) = h_{f^*(v)}$, es decir

$$f^* \circ h = h \circ f^*,$$

donde, en esta igualdad, f^* a la izquierda es el dual de f , mientras que a la derecha es el adjunto. A través de la biyección semilineal h , ambos conceptos coinciden.

Observación 5.25. Note que si V tiene dimensión finita, $f^* \circ f$ es auto-adjunta, de hecho para todo $v_1, v_2 \in V$

$$\langle f^* \circ f(v_1); v_2 \rangle = \langle f(v_1); f(v_2) \rangle = \langle v_1; f^* \circ f(v_2) \rangle$$

Proposición 5.26. Si V tiene dimensión finita, las siguientes dos propiedades son equivalentes:

1. f es auto-adjunta; y,
2. la representación matricial de f respecto a una base ortonormal es hermítica.

Dem. Proposición 5.21 implica que si f es auto-adjunta, su representación matricial respecto a una base ortogonal es hermítica. Para obtener el converso, tomamos una base ortonormal de V , $\mathcal{B} = \{u_1, \dots, u_n\}$ y asumimos que $\begin{bmatrix} f \end{bmatrix}_{\mathcal{B}}^{\mathcal{B}}$ es hermítica, es decir para todo $i, j \in \{1, \dots, n\}$

$$\langle f(u_j); u_i \rangle = \begin{bmatrix} f \end{bmatrix}_{\mathcal{B},(i,j)}^{\mathcal{B}} = \overline{\begin{bmatrix} f \end{bmatrix}_{\mathcal{B},(j,i)}^{\mathcal{B}}} = \overline{\langle f(u_i); u_j \rangle} = \langle u_j; f(u_i) \rangle,$$

luego si $v_1 = \sum_{i=1}^n z_i u_i$ y $v_2 = \sum_{j=1}^n w_j u_j$

$$\begin{aligned} \langle f(v_1); v_2 \rangle &= \sum_{i,j=1}^n z_i \overline{w_j} \langle f(u_i); u_j \rangle \\ &= \sum_{i,j=1}^n z_i \overline{w_j} \langle u_i; f(u_j) \rangle \\ &= \langle v_1; f(v_2) \rangle \end{aligned}$$

□

Observación 5.27. La descomposición de Jordan-Chevalley de los operadores auto-adjuntos sobre un espacio unitario empata con la descomposición sobre espacios euclideos pues tienen la particularidad de tener todos sus valores propios reales. Esto nos permitirá relajar las hipótesis del teorema espectral ya demostrado.

Lema 5.28. *Suponga que f es auto-adjunto, entonces*

1. *Para todo $P(t) \in \mathbb{R}[t]$, $P(f)$ es auto-adjunto;*
2. *si f es nilpotente, $f = 0$;*
3. *los valores propios de f son reales; y,*
4. *si $v_1, v_2 \in V$ son vectores propios asociados a valores propios distintos, entonces $\langle v_1; v_2 \rangle = 0$.*

Dem.

1. Note primero que para todo $v_1, v_2 \in V$, recursivamente establecemos que para $k \in \mathbb{Z}_{\geq 0}$,

$$\langle f^i(v_1); v_2 \rangle = \langle v_1; f^i(v_2) \rangle.$$

Ahora, si $P(t) = \sum_{k=0}^n a_k t^k$, para $k = 1, \dots, n$ $a_k = \overline{a_k}$, y entonces para todo $v_1, v_2 \in V$

$$\begin{aligned} \langle P(f)(v_1); v_2 \rangle &= \left\langle \sum_{k=0}^n a_k f^k(v_1); v_2 \right\rangle \\ &= \sum_{k=0}^n a_k \langle f^k(v_1); v_2 \rangle \\ &= \sum_{k=0}^n a_k \langle v_1; f^k(v_2) \rangle \\ &= \langle v_1; \sum_{k=0}^n \overline{a_k} f^k(v_2) \rangle \\ &= \langle v_1; P(f)(v_2) \rangle. \end{aligned}$$

2. Sea $r \in \mathbb{Z}_{>0}$ el grado de nilpotencia de f . Asuma por contradicción que $r > 1$, luego $r \geq 2$, y existe $v \in V$ tal que $f^{r-1}(v) \neq 0$; pero en tal caso

$$\|f^{r-1}(v)\|^2 = \langle f^{r-1}(v); f^{r-1}(v) \rangle = \langle f^r(v); f^{r-2}(v) \rangle = \langle 0; f^{r-2}(v) \rangle = 0$$

luego $f^{r-1}(v) = 0$, lo cual contradice la elección de v . Luego $r = 1$ y así $f = 0$.

3. Sean $\lambda \in \mathbb{C}$ y $v \in V$, $v \neq 0$, tales que $f(v) = \lambda v$. Así

$$\lambda \langle v; v \rangle = \langle \lambda v; v \rangle = \langle f(v); v \rangle = \langle v; f(v) \rangle = \langle v; \lambda v \rangle = \overline{\lambda} \langle v; v \rangle,$$

de donde $(\lambda - \overline{\lambda})\|v\|^2 = 0$, pero $v \neq 0$, entonces, $\lambda = \overline{\lambda}$, es decir $\lambda \in \mathbb{R}$.

4. Sean $\lambda_1, \lambda_2 \in \mathbb{R}$, $\lambda_1 - \lambda_2 \neq 0$, tales que $f(v_1) = \lambda_1 v_1$ y $f(v_2) = \lambda_2 v_2$. Así
- $$\lambda_1 \langle v_1; v_2 \rangle = \langle \lambda_1 v_1; v_2 \rangle = \langle f(v_1); v_2 \rangle = \langle v_1; f(v_2) \rangle = \langle v_1; \lambda_2 v_2 \rangle = \lambda_2 \langle v_1; v_2 \rangle,$$
- luego

$$(\lambda_1 - \lambda_2) \langle v_1; v_2 \rangle = 0,$$

pero como $\lambda_1 - \lambda_2 \neq 0$, $\langle v_1; v_2 \rangle = 0$. \square

Teorema 5.29 (Teorema Espectral). *Suponga que V tiene dimensión finita y f es auto-adjunta entonces existe una base ortonormal $\mathcal{B} = \{u_1, \dots, u_n\}$ de V tal que $\left[f \right]_{\mathcal{B}}^{\mathcal{B}}$ es diagonal.*

Dem. Como \mathbb{C} es algebraicamente cerrado,

$$P_f(t) = (t - \lambda_1)^{m_1} (t - \lambda_2)^{m_2} \dots (t - \lambda_r)^{m_r}, \quad \lambda_1, \lambda_2, \dots, \lambda_r \in \mathbb{C}.$$

Como $\lambda_1, \lambda_2, \dots, \lambda_r$ son valores propios, por el lema son valores reales y $P_f(t) \in \mathbb{R}[t]$. En particular, como $t - \lambda_1, t - \lambda_2, \dots, t - \lambda_r \in \mathbb{R}[t]$, por la demostración de Teorema 2.65 existen $P_D(t), P_N(t) \in \mathbb{R}[t]$, tales que si $f_D = P_D(f)$ y $f_N = P_N(f)$ entonces $f = f_D + f_N$ es la descomposición de Jordan-Chevalley, es decir f_D es diagonalizable y f_N nilpotente y estas conmutan. Ahora, por el lema, f_N es auto-adjunta y así, como es nilpotente, $f_N = 0$. Luego $f = f_D$ es diagonalizable. Para $i = 1, \dots, r$, denote V_i el espacio generado por los vectores propios de f asociados a λ_i , es decir

$$V_i = \{v \in V \mid f(v) = \lambda_i v\},$$

de forma que, como f es diagonalizable,

$$V = V_1 \oplus \dots \oplus V_r.$$

Por el lema también sabemos que si $v_i \in V_i$ y $v_j \in V_j$, $i \neq j$, tenemos $\langle v_i, v_j \rangle = 0$, luego si $\mathcal{B}_1, \dots, \mathcal{B}_r$ son respectivamente bases ortonormales de V_1, \dots, V_r ,

$$\mathcal{B} = \mathcal{B}_1 \cup \dots \cup \mathcal{B}_r$$

es una base ortonormal de V formada por vectores propios de f , en particular, $\left[f \right]_{\mathcal{B}}^{\mathcal{B}}$ es diagonal. \square

Corolario 5.30. *Sea V un espacio euclídeo de dimensión finita y $f \in \text{Hom}_{\mathbb{R}}(V, V)$ un operador auto-adjunto, entonces una base ortonormal $\mathcal{B} = \{u_1, \dots, u_n\}$ de V tal que $\left[f \right]_{\mathcal{B}}^{\mathcal{B}}$ es diagonal.*

Dem. Sea \mathcal{B} una base ortonormal de V y $A \in M_{n \times n}(\mathbb{R})$ la representación de f respecto a la base \mathcal{B} donde $n = \dim(V)$. Entonces A es una matriz simétrica con entradas reales. Sea

$$\begin{aligned} f_A : \mathbb{C}^n &\longrightarrow \mathbb{C}^n \\ (z_1, \dots, z_n) &\longmapsto \left(\sum_{k=1}^n a_{1k} z_k, \dots, \sum_{k=1}^n a_{nk} z_k \right) \end{aligned}$$

donde la kl -ésima entrada de es $A(k, l) = a_{kl}$. De forma que la representación matricial de f_A respecto a la base canónica de \mathbb{C}^n es A , la cual es hermítica pues es simétrica con entradas reales, luego f_A es auto-adjunta respecto al producto hermítico

$$\langle (z_1, \dots, z_n), (w_1, \dots, w_n) \rangle = \sum_{k=1}^n z_k \overline{w_k}.$$

Así

$$P_f(t) = P_{f_A}(t) = (t - \lambda_1)^{m_1} (t - \lambda_2)^{m_2} \dots (t - \lambda_r)^{m_r}, \quad \lambda_1, \lambda_2, \dots, \lambda_r \in \mathbb{R}.$$

y la conclusión se sigue ahora de Teorema 4.33. \square

5.3. Operadores unitarios

Sea V un espacio unitario y $f \in \text{Hom}_{\mathbb{C}}(V, V)$ un operador.

Definición 5.31. Decimos que f es un *operador unitario* si para todo $v_1, v_2 \in V$

$$\langle f(v_1); f(v_2) \rangle = \langle v_1; v_2 \rangle.$$

Observación 5.32. Tenemos

$$\begin{aligned} \|v_1 + v_2\|^2 &= \langle v_1 + v_2; v_1 + v_2 \rangle \\ &= \langle v_1; v_1 \rangle + \langle v_2; v_1 \rangle + \langle v_1; v_2 \rangle + \langle v_2; v_2 \rangle \\ &= \langle v_1; v_1 \rangle + \overline{\langle v_1; v_2 \rangle} + \langle v_1; v_2 \rangle + \langle v_2; v_2 \rangle \\ &= \|v_1\|^2 + 2 \operatorname{Re}(\langle v_1; v_2 \rangle) + \|v_2\|^2, \text{ y} \\ \|v_1 + iv_2\|^2 &= \langle v_1 + iv_2; v_1 + iv_2 \rangle \\ &= \langle v_1; v_1 \rangle + i\langle v_2; v_1 \rangle - i\langle v_1; v_2 \rangle + \langle v_2; v_2 \rangle \\ &= \langle v_1; v_1 \rangle - i(\langle v_1; v_2 \rangle - \overline{\langle v_1; v_2 \rangle}) + \langle v_2; v_2 \rangle \\ &= \|v_1\|^2 + 2 \operatorname{Im}(\langle v_1; v_2 \rangle) + \|v_2\|^2, \end{aligned}$$

de forma que el producto interno se puede expresar en términos de la norma:

$$\langle v_1; v_2 \rangle = \frac{\|v_1 + v_2\|^2 - (\|v_1\|^2 + \|v_2\|^2)}{2} + \frac{\|v_1 + iv_2\|^2 - (\|v_1\|^2 + \|v_2\|^2)}{2}i.$$

De esto podemos concluir que f es unitario si y solo f preserva la norma, es decir $\|f(v)\| = \|v\|$ para todo $v \in V$.

Proposición 5.33. Si V tiene dimensión finita, las siguiente propiedades son equivalentes

1. f es unitario;
2. f preserva la norma;

3. $f^* \circ f = \text{id}_V$; y,

4. la imagen por f de una base ortonormal es una base ortonormal.

Dem. El argumento es similar al de la demostración de Proposición 4.37 \square

Observación 5.34. Como en el caso de los operadores ortogonales sobre espacios euclídeos, los operadores unitarios son invertibles y envían bases ortogonales en bases ortogonales.

Definición 5.35. Decimos que $A \in M_{n \times n}(\mathbb{C})$ es una *matriz unitaria* si

$$A^* A = I_n$$

donde I_n denota la matriz con unos en la diagonal y ceros en el resto de entradas.

Proposición 5.36. Si V tiene dimensión finita y $\mathcal{B} = \{u_1, \dots, u_n\}$ es una base ortonormal de V , f es unitario si y solo si $\begin{bmatrix} f \end{bmatrix}_{\mathcal{B}}^{\mathcal{B}}$ es unitaria.

Dem. La proposición se sigue del hecho que si $A = \begin{bmatrix} f \end{bmatrix}_{\mathcal{B}}^{\mathcal{B}}$, entonces $A^* = \begin{bmatrix} f^* \end{bmatrix}_{\mathcal{B}}^{\mathcal{B}}$ y

$$A^* A = \begin{bmatrix} f^* \end{bmatrix}_{\mathcal{B}}^{\mathcal{B}} \begin{bmatrix} f \end{bmatrix}_{\mathcal{B}}^{\mathcal{B}} = \begin{bmatrix} f^* \circ f \end{bmatrix}_{\mathcal{B}}^{\mathcal{B}}.$$

Entonces $A^* A = I_n$ si y solo si $f^* \circ f = \text{id}_V$. \square

Teorema 5.37. Si V tiene dimensión finita igual a n , la colección de operadores unitarios de V está en correspondencia biyectiva con la colección de n -tuplas $(v_1, \dots, v_n) \in V \times \dots \times V$ tales que $\{v_1, \dots, v_n\}$ es una base ortonormal de V .

Dem. El argumento es similar al de la demostración de Teorema 4.41. \square

5.4. Estructura compleja

Sea V un espacio vectorial sobre \mathbb{R} .

Propiedad 5.38. Suponga que $f \in \text{Hom}_{\mathbb{R}}(V, V)$ es un operador simple, entonces: o bien,

1. $f = c \text{id}_V$, para algún $c \in \mathbb{R}$ y en tal caso $\dim(V) = 1$; o bien,
2. $f = a \text{id}_V + bj$, donde $j \in \text{Hom}_{\mathbb{R}}(V, V)$ es tal que $j^2 = -\text{id}_V$, $a, b \in \mathbb{R}$, y en tal caso $\dim(V) = \mathbb{R}^2$.

Dem. Como f es simple, su polinomio minimal es irreducible, en particular este es de la forma $t - c$, con $c \in \mathbb{R}$, ó $(t - a)^2 + b^2$, con $a, b \in \mathbb{R}$. En el primer caso

$f - c \operatorname{id}_V = 0$, es decir $f = c \operatorname{id}_V$; en el segundo, $(f - a \operatorname{id}_V)^2 + b^2 \operatorname{id}_V = 0$, y si $j = \frac{1}{b}(f - a \operatorname{id}_V)$, tenemos $f = a \operatorname{id}_V + bj$ con

$$\begin{aligned} j^2 &= \left(\frac{1}{b}(f - a \operatorname{id}_V) \right) \circ \left(\frac{1}{b}(f - a \operatorname{id}_V) \right) \\ &= \frac{1}{b^2} (f - a \operatorname{id}_V)^2 \\ &= \frac{1}{b^2} (-b^2 \operatorname{id}_V) = -\operatorname{id}_V. \end{aligned}$$

Como f es simple, en el primer caso, dado $v \in V$, $v \neq 0$, $\langle v \rangle$ es invariante bajo f y así $V = \langle v \rangle$; en el segundo, dado $v \in V$, $v \neq 0$, $\{v, j(v)\}$ es linealmente independiente pues si

$$\alpha v + \beta j(v) = 0,$$

$\alpha, \beta \in \mathbb{R}$, entonces operando por j ,

$$-\beta v + \alpha j(v) = 0;$$

combinando las dos relaciones obtenemos

$$0 = \alpha(\alpha v + \beta j(v)) - \beta(-\beta v + \alpha j(v)) = (\alpha^2 + \beta^2)v$$

pero $v \neq 0$, luego $\alpha^2 + \beta^2 = 0$, de donde $\alpha = \beta = 0$; ahora $\langle v, j(v) \rangle$ es invariante bajo f , así $V = \langle v, j(v) \rangle$. \square

Observación 5.39. Considere \mathbb{C} como espacio vectorial sobre \mathbb{R} , con la base $\mathcal{B} = \{1, i\}$ (en particular $\mathbb{C} \simeq_{\mathbb{R}} \mathbb{R}^2$). Dado $a + bi \in \mathbb{C}$, con $a, b \in \mathbb{R}$, la función:

$$\begin{aligned} m_{a+bi} : \mathbb{C} &\longrightarrow \mathbb{C} \\ z &\longrightarrow (a + bi)z \end{aligned}$$

es un operador en $\operatorname{Hom}_{\mathbb{R}}(\mathbb{C}, \mathbb{C})$, de hecho para todo $z, z_1, z_2 \in \mathbb{C}$ y $c \in \mathbb{R}$

$$(a + bi)(z_1 + z_2) = (a + bi)z_1 + (a + bi)z_2, (a + bi)cz = c(a + bi)z.$$

Tenemos entonces que si $f = m_{a+ib}$, entonces

$$\begin{bmatrix} f \end{bmatrix}_{\mathcal{B}}^{\mathcal{B}} = \begin{bmatrix} a & -b \\ b & a \end{bmatrix}$$

y $f = a \operatorname{id}_{\mathbb{C}} + bj$ donde $j = m_i$. Note que

$$\begin{bmatrix} j \end{bmatrix}_{\mathcal{B}}^{\mathcal{B}} = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$$

Observación 5.40. La observación anterior se puede generalizar a \mathbb{C}^n el cual visto como espacio vectorial sobre \mathbb{R} tiene dimensión $2n$. Tome la base

$$\mathcal{B} = \{e_1, \dots, e_n, f_1, \dots, f_n\}$$

donde $\{e_1, \dots, e_n\}$ es la base canónica de \mathbb{C}^n , visto como espacio vectorial sobre \mathbb{C} , y, $f_1 = ie_1, \dots, f_n = ie_n$. Tome $j \in \text{Hom}_{\mathbb{R}}(\mathbb{C}^n, \mathbb{C}^n)$ definida por

$$\begin{aligned} j : \mathbb{C}^n &\longrightarrow \mathbb{C}^n \\ (z_1, \dots, z_n) &\longmapsto i(z_1, \dots, z_n). \end{aligned}$$

Entonces $j^2 = -\text{id}_{\mathbb{C}^n}$ y

$$\begin{bmatrix} j \end{bmatrix}_{\mathcal{B}} = \begin{bmatrix} 0 & -I_n \\ I_n & 0 \end{bmatrix}$$

donde 0 denota el origen de $M_{n \times n}(\mathbb{R})$ y $I_n \in M_{n \times n}(\mathbb{R})$ es la matriz con unos en la diagonal y ceros en el resto de entradas.

Definición 5.41. Una *estructura compleja* en V es un operador $j \in \text{Hom}_{\mathbb{R}}(V, V)$ tal que $j^2 = -\text{id}_V$.

Propiedad 5.42. *Suponga que V tiene dimensión finita. Si V admite una estructura compleja j , entonces la dimensión de V es par. En tal caso V es un espacio vectorial sobre \mathbb{C} mediante el producto por escalar definido por*

$$(a + bi)v = (a \text{id}_V + bj)(v)$$

para todo $a, b \in \mathbb{R}$ y $v \in V$. Más aún, existe una base de la forma $T = \{v_1, \dots, v_n, j(v_1), \dots, j(v_n)\}$ donde $2n = \dim(V)$. En particular

$$\begin{bmatrix} j \end{bmatrix}_{\mathcal{B}} = \begin{bmatrix} 0 & -I_n \\ I_n & 0 \end{bmatrix}$$

Dem. Si m es la dimensión de V entonces

$$0 \leq (\det(j))^2 = \det(j^2) = \det(-\text{id}_V) = (-1)^m$$

luego m es par, es decir $m = 2n$ para algún $n \in \mathbb{Z}_{\geq 0}$. Para verificar que bajo la multiplicación por escalar definida V es un espacio vectorial bajo \mathbb{C} basta verificar que esta es unitaria, asociativa y que es distributiva. Lo cual se sigue de las siguientes igualdades

$$\begin{aligned} 1v &= \text{id}_V(v) = v \\ (a + bi)((c + di)v) &= (a \text{id}_V + bj) \circ (c \text{id}_V + dj)(v) \\ &= (ac \text{id}_V + adj + bcj + bdj^2)(v) \\ &= ((ac - bd) \text{id}_V + (ad + bc)j)(v) \\ &= ((a + bi)(c + di))v \\ (a + bi)(v + w) &= (a \text{id}_V + bj)(v + w) \\ &= (a \text{id}_V + bj)(v) + (a \text{id}_V + bj)(w) \\ &= (a + bi)v + (a + bi)w \\ ((a + bi) + (c + di))v &= ((a + c) \text{id}_V + (b + d)j)v \\ &= (a \text{id}_V + bj + c \text{id}_V + dj)v \\ &= (a \text{id}_V + bj)(v) + (c \text{id}_V + dj)(v) \\ &= (a + bi)v + (c + di)v \end{aligned}$$

validas para todo $a, b, c, d \in \mathbb{R}$ y $v, w \in V$.

Sea $\{v_1, \dots, v_{n'}\}$ una base de V como espacio vectorial sobre \mathbb{C} , donde n' es su dimensión. Entonces V es generado como espacio vectorial sobre \mathbb{R} , por $\{v_1, \dots, v_{n'}, j(v_1), \dots, j(v_{n'})\}$ pues

$$\sum_{k=1}^{n'} z_k v_k = \sum_{k=1}^{n'} a_k v_k + b_k j(v_k),$$

donde para $k = 1, \dots, n'$, $z_k = a_k + b_k i$ con $a_k, b_k \in \mathbb{R}$. Más este conjunto de generadores es linealmente independiente pues si $a_1, \dots, a_{n'}, b_1, \dots, b_{n'} \in \mathbb{R}$ son tales que $\sum_{k=1}^{n'} a_k v_k + b_k j(v_k) = 0$, entonces $\sum_{k=1}^{n'} z_k v_k = 0$, con $z_k = a_k + b_k i$; luego todo $z_k = 0$ y así todo $a_k = b_k = 0$. Tenemos $2n' = 2n$ y $\mathcal{B} = \{v_1, \dots, v_n, j(v_1), \dots, j(v_n)\}$ es una base de V como espacio vectorial sobre \mathbb{R} . \square

Propiedad 5.43. Sean j una estructura compleja en V y $f \in \text{Hom}_{\mathbb{R}}(V, V)$. Entonces, tomando V como un espacio vectorial sobre \mathbb{C} mediante el producto por escalar definido por $(a + bi)v = (a \text{id}_V + bj)(v)$, para todo $a, b \in \mathbb{R}$ y $v \in V$, tenemos $f \in \text{Hom}_{\mathbb{C}}(V, V)$ si y solo si $f \circ j = j \circ f$, o, equivalentemente, $-j \circ f \circ j = f$.

Dem. Suponga que $f \in \text{Hom}_{\mathbb{C}}(V, V)$, entonces para todo $v \in V$

$$f \circ j(v) = f(iv) = if(v) = j \circ f(v).$$

Recíprocamente, si $f \circ j = j \circ f$, para todo $a, b \in \mathbb{R}$ y $v \in V$,

$$\begin{aligned} f((a + bi)v) &= f \circ (a \text{id}_V + bj)(v) \\ &= (af + bf \circ j)(v) \\ &= (af + bj \circ f)(v) \\ &= (a \text{id}_V + bj) \circ f(v) \\ &= (a + bi)f(v) \end{aligned}$$

luego $f \in \text{Hom}_{\mathbb{C}}(V, V)$. Finalmente componemos ambos lados de la igualdad $f \circ j = j \circ f$ por $-j = j^{-1}$ para obtener $-j \circ f \circ j = f$. \square

Teorema 5.44. Suponga que V tiene dimensión finita igual a $2n$ y sea $j \in \text{Hom}_{\mathbb{R}}(\mathbb{C}^n, \mathbb{C}^n)$ el operador que corresponde a multiplicación por i . Entonces cada estructura compleja en V es de la forma

$$j_f = f \circ j \circ f^{-1}$$

para algún isomorfismo $f \in \text{Hom}_{\mathbb{R}}(\mathbb{C}^n, V)$. Más aún $j_f = j_g$ si y solo si $g^{-1} \circ f \in \text{Hom}_{\mathbb{C}}(\mathbb{C}^n, \mathbb{C}^n)$.

Dem. Dado un isomorfismo $f \in \text{Hom}_{\mathbb{R}}(\mathbb{C}^n, V)$, el operador

$$j_f = f \circ j \circ f^{-1} \in \text{Hom}_{\mathbb{R}}(V, V)$$

es una estructura compleja (ver Figura 5.1), pues

$$\begin{aligned}
 j_f^2 &= f \circ j \circ f^{-1} \circ f \circ j \circ f^{-1} \\
 &= f \circ j^2 \circ f^{-1} \\
 &= f \circ -\text{id}_{\mathbb{C}^n} \circ f^{-1} \\
 &= -f \circ f^{-1} \\
 &= -\text{id}_V.
 \end{aligned}$$

Recíprocamente, dada una estructura compleja j_V en V , usando la notación en

$$\begin{array}{ccc}
 \mathbb{C}^n & \xrightarrow{f} & V \\
 \downarrow j & & \downarrow j_f \\
 \mathbb{C}^n & \xrightarrow{f} & V
 \end{array}$$

Figura 5.1: Estructura compleja

Observación 5.40, Propiedad 5.42 implica que

$$\begin{aligned}
 f : \mathbb{C}^n &\longrightarrow V \\
 e_k &\longmapsto v_k \\
 f_k &\longmapsto j_V(v_k)
 \end{aligned}$$

donde $\{v_1, \dots, v_n, j_V(v_1), \dots, j_V(v_n)\}$ es una base de V , es un isomorfismo de espacios vectoriales sobre \mathbb{R} tal que $j_f = j_V$. Finalmente, suponga que $f, g \in \text{Hom}_{\mathbb{R}}(\mathbb{C}^n, V)$ son isomorfismos tales que $j_f = j_g$, es decir

$$f \circ j \circ f^{-1} = g \circ j \circ g^{-1}.$$

Entonces, como $(g^{-1} \circ f) \circ j = j \circ (g^{-1} \circ f)$, por Propiedad 5.43 tenemos que $g^{-1} \circ f \in \text{Hom}_{\mathbb{C}}(\mathbb{C}^n, \mathbb{C}^n)$. \square

Propiedad 5.45. *Suponga que V es un espacio euclideo con producto interno denotado por $\langle \bullet, \bullet \rangle$ y j es una estructura compleja en V . Entonces, tomando V como un espacio vectorial sobre \mathbb{C} mediante el producto por escalar definido por $(a + bi)v = (a \text{id}_V + bj)(v)$, para todo $a, b \in \mathbb{R}$ y $v \in V$,*

$$\begin{aligned}
 \langle \bullet; \bullet \rangle_j : V \times V &\longrightarrow \mathbb{C} \\
 (v_1, v_2) &\longmapsto \langle v_1; v_2 \rangle + \langle v_1; j(v_2) \rangle i
 \end{aligned}$$

es un producto de hermítico sobre V si y solo si j es ortogonal.

Dem. Suponga primero que j es ortogonal. Entonces $j^* = j^{-1} = -j$. Note que $\langle \bullet, \bullet \rangle_f$ es sesquilineal y además hermítica pues

$$\begin{aligned}\langle v_2; v_1 \rangle_j &= \langle v_2; v_1 \rangle + \langle v_2; j(v_1) \rangle i \\ &= \langle v_2; v_1 \rangle - \langle j(v_2); v_1 \rangle i \\ &= \langle v_1; v_2 \rangle - \langle v_1; j(v_2) \rangle i \\ &= \overline{\langle v_1; v_2 \rangle_j}.\end{aligned}$$

Ahora, dado $v \in V$,

$$\langle v; j(v) \rangle = -\langle j(v); v \rangle = -\langle v; j(v) \rangle$$

luego $\langle v; j(v) \rangle = 0$. Si además $v \neq 0$,

$$\langle v; v \rangle_j = \langle v; v \rangle + \langle v; j(v) \rangle i = \langle v; v \rangle > 0$$

luego $\langle \bullet, \bullet \rangle_f$ es definitivamente positiva y así es un producto hermítico. Recíprocamente suponga que $\langle \bullet, \bullet \rangle_f$ es producto hermítico, entonces para todo $v_1, v_2 \in V$

$$\begin{aligned}\langle j(v_1); j(v_2) \rangle &= \operatorname{Re}(\langle j(v_1); j(v_2) \rangle_j) \\ &= (\langle j(v_1); j(v_2) \rangle_j + \langle j(v_2); j(v_1) \rangle_j) / 2 \\ &= (\langle iv_1; iv_2 \rangle_j + \langle iv_2; iv_1 \rangle_j) / 2 \\ &= (\langle v_1; v_2 \rangle_j + \langle v_2; v_1 \rangle_j) / 2 \\ &= \operatorname{Re}(\langle v_1; v_2 \rangle_j) \\ &= \langle v_1; v_2 \rangle\end{aligned}$$

luego j es ortogonal. □

Observación 5.46. Note que la norma definida por $\langle \bullet, \bullet \rangle$ y por $\langle \bullet, \bullet \rangle_j$ coinciden.

Propiedad 5.47. Suponga que V es un espacio euclideo con producto interno denotado por $\langle \bullet, \bullet \rangle$, j es una estructura compleja ortogonal en V y $f \in \operatorname{Hom}_{\mathbb{R}}(V, V)$. Entonces, tomando V como un espacio unitario mediante el producto por escalar definido por $(a + bi)v = (a \operatorname{id}_V + bj)(v)$, para todo $a, b \in \mathbb{R}$ y $v \in V$, y el producto hermítico

$$\langle \bullet, \bullet \rangle_j = \langle \bullet, \bullet \rangle + \langle \bullet; j(\bullet) \rangle i,$$

tenemos f es unitario si y solo si $f \circ j = j \circ f$ y f es ortonormal.

Dem. Suponga primero que f es unitaria, entonces $f \in \operatorname{Hom}_{\mathbb{C}}(V, V)$ y así $f \circ j = j \circ f$; además para todo $v \in V$

$$\|j(v)\|^2 = \langle j(v); j(v) \rangle = \langle j(v); j(v) \rangle_j = \langle v; v \rangle_j = \langle v; v \rangle = \|v\|^2$$

luego f es ortonormal. Recíprocamente, si $f \circ j = j \circ f$ y f es ortogonal, para todo $v_1, v_2 \in V$,

$$\begin{aligned} \langle f(v_1); f(v_2) \rangle_j &= \langle f(v_1); f(v_2) \rangle + \langle f(v_1); j \circ f(v_2) \rangle i \\ &= \langle f(v_1); f(v_2) \rangle + \langle f(v_1); f \circ j(v_2) \rangle i \\ &= \langle v_1; v_2 \rangle + \langle v_1; j(v_2) \rangle i \\ &= \langle v_1; v_2 \rangle_j, \end{aligned}$$

es decir, f es unitaria. □

Observación 5.48. Bajo las hipótesis de la propiedad, suponga además que V tiene dimensión finita. Si f es unitaria, entonces

$$j = f^* \circ j \circ f$$

pues, como f es ortogonal, $f^* = f^{-1}$ y, como $f \in \text{Hom}_{\mathbb{C}}(V, V)$, $f \circ j = j \circ f$, así $f^* \circ j \circ f = f^{-1} \circ f \circ j = j$. Recíprocamente, suponga que $j = f^* \circ j \circ f$, entonces para todo $v_1, v_2 \in V$

$$\begin{aligned} \langle f(v_1), j \circ f(v_2) \rangle &= \langle v_1, f^* \circ j \circ f(v_2) \rangle \\ &= \langle v_1, j(v_2) \rangle \end{aligned}$$

luego f preserva la parte imaginaria del producto hermítico $\langle \bullet; \bullet \rangle_j$. Si además f es ortogonal, f preserva la parte real y como en tal caso $f^* = f^{-1}$,

$$f \circ j = f \circ f^* \circ j \circ f = j \circ f,$$

luego $f \in \text{Hom}_{\mathbb{C}}(V, V)$ y f es unitaria. Por otro lado si f conmuta con j , entonces $j = f^* \circ j \circ f = f^* \circ f \circ j$ y así $\text{id}_V = f^* \circ f$, es decir f es ortogonal y a su vez f es entonces unitaria.

El hecho que $j = f^* \circ j \circ f$ sea equivalente a que f preserve la parte imaginaria del producto hermítico, $\langle \bullet; j(\bullet) \rangle$, es la motivación para el contenido del próximo capítulo donde entraremos a estudiar este tipo de estructuras que se llaman espacios simplécticos, las cuales generalizan esta intersección entre espacios ortogonales, espacios unitarios y estructuras complejas.

Capítulo 6

Espacios simplécticos

Sea K un cuerpo y V un espacio vectorial sobre K .

6.1. Forma simpléctica

Definición 6.1. Una *forma simpléctica* en V es una función

$$\begin{aligned}\sigma : V \times V &\longrightarrow K \\ (v_1, v_2) &\longmapsto \sigma(v_1, v_2)\end{aligned}$$

tal que:

1. *es bilineal*: para todo $v, v_1, v_2 \in V$ y $c \in \mathbb{C}$

$$\begin{aligned}\sigma(v_1 + v_2, v) &= \sigma(v_1, v) + \sigma(v_2, v) \\ \sigma(cv_1, v_2) &= c\sigma(v_1, v_2) \\ \sigma(v, v_1 + v_2) &= \sigma(v, v_1) + \sigma(v, v_2) \\ \sigma(v_1, cv_2) &= c\sigma(v_1, v_2);\end{aligned}$$

2. *es alternante*: para todo $v \in V$

$$\sigma(v, v) = 0;$$

3. *es no-degenerada* Si $v \in V$ es tal que $\sigma(v, w) = 0$ para todo $w \in V$ entonces $v = 0$.

Un *espacio simpléctico* es un espacio vectorial provisto de una forma simpléctica.

Observación 6.2. Para todo $v_1, v_2 \in V$,

$$\sigma(v_2, v_1) = -\sigma(v_1, v_2).$$

De hecho, la condición alternante implica que

$$\begin{aligned} 0 &= \sigma(v_1 + v_2, v_1 + v_2) \\ &= \sigma(v_1, v_1) + \sigma(v_1, v_2) + \sigma(v_2, v_1) + \sigma(v_2, v_2) \\ &= \sigma(v_1, v_2) + \sigma(v_2, v_1) \end{aligned}$$

Ejemplo 6.3. 1. Sobre $V = K^{2n} = K^n \times K^n$

$$\sigma((\bar{q}, \bar{p}), (\bar{q}', \bar{p}')) = \sum_{i=1}^n p_i q'_i - p'_i q_i$$

donde $\bar{q} = (q_1, \dots, q_n)$, $\bar{p} = (p_1, \dots, p_n)$, $\bar{q}' = (q'_1, \dots, q'_n)$ y $\bar{p}' = (p'_1, \dots, p'_n)$.

2. Sobre $V \times V^*$

$$\sigma((v, \lambda), (w, \mu)) = \lambda(w) - \mu(v)$$

donde $v, w \in V$ y $\lambda, \mu \in V^*$.

3. Suponga que $K = \mathbb{R}$ y V es un espacio euclídeo con una estructura compleja ortogonal j . Sobre V

$$\sigma(v_1, v_2) = \langle v_1, j(v_2) \rangle$$

Observación 6.4. Una forma simpléctica σ , al ser bilineal, induce una transformación lineal

$$\begin{aligned} s : V &\longrightarrow V^* \\ v &\longmapsto s_v = \sigma(v, \bullet) : w \mapsto \sigma(v, w). \end{aligned}$$

El hecho que σ sea no-degenerada implica que s es inyectiva; y es un isomorfismo si V tiene dimensión finita. Pues, la condición de que σ sea no degenerada quiere decir que $s_v = 0$ si y solo si $v = 0$.

Con este mapa, tenemos

$$s(v)(w) = s_v(w) = \sigma(v, w)$$

En particular s es una transformación lineal $V \rightarrow V^*$ y su dual s^* es una transformación lineal $(V^*)^* \rightarrow V^*$.

Proposición 6.5. Sea V un espacio simpléctico de dimensión finita y

$$\begin{aligned} \hat{\bullet} : V &\longrightarrow (V^*)^* \\ v &\longmapsto \hat{v} : \lambda \mapsto \lambda(v). \end{aligned}$$

el isomorfismo canónico. Entonces para todo $v \in V$

$$s^*(\hat{v}) = -s(v)$$

Dem. Para todo $w \in V$

$$\begin{aligned} s^*(\widehat{v})(w) &= \widehat{v}(s(w)) \\ &= s(w)(v) \\ &= \sigma(w, v) \\ &= -\sigma(v, w) \\ &= -s(v)(w). \end{aligned}$$

□

Propiedad 6.6. Si V es un espacio simpléctico y tiene dimensión finita, entonces su dimensión es par.

Dem. Sea $T = \{v_1, \dots, v_m\}$ una base de V y $T^* = \{\lambda_1, \dots, \lambda_m\}$ la base de V^* dual de T . Tomamos la imagen de T mediante el isomorfismo canónico $V \mapsto (V^*)^*$, la cual es la base $\widehat{T} = \{\widehat{v}_1, \dots, \widehat{v}_m\}$ de $(V^*)^*$ dual de T^* . La proposición anterior implica que si tomamos las representaciones matriciales en $M_{m \times m}(K)$

$$A = \begin{bmatrix} s \end{bmatrix}_T^{T^*}, \text{ y } B = \begin{bmatrix} s^*(\bullet) \end{bmatrix}_T^{T^*} = \begin{bmatrix} s^* \end{bmatrix}_{\widehat{T}}^{T^*} \begin{bmatrix} \widehat{\bullet} \end{bmatrix}_T^{\widehat{T}} = \begin{bmatrix} s^* \end{bmatrix}_{\widehat{T}}^{T^*},$$

entonces $B = -A$, pero por otro lado $B = A^\top$, luego $A^\top = -A$. De donde

$$\det(A) = \det(A^\top) = \det(-A) = (-1)^m \det(A).$$

Ahora como $\sigma : V \rightarrow V^*$ es inyectiva, $\det(A) \neq 0$ y así $1 = (-1)^m$, en particular $m = 2n$ para algún $n \in \mathbb{Z}_{\geq 0}$. □

Definición 6.7. Sean V un espacio simpléctico y $S \subseteq V$, el conjunto σ -ortogonal a S está definido por

$$S^\sigma = \{v \in V \mid \sigma(w, v) = 0, \text{ para todo } w \in S\}$$

Observación 6.8. Note que $S^\sigma = (s(S))_0$. De hecho

$$\begin{aligned} S^\sigma &= \{v \in V \mid \sigma(w, v) = 0, \text{ para todo } w \in S\} \\ &= \{v \in V \mid s(w)(v) = 0, \text{ para todo } w \in S\} \\ &= \{v \in V \mid \lambda(v) = 0, \text{ para todo } \lambda \in s(S)\} \\ &= (s(S))_0 \end{aligned}$$

Esto implica la siguiente propiedad.

Propiedad 6.9. Sean V un espacio simpléctico y $S \subseteq V$, entonces $S^\sigma \leq V$. Si $S' \subseteq S$ entonces $S^\sigma \leq S'^\sigma$. Si V tiene dimensión finita y $U \leq V$ entonces

$$\dim(U) + \dim(U^\sigma) = \dim(V)$$

6.2. Subespacios isotrópicos y bases de Darboux

Sea V un espacio simpléctico sobre K .

Definición 6.10. Sea $U \leq V$, entonces decimos que

1. U es un *subespacio simpléctico* si la restricción de σ a $U \times U$ es una forma simpléctica;
2. U es un *subespacio isotrópico* si $U \leq U^\sigma$;
3. U es un *subespacio lagrangiano* si $U = U^\sigma$;

Observación 6.11. Suponga que V tiene dimensión finita. Si $U \leq V$ es un subespacio lagrangiano y $\dim(V) = 2n$ entonces $\dim(U) = n$, de hecho como $U = U^\sigma$,

$$2n = \dim(V) = \dim(U) + \dim(U^\sigma) = 2\dim(U).$$

Ejemplo 6.12. 1. En el espacio simpléctico de Ejemplo 6.3.1, $V = K^n \times K^n$, denote, para $i = 1, \dots, n$, e_i el elemento cuya i -ésima entrada es 1 y el resto ceros, y f_i el elemento cuya $n+i$ -ésima entrada es 1 y el resto ceros. Note que para todo $i, j \in \{1, \dots, n\}$,

$$\sigma(e_i, f_j) = -\delta_{ij} \quad \sigma(e_i, e_j) = \sigma(f_i, f_j) = 0.$$

Entonces para cualquier subconjunto de índices $J \subset I = \{1, \dots, n\}$,

$$V_J = \text{Sp}(\{e_j, f_j\}_{j \in J})$$

es un subespacio simpléctico,

$$E_J = \text{Sp}(\{e_j\}_{j \in J}), \text{ y } F_J = \text{Sp}(\{f_j\}_{j \in J})$$

son isotrópicos, y si $J = I$, E_I y F_I son subespacios lagrangianos.

2. En el espacio simpléctico de Ejemplo 6.3.2, $V \times V^*$, sea $\{v_i\}_{i \in I}$ una base de V y $\{\lambda_i\}_{i \in I'}$ una base de V^* donde $I \subseteq I'$ y $\lambda_i(v_j) = \delta_{ij}$ para todo $i, j \in I$. Note que para todo $i, j \in I$

$$\sigma((v_i, 0), (0, \lambda_j)) = -\delta_{ij} \quad \sigma((v_i, 0), (v_j, 0)) = 0$$

y para todo $i, j \in I'$

$$\sigma((0, f_i), (0, f_j)) = 0.$$

Entonces para cualquier subconjunto de índices $J \subset I$,

$$(V \times V^*)_J = \text{Sp}(\{(v_j, 0), (0, \lambda_j)\}_{j \in J})$$

es un subespacio simpléctico,

$$E_J = \text{Sp}(\{(v_j, 0)\}_{j \in J}), \text{ y } F_J = \text{Sp}(\{(0, \lambda_j)\}_{j \in J})$$

son isotrópicos, y si $J = I$, E_I y F_I son subespacios lagrangianos.

Proposición 6.13. *Sea $U \leq V$, entonces*

1. *U es un subespacio simpléctico si y solo si la restricción de s a U es inyectiva. En particular, U es un subespacio simpléctico si y solo si $U \cap U^\sigma = \{0\}$.*
2. *U es un subespacio isotrópico si y solo si $\sigma(u, u') = 0$ para todo $u, u' \in U$ (es decir $s(U) = 0$).*
3. *U es un subespacio lagrangiano si y solo si U es un subespacio isotrópico maximal.*

Dem.

1. Si U es subespacio simpléctico, entonces la restricción de σ a $U \times U$ es no-degenerada, en particular dado $u \in U$, $u \neq 0$, existe $w \in U$ tal que $\sigma(u, w) \neq 0$. Así pues la imagen en U^* , $s(u)$, es diferente de 0 pues $s(u)(w) \neq 0$. Es decir el núcleo de s restringido a U es $\{0\}$, luego es inyectiva. Recíprocamente, si la restricción s a U es inyectiva, la restricción de σ a $U \times U$ es bilineal, alternante y no-degenerada, luego U es un subespacio simpléctico. Para establecer la segunda afirmación basta con observar que $u \in U \cap U^\sigma$ si y solo si $s(u)(w) = 0$ para todo $w \in U$, es decir si y solo si u pertenece al núcleo de la restricción de s a U .
2. Suponga que U es isotrópico, luego para todo $u, u' \in U$, como $U \leq U^\sigma$, $u' \in U^\sigma$ y $\sigma(u, u') = 0$. Recíprocamente, si $\sigma(u, u') = 0$ para todo $u, u' \in U$, entonces $U \leq U^\sigma$.
3. Suponga que U es un subespacio lagrangiano, entonces U es isotrópico. Suponga que existe $U' \leq V$ isotrópico tal que $U \leq U'$. Sea $u' \in U'$, entonces $\sigma(u, u') = 0$ para todo $u \in U$, en particular $u' \in U^\sigma = U$. Luego $U' = U$. Recíprocamente, si U es isotrópico maximal, dado $u' \in U^\sigma$,

$$\sigma(u_1 + au', u_2 + bu') = \sigma(u_1, u_2) + b\sigma(u_1, u') + a\sigma(u', u_2) + \sigma(u', u') = 0,$$

luego $U + \text{Sp}(\{u'\})$ es isotrópico y así $u' \in U$. Luego $U = U^\sigma$. \square

Propiedad 6.14. *Suponga que V tiene dimensión finita y sea $U_0 \leq V$ un subespacio isotrópico. Entonces existe un subespacio lagrangiano $U \leq V$ que contiene a U_0 .*

Dem. Tenemos $U_0 \leq U_0^\sigma$. Si $U_0 = U_0^\sigma$, entonces $U = U_0$ es un subespacio lagrangiano, de lo contrario existe $u \in U_0^\sigma \setminus U_0$. Entonces, $u \neq 0$ y para todo $u_1 + au, u_2 + bu \in U' = U_0 + \text{Sp}(\{u\})$, $u_1, u_2 \in U$ y $a, b \in K$,

$$\sigma(u_1 + au, u_2 + bu) = \sigma(u_1, u_2) + b\sigma(u_1, u) + a\sigma(u, u_2) + \sigma(u, u) = 0,$$

luego U' es isotrópico y

$$U_0 < U' \leq U'^\sigma < U_0^\sigma.$$

Reemplazamos U_0 por U' y continuamos recursivamente. Por monotonía de la dimensión, como V tiene dimensión finita, eventualmente obtenemos $U = U'$ subespacio lagrangiano. \square

Observación 6.15 (Extensión de subespacios isotrópicos a lagrangianos en dimensión infinita). El mismo resultado de la proposición anterior, se puede generalizar a espacios simplécticos de dimensión infinita usando Lema de Zorn. De hecho, dado $U_0 \leq V$ subespacio isotrópico, consideramos la colección P de subespacios isotrópicos que contienen a U_0 , ordenados por contención. Como $U_0 \in P$, $P \neq \emptyset$. También, la unión de elementos en una cadena de P está en P y es una cota superior de la cadena. Entonces U maximal en P por la proposición anterior es lagrangiano y contiene a U_0 .

Proposición 6.16. *Suponga que V tiene dimensión finita y sea $V_1 \leq V$ un subespacio lagrangiano. Entonces dado $U \leq V$, isotrópico, tal que $U \cap V_1 = \{0\}$, existe $V_2 \leq V$ lagrangiano, tal que $U \leq V_2$ y $V_1 \cap V_2 = \{0\}$, en particular*

$$V = V_1 \oplus V_2.$$

Dem. Suponga que $\dim(V) = 2n$. Si $U = U^\sigma$, entonces $V_2 = U$ es lagrangiano, de lo contrario $\dim(U) = k < n$ y $\dim(U^\sigma) = 2n - k > n$, y, como $\dim(V_1) = n$,

$$\dim(U + V_1) = \dim(U) + \dim(V_1) - \dim(U \cap V_1) = k + n$$

Suponga por contradicción que $U^\sigma \subseteq U + V_1$, entonces tomando los espacios σ -ortogonales,

$$U^\sigma \cap V_1 = U^\sigma \cap V_1^\sigma = (U + V_1)^\sigma \subseteq (U^\sigma)^\sigma = U$$

de donde $U^\sigma \cap V_1 \subseteq U \cap V_1 = \{0\}$. Entonces

$$\dim(U^\sigma + V_1) = \dim(U^\sigma) + \dim(V_1) > n + n = 2n = \dim(V)$$

lo cual es una contradicción. Así pues, existe $u \in U^\sigma \setminus (U + V_1)$. Entonces, $u \notin V_1$ y para todo $u_1 + au, u_2 + bu \in U' = U + \text{Sp}(\{u\})$, $u_1, u_2 \in U$ y $a, b \in K$,

$$\sigma(u_1 + au, u_2 + bu) = \sigma(u_1, u_2) + b\sigma(u_1, u) + a\sigma(u, u_2) + \sigma(u, u) = 0,$$

luego U' es isotrópico, $U' \cap V_1 = \{0\}$ y

$$U < U' \leq U'^\sigma < U^\sigma.$$

Reemplazamos U_0 por U' y continuamos recursivamente. Por monotonía de la dimensión, como V tiene dimensión finita, eventualmente obtenemos $V_2 = U'$ lagrangiano con $V_2 \cap V_1 = \{0\}$. Ahora como $\dim(V_1) = \dim(V_2) = n$ entonces $\dim(V_1 + V_2) = 2n$ y $V = V_1 \oplus V_2$. \square

Propiedad 6.17. *Suponga que V tiene dimensión finita y sean $V_1, V_2 \leq V$ lagrangianos tales que*

$$V = V_1 \oplus V_2,$$

entonces si $\pi_1 : V \rightarrow V_1$ y $\pi_2 : V \rightarrow V_2$ son respectivamente las proyecciones sobre V_1 y V_2 dadas por la descomposición $V = V_1 \oplus V_2$,

$$\pi_1^*(V_1^*) = s(V_2) \quad y \quad \pi_2^*(V_2^*) = s(V_1)$$

Dem. (Ver Figura 6.1) Como π_1 y π_2 son sobreyectivas, π_1^* y π_2^* son inyectivas. Luego dado $\lambda \in \pi^*(V_1)$, existe un único $\lambda_1 \in V_1^*$ tal que $\pi^*(\lambda_1) = \lambda$. En particular, para todo $v_2 \in V_2$, como $\pi_1(v_2) = 0$,

$$\lambda(v_2) = \pi^*(\lambda_1)(v_2) = \lambda_1(\pi_1(v_2)) = 0.$$

Ahora, como V tiene dimensión finita $\sigma : V \rightarrow V^*$ es un isomorfismo, luego existe un único $v \in V$ tal que $\sigma(v) = \lambda$, pero $\sigma(v, v_2) = \lambda(v_2) = 0$ para todo $v_2 \in V_2$, entonces $v \in V_2^\sigma = V_2$. Así $\pi_1^*(V_1^*) \subseteq S(V_2)$. Recíprocamente, dado $v_2 \in V_2$, defina $\lambda_1 \in V_1^*$ por

$$\lambda_1 : v_1 \mapsto s(v_2)(v_1) = \sigma(v_2, v_1)$$

el mapa

$$\begin{aligned} V_2 &\longrightarrow \pi_1^*(V_1^*) \\ v &\longmapsto s(v) = \lambda = \pi_1^*(\lambda_1) \end{aligned}$$

es inyectivo, y como V_1^* y $\sigma(V_2)$ tienen la misma dimensión, es un isomorfismo. \square

Definición 6.18. Suponga que V tiene dimensión finita y sea $T = \{v_1, \dots, v_n, w_1, \dots, w_n\}$ una base de V . Decimos que T es una base de Darboux si para todo $i, j \in \{1, \dots, n\}$

$$\begin{aligned} \sigma(v_i, v_j) &= 0 &= \sigma(w_i, w_j) \\ \sigma(w_i, v_j) &= \delta_{ij}. \end{aligned}$$

Observación 6.19. Note que si $T = \{v_1, \dots, v_n, w_1, \dots, w_n\}$ es una base de Darboux entonces $V_1 = \text{Sp}(\{v_1, \dots, v_n\})$ y $V_2 = \text{Sp}(\{w_1, \dots, w_n\})$ son lagrangianos. Más aún si $T_1 = \{v_1, \dots, v_n\}$, $T_2 = \{w_1, \dots, w_n\}$ y $T_1^* = \{\lambda_1, \dots, \lambda_n\}$ y $T_2^* = \{\mu_1, \dots, \mu_n\}$ son las bases de V_1^* y V_2^* , respectivamente, duales de T_1 y T_2 . Entonces para $i = 1, \dots, n$,

$$s(w_i) = \pi_1^*(\lambda_i) \quad s(-v_i) = \pi_2^*(\mu_i).$$

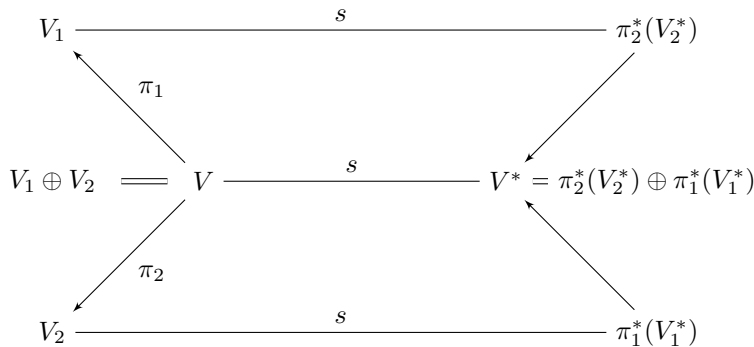


Figura 6.1: Descomposición lagrangiana

Teorema 6.20. *Suponga que V tiene dimensión finita, entonces V admite un base de Darboux.*

Dem. Como V tiene dimensión finita, existen $V_1, V_2 \leq V$ subespacios lagrangianos, tales que $V = V_1 \oplus V_2$. Sea $T_1 = \{v_1, \dots, v_n\}$ una base de V_1 y $T_1^* = \{\lambda_1, \dots, \lambda_n\}$ la base de V_1^* dual de T_1 . Como $s(V_2) = \pi_1^*(V_1^*)$, donde $\pi_1 : V \rightarrow V_1$ es la proyección en el primer sumando de la descomposición $V = V_1 \oplus V_2$, entonces existe $T_2 = \{w_1, \dots, w_n\}$ base de V_2 tal que para $i = 1, \dots, n$,

$$s(w_i) = \pi_1^*(\lambda_i).$$

Entonces como V_1 y V_2 son subespacios lagrangianos, en particular isotrópicos,

$$\sigma(v_i, v_j) = 0 = \sigma(w_i, w_j)$$

para todo $i, j \in \{1, \dots, n\}$; además

$$\sigma(w_i, v_j) = s(w_i)(v_j) = \pi_1^*(\lambda_i)(v_j) = \lambda_i(\pi_1(v_j)) = \lambda_i(v_j) = \delta_{ij}.$$

□

Propiedad 6.21. *Suponga que V tiene dimensión finita y sea $T = \{v_1, \dots, v_n, w_1, \dots, w_n\}$ una base de Darboux. Entonces para todo $v \in V$*

$$v = \sum_{i=1}^n \sigma(w_i, v) v_i - \sigma(v_i, v) w_i.$$

En particular, si $v_1, v_2 \in V$ son tales que

$$v_1 = \sum_{i=1}^n q_i v_i + p_i w_i \quad v_2 = \sum_{i=1}^n q'_i v_i + p'_i w_i$$

entonces

$$\sigma(v_1, v_2) = \sum_{i=1}^n p_i q'_i - p'_i q_i$$

Dem. Sean $a_1, \dots, a_n, b_1, \dots, b_n \in K$ tales que

$$v = \sum_{i=1}^n a_i v_i + b_i w_i,$$

de forma que para $j = 1, \dots, n$,

$$\sigma(v_j, v) = \sum_{i=1}^n a_i \sigma(v_j, v_i) + b_i \sigma(v_j, w_i) = -b_j$$

y

$$\sigma(w_j, v) = \sum_{i=1}^n a_i \sigma(w_j, v_i) + b_i \sigma(w_j, w_i) = a_j.$$

Finalmente

$$\begin{aligned}
\sigma(v_1, v_2) &= \sigma\left(\sum_{i=1}^n q_i v_i + p_i w_i, \sum_{j=1}^n q'_j v_j + p'_j w_j\right) \\
&= \sum_{i,j=1}^n q_i q'_j \sigma(v_i, v_j) + p_i q'_j \sigma(w_i, v_j) + q_i p'_j \sigma(v_i, w_j) + p_i p'_j \sigma(w_i, w_j) \\
&= \sum_{i,j=1}^n (p_i q'_j - q_i p'_j) \delta_{ij} \\
&= \sum_{i=1}^n p_i q'_i - p'_i q_i
\end{aligned}$$

□

Observaci3n 6.22. Note que si $T = \{v_1, \dots, v_n, w_1, \dots, w_n\}$, es una base de Darboux, para todo $J \subseteq \{1, \dots, n\}$,

$$V_J = \text{Sp}(\{v_j, w_j\}_{j \in J})$$

es tambi3n un subespacio simpl3ctico.

Teorema 6.23 (Ortogonalizaci3n de Gram-Schmidt). *Suponga que $\dim(V) = 2n$ y sean $U \leq V$ subespacios simpl3ctico con $\dim(U) = 2m$, $m \leq n$. Entonces existe una base de Darboux $T = \{v_1, \dots, v_n, w_1, \dots, w_n\}$ de V , tal que*

$$U = \text{Sp}(\{v_j, w_j\}_{j=1, \dots, m})$$

Dem. Sea $T' = \{v_1, \dots, v_m, w_1, \dots, w_m\}$ una base de Darboux de U . Si $m = n$ hemos terminado. De lo contrario, $m + 1 \leq n$ y tome $v'_{m+1} \in V \setminus U$. Defina

$$v_{m+1} = v'_{m+1} - \left(\sum_{i=1}^m \sigma(w_i, v'_{m+1}) v_i - \sigma(v_i, v'_{m+1}) w_i \right)$$

de forma tal que para $j = 1, \dots, m$,

$$\begin{aligned}
\sigma(v_j, v_{m+1}) &= \sigma(v_j, v'_{m+1}) - \left(\sum_{i=1}^m \sigma(w_i, v'_{m+1}) \sigma(v_j, v_i) - \sigma(v_i, v'_{m+1}) \sigma(v_j, w_i) \right) \\
&= \sigma(v_j, v'_{m+1}) - \sum_{i=1}^m \sigma(v_i, v'_{m+1}) \delta_{ij} \\
&= \sigma(v_j, v'_{m+1}) - \sigma(v_j, v'_{m+1}) \\
&= 0;
\end{aligned}$$

y,

$$\begin{aligned}
\sigma(w_j, v_{m+1}) &= \sigma(w_j, v'_{m+1}) - \left(\sum_{i=1}^m \sigma(w_i, v'_{m+1}) \sigma(w_j, v_i) - \sigma(v_i, v'_{m+1}) \sigma(w_j, w_i) \right) \\
&= \sigma(w_j, v'_{m+1}) - \sum_{i=1}^m \sigma(w_i, v'_{m+1}) \delta_{ji} \\
&= \sigma(w_j, v'_{m+1}) - \sigma(w_j, v'_{m+1}) \\
&= 0.
\end{aligned}$$

Ahora, existe $w''_{m+1} \in V$ tal que $\sigma(w''_{m+1}, v_{m+1}) \neq 0$. Sea

$$w'_{m+1} = \frac{1}{\sigma(w''_{m+1}, v_{m+1})} w''_{m+1}$$

de forma que $\sigma(w'_{m+1}, v_{m+1}) = 1$. Defina

$$w_{m+1} = w'_{m+1} - \left(\sum_{i=1}^m \sigma(w_i, w'_{m+1}) v_i - \sigma(v_i, w'_{m+1}) w_i \right),$$

así $\sigma(w_{m+1}, v_{m+1}) = 1$ y al igual que con v_{m+1} , para $j = 1, \dots, m$,

$$\sigma(w_{m+1}, v_j) = 0 = \sigma(w_{m+1}, w_j).$$

En particular $\{v_1, v_2, \dots, v_{m+1}, w_1, \dots, w_{m+1}\}$ es base de Darboux de $U' = U + \text{Sp}(\{v_{m+1}, w_{m+1}\})$. Reemplazamos U por U' y continuamos recursivamente. \square

Propiedad 6.24. *Suponga que V tiene dimensión finita y sea $U < V$ subespacio symplectico. Entonces $U^\sigma < V$ es un subespacio simpléctico tal que*

$$V = U \oplus U^\sigma.$$

Dem. Sea $v \in U$. Como U es un espacio simpléctico, si $v \neq 0$, existe $w \in U$ tal que $\sigma(w, v) \neq 0$, luego $v \notin U^\sigma$. Luego

$$U \cap U^\sigma = \{0\}$$

Ahora como $\dim(U) + \dim(U^\sigma) = \dim(V)$, entonces $V = U \oplus U^\sigma$. Finalmente, sea $T = \{v_1, \dots, v_n, w_1, \dots, w_n\}$ una base de Darboux de V , tal que

$$U = \text{Sp}(\{v_j, w_j\}_{j=1, \dots, m})$$

donde $2n = \dim(V)$ y $2m = \dim(U)$, $m < n$. Entonces si

$$U' = \text{Sp}(\{v_j, w_j\}_{j=m+1, \dots, n}),$$

$\dim(U') = \dim(U^\sigma)$ y $U' \subseteq U^\sigma$, luego $U' = U^\sigma$ es un subespacio simpléctico de V . \square

Definición 6.25. Suponga que V tiene dimensión finita y sea $U \leq V$ un subespacio simpléctico de V . Llamamos a U^σ el *complemento simpléctico de U* . A la proyección

$$p_U^\sigma : V \longrightarrow V$$

sobre U , definida por la descomposición $V = U \oplus U^\sigma$ la llamamos *proyección simpléctica sobre U* .

Propiedad 6.26. Suponga que V tiene dimensión finita. Sean $V_1, V_2 < V$ subespacios lagrangianos tales que

$$V = V_1 \oplus V_2$$

y sean $p_1 : V \rightarrow V$ y $p_2 : V \rightarrow V$ las respectivas proyecciones sobre V_1 y V_2 definidas por esta descomposición. Entonces para todo $v, w \in V$,

$$\sigma(p_1(v), w) = \sigma(v, p_2(w)).$$

Por otro lado, sea $U < V$ subespacio simpléctico, entonces para todo $v, w \in V$,

$$\sigma(p_U^\sigma(v), w) = \sigma(v, p_U^\sigma(w)).$$

Dem. Sean $v_1, w_1 \in V_1$ y $v_2, w_2 \in V_2$ tales que

$$v = v_1 + w_1 \quad w = w_1 + w_2,$$

entonces

$$\begin{aligned} \sigma(p_1(v), w) &= \sigma(v_1, w_1) + \sigma(v_1, w_2) = \sigma(v_1, w_2) \\ \sigma(v, p_2(w)) &= \sigma(v_1, w_2) + \sigma(v_2, w_2) = \sigma(v_1, w_2). \end{aligned}$$

Considere ahora $v', w' \in U^\sigma$ tales que

$$v = p_U^\sigma(v) + v' \quad w = p_U^\sigma(w) + w'.$$

Entonces

$$\begin{aligned} \sigma(p_U^\sigma(v), w) &= \sigma(p_U^\sigma(v), p_U^\sigma(w)) + \sigma(p_U^\sigma(v), w') = \sigma(p_U^\sigma(v), p_U^\sigma(w)) \\ \sigma(v, p_U^\sigma(w)) &= \sigma(p_U^\sigma(v), p_U^\sigma(w)) + \sigma(v', p_U^\sigma(w)) = \sigma(p_U^\sigma(v), p_U^\sigma(w)). \end{aligned}$$

□

6.3. Operadores adjuntos

Sea V un espacio simpléctico y $f \in \text{Hom}_{\mathbb{K}}(V, V)$ un operador.

Definición 6.27. Sea $g \in \text{Hom}_{\mathbb{K}}(V, V)$, decimos que g es un *operador adjunto de f* si para todo $v, w \in V$

$$\sigma(f(v), w) = \sigma(v, g(w)).$$

Decimos que f es *auto-adjunto* si f es un operador adjunto de f .

Observación 6.28. Note que si g es adjunto de f , entonces f es adjunto de g . De hecho

$$\sigma(g(v), w) = -\sigma(w, g(v)) = -\sigma(f(w), v) = \sigma(v, f(w)).$$

Definición 6.29. Sea $n \in \mathbb{Z}_{>0}$ y $A \in M_{2n \times 2n}(K)$, definimos la *matriz adjunta simpléctica* de A por $A^{\text{sigma}} \in M_{2n \times 2n}(K)$ tal que

$$A^\sigma = J^{-1} A^\top J$$

con $J \in M_{2n \times 2n}(K)$ tal que

$$J = \begin{bmatrix} 0 & -I_n \\ I_n & 0 \end{bmatrix}$$

donde 0 denota el origen de $M_{n \times n}(K)$ y $I_n \in M_{n \times n}(K)$ es la matriz con unos en la diagonal y ceros en el resto de entradas. Es decir si $A_{11}, A_{12}, A_{21}, A_{22} \in M_{n \times n}(K)$ son tales que

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix},$$

entonces

$$A^\sigma = \begin{bmatrix} A_{22}^\top & -A_{12}^\top \\ -A_{21}^\top & A_{11}^\top \end{bmatrix}.$$

Decimos que A es *auto-adjunta simpléctica* si $A^\sigma = A$. Es decir si $A_{11} = A_{22}^\top$, $A_{12} = -A_{12}^\top$ y $A_{21} = -A_{21}^\top$.

Proposición 6.30. Suponga que V tiene dimensión finita, entonces existe un único operador $g \in \text{Hom}_K(V, V)$ adjunto de f . Más aún, si $T = \{v_1, \dots, v_n, w_1, \dots, w_n\}$ es una base de Darboux de V , entonces

$$[g]_T^T = \left([f]_T^T \right)^\sigma$$

Dem. Defina el operador $g \in \text{Hom}_K(V, V)$ por la imagen de la base T :

$$\begin{aligned} g(v_j) &= \sum_{i=1}^n \sigma(f(w_i), v_j) v_i - \sigma(f(v_i), v_j) w_i, \\ g(w_j) &= \sum_{i=1}^n \sigma(f(w_i), w_j) v_i - \sigma(f(v_i), w_j) w_i. \end{aligned}$$

De esta forma

$$\begin{aligned} \sigma(v_i, g(v_j)) &= \sigma(f(v_i), v_j) \\ \sigma(v_i, g(w_j)) &= \sigma(f(v_i), w_j) \\ \sigma(w_i, g(v_j)) &= \sigma(f(w_i), v_j) \\ \sigma(w_i, g(w_j)) &= \sigma(f(w_i), w_j) \end{aligned}$$

y por Propiedad 6.21

$$\begin{aligned}
\sigma(v, g(w)) &= \sum_{i=1}^n \sigma(v_i, g(w)) \sigma(w_i, v) - \sigma(v_i, v) \sigma(w_i, g(w)) \\
&= \sum_{i,j=1}^n \left(\sigma(w_j, w) \sigma(v_i, g(v_j)) - \sigma(v_j, w) \sigma(v_i, g(w_j)) \right) \sigma(w_i, v) \\
&\quad - \sigma(v_i, v) \left(\sigma(w_j, w) \sigma(w_i, g(v_j)) - \sigma(v_j, w) \sigma(w_i, g(w_j)) \right) \\
&= \sum_{i,j=1}^n \left(\sigma(w_j, w) \sigma(f(v_i), v_j) - \sigma(v_j, w) \sigma(f(v_i), w_j) \right) \sigma(w_i, v) \\
&\quad - \sigma(v_i, v) \left(\sigma(w_j, w) \sigma(f(w_i), v_j) - \sigma(v_j, w) \sigma(f(w_i), w_j) \right) \\
&= \sum_{i,j=1}^n \sigma(v_j, w) \left(\sigma(v_i, v) \sigma(f(w_i), w_j) - \sigma(w_i, v) \sigma(f(v_i), w_j) \right) \\
&\quad - \left(\sigma(v_i, v) \sigma(f(w_i), v_j) - \sigma(w_i, v) \sigma(f(v_i), v_j) \right) \sigma(w_j, w) \\
&= \sum_{i,j=1}^n \sigma(v_j, w) \left(\sigma(w_i, v) \sigma(w_j, f(v_i)) - \sigma(v_i, v) \sigma(w_j, f(w_i)) \right) \\
&\quad - \left(\sigma(w_i, v) \sigma(v_j, f(v_i)) - \sigma(v_i, v) \sigma(v_j, f(w_i)) \right) \sigma(w_j, w) \\
&= \sum_{j=1}^n \sigma(v_j, w) \sigma(w_j, f(v)) - \sigma(v_j, f(v)) \sigma(w_j, w) \\
&= \sigma(f(v), w)
\end{aligned}$$

Por otro lado si, $h \in \text{Hom}_{\mathbb{K}}(V, V)$ es adjunto de f , por Propiedad 5.12,

$$\begin{aligned}
h(v_j) &= \sum_{i=1}^n \sigma(w_i, h(v_j)) v_i - \sigma(v_i, h(v_j)) w_i \\
&= \sum_{i=1}^n \sigma(f(w_i), v_j) v_i - \sigma(f(v_i), v_j) w_i \\
&= g(u_j), \\
h(w_j) &= \sum_{i=1}^n \sigma(w_i, h(w_j)) v_i - \sigma(v_i, h(w_j)) w_i \\
&= \sum_{i=1}^n \sigma(f(w_i), w_j) v_i - \sigma(f(v_i), w_j) w_i \\
&= g(w_j).
\end{aligned}$$

luego $h = g$.

Ahora, para ver que la representación matricial de g respecto a T es la adjunta

simpléctica de la de f basta observar que para $i, j = 1, \dots, n$

$$\begin{aligned}
\begin{bmatrix} g \end{bmatrix}_{T, (i, j)}^T &= \begin{bmatrix} g(v_j) \end{bmatrix}_i^T \\
&= \sigma(w_i, g(v_j)) \\
&= \sigma(f(w_i), v_j) \\
&= -\sigma(v_j, f(w_i)) \\
&= \begin{bmatrix} f(w_i) \end{bmatrix}_{n+j}^T \\
&= \begin{bmatrix} f \end{bmatrix}_{T, (n+j, n+i)}^T, \\
\begin{bmatrix} g \end{bmatrix}_{T, (n+i, n+j)}^T &= \begin{bmatrix} g(w_j) \end{bmatrix}_{n+i}^T \\
&= -\sigma(v_i, g(w_j)) \\
&= -\sigma(f(v_i), w_j) \\
&= \sigma(w_j, f(v_i)) \\
&= \begin{bmatrix} f(v_i) \end{bmatrix}_j^T \\
&= \begin{bmatrix} f \end{bmatrix}_{T, (j, i)}^T, \\
\begin{bmatrix} g \end{bmatrix}_{T, (n+i, j)}^T &= \begin{bmatrix} g(v_j) \end{bmatrix}_{n+i}^T \\
&= -\sigma(v_i, g(v_j)) \\
&= -\sigma(f(v_i), v_j) \\
&= \sigma(v_j, f(v_i)) \\
&= -\begin{bmatrix} f(v_i) \end{bmatrix}_{n+j}^T \\
&= -\begin{bmatrix} f \end{bmatrix}_{T, (n+j, i)}^T, \\
\begin{bmatrix} g \end{bmatrix}_{T, (i, n+j)}^T &= \begin{bmatrix} g(w_j) \end{bmatrix}_i^T \\
&= \sigma(w_i, g(w_j)) \\
&= \sigma(f(w_i), w_j) \\
&= -\sigma(w_j, f(w_i)) \\
&= -\begin{bmatrix} f(w_i) \end{bmatrix}_j^T \\
&= -\begin{bmatrix} f \end{bmatrix}_{T, (j, n+i)}^T.
\end{aligned}$$

□

Notación 6.31. Si V tiene dimensión finita, a la adjunta de f la denotaremos por f^* .

Observación 6.32. Note que si V tiene dimensión finita, para todo $v, w \in V$,

$$\begin{aligned} f^*(s(v))(w) &= s(v)(f(w)) \\ &= \sigma(v, f(w)) \\ &= \sigma(f^*(v), w) \\ &= s(f^*(v))(w) \end{aligned}$$

luego

$$f^* \circ s = s \circ f^*$$

donde a la izquierda en la igualdad tenemos el dual y a la derecha el adjunto.

Observación 6.33. Si V tiene dimensión finita $f^* \circ f$ es auto-adjunta, de hecho para todo $v, w \in V$

$$\sigma(v, f^* \circ f(w)) = \sigma(f(v), f(w)) = \sigma(f^* \circ f(v), w).$$

Proposición 6.34. Si V tiene dimensión finita, las siguientes propiedades son equivalentes:

1. f es auto-adjunta; y,
2. la representación matricial de f respecto a una base de Darboux es auto-adjunta symplética.

Dem. Proposición 6.30 implica que si f es auto-adjunta, su representación matricial respecto a una base de Darboux es auto-adjunta symplética. Para establecer el converso, tomamos una base de Darboux $T = \{v_1, \dots, v_n, w_1, \dots, w_n\}$ asu-

mimos que $\begin{bmatrix} f \end{bmatrix}_T^T$ es auto-adjunta simpléctica, es decir para todo $i, j \in \{1, \dots, n\}$

$$\begin{aligned}
 \sigma(w_i, f(v_j)) &= \begin{bmatrix} f \end{bmatrix}_{T, (i, j)}^T \\
 &= \begin{bmatrix} f \end{bmatrix}_{T, (n+j, n+i)}^T \\
 &= -\sigma(v_j, f(w_i)) \\
 -\sigma(v_i, f(w_j)) &= \begin{bmatrix} f \end{bmatrix}_{T, (n+i, n+j)}^T \\
 &= \begin{bmatrix} f \end{bmatrix}_{T, (j, i)}^T \\
 &= \sigma(w_j, f(v_i)) \\
 -\sigma(v_i, f(v_j)) &= \begin{bmatrix} f \end{bmatrix}_{T, (n+i, j)}^T \\
 &= -\begin{bmatrix} f \end{bmatrix}_{T, (n+j, i)}^T \\
 &= \sigma(v_j, f(v_i)) \\
 \sigma(w_i, f(w_j)) &= \begin{bmatrix} f \end{bmatrix}_{T, (i, n+j)}^T \\
 &= -\begin{bmatrix} f \end{bmatrix}_{T, (j, n+i)}^T \\
 &= -\sigma(w_j, f(w_i)),
 \end{aligned}$$

en particular

$$\begin{aligned}
 \sigma(v_i, f(v_j)) &= \sigma(f(v_i), v_j) \\
 \sigma(v_i, f(w_j)) &= \sigma(f(v_i), w_j) \\
 \sigma(w_i, f(v_j)) &= \sigma(f(w_i), v_j) \\
 \sigma(w_i, f(w_j)) &= \sigma(f(w_i), w_j).
 \end{aligned}$$

Así, por la demostración de Proposición 6.30, estas igualdades implican que el adjunto de f es él mismo. \square

Ejemplo 6.35. Suponga que $V = U \times U^*$ y

$$\sigma((v, \lambda), (w, \mu)) = \lambda(w) - \mu(v).$$

Sea ahora $g \in \text{Hom}_K(U, U)$ y tome

$$f(v, \lambda) = (g(v), g^*(\lambda))$$

de forma que

$$\begin{aligned}
 \sigma((v, \lambda), f(w, \mu)) &= \lambda(g(w)) - g^*(\mu)(v) \\
 &= g^*(\lambda)(w) - \mu(g(v)) \\
 &= \sigma(f(v, \lambda), (w, \mu)),
 \end{aligned}$$

luego f es auto-adjunto.

Observación 6.36. El operador del ejemplo anterior es de hecho la forma más general de operador auto-adjunto sobre un espacio simpléctico. Es decir, dado un operador auto-adjunto, existe una descomposición del espacio en subespacios invariantes compatibles con el operador tales que este toma la forma como el operador f en Ejemplo 6.35. El resto de este capítulo tiene como objetivo establecer ese resultado para el caso en el que el polinomio característico del operador se factoriza en factores lineales en K .

Lema 6.37. *Suponga que V tiene dimensión finita y que f es auto-adjunto, entonces:*

1. *si f es una proyección, es decir $f^2 = f$, entonces $f(V)$ es un subespacio simpléctico;*
2. *para todo $P(t) \in K[t]$, $P(f)$ es auto-adjunto.*

Dem.

1. Como f es una proyección $V = f(V) \oplus \ker(f)$, y para $v \in V$

$$v = f(v) + (v - f(v))$$

con $f(v) \in f(V)$ y $v - f(v) \in \ker(f)$. Para probar el lema basta con establecer que $\ker(f) = f(V)^\sigma$, pues en tal caso, como $f(V) \cap \ker(f) = \{0\}$ tendríamos $f(V) \cap f(V)^\sigma = \{0\}$, y la conclusión se sigue de Proposición 6.13.1. Ahora, dado $v \in \ker(f)$, para todo $w \in V$,

$$\sigma(f(w), v) = \sigma(w, f(v)) = 0$$

luego $v \in f(V)^\sigma$, y así $\ker(f) \subseteq f(V)^\sigma$. Pero $\dim(\ker(f)) = V - \dim(f(V)) = \dim(f(V)^\sigma)$, entonces $\ker(f) = f(V)^\sigma$.

2. Si $P(t) = \sum_{i=0}^d a_i t^i$, para todo $v, w \in V$ tenemos

$$\begin{aligned} \sigma(v, P(f)(w)) &= \sum_{i=0}^d a_i \sigma(v, f^i(w)) \\ &= \sum_{i=1}^d a_i \sigma(f^i(v), w) \\ &= \sigma(P(f)(v), w). \end{aligned}$$

□

Proposición 6.38. *Suponga que V tiene dimensión finita, que f es auto-adjunto y que*

$$P_f(t) = (t - \lambda_1)^{m_1} (t - \lambda_2)^{m_2} \dots (t - \lambda_r)^{m_r}, \quad \lambda_1, \lambda_2, \dots, \lambda_r \in \mathbb{K}.$$

con $\lambda_i \neq \lambda_j$ si $i \neq j$. Entonces para $i = 1, \dots, r$, $V_i = \ker(P_i(f)^{m_i})$ es un subespacio simpléctico invariante bajo f , y

$$V = V_1 \oplus \dots \oplus V_r.$$

Dem. La descomposición $V = V_1 \oplus \dots \oplus V_r$ como suma directa de subespacios invariantes bajo f es consecuencia directa de Propiedad ??, y al considerar también la afirmación 2. del lema anterior obtenemos que la proyección sobre cada uno de estos subespacios es auto-adjunta. La primera afirmación del mismo lema implica que cada uno de estos V_i , $i = 1, \dots, r$, es un subespacio simpléctico. \square

Capítulo 7

Álgebra Multilineal

Sea K un cuerpo y V, W espacios vectoriales sobre K .

Notación 7.1. Dado $k \in \mathbb{Z}_{>0}$ denotamos

$$V^k = \underbrace{V \times \cdots \times V}_{k\text{-veces}}$$

y para $k = 0$ usamos la convención $V^0 = K$.

7.1. Tensores

Definición 7.2. Sea $k \in \mathbb{Z}_{\geq 0}$ y $T : V^k \longrightarrow K$ una función. Decimos que T es *multilineal* si para $i = 1 \dots k$ tenemos

$$T(v_1, \dots, c_i v_i + c'_i v'_i, \dots, v_k) = c_i T(v_1, \dots, v_i, \dots, v_k) + c'_i T(v_1, \dots, v'_i, \dots, v_k)$$

para todo $v_1, \dots, v_i, v'_i, \dots, v_k \in V$ y $c_i, c'_i \in K$. Si T es multilineal decimos que T es un k -tensor. Al conjunto de k -tensores lo denotamos por $T^k(V)$.

Observación 7.3. Para todo $k \in \mathbb{Z}_{\geq 0}$ el conjunto $T^k(V)$ es un espacio vectorial sobre K bajo las operaciones:

$$\begin{aligned} S + T : \quad V^k &\longrightarrow K \\ (v_1, \dots, v_k) &\longmapsto S(v_1, \dots, v_k) + T(v_1, \dots, v_k) \\ \\ cT : \quad V^k &\longrightarrow K \\ (v_1, \dots, v_k) &\longmapsto cT(v_1, \dots, v_k) \end{aligned}$$

para todo $S, T \in T^k(V)$ y $c \in K$.

Observación 7.4. $T^0(V) \simeq K$ y $T^1(V) = V^*$.

Ejemplo 7.5. Sea $n \in \mathbb{Z}_{>0}$.

1. La función

$$\begin{aligned} \langle \bullet; \bullet \rangle : K^n \times K^n &\longrightarrow K \\ ((x_1, \dots, x_n), (y_1, \dots, y_n)) &\longmapsto \sum_{i=1}^n x_i y_i \end{aligned}$$

define un 2-tensor.

2. Sea $a_{ij} \in K$, $i, j = 1, \dots, n$. La función

$$\begin{aligned} T : K^n \times K^n &\longrightarrow K \\ ((x_1, \dots, x_n), (y_1, \dots, y_n)) &\longmapsto \sum_{i,j=1}^n x_i a_{ij} y_j \end{aligned}$$

define un 2-tensor.

3. La función

$$\begin{aligned} \det : (K^n)^n &\longrightarrow K \\ (\bar{x}_1, \dots, \bar{x}_n) &\longmapsto \det(x_{ij}) = \sum_{\sigma \in S_n} \text{sgn}(\sigma) x_{\sigma(1)1} \dots x_{\sigma(n)n} \end{aligned}$$

donde $\bar{x}_j = (x_{1j}, \dots, x_{nj})$ para $j = 1, \dots, n$ define un n -tensor.

Definición 7.6. Sean $k, l \in \mathbb{Z}_{\geq 0}$. Dados $S \in T^k(V)$ y $T \in T^l(V)$ definimos su *producto tensorial* $S \otimes T \in T^{k+l}(V)$ por

$$S \otimes T(v_1, \dots, v_k, v_{k+1}, \dots, v_{k+l}) = S(v_1, \dots, v_k)T(v_{k+1}, \dots, v_{k+l})$$

para todo $v_1, \dots, v_{k+l} \in V$.

Observación 7.7. El producto tensorial no es conmutativo pues no siempre es cierto que $S \otimes T$ y $T \otimes S$ coincidan.

Propiedad 7.8. El producto tensorial es bilineal y asociativo. Es decir, dados $k, l, m \in \mathbb{Z}_{\geq 0}$ y $S, S' \in T^k(V)$, $T, T' \in T^l(V)$, $U \in T^m(V)$, $c \in K$ tenemos:

1. $(S + S') \otimes T = S \otimes T + S' \otimes T$,
2. $S \otimes (T + T') = S \otimes T + S \otimes T'$,
3. $(cS) \otimes T = c(S \otimes T) = S \otimes (cT)$,
4. $(S \otimes T) \otimes U = S \otimes (T \otimes U)$.

Dem. La demostración es una verificación directa. □

Notación 7.9. Por la Proposición 7.8 4. denotamos $S \otimes T \otimes U = (S \otimes T) \otimes U$ y así podemos definir producto tensorial de más de dos tensores.

Teorema 7.10. Suponga que $\dim_K(V) = n < \infty$. Sea $\{v_1, \dots, v_n\}$ una base de V y $\{\lambda_1, \dots, \lambda_n\}$ la base dual. Para todo $k \in \mathbb{Z}_{>0}$, la colección de k -tensores

$$\left\{ \lambda_{i_1} \otimes \dots \otimes \lambda_{i_k} \right\}_{i_1, \dots, i_k=1}^n$$

es una base de $T^k(V)$. En particular

$$\dim_K(T^k) = n^k.$$

Dem. Veamos primero que

$$T^k(V) = \langle \lambda_{i_1} \otimes \cdots \otimes \lambda_{i_k} \rangle_{i_1, \dots, i_k=1}^n.$$

Note que para todo $i_1, \dots, i_k, j_1, \dots, j_k \in \{1, \dots, n\}$

$$\lambda_{i_1} \otimes \cdots \otimes \lambda_{i_k}(v_{j_1}, \dots, v_{j_k}) = \delta_{i_1 j_1} \cdots \delta_{i_k j_k} = \begin{cases} 1 & \text{si } i_1 = j_1, \dots, i_k = j_k \\ 0 & \text{si no} \end{cases}.$$

Así, si $w_1, \dots, w_k \in V$ y $w_j = \sum_{i=1}^n a_{ij} v_i$, $a_{ij} \in K$, $j = 1, \dots, k$ entonces

$$\begin{aligned} \lambda_{i_1} \otimes \cdots \otimes \lambda_{i_k}(w_1, \dots, w_k) &= \sum_{j_1, \dots, j_k=1}^n a_{j_1 1} \cdots a_{j_k k} \lambda_{i_1} \otimes \cdots \otimes \lambda_{i_k}(v_{j_1}, \dots, v_{j_k}) \\ &= \sum_{j_1, \dots, j_k=1}^n a_{j_1 1} \cdots a_{j_k k} \delta_{i_1 j_1} \cdots \delta_{i_k j_k} \\ &= a_{i_1 1} \cdots a_{i_k k} \end{aligned}$$

Ahora, dado $T \in T^k(V)$

$$\begin{aligned} T(w_1, \dots, w_k) &= \sum_{i_1, \dots, i_k=1}^n a_{i_1 1} \cdots a_{i_k k} T(v_{i_1}, \dots, v_{i_k}) \\ &= \sum_{i_1, \dots, i_k=1}^n T(v_{i_1}, \dots, v_{i_k}) \lambda_{i_1} \otimes \cdots \otimes \lambda_{i_k}(w_1, \dots, w_k) \end{aligned}$$

luego

$$T = \sum_{i_1, \dots, i_k=1}^n T(v_{i_1}, \dots, v_{i_k}) \lambda_{i_1} \otimes \cdots \otimes \lambda_{i_k},$$

y $T^k(V) = \langle \lambda_{i_1} \otimes \cdots \otimes \lambda_{i_k} \rangle_{i_1, \dots, i_k=1}^n$.

Establezcamos ahora la independencia lineal. Suponga que

$$0 = \sum_{i_1, \dots, i_k=1}^n c_{i_1 \dots i_k} \lambda_{i_1} \otimes \cdots \otimes \lambda_{i_k}.$$

Si evaluamos en $(v_{j_1}, \dots, v_{j_k})$ obtenemos

$$\begin{aligned} 0 &= \sum_{i_1, \dots, i_k=1}^n c_{i_1 \dots i_k} \lambda_{i_1} \otimes \cdots \otimes \lambda_{i_k}(v_{j_1}, \dots, v_{j_k}) \\ &= c_{j_1, \dots, j_k}. \end{aligned}$$

□

Ejemplo 7.11. Sea $n \in \mathbb{Z}_{>0}$, $\{e_i\}_{i=1}^n$ la base canónica de K^n y $\{f_1, \dots, f_n\}$ la base dual.

1. Si

$$\begin{aligned} \langle \bullet; \bullet \rangle : K^n \times K^n &\longrightarrow K \\ \left((x_1, \dots, x_n), (y_1, \dots, y_n) \right) &\longmapsto \sum_{i=1}^n x_i y_i \end{aligned}$$

entonces $\langle \bullet; \bullet \rangle = \sum_{i=1}^n f_i \otimes f_i$.

2. Sea $a_{ij} \in K$, $i, j = 1, \dots, n$. Si

$$\begin{aligned} T : K^n \times K^n &\longrightarrow K \\ \left((x_1, \dots, x_n), (y_1, \dots, y_n) \right) &\longmapsto \sum_{i,j=1}^n x_i a_{ij} y_j \end{aligned}$$

entonces $T = \sum_{i,j=1}^n a_{ij} f_i \otimes f_j$.

3. Si

$$\begin{aligned} \det : (K^n)^n &\longrightarrow K \\ (\bar{x}_1, \dots, \bar{x}_n) &\longmapsto \sum_{\sigma \in S_n} \text{sgn}(\sigma) x_{\sigma(1)1} \dots x_{\sigma(n)n} \end{aligned}$$

donde $\bar{x}_j = (x_{1j}, \dots, x_{nj})$ para $j = 1, \dots, n$ entonces

$$\det = \sum_{\sigma \in S_n} \text{sgn}(\sigma) f_{\sigma(1)} \otimes \dots \otimes f_{\sigma(n)}.$$

Definición 7.12. Sea $f \in \text{Hom}_K(V, W)$. Dado $k \in \mathbb{Z}_{\geq 0}$ definimos

$$\begin{aligned} f^* : T^k(W) &\longrightarrow T^k(V) \\ T &\longmapsto f^*T : (v_1, \dots, v_k) \mapsto T(f(v_1), \dots, f(v_k)). \end{aligned}$$

Proposición 7.13. Sea $f \in \text{Hom}_K(V, W)$ y $k, l \in \mathbb{Z}_{\geq 0}$. Para todo $S \in T^k(W)$ y $T \in T^l(W)$ tenemos

$$f^*(S \otimes T) = f^*S \otimes f^*T.$$

Dem. La demostración es una verificación inmediata. □

7.2. Tensores alternantes

Definición 7.14. Sea $k \in \mathbb{Z}_{\geq 0}$ y $\omega \in T^k(V)$, decimos que ω es *alternante* si

$$\omega(v_1, \dots, v_i, \dots, v_j, \dots, v_k) = 0, \quad v_1, \dots, v_k \in V$$

siempre que $v_i = v_j$ para algún par $i, j \in \{1, \dots, k\}$, $i \neq j$. Denotamos por $\Lambda^k(V)$ al subespacio de $T^k(V)$ de k -tensores alternantes.

Propiedad 7.15. Sea $k \in \mathbb{Z}_{\geq 0}$ y $\omega \in T^k(V)$ un k -tensor alternante, entonces

$$\omega(v_1, \dots, v_i, \dots, v_j, \dots, v_k) = -\omega(v_1, \dots, v_j, \dots, v_i, \dots, v_k)$$

para todo $v_1, \dots, v_k \in V$ y todo $1 \leq i < j \leq k$.

Dem.

$$\begin{aligned}
 0 &= \omega(v_1, \dots, v_i + v_j, \dots, v_i + v_j, \dots, v_k) \\
 &= \omega(v_1, \dots, v_i, \dots, v_i, \dots, v_k) + \omega(v_1, \dots, v_i, \dots, v_j, \dots, v_k) \\
 &\quad + \omega(v_1, \dots, v_j, \dots, v_i, \dots, v_k) + \omega(v_1, \dots, v_j, \dots, v_j, \dots, v_k) \\
 &= \omega(v_1, \dots, v_i, \dots, v_j, \dots, v_k) + \omega(v_1, \dots, v_j, \dots, v_i, \dots, v_k)
 \end{aligned}$$

□

Observación 7.16. Sea $k \in \mathbb{Z}_{\geq 0}$ y $\omega \in \Lambda^k(V)$, para toda permutación $\sigma \in S_k$ tenemos

$$\omega(v_{\sigma(1)}, \dots, v_{\sigma(k)}) = -\operatorname{sgn}(\sigma)\omega(v_1, \dots, v_k).$$

Esta última igualdad es la que justifica el nombre de alternante.

Observación 7.17. $\Lambda^0(V) = T^0(V) \simeq K$ y $\Lambda^1(V) = T^1(V) = V^*$.

Ejemplo 7.18. Sea $n \in \mathbb{Z}_{>0}$.

1. La función determinante

$$\begin{aligned}
 \det : \quad & \left(K^n\right)^n \longrightarrow K \\
 & (\bar{x}_1, \dots, \bar{x}_n) \longmapsto \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) x_{\sigma(1)1} \dots x_{\sigma(n)n}
 \end{aligned}$$

donde $\bar{x}_j = (x_{1j}, \dots, x_{nj})$ para $j = 1, \dots, n$ define un n -tensor alternante.

2. Sea $k \in \{1, \dots, n\}$ y $1 \leq i_1 < \dots < i_k \leq n$. La función determinante menor

$$\begin{aligned}
 \det_{i_1 \dots i_k} : \quad & \left(K^n\right)^k \longrightarrow K \\
 & (\bar{x}_1, \dots, \bar{x}_k) \longmapsto \sum_{\sigma \in S_k} \operatorname{sgn}(\sigma) x_{i_{\sigma(1)}1} \dots x_{i_{\sigma(k)}k}
 \end{aligned}$$

donde $\bar{x}_j = (x_{1j}, \dots, x_{nj})$ para $j = 1, \dots, n$ define un k -tensor alternante.

Observación 7.19. Sea $k, l \in \mathbb{Z}_{\geq 0}$ y $\omega \in \Lambda^k(V)$, $\eta \in \Lambda^l(V)$, el $k + l$ -tensor $\omega \otimes \eta$ no es necesariamente alternante. Para obtener un tensor alternante hace falta proyectar sobre el subespacio $\Lambda^{k+l}(V) \leq T^{k+l}(V)$.

Definición 7.20. Sea $k \in \mathbb{Z}_{\geq 0}$ tal que si $\operatorname{char}(K) > 0$ entonces $k < \operatorname{char}(K)$. Definimos $\operatorname{Alt} \in \operatorname{Hom}_K(T^k(V), T^k(V))$ por

$$\operatorname{Alt}(T)(v_1, \dots, v_k) = \frac{1}{k!} \sum_{\sigma \in S_k} \operatorname{sgn}(\sigma) T(v_{\sigma(1)}, \dots, v_{\sigma(k)})$$

Ejemplo 7.21. Sea $n \in \mathbb{Z}_{>0}$, $\{e_i\}_{i=1}^n$ la base canónica de K^n y $\{f_1, \dots, f_n\}$ la base dual.

1. Para $\text{char}(K) > 0$ suponga que $n < \text{char}(K)$. Si $T = f_1 \otimes \cdots \otimes f_n$, tenemos que para todo $\bar{x}_j = (x_{1j}, \dots, x_{nj}) \in K^n$, $j = 1, \dots, n$,

$$\begin{aligned}
 \text{Alt}(T)(\bar{x}_1, \dots, \bar{x}_n) &= \frac{1}{n!} \sum_{\sigma \in S_n} \text{sgn}(\sigma) f_1 \otimes \cdots \otimes f_n(v_{\sigma(1)}, \dots, v_{\sigma(n)}) \\
 &= \frac{1}{n!} \sum_{\sigma \in S_n} \text{sgn}(\sigma) x_{1\sigma(1)} \cdots x_{n\sigma(n)} \\
 &= \frac{1}{n!} \sum_{\sigma \in S_n} \text{sgn}(\sigma) x_{\sigma^{-1}(1)1} \cdots x_{\sigma^{-1}(n)n} \\
 &= \frac{1}{n!} \sum_{\sigma \in S_n} \text{sgn}(\sigma^{-1}) x_{\sigma^{-1}(1)1} \cdots x_{\sigma^{-1}(n)n} \\
 &= \frac{1}{n!} \sum_{\sigma \in S_n} \text{sgn}(\sigma) x_{\sigma(1)1} \cdots x_{\sigma(n)n} \\
 &= \frac{1}{n!} \det(\bar{x}_1, \dots, \bar{x}_n)
 \end{aligned}$$

así $\text{Alt}(f_1 \otimes \cdots \otimes f_n) = \frac{1}{n!} \det$.

2. Sea $k \in \{1, \dots, n\}$ y $1 \leq i_1 < \dots < i_k \leq n$. Para $\text{char}(K) > 0$ suponga que $k < \text{char}(K)$. Si $T = f_{i_1} \otimes \cdots \otimes f_{i_k}$, tenemos que para todo $\bar{x}_j = (x_{1j}, \dots, x_{nj}) \in K^n$, $j = 1, \dots, k$,

$$\begin{aligned}
 \text{Alt}(T)(\bar{x}_1, \dots, \bar{x}_k) &= \frac{1}{k!} \sum_{\sigma \in S_k} \text{sgn}(\sigma) f_{i_1} \otimes \cdots \otimes f_{i_k}(v_{\sigma(1)}, \dots, v_{\sigma(k)}) \\
 &= \frac{1}{k!} \sum_{\sigma \in S_k} \text{sgn}(\sigma) x_{i_1\sigma(1)} \cdots x_{i_k\sigma(k)} \\
 &= \frac{1}{k!} \sum_{\sigma \in S_k} \text{sgn}(\sigma) x_{i_{\sigma^{-1}(1)}1} \cdots x_{i_{\sigma^{-1}(k)}k} \\
 &= \frac{1}{k!} \sum_{\sigma \in S_k} \text{sgn}(\sigma^{-1}) x_{i_{\sigma^{-1}(1)}1} \cdots x_{i_{\sigma^{-1}(k)}k} \\
 &= \frac{1}{k!} \sum_{\sigma \in S_k} \text{sgn}(\sigma) x_{i_{\sigma(1)}1} \cdots x_{i_{\sigma(k)}k} \\
 &= \frac{1}{k!} \det_{i_1 \dots i_k}(\bar{x}_1, \dots, \bar{x}_k)
 \end{aligned}$$

así $\text{Alt}(f_{i_1} \otimes \cdots \otimes f_{i_k}) = \frac{1}{k!} \det_{i_1 \dots i_k}$.

Proposición 7.22. Para todo $k \in \mathbb{Z}_{\geq 0}$ el operador Alt es una proyección sobre $\Lambda^k(V)$. Es decir:

1. para todo $T \in T^k(V)$, $\text{Alt}(T) \in \Lambda^k(V)$,
2. para todo $\omega \in \Lambda^k(V)$, $\text{Alt}(\omega) = \omega$,

3. $\text{Alt} \circ \text{Alt} = \text{Alt}$.

Dem.

1. Dados $1 \leq i < j \leq k$, defina $\tau \in S_k$ la transposición que intercambia i y j (y deja al resto de elementos en $\{1, \dots, k\}$ fijos). Sea A_k el subgrupo de S_k formado por las permutaciones con signo 1, de forma que si $\tau A_k = \{\tau\sigma \in S_k \mid \sigma \in A_k\}$ entonces $S_k = A_k \cup \tau A_k$ es una partición de S_k . Nota que $\text{sgn}(\tau\sigma) = -\text{sgn}(\sigma)$ para todo $\sigma \in S_k$. Sean $v_1, \dots, v_k \in V$ tales que $v_i = v_j$, entonces para todo $\sigma \in S_k$ tenemos

$$(v_{\sigma(1)}, \dots, v_{\sigma(k)}) = (v_{\tau\sigma(1)}, \dots, v_{\tau\sigma(k)})$$

y así

$$\begin{aligned} \text{Alt}(T)(v_1, \dots, v_k) &= \frac{1}{k!} \sum_{\sigma \in S_k} \text{sgn}(\sigma) T(v_{\sigma(1)}, \dots, v_{\sigma(k)}) \\ &= \frac{1}{k!} \left(\sum_{\sigma \in A_k} \text{sgn}(\sigma) T(v_{\sigma(1)}, \dots, v_{\sigma(k)}) + \sum_{\sigma \in \tau A_k} \text{sgn}(\sigma) T(v_{\sigma(1)}, \dots, v_{\sigma(k)}) \right) \\ &= \frac{1}{k!} \left(\sum_{\sigma \in A_k} \text{sgn}(\sigma) T(v_{\sigma(1)}, \dots, v_{\sigma(k)}) + \sum_{\sigma \in A_k} \text{sgn}(\tau\sigma) T(v_{\tau\sigma(1)}, \dots, v_{\tau\sigma(k)}) \right) \\ &= \frac{1}{k!} \left(\sum_{\sigma \in A_k} \text{sgn}(\sigma) T(v_{\sigma(1)}, \dots, v_{\sigma(k)}) - \sum_{\sigma \in A_k} \text{sgn}(\sigma) T(v_{\tau\sigma(1)}, \dots, v_{\tau\sigma(k)}) \right) \\ &= 0 \end{aligned}$$

luego $\text{Alt}(T) \in \Lambda^k(V)$.

2. Sea $\omega \in \Lambda^k(V)$, entonces para todo $v_1, \dots, v_k \in V$ tenemos que

$$\begin{aligned} \text{Alt}(\omega)(v_1, \dots, v_k) &= \frac{1}{k!} \sum_{\sigma \in S_k} \text{sgn}(\sigma) \omega(v_{\sigma(1)}, \dots, v_{\sigma(k)}) \\ &= \frac{1}{k!} \sum_{\sigma \in S_k} \text{sgn}(\sigma) \text{sgn}(\sigma) \omega(v_1, \dots, v_k) \\ &= \frac{1}{k!} \sum_{\sigma \in S_k} \omega(v_1, \dots, v_k) \\ &= \frac{1}{k!} k! \omega(v_1, \dots, v_k) \\ &= \omega(v_1, \dots, v_k) \end{aligned}$$

luego $\text{Alt}(\omega) = \omega$.

3. Se sigue inmediatamente de 1. y 2.

□

Definición 7.23. Sea $k, l \in \mathbb{Z}_{\geq 0}$ tal que si $\text{char}(K) > 0$ entonces $k + l < \text{char}(K)$. Sea $\omega \in \Lambda^k V$ y $\eta \in \Lambda^l(V)$, definimos el *producto exterior* $\omega \wedge \eta \in \Lambda^{k+l}(V)$ por

$$\omega \wedge \eta = \frac{(k+l)!}{k! l!} \text{Alt}(\omega \otimes \eta).$$

Lema 7.24. Sea $k, l, m \in \mathbb{Z}_{\geq 0}$ tal que si $\text{char}(K) > 0$ entonces $k + l + m < \text{char}(K)$. Sea $S \in T^k(V)$, $T \in T^l(V)$ y $U \in T^m(V)$ tenemos que

1. si $\text{Alt}(S) = 0$ entonces $\text{Alt}(S \otimes T) = 0 = \text{Alt}(T \otimes S)$,
2. $\text{Alt}(\text{Alt}(S \otimes T) \otimes U) = \text{Alt}(S \otimes T \otimes U) = \text{Alt}(S \otimes \text{Alt}(T \otimes U))$

Dem.

1. Tome S_k como el subgrupo de S_{k+l} formado por las permutaciones que dejan fijos a $k+1, \dots, k+l$ y sean $\sigma_1, \dots, \sigma_N \in S_{k+l}$, $N = (k+l)/(k!)$, representantes de los coconjuntos $S_k \sigma = \{\tau \sigma \in S_{k+l} \mid \tau \in S_k\}$, $\sigma \in S_{k+l}$, de forma que

$$S_{k+l} = \cup_{i=1}^N \sigma_i S_k$$

es una partición de S_{k+l} . Entonces para todo $v_1, \dots, v_{k+l} \in V$

$$\begin{aligned} & \text{Alt}(S \otimes T)(v_1, \dots, v_{k+l}) \\ &= \frac{1}{(k+l)!} \sum_{\sigma \in S_{k+l}} \text{sgn}(\sigma) S \otimes T(v_{\sigma(1)}, \dots, v_{\sigma(k+l)}) \\ &= \frac{1}{(k+l)!} \sum_{i=1}^N \sum_{\tau \in S_k} \text{sgn}(\tau \sigma_i) S \otimes T(v_{\tau \sigma_i(1)}, \dots, v_{\tau \sigma_i(k+l)}) \\ &= \frac{1}{(k+l)!} \sum_{i=1}^N \text{sgn}(\sigma_i) \sum_{\tau \in S_k} \text{sgn}(\tau) S(v_{\tau \sigma_i(1)}, \dots, v_{\tau \sigma_i(k)}) T(v_{\tau \sigma_i(k+1)}, \dots, v_{\tau \sigma_i(k+l)}). \end{aligned}$$

Para $i = 1, \dots, N$ y $j = 1, \dots, k+l$, denote $w_j^{(i)} = v_{\sigma_i(j)}$, así, como $\text{Alt} = 0$ entonces

$$\begin{aligned} & \text{Alt}(S \otimes T)(v_1, \dots, v_{k+l}) \\ &= \frac{1}{(k+l)!} \sum_{i=1}^N \text{sgn}(\sigma_i) \sum_{\tau \in S_k} \text{sgn}(\tau) S(w_{\tau(1)}^{(i)}, \dots, w_{\tau(k)}^{(i)}) T(w_{\tau(k+1)}^{(i)}, \dots, w_{\tau(k+l)}^{(i)}) \\ &= \frac{1}{(k+l)!} \sum_{i=1}^N \text{sgn}(\sigma_i) \sum_{\tau \in S_k} \text{sgn}(\tau) S(w_{\tau(1)}^{(i)}, \dots, w_{\tau(k)}^{(i)}) T(w_{k+1}^{(i)}, \dots, w_{k+l}^{(i)}) \\ &= \frac{1}{(k+l)!} \sum_{i=1}^N \text{sgn}(\sigma_i) \left(\sum_{\tau \in S_k} \text{sgn}(\tau) S(w_{\tau(1)}^{(i)}, \dots, w_{\tau(k)}^{(i)}) \right) T(w_{k+1}^{(i)}, \dots, w_{k+l}^{(i)}) \\ &= \frac{k!}{(k+l)!} \sum_{i=1}^N \text{sgn}(\sigma_i) \text{Alt}(S)(w_1^{(i)}, \dots, w_k^{(i)}) T(w_{k+1}^{(i)}, \dots, w_{k+l}^{(i)}) \\ &= 0 \end{aligned}$$

Similarmente, si tomamos S_k como el subgrupo de S_{k+l} formado por la permutaciones que dejan fijos a $1, \dots, l$, obtenemos $\text{Alt}(T \otimes S) = 0$.

2. $\text{Alt}(\text{Alt}(S \otimes T) - S \otimes T) = \text{Alt}(S \otimes T) - \text{Alt}(S \otimes T) = 0$, luego por 1.

$$\begin{aligned} 0 &= \text{Alt}((\text{Alt}(S \otimes T) - S \otimes T) \otimes U) \\ &= \text{Alt}(\text{Alt}(S \otimes T) \otimes U) - \text{Alt}(S \otimes T \otimes U) \end{aligned}$$

y así $\text{Alt}(\text{Alt}(S \otimes T) \otimes U) = \text{Alt}(S \otimes T \otimes U)$. Similarmente obtenemos $\text{Alt}(S \otimes \text{Alt}(T \otimes U)) = \text{Alt}(S \otimes T \otimes U)$.

□

Propiedad 7.25. *El producto exterior es bilineal, asociativo y anticonmutativo. Es decir, dados $k, l, m \in \mathbb{Z}_{\geq 0}$ tal que si $\text{char}(K) > 0$ entonces $k+l+m < \text{char}(K)$ y $\omega, \omega' \in \Lambda^k(V)$, $\eta, \eta' \in \Lambda^l(V)$, $\theta \in \Lambda^m(V)$, $c \in K$ tenemos:*

$$1. (\omega + \omega') \wedge \eta = \omega \wedge \eta + \omega' \wedge \eta,$$

$$2. \omega \wedge (\eta + \eta') = \omega \wedge \eta + \omega \wedge \eta',$$

$$3. (c\omega) \wedge \eta = c(\omega \wedge \eta) = \omega \wedge (c\eta),$$

$$4. \omega \wedge \eta = (-1)^{kl} \eta \wedge \omega,$$

$$5. (\omega \wedge \eta) \wedge \theta = \omega \wedge (\eta \wedge \theta) = \frac{(k+l+m)!}{k! l! m!} \text{Alt}(\omega \otimes \eta \otimes \theta).$$

Dem. Las propiedades 1., 2. y 3. se siguen inmediatamente de la bilinearidad de \otimes y de la linealidad de Alt .

4. Sea $\tau \in S_{k+l}$ definida por

$$\tau(i) = \begin{cases} i+k & \text{si } 1 \leq i \leq l \\ i-l & \text{si } l+1 \leq i \leq l+k \end{cases}$$

Como $\text{sgn}(\tau) = (-1)^{kl}$ y $\text{sgn}(\sigma\tau) = (-1)^{kl} \text{sgn}(\sigma)$ para toda $\sigma \in S_{k+l}$

entonces para todo $v_1, \dots, v_{k+l} \in V$

$$\begin{aligned}
& \omega \wedge \eta(v_1, \dots, v_{k+l}) \\
&= \frac{(k+l)!}{k! \, l!} \text{Alt}(\omega \otimes \eta)(v_1, \dots, v_{k+l}) \\
&= \frac{1}{k! \, l!} \sum_{\sigma \in S_{k+l}} \text{sgn}(\sigma) \omega(v_{\sigma(1)}, \dots, v_{\sigma(k)}, v_{\sigma(k+1)}, \dots, v_{\sigma(k+l)}) \\
&= \frac{1}{k! \, l!} \sum_{\sigma \in S_{k+l}} \text{sgn}(\sigma) \omega(v_{\sigma(1)}, \dots, v_{\sigma(k)}) \eta(v_{\sigma(k+1)}, \dots, v_{\sigma(k+l)}) \\
&= \frac{1}{k! \, l!} \sum_{\sigma \in S_{k+l}} \text{sgn}(\sigma) \omega(v_{\sigma\tau(l+1)}, \dots, v_{\sigma\tau(l+k)}) \eta(v_{\sigma\tau(1)}, \dots, v_{\sigma\tau(l)}) \\
&= \frac{1}{k! \, l!} \sum_{\sigma \in S_{k+l}} (-1)^{kl} \text{sgn}(\sigma\tau) \eta(v_{\sigma\tau(1)}, \dots, v_{\sigma\tau(l)}) \omega(v_{\sigma\tau(l+1)}, \dots, v_{\sigma\tau(l+k)}) \\
&= (-1)^{kl} \frac{1}{k! \, l!} \sum_{\sigma\tau \in S_{k+l}} \text{sgn}(\sigma\tau) \eta(v_{\sigma\tau(1)}, \dots, v_{\sigma\tau(l)}) \omega(v_{\sigma\tau(l+1)}, \dots, v_{\sigma\tau(l+k)}) \\
&= (-1)^{kl} \frac{1}{k! \, l!} \sum_{\sigma\tau \in S_{k+l}} \text{sgn}(\sigma\tau) \eta \otimes \omega(v_{\sigma\tau(1)}, \dots, v_{\sigma\tau(k)}, v_{\sigma\tau(k+1)}, \dots, v_{\sigma\tau(k+l)}) \\
&= (-1)^{kl} \frac{(k+l)!}{k! \, l!} \text{Alt}(\eta \otimes \omega)(v_1, \dots, v_{k+l}) \\
&= (-1)^{kl} \eta \wedge \omega(v_1, \dots, v_{k+l}).
\end{aligned}$$

así $\omega \wedge \eta = (-1)^{kl} \eta \wedge \omega$.

5. Por Lema 7.24

$$\begin{aligned}
(\omega \wedge \eta) \wedge \theta &= \frac{(k+l+m)!}{(k+l)! \, m!} \text{Alt}((\omega \wedge \eta) \otimes \theta) \\
&= \frac{(k+l+m)!}{(k+l)! \, m!} \text{Alt}\left(\frac{(k+l)!}{k! \, l!} \text{Alt}(\omega \otimes \eta) \otimes \theta\right) \\
&= \frac{(k+l+m)!}{k! \, l! \, m!} \text{Alt}(\omega \otimes \eta \otimes \theta)
\end{aligned}$$

Similarmente obtenemos $\omega \wedge (\eta \wedge \theta) = \frac{(k+l+m)!}{k! \, l! \, m!} \text{Alt}(\omega \otimes \eta \otimes \theta)$.

□

Observación 7.26. Suponga que $\text{char}(K) \neq 2$ y que λ es un 1-tensor, entonces por Propiedad 7.25 4.

$$\lambda \wedge \lambda = 0.$$

De hecho $\lambda \wedge \lambda = -\lambda \wedge \lambda$ luego $2\lambda \wedge \lambda = 0$ y como $2 \neq 0$ tenemos $\lambda \wedge \lambda = 0$. Además si λ_1, λ_2 son 1-tensores entonces por definición

$$\lambda_1 \wedge \lambda_2 = \lambda_1 \otimes \lambda_2 - \lambda_2 \otimes \lambda_1,$$

y en general para $\lambda_1, \dots, \lambda_k \in T^1(V)$, $k \in \mathbb{Z}_{\geq 0}$ tal que si $\text{char}(K) > 0$ entonces $k < \text{char}(K)$,

$$\lambda_1 \wedge \dots \wedge \lambda_k = \sum_{\sigma \in S_k} \text{sgn}(\sigma) \lambda_{\sigma(1)} \otimes \dots \otimes \lambda_{\sigma(k)}.$$

Proposición 7.27. Sea $f \in \text{Hom}_K(V, W)$ y $k, l \in \mathbb{Z}_{\geq 0}$. Para todo $\omega \in \Lambda^k(W)$ y $\eta \in \Lambda^l(W)$ tenemos

$$f^*(\omega \wedge \eta) = f^*\omega \wedge f^*\eta.$$

Dem. La demostración es una verificación inmediata. \square

Notación 7.28. Por Propiedad 7.25 5. denotamos $(\omega \wedge \eta) \wedge \theta = \omega \wedge \eta \wedge \theta$ y así podemos definir producto exterior de más de dos tensores alternantes.

Corolario 7.29. Sean $k_1, \dots, k_r \in \mathbb{Z}_{\geq 0}$ tal que si $\text{char}(K) > 0$ entonces $k_1 + \dots + k_r < \text{char}(K)$ y $\omega_i \in \Lambda^{k_i}(V)$, $i = 1, \dots, r$, entonces

$$\omega_1 \wedge \dots \wedge \omega_r = \frac{(k_1 + \dots + k_r)!}{k_1! \dots k_r!} \text{Alt}(\omega_1 \otimes \dots \otimes \omega_r).$$

Dem. Se sigue de inmediatamente de Propiedad 7.25 5. por inducción en r . \square

Ejemplo 7.30. Sea $n \in \mathbb{Z}_{>0}$, $\{e_i\}_{i=1}^n$ la base canónica de K^n y $\{f_1, \dots, f_n\}$ la base dual.

1. Para $\text{char}(K) > 0$ suponga que $n < \text{char}(K)$. Como $\text{Alt}(f_1 \otimes \dots \otimes f_n) = \frac{1}{n!} \det$ entonces

$$f_1 \wedge \dots \wedge f_n = \det.$$

2. Sea $k \in \{1, \dots, n\}$ y $1 \leq i_1 < \dots < i_k \leq n$. Para $\text{char}(K) > 0$ suponga que $k < \text{char}(K)$. Como $\text{Alt}(f_{i_1} \otimes \dots \otimes f_{i_k}) = \frac{1}{k!} \det_{i_1 \dots i_k}$ entonces

$$f_{i_1} \wedge \dots \wedge f_{i_k} = \det_{i_1 \dots i_k}.$$

Teorema 7.31. Suponga que $\text{char}(K) \neq 2$ y $\dim_K(V) = n < \infty$. Sea $\{v_1, \dots, v_n\}$ una base de V y $\{\lambda_1, \dots, \lambda_n\}$ la base dual. Sea $k \in \mathbb{Z}_{>0}$ tal que si $\text{char}(K) > 0$ entonces $k < \text{char}(K)$. La colección de k -tensores

$$\left\{ \lambda_{i_1} \wedge \dots \wedge \lambda_{i_k} \right\}_{1 \leq i_1 < \dots < i_k \leq n}$$

es una base de $\Lambda^k(V)$. En particular

$$\dim_K(\Lambda^k(V)) = \binom{n}{k}.$$

Dem. Sea $k \in \mathbb{Z}_{\geq 0}$ y $\omega \in \Lambda^k(V)$, como $\omega \in T^k(V)$ y $\{\lambda_{i_1} \otimes \cdots \otimes \lambda_{i_k}\}_{i_1, \dots, i_k=1}^n$ es un base de $T^k(V)$ entonces

$$\omega = \sum_{j_1, \dots, j_k=1}^n c_{j_1 \dots j_k} \lambda_{j_1} \otimes \cdots \otimes \lambda_{j_k}$$

con $c_{j_1 \dots j_k} \in K$. Así

$$\begin{aligned} \omega &= \text{Alt}(\omega) \\ &= \sum_{j_1, \dots, j_k=1}^n c_{j_1 \dots j_k} \text{Alt}(\lambda_{j_1} \otimes \cdots \otimes \lambda_{j_k}) \\ &= \sum_{j_1, \dots, j_k=1}^n c_{j_1 \dots j_k} k! \lambda_{j_1} \wedge \cdots \wedge \lambda_{j_k}. \end{aligned}$$

Ahora, dados $j_1, \dots, j_k \in \{1, \dots, n\}$ por Propiedad 7.25 4. y Observación 7.26, si dos de los subíndices j_i coinciden entonces $\lambda_{j_1} \wedge \cdots \wedge \lambda_{j_k} = 0$, luego podemos asumir que son todos distintos; en tal caso $\lambda_{j_1} \wedge \cdots \wedge \lambda_{j_k} = \text{sgn}(\sigma_{j_1, \dots, j_k}) \lambda_{i_1} \wedge \cdots \wedge \lambda_{i_k}$ donde $i_1 < \dots < i_k$ son los mismos subíndices j_1, \dots, j_k ordenados en forma creciente y $\sigma_{j_1, \dots, j_k} \in S_k$ es la permutación que reorganiza los j_1, \dots, j_k . Luego

$$\omega = \sum_{1 \leq i_1 < \dots < i_k \leq n} a_{i_1, \dots, i_k} \lambda_{i_1} \wedge \cdots \wedge \lambda_{i_k}$$

con $a_{i_1, \dots, i_k} = k! \sum_{\sigma \in S_k} c_{i_{\sigma(1)} \dots i_{\sigma(k)}}$, y así $\Lambda^k(V) = \langle \lambda_{i_1} \wedge \cdots \wedge \lambda_{i_k} \rangle_{1 \leq i_1 < \dots < i_k \leq n}$.

Para establecer la independencia lineal, note que si $1 \leq i_1 < \dots < i_k \leq n$ y $1 \leq j_1 < \dots < j_k \leq n$ entonces (ver Ejemplo 7.30 2.)

$$\lambda_{i_1} \wedge \cdots \wedge \lambda_{i_k}(v_{j_1}, \dots, v_{j_k}) = \begin{cases} 1 & \text{si } i_1 = j_1, \dots, i_k = j_k \\ 0 & \text{si no} \end{cases}$$

luego si $0 = \sum_{1 \leq i_1 < \dots < i_k \leq n} a_{i_1, \dots, i_k} \lambda_{i_1} \wedge \cdots \wedge \lambda_{i_k}$ entonces

$$\begin{aligned} 0 &= \sum_{1 \leq i_1 < \dots < i_k \leq n} a_{i_1, \dots, i_k} \lambda_{i_1} \wedge \cdots \wedge \lambda_{i_k}(v_{j_1}, \dots, v_{j_k}) \\ &= a_{j_1, \dots, j_k}. \end{aligned}$$

□

7.3. (l, k) -Tensores

En esta sección asumiremos que V tienen dimensión finita.

Definición 7.32. Sean $l, k \in \mathbb{Z}_{\geq 0}$. Un (l, k) -tensor es una función multilineal:

$$T^{(l, k)} : (V^*)^l \times V^k = \underbrace{V^* \times \cdots \times V^*}_{l\text{-veces}} \times \underbrace{V \times \cdots \times V}_{k\text{-veces}} \longrightarrow K.$$

Al espacio de (l, k) -tensores lo denotamos $T^{(l, k)}(V)$ ó $T_k^l(V)$.

Notación 7.33. Como V tiene dimensión finita, por Teorema 3.10 tenemos $V \simeq (V^*)^*$ canónicamente, en particular todo $v \in V$ lo identificaremos con el $(1, 0)$ -tensor

$$\begin{aligned}\widehat{v}: V^* &\longrightarrow K \\ \lambda &\longmapsto \lambda(v)\end{aligned}$$

y $T(1, 0)(V) \simeq V$ canónicamente.

Definición 7.34. Sean $l_1, l_2, k_1, k_2 \in \mathbb{Z}_{\geq 0}$, $S \in T^{(l_1, k_1)}(V)$ y $T \in T^{(l_2, k_2)}(V)$ definimos su *producto tensorial* $S \otimes T \in T^{(l_1+l_2, k_1+k_2)}(V)$ por

$$\begin{aligned}S \otimes T(\lambda_1, \dots, \lambda_{l_1+l_2}, v_1, \dots, v_{k_1+k_2}) \\ = S(\lambda_1, \dots, \lambda_{l_1}, v_1, \dots, v_{k_1})T(\lambda_{l_1+1}, \dots, \lambda_{l_1+l_2}, v_{k_1+1}, \dots, v_{k_1+k_2})\end{aligned}$$

para todo $\lambda_1, \dots, \lambda_{l_1+l_2} \in V^*$, $v_1, \dots, v_{k_1+k_2} \in V$.

Propiedad 7.35. El producto tensorial es bilineal y asociativo. Es decir, dados $l_1, l_2, l_3, k_1, k_2, k_3 \in \mathbb{Z}_{\geq 0}$ y $S, S' \in T^{(l_1, k_1)}(V)$, $T, T' \in T^{(l_2, k_2)}(V)$, $U \in T^{(l_3, k_3)}(V)$, $c \in K$ tenemos:

1. $(S + S') \otimes T = S \otimes T + S' \otimes T$,
2. $S \otimes (T + T') = S \otimes T + S \otimes T'$,
3. $(cS) \otimes T = c(S \otimes T) = S \otimes (cT)$,
4. $(S \otimes T) \otimes U = S \otimes (T \otimes U)$.

Dem. La demostración es una verificación directa. \square

Notación 7.36. Por la Proposición 7.35 4. denotamos $S \otimes T \otimes U = (S \otimes T) \otimes U$ y así podemos definir producto tensorial de más de dos tensores.

Teorema 7.37. Suponga que $\dim_K(V) = n$. Sea $\{v_1, \dots, v_n\}$ una base de V y $\{\lambda_1, \dots, \lambda_n\}$ la base dual. Para todo $l, k \in \mathbb{Z}_{>0}$, la colección de (l, k) -tensores

$$\left\{ v_{i_1} \otimes \dots \otimes v_{i_l} \otimes \lambda_{j_1} \otimes \dots \otimes \lambda_{j_k} \right\}_{i_1, \dots, i_l, j_1, \dots, j_k=1}^n$$

es una base de $T^{(l, k)}(V)$. En particular

$$\dim_K \left(T^{(l, k)}(V) \right) = n^{(l+k)}.$$

Dem. Similar a la demostración de Teorema 7.10. Note que para todo $T \in T^{(l, k)}(V)$

$$T = \sum_{i_1, \dots, i_l=1}^n \sum_{j_1, \dots, j_k=1}^n T(\lambda_{i_1}, \dots, \lambda_{i_l}, v_{j_1}, \dots, v_{j_k}) v_{i_1} \otimes \dots \otimes v_{i_l} \otimes \lambda_{j_1} \otimes \dots \otimes \lambda_{j_k}.$$

\square

Ejemplo 7.38. Sea $\{e_1, \dots, e_n\}$ la base canónica de K^n y $\{f_1, \dots, f_n\}$ la base dual.

1. Suponga que $\text{char}(K) \neq 2$. Para $n = 3$ sea $T_\times \in T^{(1,2)}(K^3)$ el tensor

$$T_\times = e_1 \otimes (f_2 \wedge f_3) - e_2 \otimes (f_1 \wedge f_3) + e_3 \otimes (f_1 \wedge f_2).$$

2. Sea $A = (a_{ij}) \in M_{n \times n}(K)$, y $T_A \in T^{(1,1)}(V)$ el tensor

$$T_A = \sum_{i,j=1}^n a_{ij} e_i \otimes f_j.$$

Definición 7.39. Suponga que $\dim_K(V) = n$. Sea $\{v_1, \dots, v_n\}$ una base de V y $\{\lambda_1, \dots, \lambda_n\}$ la base dual. Sea $l, k \in \mathbb{Z}_{\geq 0}$ y

$$t = v_{\alpha_1} \otimes \dots \otimes v_{\alpha_l} \otimes \lambda_{\beta_1} \otimes \dots \otimes \lambda_{\beta_k} \in T^{(l,k)}(V).$$

Dados $l', k' \in \mathbb{Z}_{\geq 0}$ tales que $l' \geq k$ y $k' \geq l$ definimos la *contracción por t* como la transformación lineal

$$\begin{aligned} \hat{t}: T^{(l',k')}(V) &\longrightarrow T^{(l'-k,k'-l)}(V) \\ T &\longmapsto T(t) := \hat{t}(T) \end{aligned}$$

donde si

$$T = v_{i_1} \otimes \dots \otimes v_{i_{l'}} \otimes \lambda_{j_1} \otimes \dots \otimes \lambda_{j_{k'}}$$

entonces

$$T(t) = \left(\prod_{s=1}^k \lambda_{\beta_s}(v_{i_s}) \prod_{s=1}^l \lambda_{j_s}(v_{\alpha_s}) \right) v_{i_{k+1}} \otimes \dots \otimes v_{i_{l'}} \otimes \lambda_{j_{l+1}} \otimes \dots \otimes \lambda_{j_{k'}}.$$

Extendemos linealmente a todo los (l, k) -tensores $t \in T^{(l,k)}(V)$ para obtener

$$\begin{aligned} \hat{\bullet}: T^{(l,k)}(V) &\longrightarrow \text{Hom}_K \left(T^{(l',k')}(V), T^{(l'-k,k'-l)}(V) \right) \\ t &\longmapsto \hat{t} \end{aligned}$$

Observación 7.40. Note que

$$\left(\prod_{s=1}^k \lambda_{\beta_s}(v_{i_s}) \prod_{s=1}^l \lambda_{j_s}(v_{\alpha_s}) \right) = v_{i_1} \otimes \dots \otimes v_{i_k} \otimes \lambda_{j_1} \otimes \dots \otimes \lambda_{j_l}(v_{\alpha_1}, \dots, v_{\alpha_l}, \lambda_{\beta_1}, \dots, \lambda_{\beta_k}).$$

Ejemplo 7.41. Sea $\{e_1, \dots, e_n\}$ la base canónica de K^n y $\{f_1, \dots, f_n\}$ la base dual.

1. Sea $T_{\times} \in T^{(1,2)}(K^3)$ como en Ejemplo 7.38 1., y

$$v_1 = \sum_{i=1}^3 a_i e_i, \quad v_2 = \sum_{i=1}^3 b_i e_i$$

con $a_i, b_i \in K, i = 1, \dots, 3$. Entonces $v_1 \otimes v_2 \in T^{(2,0)}(K^3)$ y $T_{\times}(v_1 \otimes v_2) \in T^{(1,0)}(K^3) \simeq K^3$ con

$$\begin{aligned} T_{\times}(v_1 \otimes v_2) &= (f_2 \wedge f_3)(v_1, v_2)e_1 - (f_1 \wedge f_3)(v_1, v_2)e_2 + (f_1 \wedge f_2)(v_1, v_2)e_3 \\ &= (a_2 b_3 - a_3 b_2)e_1 - (a_1 b_3 - a_3 b_1)e_2 + (a_1 b_2 - a_2 b_1)e_3 \\ &=: v_1 \times v_2 \end{aligned}$$

2. Sea $T_A \in T^{(1,1)}(K^n)$ como en Ejemplo 7.38 2. y

$$v = \sum_{i=1}^n c_i e_i$$

con $c_i \in K, i = 1, \dots, n$. Entonces $T_A(v) \in T^{(1,0)}(K^n) \simeq K^n$ con

$$\begin{aligned} T_A(v) &= \sum_{i,j=1}^n a_{ij} f_j(v) e_i \\ &= \sum_{i,j=1}^n a_{ij} c_j e_i \\ &= \sum_{i=1}^n \left(\sum_{j=1}^n a_{ij} c_j \right) e_i \\ &= f_A(v) \end{aligned}$$

donde $f_A \in \text{Hom}_K(K^n, K^n)$ está definida por $f_A(e_j) = \sum_{i=1}^n a_{ij} e_i$.

3. Sea $T_A T^{(1,1)}(K^n)$ como en Ejemplo 7.38 2. y

$$\lambda = \sum_{j=1}^n d_j f_j$$

con $d_i \in K, i = 1, \dots, n$. Entonces $T_A(\lambda) \in T^{(0,1)}(K^n) \simeq (K^n)^*$ con

$$\begin{aligned} T_A(\lambda) &= \sum_{i,j=1}^n a_{ij} \lambda(e_i) f_j \\ &= \sum_{i,j=1}^n a_{ij} d_i f_j \\ &= \sum_{j=1}^n \left(\sum_{i=1}^n a_{ij} d_i \right) f_j \\ &= f_A^*(\lambda) \end{aligned}$$

donde $f_A \in \text{Hom}_K(K^n, K^n)$ está definida por $f_A(e_j) = \sum_{i=1}^n a_{ij}e_i$.

7.4. Convenciones en notación de tensores

En esta sección asumiremos que V tienen dimensión finita.

Notación 7.42. En esta sección usaremos las siguientes convenciones.

1. Los elementos de V los denotaremos con subíndices.
2. Los elementos de V^* los denotaremos con superíndices.
3. Al espacio de (l, k) -tensores lo denotaremos por $T_k^l(V)$.
4. Convención de Einstein: si un índice se repite como subíndice y superíndice se asume sumatoria sobre este.

Ejemplo 7.43. Sea $n = \dim(V)$, $\{v_1, \dots, v_n\}$ una base de V y $\{\lambda^1, \dots, \lambda^n\}$ la base dual.

1. Dado $v \in V$ y $\lambda \in V^*$ si denotamos

$$\begin{aligned} v^i &= \lambda^i(v) \\ \lambda_i &= \lambda(v_i) \end{aligned}$$

entonces

$$\begin{aligned} v &= v^i v_i \\ \lambda &= \lambda_i \lambda^i \end{aligned}$$

2. Dado $T \in T_k^l(V)$ si denotamos

$$T_{j_1 \dots j_k}^{i_1 \dots i_l} = T(\lambda^{i_1}, \dots, \lambda^{i_l}, v_{j_1}, \dots, v_{j_k})$$

entonces

$$T = T_{j_1 \dots j_k}^{i_1 \dots i_l} v_{i_1} \otimes \dots \otimes v_{i_l} \otimes \lambda^{j_1} \otimes \dots \otimes \lambda^{j_k}$$

3. Dado $f \in \text{Hom}_K(V, V)$ si denotamos

$$f_j^i = \lambda^i(f(v_j))$$

entonces el tensor $T_f = f_j^i v_i \otimes \lambda^j \in T_1^1(V)$ es tal que

$$\begin{aligned} T_f(v) &= f_j^i (v_i \otimes \lambda^j)(v) \\ &= f_j^i \lambda^j(v) v_i \\ &= f_j^i v^j v_i \\ &= f(v) \end{aligned}$$

y

$$\begin{aligned} T_f(\lambda) &= f_j^i (v_i \otimes \lambda^j) (\lambda) \\ &= f_j^i \lambda(v_i) \lambda^j \\ &= f_j^i \lambda_i \lambda^j \\ &= f^*(\lambda) \end{aligned}$$

Apéndice A

Cuerpos

En este apéndice cubrimos algunos resultados elementales de cuerpos y polinomios que son necesarios en nuestro estudio de operadores. Los cuerpos son las estructuras algebraicas conmutativas minimales que garantizan que toda ecuación lineal tenga una única solución dentro de la misma estructura, es decir que dado un cuerpo \mathbb{K} y $a, b \in \mathbb{K}$ cualesquiera, con $a \neq 0$, la ecuación $ax + b = 0$ tiene solución $x = -b/a$ en \mathbb{K} .

Los cuerpos más comunes son los números racionales \mathbb{Q} , los números reales \mathbb{R} , los números complejos \mathbb{C} y las fracciones racionales sobre estos cuerpos, por ejemplo $\mathbb{Q}(x_1, \dots, x_n)$ que es el conjunto cociente de polinomios con coeficientes en \mathbb{Q} en n variables.

Definición A.1. Un *cuerpo* es un conjunto \mathbb{K} junto con dos operaciones binarias $+$ y \cdot (e.d. funciones $\mathbb{K} \times \mathbb{K} \rightarrow \mathbb{K}$), que llamamos respectivamente *suma* y *producto* (o *adición* y *multiplicación*), y que contiene dos elementos distintos 0 y 1, que llamamos respectivamente *cero* y *uno*, los cuales satisfacen las siguientes propiedades.

- (i) *Commutatividad*: Para todo $a, b, c \in \mathbb{K}$, se tiene $a + b = b + a$ y $a \cdot b = b \cdot a$.
- (ii) *Asociatividad*: Para todo $a, b, c \in \mathbb{K}$, se tiene $a + (b + c) = (a + b) + c$ y $a \cdot (b \cdot c) = (a \cdot b) \cdot c$,
- (iii) *Neutralidad de 0 y 1*: Para todo $a \in \mathbb{K}$, se tiene $0 + a = a$ y $1 \cdot a = a$,
- (iv) *Existencia de opuesto y de inverso*: Para todo $a \in \mathbb{K}$, existe $-a \in \mathbb{K}$ para el cual $-a + a = 0$ y, si además $a \neq 0$, entonces existe $a^{-1} \in \mathbb{K}$ para el cual $a \cdot a^{-1} = 1$,
- (v) *Distributividad del producto sobre la suma*: Para todo $a, b, c \in \mathbb{K}$, se tiene $a \cdot (b + c) = a \cdot b + a \cdot c$.

Notación A.2. Es usual omitir el símbolo \cdot en la operación de multiplicación, de tal forma que $a \cdot b$ se denota también por ab .

Ejemplo A.3. Los siguientes conjuntos junto con sus respectivas operaciones son cuerpos.

1. El subconjunto de los números reales $\mathbb{Q}[\sqrt{2}]$, el cual está formado por los números de la forma $a + b\sqrt{2}$ donde $a, b \in \mathbb{Q}$, con las operaciones heredadas de \mathbb{R} .
2. El subconjunto de los números complejos $\mathbb{Q}[i]$, el cual está formado por los números de la forma $a + bi$ donde $a, b \in \mathbb{Q}$, con las operaciones heredadas de \mathbb{C} .
3. El conjunto \mathbb{F}_p de clases de equivalencia módulo p en los números enteros \mathbb{Z} , donde p es un número primo, con las operaciones heredadas de las operaciones usuales de suma y multiplicación de \mathbb{Z} .

Ejemplo A.4. Los siguientes conjuntos no son cuerpos.

1. El conjunto de los números naturales \mathbb{N} con sus operaciones usuales, pues los elementos diferentes de 0 no tienen opuesto.
2. El conjunto de los números enteros \mathbb{Z} con sus operaciones usuales, pues los elementos diferentes de 0, aparte de -1 y de 1 , no tienen inverso.

Proposición A.5. Sea \mathbb{K} un cuerpo.

- (i) Ley de cancelación: Dados $a, b, c \in \mathbb{K}$, la igualdad $a + b = c + b$ implica $a = c$, y cuando $b \neq 0$, la igualdad $ab = cb$ implica $a = c$.
- (ii) Unicidad de 0 y 1: Los elementos neutros de la suma y el producto son únicos.
- (iii) Unicidad del opuesto e inverso: Para todo $a \in \mathbb{K}$ su opuesto y, cuando $a \neq 0$, su inverso son únicos.

Dem.

- (i) Basta con sumar el opuesto de b a ambos lados de la igualdad en el caso de la suma, o multiplicar por el inverso de b en el caso de la multiplicación.
- (ii) Si $e \in \mathbb{K}$ es tal que se tiene $a + e = a$ para algún $a \in \mathbb{K}$, por la neutralidad de 0 se obtiene $a + 0 = a = a + e$. La ley de cancelación implica $0 = e$. Similarmente, se puede verificar la unicidad de 1 como neutro del producto.
- (iii) Para verificar la unicidad del opuesto, observe que si $b, c \in \mathbb{K}$ verifican $a + b = 0 = a + c$, la ley de cancelación implica que b y c son iguales. Similarmente se establece la unicidad del inverso, cuando este existe.

□

Notación A.6. Si $a \in \mathbb{K}$ es diferente de 0, a su inverso a^{-1} también lo denotaremos por $1/a$. Es usual denotar las operaciones $a + (-b)$ por $a - b$ y $a \cdot b^{-1}$ por $\frac{a}{b}$.

Ejemplo A.7. En \mathbb{F}_5 las operaciones de suma y producto están dadas por las siguientes tablas.

+	0	1	2	3	4
0	0	1	2	3	4
1	1	2	3	4	0
2	2	3	4	0	1
3	3	4	0	1	2
4	4	0	1	2	3

·	0	1	2	3	4
0	0	0	0	0	0
1	0	1	2	3	4
2	0	2	4	1	3
3	0	3	1	4	2
4	0	4	3	2	1

Sobre \mathbb{F}_5 , la ecuación $3x + 4 = 1$ es equivalente a

$$3x = 1 - 4, \quad 3x = 1 + 1, \quad 3x = 2, \quad x = 2/3, \quad x = 2 \cdot 2$$

luego su solución es $x = 4$.

Proposición A.8. Sea \mathbb{K} un cuerpo. Para todo $a, b \in \mathbb{K}$, se tienen las siguientes igualdades.

(i) $a \cdot 0 = 0$

(ii) $-1 \cdot a = -a$

(iii) $(-a) \cdot b = a \cdot (-b) = -(a \cdot b)$

(iv) $(-a) \cdot (-b) = a \cdot b$

Dem.

(i) Tenemos

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

y por la ley de cancelación obtenemos $0 = a \cdot 0$.

(ii) Tenemos

$$-1 \cdot a + a = -1 \cdot a + 1 \cdot a = (-1 + 1)a = 0 \cdot a = 0.$$

(iii) Por unicidad del opuesto basta verificar que $(-a)b$ y $a(-b)$ son el opuesto de ab . De las igualdades

$$0 = 0 \cdot b = (a + (-a))b = ab + (-a)b$$

se obtiene que $(-a)b$ es el opuesto de ab . Similarmente se establece $a \cdot (-b) = -(a \cdot b)$.

(iv) Usando la igualdad $-(-b) = b$ y la propiedad A.8.(iii) obtenemos

$$(-a)(-b) = a(-(-b)) = ab$$

□

Definición A.9. Sea \mathbb{K} un cuerpo. Si existe un número natural k para el cual

$$\underbrace{1 + 1 + \dots + 1}_{k \text{ sumandos}} = 0,$$

al mínimo entre estos lo llamamos la *característica* de \mathbb{K} y lo denotamos por $\text{char}(\mathbb{K})$. En caso de que no exista tal k , definimos la característica de \mathbb{K} como 0.

Observación A.10. Para todo primo p , la característica de \mathbb{F}_p es p . La característica de \mathbb{Q} , \mathbb{R} y \mathbb{C} es 0.

Definición A.11. Sea \mathbb{K} un cuerpo. Un *polinomio con coeficientes en \mathbb{K} en la variable t* es una expresión de la forma

$$a_n t^n + a_{n-1} t^{n-1} + \dots + a_1 t + a_0$$

con $a_n, \dots, a_1, a_0 \in \mathbb{K}$. Si $P(t)$ denota este polinomio, dado $c \in \mathbb{K}$, el *valor de $P(t)$ en c* es

$$a_n c^n + a_{n-1} c^{n-1} + \dots + a_1 c + a_0,$$

el cual denotaremos por $P(c)$. Cuando tenemos $P(c) = 0$ decimos que c es una *raíz* de $P(t)$. Denotamos por $\mathbb{K}[t]$ al conjunto de polinomios con coeficientes en \mathbb{K} en la variable t .

Observación A.12. Si $P(t)$ y $Q(t)$ son los polinomios $a_n t^n + \dots + a_1 t + a_0$ y $b_m t^m + \dots + b_1 t + b_0$, su suma $P(t) + Q(t)$ es el polinomio $c_k t^k + \dots + c_1 t + c_0$, con $k = \max n, m$ y $c_i = a_i + b_i$, y su producto $P(t)Q(t)$ es el polinomio $d_l t^l + \dots + d_1 t + d_0$, con $l = n + m$ y $d_i = \sum_{j=0}^i a_{i-j} b_j$, donde en la fórmulas para c_i y d_i tomamos valores $a_i = 0$ para $i > n$ y $b_i = 0$ para $i > m$.

Definición A.13. Sea \mathbb{K} un cuerpo. Decimos que \mathbb{K} es *algebraicamente cerrado* si todo polinomio no constante en $\mathbb{K}[t]$ tiene una raíz en \mathbb{K} .

Teorema A.14 (Teorema fundamental del álgebra). *El cuerpo de los números complejos \mathbb{C} es algebraicamente cerrado.*

Dem. Presentamos una prueba usando análisis complejo, en particular usamos el Teorema de Liouville que indica que si una función es analítica y acotada en todo el plano complejo, entonces es constante.

Sea $P(t) \in \mathbb{C}[t]$ con $P(t) = a_n t^n + \dots + a_1 t + a_0$, y $a_n \neq 0$. Considere la función $f(z)$ dada por

$$f(z) = P(z)/(a_n z^n) = 1 + \sum_{k=1}^{n-1} \frac{a_k}{a_n} \frac{1}{z^{n-k}}$$

la cual está definida para todo $z \in \mathbb{C} \setminus \{0\}$. Si $r = |z|$, por la desigualdad triangular, tenemos

$$|f(z)| \geq 1 - \sum_{k=1}^{n-1} \left| \frac{a_k}{a_n} \right| \frac{1}{r^{n-k}}.$$

El límite cuando r tiende a infinito del lado derecho de la desigualdad es 1, luego existe $R > 0$ tal que $|f(z)| > 1/2$ siempre que $|z| = r > R$, y así, $|P(z)| > |a_n|R^n/2 > 0$ para $|z| > R$. Si D es el disco cerrado centrado en el origen de radio R , como D es compacto y la función $|P(z)|$ es continua, esta alcanza un mínimo m en D .

Suponga que $P(t)$ no tiene raíces y veamos que esto implica que es constante. Bajo este supuesto, la función $z \mapsto 1/P(z)$ es analítica sobre todo el plano complejo y así el mínimo m es estrictamente positivo. Así $|1/P(z)|$ está acotada por $2/(|a_n|R^n)$ fuera de D y por $1/m$ en D , luego por el teorema de Liouville $z \mapsto 1/P(z)$ es una función constante y así $P(t)$ es un polinomio constante. \square

Polinomios con coeficientes en un cuerpo

Definición A.15. Sea $P(t) \in \mathbb{K}[t]$ el polinomio $a_n t^n + a_{n-1} t^{n-1} + \dots + a_1 t + a_0$, con $a_n \neq 0$. El *grado de $P(t)$* es n y lo denotamos por $\deg(P(t))$ y al coeficiente a_n lo llamamos *coeficiente líder*. Para $P(t) = 0$, definimos $\deg(0) = -\infty$ y convenimos la desigualdad $-\infty < n$ válida para todo $n \in \mathbb{Z}$.

Observación A.16. Para todo $P(t), Q(t) \in \mathbb{K}[t]$, tenemos

$$\begin{aligned} \deg(P(t) + Q(t)) &\leq \max\{\deg(P(t)), \deg(Q(t))\} \\ \deg(P(t)Q(t)) &= \deg(P(t)) + \deg(Q(t)). \end{aligned}$$

Si $P(t)$ y $Q(t)$ difieren en su grado, se cumple la igualdad

$$\deg(P(t) + Q(t)) = \max\{\deg(P(t)), \deg(Q(t))\},$$

y si $Q(t)$ no es 0, se tiene

$$\deg(P(t)Q(t)) \geq \deg(P(t)).$$

Teorema A.17 (Algoritmo de la división). *Dados $A(t), B(t) \in \mathbb{K}[t]$, con $B(t) \neq 0$, existe un único par de polinomios $(Q(t), R(t))$ de $\mathbb{K}[t]$, llamados respectivamente cociente y residuo en la división de $A(t)$ por $B(t)$, tales que*

$$A(t) = Q(t)B(t) + R(t), \text{ y } \deg(R(t)) < \deg(B(t)).$$

Dem. Sea C el conjunto $\{A(t) - S(t)B(t)\}_{S(t) \in \mathbb{K}[t]}$. Como C no es vacío, contiene un polinomio $R(t)$ de grado mínimo. Sea $Q(t) \in \mathbb{K}[t]$ tal que $R(t) = A(t) - Q(t)B(t)$. Obtenemos $\deg(R(t)) < \deg(B(t))$, pues de lo contrario, si escribimos $R(t) = at^n + \dots$ y $B(t) = bt^m + \dots$ donde $a, b \in \mathbb{K}$ son los respectivos coeficientes líderes de $R(t)$ y $B(t)$, tendríamos $n \geq m$ y así el polinomio $R(t) - \frac{a}{b}t^{n-m}B(t)$, que está en C pues es igual a $A(t) - (Q(t) + \frac{a}{b}t^{n-m})B(t)$, sería un polinomio de grado estrictamente menor que $\deg(R(t))$, contradiciendo la minimalidad del grado de este.

Para establecer la unicidad de $Q(t)$ y $R(t)$, tomamos $\tilde{Q}(t), \tilde{R}(t) \in \mathbb{K}[t]$ para los cuales también se tiene $A(t) = \tilde{Q}(t)B(t) + \tilde{R}(t)$ con $\deg(\tilde{R}(t)) < \deg(B(t))$. De las igualdades

$$Q(t)B(t) + R(t) = A(t) = \tilde{Q}(t)B(t) + \tilde{R}(t),$$

se obtiene

$$(Q(t) - \tilde{Q}(t))B(t) = \tilde{R}(t) - R(t).$$

Como

$$\deg(R(t) - \tilde{R}(t)) \leq \max\{\deg(R(t)), \deg(\tilde{R}(t))\} < \deg(B(t)),$$

entonces

$$\deg((Q(t) - \tilde{Q}(t))B(t)) = \deg(R(t) - \tilde{R}(t)) < \deg(B(t)).$$

Por ende $Q(t) - \tilde{Q}(t)$ es 0 y así $R(t) - \tilde{R}(t)$ también es 0, es decir $\tilde{R}(t) = R(t)$ y $\tilde{S}(t) = S(t)$. \square

Observación A.18. (i) Cuando en el algoritmo de la división el residuo es 0 decimos que $B(t)$ divide a $A(t)$.

(ii) Observamos que $\mathbb{K}[t]$ no tiene divisores de 0.

(iii) Una consecuencia de no tener divisor de cero es que la ley cancelativa es válida para el producto de polinomios. Pues, si $A(t)B(t) = A(t)C(t)$ con $A(t) \neq 0$, entonces $A(t)(B(t) - C(t)) = 0$ y como $A(t)$ no es un divisor de cero se tiene $B(t) - C(t) = 0$, es decir $B(t) = C(t)$.

Corolario A.19. Sea $P(t) \in \mathbb{K}[t]$. Si $\lambda \in \mathbb{K}$ es una raíz de $P(t)$, entonces $(t - \lambda)$ divide a $P(t)$.

Dem. En el algoritmo de la división tome $A(t) = P(t)$ y $B(t) = t - \lambda$ y sean $Q(t), R(t) \in \mathbb{K}[t]$ el cociente y el residuo. En particular, como $B(t)$ tiene grado 1 entonces el residuo es un polinomio constante, es decir $R(t) = a$ para algún $a \in \mathbb{K}$. Pero tenemos $0 = P(\lambda) = (\lambda - \lambda)S(\lambda) + a = a$, luego $R(t) = 0$. \square

Definición A.20. Sean $A(t), B(t) \in \mathbb{K}[t]$. Decimos que $D(t) \in \mathbb{K}[t]$ es el *máximo divisor común* de $A(t)$ y $B(t)$ si es un polinomio mónico que satisface las siguientes dos propiedades.

(i) El polinomio $D(t)$ divide a $A(t)$ y $B(t)$.

(ii) Si $D_0(t) \in \mathbb{K}[t]$ divide a $A(t)$ y $B(t)$, entonces $D_0(t)$ divide a $D(t)$.

Proposición A.21. Todo par de polinomios $(A(t), B(t))$ de $\mathbb{K}[t]$, con uno de ellos no nulo, tiene un máximo común divisor $D(t)$. Mas aún, existen $P(t), Q(t) \in \mathbb{K}[t]$ para los cuales $D(t) = P(t)A(t) + Q(t)B(t)$.

Dem. Sea C el conjunto de polinomios no nulos definido por

$$C = \{S(t)A(t) + T(t)B(t) \in \mathbb{K}[t] \setminus \{0\}\}_{S(t), T(t) \in \mathbb{K}[t]}.$$

Como C no es vacío, contiene un polinomio mónico $D_0(t)$ de grado mínimo. Sean $P(t), Q(t) \in \mathbb{K}[t]$ tales que $D(t) = P(t)A(t) + Q(t)B(t)$.

Veamos que $D(t)$ divide a $A(t)$ y $B(t)$. De hecho, si $S(t), R(t) \in \mathbb{K}[t]$ son el cociente y residuo en la división de $A(t)$ por $D(t)$, obtenemos

$$R(t) = A(t) - S(t)D(t) = (1 - S(t)P(t))A(t) - S(t)Q(t)B(t).$$

Pero como $D(t)$ tiene grado mínimo en C y $\deg(R(t)) < \deg(D(t))$, entonces $R(t)$ no pertenece a C , luego es igual a 0. Por ende, $D(t)$ divide a $A(t)$. Por un argumento similar, se obtiene que $D(t)$ divide a $B(t)$.

Veamos que $D(t)$ es maximal entre los divisores comunes de $A(t)$ y $Q(t)B(t)$. Si $D_0(t) \in \mathbb{K}[t]$ divide a $A(t)$ y $B(t)$, es decir si existen $S(t), T(t) \in \mathbb{K}[t]$ tales que $A(t) = S(t)D_0(t)$ y $B(t) = T(t)D_0(t)$, entonces $D_0(t)$ también divide a $D(t)$, pues

$$D(t) = P(t)A(t) + Q(t)B(t) = (P(t)S(t) + Q(t)T(t))D_0(t).$$

□

Notación A.22. Dados $A(t), B(t) \in \mathbb{K}[t]$, no ambos nulos, denotamos por $(A(t), B(t))$ al máximo divisor común de $A(t)$ y $B(t)$.

Definición A.23. Decimos que $A(t), B(t) \in \mathbb{K}[t]$, no ambos nulos, son *primos relativos* si $(A(t), B(t)) = 1$.

Ejemplo A.24. Si $\lambda_1, \lambda_2 \in \mathbb{K}$ son distintos, entonces $t - \lambda_1$ y $t - \lambda_2$ son primos relativos, pues se tiene la igualdad

$$\frac{1}{\lambda_2 - \lambda_1} ((t - \lambda_1) - (t - \lambda_2)) = 1.$$

y así $(t - \lambda_1, t - \lambda_2) = 1$.

Teorema A.25 (Algoritmo de Euclides). Sean $A(t), B(t) \in \mathbb{K}[t]$. Si $Q(t), R(t) \in \mathbb{K}[t]$ son el cociente y el residuo en la división de $A(t)$ por $B(t)$, entonces

$$(A(t), B(t)) = (B(t), R(t))$$

Dem. Sea $D(t) = (A(t), B(t))$. Veamos que $D(t)$ divide a $B(t)$ y $R(t)$. Para ello basta verificar que $D(t)$ divide a $R(t)$. Note que si $A(t) = S(t)D(t)$ y $B(t) = T(t)D(t)$, entonces

$$R(t) = A(t) - Q(t)B(t) = (S(t) - Q(t)T(t))D(t).$$

Veamos ahora que $D(t)$ es maximal entre los divisores comunes de $B(t)$ y $R(t)$. Como $D(t) = (A(t), B(t))$, existen $P_0(t), Q_0(t) \in \mathbb{K}[t]$ tales que $P_0(t)A(t) +$

$Q_0(t)B(t) = D(t)$. Suponga que $D_0(t)$ divide a $B(t)$ y $R(t)$, entonces si $B(t) = S_0(t)D_0(t)$ y $R(t) = T_0(t)D_0(t)$, tenemos

$$\begin{aligned} D(t) &= P_0(t)A(t) + Q_0(t)Q(t) \\ &= P_0(t)(Q(t)B(t) + R(t)) + Q_0(t)B(t) \\ &= (P_0(t)(Q(t)S_0(t) + T_0(t)) + Q_0(t)S_0(t))D_0(t) \end{aligned}$$

y obtenemos que $D_0(t)$ divide a $D(t)$. \square

Observación A.26. La utilidad del algoritmo de Euclides es que en la búsqueda del máximo común divisor $(A(t), B(t))$ podemos reemplazar esta pareja por la pareja de menor grado $(B(t), R(t))$. De esta forma seguimos iterativamente reduciendo los grados en la pareja hasta llegar al caso trivial $(R_0(t), 0) = c^{-1}R_0(t)$, donde c es el coeficiente líder de $R_0(t)$.

Ejemplo A.27. Considere los polinomios $A(t) = t^3 + t^2 - t - 1$ y $B(t) = t^3 - t^2$ en $\mathbb{Q}[t]$. Tenemos

$$\begin{aligned} A(t) &= B(t) + 2t^2 - t - 1 \\ B(t) &= \left(\frac{1}{2}t - \frac{1}{4}\right)(2t^2 - t - 1) + \frac{1}{4}t - \frac{1}{4} \\ 2t^2 - t - 1 &= 4(2t + 1)\left(\frac{1}{4}t - \frac{1}{4}\right) \end{aligned}$$

así

$$\begin{aligned} (A(t), B(t)) &= (B(t), 2t^2 - t - 1) = (2t^2 - t - 1, \frac{1}{4}t - \frac{1}{4}) = (\frac{1}{4}t - \frac{1}{4}, 0) \\ &= t - 1. \end{aligned}$$

y

$$\begin{aligned} t - 1 &= 4B(t) - (2t - 1)(2t^2 - t - 1) \\ &= 4B(t) - (2t - 1)(A(t) - B(t)) \\ &= -(2t - 1)A(t) + (2t + 3)B(t) \end{aligned}$$

Observación A.28. Similarmente a como definimos el máximo común divisor de un par de polinomios en $\mathbb{K}[t]$, podemos definir el *máximo común divisor de una familia finita de polinomios*. Si denotamos por $(A_1(t), \dots, A_n(t))$ al máximo común divisor de $\{A_1(t), \dots, A_n(t)\}$, observemos que

$$(A_1(t), \dots, A_n(t)) = ((A_1(t), \dots, A_{n-1}(t)), A_n(t)),$$

lo cual nos permite calcularlo inductivamente.

Teorema A.29 (Relación de Bézout). *Dados $A_1(t), \dots, A_n(t) \in \mathbb{K}[t]$, con uno de ellos no nulo, existen polinomios $P_1(t), \dots, P_n(t) \in \mathbb{K}[t]$ tales que*

$$P_1(t)A_1(t) + \dots + P_n(t)A_n(t) = (A_1(t), \dots, A_n(t))$$

Dem. Hacemos inducción en n , donde el caso base $n = 2$ ya fue demostrado. Para obtener el paso inductivo, asumimos que el resultado es cierto cuando nos son dados $n - 1$ polinomios, uno de ellos no nulo. Por la hipótesis de inducción existen $Q_1(t), \dots, Q_{n-1}(t) \in \mathbb{K}[t]$ tales que

$$Q_1(t)A_1(t) + \dots + Q_{n-1}(t)A_{n-1}(t) = B(t),$$

donde $B(t) = (A_1(t), \dots, A_{n-1}(t))$. Por el caso base, existen $Q(t), P_n(t) \in \mathbb{K}[t]$ para los cuales

$$Q(t)B(t) + P_n(t)A_n(t) = (B(t), A_n(t)).$$

Entonces, si definimos $P_i(t) = Q(t)Q_i(t)$ para $i \in \{1, \dots, n-1\}$, obtenemos

$$\begin{aligned} P_1(t)A_1(t) + \dots + P_{n-1}(t)A_{n-1}(t) + P_n(t)A_n(t) &= (B(t), A_n(t)) \\ &= ((A_1(t), \dots, A_{n-1}(t)), A_n(t)) \\ &= (A_1(t), \dots, A_n(t)). \end{aligned}$$

□

Ejemplo A.30. Considere los polinomios $A_1(t) = t^3 + t^2 - t - 1$, $A_2(t) = t^3 - t^2$ y $A_3(t) = t^2 + t$ en $\mathbb{Q}[t]$. Por el ejemplo A.27, tenemos

$$(A_1(t), A_2(t), A_3(t)) = ((A_1(t), A_2(t)), A_3(t)) = (t - 1, A_3(t)).$$

De las igualdades

$$\begin{aligned} A_3(t) &= t(t - 1) + 2t \\ t - 1 &= \frac{1}{2}(2t) - 1 \\ 2t &= (-2t)(-1) + 0 \end{aligned}$$

y el algoritmo de Euclides, se sigue

$$\begin{aligned} (A_3(t), t - 1) &= (t - 1, 2t) = (2t, -1) = (-1, 0) \\ &= 1 \end{aligned}$$

y obtenemos

$$\begin{aligned} 1 &= \frac{1}{2}(2t) - (t - 1) \\ &= \frac{1}{2}(A_3(t) - t(t - 1)) - (t - 1) \\ &= \frac{1}{2}A_3(t) - (\frac{1}{2}t + 1)(t - 1) \end{aligned}$$

al igual que

$$\begin{aligned} 1 &= \frac{1}{2}A_3(t) - (\frac{1}{2}t + 1)(-(2t - 1)A_1(t) + (2t + 3)A_2(t)) \\ &= (\frac{1}{2}t + 1)(2t - 1)A_1(t) - (\frac{1}{2}t + 1)(2t + 3)A_2(t) + \frac{1}{2}A_3(t). \end{aligned}$$

Definición A.31. Sea $P(t) \in \mathbb{K}[t]$, con $\deg(P(t)) > 0$. Decimos que $P(t)$ es un *polinomio irreducible* en $\mathbb{K}[t]$ si para toda factorización $P(t) = A(t)B(t)$ con $A(t), B(t) \in \mathbb{K}[t]$ tenemos que $\deg(A(t)) = 0$ ó $\deg(B(t)) = 0$, es decir $A(t)$ ó $B(t)$ es constantes.

Lema A.32. Sea $P(t) \in \mathbb{K}[t]$ un polinomio irreducible. Entonces, si $P(t)$ divide a $A(t)B(t)$ con $A(t), B(t) \in \mathbb{K}[t]$, entonces $P(t)$ divide a $A(t)$ ó $B(t)$.

Dem. Sea $Q(t)$ tal que se tiene $P(t)Q(t) = A(t)B(t)$ y suponga que $P(t)$ no divide a $A(t)$. Como $P(t)$ es irreducible tenemos $(P(t), A(t)) = 1$. Sean $S(t), T(t) \in \mathbb{K}[t]$ tales que $1 = S(t)P(t) + T(t)A(t)$, entonces

$$\begin{aligned} B(t) &= (S(t)P(t) + T(t)A(t))B(t) \\ &= S(t)P(t)B(t) + T(t)A(t)B(t) \\ &= S(t)P(t)B(t) + T(t)P(t)Q(t) \\ &= P(t)(S(t)B(t) + T(t)Q(t)) \end{aligned}$$

y así, $P(t)$ divide a $B(t)$.

Teorema A.33 (Factorización única). Si $A(t) \in \mathbb{K}[t]$ es un polinomio con $\deg(A(t)) > 0$, entonces existen polinomios irreducibles mónicos $P_1(t), \dots, P_r(t) \in \mathbb{K}[t]$ tales que

$$A(t) = cP_1(t) \cdots P_r(t),$$

donde c es el coeficiente líder de $P(t)$. Más aún, esta factorización de $A(t)$ es única en el siguiente sentido. Si $A(t) = cQ_1(t) \cdots Q_s(t)$ con $Q_1(t), \dots, Q_s(t) \in \mathbb{K}[t]$ irreducibles mónicos, entonces s es igual a r y existe una permutación σ de $\{1, \dots, r\}$ tal que, para $i \in \{1, \dots, r\}$ se tiene $Q_i(t) = P_{\sigma(i)}(t)$.

Dem. Demostremos primero la existencia de la factorización. Lo haremos por inducción en $\deg(A(t))$, siendo el caso base $\deg(A(t)) = 1$ evidente pues en tal caso $A(t)$ es irreducible y así obtenemos la factorización deseada con $r = 1$ y $P_1(t) = c^{-1}A(t)$. Suponga que la factorización por irreducibles mónicos ha sido demostrada para polinomios de grado estrictamente menor a d y sea $A(t) \in \mathbb{K}[t]$ con $\deg(A(t)) = d$. Si $A(t)$ es irreducible, no hay nada que demostrar de nuevo obtenemos la factorización con $r = 1$ y $P_1(t) = c^{-1}A(t)$. Suponga entonces que existen $B(t), C(t) \in \mathbb{K}[t]$ tales que $A(t) = B(t)C(t)$ con $\deg(B(t)) > 0$ y $\deg(C(t)) > 0$. En particular, tenemos $\deg(B(t)) < d$ y $\deg(C(t)) < d$. Por hipótesis de inducción, existen polinomios irreducibles mónicos $P_1(t), \dots, P_{r_1}(t), P_{r_1+1}(t), \dots, P_{r_1+r_2}(t) \in \mathbb{K}[t]$ para los que $B(t) = aP_1(t) \cdots P_{r_1}(t)$ y $C(t) = bP_{r_1+1}(t) \cdots P_{r_1+r_2}(t)$, donde a y b son los respectivos coeficientes líder. Así, obtenemos la factorización de $A(t)$ en irreducibles mónicos $A(t) = abP_1(t) \cdots P_{n_1+n_2}(t)$, donde ab es el coeficiente líder de $A(t)$.

Para establecer la unicidad de la factorización procedemos por inducción en el número de factores, siendo el caso base de un único factor inmediato pues en tal caso $A(t)$ es irreducible. Asuma por inducción que la unicidad ha sido demostrada cuando el número de factores es estrictamente menor a r . Si tenemos $A(t) = P_1(t) \cdots P_r(t) = Q_1(t) \cdots Q_s(t)$, entonces el lema anterior implica

que $P_r(t)$ divide a algún $Q_i(t)$, que podemos asumir, sin pérdida de generalidad, que es $Q_s(t)$. Pero como $Q_s(t)$ es irreducible entonces $Q_s(t) = cP_r(t)$ para algún $c \in \mathbb{K}$ con $c \neq 0$, y como $Q_s(t)$ y $P_r(t)$ son mónicos se tiene $c = 1$. Obtenemos así por la ley cancelativa $P_1(t) \cdots P_{r-1}(t) = Q_1(t) \cdots Q_{s-1}(t)$. Al tener un número de factores irreducibles mónicos estrictamente menor r , por la hipótesis de inducción se establece la igualdad $r - 1 = s - 1$, luego $r = s$, y la existencia de una biyección σ_0 de $\{1, \dots, r - 1\}$ tal que, para $i \in \{1, \dots, r - 1\}$, se tiene $Q_i(t) = P_{\sigma_0(i)}(t)$. El teorema se sigue cuando tomamos la biyección σ definida por $\sigma(i) = \sigma_0(i)$ para $1 \leq i \leq r - 1$ y $\sigma(r) = r$.

Teorema A.34. *Si $P(t) \in \mathbb{R}[t]$ es un polinomio mónico irreducible, entonces $P(t) = t - a$, para algún $a \in \mathbb{R}$, ó $P(t) = (t - a)^2 + b^2$, con $a, b \in \mathbb{R}$ donde $b \neq 0$.*

Dem. Sea $P(t) \in \mathbb{R}[t]$ un polinomio mónico irreducible. Por el teorema fundamental del álgebra $P(t)$ tiene una raíz $w = a + bi$ en \mathbb{C} , con $a, b \in \mathbb{R}$. Si w es un número real, es decir $b = 0$ y $w = a$, entonces $t - a$ divide a $P(t)$ y como $P(t)$ es mónico, obtenemos $P(t) = t - a$. De lo contrario, tenemos $w = a + bi$ con $b \neq 0$, y en tal caso \bar{w} , el conjugado de w , también es una raíz de $P(t)$. Luego $t - w$ y $t - \bar{w}$ dividen a $P(t)$ en $\mathbb{C}[t]$ y así, el mínimo múltiplo común $(t - w)(t - \bar{w})$ también divide a $P(t)$ (ver ejercicio 10). Por ende, como se tiene $(t - w)(t - \bar{w}) = (t - a - bi)(t - a + bi) = (t - a)^2 + b^2$ entonces $(t - w)(t - \bar{w})$ es un divisor de $P(t)$ en $\mathbb{R}[t]$ y al ser $P(t)$ mónico, obtenemos $P(t) = (t - a)^2 + b^2$.

Ejercicios

1. Demuestre que no hay ningún cuerpo propiamente contenido dentro del cuerpo de los números racionales.
2. Demuestre que en un cuerpo no hay divisores de cero.
3. Demuestre que la característica de un cuerpo es 0 ó un número primo.
4. Demuestre que en todo cuerpo de característica finita p se tiene $(a + b)^p = a^p + b^p$.
5. Demuestre que en todo cuerpo de característica p la función $x \mapsto x^p$ preserva sumas y productos, y envía a 1 y 0 en ellos mismos.
6. Demuestre que en un cuerpo finito de característica diferente a 2, la suma de todos sus elementos es igual a 0.
7. Encuentre el cociente y el residuo de las siguientes divisiones.
 - a) $t^4 - 3t^2 + 2t - 6$ entre $t^2 - t + 1$ en $\mathbb{Q}[t]$.
 - b) $t^5 + t^3 - 3t + 4$ entre $t^2 + 1$ en $\mathbb{R}[t]$.
 - c) $t^3 + 2t + 2$ entre $t + 1$ en $\mathbb{Q}[t]$.
 - d) $t^3 + 2t + 2$ entre $t + 1$ en $\mathbb{F}_3[t]$ y compare con la división anterior.

- e) $t^4 + t^2 + 1$ entre $t^2 + t + 1$ en $\mathbb{F}_2[t]$.
8. Determine si $t^4 + 3t^2 + 3t + 3$ es divisible por $t^2 + t + 3$ en $\mathbb{F}_5[t]$.
9. Para cada una de las siguientes familias de polinomios $\{A_1(t), \dots, A_n(t)\} \subseteq \mathbb{Q}[t]$:
- i) $A_1(t) = (t-3)(t-5)$, $A_2(t) = (t-1)(t-5)$, $A_3(t) = (t-1)(t-3)$;
 - ii) $A_1(t) = (t-3)$, $A_2(t) = (t-1)^2$;
 - iii) $A_1(t) = (t-1)(t^2-2)$, $A_2(t) = (t+1)(t^2-2)$, $A_3(t) = t^2-1$;
 - iv) $A_1(t) = t^2-2t+10$, $A_2(t) = t^2-2t+2$;
- (a) Encuentre el máximo común divisor $(A_1(t), \dots, A_n(t))$ en $\mathbb{Q}(t)$.
- (b) Encuentre polinomios $P_1(t), \dots, P_n(t) \in \mathbb{Q}[t]$ tales que

$$(A_1(t), \dots, A_n(t)) = P_1(t)A_1(t) + \dots + P_n(t)A_n(t).$$

10. Sea \mathbb{K} un cuerpo y sean $A(t), B(t) \in \mathbb{K}[t]$. Decimos que $M(t) \in \mathbb{K}[t]$ es el mínimo múltiplo común de $A(t)$ y $B(t)$ si es un polinomio mónico que satisface las siguientes dos propiedades.

- (i) $A(t)$ y $B(t)$ dividen a $M(t)$.
- (ii) Si $A(t)$ y $B(t)$ dividen a $M_1(t) \in \mathbb{K}[t]$, entonces $M(t)$ divide a $M_1(t)$.

Demuestre las siguientes dos afirmaciones.

- (a) Para todo $A(t), B(t) \in \mathbb{K}[t]$, con $A(t) \neq 0$ y $B(t) \neq 0$, existe el mínimo múltiplo común de $A(t)$ y $B(t)$.
- (b) Si $[A(t), B(t)]$ denota al mínimo múltiplo común de $P(t)$ y $Q(t)$, entonces

$$A(t)B(t) = ab(A(t), B(t))[A(t), B(t)]$$

donde $a, b \in \mathbb{K}$ son los respectivos coeficientes líderes de $A(t)$ y $B(t)$.

11. Sea \mathbb{K} un cuerpo y $P(t) \in \mathbb{K}[t]$ con $\deg(P(t)) > 0$. Dado $A(t) \in \mathbb{K}[t]$, demuestre que existen unos únicos $A_0(t), A_1(t), \dots, A_m(t) \in \mathbb{K}[t]$ con $\deg(A_i(t)) < \deg(P(t))$ para $i \in \{0, 1, \dots, m\}$ tales que

$$A(t) = A_0(t) + A_1(t)P(t) + \dots + A_m(t)P(t)^m.$$

12. Sea \mathbb{K} un cuerpo y sean $c_0, c_1, \dots, c_n \in \mathbb{K}$ distintos. Para $i \in \{0, 1, \dots, n\}$ defina

$$P_i(t) = \prod_{j \neq i} \frac{t - c_i}{c_j - c_i}.$$

Dado $P(t) \in \mathbb{K}[t]$ con $\deg(P(t)) \leq n$ demuestre que

$$P(t) = \sum_{i=0}^n P(c_i)P_i(t).$$