Un curso de Álgebra Lineal II

Camilo Sanabria

Universidad de los Andes
Departamento de Matemáticas
Bogotá - Colombia.

Índice general

Índice de figuras

Capítulo 1

Espacios vectoriales y transformaciones lineales

1.1. Espacios vectoriales

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

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

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

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

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

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

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

(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

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

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

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

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

(iv) *Espacio de funciones con valor en* $\mathbb{K}$: Dado un conjunto X, el conjunto $\mathbb{K}^X$ de funciones $X \to \mathbb{K}$, 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 $X \to \mathbb{K}$ 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 $+ : U \times U \to U$ 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. □

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

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

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

$$v=c_1v_1+\ldots+c_nv_n$$

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 $\left\{\sum_{i=1}^n c_iv_i \,\middle|\, c_1,\dots,c_n\in\mathbb{K},\ v_1,\dots,v_n\in C\right\}$. 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 $v=\sum_{i=1}^n a_iv_i$ y $w=\sum_{j=1}^m b_jw_j$, y así

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

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

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

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. ☐

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 $v_1, \ldots, v_n \in C$ y $c_1, \ldots, c_n \in \mathbb{K}$, la combinación lineal $\sum_i^n c_i v_i$ 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 $\langle C\rangle \leq W$. □

Ejercicios

1. Verique que los conjuntos en el ejemplo 1.2 con sus respectivas operaciones son espacios vectoriales.

2. 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.

3. Verifique que el conjunto de funciones racionales $\mathbb{C}(z_1, \ldots, z_n)$ es un espacio vectorial sobre $\mathbb{C}$.

4. Explique por qué los conjuntos $\mathbb{N}$ y $\mathbb{Z}$ no son espacios vectoriales.

5. ¿Es el conjunto $\{(x, y) \in \mathbb{R}^2 | y > 0\}$ un subespacio de $\mathbb{R}^2$?

6. ¿Es el conjunto $\{(x, y, z) \in \mathbb{K}^3 | x + y + z = 0\}$ un subespacio de $\mathbb{K}^3$?

7. ¿Es el conjunto $\{(x, -x, 2x) \in \mathbb{K}^3 | x \in \mathbb{R}\}$ un subespacio de $\mathbb{K}^3$?

8. 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.

9. 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.

10. Describa geométricamente el subconjunto $\langle(1, 0, 1), (0, 1, 0)\rangle$ de $\mathbb{R}^3$.

11. 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 indices I y un conjunto C, una *familia de elementos de C indexada por I*, o simplemente una *familia de C*, es una función $I \to C$ 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 colección de vectores de V indexados por un conjunto I y $(c_i)_{i\in I}$ una colección de escalares indexados por I, tales que $c_i = 0$ para todo indice $i \in I$, 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 $I = \{1, \ldots, n\}$.

Definición 1.14. Sea $\mathcal{B} = (v_i)_{i\in I}$ una colección indexada 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 $O = \sum_{i\in I} c_i v_i$, 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. □

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 $V = \{O\}$, 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 $(v_1, \ldots, v_n)$ y $(w_1, \ldots, w_m)$ de diferentes tamaños, es decir $n \neq m$. Sin perdida de generalidad, podemos asumir que $m > n$.
Para $j \in \{1, \ldots, m\}$, tome escalares $a_{1j}, \ldots, a_{nj}$ tales que

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

Así, dado $v \in V$ con $v = x_1 w_1 + \ldots + x_m w_m$ y $x_1, \ldots, x_m \in \mathbb{K}$, 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 homogeneo 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 $v = x_1 w_1 + \ldots + x_n w_m$ de vectores $w_1, \dots, w_n \in \{w_i\}_{i \in I}$ y obtenemos una contradicción similarmente a como la acabamos de obtener. □

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 $\dim(V)$. 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

$$v_0 = c_1 v_1 + \ldots + c_n v_n$$

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

$$O = c_1 v_1 + \ldots + c_n v_n.$$

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 $j \in \{1, \dots, n\}$. □

Teorema 1.21 (Caracterización de las bases)**.** *Sea $\mathcal{B} = (v_i)_{i\in I}$ una colección 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. □

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 $V = \left(K^I\right)_0$ (ver Ejemplo 1.2.5). Para $i \in I$, defina la función $\delta_i : I \to K$ 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 $\mathcal{C}$.

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

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

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 $\langle C \rangle$. □

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 $v_1, \ldots, v_n$ los elementos de C, enumerados de forma tal que $\{v_1, \ldots, v_m\}$ sea C_0. Definimos $m = 0$ cuando $C_0 = \emptyset$. Tenemos $m \leq n$. Iteratívamente, para $i \in \{1, \ldots, n-m\}$, 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 $C_0 \subseteq C_1 \subseteq \ldots \subseteq C_{n-m} = C'$. Para ver la maximalidad de C', tomemos $j \in \{1, \ldots, n\}$. 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 máximal respecto a las propiedades de contener a C_0 y ser linealmente independiente. El mismo argumento demuestra la contenencia de C en $\langle C' \rangle$, y la proposición 1.12.2. implica que $\langle C \rangle \leq \langle C' \rangle$. Para la contenencia opuesta, note que $C' \subseteq C$, y por lo tanto $\langle C' \rangle \leq \langle C \rangle$, y así $\langle C' \rangle = \langle C \rangle$. □

Teorema 1.26 (Extensión de un conjunto linealmente independientes 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. □

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 $\{v_1, \ldots, v_n\}$ en V. Como V tiene dimensión infinita, $(v_1, \ldots, v_n)$ no es una base de V, luego tampoco lo genera y por ende existe $v_{n+1} \in V$ fuera de $\langle v_1, \ldots, v_n \rangle$. Por lo tanto, el lema 1.24 implica que el conjunto $\{v_1, \ldots, v_n, v_{n+1}\}$ es linealmente independiente. □

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 contradiccón 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, 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, 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 extendendida 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 $U = \langle \mathcal{B}_0 \rangle = V$. □

Observación 1.29 (Existencias de bases y dimensión infinita)**.** El teorema 1.26 se puede usar para demostrar que, partiendo de un conjunto vacio, 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 $(P, \preccurlyeq)$ 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 contenencia. 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 máximal $\mathcal{B}_1$. Usando un argumento similar al del lema1.24, 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 máximal 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 máximal C' en la colección, ordenada por inclusión, de subconjunto linealmente indpendientes 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 $\{(x,y,z)\in\mathbb{R}^3|2x+y-z=0, x-y=0\}$ de $\mathbb{R}^3$.

5. Encuentre una base del subespacio $\{(x,y,z)\in\mathbb{R}^3|2x+y-z=0\}$ de $\mathbb{Q}^3$.

6. Para cada uno de las siguientes conjuntos de vectores encuentre un subconjunto linealmente independiente maximal.

   a) $\big\{(1,0,-2,1),(0,-1,1,1),(-1,2,1,0),(2,1,-4,4)\big\}\subseteq\mathbb{Q}^4$

   b) $\big\{(1,0,-2,1),(0,-1,1,1),(-1,2,1,0),(2,1,2,-2)\big\}\subseteq\mathbb{Q}^4$

   c) $\big\{(1,-1,5,-8,6),(-1,1,-5,5,-3),(1,0,3,-3,5),(2,3,4,-1,1),(0,1,0,-1,2)\big\}\subseteq\mathbb{Q}^5$

   d) $\big\{(1,-1,5,-8,6),(-1,1,-5,5,-3),(1,0,3,-3,5),(2,1,4,8,0)\big\}\subseteq\mathbb{Q}^5$

   e) $\big\{(1,1,0,0,0,0),(0,0,1,1,0,0),(0,0,0,0,1,1),(1,0,1,0,1,0),(0,1,0,1,0,1)\big\}\subseteq\mathbb{Q}^6$

1.3. Transformaciones lineales

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

Definición 1.30. Sea $f:V\to W$ 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 $\mathrm{Hom}_{\mathbb{K}}(V,W)$. A las transformaciones lineales de un espacio en él mismo las llamamos *operadores* (o *endomorfismos de espacios vectoriales*). Al conjunto $\mathrm{Hom}_{\mathbb{K}}(V,V)$ de operadores lineales de V lo denotamos por $\mathrm{End}_{\mathbb{K}}(V)$.

Ejemplo 1.31. 1. *Transforación lineal cero*: La función $\underline{O} : V \to W$ definida por $\underline{O}(v) = O$ para todo v es una transformación lineal.

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

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

1. $f(O) = O$,

2. $f(-v) = -f(v)$ *para todo* $v \in V$, *y*

3. $f(c_1v_1 + \ldots + c_nv_n) = c_1f(v_1) + \ldots + c_nf(v_n)$ *para todo* $c_1, \ldots, c_n \in K$ *y todo* $v_1, \ldots, v_n \in V$.

Dem.

1. Tenemos $f(O) = f(0O) = 0f(O) = O$.

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

3. 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 $c_1, \ldots, c_{n+1} \in K$ y todo $v_1, \ldots, v_{n+1} \in V$ tenemos

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

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

□

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

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

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* $v_1, \ldots, v_n \in V$ *tales que* $\langle v_1, \ldots, v_n \rangle = V$*. Si* $w_1, \ldots, w_n$ *son elementos en* W*, entonces tenemos las siguientes dos propiedades.*

(i) Existe a lo sumo una transformación lineal $f \in \mathrm{Hom}_{\mathbb{K}}(V, W)$ *tal que* $f(v_i) = w_i$*, para* $i \in \{1, \ldots, n\}$*.*

(ii) Si $\{v_1, \ldots, v_n\}$ *es linealmente independiente, entonces existe una transformación* $f \in \mathrm{Hom}_{\mathbb{K}}(V, W)$ *tal que* $f(v_i) = w_i$*, para* $i \in \{1, \ldots, n\}$*.*

Dem.

(i) Considere dos transformaciones lineales $f, g \in \mathrm{Hom}_{\mathbb{K}}(V, W)$ tales que $f(v_i) = w_i = g(v_i)$, para $i \in \{1, \ldots, n\}$. Dado $v \in V$, existen $c_1, \ldots, c_n \in \mathbb{K}$ para los cuales $v = c_1 v_1 + \ldots + c_n v_n$ 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 $\{v_1, \ldots, v_n\}$ es linealmente independiente, por la proposición 1.21, se sigue que $(v_1, \ldots, v_n)$ es una base de V. Dado $v \in V$, existe un único elemento $(c_1, \ldots, c_n) \in \mathbb{K}^n$ para el que se tiene $v = c_1 v_1 + \ldots + c_n v_n$. Defina la función $f : V \to W$ por

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

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

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

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

$$\begin{aligned} f(cv_1) &= (ca_1) w_1 + \ldots + (ca_n) w_n \\ &= c(a_1 w_1 + \ldots + a_n w_n) \\ &= c f(v_1) \end{aligned}$$

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

$\square$

Proposición 1.36. *Sea $f \in \mathrm{Hom}_{\mathbb{K}}(V,W)$. 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

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

y

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

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,

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

y

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

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 $f \in \mathrm{Hom}_{\mathbb{K}}(V,W)$. Si f es una biyección, entonces decimos que f es un *isomorfismo*. Si existe un isomorfismo $f \in \mathrm{Hom}_{\mathbb{K}}(V,W)$ 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 $f : V \to V$ definida por $f(v) = \lambda v$ para todo v es una transformación lineal. Recíprocamente, si $f : V \to V$ es un 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

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

Entonces, si $\Phi : \mathbb{K} \to \mathrm{Hom}_{\mathbb{K}}(V,V)$ es la función definida por $\Phi(\lambda) = m_\lambda$, donde m_λ está definida por $m_\lambda(v) = \lambda v$ para todo $v \in V$, vemos que Φ es un isomorfismo y $\mathrm{Hom}_{\mathbb{K}}(V,V) \simeq_{\mathbb{K}} \mathbb{K}$.

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

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 $f \in \mathrm{Hom}_{\mathbb{K}}(V,W)$. Tome $(v_1, \ldots, v_n)$ una base de V, donde $n = \dim(V)$. Para $i \in \{1, \ldots, n\}$, defina $w_i = f(v_i)$. Veamos que $(w_1, \ldots, w_n)$

es base de W, y así obtener que $\dim(W) = n = \dim(V)$. Si $c_1, \ldots, c_n \in \mathbb{K}$ son tales que $c_1w_1 + \ldots + c_nw_n = O$, entonces

$$f(c_1v_1 + \ldots + c_nv_n) = c_1f(v_1) + \ldots + c_nf(v_n) = c_1w_1 + \ldots + c_nw_n = O.$$

Como f es inyectiva y $f(O) = O$, se sigue que $c_1v_1 + \ldots + c_nv_n = O$, pero al ser $\{v_1, \ldots, v_n\}$ linealmente independiente, entonces $c_1 = \ldots = c_n = 0$. De esta forma, el conjunto $\{w_1, \ldots, w_n\}$ 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 $\{v_1, \ldots, v_n\}$ genera V, existen $c_1, \ldots, c_n \in \mathbb{K}$ tales que $v = c_1v_1 + \ldots + c_nv_n$. Luego

$$w = f(v) = f(c_1v_1 + \ldots + c_nv_n) = c_1f(v_1) + \ldots + c_nf(v_n) = c_1w_1 + \ldots + c_nw_n.$$

De donde $W = \langle w_1, \ldots, w_n\rangle$ y así $(w_1, \ldots, w_n)$ 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 base $(v_1, \ldots, v_n)$ y $(w_1, \ldots, w_n)$ respectivamente de V y W. Por la proposición 1.35.(ii), existe $f \in \mathrm{Hom}(V, W)$ tal que $f(v_i) = w_i$, para $i \in \{1, \ldots, n\}$. La prueba estará completa cuando establezcamos que f es biyectiva. Sean $u, v \in V$ tales que $f(u) = f(v)$. Existen $(a_1, \ldots, a_n)$ y $(b_1, \ldots, b_n)$ en $\mathbb{K}^n$ tales que $u = a_1v_1 + \ldots + a_nv_n$ y $v = b_1v_1 + \ldots + b_nv_n$. Como $f(u - v) = f(u) - f(v) = O$ y

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

entonces

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

La independencia lineal de $\{w_1, \ldots, w_n\}$ implica que, para $i \in \{1, \ldots, n\}$, $a_i - b_i = 0$, 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 $(c_1, \ldots, c_n) \in \mathbb{K}^n$ tal que $w = c_1w_1 + \ldots + c_nw_n$. Al tomar $v = c_1v_1 + \ldots + c_nv_n$, se obtiene

$$f(v) = f(c_1v_1 + \ldots + c_nv_n) = c_1f(v_1) + \ldots + c_nf(v_n) = c_1w_1 + \ldots + c_nw_n = w.$$

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

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

(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 $\mathrm{im}(f)$.

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

Dem. Para la demostración, mostramos que los dos conjuntos contienen al origen y son cerrados mediate 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 $cw \in \operatorname{im}(f)$. Luego, de la propiedad 1.7, se sigue que $\operatorname{im}(f)$ es un subespacio de V. □

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. □

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 $\mathcal{B}_0 = (v_1, \ldots, v_n) \subseteq V$ una base de $\ker(f)$ (si $n = 0$ tomamos $\mathcal{B}_0 = \emptyset$). Extendemos $\mathcal{B}_0$ a una base $\mathcal{B} = (v_1, \ldots, v_n, v_{n+1}, \ldots, v_{n+m})$ de V.
Para $i \in \{1, \ldots, m\}$, defina $w_i = f(v_{n+i})$. Basta demostrar que $\{w_1, \ldots, w_m\}$ es una base de $\operatorname{im}(f)$, pues en tal caso tendríamos que $m = \dim(im(f)) = \rho(f) = \dim(V) - \nu(f)$. Para establecerlo usaremos la proposición 1.21.
Veamos primero que $\{w_1, \ldots, w_m\}$ es linealmente independiente. Suponga que $a_1, \ldots, a_m \in \mathbb{K}$ son tales que

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

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

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

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

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

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

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

Por ende, como $\{v_1, \ldots, v_n, v_{n+1}, \ldots, v_{n+m}\}$ es linealmente independiente, la igualdad anterior implica que $a_1 = \ldots = a_m = 0$ y la independencia lineal de $\{w_1, \ldots, w_m\}$ se sigue ahora de la proposición 1.20.
Veamos que $\{w_1, \ldots, w_m\}$ genera a $im(f)$ y con eso completamos la prueba. Como $w_1, \ldots, w_m \in \operatorname{im}(f)$ entonces $\langle w_1, \ldots, w_m \rangle \subseteq \operatorname{im}(f)$. Basta entonces establecer la otra inclusión. Dado $w \in \operatorname{im}(f)$, existe $v \in V$ tal que $f(v) = w$. Sean $c_1, \ldots, c_n, c_{n+1}, \ldots, c_{n+m} \in \mathbb{K}$ tales que

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

de forma que

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

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

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

(i) $\nu(f) = 0$

(ii) $\rho(f) = \dim_{\mathbb{K}}(V)$

Dem. Se sigue inmediatamente del teorema del rango.

Ejercicios

1. Demuestre que la función $f : \mathbb{K}^2 \to \mathbb{K}^3$ dada por $f(x, y) = (x + 2y, 2x - y, x + y)$ es una transformación lineal. Asuma que $\operatorname{char}(K) \neq 2$.

2. ¿Existe alguna transformación lineal f de $\mathbb{Q}^2$ en $\mathbb{Q}^3$ tal que $f(1, -2) = (1, 1, 1)$ y $f(-2, 4) = (1, -1, -1)$?

3. ¿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.

4. Sea $f : V \to W$ 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 $d : \mathbb{R}[t] \to \mathbb{R}[t]$ e $i : \mathbb{R}[t] \to \mathbb{R}[t]$ dadas por:

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

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

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

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

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 $\{1, \ldots, m\}$ y $\{1, \ldots, n\}$.

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 $1 \times n$

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

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 $\overline{x} \in M_{n\times 1}(\mathbb{K})$ es tal que $\overline{x}(i,1) = x_i$, para $i \in \{1,\ldots,n\}$, entonces denotaremos $\overline{x} = (x_i)$ ó

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

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

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

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

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

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

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 $I_m A = A = AI_n$.*

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 AI_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 ó AB es igual a 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 $\mathcal{B}_V = (v_1, \ldots, v_n)$, $\mathcal{B}_W = (w_1, \ldots, w_m)$ y $\mathcal{B}_U = (u_1, \ldots, u_r)$ de cada uno de estos tres espacios.

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

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

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

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

.

Proposición 1.56. *La función*

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

es un isomorfismo.

Dem. Tenemos $[v]^{\mathcal{B}_V} = O$ si y solo si $v = 0$, así la proposición se sigue del corolario 1.46 del teorema del rango.

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

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

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 $\Big[f\Big]_{\mathcal{B}_V}^{\mathcal{B}_W}$.

Observación 1.58. La columna j de $\Big[f\Big]_{\mathcal{B}_V}^{\mathcal{B}_W}$ es el vector de coordenadas $\Big[f(v_j)\Big]^{\mathcal{B}_W}$, en particular tenemos

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

y $\Big[\operatorname{id}_V\Big]_{\mathcal{B}_V}^{\mathcal{B}_V} = I_n$.

Proposición 1.59. *Dado $f \in \operatorname{Hom}_{\mathbb{K}}(V, W)$ y $v \in V$, si $A = \Big[f\Big]_{\mathcal{B}_V}^{\mathcal{B}_W}$ y $\overline{x} = \Big[v\Big]^{\mathcal{B}_V}$ se tiene que*

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

Dem. Sean $\Big[f\Big]_{\mathcal{B}_V}^{\mathcal{B}_W} = (a_{ij})$ y $\Big[v\Big]^{\mathcal{B}_V} = (c_i)$. La coordenada i del vector de coordenadas $\Big[f\Big]_{\mathcal{B}_V}^{\mathcal{B}_W} \Big[v\Big]^{\mathcal{B}_V}$ es $\sum_j a_{ij} c_j$, y así, de las igualdades

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

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 $\Big[f\Big]_{\mathcal{B}_V}^{\mathcal{B}_W} \Big[v\Big]^{\mathcal{B}_V}$. $\square$

Proposición 1.60. *Dadas* $f \in \operatorname{Hom}_{\mathbb{K}}(V,W)$ *y* $g \in \operatorname{Hom}_{\mathbb{K}}(W,U)$*, si* $A = \Big[f\Big]_{\mathcal{B}_V}^{\mathcal{B}_W}$ *y* $B = \Big[g\Big]_{\mathcal{B}_W}^{\mathcal{B}_U}$ *se tiene*

$$\Big[g \circ f\Big]_{\mathcal{B}_V}^{\mathcal{B}_U} = AB = \Big[g\Big]_{\mathcal{B}_W}^{\mathcal{B}_U} \Big[f\Big]_{\mathcal{B}_V}^{\mathcal{B}_W}.$$

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

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

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 $\Big[g\Big]_{\mathcal{B}_W}^{\mathcal{B}_U} \Big[f\Big]_{\mathcal{B}_V}^{\mathcal{B}_W}$. □

Proposición 1.61. *El mapa* $\Big[\bullet\Big]_{\mathcal{B}_V}^{\mathcal{B}_W}$ *definido por*

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

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

Dem. Sean $f, g \in \operatorname{Hom}_{\mathbb{K}}(V,W)$. Como $(f+g)(v_j) = f(v_j) + g(v_j)$ y $(cf)(v_j) = cf(v_j)$ para todo $j \in \{1,\ldots,n\}$ y todo $c \in \mathbb{K}$, entonces $\Big[\bullet\Big]_{\mathcal{B}_V}^{\mathcal{B}_W}$ 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 $f \in \operatorname{Hom}_{\mathbb{K}}(V,W)$ tal que $f(v_j) = \sum_i a_{ij}w_i$.

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

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

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

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

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

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

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

$$\left[\mathrm{id}_V\right]_{\mathcal{B}}^{\mathcal{B}'} = \left[\left[v_1\right]^{\mathcal{B}'}\Big|\ldots\Big|\left[v_n\right]^{\mathcal{B}'}\right].$$

Proposición 1.64. *Dadas dos bases* $\mathcal{B} = (v_1,\ldots,v_n)$ *y* $\mathcal{B}' = (v_1',\ldots,v_n')$ *de* V*.*

(i) Para todo $v\in V$ *se tiene* $\left[v\right]^{\mathcal{B}'} = \left[\mathrm{id}_V\right]_{\mathcal{B}}^{\mathcal{B}'}\left[v\right]^{\mathcal{B}}$.

(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

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

Dem.

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

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

□

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

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

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

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

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

tenemos

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

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

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

Si $\mathcal{C}$ es la base canónica $\{(1,0),(0,1)\}$ se tiene

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

y si $\mathcal{B}$ es la base $\{(1,1),(1,-1)\}$ se tiene

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

luego obtenemos

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

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 $\Big[v\Big]^{\mathcal{B}} \in \left(\mathbb{K}^I\right)_0$ definido por

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

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 $\phi : J \to I$.

Empecemos por establecer la suficiencia. Note que los mapas (ver notación en Observación 1.68)

$$\begin{array}{rcl} V & \longrightarrow & \left(\mathbb{K}^J\right)_0 \\ v & \longmapsto & \left[v\right]^{\mathcal{B}_V} \end{array} \qquad \text{y} \qquad \begin{array}{rcl} W & \longrightarrow & \left(\mathbb{K}^I\right)_0 \\ w & \longmapsto & \left[w\right]^{\mathcal{B}_W} \end{array},$$

son isomorfismos. Ahora, dada una biyección $\phi : J \to I$ entre los indices de las bases, la transformación lineal $\Phi \in \operatorname{Hom}_{\mathbb{K}}\left(\left(\mathbb{K}^J\right)_0, \left(\mathbb{K}^I\right)_0\right)$ definida en la base $\{\delta_j\}_{j\in J}$ de $\left(\mathbb{K}^J\right)_0$ (ver el ejemplo 1.22.(iii)) por $\Phi(\delta_j) = \delta_{\phi(j)}$ es un isomorfismo. Tenemos así $V \simeq_{\mathbb{K}} \left(\mathbb{K}^J\right)_0 \simeq_{\mathbb{K}} \left(\mathbb{K}^I\right)_0 \simeq_{\mathbb{K}} W$.
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 $f \in \operatorname{Hom}_{\mathbb{K}}(W, V)$ 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 indices $j' \in J'$ para los cuales v_j tiene coordenada diferente de cero en la base $\mathcal{B}'$, es decir tenemos $J'_j = \left\{ j' \in J' \mid \left[v_j\right]^{\mathcal{B}'_V}_{j'} \neq 0 \right\}$. En particular J'_j es la mínima colección de indices $j' \in J'$ 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 $j' \in J' \backslash \bigcup_{j\in J} J'_j$ y $\{v_{j_1}, \ldots, v_{j_n}\} \subseteq \mathcal{B}_V$ es tal que $v_{j'}$ está en $\langle v_{j_1}, \ldots, v_{j_n} \rangle$, entonces $v_{j'}$ está en $\langle v_{k'} | \; k' \in \bigcup_{k=1}^n J'_{j_k} \rangle$ que es un subespacio de $\langle v_{k'} | \; k' \in J' \backslash \{j'\} \rangle$, 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 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 $f \in \operatorname{Hom}_{\mathbb{K}}(V, W)$, entonces existe una base de $\mathcal{B}$ de V tal que $\mathcal{B}$ es igual a la union 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 $\operatorname{im}(f)$. 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 $f_0 : C \to W$, existe a los sumo una transformación lineal

$f \in \mathrm{Hom}_{\mathbb{K}}(V, W)$ 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.

1.5. Suma y producto directo

Sea V un espacio vectorial sobre un cuerpo K.

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

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

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

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 $V_1 \cap V_2$ también tiene dimensión finita, por el teorema 1.28. 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

$$0 = \sum_{i=1}^{p} a_i v_i + \sum_{i=p+1}^{n_1} a_i' v_i' + \sum_{i=p+1}^{n_2} a_i'' v_i''.$$

Luego, si v es el vector $\sum_{i=1}^{p} a_i v_i + \sum_{i=p+1}^{n_1} a'_i v'_i$, entonces v está en V_1 y tenemos $v = \sum_{i=p+1}^{n_2} a''_i v''_i$, y así v está también en V_2, y por ende v está en $V_1 \cap V_2$. Sean $b_1 \ldots, b_p \in K$ para los cuales se tiene

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

entonces obtenemos

$$0 = v - v = \sum_{i=1}^{p} (a_i - b_i) v_i + \sum_{i=p+1}^{n_1} a'_i v'_i.$$

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

$$0 = \sum_{i=1}^{p} a_i v_i + \sum_{i=p+1}^{n_2} a''_i v''_i.$$

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

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 $\{v_1, \ldots, v_s\}$ es una es una colección de s vectores linealmente independientes en V basta tomar $V_1 = \langle v_1, \ldots, v_{n_1} \rangle$ y $V_2 = \langle v_1, \ldots, v_i, v_{n_1+1}, \ldots, v_s \rangle$. Más aún, si V'_1 y V'_2 son subespacios de V para los 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 $f \in \operatorname{End}_K(V)$ 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 $V_1, \ldots, V_n \leq V$, decimos que V es la *suma directa* de $V_1, \ldots, V_n$, lo cual denotamos por

$$V = V_1 \oplus \ldots \oplus V_n = \bigoplus_{i=1}^{n} V_i$$

si para cada $v \in V$ existe un único $(v_1, \ldots, v_n) \in V_1 \times \ldots \times V_n$ que cumple

$$v = v_1 + \ldots + v_n$$

Proposición 1.79. *Sean $V_1, \ldots, V_n \leq V$. Se tiene $V = \bigoplus_{i=1}^n V_i$ si y solo si $V_1, \ldots, v_n$ satisfacen las siguientes dos propiedades.*

1. *La suma $\sum_{i=1}^n V_i$ es igual a V.*

2. *Para todo $i \in \{1, \ldots, n\}$ se tiene $V_i \cap \sum_{j \neq i} V_j = \{O\}$.*

Dem. Suponga primero que tenemos $V = \bigoplus_{i=1}^n V_i$, luego por definición tenemos $V = \sum_{i=1}^n V_i$. Por otro lado, sea $i \in \{1, \ldots, n\}$ y tome $v \in V_i \cap \sum_{j \neq i} V_j$. Así, existe $(v_1, \ldots, v_n) \in V_1 \times \ldots \times V_n$ para el cual se cumple $v = -v_i = \sum_{j \neq i} v_j$ y por lo cual se tiene $O = v_1 + \ldots + v_n$. Pero por otro lado, para $(O, \ldots, O) \in V_1 \times \ldots \times V_n$ se tiene $O = O + \ldots + O$, luego, por unicidad de esta descomposición se siguen las igualdades $v_1 = \ldots = v_n = O$, $v = O$ y $V_i \cap \sum_{j \neq i} V_j = \{O\}$.
Recíprocamente, suponga que $\sum_{i=1}^n V_i$ es igual a V y que para todo $i \in \{1, \ldots, n\}$ se tiene $V_i \cap \sum_{j \neq i} V_j = \{O\}$. Sea $v \in V$. Entonces existe $(v_1, \ldots, v_n) \in V_1 \times \ldots \times V_n$ para el cual se tiene $v = v_1 + \ldots + v_n$. Veamos que esta descomposición es única. De hecho, si $(v_1', \ldots, v_n') \in V_1 \times \ldots \times V_n$ es tal que se tiene $v = v_1' + \ldots + v_n'$, dado $i \in \{1, \ldots, n\}$, 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 $V_1, \ldots, V_n \leq V$ para los cuales se tiene $V = \sum_{i=1}^n V_i$ y suponga que V tiene dimensión finita. Entonces la siguientes propiedades son equivalentes.*

1. *Para todo $i \in \{1, \ldots, n\}$ se tiene $V_i \cap \sum_{j \neq i} V_j = \{0\}$.*

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

Dem. Suponga primero que tenemos $V_i \cap \sum_{j \neq i} V_j = \{0\}$, para todo $i \in \{1, \ldots, n\}$. Por el teorema 1.74 se obtiene

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

La inclusión $\left(V_2 \cap \sum_{j>2} V_j\right) \subseteq \left(V_2 \cap \sum_{j \neq 2} V_j\right)$ implica la igualdad $\left(V_2 \cap \sum_{j>2} V_j\right) = \{O\}$. 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

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

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

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

Pero se tiene $\sum_{i=1}^{n} \dim(V_i) = \dim(V)$, 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 $\dim\left(V_1 \cap \sum_{j\neq 1} V_j\right) = 0$, es decir $V_1 \cap \sum_{j\neq 1} V_j = \{O\}$. Reordenando los subespacios V_i, $i = 1, \ldots, n$, obtenemos $V_i \cap \sum_{j\neq i} V_j = \{0\}$, para todo $i \in \{1, \ldots, n\}$. $\square$

Definición 1.81. Sea $p \in \mathrm{End}_K(V)$. Decimos que p es una *proyección* si se tiene $p \circ p = p$.

Observación 1.82. Si $p \in \mathrm{Hom}_K(V, V)$ es una proyección y V_0 es la imagen de p entonces se tiene $p(v_0) = v_0$ 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 1.83. Suponga que tenemos $V = V_1 \oplus V_2$, y defina los operadores $p_1, p_2 \in \mathrm{End}_K(V)$ 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 $p_1 + p_2 = \mathrm{id}_V$. Similarmente, si tenemos $V = \bigoplus_{i=1}^{n} V_i$, podemos definir n proyecciones $p_1, \ldots, p_n$ que satisfacen $p_i(V) = V_i$, para cada $i \in \{1, \ldots, n\}$, $\sum_{i=1}^{n} p_i = \mathrm{id}_V$, y $p_i \circ p_j = \underline{O}$ si $i \neq j$. Esto nos sugiere otra forma de caracterizar sumas directas, como lo sugiere el siguiente teorema.

Teorema 1.84. *Sean $p_1, \dots, p_n \in \mathrm{End}_K(V)$ proyecciones y $V_1, \dots, V_n$ sus respectivas imágenes. Si se tiene $\sum_{i=1}^n p_i = \mathrm{id}_V$ y $p_i \circ p_j = 0$ para $i \neq j$, entonces se tiene $V = \bigoplus_{i=1}^n V_i$.*

Dem. Usamos la propiedad 1.79 para establecer este teorema. Para ver que tenemos $V = \sum_{i=1}^n V_i$, dado $v \in V$, definimos $v_i \in V_i$ por $v_i = p_i(v)$, para $i \in \{1, \dots, n\}$, y obtenemos

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

Para establecer $V_i \cap \sum_{j\neq i} V_j = \{O\}$, tome $i \in \{1, \dots, n\}$ y $v \in V_i \cap \sum_{j\neq i} V_j$ y veamos que se sigue $v = O$. 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, existe $v'_j \in V$ que satisface $v_j = p_j(v'_j)$, para cada $j \neq i$, 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

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

□

Observación 1.85. Para terminar está 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. Cuando combinamos una colección finita de espacios, obtenemos espacio isomorfos.

Definición 1.86. Sea I una colección de indices y $\{V_i\}_{i\in I}$ una familia de espacios vectoriales sobre K.

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

   $$\prod_{i\in I} V_i = \left\{\phi : I \to \coprod_{i\in I} V_i \;\middle|\; \phi(i) \in V_i\right\},$$

   el cual es un espacio vectorial sobre K bajo las operaciones

   $$(\phi + \psi)(i) = \phi(i) + \psi(i) \qquad (a\phi)(i) = a\psi(i)$$

   para todo $\phi, \psi \in \prod_{i\in I} V_i$ y $a \in K$; y,

2. la *suma directa externa* de $\{V_i\}_{i\in I}$ por

   $$\bigoplus_{i\in I} V_i = \left\{\phi \in \prod_{i\in I} V_i \;\middle|\; \phi(i) \neq 0 \text{ únicamente para finitos indices } i \in I\right\},$$

   el cual es un subespacio de $\prod_{i\in I} V_i$.

Si además $\{W_i\}_{i\in I}$ es otra familia de espacios vectoriales sobre K, y para cada $i \in I$ tenemos un $f_i \in \mathrm{Hom}_K(V_i, W_i)$, definimos:

1. el *producto externo* de $\{f_i\}_{i\in I}$ por
$$\begin{array}{rcl}\prod_{i\in I} f_i : \prod_{i\in I} V_i & \longrightarrow & \prod_{i\in I} W_i \\ \phi & \longmapsto & \left(\prod_{i\in I} f_i\right)(\phi) : i \mapsto f_i\left(\phi(i)\right),\end{array}$$
el cual es una transformación lineal; y,

2. la *suma directa externa* de $\{f_i\}_{i\in I}$ por
$$\begin{array}{rcl}\bigoplus_{i\in I} f_i : \bigoplus_{i\in I} V_i & \longrightarrow & \bigoplus_{i\in I} W_i \\ \phi & \longmapsto & \left(\prod_{i\in I} f_i\right)(\phi) : i \mapsto f_i\left(\phi(i)\right),\end{array}$$
la cual es la transformación lineal inducida por $\prod_{i\in I} f_i$ entre los subespacios $\bigoplus_{i\in I} V_i$ y $\bigoplus_{i\in I} W_i$.

Observación 1.87. Note que si $\{V_i\}_{i\in I}$ es una colección de espacios vectoriales sobre K indexada por los indices $i \in I$; y, $f_i \in \mathrm{Hom}_K(V, V_i)$ y $g_i \in \mathrm{Hom}_K(V_i, V)$, para todo $i \in I$, podemos definir las transformaciones lineales

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

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

Observación 1.88. Note que, si $\mathcal{B} = \{v_i\}_{i\in I} \subseteq V$ es una base, entonces
$$V \simeq_K \left(K^I\right)_0 \simeq_K \bigoplus_{i\in I} K,$$
y
$$K^I \simeq_K \prod_{i\in I} K$$

1.6. Espacios cocientes

Sea K un cuerpo y V, W espacios vectoriales sobre K.

Definición 1.89. Sean $V_0 \leq V$ y $v \in V$. Definimos la *translación de V_0 por v* como el conjunto
$$v + V_0 = \{v' \in V \mid v' = v + v_0,\ v_0 \in V\}.$$

Observación 1.90. 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 $v \in v + V_0 = v' + V_0$, 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 1.91. 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 1.92. *Sean $V_0 \leq V$, $v, w, v', w' \in V$ y $a \in K$. Si $v + V_0 = w + V_0$ y $v' + V_0 = w' + V_0$ entonces $(v + v') + V_0 = (w + w') + V_0$ y $av + V_0 = aw + V_0$.*

Dem. $v + V_0 = w + V_0$ y $v' + V_0 = w' + V_0$ si y solo si $v - w \in V_0$ y $v' - w' \in V_0$, en tal caso $(v+v') - (w+w') = (v-w) + (v'-w') \in V_0$, es decir $(v+v') + V_0 = (w + w') + V_0$, y $av - aw = a(v - w) \in V_0$, es decir $av + V_0 = aw + V_0$. □

Proposición 1.93. *Sea $V_0 \leq V$. El espacio cociente V/V_0 es un espacio vectorial sobre K bajo las operaciones*

$$(v + V_0) + (v' + V_0) = (v + v') + V_0 \qquad a\,(v + V_0) = av + V_0,$$

y su origen es $0 + V_0 = V_0$. 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 estas en Definición 1.1 se heredan de las de V. La misma proposición implica la linearidad de π_{V_0}. Por definición de V/V_0, π_{V_0} es sobreyectiva. Por último, $v \in \ker(\pi_{V_0})$ si y solo si $\pi_{V_0}(v) = V_0$, es decir si y solo si $v + V_0 = V_0$, o si y solo si $v \in V_0$. □

Proposición 1.94. *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 de Teorema 1.45 y de la propiedad anterior.

Teorema 1.95. *Sean $f \in \mathrm{Hom}_K(V, W)$ y $V_0 = \ker(f)$. Entonces existe una única transformación lineal $f_{V_0} \in \mathrm{Hom}_K(V/V_0, W)$ tal que $f = f_{V_0} \circ \pi_{V_0}$. La transformación f_{V_0} es inyectiva, y, si f es además sobreyectiva, f_{V_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

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

Así, f_{V_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 $f = f_{V_0} \circ \pi_{V_0}$. Ahora si f es sobreyectiva, entonces f_{V_0} es biyectiva y así un isomorfismo. $\square$

Capítulo 2

Estructura de las transformaciones lineales

Sea K un cuerpo y V, W espacios vectoriales sobre K.

Notación 2.1. Suponga que V y W tienen dimensión finita y denote $n = \dim(V)$, $m = \dim(W)$. Sean $\mathcal{B}_V = \{v_1, \ldots, v_n\} \subseteq V$ y $\mathcal{B}_W = \{w_1, \ldots, w_m\} \subseteq W$ bases. Dada $f \in \mathrm{Hom}_K(V, W)$, la matriz $\{1, \ldots, m\} \times \{1, \ldots, n\}$

$$A = \Big[f\Big]_{\mathcal{B}_V}^{\mathcal{B}_W},$$

que representa a f respecto a las bases $\mathcal{B}_V$ y $\mathcal{B}_W$, se denota por un arreglo rectangular $m \times n = |\mathcal{B}_W| \times |\mathcal{B}_V|$, con entradas en K, cuya ij-ésima entrada es

$$a_{ij} = \Big[f(v_j)\Big]_i^{\mathcal{B}_W}.$$

De forma que

$$f(v_j) = \sum_{i=1}^{n} a_{ij} w_i.$$

En tal caso identificaremos a la matriz A con el arreglo

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

Igualmente, a las matrices $\{1, \ldots, n\} \times \{*\}$ y $\{1, \ldots, m\} \times \{*\}$ de coordenadas en las bases $\mathcal{B}_V$ y $\mathcal{B}_W$ las identificaremos con los arreglos $n \times 1$ y $m \times 1$ con entradas en K, de tal forma que para $v \in V$ y $w \in W$ escribimos

$$\Big[v\Big]^{\mathcal{B}_V} = \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix}, \qquad \Big[w\Big]^{\mathcal{B}_W} = \begin{bmatrix} d_1 \\ \vdots \\ d_m \end{bmatrix},$$

cuando $v = \sum_{j=1}^{n} c_j v_j$ y $w = \sum_{i=1}^{m} d_i w_i$. En particular

$$\Big[f(v)\Big]^{\mathcal{B}_W} = \Big[f\Big]_{\mathcal{B}_V}^{\mathcal{B}_W} \Big[v\Big]^{\mathcal{B}_V} = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix} = \begin{bmatrix} \sum_j a_{1j} c_j \\ \vdots \\ \sum_j a_{mj} c_j \end{bmatrix}.$$

A los arreglos $m \times n$ los llamaremos también *matrices* $m \times n$ y el espacio de estas lo denotamos por $M_{m\times n}(K)$.

Observación 2.2. Sean $A = (a_{ij})_{i,j=1}^{n}$ y $C = (c_{ij})_{i,j=1}^{n}$ matrices $n\times n$, entonces

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

Ahora, si V tiene dimensión finita igual a n, y $f \in \operatorname{Hom}_K(V,V)$, dadas dos bases $\mathcal{B}, \mathcal{B}' \subseteq V$, tenemos dos matrices $n \times n$ que representan a f, $A = \Big[f\Big]_{\mathcal{B}}^{\mathcal{B}}$ y $B = \Big[f\Big]_{\mathcal{B}'}^{\mathcal{B}'}$. Entonces, si además $C = \Big[\operatorname{id}_V\Big]_{\mathcal{B}'}^{\mathcal{B}}$,

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

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 la traza y el determinante de una matriz de representación de un operador lineal, respecto a la misma base para el dominio y el rango, es independiente de la base escogida.

Definición 2.3. Suponga que V tiene dimensión finita, sean $f \in \operatorname{Hom}_K(V,V)$ y $\mathcal{B} \subseteq V$ una base. Definimos el *determinante* y la *traza* de f respectivamente por

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

2.1. Descomposición directa

Definición 2.4. Sean $V_1, V_2 \leq V$, decimos que V_1 y V_2 forman una *descomposición directa* de V si $V = V_1 \oplus V_2$.

Teorema 2.5. *Sea $f \in \operatorname{Hom}_K(V,W)$. Entonces existe descomposiciones directas $V = V_0 \oplus V_1$ y $W = W_1 \oplus W_2$ tales que* $\ker(f) = V_0$, $\operatorname{im}(f) = W_1$. *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 $\mathcal{B} = \mathcal{B}_0 \cup \mathcal{B}_1$ de V. Defina $V_1 = \langle \mathcal{B}_1 \rangle$. Así pues $V_0 + V_1 = V$ y $V_0 \cap V_1 = \{0\}$, en particular $V = V_0 \oplus V_1$. 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 f es inyectiva en V_1, es decir $f(v) = 0$ con $v \in V_1$ si y solo si $v = 0$, $\mathcal{B}_1'$ 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 $\mathrm{im}(f) = W_1$; pues, dado $w \in \mathrm{im}(f)$, existe $v \in V$ tal que $w = f(v)$, si $v = v_0 + v_1$ con $(v_0, v_1) \in V_0 \times V_1$, $w = f(v) = f(v_0) + f(v_1) = f(v_1)$. Finalmente, extienda $\mathcal{B}_1'$ a una base $\mathcal{B}' = \mathcal{B}_1' \cup \mathcal{B}_2'$ de W. Si $W_2 = \langle \mathcal{B}_2' \rangle$, $V = V_0 \oplus V_1$ y $W = W_1 \oplus W_2$ son las descomposiciones directas buscadas. Como f es inyectiva en V_1 y $f(V_1) = W_1$, f induce un isomorfismo entre V_1 y W_1. $\square$

Corolario 2.6. *Suponga que V y W tienen dimensión finita, sea $f \in \mathrm{Hom}_K(V, W)$, y denote $n = \dim(V)$, $m = \dim(W)$ y $r = \dim(\mathrm{im}(f))$. Entonces existen bases $\mathcal{B} = \{v_j\}_{j=1}^n \subseteq V$ y $\mathcal{B}' = \{w_i\}_{i=1}^m \subseteq W$ tales que, si*

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

$a_{ii} = 1$ 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 $r \times r$ con unos en diagonal y ceros en el resto de entradas y 0 los orígenes de $M_{r\times(n-r)}(K)$, $M_{(m-r)\times r}(K)$ y $M_{(m-r)\times(n-r)}(K)$.

Dem. Tome $\mathcal{B}_0, \mathcal{B}_1, \mathcal{B}_1'$ y $\mathcal{B}_2'$ como en la prueba del teorema, y denote $v_1, \ldots, v_n \in V$ y $w_1, \ldots, w_m \in W$ de forma que

$$\mathcal{B}_1 = \{v_1, \ldots, v_r\}, \mathcal{B}_0 = \{v_{r+1}, \ldots, v_n\}, \mathcal{B}_1' = \{w_1, \ldots, w_r\}, \mathcal{B}_2' = \{w_{r+1}, \ldots, v_m\}.$$

Las bases $\mathcal{B} = \{v_1, \ldots, v_n\}$ y $\mathcal{B}' = \{w_1, \ldots, w_m\}$ son tales que $\Big[f\Big]_{\mathcal{B}}^{\mathcal{B}'}$ tiene la forma buscada. $\square$

2.2. Espacios invariantes y espacios propios

Definición 2.7. Sean $f \in \mathrm{Hom}_K(V, V)$ y $V_0 \leq V$. Decimos que V_0 es *invariante bajo f* si $f(V_0) \subseteq V_0$. La restricción de f a V_0 la denotamos f_{V_0}, es decir $f_{V_0} \in \mathrm{Hom}_K(V_0, V_0)$ es el operador definido por:

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

Definición 2.8. Sean I un conjunto y $A \in M_{I\times I}(K)$. Decimos que A es *diagonal* si $A(i,j) = 0$ siempre que $i \neq j$. Sea $f \in \mathrm{Hom}_K(V,V)$, decimos que f es diagonalizable si $\left[f\right]_{\mathcal{B}}^{\mathcal{B}}$ es diagonal para alguna base $\mathcal{B}$ de V.

Teorema 2.9. *Sea $f \in \mathrm{Hom}_K(V,V)$. Entonces f es diagonalizable si y solo si existe una familia $\{V_i\}_{i\in I}$ de subespacios unidimensional de V, invariantes bajo f, tal que $V = \bigoplus_{i\in I} V_i$.*

Dem. Note primero que si $\mathcal{B} = \{v_i\}_{i\in I} \subseteq V$ es una base, entonces

$$V = \bigoplus_{i\in I} \langle v_i \rangle.$$

Suponga primero que f es diagonalizable y sea $\mathcal{B} = \{v_i\}_{i\in I} \subseteq V$ base tal que $\left[f\right]_{\mathcal{B}}^{\mathcal{B}}$ es diagonal. Para cada $i \in I$ defina $V_i = \langle v_i \rangle$. Ahora, dados $i, j \in I$,

$$\left[f(v_j)\right]_i^{\mathcal{B}} = \sum_{l\in I} \left[f\right]_{\mathcal{B},(i,l)}^{\mathcal{B}} \left[v_j\right]_l^{\mathcal{B}} = \left[f\right]_{\mathcal{B},(i,j)}^{\mathcal{B}}.$$

Así, como $\left[f\right]_{\mathcal{B}}^{\mathcal{B}}$ es diagonal,

$$f(v_j) = \sum_{i\in I} \left[f(v_j)\right]_i^{\mathcal{B}} v_i = \sum_{i\in I} \left[f\right]_{\mathcal{B},(i,j)}^{\mathcal{B}} v_i = \left[f\right]_{\mathcal{B},(j,j)}^{\mathcal{B}} v_j,$$

es decir que si $\lambda_j = \left[f\right]_{\mathcal{B},(j,j)}^{\mathcal{B}}$, entonces $f(v_j) = \lambda_j v_j \in V_j$, luego V_j es invariante bajo f. De donde

$$V = \bigoplus_{j\in I} V_j$$

es una descomposición de V en espacios unidimensional invariantes bajo f.
Suponga ahora que $V = \bigoplus_{i\in I} V_i$, donde $\{V_i\}_{i\in I}$ es una familia subespacios unidimensional de V invariantes bajo f. Para cada $i \in I$ sea $v_i \in V_i$, con $v_i \neq 0$, de tal forma que $V_i = \langle v_i \rangle$. Luego

$$\mathcal{B} = \{v_i\}_{i\in I},$$

es una base de V; y, además, como cada V_i es invariante bajo f y unidimensional, $f(v_i) = \lambda_i v_i$ para algún $\lambda_i \in K$. Así pues

$$\left[f\right]_{\mathcal{B},(i,j)}^{\mathcal{B}} = \left[f(v_j)\right]_i^{\mathcal{B}} = \begin{cases} \lambda_i & \text{si } i = j \\ 0 & \text{si } i \neq j \end{cases},$$

es decir $\left[f\right]_{\mathcal{B}}^{\mathcal{B}}$ es diagonal. $\square$

Definición 2.10. Sea $f \in \mathrm{Hom}_K(V,V)$ y $V_0 \leq V$ con $\dim(V_0) = 1$. Decimos que V_0 es un *espacio propio* de f si V_0 es invariante bajo f. En tal caso, a los elementos en V_0 diferentes del origen los llamamos *vectores propios* de f. Dado un vector propio v en V_0, existe $\lambda \in K$ tal que $f(v) = \lambda v$; a este λ lo llamamos *valor propio* (asociado a V_0 o a v) de f. Igualmente en tal caso, decimos que V_0 es un espacio propio (ó v es un vector propio) asociado a λ.

Observación 2.11. Del mismo modo en que definimos arreglos $m \times n$, donde n y m son enteros positivos, con entradas en K, podemos definir arreglos $m \times n$ con entradas en conjunto de polinomios con coeficientes en K en la variable t. A este conjunto lo denotaremos $M_{m\times n}(K[t])$. Los elementos en $K[t]$ se pueden multiplicar y sumar entre si en base a operaciones de multiplicación y suma de K. De esta forma podemos igualmente hablar del determinante y de la traza de un matriz $n \times n$ con entradas en $K[t]$, los cuales serán igualmente polinomios en $K[t]$.

Observación 2.12. Sea $n \in \mathbb{Z}_{>0}$ y $A \in M_{n\times n}(K)$. Dada cualquier $C \in M_{n\times n}(K)$, invertible, tenemos

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

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

Definición 2.13. Suponga que V tiene dimensión finita y denote $n = \dim(V)$. Dado $f \in \mathrm{Hom}_K(V,V)$, definimos el *polinomio caraterístico* de f por

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

donde $A = \Big[f\Big]_{\mathcal{B}}^{\mathcal{B}}$, y $\mathcal{B} \subseteq V$ es una base.

Teorema 2.14. *Suponga que V tiene dimensión finita. Sean $f \in \mathrm{Hom}_K(V,V)$ y $\lambda \in K$. Entonces, λ es un valor propio de f si y solo si $P_f(\lambda) = 0$.*

Dem. Sea $\mathcal{B} \subseteq V$ una base. El escalar $\lambda \in K$ 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 $(\lambda \,\mathrm{id}_V - f)(v) = 0$. Es decir $\lambda \in K$ es un valor propio de f si y solo si $\lambda \,\mathrm{id}_V - f$ no es inyectiva, lo que equivale a

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

□

Definición 2.15. Sean $P(t) \in K[t]$ y $f \in \mathrm{Hom}_K(V,V)$. Definimos el operador $P(f) \in \mathrm{Hom}_K(V,V)$ por

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

cuando $P(t) = a_n t^n + a_{n-1} t^{n-1} + \ldots + a_1 t + a_0$, donde para todo entero positivo k

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

Observación 2.16. 1. Sea $C \in M_{m\times n}(K[t])$, donde m y n son enteros positivos, cuya ij-ésima entrada denotamos $c_{ij}(t)$. Dado $f \in \mathrm{Hom}_K(V, V)$ definimos la transformación lineal

$$C_f : \underbrace{V \times \ldots \times V}_{n-\text{veces}} \longrightarrow \underbrace{V \times \ldots \times V}_{m-\text{veces}} \tag{2.1}$$

por

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

2. Note que si $C_1 \in M_{m\times n}(K[t])$ y $C_2 \in M_{l\times m}(K[t])$, donde l, m y n son enteros positivos, dado $f \in \mathrm{Hom}_K(V, V)$,

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

3. Dado $B \in M_{n\times n}(K[t])$, donde n es un entero positivo, cuya ij-ésima entrada es $b_{ij}(t)$, denotamos por $\tilde{B}$ su matriz de cofactores, es decir la matriz $n \times n$ con entradas en $K[t]$ cuya ij-ésima entrada es

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

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

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

donde $\tilde{B}^{\intercal}$ es la transpuesta de $\tilde{B}$, es decir la matriz $n \times n$ cuya ij-ésima entrada es la entrada ji-ésima de $\tilde{B}$; similarlmente para $B^{\intercal}$ y $(B\tilde{B}^{\intercal})^{\intercal}$.

Teorema 2.17 (Caley-Hamilton)**.** *Suponga que V tiene dimensión finita y sea $f \in \mathrm{Hom}_K(V, V)$. Entonces $P_f(f) = 0$.*

Dem. Sean $n = \dim(V)$ y $\mathcal{B} = \{v_1, \ldots, v_n\} \subseteq V$ una base. Defina

$$\Big[f\Big]_{\mathcal{B}}^{\mathcal{B}} = A = (a_{ij})_{i,j=1}^{n}$$

de forma que

$$f(v_j) = \sum_{i=1}^{n} a_{ij} v_i.$$

Considere la matriz $B = tI_n - A \in M_{n\times n}(K[t])$. Entonces $\tilde{B}B^{\intercal} = P_f(t)I_n$. Ahora

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

por un lado; pero, por el otro

$$\begin{aligned}(P_f(f)(v_1),\ldots,P_f(f)(v_n)) &= (P_f(t)I_n)_f\,(v_1,\ldots,v_n)\\ &= \left(\tilde{B}B^T\right)_f(v_1,\ldots,v_n)\\ &= \tilde{B}_f\circ(B^{\intercal})_f\,(v_1,\ldots,v_n)\\ &= \tilde{B}_f(0,\ldots,0)\\ &= (0,\ldots,0).\end{aligned}$$

Luego $\mathcal{B}\subseteq \ker(P_f(f))$ y así $P_f(f)=0$. □

Observación 2.18. Note que si $P_1(t),P_2(t)\in K[t]$, $P(t)=P_1(t)P_2(t)$ y $f\in \mathrm{Hom}_K(V,V)$, entonces $P(f)=P_1(f)\circ P_2(f)=P_2(f)\circ P_1(f)$, pues $(af^m)\circ(bf^n)=(bf^n)\circ(af^m)$ para todo $n,m\in\mathbb{Z}_{\geq 0}$ y $a,b\in K$.

Propiedad 2.19. *Sea $P(t)\in K[t]$ y $f\in \mathrm{Hom}_K(V,V)$, entonces $V_0=\ker(P(f))$ es invariante bajo f.*

Dem. Sea $v\in V_0$, luego $P(f)\,(f(v))=P(f)\circ f(v)=f\circ P(f)(v)=f(0)=0$. Es decir $f(v)\in\ker(P(f))=V_0$. □

Propiedad 2.20. *Sean $P(t)\in K[t]$ y $f\in\mathrm{Hom}_K(V,V)$ tales que $P(f)=0$. Si $P_1(t),P_2(t)\in K[t]$ son tales que $P(t)=P_1(t)P_2(t)$ y $(P_1(t),P_2(t))=1$, entonces*

$$V=V_1\oplus V_2$$

donde $V_1=\ker(P_1(f))$ y $V_2=\ker(P_2(f))$. Más aún V_1 y V_2 son invariantes bajo f y existen polinomios $\Pi_1(t),\Pi_2(t)\in K[t]$, tales que

$$\Pi_1(f)=p_1 \qquad y \qquad \Pi_2(f)=p_2$$

son las proyecciones en V_1 y V_2.

Dem. Sean $Q_1,Q_2\in K[t]$ tales que $Q_1(t)P_1(t)+P_2(t)Q_2(t)=1$, luego

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

en particular, dado $v\in V$

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

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 $V = V_1 + V_2$. Ahora si asumimos que $v \in V_1 \cap V_2$, $P_1(f)(v) = 0 = P_2(f)(v)$, entonces $v_1 = 0 = v_2$, luego $v = 0$.
Por la propiedad anterior V_1 y V_2 son invariantes bajo f. 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

$$\Pi_2(f) + \Pi_1(f) = \mathrm{id}_V .$$

Ahora,

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

luego

$$\Pi_1(f) \circ \Pi_2(f) = 0,$$

y, como

$$\Pi_2(t) = \Pi_2(t)\left(\Pi_2(t) + \Pi_1(t)\right) = \left(\Pi_2(t)\right)^2 + \Pi_2(t)\Pi_1(t)$$

entonces

$$\Pi_2(f) = \left(\Pi_2(f)\right)^2 .$$

Similarmente obtenemos

$$\Pi_1(f) = \left(\Pi_1(f)\right)^2 .$$

Luego, si $\Pi_1(f) = p_1$ y $\Pi_2(f) = p_2$, por Teorema 1.84, p_1 y p_2 son proyecciones sobre V_1 y V_2 respectivamente. $\square$

Ejemplo 2.21. Sea $p \in \mathrm{Hom}_K(V, V)$ una proyección, es decir $p^2 = p$. Si $P(t) = t^2 - t$ entonces $P(p) = p^2 - p = 0$ y $P(t) = t(t-1) = P_1(t)P_2(t)$ donde $P_1(t) = t - 1$ 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 si

$$V_1 = \ker\left(P_1(p)\right) = \ker\left(p - \mathrm{id}_V\right)$$
$$V_2 = \ker\left(P_2(p)\right) = \ker\left(p\right)$$

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

Ejemplo 2.22. Suponga que $\mathrm{char}(K) \neq 2$, de forma que $-1 \neq 1$. Sea $f \in \mathrm{Hom}_K\left(K^2, K^2\right)$ el operador definido por

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

Si $\mathcal{C} = \{(1,0),(0,1)\}$ es la base canónica entonces

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

y $P_f(t) = t^2 - 1 = (t-1)(t+1)$. Por el teorema del Caley-Hamilton $P_f(f) = 0$, entonces si $P_1(t) = t-1$ y $P_2(t) = t+1$, por la propiedad anterior, $K^2 = V_1 \oplus V_2$ donde $V_1 = \ker(f - \mathrm{id}_{K^2})$ y $V_2 = \ker(f + \mathrm{id}_{K^2})$. Como

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

entonces

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

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 2.23. Note que bajo las condiciones de la propiedad anterior, si denotamos por $f_i \in \mathrm{Hom}_K(V_i, V_i)$, para $i = 1,2$ la restricción de f a V_i, es decir $f_i(v_i) = f(v_i) \in V_i$ para todo $v_i \in V_i$, tenemos que $P_i(f_i) = 0$, pues $V_i = \ker(P_i(f))$ así que $P_i(f_i)(v_i) = P_i(f)(v_i) = 0$. Así, inductivamente, podemos aplicar la propiedad a cualquier descomposición de $P_i(t)$ en factores primos relativos para obtener el siguiente resultado.

Propiedad 2.24. *Sean $P(t) \in K[t]$ y $f \in \mathrm{Hom}_K(V,V)$ tales que $P(f) = 0$. Si $P_1(t), P_2(t), \ldots, P_n(t) \in K[t]$ son tales que $P(t) = P_1(t)P_2(t)\ldots P_n(t)$ y $(P_i(t), P_j(t)) = 1$ siempre que $i \neq j$, entonces*

$$V = V_1 \oplus V_2 \oplus \ldots \oplus V_n$$

donde $V_i = \ker(P_i(f))$, $i = 1,\ldots,n$. Más aún cada V_i es invariante bajo f y existen polinomios $\Pi_1(t),\ldots,\Pi_n(t) \in K[t]$, tales que

$$\Pi_1(f) = p_1 \qquad ,\ldots, \qquad \Pi_n(f) = p_n$$

son las proyecciones sobre $V_1,\ldots,V_n$.

Dem. Falta mostrar la existencia de $\Pi_1(t),\ldots,\Pi_n(t) \in K[t]$. De hecho, para $i = 1,\ldots,n$ sea

$$R_i(t) = \prod_{j \neq i} P_j(t),$$

Entonces $(R_1(t), \dots, R_n(t)) = 1$. Tome $Q_1(t), \dots, Q_n(t) \in K[t]$ tales que

$$Q_1(t)R_1(t) + \dots Q_n(t)R_n(t) = 1.$$

De forma que, si $\Pi_i(t) = Q_i(t)R_i(t)$, $i = 1, \dots, n$. Entonces, similarmente a la demostración anterior obtenemos

$$\Pi_1(f) + \dots + \Pi_n(f) = \mathrm{id}_V,$$

$\Pi_i(f) \circ \Pi_j(f) = 0$, si $i \neq j$, y $(\Pi_i(f))^2 = \Pi_i(f)$. El resultado se sigue de Teorema 1.84. □

Ejemplo 2.25. Sea $f \in \mathrm{Hom}_{\mathbb{Q}}(\mathbb{Q}^4, \mathbb{Q}^4)$ el operador definido por

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

Si $\mathcal{C}$ la base canónica de $\mathbb{Q}^4$ entonces:

$$\left[f\right]_{\mathcal{C}}^{\mathcal{C}} = \begin{bmatrix} 1 & -1 & 0 & 1 \\ -1 & 0 & -1 & 2 \\ 2 & -1 & -1 & -1 \\ 2 & -1 & 0 & 0 \end{bmatrix}$$

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

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

Si usamos la misma notación de la demostración anterior, tenemos $R_1(t) = P_2(t)P3(t) = (t-1)(t^2-2)$, $R_2(t) = P_1(t)P_3(t) = (t+1)(t^2-2)$, $R_3 = (t-1)(t+1)$, y como

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

si

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

entonces $p_i = \Pi_i(f)$, para $i = 1, 2, 3$, definen 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:

$$\left[p_1\right]_{\mathcal{C}}^{\mathcal{C}} = \begin{bmatrix} 1 & 0 & 0 & -1 \\ 2 & 0 & 0 & -2 \\ 1 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

$$\left[p_2\right]_{\mathcal{C}}^{\mathcal{C}} = \begin{bmatrix} -3 & 2 & -1 & 2 \\ -3 & 2 & -1 & 2 \\ 0 & 0 & 0 & 0 \\ -3 & 2 & -1 & 2 \end{bmatrix}$$

$$\left[p_3\right]_{\mathcal{C}}^{\mathcal{C}} = \begin{bmatrix} 3 & -2 & 1 & -1 \\ 1 & -1 & 1 & 0 \\ -1 & 0 & 1 & 1 \\ 3 & -2 & 1 & -1 \end{bmatrix}$$

así pues $V_1 = \mathrm{im}(p_1) = \langle(1,2,1,0)\rangle$, $V_2 = \mathrm{im}(p_2) = \langle(1,1,0,1)\rangle$ y $V_3 = \mathrm{im}(p_3) = \langle(1,1,1,1),(1,0,-1,1)\rangle$. Sea $\mathcal{B} = \{(1,2,1,0),(1,1,0,1),(1,1,1,1),(1,0,-1,1)\}$, de forma que la representación matricial de f en esta base

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

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

2.3. Operadores nilpotentes, espacios cíclicos y forma de Jordan

Sea $f \in \mathrm{Hom}_K(V,V)$ un operador.

Observación 2.26. Note que si V tiene dimensión finita y tomamos $f \in \mathrm{Hom}_K(V,V)$, $P_f(f) = 0$. Ahora suponga que $P_f(t)$ se descompone en factores lineales

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

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

$$V = V_1 \oplus V_2 \oplus \dots \oplus V_n$$

y si además denotamos $g_i \in \mathrm{Hom}_K(V_i,V_i)$ a la restricción de $f - \lambda_i\,\mathrm{id}_V$ a V_i, tenemos $g_i^{m_i} = 0$. Este tipo de operadores, cuya potencia se anula, motivan la siguiente definición.

Definición 2.27. Decimos que f es nilpotente si existe $r \in \mathbb{Z}_{>0}$ tal que $f^r = 0$, y al mínimo entre estos lo llamamos el grado de f.

Propiedad 2.28. *Suponga que f es nilpotente de grado r, y $V \neq \{0\}$, entonces tenemos una cadena de contenencias estrictas*

$$\{0\} < \ker(f) < \ker(f^2) < \dots < \ker(f^r) = V.$$

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

Dem. Note primero que para todo $i \in \mathbb{Z}_{>0}$, si $v \in V$ es tal que $f^i(v) = 0$, entonces $f^{i+1}(v) = 0$, luego $\ker(f^i) \leq \ker(f^{i+1})$.
Si $r = 1$, no hay nada que demostrar pues $f = 0$ y así la cadena corresponde a $\{0\} < V$. Ahora suponga que $r > 1$, luego $f^{r-1} \neq 0$ y así existe $v \in V$ tal que $f^{r-1}(v) \neq 0$. Note que para $i = 1, \ldots, r-1$

$$\begin{aligned} f^{i-1}\left(f^{r-i}(v)\right) &= f^{r-1}(v) \neq 0, \text{ y} \\ f^{i}\left(f^{r-i}(v)\right) &= f^{r}(v) = 0 \end{aligned}$$

así $f^{r-i}(v) \in \ker(f^i) \setminus \ker(f^{i-1})$ y tenemos una contenencia estricta $\ker(f^{i-1}) < \ker(f^i)$.
Suponga ahora que V tiene dimensión finita y denote, para $i = 1, \ldots, r$, $n_i = \dim(\ker(f^i))$. Entonces

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

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

Ejemplo 2.29. Sea $f \in \mathrm{Hom}_K(K^4, K^4)$ 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 si $n_i = \dim(\ker(f^i))$ entonces

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

La representación matricial de f en la base canónica $\mathcal{C}$ 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 el polinomio característico es $P_f(t) = t^4$.

Ejemplo 2.30. Sea $f \in \mathrm{Hom}_K(K^4, K^4)$ 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 si $n_i = \dim(\ker(f^i))$ entonces

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

La representación matricial de f en la base canónica $\mathcal{C}$ 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 el polinomio característico es $P_f(t) = t^4$.

Ejemplo 2.31. Sea $f \in \mathrm{Hom}_K(K^4, K^4)$ definido por

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

Así

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

y si $n_i = \dim(\ker(f^i))$ entonces

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

La representación matricial de f en la base canónica $\mathcal{C}$ 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 2.32. Sea $f \in \mathrm{Hom}_K(K^4, K^4)$ definido por

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

Así

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

y si $n_i = \dim(\ker(f^i))$ entonces

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

La representación matricial de f en la base canónica $\mathcal{C}$ es

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

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

Definición 2.33. Sea $v \in V$, si existe $k \in \mathbb{Z}_{>0}$ tal que $f^k(v) = 0$, al mínimo entre estos los llamamos el orden de v bajo f y lo denotamos por $\mathrm{ord}_f(v)$.

Propiedad 2.34. *Sea $v \in V$, $v \neq 0$, y suponga que $k = \mathrm{ord}_f(v)$, entonces $S = \{v, f(v), \ldots, f^{k-1}(v)\}$ es linealmente independiente.*

Dem. Suponga que $a_0, a_1, \ldots, a_{k-1} \in K$ son tales que

$$a_0 v + a_1 f(v) + \ldots + a_{k-1} f^{k-1}(v) = 0.$$

Aplicando f^{k-1} a esta igualdad obtenemos $a_0 f^{k-1}(v) = 0$, pero $f^{k-1}(v) \neq 0$ luego $a_0 = 0$. Inductivamente, si hemos establecido que $a_0 = a_1 = \ldots = a_{i-1} = 0$ para $0 < i < k-1$, aplicando f^{k-i-1} a la misma igualdad, obtenemos $a_i f^{k-1}(v) = 0$, luego $a_i = 0$. Así $a_0 = a_1 = \ldots = a_{k-1} = 0$. $\square$

Observación 2.35. Suponga que V tiene dimensión finita y f es nilpotente de grado $r = \dim(V)$. Si $v \in V$ es tal que $v \notin \ker(f^{r-1})$ entonces $\mathrm{ord}_f(v) = r$, luego si $v_i = f^{r-i}(v)$ para $i = 1, \ldots, r$, por la propiedad anterior

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

es una base de V; más aún

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

Definición 2.36. Decimos que V es cíclico bajo f si

$$S = \left\{f^i(v)\right\}_{i \in \mathbb{Z}_{>0}}$$

genera a V, es decir $\langle S \rangle = V$, para algún $v \in V$. En tal caso decimos que v es un *vector cíclico* relativo a f.

Observación 2.37. Si V tiene dimensión finita y $f \neq 0$ es nilpotente de grado $r = \dim(V)$, la observación anterior explica que V es cíclico bajo f.

Definición 2.38. Suponga que V tiene dimensión finita y $f \neq 0$ es nilpotente de grado $r = \dim(V)$, una base de la forma

$$\mathcal{B} = \{v_1, \ldots, v_r\}, \qquad v_i = f^{r-i}(v_r)$$

se llama una *base de Jordan de V relativa a f*.

Observación 2.39. En caso de que f sea nilpotente de grado inferior, V no es cíclico, pero se puede descomponer en subespacios invariantes bajo f y cíclicos bajo la restricción de f a ellos. Esto es el contenido del siguiente teorema.

Teorema 2.40. *Suponga que V tiene dimensión finita y que $f \neq 0$ es nilpotente de grado r. Sea $n = \dim(\ker(f))$. Entonces existen n subespacios invariantes bajo f, $V_1, \ldots, V_n$ tales que*

$$V = V_1 \oplus \ldots \oplus V_n$$

y si $f_i \in \operatorname{Hom}_K(V_i, V_i)$ es la restricción de f a V_i, para $i = 1, \ldots, n$, entonces V_i cíclico bajo f_i.

Dem. Denotemos $K_i = \ker(f^i)$, de forma que $K_0 = \{0\}$ y $K_r = V$. Note que para $i = 1, \ldots, r-1$, $K_i < K_{i+1}$. Podemos así descomponer para cada $i = 2, \ldots r$

$$K_i = K_{i-1} \oplus K_i'.$$

De forma que si $v \in K_i'$, $v \neq 0$, entonces $\operatorname{ord}_f(v) = i$. Por lo tanto

$$f(K_i') \leq K_{i-1}'.$$

Tenemos entonces

$$\begin{aligned} V &= K_r \\ &= K_{r-1} \oplus K_r' \\ &\vdots \\ &= K_1 \oplus K_2' \oplus \ldots \oplus K_r' \end{aligned}$$

y

$$K_r' \xrightarrow{f} K_{r-1}' \xrightarrow{f} \ldots \xrightarrow{f} K_2' \xrightarrow{f} K_1 \xrightarrow{f} \{0\}$$

Se trata entonces de escoger una base de V que sea compatible con esta descomposición y esta cadena de imágenes bajo f. Denote $n_i = \dim(K_i)$ y $n_i' = \dim(K_i')$, para $i = 2, \ldots, r$, y $n_1 = n = \dim(K_1)$ de forma que

$$n_i = n_i' + n_{i-1}$$

y

$$\begin{aligned} \dim(V) &= n_r \\ &= n_r' + n_{r-1} \\ &\vdots \\ &= n_r' + n_{r-1}' + \ldots + n_2 \\ &= n_r' + n_{r-1}' + \ldots + n_2' + n_1 \\ &= n_r' + n_{r-1}' + \ldots + n_2' + n \end{aligned}$$

Sea $\mathcal{B}_r = \{v_{r,1}, \ldots, v_{r,n_r'}\} \subseteq V$ una base de K_r'. Para $i = 1, \ldots n_r'$, sea

$$v_{r-1,i} = f(v_{r,i}).$$

Note que $f(\mathcal{B}_r) = \{v_{r-1,1}, \ldots, v_{r-1,n'_r}\} \subseteq K'_{r-1}$ es linealmente independiente. De hecho si

$$a_1 v_{r-1,1} + \ldots + a_{n'_r} v_{r-1,n'_r} = 0,$$

entonces

$$a_1 f(v_{r,1}) + \ldots + a_{n'_r} f(v_{r,n'_r}) = 0$$

luego $a_1 v_{r,1} + \ldots + a_{n'_r} v_{r,n'_r} \in K_1 \cap K'_r$; por lo tanto $a_1 v_{r,1} + \ldots + a_{n'_r} v_{r,n'_r} = 0$ y $a_1 = \ldots = a_{n'_r} = 0$.
Sea $\mathcal{B}_{r-1} = \{v_{r-1,1}, \ldots, v_{r-1,n'_{r-1}}\}$ un base de K'_{r-1} que contiene a $f(\mathcal{B}_r)$. Para $i = 1, \ldots n'_{r-1}$, sea

$$v_{r-2,i} = f(v_{r-1,i}).$$

Similarmente, note que $f(\mathcal{B}_{r-1}) = \{v_{r-2,1}, \ldots, v_{r-2,n'_{r-1}}\} \subseteq K'_{r-2}$ es linealmente independiente.
Iterativamente obtenemos bases $\mathcal{B}_1, \mathcal{B}_2, \ldots, \mathcal{B}_r$ respectivamente de $K_1, K'_2, \ldots, K'_r$ con $f(\mathcal{B}_{i+1}) \subseteq \mathcal{B}_i$ para $i = 1, \ldots, r-1$. En particular

$$\mathcal{B} = \mathcal{B}_1 \cup \mathcal{B}_2 \cup \ldots \cup \mathcal{B}_r$$

es una base de V. Defina (ver Figura 2.1)

$$\begin{aligned} V_1 &= \langle v_{j,1} \in \mathcal{B} \mid 1 \leq \dim(K'_j) \rangle \\ V_2 &= \langle v_{j,2} \in \mathcal{B} \mid 2 \leq \dim(K'_j) \rangle \\ &\vdots \\ V_n &= \langle v_{j,n} \in \mathcal{B} \mid n \leq \dim(K'_j) \rangle \end{aligned}$$

de esta forma por construcción cada V_i, $i = 1, \ldots, n$, son invariantes bajo f y cíclicos bajo f_i. $\square$

Ejemplo 2.41. Si $f \in \mathrm{Hom}_K(K^4, K^4)$ está definido como en Ejemplo 2.29

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

entonces

$$n = 1, \quad n'_2 = 1, \quad n'_3 = 1, \quad n'_4 = 1.$$

Ejemplo 2.42. Si $f \in \mathrm{Hom}_K(K^4, K^4)$ está definido como en Ejemplo 2.30

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

entonces

$$n = 2, \quad n'_2 = 1, \quad n'_3 = 1.$$

Ejemplo 2.43. Si $f \in \mathrm{Hom}_K(K^4, K^4)$ está definido como en Ejemplo 2.31

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

entonces

$$n = 2, \quad n'_2 = 2.$$

Ejemplo 2.44. Si $f \in \mathrm{Hom}_K(K^4, K^4)$ está definido como en Ejemplo 2.32

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

entonces

$$n = 3, \quad n_2' = 1.$$

Observación 2.45. Bajo la hipótesis del teorema, y usando la notación en él, obtenemos que para cada V_i, $i = 1, \ldots, n$, tenemos una base de Jordan $\mathcal{B}_i$ relativa a f_i. De esta forma la unión de ella forma una base $\mathcal{B}$ de V. La representación matricial de f en la base T es una matriz diagonal por bloques:

$$\big[f\big]_{\mathcal{B}}^{\mathcal{B}} = \left[\begin{array}{c|c|c|c} J_1 & 0 & \cdots & 0 \\ \hline 0 & J_2 & \cdots & 0 \\ \hline \vdots & \vdots & \ddots & \vdots \\ \hline 0 & 0 & \cdots & J_n \end{array}\right]$$

donde cada $J_i = \big[f\big]_{\mathcal{B}_i}^{\mathcal{B}_i}$ es una matriz $\dim(V_i) \times \dim(V_i)$ de la forma en Observación 2.35.

Observación 2.46. Como corolario de la prueba del teorema tenemos que cuando V tiene dimensión finita y f es nilpotente, la información suministrada por las cantidades

$$\begin{array}{rcl} \dim(K_1) & = & n \\ \dim(K_i) - \dim(K_{i-1}) & = & n_i', \qquad i = 2, \ldots, r \end{array}$$

$$\begin{array}{cccccccccccc}
K_r' & v_{r,1} & v_{r,2} & \cdots & v_{r,n_r'} & & & & & & \\
\downarrow & \downarrow & \downarrow & & \downarrow & & & & & & \\
K_{r-1}' & v_{r-1,1} & v_{r-1,2} & \cdots & v_{r-1,n_r'} & \cdots & v_{r-1,n_{r-1}'} & & & & \\
\downarrow & \downarrow & \downarrow & & \downarrow & & \downarrow & & & & \\
\vdots & \vdots & \vdots & \cdots & \vdots & \cdots & \vdots & & & & \\
\downarrow & \downarrow & \downarrow & & \downarrow & & \downarrow & & & & \\
K_2' & v_{2,1} & v_{2,2} & \cdots & v_{2,n_r'} & \cdots & v_{2,n_{r-1}'} & \cdots & v_{2,n_2'} & & \\
\downarrow & \downarrow & \downarrow & & \downarrow & & \downarrow & & \downarrow & & \\
K_1 & v_{1,1} & v_{1,2} & \cdots & v_{1,n_r'} & \cdots & v_{1,n_{r-1}'} & \cdots & v_{1,n_2'} & \cdots & v_{1,n} \\
\downarrow & \downarrow & \downarrow & & \downarrow & & \downarrow & & \downarrow & & \downarrow \\
\{0\} & 0 & 0 & \cdots & 0 & \cdots & 0 & \cdots & 0 & \cdots & 0 \\
 & & & & & & & & & & \\
 & V_1 & V_2 & \cdots & V_{n_r'} & \cdots & V_{n_{r-1}'} & \cdots & V_{n_2'} & \cdots & V_n
\end{array}$$

Figura 2.1: Edificios colapsando

son tales que $n \geq n'_2 \geq \ldots \geq n'_r$ y determinan univocamente 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, n indica el número de bloques de Jordan y n'_i el número de bloques de Jordan de tamaño mayor o igual a i.

Definición 2.47. Se le llama *matriz en bloque de Jordan* a una matriz cuadrada $n \times n$ de la forma

$$J_{\lambda,n} = \begin{bmatrix} \lambda & 1 & 0 & \cdots & 0 \\ 0 & \lambda & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ 0 & 0 & 0 & \cdots & \lambda \end{bmatrix}.$$

Lema 2.48. *Suponga que V tiene dimensión finita y que $P(t) = (t-\lambda)^m \in K[t]$, es tal que $P(f) = 0$. Entonces existe una base $\mathcal{B}$ de V tal que la representación matricial de f en esta base es es una matriz diagonal por bloques:*

$$\Big[f\Big]_{\mathcal{B}}^{\mathcal{B}} = \left[\begin{array}{c|c|c|c} J_1 & 0 & \cdots & 0 \\ \hline 0 & J_2 & \cdots & 0 \\ \hline \vdots & \vdots & \ddots & \vdots \\ \hline 0 & 0 & \cdots & J_n \end{array}\right]$$

donde cada J_i, $i = 1, \ldots, n$ es una matriz en bloque de Jordan.

Dem. Tenemos $P(f) = (f - \lambda\,\mathrm{id}_V)^m = 0$. Luego el operador $g = f - \lambda\,\mathrm{id}_V$ es nilpotente. Por Teorema 2.40,

$$V = V_1 \oplus V_2 \oplus \ldots \oplus V_n$$

donde para cada V_i, $i = 1, \ldots, n$, hay una base de la forma $\mathcal{B}_i = \{v_{1,i}, \ldots, v_{m_i,i}\}$, con $\dim(V_i) = m_i$ y

$$\begin{array}{rcccl} v_{m_i-1,i} & = & g(v_{m_i,i}) & = & f(v_{m_i,i}) - \lambda v_{m_i,i} \\ v_{m_i-2,i} & = & g(v_{m_i-1,i}) & = & f(v_{m_i-1,i}) - \lambda v_{m_i-1,i} \\ & \vdots & & \vdots & \\ v_{1,i} & = & g(v_{2,i}) & = & f(v_{2,i}) - \lambda v_{2,i} \\ 0 & = & g(v_{1,i}) & = & f(v_{1,i}) - \lambda v_{1,i}. \end{array}$$

Así,

$$\begin{array}{rcl} f(v_{m_i,i}) & = & v_{m_i-1,i} + \lambda v_{m_i,i} \\ f(v_{m_i-1,i}) & = & v_{m_i-2,i} + \lambda v_{m_i-1,i} \\ & \vdots & \\ f(v_{2,i}) & = & v_{1,i} + \lambda v_{2,i} \\ f(v_{1,i}) & = & \lambda v_{1,i}. \end{array}$$

En particular, cada V_i es invariante bajo f, luego, si $f_i \in \mathrm{Hom}_K(V_i, V_i)$ denota la restricción de f a V_i, $\left[f_i\right]_{\mathcal{B}_i}^{\mathcal{B}_i} = J_i$ es una matriz en bloque de Jordan. De esto, si $\mathcal{B} = \mathcal{B}_1 \cup \ldots \cup \mathcal{B}_n$, la representación matricial $\left[f\right]_{\mathcal{B}}^{\mathcal{B}}$ tiene la forma buscada. □

Teorema 2.49 (Teorema de Jordan)**.** *Suponga que V tiene dimensión finita y que*

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

Entonces existe una base $\mathcal{B}$ de V tal que la representación matricial de f en esta base es es una matriz diagonal por bloques de Jordan.

Dem. Sin perdida 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 Caley-Hamilton $P_f(f) = 0$, luego por Propiedad 2.24,

$$V = V_1 \oplus \ldots \oplus V_r$$

donde cada $V_i = \ker((f - \lambda_i \, \mathrm{id}_V)^{m_i})$, $i = 1, \ldots, r$, es invariante bajo f. En particular, si $f_i \in \mathrm{Hom}_K(V_i, V_i)$ es la restricción de f a V_i, $i = 1, \ldots, r$, $P_i(f_i) = 0$, donde $P_i(t) = (t - \lambda_i)^{m_i}$. Por lo tanto, el lema implica que existe una base $\mathcal{B}_i$ de V_i para la cual $\left[f_i\right]_{\mathcal{B}_i}^{\mathcal{B}_i}$ es una matriz diagonal por bloques de Jordan. Finalmente si $\mathcal{B} = \mathcal{B}_1 \cup \ldots \cup \mathcal{B}_n$, la representación matricial $\left[f\right]_{\mathcal{B}}^{\mathcal{B}}$ tiene la forma afirmada. □

Definición 2.50. Generalizamos la definición anterior 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 representación matricial de este operador en aquella base es diagonal en bloques de Jordan.

Lema 2.51. *Sea $f \in \mathrm{Hom}_K(V, V)$. Si $\lambda_1, \ldots, \lambda_n \in K$ son valores propios, todos distintos, de f, y, para $i = 1, \ldots, n$, $v_i \in V$ es un vector propio de λ_i, entonces $\{v_1, \ldots, v_n\}$ es linealmente independiente.*

Dem. Por inducción en n, siendo el caso base $n = 1$ inmediato, pues $\{v_1\}$ es linealmente independiente si $v_1 \neq 0$, la cual se cumple pues v_1 es vector propio. Para el paso inductivo, si $a_1, \ldots, a_n$ son tales que $a_1 v_1 + \ldots + a_n v_n = 0$,por contradicción podemos asumir que cada $a_i \neq 0$, o de lo contrario, por hipotesis de inducción, si algún a_i es 0 el resto también lo son. Entonces

$$0 = (f - \lambda_n \, \mathrm{id}_V)(a_1 v_1 + \ldots + a_n v_n) = a_1(\lambda_1 - \lambda_n)v_1 + \ldots + a_{n-1}(\lambda_{n-1} - \lambda_n)v_{n-1};$$

y así, por hipótesis de inducción, para $i = 1, \ldots, n-1$, $a_i(\lambda_i - \lambda_n) = 0$. Pero $a_i \neq 0$ y $\lambda_i - \lambda_n \neq 0$, si $i \in \{1, \ldots, n-1\}$, lo cual es una contradicción. □

Lema 2.52. *Suponga que $f \in \mathrm{Hom}_K(V,V)$ es diagonalizable, entonces:*

1. *Si $V_0 \neq \{0\}$ es invariante bajo f, su restricción a V_0, $f_0 \in \mathrm{Hom}_K(V_0, V_0)$, también es diagonalizable.*

2. *Si $g \in \mathrm{Hom}_K(V,V)$ es diagonalizable y $f \circ g = g \circ f$, entonces existe una familia $\{V_i\}_{i\in I}$ de espacios propios simultáneamente de f y g tal que $V = \bigoplus_{i\in I} V_i$. En particular si v es un vector propio simultáneamente de f y g, v es un vector propio de $af + bg$ para todo $a, b \in K$.*

Dem.

1. Dado un valor propio $\lambda \in K$ de f, definimos $E_\lambda \leq V$ como el subespacio generado por los vectores propios de f asociados a λ, es decir $E_\lambda = \ker(f - \lambda\,\mathrm{id}_V)$, y $F_\lambda = V_0 \cap E_\lambda$. Note que, como f es diagonalizable, por el lema anterior,
$$V = \bigoplus_{i\in I} E_{\lambda_i}$$
donde $\{\lambda_i\}_{i\in I}$ es la colección de valores propios de f. Sea $v \in V_0$, $v \neq 0$, y
$$v = v_1 + \ldots + v_n$$
una descomposición en vectores propios asociados respectivamente a valores propios $\lambda_1, \ldots, \lambda_n \in K$. Por inducción en n veamos que $v_1, \ldots, v_n \in V_0$ siendo el caso base $n = 1$ inmediato pues en tal caso $v_0 = v_1$. Para el paso inductivo, como V_0 es invariante bajo f
$$(f - \lambda_n\,\mathrm{id}_V)(v) = (\lambda_1 - \lambda_n)v_1 + \ldots + (\lambda_{n-1} - \lambda_n)v_{n-1}$$
también pertenece a V_0. Luego por hipótesis inductiva , $(\lambda_1 - \lambda_n)v_1, \ldots, (\lambda_{n-1} - \lambda_n)v_{n-1} \in V_0$, así pues $v_1, \ldots, v_{n-1} \in V_0$ y $v_n = v - v_1 - \ldots - v_{n-1} \in V_0$. De donde
$$V_0 = \bigoplus_{i\in J} F_{\lambda_i},$$
donde J es la colección de $i \in I$ tales que $F_{\lambda_i} \neq \{0\}$. Entonces f_0 es diagonalizable tomando bases de cada F_{λ_i}, $i \in J$.

2. Usando la notación de la demostración de la primera afirmación del lema, si $v \in E_{\lambda_i}$, $i \in I$,
$$f\left(g(v)\right) = g\left(f(v)\right) = \lambda_i g(v),$$
luego E_{λ_i} es invariante bajo g. Por la primera parte del lema, la restricción de g a E_{λ_i}, $g_i \in \mathrm{Hom}_K(E_{\lambda_i}, E_{\lambda_i})$ es diagonalizable. Luego g es diagonalizable tomando bases de cada E_{λ_i}, $i \in I$. Los espacios propios generados por cada uno de estos elementos de estas bases forman una colección de espacios propios simultáneos cuya suma es una suma directa igual a V. $\square$

Teorema 2.53 (Descomposición de Jordan-Chevalley)**.** *Suponga que V tiene dimensión finita y que*

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

donde $\lambda_i \neq \lambda_j$ si $i \neq j$. Entonces existen operadores $f_N, f_D \in \mathrm{Hom}_K(V,V)$, tales que

1. *f_D es diagonalizable y f_N es nilpotente;*
2. *$f_D + f_N = f$; y,*
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 $P_D(t), P_N(t) \in K[t]$ tales que $f_N = P_N(f)$ y $f_D = P_D(f)$.

Dem. Defina $P_i(t) = (t-\lambda_i)^{m_i}$, $i = 1,\ldots,n$. Por Propiedad 2.24, existen $\Pi_1(t),\ldots,\Pi_n(t) \in K[t]$ tales que $\Pi_i(f) = p_i$, $i = 1,\ldots,n$, son las proyecciones sobre $V_i = \ker\left((f-\lambda_i\,\mathrm{id}_V)^{m_i}\right)$ respecto a la descomposición

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

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

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

y así f_D es diagonalizable por Teorema 2.9. Defina $P_N(t) = t - P_D(t)$ y $f_N = P_N(f) = f - f_D$. De esta forma, $f_D + f_N = f$, y si $v_i \in V_i$

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

luego la restricción de f_N a V_i es nilpotente de grado $\leq m_i$. De donde f_N es nilpotente de grado $\leq \max\{m_1,\ldots,m_n\}$. Finalmente,

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

Si $f'_D, f'_N \in \mathrm{Hom}_K(V,V)$ conmutan y son respectivamente diagonalizable y nilpotente tales que $f = f'_D + f'_N$, entonces

$$f \circ f'_D = (f'_D + f'_N)f'_D = f'_D \circ f'_D + f'_N \circ f'_D = f'_D \circ f'_D + f'_D \circ f'_N = f'_D \circ f,$$

es decir f y f'_D conmutan. Por lo cual, $P_D(f) = f_D$ y f'_D también lo hacen. Entonces f_D y f'_D son diagonalizables y conmutan. Ahora, si v es un vector propio común, entonces v es un vector propio de $f_D - f'_D$. Pero $f_D - f'_D = f'_N - f_N$, y, como f'_N y f_N igualmente conmutan, $f'_N - f_N$ es igualmente nilpotente. Así $f_D - f'_D$ es diagonalizable y, a su vez, nilpotente, el valor propio asociado a v es 0. Por el lema anterior existe una base de V de vectores propios simultáneos de f_D y f'_D, luego todos los valores propios de $f_D - f'_D$ son 0. Es decir $f_D - f'_D = 0 = f'_N - f_N$; y, $f'_D = f_D$ y $f'_N = f_N$. □

2.4. Polinomio minimal y transformaciones semi-simples

Capítulo 3

Espacio dual

Sea K un cuerpo y V, W espacios vectoriales sobre K.

Notación 3.1. Dada una colección de indices 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 $V^* = \operatorname{Hom}_K(V, K)$, 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 $\{v_1, \ldots, v_n\}$ una base de V, donde $n = \dim(V)$. Por Proposición 1.35.2, $\lambda \in V^*$ está unívocamente por los valores $\lambda(v_1), \ldots, \lambda(v_n)$. Defina $\lambda_1, \ldots, \lambda_n \in V*$ por

$$\lambda_i(v_j) = \delta_{ij}.$$

Veamos que $\{\lambda_1, \ldots, \lambda_n\}$ 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 $a_1, \ldots, a_n \in K$ tales que

$$\sum_{i=1}^{n} a_i \lambda_i = 0.$$

De forma que, para $j = 1, \ldots, n$,

$$0 = \left(\sum_{i=1}^{n} a_i \lambda_i \right)(v_j) = \sum_{i=1}^{n} a_i \lambda_i(v_j) = \sum_{i=1}^{j} a_i \delta_{ij} = a_j.$$

Para ver que $V^* = \langle \lambda_1, \ldots, \lambda_n \rangle$, 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 $j = 1, \ldots, n$,

$$\mu(v_j) = \left(\sum_{i=1}^{n} a_i \lambda_i \right)(v_j) = \sum_{i=1}^{j} a_i \delta_{ij} = a_j = \lambda(v_j),$$

luego $\mu = \lambda$. □

Definición 3.4. Suponga que V tiene dimensión finita y sea $\mathcal{B} = \{v_1, \ldots, v_n\}$ una base de V, donde $n = \dim(V)$. A la base $\mathcal{B}^* = \{\lambda_1, \ldots, \lambda_n\}$ de V^* donde

$$\lambda_i(v_j) = \delta_{ij}.$$

la llamamos *base dual* de $\mathcal{B}$.

Observación 3.5. Si V tiene dimensión finita, $\mathcal{B} = \{v_1, \ldots, v_n\}$ es una base de V y $\mathcal{B}^* = \{\lambda_1, \ldots, \lambda_n\}$ es la base dual, entonces para todo $v \in V$

$$v = \sum_{i=1}^{n} \lambda_i(v) v_i.$$

De hecho si $a_1, \ldots, a_n \in K$ son tales que $a_1 v_1 + \ldots + a_n v_n = v$,

$$\lambda_i(v) = \lambda_i \left(\sum_{j=1}^{n} a_j v_j \right) = \sum_{j=1}^{n} a_j \delta_{ij} = a_i.$$

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 indices. 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*

$$S_1 \subseteq S_2 \qquad y \qquad L_1 \subseteq L_2,$$

entonces

$$S_2^0 \leq S_1^0 \qquad y \qquad (L_2)_0 \leq (L_1)_0\,;$$

3. *$\langle S^0\rangle_0 = \mathrm{Sp}(S)$;*

4. *si $V_1, V_2 \leq V$ y $V_1^*, V_2^* \leq V^*$, entonces*

$$(V_1 + V_2)^0 = V_1^0 \cap V_2^0, \ y \ (V_1^* + V_2^*)_0 = (V_1^*)_0 \cap (V_2^*)_0\,;$$

5. *si V tiene dimensión finita,*

$$\dim(\langle S\rangle) + \dim(S^0) = \dim(V), \ y \ \dim(L_0) + \dim(\langle L\rangle) = \dim(V^*).$$

Dem.

1. Si $\lambda_1, \lambda_2 \in S^0$ y $c \in K$, entonces para todo $v \in S$

$$(\lambda_1 + \lambda_2)(v) = \lambda_1(v) + \lambda_2(v) = 0, \text{ y } (c\lambda_1)(v) = c\lambda_1(v) = 0,$$

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 $v_1, \ldots, v_m \in S$ y $a_1, \ldots, a_m \in K$ tales que

$$v = a_1 v_1 + \ldots + a_m v_m$$

así, si $\lambda \in S^0$, $\lambda(v) = a_1\lambda(v_1) + \ldots + a_m\lambda(v_m) = 0$. Luego $\langle S\rangle \leq \left(S^0\right)_0$. 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 indices definida por $j \in J$ si $v_j \in S'$. Entonces $L = \{\lambda_i\}_{i \in I \setminus J} \subseteq S^0$ y $\left(S^0\right)_0 \leq L_0$. 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 $\left(S^0\right)_0 \leq \langle S' \rangle = \langle S \rangle$.

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 $\{v_1, \ldots, v_k\} = S' \subseteq S$ linealmente independiente tal que $\langle S' \rangle = \langle S \rangle$, el cual extendemos a una base $\mathcal{B} = \{v_1 \ldots, v_n\}$ de V. Sea $\mathcal{B}^* = \{\lambda_1 \ldots, \lambda_n\}$ la base dual a $\mathcal{B}$. Defina $f \in \operatorname{Hom}_K(V, V)$ por

$$f(v) = \sum_{i=1}^{k} \lambda_i(v) v_i.$$

   Por construcción, $\operatorname{im}(f) = \langle S' \rangle = \langle S \rangle$ y $\ker(f) = \langle \lambda_{k+1}, \ldots, \lambda_n \rangle_0$. Pero $\langle \lambda_{k+1} \ldots \lambda_n \rangle = S^0$, luego $\dim(\langle S \rangle) + \dim(S^0) = n$.
   Similarmente, tome un subconjunto $\{\lambda_1, \ldots, \lambda_k\} = L' \subseteq L$ linealmente independiente tal que $\langle L' \rangle = \langle L \rangle$, el cual extendemos a una base $\mathcal{B}^* = \{\lambda_1, \ldots, \lambda_n\}$ de V. Sea $\{v_1, \ldots, v_n\}$ una base de V. Defina $f \in \operatorname{Hom}_K(V, V)$ por

$$f(v) = \sum_{i=1}^{k} \lambda_i(v) v_i.$$

   Por construcción, $\ker(f) = L'_0 = L_0$ y $\operatorname{im}(f) = \langle v_1, \ldots, v_k \rangle$. Luego $\dim \operatorname{im}(f) = \dim(\langle L \rangle)$ y $\dim(L_0) + \dim(\langle L \rangle) = n$. □

Proposición 3.9. *Sea* $V_1 \leq V$ *entonces* $(V/V_1)^* = V_1^0$.

Dem. Tome $\pi_{V_1} : V \to V/V_1$ con $\pi_{V_1}(v) = v + V_1$; y, defina la transformación lineal $f : (V/V_1)^* \to V^*$ 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 $\lambda : V/V_1 \to K$ 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*

$$\begin{array}{rcl} \widehat{\bullet} : V & \longrightarrow & (V^*)^* \\ v & \longmapsto & \widehat{v} : \lambda \mapsto \lambda(v). \end{array}$$

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, $\dim(V) = \dim(V^*) = \dim\left((V^*)^*\right)$, entonces $\widehat{\bullet}$ es un isomorfismo. $\square$

3.2. Transformación dual

Definición 3.11. Sea $f \in \mathrm{Hom}_K(V, W)$. Definimos la *transformación dual*, $f^* \in \mathrm{Hom}_K(W^*, V^*)$, por

$$f^*(\lambda) = \lambda \circ f$$

para todo $\lambda \in W^*$.

$$\begin{tikzcd} V \arrow[r, "f"] \arrow[dr, dashed, "f^*(\lambda)"'] & W \arrow[d, "\lambda"] \\ & K \end{tikzcd}$$

Figura 3.1: Transformación dual

Observación 3.12. La linearidad del mapa f^* se sigue de las siguientes igualdades, validas para todo $f \in \mathrm{Hom}_K(V, W)$, $\lambda_1, \lambda_2 \in W^*$, $c \in K$:

$$\begin{array}{rcl} (\lambda_1 + \lambda_2) \circ f & = & \lambda_1 \circ f + \lambda_2 \circ f \\ (c\lambda_1) \circ f & = & c(\lambda_1 \circ f) \end{array}$$

Propiedad 3.13. *Sean U un espacio vectorial sobre K y $f \in \mathrm{Hom}_K(V, W)$ y $g \in \mathrm{Hom}_K(W, U)$, entonces*

$$(g \circ f)^* = f^* \circ g^*$$

Dem. Para $\lambda \in U^*$, tenemos

$$(g \circ f)^* \lambda = \lambda \circ (g \circ f) = (\lambda \circ g) \circ f = g^*(\lambda) \circ f = f^* \circ g^*(\lambda)$$

$\square$

Propiedad 3.14. *Sea $f \in \mathrm{Hom}_K(V, W)$, 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
$$\lambda(w) = \lambda\left(f(v)\right) = f^*(\lambda)(v) = 0$$
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 $\lambda : W \to K$ 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 contrucción $\mu = f^*\lambda$ pues
$$f^*(\lambda)(v_1) = \lambda(f(v_1)) = \lambda(w_1) = \mu(v_1),$$
luego f^* es sobreyectiva. $\square$

Propiedad 3.15. *Sean $V_1, V_2 \leq V$ tales que $V = V_1 \oplus V_2$; y, $\pi_1 : V \to V_1$ y $\pi_2 : V \to V_2$ 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

$$\lambda_1(v_1) = \lambda(v_1) \qquad \lambda_2(v_2) = \lambda(v_2).$$

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

$$\begin{aligned} \lambda(v) &= \lambda(v_1) + \lambda(v_2) \\ &= \lambda_1\left(\pi_1(v)\right) + \lambda_2\left(\pi_2(v)\right) \\ &= \left(\pi_1^*(\lambda_1) + \pi_2^*(\lambda_2)\right)(v) \end{aligned}$$

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 $v_2 = \pi_2(v)$

$$\lambda(v) = \lambda(v_1) + \lambda(v_2) = \lambda_2\left(\pi_2(v_1)\right) + \lambda_1\left(\pi_1(v_2)\right) = \lambda_2(0) + \lambda_1(0) = 0.$$

Luego $\pi_1^*(V_1) \cap \pi_2^*(V_2) = \{0\}$. $\square$

Teorema 3.16. *El mapa*

$$\begin{array}{rcl} \bullet^* : \mathrm{Hom}_K(V,W) & \longrightarrow & \mathrm{Hom}_K(W^*,V^*) \\ f & \longmapsto & f^* \end{array}$$

es una transformación lineal inyectiva. Si W tiene dimensión finita entonces es un isomorfismo.

Dem. Sean $f, g \in \mathrm{Hom}_K(V,W)$, y $c \in K$. Dado $\lambda \in W^*$ tenemos:

$$\begin{array}{rcl} (f+g)^*(\lambda) & = & \lambda \circ (f+g) \\ & = & \lambda \circ f + \lambda \circ g \\ & = & f^*(\lambda) + g^*(\lambda) \\ & = & (f^* + g^*)(\lambda) \\ (cf)^*(\lambda) & = & \lambda \circ (cf) \\ & = & c(\lambda \circ f) \\ & = & cf^*(\lambda). \end{array}$$

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 $v \in V$

$$f(v) = \sum_{i\in I} \lambda_i \left(f(v)\right) w_i = \sum_{i\in I} f^*(\lambda_i)(v) w_i = \sum_{i\in I} 0 w_i = 0.$$

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 $\phi \in \mathrm{Hom}^*(W^*, V^*)$ está unívocamente determinado por la imágen $\{\phi(\lambda_i)\}_{i\in I} \subseteq V^*$. Defina $f \in \mathrm{Hom}_K(V,W)$ por la suma finita

$$f(v) = \sum_{j\in I} \left[\phi(\lambda_j)(v)\right] w_j,$$

de forma tal que para todo $i \in I$, $v \in V$,

$$\begin{array}{rcl} f^*(\lambda_i)(v) & = & \lambda_i(f(v)) \\ & = & \lambda_i \left(\displaystyle\sum_{j\in I} \left[\phi(\lambda_j)(v)\right] w_j \right) \\ & = & \displaystyle\sum_{j\in I} \left[\phi(\lambda_j)(v)\right] \lambda_i(w_j) \\ & = & \displaystyle\sum_{i\in I} \left[\phi(\lambda_j)(v)\right] \delta_{ij} = \phi(\lambda_i)(v). \end{array}$$

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. □

Observación 3.17. Suponga que W tiene dimensión finita, sea $\mathcal{B}_W = \{w_1, \ldots, w_m\}$ una base de W y $\lambda \in W^* = \operatorname{Hom}_K(W, K)$. Si tomamos la base $\{1\}$ de K, entonces la representación matricial de λ respecto a las base $\mathcal{B}_W$ y $\{1\}$ es

$$\Big[\lambda\Big]_{\mathcal{B}_W}^{\{1\}} = \Big[\lambda(w_1) \cdots \lambda(w_m)\Big].$$

Ahora, si $\mathcal{B}_V$ es una base de V y $f \in \operatorname{Hom}_K(V, W)$, tenemos

$$\Big[f^*(\lambda)\Big]_{\mathcal{B}_V}^{\{1\}} = \Big[\lambda \circ f\Big]_{\mathcal{B}_V}^{\{1\}} = \Big[\lambda\Big]_{S}^{\{1\}} \Big[f\Big]_{\mathcal{B}_W}^{\mathcal{B}_V}.$$

Por otro lado, podemos tomar las coordenadas de λ en la base $\mathcal{B}_W^* = \{\lambda_1, \ldots, \lambda_m\}$ de W^* dual de $\mathcal{B}_W$:

$$\Big[\lambda\Big]^{\mathcal{B}_W^*} = \begin{bmatrix} \lambda(w_1) \\ \vdots \\ \lambda(w_m) \end{bmatrix}$$

Si ademas asumimos que V tiene también dimensión finita y tomamos la base $\mathcal{B}_V^*$ de V^* dual de T, entonces

$$\Big[f^*(\lambda)\Big]^{\mathcal{B}_V^*} = \Big[f^*\Big]_{\mathcal{B}_W^*}^{\mathcal{B}_V^*} \Big[\lambda\Big]^{\mathcal{B}_W^*}.$$

La pregunta inmediata es: ¿Cuál es la relación entre $\Big[f\Big]_{\mathcal{B}_V}^{\mathcal{B}_W}$ y $\Big[f^*\Big]_{\mathcal{B}_W^*}^{\mathcal{B}_V^*}$?

Definición 3.18. Sean I, J conjuntos y $A \in M_{I \times J}(K)$, definimos la *matriz traspuesta* de A por $A^\intercal \in M_{J \times I}(K)$ tal que

$$A^\intercal(j, i) = A(i, j)$$

para todo $(j, i) \in J \times I$. Es decir el valor en (j, i) de $A^\intercal$ 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 $A^\intercal \in M_{n \times m}(K)$ tal que

$$A^\intercal(j, i) = A(i, j)$$

Sea $A \in M_{I \times I}(K)$, o $A \in M_{n \times n}(K)$, decimos que A es *simétrica* si $A^\intercal = A$.

Teorema 3.19. *Suponga que V y W tienen dimensión finita, $n = \dim(V) > 0$ y $m = \dim(W) > 0$, y $f \in \operatorname{Hom}_K(V, W)$. Sean $\mathcal{B}_W = \{w_1, \ldots, w_m\}$ y $\mathcal{B}_V = \{v_1, \ldots, v_n\}$ respectivamente bases de W y V; y, sean $\mathcal{B}_W^* = \{\lambda_1, \ldots, \lambda_m\}$ y $\mathcal{B}_V^* = \{\mu_1, \ldots, \mu_n\}$ 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 $A^\intercal \in M_{n \times m}(K)$ es la representación matricial de f^* respecto a las bases $\mathcal{B}_V^*$ y $\mathcal{B}_W^*$. Es decir,*

$$\Big[f^*\Big]_{\mathcal{B}_W^*}^{\mathcal{B}_V^*} = \left(\Big[f\Big]_{\mathcal{B}_V}^{\mathcal{B}_W}\right)^\intercal$$

Dem. Si $a_{ij} \in K$ es la ij-ésima entrada de $A = \left[f\right]_{\mathcal{B}_V}^{\mathcal{B}_W}$, entonces

$$a_{ij} = \lambda_i \left(f(v_j)\right) = f^*(\lambda_i)(v_j);$$

y,

$$f^*(\lambda_i) = \sum_{l=1}^{n} \left[f^*(\lambda_i)(v_l)\right] \mu_l,$$

luego la ji-esima entrada de $\left[f^*\right]_{\mathcal{B}_W^*}^{\mathcal{B}_V^*}$ es $f^*(\lambda_i)(v_j) = a_{ij}$. $\square$

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

$$\begin{array}{rcl} \langle \bullet, \bullet \rangle : V \times V & \longrightarrow & \mathbb{R} \\ (v_1, v_2) & \longmapsto & \langle v_1; v_2 \rangle \end{array}$$

tal que:

1. *es bilineal*: para todo $v, v_1, v_2 \in V$ y $c \in \mathbb{R}$

$$\begin{array}{rcl} \langle v_1 + v_2; v \rangle & = & \langle v_1; v \rangle + \langle v_2; v \rangle \\ \langle cv_1; v_2 \rangle & = & c\langle v_1; v_2 \rangle \\ \langle v; v_1 + v_2 \rangle & = & \langle v; v_1 \rangle + \langle v; v_2 \rangle \\ \langle v_1; cv_2 \rangle & = & c\langle v_1; v_2 \rangle; \end{array}$$

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 $V = \mathbb{R}^n$,

$$\langle (x_1, \ldots, x_n); (y_1, \ldots, y_n) \rangle = \sum_{i=1}^{n} x_i y_i.$$

2. Sobre $V = M_{n\times n}(\mathbb{R})$,

$$\langle A; B \rangle = \operatorname{tr}(B^\intercal A).$$

3. Sea $[a,b] \subseteq \mathbb{R}$ un intervalo cerrado. Sobre $V = C^0[a,b]$, el conjunto de funciones continuas $[a,b] \to \mathbb{R}$,

$$\langle f; g \rangle = \int_a^b f(x)g(x)dx.$$

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, pósitiva 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 discrimante 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 $b = \|v_1\|$. □

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 está 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 $v_1, v_2 \in T$ 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 $v_1, \ldots, v_n \in S$ y $a_1, \ldots, a_n \in S$ son tales que

$$a_1v_1 + \ldots + a_nv_n = 0.$$

Entonces, para $i = 1, \ldots, n$,

$$0 = \langle 0; v_i\rangle = \langle a_1v_1 + \cdots + a_nv_n; v_i\rangle = a_1\langle v_1; v_i\rangle + \ldots + a_n\langle v_n; v_i\rangle = a_i\|v_i\|^2,$$

pero $v_i \neq 0$ pues $0 \notin S$, luego $\|v_i\|^2 \neq 0$, y así $a_i = 0$. □

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 $\{v_1, \ldots, v_n\}$ es una base de V, existe una base ortonormal $\{u_1, \ldots, u_n\}$ de V tal que para $k = 1, \ldots, n$*

$$\langle v_1, \ldots, v_k\rangle = \langle u_1, \ldots, u_k\rangle.$$

Dem. Sea $\{v_1, \ldots, v_n\}$ una base de V, definimos recursivamente $\{v_1', \ldots, v_n'\}$ 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 $\{v_1', \ldots, v_n'\}$ 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

$$\begin{aligned} \langle v_1 \rangle &= \langle v_1' \rangle \\ \langle v_1, \ldots, v_{k+1} \rangle &= \langle v_1', \ldots, v_{k+1}' \rangle \end{aligned}$$

pues $v_{k+1} - v_{k+1}' \in \langle v_1', \ldots, v_k' \rangle = \langle v_1, \ldots, v_k \rangle$. Note que, como $\{v_1, \ldots, v_n\}$ es linealmente independiente, entonces $v_{k+1} \notin \langle v_1, \ldots, v_k \rangle = \langle v_1', \ldots, v_k' \rangle$, luego $v_{k+1}' \notin \langle v_1', \ldots, v_k' \rangle$ y así, para $i = 1, \ldots, n$, $v_i' \neq 0$. De donde, como $\{v_1', \ldots, v_n'\}$ 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 $i = 1, \ldots, n$,

$$u_i = \frac{1}{\|v_i'\|} v_i'.$$

Tenemos para $k = 1, \ldots, n$

$$\langle v_1, \ldots, v_k \rangle = \langle v_1', \ldots, v_k' \rangle = \langle u_1, \ldots, u_k \rangle.$$

□

Propiedad 4.11. *Sea V un espacio euclídeo. Suponga que V tiene dimensión finita y $\{u_1, \ldots, u_n\}$ 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

$$v_1 = \sum_{i=1}^{n} x_i u_i \qquad v_2 = \sum_{i=1}^{n} y_i u_i,$$

entonces

$$\langle v_1; v_2 \rangle = \sum_{i=1}^{n} x_i y_i.$$

Dem. Como $\{u_1, \ldots, u_n\}$ es base, existen $a_1, \ldots, a_n \in K$ tal que $v = \sum_{i=1}^{n} a_i u_i$. De esta forma, para $j = 1, \ldots, n$,

$$\langle v; u_j \rangle = \sum_{i=1}^{n} a_i \langle u_i; u_j \rangle = \sum_{i=1}^{n} a_i \delta_{ij} = a_j$$

(ver Observación 4.8). Finalmente,

$$\langle v_1; v_2 \rangle = \langle \sum_{i=1}^{n} x_i u_i; \sum_{j=1}^{n} y_j u_j \rangle = \sum_{i=1}^{n} \sum_{j=1}^{n} x_i y_j \delta_{ij} = \sum_{i=1}^{n} x_i y_i.$$

□

Observación 4.12. Si $\mathcal{B} = \{u_1, \ldots, u_n\}$ es una base ortonormal de V, la propiedad anterior implica que para todo $v_1, v_2 \in V$ tenemos

$$\langle v_1; v_2 \rangle = \left(\left[v_2 \right]^{\mathcal{B}} \right)^{\mathsf{T}} \left[v_1 \right]^{\mathcal{B}}$$

Definición 4.13. Sean V un espacio euclídeo y $S \subseteq V$, el *conjunto ortogonal* a S está definido por

$$S^{\perp} = \{ v \in V | \ \langle v; u \rangle = 0, \text{ para todo } u \in S \}$$

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 $\{v_1, \ldots, v_n\}$ una base de V tal que $\{v_1, \ldots, v_m\}$ es una base de U. Sea $\{u_1, \ldots, u_n\}$ la base obtenida mediante ortogonalización a partir de $\{v_1, \ldots, v_n\}$. En particular $U = \langle u_1, \ldots, u_m\rangle$ y $\langle u_{m+1}, \ldots, u_n\rangle \leq U^{\perp}$. 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 $v \in \langle u_{m+1}, \ldots, u_n\rangle$. Entonces $U^{\perp} = \langle u_{m+1}, \ldots, u_n\rangle$ y $V = U \oplus U^{\perp}$. □

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 $\langle v_1, \ldots, v_m\rangle = U \leq V$. Si $\mathcal{B}$ es una base ortonormal de V y $\{v_1, \ldots, v_m\}$ es linealmente independiente entonces*

$$\left[p_U^{\perp}\right]_{\mathcal{B}}^{\mathcal{B}} = A(A^{\mathsf{T}}A)^{-1}A^{\mathsf{T}}$$

donde $A \in M_{n\times m}(\mathbb{R})$ es la matrix cuya j-ésima columna es $\left[v_j\right]^{\mathcal{B}}$, $j = 1, \ldots, n$.

Dem. Para todo $v \in V$ tenemos $p_U^{\perp}(v) \in U$ luego existe un único $\overline{c}_v \in M_{m\times 1}(\mathbb{R})$ tal que

$$\left[p_U^{\perp}(v)\right]^{\mathcal{B}} = Ac_v$$

Si $P = \left[p_U^{\perp}\right]^{\mathcal{B}}$, y $\overline{x} = \left[v\right]^{\mathcal{B}}$ entonces

$$P\overline{x} = A\overline{c}_v.$$

Ahora como $v - p_U^{\perp}(v) \in U^{\perp}$ entonces $\langle v - p_U^{\perp}(v); v_j\rangle = 0$ para $j = 1, \ldots, n$ luego (ver Observación 4.12)

$$\left(\left[v_j\right]^{\mathcal{B}}\right)^{\mathsf{T}} (\overline{x} - A\overline{c}_v) = 0$$

y

$$\begin{aligned} 0 &= A^{\mathsf{T}}(\overline{x} - A\overline{c}_v) \\ &= A^{\mathsf{T}}\overline{x} - A^{\mathsf{T}}A\overline{c}_v \end{aligned}$$

Veamos que $A^\intercal A \in M_{m\times m}(\mathbb{R})$ es invertible. De hecho, si $\{u_1,\ldots,u_m\}$ es una base ortonormal de U y $c_{ij} \in \mathbb{R}$, $i,j = 1,\ldots,m$, son tales que

$$u_j = c_{1j}v_1 + \ldots + c_{mj}v_m, \quad j = 1,\ldots,m$$

y $C = (c_{ij})$ entonces $C \in M_{m\times m}(\mathbb{R})$ es invertible y la j-ésima columna de AC es $\left[u_j\right]^{\mathcal{B}}$. Así

$$I_m = (AC)^\intercal AC = C^\intercal A^\intercal AC$$

y $A^\intercal A = (CC^\intercal)^{-1}$, luego $A^\intercal A$ es invertible. Tenemos entonces

$$\begin{aligned} \overline{c}_v &= (A^\intercal A)^{-1}A\overline{x} \\ A\overline{c}_v &= A(A^\intercal A)^{-1}A\overline{x} \\ P\overline{x} &= A(A^\intercal A)^{-1}A\overline{x} \end{aligned}$$

y se sigue $P = A(A^\intercal A)^{-1}A^\intercal$. □

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 $\{u_1,\ldots,u_m\}$ 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

$$v_1 = p_U^\perp(v_1) + v_1' \qquad v_2 = p_U^\perp(v_2) + v_2'.$$

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 $\{u_1,\ldots,u_m\}$ de U a una base ortonormal $\{u_1,\ldots,u_n\}$ 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 $f \in \mathrm{Hom}_{\mathbb{R}}(V, V)$ un operador.

Definición 4.20. Sea $g \in \mathrm{Hom}_{\mathbb{R}}(V, V)$, 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 $g \in \mathrm{Hom}_R(V, V)$ adjunto de f. Más aún, si $\mathcal{B} = \{u_1, \dots, u_n\}$ es una base ortonormal de V, entonces*

$$\Big[g\Big]_{\mathcal{B}}^{\mathcal{B}} = \left(\Big[f\Big]_{\mathcal{B}}^{\mathcal{B}}\right)^{\mathsf{T}}$$

Dem. Defina el operador $g \in \mathrm{Hom}_{\mathbb{R}}(V, V)$ 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, $h \in \mathrm{Hom}_{\mathbb{R}}(V, V)$ 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

$$\begin{aligned}
\Big[g\Big]^{\mathcal{B}}_{\mathcal{B},(j,i)} &= \Big[g(u_i)\Big]^{\mathcal{B}}_{j} \\
&= \langle g(u_i); u_j\rangle \\
&= \langle u_i; f(u_j)\rangle \\
&= \langle f(u_j); u_i\rangle \\
&= \Big[f(u_j)\Big]^{\mathcal{B}}_{i} \\
&= \Big[f\Big]^{\mathcal{B}}_{\mathcal{B},(i,j)}
\end{aligned}$$

□

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*

$$\begin{aligned}
\imath : V &\longrightarrow V^* \\
v &\longmapsto \imath_v = \langle \bullet; v\rangle : v' \mapsto \langle v'; v\rangle
\end{aligned}$$

es una transformación lineal inyectiva, la cual es un isomorfismo si V tiene dimensión finita.

Dem. Por la linearidad en el segundo factor del producto interno, $\imath$ es una transformación lineal. Suponga ahora que $\imath_v = 0$, en particular $0 = \imath_v(v) = \langle v, v\rangle = \|v\|^2$, luego $v = 0$ y así $\imath$ es inyectiva. Finalmente como $\dim(V) = \dim(V^*)$ cuando V tiene dimensión finita, entonces $\imath$ en este caso también es sobreyectiva, y es un isomorfismo.

Proposición 4.26. *Suponga que V tiene dimensión finita y sea*

$$\begin{aligned}
\widehat{\bullet} : V &\longrightarrow (V^*)^* \\
v &\longmapsto \widehat{v} : \lambda \mapsto \lambda(v).
\end{aligned}$$

el isomorfismo canónico. Entonces para todo $v \in V$

$$\imath^*(\widehat{v}) = \imath(v).$$

Dem. Para todo $v' \in V$

$$\begin{aligned}
\imath^*(\widehat{v})(v') &= \widehat{v}(\imath(v')) \\
&= \imath(v')(v) \\
&= \langle v; v'\rangle \\
&= \langle v'; v\rangle \\
&= \imath(v)(v').
\end{aligned}$$

□

Observación 4.27. Note que si V tiene dimensión finita, para todo $v' \in V$,

$$\begin{aligned} f^*(\imath_v)(v') &= \imath_v\left(f(v')\right) \\ &= \langle f(v'); v \rangle \\ &= \langle v'; f^*(v) \rangle \\ &= \imath_{f^*(v)}(v') \end{aligned}$$

luego $f^*(\imath_v) = \imath_{f^*(v)}$, es decir

$$f^* \circ \imath = \imath \circ f^*,$$

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 $\imath$, 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 $M_{m\times m}(\mathbb{R})$

$$A = \Big[\imath\Big]_{\mathcal{B}}^{\mathcal{B}^*}, \text{ y } B = \Big[\imath^*(\widehat{\bullet})\Big]_{\mathcal{B}}^{\mathcal{B}^*} = \Big[\imath^*\Big]_{\widehat{\mathcal{B}}}^{\mathcal{B}^*} \Big[\widehat{\bullet}\Big]_{\mathcal{B}}^{\widehat{\mathcal{B}}} = \Big[\imath^*\Big]_{\widehat{\mathcal{B}}}^{\mathcal{B}^*},$$

entonces $B = A$, pero por otro lado $B = A^\intercal$, luego $A^\intercal = A$. Es decir, la representación matricial de $\imath$ 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 $\Big[f\Big]_{\mathcal{B}}^{\mathcal{B}}$ es simétrica, es decir para todo $i, j \in \{1, \dots, n\}$

$$\langle f(u_j); u_i \rangle = \Big[f\Big]_{\mathcal{B},(i,j)}^{\mathcal{B}} = \Big[f\Big]_{\mathcal{B},(j,i)}^{\mathcal{B}} = \langle f(u_i); u_j \rangle,$$

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 $v_1, v_2 \in V$
$$\begin{aligned}\langle P(f)(v_1); v_2\rangle &= \langle \sum_{k=0}^{n} a_k f^k(v_1); v_2\rangle \\ &= \sum_{k=0}^{n} a_k\langle f^k(v_1); v_2\rangle \\ &= \sum_{k=0}^{n} a_k\langle v_1; f^k(v_2)\rangle \\ &= \langle v_1; \sum_{k=0}^{n} a_k f^k(v_2)\rangle \\ &= \langle v_1; P(f)(v_2)\rangle.\end{aligned}$$
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 $\lambda_1, \lambda_2 \in \mathbb{R}$, $\lambda_1 - \lambda_2 \neq 0$, 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 $\lambda_1 - \lambda_2 \neq 0$, $\langle v_1; v_2 \rangle = 0$. □

Teorema 4.33 (Teorema Espectral)**.** *Suponga que V tiene dimensión finita, f es auto-adjunta y que*

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

entonces existe una base ortonormal $\mathcal{B} = \{u_1, \ldots, u_n\}$ de V tal que $\left[f\right]_{\mathcal{B}}^{\mathcal{B}}$ es diagonal.

Dem. Por Teorema 2.53 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 nilpontente 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 $i = 1, \ldots, r$, denote V_i el espacio generado por los vectores propios de f asociados a λ_i, es decir

$$V_i = \{v \in V |\ f(v) = \lambda_i v\},$$

de forma que, como f es diagonalizable,

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

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 $\mathcal{B}_1, \ldots, \mathcal{B}_r$ son respectivamente bases ortonormales de $V_1, \ldots, V_r$,

$$\mathcal{B} = \mathcal{B}_1 \cup \ldots \cup \mathcal{B}_r$$

es una base ortonormal de V formada por vectores propios de f, en particular, $\left[f\right]_{\mathcal{B}}^{\mathcal{B}}$ es diagonal. □

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 $f \in \mathrm{Hom}_{\mathbb{R}}(V, V)$ 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:

$$\langle v_1; v_2\rangle = \frac{\|v_1 + v_2\|^2 - \left(\|v_1\|^2 + \|v_2\|^2\right)}{2}.$$

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 siguiente propiedades son equivalentes*

1. *f es ortogonal;*
2. *f preserva la norma;*
3. *$f^* \circ f = \mathrm{id}_V$; y,*
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 $\mathcal{B} = \{u_1, \ldots, u_n\}$ 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 $f^* \circ f = \mathrm{id}_V$, 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

$$v_1 = \sum_{i=1}^{n} x_i u_i \qquad v_2 = \sum_{i=1}^{n} y_i u_i,$$

entonces

$$f(v_1) = \sum_{i=1}^{n} x_i f(u_i) \qquad v_2 = \sum_{i=1}^{n} y_i f(u_i)$$

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

$$A^\intercal A = I_n$$

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 $\left[f\right]_{\mathcal{B}}^{\mathcal{B}}$ es ortogonal.*

Dem. La proposición se sigue del hecho que si $A = \left[f\right]_{\mathcal{B}}^{\mathcal{B}}$, entonces $A^\intercal = \left[f^*\right]_{\mathcal{B}}^{\mathcal{B}}$ y

$$A^\intercal A = \left[f^*\right]_{\mathcal{B}}^{\mathcal{B}} \left[f\right]_{\mathcal{B}}^{\mathcal{B}} = \left[f^* \circ f\right]_{\mathcal{B}}^{\mathcal{B}}.$$

Entonces $A^\intercal A = I_n$ si y solo si $f^* \circ f = \mathrm{id}_V$. □

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 $g \in \mathrm{Hom}_{\mathbb{R}}(V, V)$, 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 colleción de operadores ortogonales de V, X la de n-tuplas cuyas componentes formán bases ortogonales y

$$\begin{array}{rcl} \Phi_T : G & \longrightarrow & X \\ g & \longmapsto & (g(u_1), \dots, g(u_n)) \end{array}$$

es la correspondencia biyectiva definida por T (junto con un ordenamiento), entonces

$$\begin{array}{rcl} \Phi_T(\mathrm{id}_V) & = & (u_1, \dots u_n) \\ \Phi_T(g \circ h) & = & g\,(\Phi_T(h)) \end{array}$$

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

$$\begin{array}{rcl} \Phi : X \times G & \longmapsto & X \times X \\ (x, g) & \longmapsto & (x, gx) \end{array}$$

es una biyección, en la cual, si fijamos el primer factor en $x = (u_1, \dots u_n)$, obtenemos en el segundo la biyeccion 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 $\overline{c}$ a su conjugado el cual está definido por

$$\overline{c} = a - bi, \text{ si } c = a + bi, \ a, b \in \mathbb{R},$$

por $|c| = \sqrt{a^2 + b^2} = \sqrt{c\overline{c}}$ a su norma, por $\mathrm{Re}(c) = a = (c + \overline{c})/2$ su parte real y por $\mathrm{Im}(c) = b = (c - \overline{c})/2i$ 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

$$\begin{array}{rcl} \langle \bullet, \bullet \rangle : V \times V & \longrightarrow & \mathbb{C} \\ (v_1, v_2) & \longmapsto & \langle v_1; v_2 \rangle \end{array}$$

tal que:

1. *es sesquilineal*: para todo $v, v_1, v_2 \in V$ y $c \in \mathbb{C}$

$$\begin{array}{rcl} \langle v_1 + v_2; v \rangle & = & \langle v_1; v \rangle + \langle v_2; v \rangle \\ \langle cv_1; v_2 \rangle & = & c\langle v_1; v_2 \rangle \\ \langle v; v_1 + v_2 \rangle & = & \langle v; v_1 \rangle + \langle v; v_2 \rangle \end{array}$$

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 $c \in \mathbb{C}$

$$\langle v_1; cv_2\rangle = \overline{c}\langle v_1; v_2\rangle$$

Ejemplo 5.5. 1. Sobre $V = \mathbb{C}^n$,

$$\langle (z_1, \ldots, z_n); (w_1, \ldots, w_n)\rangle = \sum_{i=1}^{n} z_i \overline{w_i}.$$

2. Sobre $V = M_{n\times n}(\mathbb{C})$,

$$\langle A; B\rangle >= \mathrm{tr}(B^* A).$$

donde $B^* = \overline{B}^{\mathsf{T}}$ es la matriz cuyas entradas son las conjugadas de las entradas de las matriz traspuesta de B.

3. Sea $[a, b] \subseteq \mathbb{R}$ un intervalo cerrado. Sobre $V = C^0_{\mathbb{C}}[a, b]$, el conjunto de funciones continuas $[a, b] \to \mathbb{C}$,

$$\langle f; g\rangle = \int_a^b f(x)\overline{g(x)}dx.$$

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 $v \in V$

$$\|cv\| = \sqrt{\langle cv; cv\rangle} = \sqrt{c\overline{c}\langle v; v\rangle} = |c|\sqrt{\langle v; v\rangle} = |c|\|v\|.$$

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 $a, b \in \mathbb{C}$,

$$\begin{aligned} 0 \leq \|av_1 - bv_2\|^2 &= \langle av_1 - bv_2, av_1 - bv_2\rangle \\ &= a\overline{a}\langle v_1; v_1\rangle - b\overline{a}\langle v_2; v_1\rangle - a\overline{b}\langle v_1; v_2\rangle + b\overline{b}\langle v_2; v_2\rangle \\ &= |a|^2\|v_1\|^2 - \left(\overline{a}b\overline{\langle v_1; v_2\rangle} + a\overline{b}\langle v_1; v_2\rangle\right) + |b|^2\|v_2\|^2. \end{aligned}$$

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

$$\|av_1 - bv_2\|^2 = |a|^2\|v_1\|^2 - 2\operatorname{Re}\left(a\overline{b}\langle v_1; v_2\rangle\right) + |b|^2\|v_2\|^2.$$

$\square$

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 $v_1, v_2 \in T$ 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. $\square$

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 $\{v_1, \ldots, v_n\}$ es una base de V, existe una base ortonormal $\{u_1, \ldots, u_n\}$ de V tal que para $k = 1, \ldots, n$*

$$\langle v_1, \ldots, v_k \rangle = \langle u_1, \ldots, u_k \rangle.$$

Dem. El argumento es similar al de la demostración de Teorema 4.10. □

Propiedad 5.12. *Sea V un espacio unitario. Suponga que V tiene dimensión finita y $\{u_1, \ldots, u_n\}$ 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

$$v_1 = \sum_{i=1}^{n} z_i u_i \qquad v_2 = \sum_{i=1}^{n} w_i u_i,$$

entonces

$$\langle v_1; v_2 \rangle = \sum_{i=1}^{n} z_i \overline{w_i}.$$

Dem. Como $\{u_1, \ldots, u_n\}$ es base, existen $a_1, \ldots, a_n \in K$ tal que $v = \sum_{i=1}^{n} a_i u_i$. De esta forma, para $j = 1, \ldots, n$,

$$\langle v; u_j \rangle = \sum_{i=1}^{n} a_i \langle u_i, u_j \rangle = \sum_{i=1}^{n} a_i \delta_{ij} = a_j$$

(ver Observación 5.9). Finalmente,

$$\langle v_1; v_2 \rangle = \langle \sum_{i=1}^{n} z_i u_i, \sum_{j=1}^{n} w_j u_j \rangle = \sum_{i=1}^{n} \sum_{j=1}^{n} z_i \overline{w_j} \delta_{ij} = \sum_{i=1}^{n} z_i \overline{w_i}.$$

□

Definición 5.13. Sean V un espacio unitario y $S \subseteq V$, el *conjunto ortogonal* a S está definido por

$$S^{\perp} = \{v \in V | \ \langle v, u \rangle = 0, \text{ para todo } u \in S\}$$

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. □

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. □

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 $\{u_1, \ldots, u_m\}$ 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 demostracción de Propiedad 4.18. □

5.2. Operador adjunto

Sea V un espacio unitario y $f \in \mathrm{Hom}_{\mathbb{C}}(V, V)$ un operador.

Definición 5.18. Sea $g \in \mathrm{Hom}_{\mathbb{C}}(V, V)$, 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 $g \in \mathrm{Hom}_{\mathbb{C}}(V,V)$ adjunto de f. Más aún, si $\mathcal{B} = \{u_1, \dots, u_n\}$ es una base ortonormal de V, entonces*

$$\Big[g\Big]_{\mathcal{B}}^{\mathcal{B}} = \left(\Big[f\Big]_{\mathcal{B}}^{\mathcal{B}}\right)^*$$

Dem. Defina el operador $g \in \mathrm{Hom}_{\mathbb{C}}(V,V)$ 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, $h \in \mathrm{Hom}_{\mathbb{C}}(V,V)$ 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

$$\begin{aligned}
\Big[g\Big]^{\mathcal{B}}_{\mathcal{B},(j,i)} &= \Big[g(u_i)\Big]^{\mathcal{B}}_{j} \\
&= \langle g(u_i); u_j\rangle \\
&= \langle u_i; f(u_j)\rangle \\
&= \overline{\langle f(u_j); u_i\rangle} \\
&= \overline{\Big[f(u_j)\Big]^{\mathcal{B}}_{i}} \\
&= \overline{\Big[f\Big]}^{T}_{\mathcal{B},(i,j)}
\end{aligned}$$

□

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 $c \in \mathbb{C}$*

$$\begin{aligned}
h(v_1 + v_2) &= h(v_1) + h(v_2) \\
h(cv) &= \overline{c}h(v),
\end{aligned}$$

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 $c \in \mathbb{C}$

$$\begin{aligned}
h(v_1 + v_2)(v') &= \langle v'; v_1 + v_2\rangle \\
&= \langle v'; v_1\rangle + \langle v'; v_2\rangle \\
&= h(v_1)(v') + h(v_2)(v') \\
&= \left(h(v_1) + h(v_2)\right)(v), \text{ y} \\
h(cv)(v') &= \langle v'; cv\rangle \\
&= \overline{c}\langle v'; v\rangle \\
&= \overline{c}h(v)(v')
\end{aligned}$$

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 $\imath$ es inyectiva. Finalmente suponga que V tiene dimensión finita y sea $\mathcal{B} = \{u_1, \ldots, u_n\}$ una base de V, entonces $h(T) = \{h(u_1), \ldots, h(u_n)\}$ 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 $a_1, \ldots, a_n \in \mathbb{C}$ son tales que $\lambda = \sum_{i=1}^n a_i h(u_i)$ entonces $h(\sum_{i=1}^n \overline{a_i} u_i) = \lambda$, luego h es también sobreyectiva y así biyectiva. □

Observación 5.24. Note que si V tiene dimensión finita, para todo $v' \in V$,

$$\begin{array}{rcl} f^*(h_v)(v') & = & h_v\left(f(v')\right) \\ & = & \langle f(v'); v\rangle \\ & = & \langle v'; f^*(v)\rangle \\ & = & h_{f^*(v)}(v') \end{array}$$

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, $\mathcal{B} = \{u_1, \ldots, u_n\}$ y asumimos que $\left[f\right]_{\mathcal{B}}^{\mathcal{B}}$ es hermítica, es decir para todo $i, j \in \{1, \ldots, n\}$

$$\langle f(u_j); u_i\rangle = \left[f\right]_{\mathcal{B},(i,j)}^{\mathcal{B}} = \overline{\left[f\right]_{\mathcal{B},(j,i)}^{\mathcal{B}}} = \overline{\langle f(u_i); u_j\rangle} = \langle u_j; f(u_i)\rangle,$$

luego si $v_1 = \sum_{i=1}^n z_i u_i$ y $v_2 = \sum_{j=1}^n w_j u_j$

$$\begin{array}{rcl} \langle f(v_1); v_2\rangle & = & \displaystyle\sum_{i,j=1}^n z_i \overline{w_j} \langle f(u_i); u_j\rangle \\ & = & \displaystyle\sum_{i,j=1}^n z_i \overline{w_j} \langle u_i; f(u_j)\rangle \\ & = & \langle v_1; f(v_2)\rangle \end{array}$$

□

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 $P(t) = \sum_{k=0}^{n} a_k x^k$, para $k = 1, \ldots, n$ $a_k = \overline{a_k}$, y entonces para todo $v_1, v_2 \in V$
$$\begin{aligned}\langle P(f)(v_1); v_2\rangle &= \langle\sum_{k=0}^{n} a_k f^k(v_1); v_2\rangle \\ &= \sum_{k=0}^{n} a_k \langle f^k(v_1); v_2\rangle \\ &= \sum_{k=0}^{n} a_k \langle v_1; f^k(v_2)\rangle \\ &= \langle v_1; \sum_{k=0}^{n} \overline{a_k} f^k(v_2)\rangle \\ &= \langle v_1; P(f)(v_2)\rangle.\end{aligned}$$
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 $\lambda_1 - \lambda_2 =\neq 0$, $\langle v_1; v_2 \rangle = 0$. □

Teorema 5.29 (Teorema Espectral). *Suponga que V tiene dimensión finita y f es auto-adjunta entonces existe una base ortonormal $\mathcal{B} = \{u_1, \ldots, u_n\}$ de V tal que $\Big[f\Big]_{\mathcal{B}}^{\mathcal{B}}$ es diagonal.*

Dem. Como $\mathbb{C}$ es algebraicamente cerrado,

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

Como $\lambda_1, \lambda_2, \ldots, \lambda_r$ son valores propios, por el lema son valores reales y $P_f(t) \in \mathbb{R}[t]$. En particular, como $t - \lambda_1, t - \lambda_2, \ldots, t - \lambda_r \in \mathbb{R}[t]$, por la demostración de Teorema 2.53 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 nilpontente 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 $i = 1, \ldots, r$, denote V_i el espacio generado por los vectores propios de f asociados a λ_i, es decir

$$V_i = \{v \in V |\ f(v) = \lambda_i v\},$$

de forma que, como f es diagonalizable,

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

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 $\mathcal{B}_1, \ldots, \mathcal{B}_r$ son respectivamente bases ortonormales de $V_1, \ldots, V_r$,

$$\mathcal{B} = \mathcal{B}_1 \cup \ldots \cup \mathcal{B}_r$$

es una base ortonormal de V formada por vectores propios de f, en particular, $\Big[f\Big]_{\mathcal{B}}^{\mathcal{B}}$ es diagonal. □

Corolario 5.30. *Sea V un espacio euclídeo de dimensión finita y $f \in \mathrm{Hom}_{\mathbb{R}}(V, V)$ un operador auto-adjunto, entonces una base ortonormal $\mathcal{B} = \{u_1, \ldots, u_n\}$ de V tal que $\Big[f\Big]_{\mathcal{B}}^{\mathcal{B}}$ 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

$$\begin{array}{rcl} f_A : \mathbb{C}^n & \longrightarrow & \mathbb{C}^n \\ (z_1, \ldots, z_n) & \longmapsto & \left(\displaystyle\sum_{k=1}^{n} a_{1k} z_k, \ldots, \sum_{k=1}^{n} a_{nk} z_k \right) \end{array}$$

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

$$\langle (z_1, \ldots, z_n), (w_1, \ldots, w_n) \rangle = \sum_{k=1}^{n} z_k \overline{w_k}.$$

Así

$$P_f(t) = P_{f_A}(t) = (t-\lambda_1)^{m_1}(t-\lambda_2)^{m_2} \ldots (t-\lambda_r)^{m_r}, \quad \lambda_1, \lambda_2, \ldots, \lambda_r \in \mathbb{R}.$$

y la conclusión se sigue ahora de Teorema 4.33. □

5.3. Operadores unitarios

Sea V un espacio unitario y $f \in \mathrm{Hom}_{\mathbb{C}}(V,V)$ 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

$$\begin{aligned}
\|v_1 + v_2\|^2 &= \langle v_1 + v_2; v_1 + v_2 \rangle \\
&= \langle v_1; v_1 \rangle + \langle v_2; v_1 \rangle + \langle v_1; v_2 \rangle + \langle v_2; v_2 \rangle \\
&= \langle v_1; v_1 \rangle + \overline{\langle v_1; v_2 \rangle} + \langle v_1; v_2 \rangle + \langle v_2; v_2 \rangle \\
&= \|v_1\|^2 + 2\,\mathrm{Re}\,(\langle v_1; v_2 \rangle) + \|v_2\|^2, \text{ y} \\
\|v_1 + iv_2\|^2 &= \langle v_1 + iv_2; v_1 + iv_2 \rangle \\
&= \langle v_1; v_1 \rangle + i\langle v_2; v_1 \rangle - i\langle v_1; v_2 \rangle + \langle v_2; v_2 \rangle \\
&= \langle v_1; v_1 \rangle - i\left(\langle v_1; v_2 \rangle - \overline{\langle v_1; v_2 \rangle}\right) + \langle v_2; v_2 \rangle \\
&= \|v_1\|^2 + 2\,\mathrm{Im}\,(\langle v_1; v_2 \rangle) + \|v_2\|^2,
\end{aligned}$$

de forma que el producto interno se puede expresar en términos de la norma:

$$\langle v_1; v_2 \rangle = \frac{\|v_1 + v_2\|^2 - \left(\|v_1\|^2 + \|v_2\|^2\right)}{2} + \frac{\|v_1 + iv_2\|^2 - \left(\|v_1\|^2 + \|v_2\|^2\right)}{2} i.$$

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. *$f^* \circ f = \mathrm{id}_V$; 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 $\mathcal{B} = \{u_1, \ldots, u_n\}$ es una base ortonormal de V, f es unitario si y solo si $\left[f\right]_{\mathcal{B}}^{\mathcal{B}}$ es unitaria.*

Dem. La proposición se sigue del hecho que si $A = \left[f\right]_{\mathcal{B}}^{\mathcal{B}}$, entonces $A^* = \left[f^*\right]_{\mathcal{B}}^{\mathcal{B}}$ y

$$A^*A = \left[f^*\right]_{\mathcal{B}}^{\mathcal{B}}\left[f\right]_{\mathcal{B}}^{\mathcal{B}} = \left[f^* \circ f\right]_{\mathcal{B}}^{\mathcal{B}}.$$

Entonces $A^*A = I_n$ si y solo si $f^* \circ f = \mathrm{id}_V$. $\square$

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 $(v_1, \ldots, v_n) \in V \times \ldots \times V$ tales que $\{v_1, \ldots, v_n\}$ 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 $f \in \mathrm{Hom}_{\mathbb{R}}(V, V)$ es un operador simple, entonces: o bien,*

1. *$f = c\,\mathrm{id}_V$, para algún $c \in \mathbb{R}$ y en tal caso $\dim(V) = 1$; o bien,*

2. *$f = a\,\mathrm{id}_V + bj$, donde $j \in \mathrm{Hom}_{\mathbb{R}}(V, V)$ es tal que $j^2 = -\mathrm{id}_V$, $a, b \in \mathbb{R}$, y en tal caso $\dim(V) = R^2$.*

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

$f - c\,\mathrm{id}_V = 0$, es decir $f = c\,\mathrm{id}_V$; en el segundo, $(f - a\,\mathrm{id}_V)^2 + b^2\,\mathrm{id}_V = 0$, y si $j = \frac{1}{b}(f - a\,\mathrm{id}_V)$, tenemos $f = a\,\mathrm{id}_V + bj$ con

$$\begin{aligned} j^2 &= \left(\frac{1}{b}(f - a\,\mathrm{id}_V)\right) \circ \left(\frac{1}{b}(f - a\,\mathrm{id}_V)\right) \\ &= \frac{1}{b^2}\,(f - a\,\mathrm{id}_V)^2 \\ &= \frac{1}{b^2}\left(-b^2\,\mathrm{id}_V\right) = -\,\mathrm{id}_V\,. \end{aligned}$$

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 $V = \langle v, j(v) \rangle$. □

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 $\mathrm{Hom}_{\mathbb{R}}(\mathbb{C}, \mathbb{C})$, 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

$$\left[f\right]_{\mathcal{B}}^{\mathcal{B}} = \begin{bmatrix} a & -b \\ b & a \end{bmatrix}$$

y $f = a\,\mathrm{id}_{\mathbb{C}} + bj$ donde $j = m_i$. Note que

$$\left[j\right]_{\mathcal{B}}^{\mathcal{B}} = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$$

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

$$\mathcal{B} = \{e_1, \ldots, e_n, f_1, \ldots, f_n\}$$

donde $\{e_1, \ldots, e_n\}$ es la base canónica de $\mathbb{C}^n$, visto como espacio vectorial sobre $\mathbb{C}$, y, $f_1 = ie_i, \ldots, f_n = ie_n$. Tome $j \in \operatorname{Hom}_{\mathbb{R}}(\mathbb{C}^n, \mathbb{C}^n)$ definida por

$$\begin{array}{rcl} j : \mathbb{C}^n & \longrightarrow & \mathbb{C}^n \\ (z_1, \ldots, z_n) & \longmapsto & i(z_1, \ldots, z_n). \end{array}$$

Entonces $j^2 = -\operatorname{id}_{\mathbb{C}^n}$ y

$$\left[j\right]_{\mathcal{B}}^{\mathcal{B}} = \begin{bmatrix} 0 & -I_n \\ I_n & 0 \end{bmatrix}$$

donde 0 denota el origen de $M_{n\times n}(\mathbb{R})$ y $I_n \in \mathbb{M}_{n\times n}(\mathbb{R})$ 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 $j \in \operatorname{Hom}_{\mathbb{R}}(V, V)$ tal que $j^2 = -\operatorname{id}_V$.

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*

$$(a + bi)v = (a \operatorname{id}_V + bj)(v)$$

para todo $a, b \in \mathbb{R}$ y $v \in V$. Más aún, existe una base de la forma $T = \{v_1, \ldots, v_n, j(v_1), \ldots, j(v_n)\}$ donde $2n = \dim(V)$. En particular

$$\left[j\right]_{\mathcal{B}}^{\mathcal{B}} = \begin{bmatrix} 0 & -I_n \\ I_n & 0 \end{bmatrix}$$

Dem. Si m es la dimensión de V entonces

$$0 \leq (\det(j))^2 = \det\left(j^2\right) = \det(-\operatorname{id}_V) = (-1)^m$$

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

$$\begin{array}{rcl}
1v & = & \operatorname{id}_V(v) = v \\
(a+bi)\left((c+di)v\right) & = & (a \operatorname{id}_V + bj) \circ (c \operatorname{id}_V + dj)(v) \\
& = & \left(ac \operatorname{id}_V + adj + bcj + bdj^2\right)(v) \\
& = & \left((ac-bd)\operatorname{id}_V + (ad+bc)j\right)(v) \\
& = & \left((a+bi)(c+di)\right)v \\
(a+bi)(v+w) & = & (a \operatorname{id}_V + bj)(v+w) \\
& = & (a \operatorname{id}_V + bj)(v) + (a \operatorname{id}_V + bj)(w) \\
& = & (a+bi)v + (a+bi)w \\
\left((a+bi)+(c+di)\right)v & = & \left((a+c)\operatorname{id}_V + (b+d)j\right)v \\
& = & (a \operatorname{id}_V + bj + c \operatorname{id}_V + dj)\, v \\
& = & (a \operatorname{id}_V + bj)(v) + (c \operatorname{id}_V + dj)(v) \\
& = & (a+bi)v + (c+di)v
\end{array}$$

validas para todo $a, b, c, d \in \mathbb{R}$ y $v, w \in V$.
Sea $\{v_1, \ldots, v_{n'}\}$ 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 $\{v_1, \ldots, v_{n'}, j(v_1), \ldots, j(v_{n'})\}$ pues

$$\sum_{k=1}^{n'} z_k v_k = \sum_{k=1}^{n'} a_k v_k + b_k j(v_k),$$

donde para $k = 1, \ldots, n'$, $z_k = a_k + b_k i$ con $a_k, b_k \in \mathbb{R}$. Más este conjunto de generadores es linealmente independiente pues si $a_1, \ldots, a_{n'}, b_1, \ldots, b_{n'} \in \mathbb{R}$ 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 $\mathcal{B} = \{v_1, \ldots, v_n, j(v_1), \ldots, j(v_n)\}$ es una base de V como espacio vectorial sobre $\mathbb{R}$. $\square$

Propiedad 5.43. *Sean j una estructura compleja en V y $f \in \mathrm{Hom}_{\mathbb{R}}(V, V)$. Entonces, tomando V como un espacio vectorial sobre $\mathbb{C}$ mediante el producto por escalar definido por $(a + bi)v = (a \,\mathrm{id}_V + bj)(v)$, para todo $a, b \in \mathbb{R}$ y $v \in V$, tenemos $f \in \mathrm{Hom}_{\mathbb{C}}(V, V)$ si y solo si $f \circ j = j \circ f$, o, equivalentemente, $-j \circ f \circ j = f$.*

Dem. Suponga que $f \in \mathrm{Hom}_{\mathbb{C}}(V, V)$, entonces para todo $v \in V$

$$f \circ j(v) = f(iv) = i f(v) = j \circ f(v).$$

Recíprocamente, si $f \circ j = j \circ f$, para todo $a, b \in \mathbb{R}$ y $v \in V$,

$$\begin{aligned} f((a + bi)v) &= f \circ (a \,\mathrm{id}_V + bj)(v) \\ &= (af + bf \circ j)(v) \\ &= (af + bj \circ f)(v) \\ &= (a \,\mathrm{id}_V + bj) \circ f(v) \\ &= (a + bi) f(v) \end{aligned}$$

luego $f \in \mathrm{Hom}_{\mathbb{C}}(V, V)$. 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 $j \in \mathrm{Hom}_{\mathbb{R}}(\mathbb{C}^n, \mathbb{C}^n)$ el operador que corresponde a multiplicación por i. Entonces cada estructura complejas en V es de la forma*

$$j_f = f \circ j \circ f^{-1}$$

para algún isomorfismos $f \in \mathrm{Hom}_{\mathbb{R}}(\mathbb{C}^n, V)$. Más aún $j_f = j_g$ si y solo si $g^{-1} \circ f \in \mathrm{Hom}_{\mathbb{C}}(\mathbb{C}^n, \mathbb{C}^n)$.

Dem. Dado un isomorfismo $f \in \mathrm{Hom}_{\mathbb{R}}(\mathbb{C}^n, V)$, el operador

$$j_f = f \circ j \circ f^{-1} \in \mathrm{Hom}_{\mathbb{R}}(V, V)$$

es una estructura compleja (ver Figura 5.1), pues

$$\begin{aligned} j_f^2 &= f \circ j \circ f^{-1} \circ f \circ j \circ f^{-1} \\ &= f \circ j^2 \circ f^{-1} \\ &= f \circ -\operatorname{id}_{\mathbb{C}^n} \circ f^{-1} \\ &= -f \circ f^{-1} \\ &= -\operatorname{id}_V . \end{aligned}$$

Recíprocamente, dada una estructura compleja j_V en V, usando la notación en

$$\begin{tikzcd} \mathbb{C}^n \arrow[r, "f"] \arrow[d, "j"] & V \arrow[d, dashed, "j_f"] \\ \mathbb{C}^n \arrow[r, "f"] & V \end{tikzcd}$$

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 $\{v_1, \ldots, v_n, j_V(v_1), \ldots, j_V(v_n)\}$ es una base de V, es un isomorfismo de espacios vectoriales sobre $\mathbb{R}$ tal que $j_f = j_V$. Finalmente, suponga que $f, g \in \operatorname{Hom}_{\mathbb{R}}(\mathbb{C}^n, V)$ 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 $g^{-1} \circ f \in \operatorname{Hom}_{\mathbb{C}}(\mathbb{C}^n, \mathbb{C}^n)$. $\square$

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 $(a + bi)v = (a \operatorname{id}_V + bj)(v)$, 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 $v_1, v_2 \in V$

$$\begin{aligned}
\langle j(v_1); j(v_2)\rangle &= \operatorname{Re}\left(\langle j(v_1); j(v_2)\rangle_j\right) \\
&= \left(\langle j(v_1); j(v_2)\rangle_j + \langle j(v_2); j(v_1)\rangle_j\right)/2 \\
&= \left(\langle iv_1; iv_2\rangle_j + \langle iv_2; iv_1\rangle_j\right)/2 \\
&= \left(\langle v_1; v_2\rangle_j + \langle v_2; v_1\rangle_j\right)/2 \\
&= \operatorname{Re}\left(\langle v_1; v_2\rangle_j\right) \\
&= \langle v_1; v_2\rangle
\end{aligned}$$

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. $\square$

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 $f \in \mathrm{Hom}_{\mathbb{C}}(V, V)$, $f \circ j = j \circ f$, 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 ademas 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 $f \in \mathrm{Hom}_{\mathbb{C}}(V, V)$ 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 $\mathrm{id}_V = f^* \circ f$, 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 capitulo donde entraremos a estudiar este tipo de estructuras que se llaman espacios simplécticos, las cuales generalizan está 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

$$\begin{array}{rcl} \sigma : V \times V & \longrightarrow & K \\ (v_1, v_2) & \longmapsto & \sigma(v_1, v_2) \end{array}$$

tal que:

1. *es bilineal*: para todo $v, v_1, v_2 \in V$ y $c \in \mathbb{C}$

$$\begin{array}{rcl} \sigma(v_1 + v_2, v) & = & \sigma(v_1, v) + \sigma(v_2, v) \\ \sigma(cv_1, v_2) & = & c\sigma(v_1, v_2) \\ \sigma(v, v_1 + v_2) & = & \sigma(v, v_1) + \sigma(v, v_2) \\ \sigma(v_1, cv_2) & = & c\sigma(v_1, v_2); \end{array}$$

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 $V = K^{2n} = K^n \times K^n$

$$\sigma\left((\overline{q},\overline{p}),(\overline{q}',\overline{p}')\right) = \sum_{i=1}^{n} p_i q_i' - p_i' q_i$$

donde $\overline{q} = (q_1,\dots,q_n)$, $\overline{p} = (p_1,\dots,p_n)$, $\overline{q}' = (q_1',\dots,q_n')$ y $\overline{p}' = (p_1',\dots,p_n')$.

2. Sobre $V \times V^*$

$$\sigma\left((v,\lambda),(w,\mu)\right) = \lambda(w) - \mu(v)$$

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 $V \to V^*$ y su dual s^* es una transformación lineal $(V^*)^* \to V^*$.

Proposición 6.5. *Sea V un espacio simpléctico de dimensión finita y*

$$\begin{aligned} \widehat{\bullet}: V &\longrightarrow (V^*)^* \\ v &\longmapsto \widehat{v}: \lambda \mapsto \lambda(v). \end{aligned}$$

el isomorfismo canónico. Entonces para todo $v \in V$

$$s^*(\widehat{v}) = -s(v)$$

Dem. Para todo $w \in V$

$$\begin{aligned} s^*(\widehat{v})(w) &= \widehat{v}(s(w)) \\ &= s(w)(v) \\ &= \sigma(w,v) \\ &= -\sigma(v,w) \\ &= -s(v)(w). \end{aligned}$$

$\square$

Propiedad 6.6. *Si V es un espacio simpléctico y tiene dimensión finita, entonces su dimensión es par.*

Dem. Sea $T = \{v_1, \ldots, v_m\}$ una base de V y $T^* = \{\lambda_1, \ldots, \lambda_m\}$ la base de V^* dual de T. Tomamos la imagen de T mediante el isomorfismo canónico $V \mapsto (V^*)^*$, la cual es la base $\widehat{T} = \{\widehat{v_1}, \ldots, \widehat{v_m}\}$ de $(V^*)^*$ dual de T^*. La proposición anterior implica que si tomamos las representaciones matriciales en $M_{m\times m}(K)$

$$A = \Big[s\Big]_T^{T^*}, \text{ y } B = \Big[s^*(\widehat{\bullet})\Big]_T^{T^*} = \Big[s^*\Big]_{\widehat{T}}^{T^*}\Big[\widehat{\bullet}\Big]_T^{\widehat{T}} = \Big[s^*\Big]_{\widehat{T}}^{T^*},$$

entonces $B = -A$, pero por otro lado $B = A^\intercal$, luego $A^\intercal = -A$. De donde

$$\det(A) = \det(A^\intercal) = \det(-A) = (-1)^m \det(A).$$

Ahora como $\sigma : V \to V^*$ es inyectiva, $\det(A) \neq 0$ y así $1 = (-1)^m$, en particular $m = 2n$ para algún $n \in \mathbb{Z}_{\geq 0}$. $\square$

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 $i = 1, \ldots, n$, 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 $i, j \in \{1, \ldots, n\}$,

$$\sigma(e_i, f_j) = -\delta_{ij} \quad \sigma(e_i, e_j) = \sigma(f_i, f_j) = 0.$$

Entonces para cualquier subconjunto de indices $J \subset I = \{1, \ldots, n\}$,

$$V_J = \mathrm{Sp}\left(\{e_j, f_j\}_{j \in J}\right)$$

es un subespacio simpléctico,

$$E_J = \mathrm{Sp}\left(\{e_j\}_{j \in J}\right), \text{ y } F_J = \mathrm{Sp}\left(\{f_j\}_{j \in J}\right)$$

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 $i, j \in I$

$$\sigma\left((v_i, 0), (0, \lambda_j)\right) = -\delta_{ij} \quad \sigma\left((v_i, 0), (v_j, 0)\right) = 0$$

y para todo $i, j \in I'$

$$\sigma\left((0, f_i), (0, f_j)\right) = 0.$$

Entonces para cualquier subconjunto de indices $J \subset I$,

$$(V \times V^*)_J = \mathrm{Sp}\left(\{(v_j, 0), (0, \lambda_j)\}_{j \in J}\right)$$

es un subespacio simpléctico,

$$E_J = \mathrm{Sp}\left(\{(v_j, 0)\}_{j \in J}\right), \text{ y } F_J = \mathrm{Sp}\left(\{(0, \lambda_j)\}_{j \in J}\right)$$

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 subesapcio 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. Reciprocamente, 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 simplectico. 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 $\sigma(u, u') = 0$. Reciprocamente, 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 máximal, 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 $U + \mathrm{Sp}(\{u'\})$ 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 $u_1 + au, u_2 + bu \in U' = U_0 + \mathrm{Sp}(\{u\})$, $u_1, u_2 \in U$ 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 monotonia 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 contenencia. 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 máximal en P por la proposión 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 $U \cap V_1 = \{0\}$, exite $V_2 \leq V$ 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 $u_1 + au, u_2 + bu \in U' = U + \mathrm{Sp}(\{u\})$, $u_1, u_2 \in U$ 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 monotonia 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 $V = V_1 \oplus V_2$. □

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 $\pi_1 : V \to V_1$ y $\pi_2 : V \to V_2$ son respectivamente las proyecciones sobre V_1 y V_2 dadas por la descomposición $V = V_1 \oplus V_2$,

$$\pi_1^*(V_1^*) = s(V_2) \qquad y \qquad \pi_2^*(V_2^*) = s(V_1)$$

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 $\sigma : V \to V^*$ 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 $\pi_1^*(V_1^*) \subseteq S(V_2)$. Reciprocamente, 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

$$\begin{array}{ccc} V_2 & \longrightarrow & \pi_1^*(V_1^*) \\ v & \longmapsto & s(v) = \lambda = \pi_1^*(\lambda_1) \end{array}$$

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 $T = \{v_1, \ldots, v_n, w_1, \ldots, w_n\}$ una base de V. Decimos que T es una base de Darboux si para todo $i, j \in \{1, \ldots, n\}$

$$\begin{array}{rclcl} \sigma(v_i, v_j) & = & 0 & = & \sigma(w_i, w_j) \\ \sigma(w_i, v_j) & = & \delta_{ij}. & & \end{array}$$

Observación 6.19. Note que si $T = \{v_1, \ldots, v_n, w_1, \ldots, w_n\}$ es una base de Darboux entonces $V_1 = \mathrm{Sp}(\{v_1, \ldots, v_n\})$ y $V_2 = \mathrm{Sp}(\{w_1, \ldots, w_n\})$ son lagrangianos. Más aún si $T_1 = \{v_1, \ldots, v_n\}$, $T_2 = \{w_1, \ldots, w_n\}$ y $T_1^* = \{\lambda_1, \ldots, \lambda_n\}$ y $T_2^* = \{\mu_1, \ldots, \mu_n\}$ son las bases de V_1^* y V_2^*, respectivamente, duales de T_1 y T_2. Entonces para $i = 1, \ldots, n$,

$$s(w_i) = \pi_1^*(\lambda_i) \qquad s(-v_i) = \pi_2^*(\mu_i).$$

$$\begin{tikzcd}
V_1 \arrow[rrr, dash, "s"] & & & \pi_2^*(V_2^*) \arrow[dl] \\
V_1 \oplus V_2 \arrow[r, equal] & V \arrow[ul, "\pi_1"] \arrow[dl, "\pi_2"'] \arrow[r, dash, "s"] & V^* = \pi_2^*(V_2^*) \oplus \pi_1^*(V_1^*) & \\
V_2 \arrow[rrr, dash, "s"] & & & \pi_1^*(V_1^*) \arrow[ul]
\end{tikzcd}$$

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 $T_1 = \{v_1, \ldots, v_n\}$ una base de V_1 y $T_1^* = \{\lambda_1, \ldots, \lambda_n\}$ la base de V_1^* dual de T_1. Como $s(V_2) = \pi_1^*(V_1^*)$, donde $\pi_1 : V \to V_1$ es la proyección en el primer sumando de la descomposición $V = V_1 \oplus V_2$, entonces existe $T_2 = \{w_1, \ldots, w_n\}$ base de V_2 tal que para $i = 1, \ldots, n$,

$$s(w_i) = \pi_1^*(\lambda_i).$$

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 $i, j \in \{1, \ldots, n\}$; además

$$\sigma(w_i, v_j) = s(w_i)(v_j) = \pi_1^*(\lambda_i)(v_j) = \lambda_i\left(\pi_1(v_j)\right) = \lambda_i(v_j) = \delta_{ij}.$$

□

Propiedad 6.21. *Suponga que V tiene dimensión finita y sea $T = \{v_1, \ldots, v_n, w_1, \ldots, w_n\}$ 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

$$v_1 = \sum_{i=1}^{n} q_i v_i + p_i w_i \qquad v_2 = \sum_{i=1}^{n} q_i' v_i + p_i' w_i$$

entonces

$$\sigma(v_1, v_2) = \sum_{i=1}^{n} p_i q_i' - p_i' q_i$$

Dem. Sean $a_1, \ldots, a_n, b_1, \ldots, b_n \in K$ tales que

$$v = \sum_{i=1}^{n} a_i v_i + b_i w_i,$$

de forma que para $j = 1, \ldots, n$,

$$\sigma(v_j, v) = \sum_{i=1}^{n} a_i \sigma(v_j, v_i) + b_i \sigma(v_j, w_i) = -b_j$$

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

$$\begin{aligned}
\sigma(v_1, v_2) &= \sigma(\sum_{i=1}^{n} q_i v_i + p_i w_i, \sum_{j=1}^{n} q'_j v_j + p'_j w_j) \\
&= \sum_{i,j=1}^{n} q_i q'_j \sigma(v_i, v_j) + p_i q'_j \sigma(w_i, v_j) + q_i p'_j \sigma(v_i, w_j) + p_i p'_j \sigma(w_i, w_j) \\
&= \sum_{i,j=1}^{n} (p_i q'_j - q_i p'_j)\delta_{ij} \\
&= \sum_{i=1}^{n} p_i q'_i - p'_i q_i
\end{aligned}$$

□

Observación 6.22. Note que si $T = \{v_1, \ldots, v_n, w_1, \ldots, w_n\}$, es una base de Darboux, para todo $J \subseteq \{1, \ldots, n\}$,

$$V_J = \mathrm{Sp}\left(\{v_j, w_j\}_{j \in J}\right)$$

es también un subespacio simpléctico.

Teorema 6.23 (Ortogonalización de Gram-Schmidt)**.** *Suponga que* $\dim(V) = 2n$ *y sean* $U \leq V$ *subespacios simpléctico con* $\dim(U) = 2m$, $m \leq n$. *Entonces existe una base de Darboux* $T = \{v_1, \ldots, v_n, w_1, \ldots, w_n\}$ *de* V, *tal que*

$$U = \mathrm{Sp}\left(\{v_j, w_j\}_{j=1,\ldots,m}\right)$$

Dem. Sea $T' = \{v_1, \ldots, v_m, w_1, \ldots, w_m\}$ 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 $j = 1, \ldots, m$,

$$\begin{aligned}
\sigma(v_j, v_{m+1}) &= \sigma(v_j, v'_{m+1}) - \left(\sum_{i=1}^{m} \sigma(w_i, v'_{m+1}) \sigma(v_j, v_i) - \sigma(v_i, v'_{m+1}) \sigma(v_j, w_i)\right) \\
&= \sigma(v_j, v'_{m+1}) - \sum_{i=1}^{m} \sigma(v_i, v'_{m+1}) \delta_{ij} \\
&= \sigma(v_j, v'_{m+1}) - \sigma(v_j, v'_{m+1}) \\
&= 0;
\end{aligned}$$

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 $U' = U + \mathrm{Sp}\left(\{v_{m+1}, w_{m+1}\right)$. Reemplazamos U por U' y continuamos recursivamente. $\square$

Propiedad 6.24. *Suponga que V tiene dimensión finita y sea $U < V$ subespacio simplectico. 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

$$U = \mathrm{Sp}\left(\{v_j, w_j\}_{j=1,\dots,m}\right)$$

donde $2n = \dim(V)$ y $2m = \dim(U)$, $m < n$. Entonces si

$$U' = \mathrm{Sp}\left(\{v_j, w_j\}_{j=m+1,\dots,n}\right),$$

$\dim(U') = \dim(U^\sigma)$ 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 $p_1 : V \to V$ *y* $p_2 : V \to V$ *las respectivas proyecciones sobre* V_1 *y* V_2 *definidas por esta descomposición. Entonces para todo* $v, w \in V$,
$$\sigma\left(p_1(v), w\right) = \sigma\left(v, p_2(w)\right).$$
Por otro lado, sea $U < V$ *subespacio simpléctico, entonces para todo* $v, w \in V$,
$$\sigma\left(p_U^\sigma(v), w\right) = \sigma\left(v, p_U^\sigma(w)\right).$$

Dem. Sean $v_1, w_1 \in V_1$ y $v_2, w_2 \in V_2$ tales que
$$v = v_1 + w_1 \qquad w = w_1 + w_2,$$
entonces
$$\begin{aligned} \sigma\left(p_1(v), w\right) &= \sigma(v_1, w_1) + \sigma(v_1, w_2) = \sigma(v_1, w_2) \\ \sigma\left(v, p_2(w)\right) &= \sigma(v_1, w_2) + \sigma(v_2, w_2) = \sigma(v_1, w_2). \end{aligned}$$
Considere ahora $v', w' \in U^\sigma$ tales que
$$v = p_U^\sigma(v) + v' \qquad w = p_U^\sigma(w) + w'.$$
Entonces
$$\begin{aligned} \sigma\left(p_U^\sigma(v), w\right) &= \sigma\left(p_U^\sigma(v), p_U^\sigma(w)\right) + \sigma\left(p_U^\sigma(v), w'\right) = \sigma\left(p_U^\sigma(v), p_U^\sigma(w)\right) \\ \sigma\left(v, p_U^\sigma(w)\right) &= \sigma\left(p_U^\sigma(v), p_U^\sigma(w)\right) + \sigma\left(v', p_U^\sigma(w)\right) = \sigma\left(p_U^\sigma(v), p_U^\sigma(w)\right). \end{aligned}$$
□

6.3. Operadores adjuntos

Sea V un espacio simpléctico y $f \in \mathrm{Hom}_{\mathbb{K}}(V, V)$ un operador.

Definición 6.27. Sea $g \in \mathrm{Hom}_{\mathbb{K}}(V, V)$, decimos que g es un *operador adjunto de* f si para todo $v, w \in V$
$$\sigma\left(f(v), w\right) = \sigma\left(v, g(w)\right).$$
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

$$\sigma\left(g(v),w\right)=-\sigma\left(w,g(v)\right)=-\sigma\left(f(w),v\right)=\sigma\left(v,f(w)\right).$$

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 $A^{sigma}\in M_{2n\times 2n}(K)$ tal que

$$A^{\sigma}=J^{-1}A^{\intercal}J$$

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 $I_n\in\mathbb{M}_{n\times n}(K)$ 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

$$A^{\sigma}=\begin{bmatrix} A_{22}^{\intercal} & -A_{12}^{\intercal} \\ -A_{21}^{\intercal} & A_{11}^{\intercal} \end{bmatrix}.$$

Decimos que A es *auto-adjunta simpléctica* si $A^{\sigma}=A$. Es decir si $A_{11}=A_{22}^{\intercal}$, $A_{12}=-A_{12}^{\intercal}$ y $A_{21}=-A_{12}^{\intercal}$.

Proposición 6.30. *Suponga que V tiene dimensión finita, entonces existe un único operador $g\in\mathrm{Hom}_K(V,V)$ adjunto de f. Más aún, si $T=\{v_1,\ldots,v_n,w_1,\ldots,w_n\}$ es una base de Darboux de V, entonces*

$$\Big[g\Big]_T^T=\left(\Big[f\Big]_T^T\right)^{\sigma}$$

Dem. Defina el operador $g\in\mathrm{Hom}_K(V,V)$ 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

$$\begin{aligned} \sigma\left(v_i,g(v_j)\right) &= \sigma\left(f(v_i),v_j\right) \\ \sigma\left(v_i,g(w_j)\right) &= \sigma\left(f(v_i),w_j\right) \\ \sigma\left(w_i,g(v_j)\right) &= \sigma\left(f(w_i),v_j\right) \\ \sigma\left(w_i,g(w_j)\right) &= \sigma\left(f(w_i),w_j\right) \end{aligned}$$

y por Propiedad 6.21

$$
\begin{aligned}
\sigma\left(v, g(w)\right) &= \sum_{i=1}^{n} \sigma\left(v_i, g(w)\right)\sigma\left(w_i, v\right) - \sigma\left(v_i, v\right)\sigma\left(w_i, g(w)\right) \\
&= \sum_{i,j=1}^{n} \Big(\sigma\left(w_j, w\right)\sigma\left(v_i, g(v_j)\right) - \sigma\left(v_j, w\right)\sigma\left(v_i, g(w_j)\right)\Big)\sigma\left(w_i, v\right) \\
&\quad -\sigma\left(v_i, v\right)\Big(\sigma\left(w_j, w\right)\sigma\left(w_i, g(v_j)\right) - \sigma\left(v_j, w\right)\sigma\left(w_i, g(w_j)\right)\Big) \\
&= \sum_{i,j=1}^{n} \Big(\sigma\left(w_j, w\right)\sigma\left(f(v_i), v_j\right) - \sigma\left(v_j, w\right)\sigma\left(f(v_i), w_j\right)\Big)\sigma\left(w_i, v\right) \\
&\quad -\sigma\left(v_i, v\right)\Big(\sigma\left(w_j, w\right)\sigma\left(f(w_i), v_j\right) - \sigma\left(v_j, w\right)\sigma\left(f(w_i), w_j\right)\Big) \\
&= \sum_{i,j=1}^{n} \sigma\left(v_j, w\right)\Big(\sigma\left(v_i, v\right)\sigma\left(f(w_i), w_j\right) - \sigma\left(w_i, v\right)\sigma\left(f(v_i), w_j\right)\Big) \\
&\quad -\Big(\sigma\left(v_i, v\right)\sigma\left(f(w_i), v_j\right) - \sigma\left(w_i, v\right)\sigma\left(f(v_i), v_j\right)\Big)\sigma\left(w_j, w\right) \\
&= \sum_{i,j=1}^{n} \sigma\left(v_j, w\right)\Big(\sigma\left(w_i, v\right)\sigma\left(w_j, f(v_i)\right) - \sigma\left(v_i, v\right)\sigma\left(w_j, f(w_i)\right)\Big) \\
&\quad -\Big(\sigma\left(w_i, v\right)\sigma\left(v_j, f(v_i)\right) - \sigma\left(v_i, v\right)\sigma\left(v_j, f(w_i)\right)\Big)\sigma\left(w_j, w\right) \\
&= \sum_{j=1}^{n} \sigma\left(v_j, w\right)\sigma\left(w_j, f(v)\right) - \sigma\left(v_j, f(v)\right)\sigma\left(w_j, w\right) \\
&= \sigma\left(f(v), w\right)
\end{aligned}
$$

Por otro lado si, $h \in \mathrm{Hom}_{\mathbb{K}}(V, V)$ es adjunto de f, por Propiedad 5.12,

$$
\begin{aligned}
h(v_j) &= \sum_{i=1}^{n} \sigma\left(w_i, h(v_j)\right) v_i - \sigma\left(v_i, h(v_j)\right) w_i \\
&= \sum_{i=1}^{n} \sigma\left(f(w_i), v_j\right) v_i - \sigma\left(f(v_i), v_j\right) w_i \\
&= g(u_j), \\
h(w_j) &= \sum_{i=1}^{n} \sigma\left(w_i, h(w_j)\right) v_i - \sigma\left(v_i, h(w_j)\right) w_i \\
&= \sum_{i=1}^{n} \sigma\left(f(w_i), w_j\right) v_i - \sigma\left(f(v_i), w_j\right) w_i \\
&= g(w_j).
\end{aligned}
$$

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 $i, j = 1, \ldots, n$

$$\begin{aligned}
\Big[g\Big]^T_{T,(i,j)} &= \Big[g(v_j)\Big]^T_i \\
&= \sigma\left(w_i, g(v_j)\right) \\
&= \sigma\left(f(w_i), v_j\right) \\
&= -\sigma\left(v_j, f(w_i)\right) \\
&= \Big[f(w_i)\Big]^T_{n+j} \\
&= \Big[f\Big]^T_{T,(n+j,n+i)}, \\
\Big[g\Big]^T_{T,(n+i,n+j)} &= \Big[g(w_j)\Big]^T_{n+i} \\
&= -\sigma\left(v_i, g(w_j)\right) \\
&= -\sigma\left(f(v_i), w_j\right) \\
&= \sigma\left(w_j, f(v_i)\right) \\
&= \Big[f(v_i)\Big]^T_j \\
&= \Big[f\Big]^T_{T,(j,i)}, \\
\Big[g\Big]^T_{T,(n+i,j)} &= \Big[g(v_j)\Big]^T_{n+i} \\
&= -\sigma\left(v_i, g(v_j)\right) \\
&= -\sigma\left(f(v_i), v_j\right) \\
&= \sigma\left(v_j, f(v_i)\right) \\
&= -\Big[f(v_i)\Big]^T_{n+j} \\
&= -\Big[f\Big]^T_{T,(n+j,i)}, \\
\Big[g\Big]^T_{T,(i,n+j)} &= \Big[g(w_j)\Big]^T_i \\
&= \sigma\left(w_i, g(w_j)\right) \\
&= \sigma\left(f(w_i), w_j\right) \\
&= -\sigma\left(w_j, f(w_i)\right) \\
&= -\Big[f(w_i)\Big]^T_j \\
&= -\Big[f\Big]^T_{T,(j,n+i)}.
\end{aligned}$$

□

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 $v, w \in V$,

$$\begin{aligned} f^*\left(s(v))(w)\right) &= s(v)\left(f(w)\right) \\ &= \sigma(v, f(w)) \\ &= \sigma(f^*(v), w) \\ &= s\left(f^*(v)\right)(w) \end{aligned}$$

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 hecha 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 sympléctica.*

Dem. Proposión 6.30 implica que si f es auto-adjunta, su representación matricial respecto a una base de Darboux es auto-adjunta sympléctica. Para establecer el converso, tomamos una base de Darboux $T = \{v_1, \ldots, v_n, w_1, \ldots, w_n\}$ asumimos que $\left[f\right]_T^T$ es auto-adjunta simpléctica, es decir para todo $i, j \in \{1, \ldots, n\}$

$$\begin{array}{rcl}
\sigma\left(w_i, f(v_j)\right) & = & \left[f\right]_{T,(i,j)}^T \\
& = & \left[f\right]_{T,(n+j,n+i)}^T \\
& = & -\sigma\left(v_j, f(w_i)\right) \\
-\sigma\left(v_i, f(w_j)\right) & = & \left[f\right]_{T,(n+i,n+j)}^T \\
& = & \left[f\right]_{T,(j,i)}^T \\
& = & \sigma\left(w_j, f(v_i)\right) \\
-\sigma\left(v_i, f(v_j)\right) & = & \left[f\right]_{T,(n+i,j)}^T \\
& = & -\left[f\right]_{T,(n+j,i)}^T \\
& = & \sigma\left(v_j, f(v_i)\right) \\
\sigma\left(w_i, f(w_j)\right) & = & \left[f\right]_{T,(i,n+j)}^T \\
& = & -\left[f\right]_{T,(j,n+i)}^T \\
& = & -\sigma\left(w_j, f(w_i)\right),
\end{array}$$

en particular

$$\begin{array}{rcl}
\sigma\left(v_i, f(v_j)\right) & = & \sigma\left(f(v_i), v_j\right) \\
\sigma\left(v_i, f(w_j)\right) & = & \sigma\left(f(v_i), w_j\right) \\
\sigma\left(w_i, f(v_j)\right) & = & \sigma\left(f(w_i), v_j\right) \\
\sigma\left(w_i, f(w_j)\right) & = & \sigma\left(f(w_i), w_j\right).
\end{array}$$

Así, por la demostración de Proposición 6.30, estás igualdades implican que el adjunto de f es él mismo. $\square$

Ejemplo 6.35. Suponga que $V = U \times U^*$ y

$$\sigma\left((v, \lambda), (w, \mu)\right) = \lambda(w) - \mu(v).$$

Sea ahora $g \in \mathrm{Hom}_K(U, U)$ y tome

$$f(v, \lambda) = (g(v), g^*(\lambda))$$

de forma que

$$\begin{array}{rcl}
\sigma\left((v, \lambda), f(w, \mu)\right) & = & \lambda(g(w)) - g^*(\mu)(v) \\
& = & g^*(\lambda)(w) - \mu(g(v)) \\
& = & \sigma\left(f(v, \lambda), (w, \mu)\right),
\end{array}$$

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
$$\begin{aligned} \sigma\big(v, P(f)(w)\big) &= \sum_{i=0}^d a_i \sigma\big(v, f^i(w)\big) \\ &= \sum_{i=1}^d a_i \sigma\big(f^i(v), w\big) \\ &= \sigma\big(P(f)(v), w\big). \end{aligned}$$

□

Proposición 6.38. *Suponga que V tiene dimensión finita, que f es auto-adjunto y que*
$$P_f(t) = (t-\lambda_1)^{m_1}(t-\lambda_2)^{m_2} \ldots (t-\lambda_r)^{m_r}, \quad \lambda_1, \lambda_2, \ldots, \lambda_r \in \mathbb{K}.$$
con $\lambda_i \neq \lambda_j$ si $i \neq j$. Entonces para $i = 1, \ldots, r$, $V_i = \ker(P_i(f)^{m_i})$ es un subespacio simpléctico invariante bajo f, y
$$V = V_1 \oplus \ldots \oplus V_r.$$

Dem. La descomposición $V = V_1 \oplus \ldots \oplus V_r$ como suma directa de subespacios invariantes bajo f es consecuencia directa de Propiedad 2.24, 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, $i = 1, \ldots, r$, 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 $i = 1 \ldots k$ tenemos

$$T(v_1, \ldots, c_i v_i + c_i' v_i', \ldots, v_k) = c_i T(v_1, \ldots, v_i, \ldots, v_k) + c_i' T(v_1, \ldots, v_i', \ldots, v_k)$$

para todo $v_1, \ldots, v_i, v_i', \ldots, v_k \in V$ 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:

$$\begin{array}{rccl} S + T : & V^k & \longrightarrow & K \\ & (v_1, \ldots, v_k) & \longmapsto & S(v_1, \ldots, v_k) + T(v_1, \ldots, v_k) \end{array}$$

$$\begin{array}{rccl} cT : & V^k & \longrightarrow & K \\ & (v_1, \ldots, v_k) & \longmapsto & cT(v_1, \ldots, v_k) \end{array}$$

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
$$\begin{array}{rccl} \langle \bullet ; \bullet \rangle : & K^n \times K^n & \longrightarrow & K \\ & \Big((x_1, \dots, x_n), (y_1, \dots, y_n)\Big) & \longmapsto & \sum_{i=1}^n x_i y_i \end{array}$$
define un 2-tensor.

2. Sea $a_{ij} \in K$, $i, j = 1, \dots, n$. La función
$$\begin{array}{rccl} T : & K^n \times K^n & \longrightarrow & K \\ & \Big((x_1, \dots, x_n), (y_1, \dots, y_n)\Big) & \longmapsto & \sum_{i,j=1}^n x_i a_{ij} y_j \end{array}$$
define un 2-tensor.

3. La función
$$\begin{array}{rccl} \det : & \Big(K^n\Big)^n & \longrightarrow & K \\ & \Big(\overline{x}_1, \dots, \overline{x}_n\Big) & \longmapsto & \det(x_{ij}) = \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) x_{\sigma(1)1} \dots x_{\sigma(n)n} \end{array}$$
donde $\overline{x}_j = (x_{1j}, \dots, x_{nj})$ 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 $S \otimes T$ 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*
$$\Big\{\lambda_{i_1} \otimes \dots \otimes \lambda_{i_k}\Big\}_{i_1, \dots, i_k = 1}^n$$
es una base de $T^k(V)$. *En particular*
$$\dim_K(T^k) = n^k.$$

Dem. Veamos primero que

$$T^k(V) = \langle \lambda_{i_1} \otimes \cdots \otimes \lambda_{i_k} \rangle_{i_1,\ldots,i_k=1}^n.$$

Note que para todo $i_1 \ldots, i_k, j_1, \ldots, j_k \in \{1, \ldots, n\}$

$$\lambda_{i_1} \otimes \cdots \otimes \lambda_{i_k}(v_{j_1}, \ldots, v_{j_k}) = \delta_{i_1 j_1} \cdots \delta_{i_k j_k} = \begin{cases} 1 & \text{si } i_1 = j_1, \ldots, i_k = j_k \\ 0 & \text{si no} \end{cases}.$$

Así, si $w_1, \ldots, w_k \in V$ y $w_j = \sum_{i=1}^n a_{ij} v_i$, $a_{ij} \in K$, $j = 1, \ldots, k$ entonces

$$\begin{aligned} \lambda_{i_1} \otimes \cdots \otimes \lambda_{i_k}(w_1, \ldots, w_k) &= \sum_{j_1,\ldots,j_k=1}^n a_{j_1 1} \cdots a_{j_k k} \lambda_{i_1} \otimes \cdots \otimes \lambda_{i_k}(v_{j_1}, \ldots, v_{j_k}) \\ &= \sum_{j_1,\ldots,j_k=1}^n a_{j_1 1} \cdots a_{j_k k} \delta_{i_1 j_1} \cdots \delta_{i_k j_k} \\ &= a_{i_1 1} \cdots a_{i_k k} \end{aligned}$$

Ahora, dado $T \in T^k(V)$

$$\begin{aligned} T(w_1, \ldots, w_k) &= \sum_{i_1,\ldots,i_k=1}^n a_{i_1 1} \cdots a_{i_k k} T(v_{i_1}, \ldots, v_{i_k}) \\ &= \sum_{i_1,\ldots,i_k=1}^n T(v_{i_1}, \ldots, v_{i_k}) \lambda_{i_1} \otimes \cdots \otimes \lambda_{i_k}(w_1, \ldots, w_k) \end{aligned}$$

luego

$$T = \sum_{i_1,\ldots,i_k=1}^n T(v_{i_1}, \ldots, v_{i_k}) \lambda_{i_1} \otimes \cdots \otimes \lambda_{i_k},$$

y $T^k(V) = \langle \lambda_{i_1} \otimes \cdots \otimes \lambda_{i_k} \rangle_{i_1,\ldots,i_k=1}^n$.

Establezcamos ahora la independencia lineal. Suponga que

$$0 = \sum_{i_1,\ldots,i_k=1}^n c_{i_1 \ldots i_k} \lambda_{i_1} \otimes \cdots \otimes \lambda_{i_k}.$$

Si evaluamos en $(v_{j_1}, \ldots, v_{j_k})$ obtenemos

$$\begin{aligned} 0 &= \sum_{i_1,\ldots,i_k=1}^n c_{i_1 \ldots i_k} \lambda_{i_1} \otimes \cdots \otimes \lambda_{i_k}(v_{j_1}, \ldots, v_{j_k}) \\ &= c_{j_1,\ldots,j_k}. \end{aligned}$$

□

Ejemplo 7.11. Sea $n \in \mathbb{Z}_{>0}$, $\{e_i\}_{i=1}^n$ la base canónica de K^n y $\{f_1, \ldots, f_n\}$ la base dual.

1. Si
$$\begin{array}{rccl} \langle \bullet ; \bullet \rangle : & K^n \times K^n & \longrightarrow & K \\ & \Big((x_1, \dots, x_n), (y_1, \dots, y_n)\Big) & \longmapsto & \sum_{i=1}^n x_i y_i \end{array}$$
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
$$\begin{array}{rccl} T : & K^n \times K^n & \longrightarrow & K \\ & \Big((x_1, \dots, x_n), (y_1, \dots, y_n)\Big) & \longmapsto & \sum_{i,j=1}^n x_i a_{ij} y_j \end{array}$$
entonces $T = \sum_{i,j=1}^n a_{ij} f_i \otimes f_j$.
3. Si
$$\begin{array}{rccl} \det : & \Big(K^n\Big)^n & \longrightarrow & K \\ & \big(\overline{x}_1, \dots, \overline{x}_n\big) & \longmapsto & \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) x_{\sigma(1)1} \dots x_{\sigma(n)n} \end{array}$$
donde $\overline{x}_j = (x_{1j}, \dots, x_{nj})$ para $j = 1, \dots, n$ entonces
$$\det = \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) f_{\sigma(1)} \otimes \dots \otimes f_{\sigma(n)}.$$

Definición 7.12. Sea $f \in \operatorname{Hom}_K(V, W)$. Dado $k \in \mathbb{Z}_{\geq 0}$ definimos
$$\begin{array}{rcl} f^* : T^k(W) & \longrightarrow & T^k(V) \\ T & \longmapsto & f^*T : (v_1, \dots, v_k) \mapsto T\big(f(v_1), \dots, f(v_k)\big). \end{array}$$

Proposición 7.13. *Sea $f \in \operatorname{Hom}_K(V, W)$ 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. $\square$

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.

$$\begin{aligned}
0 =& \omega(v_1, \ldots, v_i + v_j, \ldots, v_i + v_j, \ldots, v_k) \\
=& \omega(v_1, \ldots, v_i, \ldots, v_i, \ldots, v_k) + \omega(v_1, \ldots, v_i, \ldots, v_j, \ldots, v_k) \\
& + \omega(v_1, \ldots, v_j, \ldots, v_i, \ldots, v_k) + \omega(v_1, \ldots, v_j, \ldots, v_j, \ldots, v_k) \\
=& \omega(v_1, \ldots, v_i, \ldots, v_j, \ldots, v_k) + \omega(v_1, \ldots, v_j, \ldots, v_i, \ldots, v_k)
\end{aligned}$$

□

Observación 7.16. Sea $k \in \mathbb{Z}_{\geq 0}$ y $\omega \in \Lambda^k(V)$, para toda permutación $\sigma \in S_k$ tenemos

$$\omega(v_{\sigma(1)}, \ldots, v_{\sigma(k)}) = -\operatorname{sgn}(\sigma)\omega(v_1, \ldots, v_k).$$

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

$$\begin{array}{rccl}
\det: & \left(K^n\right)^n & \longrightarrow & K \\
& (\overline{x}_1, \ldots, \overline{x}_n) & \longmapsto & \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) x_{\sigma(1)1} \ldots x_{\sigma(n)n}
\end{array}$$

   donde $\overline{x}_j = (x_{1j}, \ldots, x_{nj})$ para $j = 1, \ldots, n$ define un n-tensor alternante.

2. Sea $k \in \{1, \ldots, n\}$ y $1 \leq i_1 < \ldots < i_k \leq n$. La función determinante menor

$$\begin{array}{rccl}
\det_{i_1 \ldots i_k}: & \left(K^n\right)^k & \longrightarrow & K \\
& (\overline{x}_1, \ldots, \overline{x}_k) & \longmapsto & \sum_{\sigma \in S_k} \operatorname{sgn}(\sigma) x_{i_{\sigma(1)}1} \ldots x_{i_{\sigma(k)}k}
\end{array}$$

   donde $\overline{x}_j = (x_{1j}, \ldots, x_{nj})$ para $j = 1, \ldots, n$ 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 $\operatorname{Alt} \in \operatorname{Hom}_K\left(T^k(V), T^k(V)\right)$ por

$$\operatorname{Alt}(T)(v_1, \ldots, v_k) = \frac{1}{k!} \sum_{\sigma \in S_k} \operatorname{sgn}(\sigma) T(v_{\sigma(1)}, \ldots, v_{\sigma(k)})$$

Ejemplo 7.21. Sea $n \in \mathbb{Z}_{>0}$, $\{e_i\}_{i=1}^n$ la base canónica de K^n y $\{f_1, \ldots, f_n\}$ la base dual.

1. Para $\mathrm{char}(K) > 0$ suponga que $n < \mathrm{char}(K)$. Si $T = f_1 \otimes \cdots \otimes f_n$, tenemos que para todo $\overline{x}_j = (x_{1j}, \ldots, x_{nj}) \in K^n$, $j = 1, \ldots, n$,

$$\begin{aligned}
\mathrm{Alt}(T)(\overline{x}_1, \ldots, \overline{x}_n) &= \frac{1}{n!} \sum_{\sigma \in S_n} \mathrm{sgn}(\sigma) f_1 \otimes \cdots \otimes f_n(v_{\sigma(1)}, \ldots, v_{\sigma(n)}) \\
&= \frac{1}{n!} \sum_{\sigma \in S_n} \mathrm{sgn}(\sigma) x_{1\sigma(1)} \cdots x_{n\sigma(n)} \\
&= \frac{1}{n!} \sum_{\sigma \in S_n} \mathrm{sgn}(\sigma) x_{\sigma^{-1}(1)1} \cdots x_{\sigma^{-1}(n)n} \\
&= \frac{1}{n!} \sum_{\sigma \in S_n} \mathrm{sgn}(\sigma^{-1}) x_{\sigma^{-1}(1)1} \cdots x_{\sigma^{-1}(n)n} \\
&= \frac{1}{n!} \sum_{\sigma \in S_n} \mathrm{sgn}(\sigma) x_{\sigma(1)1} \cdots x_{\sigma(n)n} \\
&= \frac{1}{n!} \det\big(\overline{x}_1, \ldots, \overline{x}_n\big)
\end{aligned}$$

así $\mathrm{Alt}(f_1 \otimes \cdots \otimes f_n) = \frac{1}{n!} \det$.

2. Sea $k \in \{1, \ldots, n\}$ y $1 \leq i_1 < \ldots < i_k \leq n$. Para $\mathrm{char}(K) > 0$ suponga que $k < \mathrm{char}(K)$. Si $T = f_{i_1} \otimes \cdots \otimes f_{i_k}$, tenemos que para todo $\overline{x}_j = (x_{1j}, \ldots, x_{nj}) \in K^n$, $j = 1, \ldots, k$,

$$\begin{aligned}
\mathrm{Alt}(T)(\overline{x}_1, \ldots, \overline{x}_k) &= \frac{1}{k!} \sum_{\sigma \in S_k} \mathrm{sgn}(\sigma) f_{i_1} \otimes \cdots \otimes f_{i_k}(v_{\sigma(1)}, \ldots, v_{\sigma(k)}) \\
&= \frac{1}{k!} \sum_{\sigma \in S_k} \mathrm{sgn}(\sigma) x_{i_1\sigma(1)} \cdots x_{i_k\sigma(k)} \\
&= \frac{1}{k!} \sum_{\sigma \in S_k} \mathrm{sgn}(\sigma) x_{i_{\sigma^{-1}(1)}1} \cdots x_{i_{\sigma^{-1}(k)}k} \\
&= \frac{1}{k!} \sum_{\sigma \in S_k} \mathrm{sgn}(\sigma^{-1}) x_{i_{\sigma^{-1}(1)}1} \cdots x_{i_{\sigma^{-1}(k)}k} \\
&= \frac{1}{k!} \sum_{\sigma \in S_k} \mathrm{sgn}(\sigma) x_{i_{\sigma(1)}1} \cdots x_{i_{\sigma(k)}k} \\
&= \frac{1}{k!} \det{}_{i_1 \ldots i_k}\big(\overline{x}_1, \ldots, \overline{x}_k\big)
\end{aligned}$$

así $\mathrm{Alt}(f_{i_1} \otimes \cdots \otimes f_{i_k}) = \frac{1}{k!} \det_{i_1 \ldots i_k}$.

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* $T \in T^k(V)$, $\mathrm{Alt}(T) \in \Lambda^k(V)$,
2. *para todo* $\omega \in \Lambda^k(V)$, $\mathrm{Alt}(\omega) = \omega$,

3. $\mathrm{Alt} \circ \mathrm{Alt} = \mathrm{Alt}$.

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 $\{1, \ldots, k\}$ fijos). Sea A_k el subgrupo de S_k formado por las permutaciones con signo 1, de forma que si $\tau A_k = \{\tau\sigma \in S_k \| \ \sigma \in A_k\}$ entonces $S_k = A_k \cup \tau A_k$ es una partición de S_k. Nota que $\mathrm{sgn}(\tau\sigma) = -\mathrm{sgn}(\sigma)$ para todo $\sigma \in S_k$. Sean $v_1, \ldots, v_k \in V$ tales que $v_i = v_j$, entonces para todo $\sigma \in S_k$ tenemos

$$(v_{\sigma(1)}, \ldots, v_{\sigma(k)}) = (v_{\tau\sigma(1)}, \ldots, v_{\tau\sigma(k)})$$

y así

$$\begin{aligned}
&\mathrm{Alt}(T)(v_1, \ldots, v_k) \\
&= \frac{1}{k!} \sum_{\sigma \in S_k} \mathrm{sgn}(\sigma) T(v_{\sigma(1)} \ldots, v_{\sigma(k)}) \\
&= \frac{1}{k!} \left(\sum_{\sigma \in A_k} \mathrm{sgn}(\sigma) T(v_{\sigma(1)}, \ldots, v_{\sigma(k)}) + \sum_{\sigma \in \tau A_k} \mathrm{sgn}(\sigma) T(v_{\sigma(1)}, \ldots, v_{\sigma(k)}) \right) \\
&= \frac{1}{k!} \left(\sum_{\sigma \in A_k} \mathrm{sgn}(\sigma) T(v_{\sigma(1)}, \ldots, v_{\sigma(k)}) + \sum_{\sigma \in A_k} \mathrm{sgn}(\tau\sigma) T(v_{\tau\sigma(1)}, \ldots, v_{\tau\sigma(k)}) \right) \\
&= \frac{1}{k!} \left(\sum_{\sigma \in A_k} \mathrm{sgn}(\sigma) T(v_{\sigma(1)}, \ldots, v_{\sigma(k)}) - \sum_{\sigma \in A_k} \mathrm{sgn}(\sigma) T(v_{\tau\sigma(1)}, \ldots, v_{\tau\sigma(k)}) \right) \\
&= 0
\end{aligned}$$

luego $\mathrm{Alt}(T) \in \Lambda^k(V)$.

2. Sea $\omega \in \Lambda^k(V)$, entonces para todo $v_1, \ldots, v_k \in V$ tenemos que

$$\begin{aligned}
\mathrm{Alt}(\omega)(v_1, \ldots, v_k) &= \frac{1}{k!} \sum_{\sigma \in S_k} \mathrm{sgn}(\sigma) \omega(v_{\sigma(1)}, \ldots, v_{\sigma(k)}) \\
&= \frac{1}{k!} \sum_{\sigma \in S_k} \mathrm{sgn}(\sigma) \, \mathrm{sgn}(\sigma) \omega(v_1, \ldots, v_k) \\
&= \frac{1}{k!} \sum_{\sigma \in S_k} \omega(v_1, \ldots, v_k) \\
&= \frac{1}{k!} k! \ \omega(v_1, \ldots, v_k) \\
&= \omega(v_1, \ldots, v_k)
\end{aligned}$$

luego $\mathrm{Alt}(\omega) = \omega$.

3. Se sigue inmediatamente de 1. y 2.

□

Definición 7.23. Sea $k, l \in \mathbb{Z}_{\geq 0}$ tal que si $\operatorname{char}(K) > 0$ entonces $k + l < \operatorname{char}(K)$. 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

$$\omega \wedge \eta = \frac{(k+l)!}{k!\ l!} \operatorname{Alt}(\omega \otimes \eta).$$

Lema 7.24. *Sea $k, l, m \in \mathbb{Z}_{\geq 0}$ tal que si $\operatorname{char}(K) > 0$ entonces $k + l + m < \operatorname{char}(K)$. Sea $S \in T^k(V)$, $T \in T^l(V)$ y $U \in T^m(V)$ tenemos que*

1. *si* $\operatorname{Alt}(S) = 0$ *entonces* $\operatorname{Alt}(S \otimes T) = 0 = \operatorname{Alt}(T \otimes S)$,
2. $\operatorname{Alt}(\operatorname{Alt}(S \otimes T) \otimes U) = \operatorname{Alt}(S \otimes T \otimes U) = \operatorname{Alt}(S \otimes \operatorname{Alt}(T \otimes U))$

Dem.

1. Tome S_k como el subgrupo de S_{k+l} formado por la permutaciones que dejan fijos a $k+1, \ldots, k+l$ y sean $\sigma_1, \ldots, \sigma_N \in S_{k+l}$, $N = (k+l)!/(k!)$, representantes de los coconjuntos $S_k\sigma = \{\tau\sigma \in S_{k+l} | \ \tau \in S_k\}$, $\sigma \in S_{k+l}$, de forma que
$$S_{k+l} = \cup_{i=1}^N \sigma_i S_k$$
es una partición de S_{k+l}. Entonces para todo $v_1, \ldots, v_{k+l} \in V$
$$\begin{aligned}
&\operatorname{Alt}(S \otimes T)(v_1, \ldots, v_{k+l}) \\
&= \frac{1}{(k+l)!} \sum_{\sigma \in S_{k+l}} \operatorname{sgn}(\sigma) S \otimes T(v_{\sigma(1)}, \ldots, v_{\sigma(k+l)}) \\
&= \frac{1}{(k+l)!} \sum_{i=1}^N \sum_{\tau \in S_k} \operatorname{sgn}(\tau\sigma_i) S \otimes T(v_{\tau\sigma_i(1)}, \ldots, v_{\tau\sigma_i(k+l)}) \\
&= \frac{1}{(k+l)!} \sum_{i=1}^N \operatorname{sgn}(\sigma_i) \sum_{\tau \in S_k} \operatorname{sgn}(\tau) S(v_{\tau\sigma_i(1)}, \ldots, v_{\tau\sigma_i(k)}) T(v_{\tau\sigma_i(k+1)}, \ldots, v_{\tau\sigma_i(k+l)}).
\end{aligned}$$
Para $i = 1, \ldots, N$ y $j = 1, \ldots, k+l$, denote $w_j^{(i)} = v_{\sigma_i(j)}$, así, como $\operatorname{Alt} = 0$ entonces
$$\begin{aligned}
&\operatorname{Alt}(S \otimes T)(v_1, \ldots, v_{k+l}) \\
&= \frac{1}{(k+l)!} \sum_{i=1}^N \operatorname{sgn}(\sigma_i) \sum_{\tau \in S_k} \operatorname{sgn}(\tau) S(w_{\tau(1)}^{(i)}, \ldots, w_{\tau(k)}^{(i)}) T(w_{\tau(k+1)}^{(i)}, \ldots, w_{\tau(k+l)}^{(i)}) \\
&= \frac{1}{(k+l)!} \sum_{i=1}^N \operatorname{sgn}(\sigma_i) \sum_{\tau \in S_k} \operatorname{sgn}(\tau) S(w_{\tau(1)}^{(i)}, \ldots, w_{\tau(k)}^{(i)}) T(w_{k+1}^{(i)}, \ldots, w_{k+l}^{(i)}) \\
&= \frac{1}{(k+l)!} \sum_{i=1}^N \operatorname{sgn}(\sigma_i) \left(\sum_{\tau \in S_k} \operatorname{sgn}(\tau) S(w_{\tau(1)}^{(i)}, \ldots, w_{\tau(k)}^{(i)}) \right) T(w_{k+1}^{(i)}, \ldots, w_{k+l}^{(i)}) \\
&= \frac{k!}{(k+l)!} \sum_{i=1}^N \operatorname{sgn}(\sigma_i) \operatorname{Alt}(S)(w_1^{(i)}, \ldots, w_k^{(i)}) T(w_{k+1}^{(i)}, \ldots, w_{k+l}^{(i)}) \\
&= 0
\end{aligned}$$

Similarmente, si tomamos S_k como el subgrupo de S_{k+l} formado por la permutaciones que dejan fijos a $1, \ldots, l$, obtenemos $\mathrm{Alt}(T \otimes S) = 0$.

2. $\mathrm{Alt}\left(\mathrm{Alt}(S \otimes T) - S \otimes T\right) = \mathrm{Alt}(S \otimes T) - \mathrm{Alt}(S \otimes T) = 0$, luego por 1.

$$\begin{aligned} 0 &= \mathrm{Alt}\left((\mathrm{Alt}(S \otimes T) - S \otimes T) \otimes U\right) \\ &= \mathrm{Alt}\left(\mathrm{Alt}(S \otimes T) \otimes U\right) - \mathrm{Alt}(S \otimes T \otimes U) \end{aligned}$$

y así $\mathrm{Alt}\left(\mathrm{Alt}(S \otimes T) \otimes U\right) = \mathrm{Alt}(S \otimes T \otimes U)$. Similarmente obtenemos $\mathrm{Alt}(S \otimes \mathrm{Alt}(T \otimes U)) = \mathrm{Alt}(S \otimes T \otimes U)$.

□

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 $\mathrm{char}(K) > 0$ entonces $k+l+m < \mathrm{char}(K)$ 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. $(\omega \wedge \eta) \wedge \theta = \omega \wedge (\eta \wedge \theta) = \frac{(k+l+m)!}{k!\ l!\ m!} \mathrm{Alt}(\omega \otimes \eta \otimes \theta)$.

Dem. Las propiedades 1., 2. y 3. se siguen inmediatamente de la bilinearidad de $\otimes$ y de la linearidad 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 $\mathrm{sgn}(\tau) = (-1)^{kl}$ y $\mathrm{sgn}(\sigma\tau) = (-1)^{kl} \,\mathrm{sgn}(\sigma)$ para toda $\sigma \in S_{k+l}$

entonces para todo $v_1, \ldots, v_{k+l} \in V$

$$
\begin{aligned}
&\omega \wedge \eta(v_1, \ldots, v_{k+l}) \\
&= \frac{(k+l)!}{k!\ l!} \operatorname{Alt}(\omega \otimes \eta)(v_1, \ldots, v_{k+l}) \\
&= \frac{1}{k!\ l!} \sum_{\sigma \in S_{k+l}} \operatorname{sgn}(\sigma) \omega \otimes \eta(v_{\sigma(1)}, \ldots, v_{\sigma(k)}, v_{\sigma(k+1)}, \ldots, v_{\sigma(k+l)}) \\
&= \frac{(1}{k!\ l!} \sum_{\sigma \in S_{k+l}} \operatorname{sgn}(\sigma) \omega(v_{\sigma(1)}, \ldots, v_{\sigma(k)}) \eta(v_{\sigma(k+1)}, \ldots, v_{\sigma(k+l)}) \\
&= \frac{1}{k!\ l!} \sum_{\sigma \in S_{k+l}} \operatorname{sgn}(\sigma) \omega(v_{\sigma\tau(l+1)}, \ldots, v_{\sigma\tau(l+k)}) \eta(v_{\sigma\tau(1)}, \ldots, v_{\sigma\tau(l)}) \\
&= \frac{1}{k!\ l!} \sum_{\sigma \in S_{k+l}} (-1)^{kl} \operatorname{sgn}(\sigma\tau) \eta(v_{\sigma\tau(1)}, \ldots, v_{\sigma\tau(l)}) \omega(v_{\sigma\tau(l+1)}, \ldots, v_{\sigma\tau(l+k)}) \\
&= (-1)^{kl} \frac{1}{k!\ l!} \sum_{\sigma\tau \in S_{k+l}} \operatorname{sgn}(\sigma\tau) \eta(v_{\sigma\tau(1)}, \ldots, v_{\sigma\tau(l)}) \omega(v_{\sigma\tau(l+1)}, \ldots, v_{\sigma\tau(l+k)}) \\
&= (-1)^{kl} \frac{1}{k!\ l!} \sum_{\sigma\tau \in S_{k+l}} \operatorname{sgn}(\sigma\tau) \eta \otimes \omega(v_{\sigma\tau(1)}, \ldots, v_{\sigma\tau(k)}, v_{\sigma\tau(k+1)}, \ldots, v_{\sigma\tau(k+l)}) \\
&= (-1)^{kl} \frac{(k+l)!}{k!\ l!} \operatorname{Alt}(\eta \otimes \omega)(v_1, \ldots, v_{k+l}) \\
&= (-1)^{kl} \eta \wedge \omega(v_1, \ldots, v_{k+l}).
\end{aligned}
$$

así $\omega \wedge \eta = (-1)^{kl} \eta \wedge \omega$.

5. Por Lema 7.24

$$
\begin{aligned}
(\omega \wedge \eta) \wedge \theta &= \frac{(k+l+m)!}{(k+l)!\ m!} \operatorname{Alt}\left((\omega \wedge \eta) \otimes \theta\right) \\
&= \frac{(k+l+m)!}{(k+l)!\ m!} \operatorname{Alt}\left(\frac{(k+l)!}{k!\ l!} \operatorname{Alt}(\omega \otimes \eta) \otimes \theta\right) \\
&= \frac{(k+l+m)!}{k!\ l!\ m!} \operatorname{Alt}(\omega \otimes \eta \otimes \theta)
\end{aligned}
$$

Similarmente obtenemos $\omega \wedge (\eta \wedge \theta) = \frac{(k+l+m)!}{k!\ l!\ m!} \operatorname{Alt}(\omega \otimes \eta \otimes \theta)$.

$\square$

Observación 7.26. Suponga que $\operatorname{char}(K) \neq 2$ 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 $\mathrm{char}(K) > 0$ entonces $k < \mathrm{char}(K)$,

$$\lambda_1 \wedge \ldots \wedge \lambda_k = \sum_{\sigma \in S_k} \mathrm{sgn}(\sigma) \lambda_{\sigma(1)} \otimes \cdots \otimes \lambda_{\sigma(k)}.$$

Proposición 7.27. *Sea $f \in \mathrm{Hom}_K(V, W)$ 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. □

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 $\mathrm{char}(K) > 0$ entonces $k_1 + \ldots + k_r < \mathrm{char}(K)$ y $\omega_i \in \Lambda^{k_i}(V)$, $i = 1, \dots, r$, entonces*

$$\omega_1 \wedge \cdots \wedge \omega_r = \frac{(k_1 + \ldots + k_r)!}{k_1! \cdots k_r!} \mathrm{Alt}(\omega_1 \otimes \cdots \otimes \omega_r).$$

Dem. Se sigue de inmediatamente de Propiedad 7.25 5. por inducción en r. □

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 $\mathrm{char}(K) > 0$ suponga que $n < \mathrm{char}(K)$. Como $\mathrm{Alt}(f_1 \otimes \cdots \otimes f_n) = \frac{1}{n!} \det$ entonces
$$f_1 \wedge \cdots \wedge f_n = \det.$$

2. Sea $k \in \{1, \dots, n\}$ y $1 \leq i_1 < \ldots < i_k \leq n$. Para $\mathrm{char}(K) > 0$ suponga que $k < \mathrm{char}(K)$. Como $\mathrm{Alt}(f_{i_1} \otimes \cdots \otimes f_{i_k}) = \frac{1}{k!} \det_{i_1 \ldots i_k}$ entonces
$$f_{i_1} \wedge \cdots \wedge f_{i_k} = \det_{i_1 \ldots i_k}.$$

Teorema 7.31. *Suponga que $\mathrm{char}(K) \neq 2$ 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 $\mathrm{char}(K) > 0$ entonces $k < \mathrm{char}(K)$. La colección de k-tensores*

$$\left\{\lambda_{i_1} \wedge \cdots \wedge \lambda_{i_k}\right\}_{1 \leq i_1 < \ldots < i_k \leq n}$$

es una base de $\Lambda^k(V)$. En particular

$$\dim_K\left(\Lambda^k(V)\right) = \binom{n}{k}.$$

Dem. Sea $k \in \mathbb{Z}_{\geq 0}$ y $\omega \in \Lambda^k(V)$, como $\omega \in T^k(V)$ y $\left\{\lambda_{i_1} \otimes \cdots \otimes \lambda_{i_k}\right\}_{i_1,\ldots,i_k=1}^n$ es un base de $T^k(V)$ entonces

$$\omega = \sum_{j_1,\ldots,j_k=1}^{n} c_{j_1\ldots j_k} \lambda_{j_1} \otimes \cdots \otimes \lambda_{j_k}$$

con $c_{j_1\ldots j_k} \in K$. Así

$$\begin{aligned} \omega &= \mathrm{Alt}(\omega) \\ &= \sum_{j_1,\ldots,j_k=1}^{n} c_{j_1\ldots j_k} \mathrm{Alt}(\lambda_{j_1} \otimes \cdots \otimes \lambda_{j_k}) \\ &= \sum_{j_1,\ldots,j_k=1}^{n} c_{j_1\ldots j_k} k! \lambda_{j_1} \wedge \cdots \wedge \lambda_{j_k}. \end{aligned}$$

Ahora, dados $j_1, \ldots, j_k \in \{1, \ldots, n\}$ por Propiedad 7.25 4. y Observación 7.26, si dos de los subindices j_i coinciden entonces $\lambda_{j_1} \wedge \cdots \wedge \lambda_{j_k} = 0$, luego podemos asumir que son todos distintos; en tal caso $\lambda_{j_1} \wedge \cdots \wedge \lambda_{j_k} = \mathrm{sgn}(\sigma_{j_1,\ldots,j_k})\lambda_{i_1} \wedge \cdots \wedge \lambda_{i_k}$ donde $i_1 < \ldots < i_k$ son los mismos subindices $j_1, \ldots, j_k$ ordenados en forma creciente y $\sigma_{j_1,\ldots,j_k} \in S_k$ es la permutación que reorganiza los $j_1, \ldots, j_k$. Luego

$$\omega = \sum_{1 \leq i_1 < \ldots < i_k \leq n} a_{i_1,\ldots,i_k} \lambda_{i_1} \wedge \cdots \wedge \lambda_{i_k}$$

con $a_{i_1,\ldots,i_k} = k! \sum_{\sigma \in S_k} c_{i_{\sigma(1)}\ldots i_{\sigma(k)}}$, y así $\Lambda^k(V) = \langle \lambda_{i_1} \wedge \cdots \wedge \lambda_{i_k} \rangle_{1 \leq i_1 < \ldots < i_k \leq n}$.

Para establecer la independencia lineal, note que si $1 \leq i_1 < \ldots < i_k \leq n$ y $1 \leq j_1 < \ldots < j_k \leq n$ entonces (ver Ejemplo 7.30 2.)

$$\lambda_{i_1} \wedge \cdots \wedge \lambda_{i_k}(v_{j_1}, \ldots, v_{j_k}) = \begin{cases} 1 & \text{si } i_1 = j_1, \ldots, i_k = j_k \\ 0 & \text{si no} \end{cases}$$

luego si $0 = \sum_{1 \leq i_1 < \ldots < i_k \leq n} a_{i_1,\ldots,i_k} \lambda_{i_1} \wedge \cdots \wedge \lambda_{i_k}$ entonces

$$\begin{aligned} 0 &= \sum_{1 \leq i_1 < \ldots < i_k \leq n} a_{i_1,\ldots,i_k} \lambda_{i_1} \wedge \cdots \wedge \lambda_{i_k}(v_{j_1}, \ldots, v_{j_k}) \\ &= a_{j_1,\ldots,j_k}. \end{aligned}$$

□

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 $T^l_k(V)$.

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

$$\begin{array}{rcl} \widehat{v} : V^* & \longrightarrow & K \\ \lambda & \longmapsto & \lambda(v) \end{array}$$

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

$$\begin{aligned} &S \otimes T(\lambda_1, \ldots, \lambda_{l_1+l_2}, v_1, \ldots, v_{k_1+k_2}) \\ &= S(\lambda_1, \ldots, \lambda_{l_1}, v_1, \ldots, v_{k_1}) T(\lambda_{l_1+1}, \ldots, \lambda_{l_1+l_2}, v_{k_1+1}, \ldots, v_{k_1+k_2}) \end{aligned}$$

para todo $\lambda_1, \ldots, \lambda_{l_1+l_2} \in V^*$, $v_1, \ldots, v_{k_1+k_2} \in V$.

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. □

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* $\{v_1, \ldots, v_n\}$ *una base de* V *y* $\{\lambda_1, \ldots, \lambda_n\}$ *la base dual. Para todo* $l, k \in \mathbb{Z}_{>0}$, *la colección de* (l,k)*-tensores*

$$\left\{ v_{i_1} \otimes \cdots \otimes v_{i_l} \otimes \lambda_{j_1} \otimes \cdots \otimes \lambda_{j_k} \right\}_{i_1,\ldots,i_l,j_1,\ldots,j_k=1}^{n}$$

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 $T \in T^{(l,k)}(V)$

$$T = \sum_{i_1,\ldots,i_l=1}^{n} \sum_{j_1,\ldots,j_k=1}^{n} T(\lambda_{i_1}, \ldots, \lambda_{i_l}, v_{j_1}, \ldots, v_{j_k}) v_{i_1} \otimes \cdots \otimes v_{i_l} \otimes \lambda_{j_1} \otimes \cdots \otimes \lambda_{j_k}.$$

□

Ejemplo 7.38. Sea $\{e_1, \ldots, e_n\}$ la base canónica de K^n y $\{f_1, \ldots, f_n\}$ la base dual.

1. Suponga que $\mathrm{char}(K) \neq 2$. 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 $\{v_1, \ldots, v_n\}$ una base de V y $\{\lambda_1, \ldots, \lambda_n\}$ la base dual. Sea $l, k \in \mathbb{Z}_{\geq 0}$ y

$$t = v_{\alpha_1} \otimes \cdots \otimes v_{\alpha_l} \otimes \lambda_{\beta_1} \otimes \cdots \otimes \lambda_{\beta_k} \in T^{(l,k)}(V).$$

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

$$\begin{array}{rcl} \widehat{t} : T^{(l',k')}(V) & \longrightarrow & T^{(l'-k,k'-l)}(V) \\ T & \longmapsto & T(t) := \widehat{t}(T) \end{array}$$

donde si

$$T = v_{i_1} \otimes \cdots \otimes v_{i_{l'}} \otimes \lambda_{j_1} \otimes \cdots \otimes \lambda_{j_{k'}}$$

entonces

$$T(t) = \left(\prod_{s=1}^{k} \lambda_{\beta_s}(v_{i_s}) \prod_{s=1}^{l} \lambda_{j_s}(v_{\alpha_s}) \right) v_{i_{k+1}} \otimes \cdots \otimes v_{i_{l'}} \otimes \lambda_{j_{l+1}} \otimes \cdots \otimes \lambda_{j_{k'}}.$$

Extendemos linealmente a todo los (l,k)-tensores $t \in T^{(l,k)}(V)$ para obtener

$$\begin{array}{rcl} \widehat{\bullet} : T^{(l,k)}(V) & \longrightarrow & \mathrm{Hom}_K\left(T^{(l',k')}(V), T^{(l'-k,k'-l)}(V)\right) \\ t & \longmapsto & \widehat{t} \end{array}$$

Observación 7.40. Note que

$$\left(\prod_{s=1}^{k} \lambda_{\beta_s}(v_{i_s}) \prod_{s=1}^{l} \lambda_{j_s}(v_{\alpha_s}) \right) = v_{i_1} \otimes \cdots \otimes v_{i_k} \otimes \lambda_{j_1} \otimes \cdots \otimes \lambda_{j_l}(v_{\alpha_1}, \ldots, v_{\alpha_l}, \lambda_{\beta_1}, \ldots, \lambda_{\beta_k}).$$

Ejemplo 7.41. Sea $\{e_1, \ldots, e_n\}$ la base canónica de K^n y $\{f_1, \ldots, f_n\}$ 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
$$\begin{aligned} T_\times(v_1 \otimes v_2) &= (f_2 \wedge f_3)(v_1, v_2)e_1 - (f_1 \wedge f_3)(v_1, v_2)e_2 + (f_1 \wedge f_2)(v_1, v_2)e_3 \\ &= (a_2b_3 - a_3b_2)e_1 - (a_1b_3 - a_3b_1)e_2 + (a_1b_2 - a_2b_1)e_3 \\ &=: v_1 \times v_2 \end{aligned}$$

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 $f_A \in \mathrm{Hom}_K(K^n, K^n)$ 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 $f_A \in \mathrm{Hom}_K(K^n, K^n)$ 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 $T^l_k(V)$.
4. Convención de Einstein: si un indice 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 $T \in T^l_k(V)$ si denotamos

$$T^{i_1 \dots i_l}_{j_1 \dots j_k} = T(\lambda^{i_1}, \dots, \lambda^{i_l}, v_{j_1}, \dots, v_{j_k})$$

entonces

$$T = T^{i_1 \dots i_l}_{j_1 \dots j_k} v_{i_1} \otimes \cdots \otimes v_{i_l} \otimes \lambda^{j_1} \otimes \cdots \otimes \lambda^{j_k}$$

3. Dado $f \in \mathrm{Hom}_K(V, V)$ si denotamos

$$f^i_j = \lambda^i(f(v_j))$$

entonces el tensor $T_f = f^i_j v_i \otimes \lambda^j \in T^1_1(V)$ es tal que

$$\begin{aligned} T_f(v) &= f^i_j \left(v_i \otimes \lambda^j\right)(v) \\ &= f^i_j \lambda^j(v) v_i \\ &= f^i_j v^j v_i \\ &= f(v) \end{aligned}$$

y

$$\begin{aligned} T_f(\lambda) &= f^i_j \left(v_i \otimes \lambda^j\right)(\lambda) \\ &= f^i_j \lambda(v_i)\lambda^j \\ &= f^i_j \lambda_i \lambda^j \\ &= f^*(\lambda) \end{aligned}$$

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 $\mathbb{Q}$, las números reales $\mathbb{R}$, los números complejos $\mathbb{C}$ y las fracciones racionales sobre estos cuerpos, por ejemplo $\mathbb{Q}(x_1, \ldots, x_n)$ 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 $\mathbb{K} \times \mathbb{K} \to \mathbb{K}$), 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 heredades 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 $a = b$, 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.

$\square$

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

$$3x = 1 - 4, \qquad 3x = 1 + 1, \qquad 3x = 2, \qquad x = 2/3, \qquad x = 2 \cdot 2$$

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 $\mathbb{K}$ para el cual

$$\underbrace{1+1+\ldots+1}_{k \text{ sumandos}} = 0,$$

al mínimo entre estos lo llamamos la *característica* de $\mathbb{K}$ y lo denotamos por $\mathrm{char}(K)$. 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

$$a_n t^n + a_{n-1} t^{n-1} + \ldots + a_1 t + a_0$$

con $a_n, \ldots, a_1, a_0 \in \mathbb{K}$. Si $P(t)$ denota este polinomio, dado $c \in \mathbb{K}$, el *valor de* $P(t)$ *en* c es

$$a_n c^n + a_{n-1} c^{n-1} + \ldots + a_1 c + a_0,$$

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 $a_n t^n + \ldots + a_1 t + a_0$ y $b_m t^m + \ldots + b_1 t + b_0$, su suma $P(t) + Q(t)$ es el polinomio $c_k t^k + \ldots + c_1 t + c_0$, con $k = \max n, m$ y $c_i = a_i + b_i$, y su producto $P(t)Q(t)$ es el polinomio $d_l t^l + \ldots + d_1 t + d_0$, 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 $P(t) = a_n t^n + \ldots + a_1 t + a_0$, 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

$$|f(z)| \geq 1 - \sum_{k=0}^{n-1} \left| \frac{a_k}{a_n} \right| \frac{1}{r^{n-k}}.$$

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 Louiville $z \mapsto 1/P(z)$ 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 $a_n t^n + a_{n-1}t^{n-1} + \ldots + a_1 t + a_0$, 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

$$\deg(P(t) + Q(t)) \leq \text{máx}\{\deg(P(t)), \deg(Q(t))\}$$
$$\deg(P(t)Q(t)) = \deg(P(t)) + \deg(Q(t)).$$

Si $P(t)$ y $Q(t)$ difieren en su grado, se cumple la igualdad

$$\deg(P(t) + Q(t)) = \text{máx}\{\deg(P(t)), \deg(Q(t))\},$$

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*

$$A(t) = Q(t)B(t) + R(t), \;\; y \;\; \deg(R(t)) < \deg(B(t)).$$

Dem. Sea C el conjunto $\{A(t) - S(t)B(t)\}_{S(t) \in \mathbb{K}[t]}$. Como C no es vacio, 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 $R(t) = at^n + \ldots$ y $B(t) = bt^m + \ldots$ 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 $A(t) - \left(Q(t) + \frac{a}{b}t^{n-m}\right)B(t)$, 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 $\deg\left(\tilde{R}(t)\right) < \deg(B(t))$. De las igualdades

$$Q(t)B(t) + R(t) = A(t) = \tilde{Q}(t)B(t) + \tilde{R}(t),$$

se obtiene

$$\left(Q(t) - \tilde{Q}(t)\right)B(t) = \tilde{R}(t) - R(t).$$

Como

$$\deg\left(R(t) - \tilde{R}(t)\right) \leq \text{máx}\{\deg(R(t)), \deg\left(\tilde{R}(t)\right)\} < \deg(B(t)),$$

entonces

$$\deg\left((Q(t) - \tilde{Q}(t))B(t)\right) = \deg\left(R(t) - \tilde{R}(t)\right) < \deg(B(t)).$$

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 $\tilde{S}(t) = S(t)$. □

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 $R(t) = 0$. □

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 siguiente 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 $D(t)$ es 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 máximal entre los divisores comunes de $A(t)$ y $QBt)$. 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 $B(t)$.

Ejemplo A.23. Si $\lambda_1, \lambda_2 \in \mathbb{K}$ son distintos, entonces $(t - \lambda_1, t - \lambda_2) = 1$, pues se tiene la igualdad

$$\frac{1}{\lambda_2 - \lambda_1}\left((t - \lambda_1) - (t - \lambda_2)\right) = 1.$$

Teorema A.24 (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 $D(t)$. □

Observación A.25. La utilidad del álgoritmo de Euclides es que en la busqueda 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.26. 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

$$\begin{aligned} t - 1 &= 4B(t) - (2t-1)(2t^2 - t - 1) \\ &= 4B(t) - (2t-1)\big(A(t) - B(t)\big) \\ &= -(2t-1)A(t) + (2t+3)B(t) \end{aligned}$$

Observación A.27. 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 $(A_1(t), \ldots, A_n(t))$ al máximo común divisor de $\{A_1(t), \ldots, A_n(t)\}$, observemos que

$$(A_1(t), \ldots, A_n(t)) = \left((A_1(t), \ldots, A_{n-1}(t)), A_n(t)\right),$$

lo cual nos permite calcularlo inductivamente.

Teorema A.28 (Relación de Bézout)**.** *Dados $A_1(t), \ldots, A_n(t) \in \mathbb{K}[t]$, con uno de ellos no nulo, existen polinomios $P_1(t), \ldots, P_n(t) \in \mathbb{K}[t]$ tales que*

$$P_1(t)A_1(t) + \ldots + P_n(t)A_n(t) = (A_1(t), \ldots, A_n(t))$$

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 $Q_1(t), \ldots, Q_{n-1}(t) \in \mathbb{K}[t]$ tales que

$$Q_1(t)A_1(t) + \ldots + Q_{n-1}(t)A_{n-1}(t) = B(t),$$

donde $B(t) = (A_1(t), \ldots, A_{n-1}(t))$. 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 $i \in \{1, \ldots, n-1\}$, obtenemos

$$\begin{aligned} P_1(t)A_1(t) + \ldots + P_{n-1}(t)A_{n-1}(t) + P_n(t)A_n(t) &= (B(t), A_n(t)) \\ &= ((A_1(t), \ldots, A_{n-1}(t))\,, A_n(t)) = (A_1(t), \ldots, A_n(t))\,. \end{aligned}$$

$\square$

Ejemplo A.29. 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.26, 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 álgoritmo 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

$$\begin{aligned} 1 &= \frac{1}{2}A_3(t) - (\frac{1}{2}t + 1)\big(-(2t-1)A_1(t) + (2t+3)A_2(t)\big) \\ &= (\frac{1}{2}t + 1)(2t-1)A_1(t) - (\frac{1}{2}t + 1)(2t+3)A_2(t) + \frac{1}{2}A_3(t). \end{aligned}$$

Definición A.30. 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.31. *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.32 (Factorización única)**.** *Si $A(t) \in \mathbb{K}[t]$ es un polinomio con $\deg(A(t)) > 0$, entonces existen polinomios irreducibles mónicos $P_1(t), \ldots, P_r(t) \in \mathbb{K}[t]$ 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 $A(t) = cQ_1(t) \cdots Q_m(t)$ con $Q_1(t), \ldots, Q_s(t) \in \mathbb{K}[t]$ irreducibles mónicos, entonces s es igual a r y existe una permutación σ de $\{1, \ldots, r\}$ tal que, para $i \in \{1, \ldots, r\}$ 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 $P_1(t), \ldots, P_{r_1}(t)$, $P_{r_1+1}(t), \ldots, P_{r_1+r_2}(t) \in \mathbb{K}[t]$ 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 $\{1,\ldots,r-1\}$ tal que, para $i\in\{1,\ldots,r-1\}$, se tiene $Q_i(t)=P_{\sigma(i)}(t)$. 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.33. *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 $\overline{w}$, el conjugado de w, también es una raíz de $P(t)$. Luego $t-w$ y $t-\overline{w}$ dividen a $P(t)$ en $\mathbb{C}[t]$ y así, el mínimo multiplo común $(t-w)(t-\overline{w})$ también divide a $P(t)$ (ver ejercicio 10). Por ende, como se tiene $(t-w)(t-\overline{w})=(t-a-bi)(t-a+bi)=(t-a)^2+b^2$ entonces $(t-w)(t-\overline{w})$ 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 $\{A_1(t), \ldots, A_n(t)\} \subseteq \mathbb{Q}[t]$:

   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 $(A_1(t), \ldots, A_n(t))$ en $\mathbb{Q}(t)$.
   (b) Encuentre polinomios $P_1(t), \ldots, P_n(t) \in \mathbb{Q}[t]$ tales que

   $$(A_1(t), \ldots, A_n(t)) = P_1(t)A_1(t) + \ldots + P_n(t)A_n(t).$$

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 $A_0(t), A_1(t), \ldots, A_m(t) \in \mathbb{K}[t]$ con $\deg(A_i(t)) < \deg(P(t))$ para $i \in \{0, 1, \ldots, m\}$ tales que

    $$A(t) = A_0(t) + A_1(t)P(t) + \ldots + A_m(t)P(t)^m.$$

12. Sea $\mathbb{K}$ un cuerpo y sean $c_0, c_1, \ldots, c_n \in \mathbb{K}$ distintos. Para $i \in \{0, 1, \ldots, n\}$ 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

    $$P(t) = \sum_{i=0}^{n} P(c_i)P_i(t).$$