Notas de álgebra lineal

Pedro Massey

1er semestre 2020

Contents

1 Repaso

Sea $\mathcal{V}$ un K-ev:

1. $G = \{v_1, \ldots, v_m\} \subset \mathcal{V}$ es un sistema de generadores si $\overline{G} = \mathcal{V}$ (el subespacio generado por G coincide con $\mathcal{V}$). En otras palabras, G genera a $\mathcal{V}$ si para todo $v \in \mathcal{V}$ existen $\alpha_1, \ldots, \alpha_m \in K$ tales que
$$v = \sum_{j=1}^{m} \alpha_j \ v_j \,.$$

2. $F = \{w_1, \ldots, w_r\} \subset \mathcal{V}$ es linealmente independiente (l.i) si para todos $\beta_1, \ldots, \beta_r \in K$ tales que
$$\sum_{j=1}^{r} \beta_j \; w_j = \vec{0} \implies \beta_1 = \ldots = \beta_r = 0 \,.$$
3. Recordemos que hemos probado que si G y F son como arriba, entonces $r \leq m$.
4. Un conjunto ordenado $B = \{z_1, \ldots, z_n\} \subset \mathcal{V}$ es base de $\mathcal{V}$ si B es sistema de generadores linealmente independiente. En este caso, notamos $n = \dim_K \mathcal{V}$.

Recordemos que la dimensión de un espacio vectorial (de dimensión finita) está bien definida!

2 Sobre subespacios y dimensiones

2.1 Subespacios

Comenzamos con un lema que será de mucha utilidad en lo que sigue.

Lema 2.1. *Sea $\mathcal{V}$ un K-ev y sea $F = \{v_1, \ldots, v_k\} \subset \mathcal{V}$ un conjunto l.i. Supongamos que $v \in \mathcal{V} \setminus \overline{\{v_1, \ldots, v_k\}}$ (es decir, v es un vector de $\mathcal{V}$ que no está en el subespacio generado por F) entonces*
$$\{v_1, \ldots, v_k, \, v\} \quad \textit{es l.i.}$$

Demostración. Sean $\alpha_1, \ldots, \alpha_k, \, \alpha_{k+1} \in K$ tales que
$$\alpha_1 \, v_1 + \ldots + \alpha_k \; v_k + \alpha_{k+1} \; v = \vec{0} \,. \tag{1}$$
En este caso, queremos ver que $\alpha_1 = \ldots = \alpha_k = \alpha_{k+1} = 0$.

Supongamos que $\alpha_{k+1} \neq 0$. Entonces $-\alpha_{k+1} \neq 0$, pues en cualquier cuerpo se verifica que $-(-a) = a$ y $-0 = 0$. En este caso, la Eq. (1) implica que
$$-\alpha_{k+1} \; v = \alpha_1 \, v_1 + \ldots + \alpha_k \; v_k$$
Como $-\alpha_{k+1} \neq 0$ podemos multiplicar a ambos lados de la igualdad vectorial anterior por $(-\alpha_{k+1})^{-1} \in K$ y aplicar las propiedades de la acción escalar para obtener
$$v = (-\alpha_{k+1})^{-1} \; (\alpha_1 \, v_1 + \ldots + \alpha_k \; v_k) = ((-\alpha_{k+1})^{-1}\alpha_1) \; v_1 + \ldots + ((-\alpha_{k+1})^{-1}\alpha_k) \; v_k \,.$$
La igualdad entre el primer y tercer miembro de la ecuación anterior implica que v es combinación lineal de $\{v_1, \ldots, v_k\}$, es decir, $v \in \overline{\{v_1, \ldots, v_k\}}$, que contradice la hipótesis.

El argumento anterior muestra que $\alpha_{k+1} = 0$. Así, la Eq. (1) se transforma en
$$\alpha_1 \, v_1 + \ldots + \alpha_k \; v_k = \vec{0} \,.$$
Como $\{v_1, \ldots, v_k\}$ son l.i. por hipótesis, ahora podemos ver que el resto de los alfas satisfacen $\alpha_1 = \ldots = \alpha_k = 0 = \alpha_{k+1}$.

□

Corolario 2.2. *Sea $\mathcal{V}$ un K-ev tal que $\dim_K \mathcal{V} = n$. Si el conjunto $L = \{v_1, \ldots, v_n\} \subset \mathcal{V}$ es l.i., entonces L es base de $\mathcal{V}$.*

Demostración. Por hipótesis L es l.i. Si suponemos que L no es un sistema de generadores de $\mathcal{V}$, es decir $\overline{L} \subsetneq \mathcal{V}$ (la inclusión es estricta) entonces existe $v \in \mathcal{V} \setminus \overline{L}$. Por el lema anterior, $\{v_1, \ldots, v_n, v\}$ es l.i.

Por otro lado, como $\dim_K \mathcal{V} = n$, entonces existe una base $B = \{w_1, \ldots, w_n\}$ de $\mathcal{V}$. Como B es, en particular, un sistema de generadores, aplicando el ítem 3 de la sección Repaso, debemos tener que $n+1 \leq n$, lo que es claramente una contradicción. Concluimos entonces que L debe ser un sistema de generadores de $\mathcal{V}$, de forma que L resulta una base de $\mathcal{V}$.

□

El siguiente resultado juega un papel fundamental en la teoría de espacios vectoriales.

Teorema 2.3. *Sea $\mathcal{V}$ un K-ev con $\dim_K \mathcal{V} = n \in \mathbb{N}$. Entonces, todo conjunto l.i. de $\mathcal{V}$ puede extenderse a una base de $\mathcal{V}$.*

Demostración Sea $S = \{v_1, \ldots, v_k\}$ un conjunto l.i. Como $\dim_K \mathcal{V} = n$ concluimos que $k \leq n$. Consideramos los siguientes casos:

Si $k = n$, entonces S es una base de $\mathcal{V}$, por el Corolario anterior.

Si $k < n$ entonces S no es base de $\mathcal{V}$ (recordemos que la cantidad de elementos (ó *cardinal*) de cualquier base de $\mathcal{V}$ debe ser n). Como S es l.i., entonces $\overline{S} \subsetneq \mathcal{V}$ (la inclusión es estricta) de forma que existe $v_{k+1} \in \mathcal{V} \setminus \overline{S}$ (diferencia de conjuntos). Por el Lema 2.1 concluímos que $S_1 = \{v_1, \ldots, v_k, v_{k+1}\}$ es l.i. Si $k+1 = n$ entonces S_1 es base de $\mathcal{V}$, por el Corolario 2.2. De esta forma, hemos extendido el conjunto l.i. S a la base S_1 (agregando el vector v_{k+1}).

Si $k+1 < n$ entonces repetimos el argumento anterior y hallamos $v_{k+2} \in \mathcal{V} \setminus \overline{S_1}$ de forma que $S_2 = \{v_1, \ldots, v_{k+2}\}$ es l.i.

Así, repitiendo el argumento de los casos anteriores $n-k$ veces, construimos el conjunto

$$S_{n-k} = \{v_1, \ldots, v_k, v_{k+1}, \ldots, v_n\}$$

que se obtiene agregando los $(n-k)$ vectores $v_{k+1}, \ldots, v_n$ al conjunto inicial S, de forma tal que S_{n-k} es l.i. Si aplicamos nuevamente el Corolario 2.2, concluimos que S_{n-k} es base de $\mathcal{V}$ (que extiende al conjunto S). □

El siguiente resultado es de interés y muestra, en particular, que *ser de dimensión finita es una propiedad que se hereda.*

Corolario 2.4. *Sea $\mathcal{V}$ un K-ev tal que $\dim_K \mathcal{V} = n$. Si $\mathcal{W} \subset \mathcal{V}$ es subespacio propio de $\mathcal{V}$ entonces $1 \leq \dim_K \mathcal{W} < n$. En particular, los subespacios de $\mathcal{V}$ son de dimensión finita.*

Demostración. Recordemos que un subespacio $\mathcal{W} \subset \mathcal{V}$ es propio si $\mathcal{W} \neq \{0\}$ y $\mathcal{W} \neq \mathcal{V}$. Además, en tanto subespacio, $\mathcal{W}$ con la suma y la acción escalar de $\mathcal{V}$ restringidas es un K-ev, y podemos aplicar los resultados vistos en este caso.

Así, como $\mathcal{W} \neq \{0\}$, existe $v_1 \in \mathcal{W}$ de forma que $v_1 \neq \vec{0}$. En este caso, $S_1 = \{v_1\}$ es un conjunto l.i. (porqué?). Argumentamos como en la prueba del teorema anterior. Si S_1 es base de $\mathcal{W}$, entonces detenemos este proceso y vamos al párrafo final de la prueba.

Si S_1 no es base de $\mathcal{W}$ entonces $\overline{S} \subsetneq \mathcal{W}$ y existe $v_2 \in \mathcal{W} \setminus \overline{S}$ de forma que $S_2 = \{v_1, v_2\} \subset \mathcal{W}$ es l.i.

Notemos que podemos repetir el argumento k veces en la medida de que el conjunto l.i. $S_{k-1} = \{v_1, \ldots, v_{k-1}\} \subset \mathcal{W}$ no es base de $\mathcal{W}$ (es decir, no genera a $\mathcal{W}$).

Afirmamos ahora que este proceso de debe detener en a lo sumo $n = \dim_K \mathcal{V}$ pasos. En efecto, si el proceso no se detiene en a lo sumo n pasos, entonces podemos realizar el paso $n+1$. En este paso terminamos construyendo $S_{n+1} = \{v_1, \ldots, v_{n+1}\} \subset \mathcal{W}$ que es l.i. Ahora, recordamos que $\mathcal{W} \subset \mathcal{V}$ de forma que $S_{n+1} \subset \mathcal{V}$ es l.i. como subconjunto de $\mathcal{V}$ (aquí debemos recordar que las operaciones de espacio vectorial de $\mathcal{W}$ son las mismas que las operaciones de $\mathcal{V}$). Esto último contradice el item 3 de la sección Repaso (pues $\mathcal{V}$ admite un sistema de generadores (una base!) con n vectores).

Entonces el proceso se detiene en k pasos, con $k \leq n$. Recordemos que en este caso, se debe tener que $S_k = \{v_1, \ldots, v_k\}$ es l.i. y además, que $\overline{S_k} = \mathcal{W}$ (de otra forma, se puede continuar con el proceso). Estos hechos muestran que S_k es base de $\mathcal{W}$ y $\dim_K \mathcal{W} = k$. Finalmente, se debe tener que $k < n$: de otra forma $k = n$ y $S_n = \{v_1, \ldots, v_n\} \subset \mathcal{W}$ base de $\mathcal{W}$. Como S_n es l.i. y $\dim_K \mathcal{V} = n$ entonces, por el Corolario 2.2, S_n también resulta base de $\mathcal{V}$ y luego $\mathcal{W} = \overline{S_n} = \mathcal{V}$, que contradice que $\mathcal{W}$ es propio.

□

Obs 2.5. Sea $\mathcal{V}$ un K-ev tal que $\dim_K \mathcal{V} = n$. Si $\mathcal{W} \subset \mathcal{V}$ es subespacio de $\mathcal{V}$ entonces $0 \leq \dim_K \mathcal{W} \leq n$. En efecto, si $\mathcal{W}$ es propio, entonces $1 \leq \dim_K \mathcal{W} < n$, por el resultado anterior. Si $\mathcal{W}$ no es propio entonces $\mathcal{W} = \{0\}$ en cuyo caso $\dim_K \mathcal{W} = 0$ (recordemos que $\emptyset$ es base de $\{0\}$) ó $\mathcal{W} = \mathcal{V}$ en cuyo caso $\dim_K \mathcal{W} = n$.

□

El siguiente resultado es el punto de partida de una familia de desigualdades fundamentales para las aplicaciones del álgebra lineal (en el caso de operadores autoadjuntos en dimensión finita).

Teorema 2.6. *Sea $\mathcal{V}$ un K-ev y sean W_1, W_2 subespacios. Recordemos que en este caso*

$$W_1 + W_2 = \{u + v : u \in W_1,\, v \in W_2\} \quad y \quad W_1 \cap W_2$$

son subespacios de $\mathcal{V}$. Entonces se verifica

$$\dim_K(W_1 + W_2) = \dim_K(W_1) + \dim_K(W_2) - \dim_K(W_1 \cap W_2)\,.$$

Demostración. Como $W_1 \cap W_2 \subseteq \mathcal{V}$ es subespacio, entonces la Observación 2.5 concluimos que existe $\{v_1, \ldots, v_k\}$ base de $W_1 \cap W_2$. En este caso, $\dim_K W_1 \cap W_2 = k$. Por el Teorema 2.3, como

$$\{v_1, \ldots, v_k\} \subset W_1 \cap W_2 \subset \mathcal{W}_1$$

es l.i., se puede extender a una base $B_1 = \{v_1, \ldots, v_k, w_1, \ldots, w_r\}$ de $\mathcal{W}_1$: en este caso, $\dim_K W_1 = k+r$. De forma análoga, podemos extender a $\{v_1, \ldots, v_k\}$ de forma que $B_2 = \{v_1, \ldots, v_k, z_1, \ldots, z_s\}$ sea base de $\mathcal{W}_2$: en este caso $\dim_K W_2 = k + s$.

Afirmamos que $B = \{v_1, \ldots, v_k, w_1, \ldots, w_r, z_1, \ldots, z_s\}$ es base de $W_1 + W_2$.

Aceptemos por un instante la afirmación anterior: en este caso concluimos que $\dim_K(W_1 + W_2) = k + r + s$ (que es la cantidad de vectores que hay en B). Entonces

$$\dim_K(W_1 + W_2) = k + r + s = (k + r) + (k + s) - k = \dim_K W_1 + \dim_K W_2 - \dim_K(W_1 \cap W_2)\,,$$

que prueba el teorema.

En lo que sigue verificamos que B es base de $W_1 + W_2$ es decir, mostramos que B genera a $W_1 + W_2$ y que es l.i.

B genera a $W_1 + W_2$: recordemos que un vector en $W_1 + W_2$ es de la forma $u + v$, con $u \in W_1$ y $v \in W_2$. En este caso, como B_1 es base de W_1 existen $\alpha_1, \ldots, \alpha_k,\, \beta_1, \ldots, \beta_r \in K$ tales que

$$u = \alpha_1\, v_1 + \ldots + \alpha_k\, v_k + \beta_1\, w_1 + \ldots + \beta_r\, w_r\,.$$

De forma similar, existen $\gamma_1, \ldots, \gamma_k$, $\delta_1, \ldots, \delta_s \in K$ tales que

$$v = \gamma_1\, v_1 + \ldots + \gamma_k\, v_k + \delta_1\, z_1 + \ldots + \delta_s\, z_s\,.$$

Así, utilizando las propiedades de la suma y la acción escalar, vemos que

$$\begin{aligned} u + v &= (\alpha_1\, v_1 + \ldots + \alpha_k\, v_k + \beta_1\, w_1 + \ldots + \beta_r\, w_r) + (\gamma_1\, v_1 + \ldots + \gamma_k\, v_k + \delta_1\, z_1 + \ldots + \delta_s\, z_s) \\ &= (\alpha_1 + \gamma_1)\, v_1 + \ldots + (\alpha_k + \gamma_k)\, v_k + \beta_1\, w_1 + \ldots + \beta_r\, w_r + \delta_1\, z_1 + \ldots + \delta_s\, z_s \end{aligned}$$

que muestra que B es sistema de generadores para $W_1 + W_2$.

B es l.i.: supongamos que $\alpha_1, \ldots, \alpha_k$, $\beta_1, \ldots, \beta_r$, $\delta_1, \ldots, \delta_s \in K$ son tales que

$$\alpha_1\, v_1 + \ldots + \alpha_k\, v_k + \beta_1\, w_1 + \ldots + \beta_r\, w_r + \delta_1\, z_1 + \ldots + \delta_s\, z_s = \vec{0}\,. \tag{2}$$

Entonces, se verifica la ecuación vectorial

$$\alpha_1\, v_1 + \ldots + \alpha_k\, v_k + \beta_1\, w_1 + \ldots + \beta_r\, w_r = -\delta_1\, z_1 + \ldots + -\delta_s\, z_s\,.$$

El vector que figura a la izquierda en la identidad de arriba está en W_1, pues es combinación lineal de elementos de W_1 (y W_1 es subespacio, de forma que es cerrado bajo combinaciones lineales). Por otro lado, el vector que figura a la derecha en la identidad de arriba está en W_2, pues es combinación lineal de elementos de W_2 ($z_1, \ldots, z_s$ son vectores en W_2). Como ambas expresiones son el mismo vector (por la igualdad), se verifica que

$$-\delta_1\, z_1 + \ldots + -\delta_s\, z_s \in W_1 \cap W_2$$

ya que está simultáneamente en W_1 y W_2.

Entonces, el vector anterior se debe poder escribir como combinación lineal de los elmentos de la base $\{v_1, \ldots, v_k\}$ de $W_1 \cap W_2$, es decir, existen $\rho_1, \ldots, \rho_k \in K$ tales que

$$-\delta_1\, z_1 + \ldots + -\delta_s\, z_s = \rho_1\; v_1 + \ldots, \rho_k\, v_k\,.$$

La ecuación anterior puede escribirse como

$$\rho_1\; v_1 + \ldots, \rho_k\, v_k + \delta_1\, z_1 + \ldots + \delta_s\, z_s = \vec{0}$$

que muestra una combinación lineal de los vectores de la base B_2 de W_2 igualada a $\vec{0}$. Por la indepencia lineal de B_2, concluimos que $\rho_1 = \ldots = \rho_k = \delta_1 = \ldots = \delta_s = 0$. Con estas últimas identidades de los delta's (son todos ceros) vemos que la ecuación (2) se transforma en

$$\alpha_1\, v_1 + \ldots + \alpha_k\, v_k + \beta_1\, w_1 + \ldots + \beta_r\, w_r = \vec{0}\,.$$

Usando que B_1 es base de W_1, de forma que en particular B_1 es l.i., ahora vemos que se debe tener que $\alpha_1 = \ldots = \alpha_k = \beta_1 = \ldots = \beta_r = 0$. Estas identidades de los coeficientes, junto con las previas para los delta's muestran que B es l.i. $\square$

2.2 Coordenadas

Obs 2.7. En lo que sigue, dado un cuerpo K, consideramos el espacio vectorial $K^n = K^{n\times 1}$ de vectores *columna*. Por una cuestión de espacio en el texto, en muchos casos escribimos $(\alpha_1, \ldots, \alpha_n) \in K^n$, pero debemos recordar que se tratan de vectores columna (y no vectores fila).

Proposición 2.8. *Sea $\mathcal{V}$ un K-ev y sea $B = \{v_1, \ldots, v_n\}$ una base de $\mathcal{V}$. Si $v \in \mathcal{V}$ entonces existe una única n-upla $(\alpha_1, \ldots, \alpha_n) \in K^n$ tal que*

$$v = \alpha_1\ v_1 + \ldots + \alpha_n\ v_n\,.$$

Demostración. Sea $v \in \mathcal{V}$. Como B es un sistema de generadores, entonces existe $(\alpha_1, \ldots, \alpha_n) \in K^n$ tal que $v = \alpha_1\ v_1 + \ldots + \alpha_n\ v_n$.

Si suponemos que $(\beta_1, \ldots, \beta_n) \in K^n$ es tal que $v = \beta_1\ v_1 + \ldots + \beta_n\ v_n$, entonces

$$\alpha_1\ v_1 + \ldots + \alpha_n\ v_n = \beta_1\ v_1 + \ldots + \beta_n\ v_n \implies$$

$$(\alpha_1 - \beta_1)\ v_1 + \ldots + (\alpha_n - \beta_n)\ v_n = \vec{0}\,.$$

Usando que B es l.i., concluimos que

$$\alpha_1 - \beta_1 = \ldots = \alpha_n - \beta_n = 0 \implies \alpha_1 = \beta_1\ ,\ \ldots\ ,\ \alpha_n = \beta_n\,,$$

que demuestra la unicidad de la n-upla.

□

Definición 2.9. *Sea $\mathcal{V}$ un K-ev y sea $B = \{v_1, \ldots, v_n\}$ una base de $\mathcal{V}$. Dado $v \in \mathcal{V}$, definimos el vector de coordenadas de v con respecto a (la base ordenada) B, notado $[v]_B$, como la única n-upla de la proposición anterior. Es decir, $[v]_B = (\alpha_1, \ldots, \alpha_n) \in K^n$ es la única n-upla tal que*

$$v = \alpha_1\ v_1 + \ldots + \alpha_n\ v_n\,.$$

□

Con las notaciones de la definición anterior, notemos que a partir de las coordenadas $[v]_B$ y los elementos de la base B podemos recuperar el vector $v \in \mathcal{V}$. De esta forma, las coordenadas de un vector son una forma de codificar el vector a través de n-uplas.

Notemos que el orden de los vectores en una base ordenada resulta fundamental en este contexto: si sabemos que $[v]_B = (\alpha_1, \alpha_2)$ ($n = 2$ en este caso) entonces

$$v = \alpha_1\ v_1 + \alpha_2\ v_2 \quad \text{que resulta distinto del vector} \quad \alpha_1\ v_2 + \alpha_2\ v_1\ !!$$

Ejemplos 2.10. Sea $\mathcal{V}$ un K-ev y sea $B = \{v_1, \ldots, v_n\}$ una base de $\mathcal{V}$.

1. Por definición $[v_1]_B = (1, 0, \ldots, 0)$. En efecto, $v_1 = 1\ v_1 + 0\ v_2 + \ldots + 0\ v_n$. En este caso escribimos $[v_1]_B = e_1$, donde $\{e_1, \ldots, e_n\}$ denotan los vectores de la base canónica de K^n. De hecho, recordemos que

$$e_j = (0, \ldots, 0, 1, 0 \ldots, 0) \quad \text{donde el 1 aparece en la } j\text{-ésima coordenada} \quad .$$

2. De forma similar $[v_j]_B = e_j$ para $j = 1, \ldots, n$.

3. Es sencillo verificar que $[0]_B = (0, \ldots, 0)$.

□

Hagamos el cálculo de coordenas con respecto a alguna base en $\mathbb{R}^3$.

Ejemplo 2.11. En el $\mathbb{R}$-ev $\mathbb{R}^3$, consideramo los vectores

$$v_1 = (1,0,0) \quad , \quad v_2 = (1,1,0) \quad \text{y} \quad v_3 = (1,1,1)\,.$$

Entonces $B = \{v_1,\, v_2,\, v_3\}$ es base ordenada de $\mathbb{R}^3$ (verificar, ejercicio).

Dado $v = (a,b,c) \in \mathbb{R}^3$, calculemos las coordenadas $[v]_B$: para eso, planteamos la ecuación vectorial

$$(a,b,c) = \alpha_1\, v_1 + \alpha_2\; v_2 + \alpha_3\; v_3 = \alpha_1\,(1,0,0) + \alpha_2\;(1,1,0) + \alpha_3\;(1,1,1)\,.$$

Notemos que las incógnitas de esta ecuación son los coeficientes $\alpha_1,\, \alpha_2,\, \alpha_3 \in \mathbb{R}^3$.

La ecuación se verifica si los vectores a ambos lados tienen las mismas coordenadas. De esta forma, realizando las operaciones correspondientes en $\mathbb{R}^3$ vemos que la ecuación vectorial anterior equivale al sistema de ecuaciones lineales

$$\begin{aligned} a &= \alpha_1 + \alpha_2 + \alpha_3 \\ b &= \alpha_2 + \alpha_3 \\ c &= \alpha_3 \end{aligned}$$

Notemos que la matriz ampliada del sistema es

$$\begin{pmatrix} 1 & 1 & 1 & | & a \\ 0 & 1 & 1 & | & b \\ 0 & 0 & 1 & | & c \end{pmatrix}$$

que ya es escalonada. Si reducimos por filas llegamos a la matrix ampliada

$$\begin{pmatrix} 1 & 0 & 0 & | & a-b \\ 0 & 1 & 0 & | & b-c \\ 0 & 0 & 1 & | & c \end{pmatrix}$$

Así, la solución del sistema es $(a-b, b-c, c)$ es decir,

$$[(a,b,c)]_B = (a-b\,,\; b-c\,,\; c) \in \mathbb{R}^3\,.$$

Por otro lado, si consideramos la base canónica $B_C = \{e_1,\, e_2,\, e_3\}$ de $\mathbb{R}^3$, es sencillo verificar que

$$[(a\,,\, b\,,\, c)]_{B_C} = (a\,,\, b\,,\, c)$$

pues $(a,b,c) = a\, e_1 + b\, e_2 + c\, e_3$.

□

2.3 Matriz cambio de base

En el Ejemplo 2.11 observamos que las coordenadas de un vector dependen de la base del espacio considerada (en general, si cambia la base entonces cambian las coordenadas del vector). En lo que sigue, estudiamos cómo se relacionan las coordenadas de un vector v con respecto a dos bases distintas B_1 y B_2 del espacio.

Definición 2.12. Sea $\mathcal{V}$ un K-ev y sean $B_1 = \{v_1, \ldots, v_n\}$ y $B_2 = \{w_1, \ldots, w_n\}$ dos bases de $\mathcal{V}$. En este caso construimos la *matriz cambio de base de la base* B_1 *a la base* B_2, notada $M_{B_2\,,\,B_1} \in K^{n\times n}$ como sigue: las columnas de la matriz $M_{B_2\,,\,B_1}$ son los vectores (columna) de coordenadas $[v_j]_{B_2} \in K^n$, para $j = 1, \ldots, n$. Es decir

$$M_{B_2\,,\,B_1} = \begin{pmatrix} | & \cdots & | \\ [v_1]_{B_2} & \cdots & [v_n]_{B_2} \\ | & \cdots & | \end{pmatrix} \in K^{n\times n}\,.$$

□

Queda claro de la definición anterior que el rol de las bases B_1 y B_2 en la construcción de $M_{B_2\,,\,B_1}$ es distinto, de forma que en general $M_{B_2\,,\,B_1} \neq M_{B_1\,,\,B_2}$.

Ejemplo 2.13. Consideremos las bases $B_C = \{e_1,\, e_2,\, e_3\}$ y $B = \{v_1,\, v_2,\, v_3\}$ del Ejemplo 2.11. Calculamos $M_{B\,,\,B_C}$ (es decir, en la definición anterior consideramos $B_1 = B_C$ y $B_2 = B$). Aprovechamos los cálculos hechos anteriormente, que indican $[(a,b,c)]_B = (a-b, b-c, c)$. En particular

$$[e_1]_B = [(1,0,0)]_B = (1,0,0)$$

$$[e_2]_B = [(0,1,0)]_B = (-1,1,0)$$

$$[e_3]_B = [(0,0,1)]_B = (0,-1,1)$$

Así, $M_{B\,,\,B_C}$ es la matriz cuya primer columna es $[e_1]_B \in \mathbb{R}^3$, cuya segunda columna es $[e_2]_B \in \mathbb{R}^3$ y tercer columna es $[e_3]_B \in \mathbb{R}^3$, es decir

$$M_{B\,,\,B_C} = \begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{pmatrix} \in \mathbb{R}^{3\times 3}\,.$$

□

En lo que sigue vamos a necesitar los siguientes hechos:

Obs 2.14. Sea K un cuerpo y sea $A \in K^{m\times n}$. Podemos pensar a la matriz A en términos de sus columnas $\vec{a}_1, \ldots, \vec{a}_n \in K^m$, de forma que

$$A = \begin{pmatrix} | & \cdots & | \\ \vec{a}_1 & \cdots & \vec{a}_n \\ | & \cdots & | \end{pmatrix} \in K^{m\times n}\,.$$

Si pensamos en términos de las coordenadas de $A = (A_{ij})$ y de los vectores columna $a_j \in K^m$ se tiene

$$\vec{a}_j = \begin{pmatrix} A_{1j} \\ A_{2j} \\ \vdots \\ A_{mj} \end{pmatrix} \in K^m \quad \text{es decir} \quad (\vec{a}_j)_i = A_{ij}\,. \tag{3}$$

□

Proposición 2.15. *Con las notaciones de la Obs. 2.14, se verifica que: dado* $\vec{x} = (x_1, \ldots, x_n) \in K^n$ *(vector columna) vale la identidad*

$$A \cdot \vec{x} = \sum_{j=1}^{n} x_j\ \vec{a}_j\,.$$

En particular, si $\{e_1, \ldots, e_n\}$ *es la base canónica de* K^n *deducimos que*

$$A\, e_j = \vec{a}_j \quad \textit{para} \quad j = 1, \ldots, n\,.$$

Demostración. Notemos que la primera identidad propuesta en el enunciado compara el producto matriz-vector $A \cdot \vec{x}$ con la combinación lineal de los vectores columna $\vec{a}_1, \ldots, \vec{a}_n$ con coeficientes dados por las entradas del vector $\vec{x}$.

Por un lado, si definimos $A\,\vec{x} = \vec{y} \in K^m$ entonces las entradas del vector $\vec{y}$ son

$$y_i = \sum_{j=1}^{n} A_{ij}\ x_j \qquad \text{para} \qquad i = 1, \ldots, m\,.$$

Por otro lado, si definimos $\sum_{j=1}^{n} x_j\ \vec{a}_j = \vec{z} \in K^m$ entonces las entradas del vector $\vec{z}$ son

$$z_i = (\sum_{j=1}^{n} x_j\ \vec{a}_j)_i = \sum_{j=1}^{n} x_j\ (\vec{a}_j)_i = \sum_{j=1}^{n} x_j\ A_{ij} = \sum_{j=1}^{n} A_{ij}\ x_j = y_i$$

donde hemos usado la Eq. (3). Así, los vectores $\vec{y}$ y $\vec{z}$ son iguales (porque tienen las mismas entradas) de forma que se verifica la primer identidad de la proposición.

Para verificar la segunda identidad, aplicamos la primer identidad, tendiendo en cuenta que $e_j = (0, \ldots, 0, 1, 0, \ldots, 0)$ donde el único 1 aparece en la entrada j-ésima: entonces

$$A\,e_j = 0\,\vec{a}_1 + \ldots 0\,\vec{a}_{j-1} + 1\,\vec{a}_j + 0\,\vec{a}_{j+1} + \ldots + 0\,\vec{a}_n = \vec{a}_j\,.$$

□

Obs 2.16. Con la notación de la proposición anterior, notemos que podemos describir a la matriz A por columnas $\vec{a}_j = A\,e_j$: así se tiene la identidad

$$A = \begin{pmatrix} | & \cdots & | \\ A\,e_1 & \cdots & A\,e_n \\ | & \cdots & | \end{pmatrix}$$

□

Veamos ahora las primeras propiedades fundamentales de la matriz cambio de base

Teorema 2.17. *Sea $\mathcal{V}$ un K-ev y sean $B_1 = \{v_1, \ldots, v_n\}$ y $B_2 = \{w_1, \ldots, w_n\}$ dos bases de $\mathcal{V}$. Consideremos la matriz de cambio de base $M_{B_2\,,B_1}$ (de la base B_1 a la base B_2). Entonces se verifica que para todo $v \in \mathcal{V}$*

$$[v]_{B_2} = M_{B_2\,,B_1} \cdot [v]_{B_1}\,. \tag{4}$$

Más aún, $M_{B_2\,,B_1} \in K^{n\times n}$ es la única matriz que verifica la identidad en Eq. (4), para todo $v \in \mathcal{V}$.

Demostración. Veamos que vale la identidad de Eq. (4). En efecto, por definición se tiene que si denotamos las entradas de la matriz

$$M_{B_2\,,B_1} = (M_{ij}) \implies v_j = \sum_{i=1}^{n} M_{ij}\ w_i \qquad \text{para} \qquad j = 1, \ldots, n\,. \tag{5}$$

La afirmación anterior corresponde al hecho de que el vector de coordenadas $[v_j]_{B_2} \in K^n$ coincide con la j-ésima columna de $M_{B_2\,,B_1}$, es decir,

$$[v_j]_{B_2} = \begin{pmatrix} M_{1j} \\ \vdots \\ M_{nj} \end{pmatrix}$$

lo que es cierto, por construcción de la matriz de cambio de base $M_{B_2\,,B_1}$. Usando las representaciones de la Eq. (5) vemos que si $v \in \mathcal{V}$, y $[v]_{B_1} = (\alpha_1, \ldots, \alpha_n)$ entonces

$$\begin{aligned} v &= \sum_{j=1}^n \alpha_j \ v_j = \sum_{j=1}^n \alpha_j \left(\sum_{i=1}^n M_{ij} \ w_i \right) = \sum_{j=1}^n \sum_{i=1}^n \alpha_j M_{ij} \ w_i \\ &= \sum_{i=1}^n \sum_{j=1}^n \alpha_j M_{ij} \ w_i = \sum_{i=1}^n \ \left(\sum_{j=1}^n \alpha_j M_{ij} \right) \ w_i \end{aligned}$$

donde hemos usando: la propiedad de la acción escalar relacionada con la suma de vectores, las propiedades asociativa y conmutativa al momento de cambiar el orden de sumación (vista en clase), y la identidad

$$\sum_{j=1}^n \alpha_j M_{ij} \ w_i = \left(\sum_{j=1}^n \alpha_j M_{ij} \right) \ w_i = \left(\sum_{j=1}^n M_{ij} \ \alpha_j \right) \ w_i$$

que corresponde a propiedad de la acción escalar (y la conmutatividad del producto del cuerpo), notando que el vector w_i está fijo en la suma (el índice i no depende del índice de sumación j de la sumatoria). Lo anterior indica que

$$v = \sum_{i=1}^n \ \left(\sum_{j=1}^n M_{ij} \ \alpha_j \right) \ w_i \, .$$

Por unicidad de las coordenadas $[v]_{B_2}$ deducimos que

$$[v]_{B_2} = (\sum_{j=1}^n M_{1j} \ \alpha_j \, , \, \ldots \, , \, \sum_{j=1}^n M_{nj} \ \alpha_j) = M_{B_2\,,B_1} \cdot [v]_{B_1}$$

ya que $[v]_{B_1} = (\alpha_1, \ldots, \alpha_n)$, donde usamos la fórmula para el producto matriz-vector para $M_{B_2\,,B_1} \cdot [v]_{B_1}$.

Con respecto a la unicidad, supongamos que $P = (p_{ij}) \in K^{n\times n}$ también satisface la Eq. (4), es decir: para todo $v \in \mathcal{V}$,

$$[v]_{B_2} = P \cdot [v]_{B_1} \, .$$

Si recordamos el primer ejemplo de Ejemplos (2.10), entonces $[v_j]_{B_1} = e_j$ para $j = 1, \ldots, n$. Usando la hipótesis sobre P y la segunda identidad de la Proposicion 2.15 vemos que la j-ésima columna de la matriz P, es decir $P\, e_j$ satisface

$$P\, e_j = P\, [v_j]_{B_1} = [v_j]_{B_2} = M_{B_2\,,B_1}\, e_j$$

donde la última identidad vale por construcción de $M_{B_2\,,B_1}$ (i.e. la j-ésima columna de $M_{B_2\,,B_1}$ es $[v_j]_{B_2}$). Así, las matrices P y $M_{B_2\,,B_1}$ coinciden columna a columna y por lo tanto son iguales.

□

Ejemplo 2.18. Consideremos una vez más el Ejemplo 2.11 con $\mathcal{V} = \mathbb{R}^3$ y las bases $B_C = \{e_1, \, e_2, \, e_3\}$ y $B = \{v_1, \, v_2, \, v_3\}$. En este caso, hemos calculado que

$$M_{B\,,B_C} = \begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{pmatrix} \in \mathbb{R}^{3\times 3} \, .$$

Como ya hemos indicado, si $v = (a, b, c) \in \mathbb{R}^3$ entonces se verifica que

$$[v]_B = (a - b \, , \, b - c \, , \, c) \in \mathbb{R}^3 \quad \text{ y } \quad [v]_{B_C} = (a, b, c) \in \mathbb{R}^3 \, .$$

Por otro lado, el resultado anterior indica que

$$[v]_B = M_{B\,,B_C} \cdot [v]_{B_C} \qquad \text{es decir} \qquad [v]_B = \begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} a \\ b \\ c \end{pmatrix} .$$

Un cálculo sencillo muestra que

$$\begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} a-b \\ b-c \\ c \end{pmatrix}$$

de forma concluímos una vez más (pero por otro camino) que $[v]_B = (a-b\,,\, b-c\,,\, c)$.

□

Las matrices cambio de base tienen una serie de propiedades fundamentales adicionales. Para verificar estas propiedades consideramos la siguente

Obs 2.19. Sea K un cuerpo, y sea $A \in K^{n\times n}$ una matriz cuadrada. Sea $\{e_1, \ldots, e_n\}$ la base canónica de K^n y supongamos que A satisface

$$A\, e_j = e_j \qquad \text{para todo} \qquad j = 1, \ldots, n\,.$$

Entonces $A = I_n$ es la matriz identidad $n \times n$. En efecto, usando la Proposición 2.15 y la identidad anterior, podemos ver cuales son las columnas de la matriz A $(A\, e_j = e_j)$:

$$A = \begin{pmatrix} | & \cdots & | \\ A\, e_1 & \cdots & A\, e_n \\ | & \cdots & | \end{pmatrix} = \begin{pmatrix} | & \cdots & | \\ e_1 & \cdots & e_n \\ | & \cdots & | \end{pmatrix} = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix} = I_n \in K^{m\times n}\,.$$

□

Teorema 2.20. *Sea $\mathcal{V}$ un K-ev y sean $B_1 = \{v_1, \ldots, v_n\}$ y $B_2 = \{w_1, \ldots, w_n\}$ dos bases de $\mathcal{V}$. Consideremos la matriz de cambio de base $M_{B_2\,,B_1}$ (de la base B_1 a la base B_2). Entonces se verifica*

1. *$M_{B_2\,,B_1}$ es una matriz inversible*
2. *Más aún $(M_{B_2\,,B_1})^{-1} = M_{B_1\,,B_2}$.*

Demostración. Veamos que se verifica la identidad

$$M_{B_1\,,B_2} \cdot M_{B_2\,,B_1} = I_n$$

Si verificamos la fórmula anterior, entonces podemos concluir la validez de los ítems 1 y 2 a la vez (por resultados de Álgebra I). Recordemos la propiedad de $M_{B_1\,,B_2}$: para todo $v \in \mathcal{V}$,

$$[v]_{B_2} = M_{B_2\,,B_1} \cdot [v]_{B_1}\,.$$

De forma análoga, $M_{B_2\,,B_1}$ que es la matriz de cambio de base de la base B_2 a la base B_1 satisface: para todo $v \in \mathcal{V}$

$$[v]_{B_1} = M_{B_1\,,B_2} \cdot [v]_{B_2}\,.$$

Entonces, si elegimos $v = v_j \in B_1$ (para $1 \leq j \leq n$) vemos que, como $[v_j]_{B_1} = e_j$,

$$(M_{B_1\,,B_2} \cdot M_{B_2\,,B_1}) \cdot e_j = M_{B_1\,,B_2} \cdot (\, M_{B_2\,,B_1} \cdot [v_j]_{B_1}) = M_{B_1\,,B_2} \cdot [v_j]_{B_2} = [v_j]_{B_1} = e_j$$

donde hemos usado la propiedad asociativa del producto de matrices, la propiedad de $M_{B_2\,,B_1}$, la propiedad de $M_{B_1\,,B_2}$, y la identidad $[v_j]_{B_1} = e_j$. La conclusión es

$$(M_{B_1\,,B_2} \cdot M_{B_2\,,B_1}) \cdot e_j = e_j \quad \text{para} \quad j = 1, \ldots, n\,.$$

Por la Observación anterior, concluimos que

$$M_{B_1\,,B_2} \cdot M_{B_2\,,B_1} = I_n \in K^{n\times n}$$

□

Concluimos esta clase con un resultado que explota la unicidad de la matriz cambio de base.

Teorema 2.21. *Sea $\mathcal{V}$ un K-ev tal que $\dim_K \mathcal{V} = n \in \mathbb{N}$. Sean B_1, B_2 y B_3 bases de $\mathcal{V}$. Entonces se verifica que*

$$M_{B_3\,,B_1} = M_{B_3\,,B_2} \cdot M_{B_2\,,B_1}\,.$$

Demostración. Por la unicidad de la matriz de cambio de base $M_{B_3\,,B_1}$, basta ver que el producto $M_{B_3\,,B_2} \cdot M_{B_2\,,B_1} \in K^{n\times n}$ transforma coordenadas con respecto B_1 a coordenadas con respecto a B_3. A tal fin, sea $v \in \mathcal{V}$. Por un lado,

$$[v]_{B_3} = M_{B_3\,,B_2} \cdot [v]_{B_2}$$

mientras que

$$[v]_{B_2} = M_{B_2\,,B_1} \cdot [v]_{B_1}$$

Combinando las dos identidades anteriores y usando la asociatividad del producto de matrices

$$[v]_{B_3} = M_{B_3\,,B_2} \cdot [v]_{B_2} = M_{B_3\,,B_2} \cdot (M_{B_2\,,B_1} \cdot [v]_{B_1}) = (M_{B_3\,,B_2} \cdot M_{B_2\,,B_1}) \cdot [v]_{B_1}\,.$$

Esta última identidad muestra que $M_{B_3\,,B_2} \cdot M_{B_2\,,B_1}$ transforma coordenadas con respecto B_1 a coordenadas con respecto a B_3, de forma que se verifica la identidad propuesta en el teorema (por unicidad de la matriz cambio de base).

□

3 Transformaciones Lineales

Definición 3.1. Sean $\mathcal{V}$ y $\mathcal{W}$ dos K-ev's. Una función $T : \mathcal{V} \to \mathcal{W}$ es **transformación lineal** si satisface:

1. $T(u + v) = T(u) + T(v)$, para todos u, $v \in \mathcal{V}$;
2. $T(\alpha\, u) = \alpha\, T(u)$, para todos $u \in \mathcal{V}$ y $\alpha \in K$.

□

Obs 3.2. Con las notaciones de la definición anterior, notemos que las sumas que aparecen en el ítem 1 no son las mismas: formalmente, se debe verificar

$$T(u +_{\mathcal{V}} v) = T(u) +_{\mathcal{W}} T(v)$$

es decir, la suma entre u y v se realiza en $\mathcal{V}$, mientras que la suma de $T(u)$ y $T(v)$ se realiza en $\mathcal{W}$. Sin embargo, en adelante omitimos los subíndices para las sumas (pero tenemos en cuenta esta observación). Si una función T satisface el ítem 1. decimos que T respeta sumas.

De forma análoga, si una función T satisface el ítem 2. decimos que T respeta la acción escalar. Notemos que para que la afirmación del ítem 2 tenga sentido, $\mathcal{V}$ y $\mathcal{W}$ deben ser espacios vectoriales sobre el **mismo cuerpo**.

Finalmente, notamos que si T es transformación lineal, entonces usando inducción en la cantidad de términos, se verifica que

$$T(\alpha_1\ v_1 + \ldots + \alpha_n\ v_n) = \alpha_1\ T(v_1) + \ldots + \alpha_n\, T(v_n)$$

para todos $v_1, \ldots, v_n \in \mathcal{V}$ y $\alpha_1, \ldots, \alpha_n \in K$ (ejercicio). □

Ejemplos 3.3. Sea $\mathcal{V}$ un K-ev. Entonces se pueden verificar las siguientes afirmaciones (ejercicio)

1. La función identidad $I : \mathcal{V} \to \mathcal{V}$ se una transfomación lineal.

2. Si $K[x]$ denota el anillo de polinomios sobre K entonces la *derivación* $D : K[x] \to K[x]$ dada formalmente por

$$D(a_0 + a_1\, x + a_2\, x^2 \ldots + a_n\, x^n) = a_1 + 2\, a_2\, x + \ldots + n\, a_n\, x^{n-1}$$

   es una transformación lineal (notación: si $n \in \mathbb{N}$ y $a \in K$ entonces $n\, a = a + \ldots + a \in K$ donde hay n términos en la suma).

3. Dados $n,\, m \in \mathbb{N}$, si consideramos $K^{m\times n}$, $K^n = K^{n\times 1}$ y $K^m = K^{m\times 1}$ como K-ev's con la estructura usual, entonces: dada $A \in K^{m\times n}$ podemos definir la función $T_A : K^n \to K^m$ dada por

$$T_A(\vec{x}) = A\, \vec{x} \in K^m \quad \text{para} \quad \vec{x} \in K^n$$

   (producto matriz-vector). En este caso T_A es una transformación lineal. Notemos que la afirmación anterior (que implica verificar los ítems 1 y 2 de la Definición 3.1) es equivalente a las identidades: $\forall \vec{x},\, \vec{y} \in K^n$, $\forall \alpha \in K$,

$$A\,(\vec{x} + \vec{y}) = A\, \vec{x} + A\, \vec{y} \quad \text{y} \quad A(\alpha\, \vec{x}) = \alpha\, A\vec{x}$$

   que son propiedades del producto matriz vector conocidas de Álgebra I.

4. De forma similar, dada $A \in K^{m\times n}$ podemos definir las funciones $L_A : K^{n\times p} \to K^{m\times p}$ dada por

$$L_A(B) = A \cdot B \in K^{m\times p} \quad \text{para} \quad B \in K^{n\times p}\,,$$

   y $R_A : K^{p\times m} \to K^{p\times n}$ dadas por

$$R_A(C) = C \cdot A \in K^{p\times n} \quad \text{para} \quad C \in K^{p\times m}\,,$$

   donde hemos considerado en los dos casos productos matriz-matriz. Notemos entonces que el dominio y codominio de L_A (multiplicación a izquierda (*left* del inglés) por A) son los que corresponden para que el producto de matrices esté bien definido. Una observación similar se aplica a R_A (multiplicación a derecha (*right* del inglés) por A).

□

Lema 3.4. *Sea $\mathcal{V}$ un K-ev, con base ordenada $B = \{v_1, \ldots, v_n\}$. Consideremos la función C_B : $\mathcal{V} \to K^n$ dada por*

$$C_B(v) = [v]_B \in K^n \quad \textit{para} \quad v \in \mathcal{V}\,.$$

Entonces C_B es una transformación lineal.

Demostración Recordemos que la notación $[v]_B \in K^n$ se refiere al vector de coordenadas de v con respecto a la base B. Recordemos también que las coordenadas de un vector con respecto a una base son únicas (ver el segundo apunte sobre bases).

Para verificar que C_B es transformación lineal, verificamos los ítemos 1 y 2 de la Definición 3.1. Así, consideramos $u,\, v \in \mathcal{V}$: si $C_B(u) = [u]_B = (\alpha_j)_{j=1}^n \in K^n$ entonces, por definición

$$u = \alpha_1\, v_1 + \ldots + \alpha_n\, v_n\,.$$

Análogamente, sea $C_B(v) = [v]_B = (\beta_j)_{j=1}^n \in K^n$ de forma que $v = \beta_1\, v_1 + \ldots + \beta_n\, v_n$. Así, por las propiedades de la suma y la accion escalar en $\mathcal{V}$

$$u + v = (\alpha_1\, v_1 + \ldots + \alpha_n\, v_n) + (\beta_1\, v_1 + \ldots + \beta_n\, v_n) = (\alpha_1 + \beta_1)\, v_1 + \ldots + (\alpha_n + \beta_n)\, v_n\,.$$

A partir de la identidad anterior, y por la unicidad de las coordenadas con respecto a una base (en este caso, la base B) concluimos que

$$[u + v]_B = (\alpha_1 + \beta_1, \ldots, \alpha_n + \beta_n) \in K^n\,.$$

Finalmente, notemos que

$$C_B(u + v) = (\alpha_1 + \beta_1, \ldots, \alpha_n + \beta_n) = (\alpha_1, \ldots, \alpha_n) + (\beta_1, \ldots, \beta_n) = C_B(u) + C_B(v)\,.$$

Notemos que usamos la definición de suma en K^n (suma entrada a entrada) en una de las identidades de más arriba.

De forma análoga, si $\alpha \in K$, se verifica que $C_B(\alpha\, u) = [\alpha\ u]_B = (\alpha\, \alpha_1, \ldots, \alpha\, \alpha_n)$ (ejercicio). Así,

$$C_B(\alpha\, u) = (\alpha\, \alpha_1, \ldots, \alpha\, \alpha_n) = \alpha\, (\alpha_1, \ldots, \alpha_n) = \alpha\, C_B(u)\,.$$

Notemos que usamos la definición de acción escalar en K^n (producto entrada a entrada) en una de las identidades de más arriba. $\square$

Definición 3.5. En lo que sigue vamos a considerar una notación conveniente en este contexto, conocida como la *delta de Kronecker*. En este contexto, dados dos índices "i" y "j" consideramos la expresión

$$\delta_{ij} = \begin{cases} 1 & \text{si } \ i = j \\ 0 & \text{si } i \neq j \end{cases} \tag{6}$$

$\square$

La delta de Kronecker tiene cierto sentido simbólico. Por ejemplo, el 0 y 1 resultan, en general, ser elementos del cuerpo en consideración, sin más indicación.

Como un ejemplo de la utilidad ó conveniencia de la notación de Kronecker, consideremos un cuerpo K y la base canónica de K^n dada por $B_c = \{e_1, \ldots, e_n\}$. En este caso, con la notación de Kronecker podemos dar una descripción sencilla de los elementos de B_c: en efecto, si $1 \leq i \leq n$ entonces (recordar que e_i tiene un uno en la i-ésima entrada y todas las demás entradas son nulas)

$$e_i = (\delta_{ij})_{j=1}^n = (\delta_{i1}, \ldots, \delta_{in})\,.$$

Hacer ejemplos en $\mathbb{R}^3$ (ejercicio).

El siguiente resultado es fundamental en el estudio de las transformaciones lineales. En particular, muestra la abundancia de tales transformaciones.

Teorema 3.6. *Sean $\mathcal{V}$ y $\mathcal{W}$ K-ev's, con $\dim_K \mathcal{V} = n$. Sea $B = \{v_1, \ldots, v_n\}$ una base de $\mathcal{V}$ y sean $w_1, \ldots, w_n \in \mathcal{W}$ vectores cualesquiera. Entonces existe una única transformación lineal $T : \mathcal{V} \to \mathcal{W}$ tal que $T(v_i) = w_i$ para todo $1 \leq i \leq n$.*

Demostración. Comenzamos probando la existencia de T como arriba. Recordemos que una transformación lineal debe ser una *función*, que además tiene ciertas propiedades. Comenzamos entonces construyendo una función. Así, dado $v \in \mathcal{V}$ arbitrario, consideramos $[v]_B = (\alpha_j)_{j=1}^n \in K^n$ (el único vector de coordenadas de v con respecto a la base B del enunciado). En este caso, definimos el valor de T en v como sigue:

$$T(v) = \sum_{j=1}^{n} \alpha_j \ w_j = \alpha_1 \, w_1 + \ldots + \alpha_n \, w_n \in \mathcal{W}$$

Notemos que esto determina una función $T : \mathcal{V} \to \mathcal{W}$ bien definida. Nos queda verificar que T satisface todas las propiedades requeridas.

T es transformación lineal: en efecto, veamos que T verifica los ítems 1 y 2. de la Definición 3.1. Así, sean $u, v \in \mathcal{V}$ arbitrarios (es decir, cualesquiera) y sean $[u]_B = (\alpha_j)_{j=1}^n$, $[v]_B = (\beta_j)_{j=1}^n$ y $[u+v]_B \in K^n$ los vectores de coordenadas de u, v, $u+v$ con respecto a la base B. Por el lema anterior, se verifica que

$$[u+v]_B = [u]_B + [v]_B = (\alpha_j + \beta_j)_{j=1}^n \, .$$

Por definición de T, se tiene que

$$T(u+v) = \sum_{j=1}^{n} (\alpha_j + \beta_j) \, w_j = \sum_{j=1}^{n} (\ \alpha_j \, w_j + \beta_j \, w_j \ ) = \sum_{j=1}^{n} \alpha_j \, w_j + \sum_{j=1}^{n} \beta_j \, w_j = T(u) + T(v)$$

donde hemos usado la propiedad de la acción escalar, la asociatividad y conmutatividad de la suma de $\mathcal{W}$ y la definición de T para los valores $T(u)$ y $T(v)$ en la última identidad.

De forma similar dado $\alpha \in K$ entonces, por el lema anterior, se verifica $[\alpha \, u]_B = (\alpha \, \alpha_j)_{j=1}^n$. Entonces,

$$T(\alpha \, u) = \sum_{j=1}^{n} (\alpha \, \alpha_j) \ w_j = \sum_{j=1}^{n} \alpha \, (\alpha_j \ w_j) = \alpha \sum_{j=1}^{n} \alpha_j \ w_j = \alpha \, T(u)$$

donde hemos usado propiedades de la acción escalar y la definición de T.

Además, si $1 \leq i \leq n$ entonces: $[v_i]_B = e_i = (\delta_{ij})_{j=1}^n$. Entonces, por definición de T

$$T(v_i) = \sum_{j=1}^{n} \delta_{ij} \ w_j = 0 \, w_1 + \ldots + 0 \, w_{i-1} + 1 \, w_i \ + 0 \ w_{i+1} + \ldots + 0 \, w_n = w_i$$

Finalmente, probamos la unicidad de T. Para ello, supongamos que $T' : \mathcal{V} \to \mathcal{W}$ es otra transformación lineal tal que $T'(v_i) = w_i$ para todo $1 \leq i \leq n$ y verificamos que $T = T'$. Notemos que verificar que las funciones T y T' son iguales corresponde a probar que para todo $v \in \mathcal{V}$ vale que $T'(v) = T(v)$ (dado que T' y T tienen el mismo dominio y codominio por hipótesis). Sea $v \in \mathcal{V}$ arbitrario, y consideremos $[v]_B = (\beta_j)_{j=1}^n \in K^n$. Entonces

$$T'(v) = T'(\sum_{j=1}^{n} \beta_j \ v_j) = \sum_{j=1}^{n} \beta_j \ T'(v_j) = \sum_{j=1}^{n} \beta_j \ w_j$$

y de forma similar

$$T(v) = T(\sum_{j=1}^{n} \beta_j \ v_j) = \sum_{j=1}^{n} \beta_j \ T(v_j) = \sum_{j=1}^{n} \beta_j \ w_j \, ,$$

donde hemos usado la linealidad de T' y T, la hipótesis sobre T' y la propiedad de T: $T'(v_j) = w_j = T(v_j)$. Las identidades anteriores muestran que $T'(v) = T(v)$, para $v \in \mathcal{V}$ arbitrario. Entonces concluimos que $T' = T$. $\square$

Corolario 3.7. *Sean $\mathcal{V}$ y $\mathcal{W}$ K-ev, con $\dim_K \mathcal{V} = n$. Sea $B_{\mathcal{V}} = \{v_1, \ldots, v_n\}$ una base de $\mathcal{V}$ y sean $T_1, T_2 : \mathcal{V} \to \mathcal{W}$ transformaciones lineales tales que $T_1(v_j) = T_2(v_j)$, para $1 \leq j \leq n$. Entonces $T_1 = T_2$ es decir, para todo $v \in \mathcal{V}$ se verifica que $T_1(v) = T_2(v)$.*

Demostración Se deduce de la unicidad del teorema anterior (completar los dealles como ejercicio). $\square$

Un hecho importante acerca de la primera parte de la prueba del Teorema 3.6 es que nos indica un método para construir transformaciones lineales. Consideremos el siguiente

Ejemplo 3.8. Sea $\mathcal{V} = \mathbb{R}^2$ como $\mathbb{R}$-ev. Sea $B = \{v_1, v_2\}$ la base ordenada dada por

$$v_1 = (1, 1) \quad \text{y} \quad v_2 = (0, 2)\,.$$

Verificar que B es base (ejercicio). Consideremos $\mathcal{W} = \mathbb{R}^3$ y los vectores $w_1 = (1, 2, 3)$, $w_2 = (4, 5, 6)$. Entonces podemos calcular la única transformación lineal $T : \mathbb{R}^2 \to \mathbb{R}^3$ tal que $T(v_i) = w_i$, $i = 1, 2$, como sigue.

Dado $v = (x, y) \in \mathbb{R}^2$ entonces vale que $[(x, y)]_B = (x\,, \frac{y}{2} - x) \in \mathbb{R}^2$ (verificar, ejercicio). Así, construimos T siguiendo el argumento de la primer parte del Teorema 3.6:

$$\begin{aligned}
T(x\,, y) &= x\ (1,2,3) + (\frac{y}{2} - x)\ (4,5,6) = (x\,, 2x\,, 3x) + (4(\frac{y}{2} - x)\,, 5\,(\frac{y}{2} - x)\,, 6\,(\frac{y}{2} - x)) \\
&= (x + 4(\frac{y}{2} - x)\,, 2x + 5(\frac{y}{2} - x)\,, 3x + 6(\frac{y}{2} - x)) \\
&= (2y - 3x\,, \frac{5}{2}y - 3x\,, 3y - 3 - x)\,.
\end{aligned}$$

Ahora podemos verificar que T es una transformación lineal tal que $T(v_i) = w_i$, $i = 1, 2$ (ejercicio). $\square$

Cada transformación lineal determina dos subespacios distinguidos, que definimos a continuación.

Definición 3.9. *Sean $\mathcal{V}$ y $\mathcal{W}$ K-ev's y sea $T : \mathcal{V} \to \mathcal{W}$ una transformación lineal.*

1. *Definimos el núcleo de T, notado $N(T) \subseteq \mathcal{V}$, dado por*

$$N(T) = \{v \in \mathcal{V} :\ T(v) = 0_{\mathcal{W}}\}\,.$$

2. *La imagen de T, notada $Im(T) \subseteq \mathcal{W}$, dada por*

$$Im(T) = \{Tv :\ v \in \mathcal{V}\}\,.$$

$\square$

Con las notaciones de la definición anterior, vemos que $\mathrm{Im}(T)$ es la imágen de T como función (en el sentido de Álgebra I). Por otro lado $N(T) = T^{-1}(0_{\mathcal{W}})$.

Proposición 3.10. *Sean $\mathcal{V}$ y $\mathcal{W}$ K-ev's y sea $T : \mathcal{V} \to \mathcal{W}$ una transformación lineal. Entonces $N(T)$ es subespacio de $\mathcal{V}$ y $Im(T)$ es subespacio de $\mathcal{W}$.*

Demostración. Veamos que $N(T)$ es subespacio de $\mathcal{V}$ (verificamos las tres propiedades de subespacio). En efecto, notemos que

$$T(0_{\mathcal{V}}) = T(0_{\mathcal{V}} + 0_{\mathcal{V}}) = T(0_{\mathcal{V}}) + T(0_{\mathcal{V}}) \implies 0_{\mathcal{W}} = T(0_{\mathcal{V}})$$

donde la cancelación se realiza en $\mathcal{W}$ (en adelante no vamos a incluir la subíndices, pero hay que tener en cuenta cual es el cero en consideración). Esto muestra que $T(0) = 0$ con lo que $0 \in N(T)$.

Sean u, $v \in N(T)$. Así, por hipótesis, $T(u) = 0$ y $T(v) = 0$. Entonces

$$T(u + v) = T(u) + T(v) = 0 + 0 = 0$$

donde hemos usado la linealidad de T y la hipótesis. Así, $T(u + v) = 0$ y $u + v \in N(T)$.

Si $u \in N(T)$ y $\alpha \in K$ entonces

$$T(\alpha\, u) = \alpha\, T(u) = \alpha\, 0 = 0$$

de forma que $\alpha\, u \in N(T)$. Las observaciones anteriores muestran que $N(T) \subseteq \mathcal{V}$ es subespacio.

La prueba de que $\mathrm{Im}(T)$ es subespacio de $\mathcal{W}$ es similar y queda como ejercicio! □

Definición 3.11. Sean $\mathcal{V}$ y $\mathcal{W}$ K-ev's y sea $T : \mathcal{V} \to \mathcal{W}$ una transformación lineal. Si $\dim_K \mathcal{V} \in \mathbb{N}$ definimos la *nulidad* de T, notada $\mathrm{nul}(T)$ dada por

$$\mathrm{nul}(T) = \dim_K \mathrm{N}(T)$$

y el *rango* de T, notado $\mathrm{rg}(T)$ dado por

$$\mathrm{rg}(T) = \dim_K \mathrm{Im}(T)\,.$$

Con las notaciones de la definición anterior, el próximo resultado muestra que tanto la nulidad como el rango de T están bien definidas, en el sentido que los subespacios en cuestión tienen dimensión finita.

Teorema 3.12. *Sean $\mathcal{V}$ y $\mathcal{W}$ K-ev's tal que $\dim_K \mathcal{V} \in \mathbb{N}$. Dada $T : \mathcal{V} \to \mathcal{W}$ una transformación lineal, se tiene que*

$$\dim_K \mathcal{V} = \dim_K N(T) + \dim_K Im(T)\,.$$

Demostración. En lo que sigue, vamos a construir bases (finitas) del núcleo de T y de la imágen de T. Una vez hecho esto, podemos verificar la afirmación del teorema contando la cantidad de elementos en las bases correspondientes.

Sea $\dim_K \mathcal{V} = n \in \mathbb{N}$; como $\mathrm{N}(T) \subseteq \mathcal{V}$ es subespacio, hemos visto que $\dim \mathrm{N}(T) = k \leq n$. En este caso, podemos considerar una base $B_1 = \{v_1, \ldots, v_k\} \subset \mathrm{N}(T)$ de $\mathrm{N}(T)$. Como, en particular, $B_1 = \{v_1, \ldots, v_k\} \subset \mathcal{V}$ es conjunto l.i., podemos extender B_1 a una base $B = \{v_1, \ldots, v_n\}$ de $\mathcal{V}$.

Afirmamos que el conjunto ordenado $B_2 = \{T(v_{k+1}), \ldots, T(v_n)\} \subset \mathrm{Im}(T)$ es una base de $\mathrm{Im}(T)$. Para ello, verificamos que B_2 es l.i. y que genera a $\mathrm{Im}(T)$:

B_2 es l.i.: en efecto, supongamos que $\alpha_{k+1}, \ldots, \alpha_n \in K$ son tales que

$$\alpha_{k+1}\, T(v_{k+1}) + \ldots + \alpha_n\, T(v_n) = 0\,.$$

Entonces, por la linealidad de T,

$$0 = \alpha_{k+1}\, T(v_{k+1}) + \ldots + \alpha_n\, T(v_n) = T(\alpha_{k+1}\, v_{k+1} + \ldots + \alpha_n\, v_n)$$

La identidad anterior indica que el vector

$$v = \alpha_{k+1}\, v_{k+1} + \ldots + \alpha_n\, v_n \in \mathrm{N}(T)\,.$$

Entonces, existen coeficientes $\alpha_1, \ldots, \alpha_k \in K$ tales que

$$v = \alpha_1\, v_1 + \ldots + \alpha_k\, v_k$$

pues $B_1 = \{v_1, \ldots, v_k\}$ es base de $\mathrm{N}(T)$. Pero entonces

$$\alpha_{k+1}\, v_{k+1} + \ldots + \alpha_n\, v_n = v = \alpha_1\, v_1 + \ldots + \alpha_k\, v_k$$

$$\implies (-\alpha_1)\, v_1 + \ldots + (-\alpha_k)\, v_k + \alpha_{k+1}\, v_{k+1} + \ldots + \alpha_n\, v_n = 0\,.$$

Como $B = \{v_1, \ldots,\, v_n\}$ es base de $\mathcal{V}$, la independencia lineal de B y la ecuación anterior muestran que

$$-\alpha_1 = \ldots = -\alpha_k = \alpha_{k+1} = \ldots = \alpha_n = 0\,.$$

En particular, $\alpha_{k+1} = \ldots = \alpha_n = 0$ lo que indica que B_2 es l.i.

B_2 genera $\mathrm{Im}(T)$: en efecto, sea $w \in \mathrm{Im}(T)$ arbitrario. Entonces, por definición de imágen, existe $v \in \mathcal{V}$ tal que $w = Tv$. En este caso, como B es base de $\mathcal{V}$, existen coeficientes $\beta_1, \ldots, \beta_n \in K$ tales que

$$v = \sum_{j=1}^{n} \beta_j\, v_j \implies w = T(v) = T(\sum_{j=1}^{n} \beta_j\, v_j) = \sum_{j=1}^{n} \beta_j\, T(v_j) = \sum_{j=k+1}^{n} \beta_j\, T(v_j)\,,$$

donde hemos usado la representación de v, la linealidad de T y el hecho de que $B_1 = \{v_1, \ldots, v_k\} \subset \mathrm{N}(T)$ (porque es base de $\mathrm{N}(T)$), de forma

$$T(v_1) = \ldots = T(v_k) = 0\,.$$

Lo anterior muestra que $w = \sum_{j=k+1}^{n} \beta_j\, T(v_j)$ y B_2 resulta un sistema de generadores de $\mathrm{Im}(T)$.

De esta forma B_2 es base de $\mathrm{Im}(T)$. Recapitulando: hemos construido $B_1 = \{v_1, \ldots, v_k\}$ base de $\mathrm{N}(T)$ y $B_2 = \{T(v_{k+1}), \ldots,\, T(v_n)\} \subset \mathrm{Im}(T)$ base de $\mathrm{Im}(T)$. Ahora vemos que

$$\dim_K(\mathrm{N}(T)) = k \quad \text{y} \quad \dim_K(\mathrm{Im}(T)) = n - k$$

$$\implies \dim_k \mathcal{V} = n = \dim_K \mathrm{N}(T) + \dim_K \mathrm{Im}(T)\,.$$

□

Obs 3.13. Con las notaciones del Teorema 3.12, si además suponemos que $\dim_K \mathcal{W} = m$ entonces, se tienen los siguientes hechos (cuya verificación es ejercicio):

1. Si $n > m$ entonces $\mathrm{N}(T) \neq \{0\}$ es decir, $\dim_K(\mathrm{N}(T)) \geq 1$.
2. Si $n < m$ entonces no existe una transformación lineal suryectiva (como función), es decir, $\mathrm{Im}(T) \subsetneq \mathcal{W}$ (inclusión propia) ó equivalentemente $\dim_K \mathrm{Im}(T) < m$.

□

3.1 El espacio de las transformaciones lineales

En lo que sigue, dados dos espacios vectoriales sobre un mismo cuerpo construimos un nuevo espacio, formado por las transformaciones lineales de uno de estos espacios en el otro.

Definición 3.14. Sean $\mathcal{V}$ y $\mathcal{W}$ K-ev's tal que $\dim_K \mathcal{V} = n \in \mathbb{N}$. Notamos $L(\mathcal{V}, \mathcal{W})$ al conjunto de todas las transformaciones lineales $T : \mathcal{V} \to \mathcal{W}$, es decir

$$L(\mathcal{V}, \mathcal{W}) = \{T : \mathcal{V} \to \mathcal{W} \;/\; T \text{ es transformación lineal } \}$$

□

En lo que sigue, le damos estructura de espacio vectorial a $L(\mathcal{V}, \mathcal{W})$

Teorema 3.15. *Sean $\mathcal{V}$ y $\mathcal{W}$ K-ev's tal que $\dim_K \mathcal{V} = n \in \mathbb{N}$. Dotamos a $L(\mathcal{V}, \mathcal{W})$ de una suma y una acción escalar (puntual) como sigue: dados $S, T \in L(\mathcal{V}, \mathcal{W})$ y $\alpha \in K$ entonces definimos*

1. $S + T : \mathcal{V} \to \mathcal{W}$ *dada por* $S + T\,(v) = S(v) + T(v) \quad , \quad v \in \mathcal{V}$;
2. $\alpha\, S : \mathcal{V} \to \mathcal{W}$ *dada por* $\alpha S\,(v) = \alpha\, S(v) \quad , \quad v \in \mathcal{V}$.

La suma y la acción escalar definidas más arriba hacen de $L(\mathcal{V}, \mathcal{W})$ un K-ev.

Antes de comenzar con la prueba del teorema, consideremos con más detalle las definiciones propuestas y el objetivo del teorema. Dadas $S, T \in L(\mathcal{V}, \mathcal{W})$ entonces definimos una nueva función $S + T : \mathcal{V} \to \mathcal{W}$; para determinar la función $S + T$ debemos indicar su valor en los elementos de su dominio $\mathcal{V}$. Para eso, dado $v \in \mathcal{V}$ definimos el valor de $S + T$ en v, notado $S + T(v)$, como el vector de $\mathcal{W}$ dado por la suma de los valores $S(v) + T(v)$. De esta forma, determinamos una función $S + T$ bien definida. El primer paso a dar es verificar que la función $S + T$ es una transformación lineal, es decir $S + T \in L(\mathcal{V}, \mathcal{W})$. Luego, como S y T eran cualesquiera transformaciones, lo anterior define una operación binaria en $L(\mathcal{V}, \mathcal{W})$. Formalmente, debiéramos escribir algo así como $S +_{L(\mathcal{V},\mathcal{W})} T$, porque se trata de una operación binaria en $L(\mathcal{V}, \mathcal{W})$, pero como esto más bien engorroso, omitimos el subíndice (pero tenemos en cuenta esta consideración).

El teorema afirma que la operación binaria definida más arriba tiene las propiedades que corresponden a la suma de vectores en un espacio vectorial, que es lo que debemos verificar.

Consideraciones similares se aplican en el caso de la definición de la acción escalar en $L(\mathcal{V}, \mathcal{W})$, que el lector debe pensar. Ahora sí, vamos a la prueba

Demostración. Antes que nada, debemos verificar que la operación binaria definida más arriba es en realidad *cerrada* en $L(\mathcal{V}, \mathcal{W})$; esto es, si $S, T \in L(\mathcal{V}, \mathcal{W})$ entonces $S + T \in L(\mathcal{V}, \mathcal{W})$. Para ello, si $u, v \in \mathcal{V}$ notemos

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

donde usamos la definición de $S + T$ en $L(\mathcal{V}, \mathcal{W})$, el hecho de que S y T son transformaciones lineales, propiedades de la suma en $\mathcal{W}$ y finalmente, otra vez la definición de $S + T$ en $L(\mathcal{V}, \mathcal{W})$: en conclusión

$$S + T(u + v) = \; S + T\,(u) + S + T\,(v)\,.$$

De forma análoga se verifica que $S + T(\alpha\, u) = \alpha\; ((S + T)(u))$ (ejercicio); lo anterior muestra que $S + T \in L(\mathcal{V}, \mathcal{W})$.

Consideraciones similares valen para la acción escalar, que resulta una acción bien definida $(\alpha\,,\, T) \mapsto \alpha\, T \in L(\mathcal{V}, \mathcal{W})$ (ejercicio para el lector).

Vamos a repasar algunas de las propiedades que debe satisfacer una operación binaria para ser la suma de vectores en un espacio vectorial y verificar que se satisfacen en el caso de la operación binaria definida en $L(\mathcal{V}, \mathcal{W})$.

Por un lado, la suma debe ser asociativa: es decir, dadas $R, S, T \in L(\mathcal{V}, \mathcal{W})$ entonces se debe verificar

$$R + (S + T) = (R + S) + T$$

Como los elementos de $L(\mathcal{V}, \mathcal{W})$ son funciones, se debe verificar una igualdad de funciones. Como las funciones en cuestión tienen el mismo dominio y codominio, debemos verificar que las funciones tienen el mismo valor en cada elemento de $\mathcal{V}$. Para eso, tomamos $v \in \mathcal{V}$: notemos que por definición de la operación binaria en $L(\mathcal{V}, \mathcal{W})$

$$R + (S + T)\,(v) = R(v) + (S + T)(v) = R(v) + (S(v) + T(v))\,.$$

De forma similar

$$(R+S)+T\,(v)=(R+S)(v)+T(v)=(R(v)+S(v))+T(v)\,.$$

Como la suma de vectores en $\mathcal{W}$ es asociativa, a partir de las identidades anteriores concluimos que

$$R+(S+T)\,(v)=(R+S)+T\,(v) \quad \text{para} \quad v\in\mathcal{V}\,,$$

es decir, vale la igualdad entre funciones $R+(S+T)=(R+S)+T$.

De forma análoga se prueba la conmutatividad.

Con respecto a la existencia de elemento neutro para la suma, definimos la transformación lineal nula

$$0:\mathcal{V}\to\mathcal{W} \quad \text{dada por} \quad 0(v)=0_{\mathcal{W}} \quad , \quad v\in\mathcal{V}\,.$$

Es sencillo verificar que $0\in L(\mathcal{V},\mathcal{W})$. Además, si $S\in L(\mathcal{V},\mathcal{W})$ entonces

$$0+S\,(v)=0(v)+S(v)=0+S(v)=S(v) \quad , \quad v\in\mathcal{V}\,.$$

Entonces, vemos que $0+S=S$ como funciones (porque tienen el mismo valor en cada elemento de $\mathcal{V}$). Como antes, debiéramos escribir algo así como $0_{L(\mathcal{V},\mathcal{W})}$, pero omitimos el subíndice; en este caso, hay que saber determinar del contexto cual es el cero que estamos considerando (transformación nula ó vector nulo).

Con respecto al opuesto: dada $T\in L(\mathcal{V},\mathcal{W})$, definimos $-T:\mathcal{V}\to\mathcal{W}$ dada por

$$(-T)(v)=-T(v)\ ,\ \ v\in\mathcal{V}\,.$$

En palabras, la transformación $-T$ evaluada en el vector v tiene el valor $-T(v)$ (es decir, el opuesto del vector $T(v)$ en $\mathcal{W}$). Es sencillo verificar que $-T\in L(\mathcal{V},\mathcal{W})$ y que además satisface $T+-T=0$ (lo que queda como ejercicio para el lector).

Para concluir, verificamos una de las propiedades del producto escalar, y las propiedades restantes quedan como ejercicio para el lector. Dado $\alpha\in K$ y $S,\,T\in L(\mathcal{V},\mathcal{W})$, veamos que se verifica

$$\alpha\,(S+T)=\alpha\,S+\alpha\,T$$

Como antes, se trata de verificar una igualdad entre funciones: para eso, tomemos $v\in\mathcal{V}$ arbitrario. Entonces

$$\begin{aligned}[\alpha\,(S+T)]\ (v) &= \alpha\,[(S+T)(v)]=\alpha\,(S(v)+T(v))=\alpha\,S(v)+\alpha\,T(v)\\ &= [\alpha\,S](v)+[\alpha\,T](v)=(\alpha\,S+\alpha\,T)(v)\,,\end{aligned}$$

donde hemos usado la definción de la acción en $L(\mathcal{V},\mathcal{W})$, la definición de la suma en $L(\mathcal{V},\mathcal{W})$, propiedades de la suma y la acción escalar en $\mathcal{W}$, la definición de la acción escalar en $L(\mathcal{V},\mathcal{W})$ y finalmente, la definición de la suma en en $L(\mathcal{V},\mathcal{W})$ (otra vez). □

3.2 Sobre la dimensión de $L(\mathcal{V},\mathcal{W})$

Con las notaciones del teorema anterior, una vez que dotamos a $L(\mathcal{V},\mathcal{W})$ de una estructura de K-ev, surge una pregunta natural. Si $\dim_K\mathcal{V}=n$ y $\dim_K\mathcal{W}=m$, cual es la dimensión $\dim_K L(\mathcal{V},\mathcal{W})$? El siguiente resultado responde esta pregunta

Teorema 3.16. *Sean* $\mathcal{V}$, $\mathcal{W}$ K*-ev's tales que* $\dim_K\mathcal{V}=n$ *y* $\dim_K\mathcal{W}=m$. *Entonces* $\dim_K L(\mathcal{V},\mathcal{W})=n\cdot m$.

Demostración. Para calcular $\dim_K L(\mathcal{V},\mathcal{W})$ construimos una base de este espacio. Para ello, comenzamos considerando bases $B_{\mathcal{V}} = \{v_1, \ldots, v_n\}$ y $B_{\mathcal{W}} = \{w_1, \ldots, w_m\}$ de $\mathcal{V}$ y $\mathcal{W}$ respectivamente. En el argumento que sigue vamos a usar sistemáticamente el Teorema 3.6 de más arriba, sobre la existencia (y unicidad) de transformaciones lineales (notar que por hipótesis $B_{\mathcal{V}} = \{v_1, \ldots, v_n\}$ es base de $\mathcal{V}$).

Dados $1 \le p \le m$ y $1 \le q \le n$ definimos $E^{(p,q)} \in L(\mathcal{V},\mathcal{W})$ como la única transformación lineal que satisface: para $1 \le j \le n$,

$$E^{(p,q)}(v_j) = \begin{cases} w_p & \text{si } \ j = q\,; \\ 0_{\mathcal{W}} & \text{si } j \neq q\,. \end{cases} \tag{7}$$

Por ejemplo, $E^{(2,1)} \in L(\mathcal{V},\mathcal{W})$ es la única transformación que satisface

$$E^{(2,1)}(v_1) = w_2\,,\ E^{(2,1)}(v_2) = 0\,,\ \ldots\,,\ E^{(2,1)}(v_n) = 0\,.$$

Con la (conveniente) notación de Kronecker (ver la 1er parte del apunte para el tp3), también podemos escribir

$$E^{(p,q)}(v_j) = \delta_{jq}\ w_p \quad \text{para} \quad 1 \le j \le n\,, \tag{8}$$

dado que si $j = q$ entonces $\delta_{jq}\ w_p = 1 \cdot w_p = w_p$, mientras que si $j \neq q$ entonces $\delta_{jq}\ w_p = 0 \cdot w_p = 0$.

Notemos que el proceso anterior permite construir $m \cdot n$ transformaciones lineales. Definimos

$$B = \{E^{(p,q)}\ :\ 1 \le p \le m\,,\ 1 \le q \le n\ \} \subset L(\mathcal{V},\mathcal{W})\,,$$

que es un conjunto con $m \cdot n$ elementos. En lo que sigue, probamos que B es una base de $L(\mathcal{V},\mathcal{W})$. Notemos que una vez que sepamos que B es base, entonces podemos calcular la dimensión de $L(\mathcal{V},\mathcal{W})$ como la cantidad de elementos de B, lo que demuestra el teorema. Ahora sí, seguimos con el argumento.

B genera a $L(\mathcal{V},\mathcal{W})$: en efecto, sea $T \in L(\mathcal{V},\mathcal{W})$ arbitraria, pero fija. Para cada $1 \le j \le n$ consideramos

$$[T(v_j)]_{B_{\mathcal{W}}} = (a_{1j}\,,\ a_{2j}\,,\ \ldots\,,\ a_{mj}) \in K^m \tag{9}$$

el vector de coordenadas del valor $T(v_j) \in \mathcal{W}$ con respecto a la base $B_{\mathcal{W}}$. Notemos que por definición, se verifica que

$$T(v_j) = a_{1j}\, w_1 + \ldots + a_{mj}\, w_m = \sum_{p=1}^{m} a_{p\,j}\, w_p \quad \text{para} \quad 1 \le j \le n\,.$$

Consideramos ahora, la combinación lineal (en el espacio $L(\mathcal{V},\mathcal{W})$) dada por

$$\sum_{p=1}^{m}\ \sum_{q=1}^{n}\ a_{p\,q}\ E^{(p,q)} \in L(\mathcal{V},\mathcal{W})$$

con la estructura de espacio vectorial de $L(\mathcal{V},\mathcal{W})$ que hemos definido en las notas anteriores. Afirmamos que

$$T = \sum_{p=1}^{m}\ \sum_{q=1}^{n}\ a_{p\,q}\ E^{(p,q)}\,. \tag{10}$$

Como la igualdad propuesta es entre transformaciones lineales, entonces basta ver que las transformaciones tienen el mismo valor en los vectores de una base de $\mathcal{V}$ (por un corolario en las 1eras notas para el tp3).

Para ello, consideramos la base $B_{\mathcal{V}} = \{v_1, \ldots, v_n\}$. Si $1 \le j \le n$ entonces

$$[\sum_{p=1}^{m}\ \sum_{q=1}^{n}\ a_{p\,q}\ E^{(p,q)}](v_j) = \sum_{p=1}^{m}\ \sum_{q=1}^{n}\ a_{p\,q}\ E^{(p,q)}(v_j) = \sum_{p=1}^{m}\ \sum_{q=1}^{n}\ a_{p\,q}\ \delta_{jq}\ w_p \tag{11}$$

donde hemos usado la definición de las operaciones en $L(\mathcal{V},\mathcal{W})$ en la primer evaluación y la Eq. (8) para la segunda igualdad. Notemos que para la sumatoria interior en Eq. (11) vale que

$$\sum_{q=1}^{n} a_{p\,q}\ \delta_{jq}\ w_p = a_{pj}\ w_p\,.$$

En efecto: en la sumatoria el índice j está fijo, mientras que el índice q de sumación barre todos los valores entre 1 y n. Siempre que $q \neq j$ entonces el término correspondiente se anula, porque el escalar $\delta_{jq} = 0$; sólo el término correspondiente a $q = j$ en la sumatoria satisface $\delta_{jq} = 1$ y en ese caso, el término es $a_{p\,q}\ \delta_{jq}\ w_p = a_{pj}\ w_p$ (porque $q = j$!). La idea detallada en este párrafo muestra la aplicación más frecuente de la notación de la delta de Kronecker: en lo que sigue aplicaremos este razonamiento (repetidamente) sin hacer más referencia.

Reemplanzando la última identidad en la expresión Eq. (11) vemos que

$$[\sum_{p=1}^{m}\sum_{q=1}^{n} a_{p\,q}\ E^{(p,q)}](v_j) = \sum_{p=1}^{m}\sum_{q=1}^{n} a_{p\,q}\ \delta_{jq}\ w_p = \sum_{p=1}^{m} a_{p\,j}\ w_p = T(v_j) \tag{12}$$

donde la última igualdad se deduce de Eq. (9). Como la identidad en la Eq. (12) se verifica para todo $1 \leq j \leq n$, concluimos la validez de la identidad de la Eq. (10). Como $T \in L(\mathcal{V},\mathcal{W})$ era arbitraria, concluimos que B es un sistema de generadores.

B es l.i. en $L(\mathcal{V},\mathcal{W})$. Para verificar la indepencia lineal de B consideramos coeficientes $b_{p\,q} \in K$ para $1 \leq p \leq m$ y $1 \leq q \leq n$ tales que

$$\sum_{p=1}^{m}\sum_{q=1}^{n} b_{p\,q}\ E^{(p,q)} = 0\,. \tag{13}$$

En este caso, queremos verificar que los b's son todos nulos. Notemos que la identidad en la Eq. (13) es entre transformaciones lineales, donde $0 \in L(\mathcal{V},\mathcal{W})$ es la transformación nula (que satisface $0(v) = 0$ para $v \in \mathcal{V}$). Consideramos $1 \leq j \leq n$ y evaluamos

$$\begin{aligned} 0 &= 0(v_j) = [\sum_{p=1}^{m}\sum_{q=1}^{n} b_{p\,q}\ E^{(p,q)}](v_j) = \sum_{p=1}^{m}\sum_{q=1}^{n} b_{p\,q}\ E^{(p,q)}(v_j) \\ &= \sum_{p=1}^{m}\sum_{q=1}^{n} b_{p\,q}\ \delta_{jq}\ w_p = \sum_{p=1}^{m} b_{p\,j}\ w_p\,. \end{aligned}$$

Como $B_{\mathcal{W}} = \{w_1, \ldots, w_m\}$ es base de $\mathcal{W}$, en particular $B_{\mathcal{W}}$ es l.i.. Así, la ecuación

$$0 = \sum_{p=1}^{m} b_{p\,j}\ w_p$$

implica que

$$b_{1j} = \ldots = b_{mj} = 0 \quad \text{para} \quad 1 \leq j \leq n\,.$$

Esto muestra que todos los b's son nulos y B es l.i.

Como adelantamos, ahora podemos ver que $\dim_K L(\mathcal{V},\mathcal{W}) = m \cdot n$, que es la cantidad de elementos de B.

$\square$

El siguiente lema muestra que la composición de transformaciones lineales (como funciones) resulta una transformación lineal.

Lema 3.17. *Sean $\mathcal{V}_1$, $\mathcal{V}_2$ y $\mathcal{V}_3$ K-ev's y sean $S : \mathcal{V}_1 \to V_2$ y $T : \mathcal{V}_2 \to \mathcal{V}_3$ transformaciones lineales. Entonces la composición $T \circ S : \mathcal{V}_1 \to \mathcal{V}_3$ dada de la forma usual*

$$T \circ S(v) = T(S(v)) \in \mathcal{V}_3 \quad \textit{para} \quad v \in \mathcal{V}_1$$

es transformación lineal.

Demostración. Para verificar el lema probamos que la composición $T \circ S$ respeta la suma y la acción escalar. En efecto, si u, $v \in \mathcal{V}_1$ entonces

$$T \circ S(u+v) = T(S(u+v)) = T(S(u)+S(v)) = T(S(u)) + T(S(v)) = T \circ S(u) + T \circ S(v)$$

donde hemos usado la definición de composición, la linealidad de S y T, y nuevamente la definición de composición. Notemos que las sumas que aparecen más arriba en distintos miembros corresponden a distintos espacios: formalmente, debiéramos escribir

$$T(S(u +_{\mathcal{V}_1} v)) = T(S(u) +_{\mathcal{V}_2} S(v)) = T(S(u)) +_{\mathcal{V}_3} T(S(v))\,.$$

En adelante omitimos los subíndices (pero tenemos en cuenta esta observación).

Con respecto a la acción escalar: si $u \in \mathcal{V}_1$ y $\alpha \in K$,

$$T \circ S(\alpha\, u) = T(S(\alpha\, u)) = T(\alpha\, S(u)) = \alpha\, T(S(u)) = \alpha\ T \circ S(u)\,.$$

Así $T \circ S \in L(\mathcal{V}_1, \mathcal{V}_3)$. □

3.3 El álgebra $L(\mathcal{V})$

En lo que sigue consideramos un tipo importante de transformaciones, que corresponde al caso de transformaciones lineales en un mismo espacio.

Definición 3.18. *Sea $\mathcal{V}$ un K-ev. Si $T \in L(\mathcal{V}, \mathcal{V})$ decimos que T es operador lineal sobre $\mathcal{V}$ (ó que actúa en $\mathcal{V}$). En este caso notamos $L(\mathcal{V}) = L(\mathcal{V}, \mathcal{V})$ y lo llamamos el espacio de operadores lineales sobre $\mathcal{V}$.*

Como veremos más adelante, $L(\mathcal{V})$ tiene estructura adicional, de forma que $L(\mathcal{V})$ no sólo es un K-ev, sino que resulta una K-álgebra, es decir, admite además un producto. De hecho, el producto de operadores será definido como la composición. Para ello, consideramos el siguiente resultado.

Proposición 3.19. *Sea $\mathcal{V}$ un K-ev y sean S, T_1, $T_2 \in L(\mathcal{V})$ y $\alpha \in K$. Entonces se verifica:*

1. *Si $I \in L(\mathcal{V})$ denota la función identidad, entonces $S \circ I = I \circ S = S$.*
2. *$(S \circ T_1) \circ T_2 = S \circ (T_1 \circ T_2)$*
3. *$S \circ (T_1 + T_2) = S \circ T_1 + S \circ T_2$, $(T_1 + T_2) \circ S = T_1 \circ S + T_2 \circ S$.*
4. *$\alpha\,(S \circ T) = (\alpha\, S) \circ T = S \circ (\alpha\, T)$.*

Demostración. Los ítems 1. y 2. son hechos generales de funciones (notemos que por definición $I(v) = v$ para $v \in \mathcal{V}$ y la composición de funciones es asociativa).

Veamos una de las identidades del ítem 3: como se trata de una igualdad de funciones, lo verificamos tomando un $v \in \mathcal{V}$ arbitrario y evaluamos a ambos lados. En este caso

$$\begin{aligned} S \circ (T_1 + T_2)(v) &= S((T_1+T_2)(v)) = S(T_1(v) + T_2(v)) = S(T_1(v)) + S(T_2(v)) \\ &= S \circ T_1(v) + S \circ T_2(v) = (S \circ T_1 + S \circ T_2)(v) \end{aligned}$$

donde hemos usado la definición de la composición, la linealidad de S así como la definición de la suma de transformaciones lineales en dos oportunidades (cuáles?).

El ítem 4 se verifica de forma similar (ejercicio).

$\square$

Obs 3.20. En adelante, dado $\mathcal{V}$ un K-ev y S, $T \in L(\mathcal{V})$ notamos

$$T \cdot S = T \circ S$$

incluso, también notamos simplemente TS, para indicar la composición de las transformaciones. Con esta notación, las identidades de la proposición anterior pueden escribirse

$$I\,S = S\,I = S \quad , \quad S\,(T_1 + T_2) = S\,T_1 + S\,T_2$$

etc. En este sentido, la composición resulta una operación binaria bien definida en $L(\mathcal{V})$ que es compatible con la suma y que le da a $L(\mathcal{V})$ la estructura de una K-**álgebra** (ó de álgebra sobre K) con unidad. Más adelante vamos a usar sistemáticamente esta estructura para poder estudiar a los operadores $T \in L(\mathcal{V})$.

$\square$

En álgebra I hemos estudiado otros ejemplos K-álgebras con unidad sobre un cuerpo K. De hecho, éste es el caso de $K[x]$ el álgebra de polinomio sobre el cuerpo K. También hemos estudiado el álgebra de matrices cuadradas $K^{n\times n}$.

Sin embargo, $K[x]$ y $L(\mathcal{V})$ tienen una **gran diferencia**: $K[x]$ es un álgebra conmutativa, en donde se verifica que $p\,q = q\,p$ para todos p, $q \in K[x]$. Sin embargo, el producto de $L(\mathcal{V})$ no es conmutativo en general (como tampoco lo es en el álgebra de matrices $K^{n\times n}$). Veamos un

Ejemplo 3.21. Sea K un cuerpo, y sea $\mathcal{V} = K^2$, K-ev de pares ordenados con entradas en K. Sean

$$A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \quad \text{y} \quad B = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$$

Consideramos T_A, $T_B \in L(K^2)$ dados por

$$T_A(\vec{x}) = A\,\vec{x} \quad \text{y} \quad T_B(\vec{x}) = B\,\vec{x} \quad \text{para} \quad \vec{x} \in K^2\,.$$

Entonces, el lector puede verificar que

$$T_A\,T_B = T_B \quad \text{mientras que} \quad T_B\,T_A = 0\,.$$

De hecho, las identidades anteriores se pueden verificar observando que $T_A\,T_B = T_{A\,B}$ mientras que $T_B\,T_A = T_{B\,A}$ (y $AB = B$ pero $BA = 0$).

$\square$

En este sentido, $L(\mathcal{V})$ es más parecido a $K^{n\times n}$, que cuyo producto tampoco es conmutativo. Ya veremos que cuando $n = \dim_K \mathcal{V}$, $L(\mathcal{V})$ y $K^{n\times n}$ son en cierto sentido *la misma álgebra* (ver la sección que sigue y las - futuras - notas para el tp4).

3.4 Epis, monos e isos

En lo que sigue vamos a considerar como comparar espacios vectoriales. La comparación se realiza a través de transformaciones lineales que además, en tanto funciones, resultan inyectivas, ó suryectivas ó biyectivas. Repasemos brevemente estas nociones de Álgebra I.

Definición 3.22. *Sean $\mathcal{A}$, $\mathcal{B}$ conjuntos y sea $f : \mathcal{A} \to \mathcal{B}$ función. Decimos que f es:*

1. *Suryectiva si para todo $b \in \mathcal{B}$, existe $a \in \mathcal{A}$ tal que $f(a) = b$;*

2. *Inyectiva si para todos a_1, $a_2 \in \mathcal{A}$ vale que: si $f(a_1) = f(a_2)$ entonces $a_1 = a_2$;*

3. *Biyectiva si f es suryectiva e inyectiva a la vez.*

$\square$

Obs 3.23. Con las notaciones de la definición anterior, sea $\mathcal{C}$ conjunto y $g : \mathcal{B} \to \mathcal{C}$. Recordemos que:

1. Si f y g son suryectivas, entonces $g \circ f : \mathcal{A} \to \mathcal{C}$ es suryectiva;

2. Si f y g son inyectivas, entonces $g \circ f : \mathcal{A} \to \mathcal{C}$ es inyectiva;

3. Si f y g son biyectivas, entonces $g \circ f : \mathcal{A} \to \mathcal{C}$ es biyectiva.

$\square$

En el contexto del álgebra lineal, vamos a considerar la siguiente terminología

Definición 3.24. *Sean $\mathcal{V}$, $\mathcal{W}$ Kev's y $T \in L(\mathcal{V}, \mathcal{W})$. Decimos que T es:*

1. *Epimorfismo (ó simplemente epi) si T es suryectiva;*

2. *Monomorfismo (ó simplemente mono) si T es inyectiva;*

3. *Isomorfismo (ó simplemente iso) si T es biyectiva.*

$\square$

El siguiente resultado caracteriza los monomorfismos entre espacios vectoriales

Teorema 3.25. *Sean $\mathcal{V}$, $\mathcal{W}$ K-ev's y $T \in L(\mathcal{V}, \mathcal{W})$. Son equivalentes*

1. *T es mono;*

2. *Siempre que $\{v_1, \ldots, v_k\} \subset \mathcal{V}$ es l.i. entonces $\{T(v_1), \ldots, T(v_k)\} \subset \mathcal{W}$ es l.i.;*

3. *$N(T) = \{0\}$.*

Demostración. 1. $\implies$ 2. Supongamos que T es mono y que $\{v_1, \ldots, v_k\} \subset \mathcal{V}$ es l.i. Sean $\alpha_1, \ldots, \alpha_k \in K$ tales que

$$0 = \sum_{j=1}^{k} \alpha_j \ T(v_j) \implies 0 = T(\sum_{j=1}^{k} \alpha_j \ v_j)$$

por la linealidad de T. Recordemos que como $T \in L(\mathcal{V}, \mathcal{W})$ se verifica que $T(0) = 0$. Así, como T es inyectiva,

$$T(0) = 0 = T(\sum_{j=1}^{k} \alpha_j \ v_j) \implies 0 = \sum_{j=1}^{k} \alpha_j \ v_j \, .$$

Como $\{v_1, \ldots, v_k\}$ es l.i., deducimos que $\alpha_1 = \ldots = \alpha_k = 0$, de forma que $\{T(v_1), \ldots, T(v_k)\}$ es l.i. y vale 2.

2. $\implies$ 3. Supongamos que vale 2. y sea $v \in$N(T). Si $v \neq 0$ entonces $\{v\} \subset \mathcal{V}$ es l.i. (ejercicio). Entonces $\{T(v)\}$ resulta l.i. por 2. En particular, $T(v) \neq 0$ (de otra forma $\{T(v)\} = \{0\}$ que es l.d.) y $v \notin$N(T), lo que contradice lo anterior. En conclusión, $v = 0$; como $v \in$N(T) era arbitrario, entonces el núcleo se reduce a N$(T) = \{0\}$.

3. $\Longrightarrow$ 1. Supongamos 3. es decir, si $z \in \mathcal{V}$ es tal que $T(z) = 0$ ($z \in$N(T)) entonces $z = 0$ (en palabras, el único vector en el núcleo de T es $\vec{0}$). Veamos que T es mono: en efecto, si u, $v \in \mathcal{V}$ son tales que $T(u) = T(v)$ entonces $T(u - v) = T(u) - T(v) = 0$. Luego, $u - v \in$N(T) lo que muestra que $u - v = 0$, es decir, $u = v$.

□

Corolario 3.26. *Sean $\mathcal{V}$, $\mathcal{W}$ K-ev's, $\dim_K \mathcal{V} = n$ y $T \in L(\mathcal{V}, \mathcal{W})$. Son equivalentes*

1. *T es mono;*
2. *Siempre que $\{v_1, \ldots, v_n\}$ es base de $\mathcal{V}$ entonces $\{T(v_1), \ldots, T(v_n)\}$ es l.i.*
3. *Existe una base $\{v_1, \ldots, v_n\}$ es base de $\mathcal{V}$ tal que $\{T(v_1), \ldots, T(v_n)\}$ es l.i.*

Demostración. Como toda base es, en particular, l.i., 1. $\Longrightarrow$ 2. es consecuencia del teorema anterior.

2. $\Longrightarrow$ 3. Como $\dim_K \mathcal{V} = n$, entonces existe $\{v_1, \ldots, v_n\}$ base de $\mathcal{V}$; por 2. concluimos que $\{T(v_1), \ldots, T(v_n)\}$ es l.i.

3. $\Longrightarrow$ 1. Supongamos que $\{v_1, \ldots, v_n\}$ es una base (particular, y fija) de $\mathcal{V}$ tal que $\{T(v_1), \ldots, T(v_n)\}$ es l.i. Para verificar 1. basta ver que N$(T) = \{0\}$, por el teorema anterior. Si $v \in$N(T), en particular, $v \in \mathcal{V}$ y existen $\alpha_1, \ldots, \alpha_n \in K$ tales que

$$v = \sum_{j=1}^{n} \alpha_j \ v_j \implies 0 = T(v) = \sum_{j=1}^{n} \alpha_j \ T(v_j) \,,$$

(pues $v \in$N(T)). Como, por hipótesis, $\{T(v_1), \ldots, T(v_n)\}$ es l.i., vemos que $\alpha_1 = \ldots = \alpha_n = 0$. Finalmente, notamos que $v = \sum_{j=1}^{n} \alpha_j \ v_j = 0$. Así, dado $v \in$N(T) arbitrario, hemos probado que $v = 0$, lo que muestra que N$(T) = \{0\}$.

□

Corolario 3.27. *Sean $\mathcal{V}$, $\mathcal{W}$ K-ev's tales que $\dim_K \mathcal{V} = n$, $\dim_K \mathcal{W} = m$, n, $m \in \mathbb{N}$.*

1. *Si $T \in L(\mathcal{V}, \mathcal{W})$ es mono, entonces $n \leq m$.*
2. *Si $n \leq m$ entonces existe $T \in L(\mathcal{V}, \mathcal{W})$ mono.*

Demostración. 1. Ejercicio.

Veamos 2: sean $B_{\mathcal{V}} = \{v_1, \ldots, v_n\}$ base de $\mathcal{V}$ y $B_{\mathcal{W}} = \{w_1, \ldots, w_m\}$ base de $\mathcal{W}$. Por hipótesis, $n \leq m$ de forma que podemos definir

$$T \in L(\mathcal{V}, \mathcal{W}) \quad \text{única tal que} \quad T(v_j) = w_j \,,\ 1 \leq j \leq n \,.$$

(acá hemos usado el Teorema 3.6). Como T transforma una base de $\mathcal{V}$ en el conjunto $\{w_1, \ldots, w_n\} \subset B_{\mathcal{W}}$ que es l.i. (ejercicio), el corolario anterior muestra que T es mono.

□

Obs 3.28. En lo que sigue vamos a estudiar con detalle el caso de los isomorfismos. Para ello recordemos el siguiente hecho simple de Álgebra I: dada una función $f : \mathcal{A} \to \mathcal{B}$ entonces: f es biyectiva si y solo si existe una función $g : \mathcal{B} \to \mathcal{A}$ tal que

$$g \circ f = I_{\mathcal{B}} \quad \text{y} \quad f \circ g = I_A$$

es decir $g(f(b)) = b$ para todo $b \in \mathcal{B}$ y $f(g(a)) = a$ para todo $a \in \mathcal{A}$. Más aún, en el caso en que tal función g exista, entonces es única y se nota $g = f^{-1}$; en este caso $f^{-1} : \mathcal{B} \to \mathcal{A}$ es la llamada función inversa de f. Notemos que por lo mencionado más arriba, si f es biyectiva entonces: dados $a \in \mathcal{A}$ y $b \in \mathcal{B}$,

$$f^{-1}(b) = a \quad \text{si y solo si} \quad f(a) = b \,.$$

Teorema 3.29. *Sean $\mathcal{V}$, $\mathcal{W}$ K-ev's, y supongamos que $T \in L(\mathcal{V}, \mathcal{W})$ es isomorfismo. Entonces la función inversa $T^{-1} : \mathcal{W} \to \mathcal{V}$ es transformación lineal.*

Demostración. Por hipótesis $T : \mathcal{V} \to \mathcal{W}$ es una tranformación lineal que, como función, es biyectiva. Por los comentarios previos, existe la función inversa $T^{-1} : \mathcal{W} \to \mathcal{V}$. Recordemos que dados $v \in \mathcal{V}$ y $w \in \mathcal{W}$ entonces

$$T^{-1}(w) = v \quad \text{si y solo si} \quad Tv = w\,.$$

Veamos que T^{-1} es transformación lineal. Para ello, verificamos que $T^{-1} : \mathcal{W} \to \mathcal{V}$ respeta la suma y la acción escalar. Así, sea w_1, $w_2 \in \mathcal{W}$. Como T es suryectiva, existen v_1, $v_2 \in \mathcal{V}$ tales que

$$w_j = T(v_j) \implies T^{-1}(w_j) = v_j\,,\ j = 1, 2\,.$$

Como T es lineal

$$T(v_1 + v_2) = T(v_1) + T(v_2) = w_1 + w_2 \in \mathcal{W} \implies T^{-1}(w_1 + w_2) = v_1 + v_2 = T^{-1}(w_1) + T^{-1}(w_2)$$

de forma que T^{-1} respeta la suma.

De forma similar, si $\alpha \in K$: como T es lineal

$$T(\alpha\, v_1) = \alpha\, T(v_1) = \alpha\, w_1 \implies T^{-1}(\alpha\, w_1) = \alpha\, v_1 = \alpha\, T^{-1}(w_1)$$

y T^{-1} respeta la acción escalar.

□

Obs 3.30. Sean $\mathcal{V}$ K-ev y $T \in L(\mathcal{V})$ (operador en $\mathcal{V}$).

Si T es isomorfismo entonces, por el teorema anterior, existe $T^{-1} \in L(\mathcal{V})$ tal que

$$T\,T^{-1} = T^{-1}\,T = I$$

donde recordemos que el producto en $L(\mathcal{V})$ simboliza la composición de funciones y I denota la transformación identidad en $\mathcal{V}$.

Recíprocamente, si $T \in L(\mathcal{V})$ es tal que existe $S \in L(\mathcal{V})$ tal que $T\,S = S\,T = I$ entonces T es ismorfismo, y $S = T^{-1}$ (ver Obs. 3.28 sobre inversas de funciones).

Como I es el elemento neutro para el producto, concluimos que $T \in L(\mathcal{V})$ **es iso si y solo si** T **es inversible como elemento en el álgebra** $L(\mathcal{V})$. En adelante, vamos a usar como sinónimos las expresiones *iso* e *inversible* en el caso de operadores $T \in L(\mathcal{V})$.

□

Teorema 3.31. *Sean $\mathcal{V}$, $\mathcal{W}$ K-ev's tales que $\dim_K \mathcal{V} = \dim_K \mathcal{W} = n$. Si $T \in L(\mathcal{V}, \mathcal{W})$, son equivalentes:*

1. *T es isomorfismo;*
2. *T es epimorfismo;*
3. *T es monomorfismo;*
4. *Para toda $\{v_1, \ldots, v_n\}$ base de $\mathcal{V}$ se verifica que $\{T(v_1), \ldots, T(v_n)\}$ es base de $\mathcal{W}$;*
5. *Existe $\{v_1, \ldots, v_n\}$ base de $\mathcal{V}$ tal que $\{T(v_1), \ldots, T(v_n)\}$ es base de $\mathcal{W}$.*

Demostración. Para verificar la equivalencia de estas afirmaciones, probamos una cadena de implicaciones que comienza y termina con el ítem 1. Notemos que 1. $\implies$ 2. de forma evidente (por definición).

2. $\implies$ 3. Como T es transformación lineal, se verifica que

$$\dim_K \mathcal{V} = \dim_K \mathrm{N}(T) + \dim_K \mathrm{Im}(T)\,. \tag{14}$$

Como T es epi, entonces $\dim_K \mathrm{Im}(T) = \dim_K \mathcal{W} = \dim_K \mathcal{V}$. Así, las identidades anteriores implican que $\dim_K \mathrm{N}(T) = 0$, es decir, $\mathrm{N}(T) = \{0\}$. Por un resultado previo concluimos que T es mono.

3. $\implies$ 4. Supongamos que T es mono y que $\{v_1, \ldots, v_n\}$ base de $\mathcal{V}$. Entonces, por el Corolario 3.26, $B = \{T(v_1), \ldots, T(v_n)\} \subset \mathcal{W}$ es l.i. Como $\dim_K \mathcal{W} = n$ entonces concluimos que B es base de $\mathcal{W}$.

4. $\implies$ 5. Por el Corolario 3.26, concluimos que estos ítems son equivalentes! (pues todo base es, en particular, l.i).

5. $\implies$ 1. Por el Corolario 3.26 concluimos que T es mono. En particular, $\mathrm{N}(T) = \{0\}$, de forma que $\dim_K \mathrm{N}(T) = 0$. Consideramos ahora la identidad en la Eq. (14) y concluimos que

$$\dim_K \mathrm{Im}(T) = \dim_K \mathcal{V} = \dim_K \mathcal{W}$$

esta última igualdad de dimensiones prueba que $\mathrm{Im}(T) = \mathcal{W}$ y T es también epimorfismo. Como T es mono y epi, entonces es iso (por definición).

□

Ejemplo 3.32. Sea $\mathcal{V}$ un K-ev, y sea $B = \{v_1, \ldots, v_n\}$ una base de $\mathcal{V}$. Recordemos la transformación *tomar coordenadas con respecto a* B, notada $C_B \in L(\mathcal{V}, K^n)$ dada por

$$C_B(u) = [u]_B \in K^n \quad \text{para} \quad u \in \mathcal{V}\,.$$

Entonces C_B es un iso. En efecto, veamos primero que C_B es mono: si $u \in \mathrm{N}(C_B)$ entonces $[u]_B = \vec{0} \in K^n$. Pero entonces $u = \sum_{j=1}^n 0\, v_j = 0$. Así, $\mathrm{N}(C_B) = \{0\}$ y T es mono. Como $\dim_K \mathcal{V} = n = \dim_K K^n$, el Teorema 3.31 muestra que C_B es iso.

De hecho, en este caso podemos calcular explícitamente la inversa C_B^{-1} que viene dada por

$$C_B^{-1} : K^n \to \mathcal{V} \quad , \quad C_B^{-1}((\alpha_j)_{j=1}^n) = \sum_{j=1}^n \alpha_j\, v_j \in \mathcal{V} \quad \text{para} \quad (\alpha_j)_{j=1}^n \in K^n\,.$$

El lector puede verificar que la transformación descripta más arriba es en realidad la inversa de C_B (ejercicio).

□

Definición 3.33. *Sean $\mathcal{V}$, $\mathcal{W}$ K-ev's. Decimos que $\mathcal{V}$ y $\mathcal{W}$ con isomorfos si existe un isomorfismo $T \in L(\mathcal{V}, \mathcal{W})$.* □

Sean $\mathcal{V}$, $\mathcal{W}$ K-ev's isomorfos y sea $T \in L(\mathcal{V}, \mathcal{W})$ un isomorfismo entre los espacios. Entonces $T : \mathcal{V} \to \mathcal{W}$ es, en particular, una biyección. De esta forma, T permite identificar elementos de $\mathcal{V}$ con elementos de $\mathcal{W}$ de la siguiente forma: $v \in \mathcal{V}$ se identifica con $w \in \mathcal{W}$ si $T(v) = w$, y en este caso podemos escribir $v \approx_T w$. Como T es lineal, entonces la identificación anterior respeta la suma y la acción escalar: es decir, si

$$v_1 \approx_T w_1 \quad \text{y} \quad v_2 \approx_T w_2 \implies v_1 + v_2 \approx_T w_1 + w_2$$

(la afirmación anterior es lo mismo que decir: si $T(v_1) = w_1$ y $T(v_2) = w_2$ entonces $T(v_1 + v_2) = w_1 + w_2$, lo que es cierto gracias a la linealidad de T). De forma similar,

$$v \approx_T w \quad \text{y} \quad \alpha \in K \implies \alpha v \approx_T \alpha w$$

(verificar esta última afirmación: ejercicio). En este sentido, los espacios $\mathcal{V}$ y $\mathcal{W}$ *son esencialmente el mismo espacio*, al menos desde el punto de vista del álgebra lineal.

Los comentarios anteriores sugieren considerar el siguiente problema: dados $\mathcal{V}$ y $\mathcal{W}$ dos K-ev's, determinar si los espacios son isomorfos (o no). Este es el llamado problema del isomorfismo para espacios vectoriales. La solución de este problema, en el caso de espacios de dimensión finita, está dada por el siguiente

Teorema 3.34. *$\mathcal{V}$ y $\mathcal{W}$ dos K-ev's de dimensión finita. Entonces los espacios son isomorfos si y solo si tienen la misma dimensión sobre el cuerpo K.*

Demostración. Supongamos que existe $T \in L(\mathcal{V}, \mathcal{W})$ isomorfismo entre $\mathcal{V}$ y $\mathcal{W}$. Por el ítem 1 del Corolario 3.27 vemos que $\dim_K \mathcal{V} \leq \dim_K \mathcal{W}$, pues T es mono. Por el Teorema 3.29, $T^{-1} \in L(\mathcal{W}, \mathcal{V})$; pero, como T^{-1} también es biyectiva, vemos que T^{-1} es mono. Otra vez por el ítem 1 del Corolario 3.27 vemos que $\dim_K \mathcal{W} \leq \dim_K \mathcal{V}$ y vale la igualdad $\dim_K \mathcal{V} = \dim_K \mathcal{W}$.

Recíprocamente, si $\dim_K \mathcal{V} = \dim_K \mathcal{W}$ entonces, por el Corolario 3.27 vemos que existe $T \in L(\mathcal{V}, \mathcal{W})$ que es monomorfismo. Finalmente, por el Teorema 3.31, concluimos que T es isomorfismo. □

4 Matriz asociada a una transformación lineal

En lo que sigue extendemos nuestro estudio de coordenadas de vectores con respecto a una base (es decir, representaciones de un vector en términos de una n-upla) al contexto de representaciones de transformaciones lineales en términos de matrices.

4.1 Matriz asociada a una transformación lineal y un par de bases

Comenzamos con la definición de matriz de T con respecto a un par de bases.

Definición 4.1. *Sean $\mathcal{V}$, $\mathcal{W}$ K-ev's tales que $\dim_K \mathcal{V} = n$ y $\dim_K \mathcal{W} = m$. Sean $B_\mathcal{V} = \{v_1, \ldots, v_n\}$ y $B_\mathcal{W} = \{w_1, \ldots, w_m\}$ bases de $\mathcal{V}$ y $\mathcal{W}$. Si $T \in L(\mathcal{V}, \mathcal{W})$ definimos* **la matriz de T con respecto a las bases $B_\mathcal{V}$ y $B_\mathcal{W}$**, *notada*

$$[T]_{B_\mathcal{W}, B_\mathcal{V}} \in K^{m \times n}$$

como la matriz cuya j-esima columna viene dada por el vector de coordenadas $[T(v_j)]_{B_\mathcal{W}} \in K^m$, para $1 \leq j \leq n$. Gráficamente,

$$[T]_{B_\mathcal{W}, B_\mathcal{V}} = \begin{pmatrix} | & | & \cdots & | \\ [T(v_1)]_{B_\mathcal{W}} & [T(v_2)]_{B_\mathcal{W}} & \cdots & [T(v_n)]_{B_\mathcal{W}} \\ | & | & \cdots & | \end{pmatrix} \in K^{m \times n} . \tag{15}$$

□

Obs 4.2. Con las notaciones de la Definición 4.1, si $1 \leq j \leq n$ consideramos

$$[T(v_j)]_{B_\mathcal{W}} = (a_{1j}\,,\, a_{2j}\,,\, \ldots\,,\, a_{mj}) \in K^m$$

el vector de coordenadas del valor $T(v_j) \in \mathcal{W}$ con respecto a la base $B_\mathcal{W}$, entonces

$$[T]_{B_\mathcal{W}, B_\mathcal{V}} = (a_{ij})_{1 \leq i \leq m\ ,\ 1 \leq j \leq n} \in K^{m \times n} .$$

△

En lo que sigue vamos a describir una serie de propiedades fundamentales que tiene la matriz de una transformación con respecto a un par de bases.

Teorema 4.3. *Sean $\mathcal{V}$, $\mathcal{W}$ K-ev's tales que* $\dim_K \mathcal{V} = n$ *y* $\dim_K \mathcal{W} = m$. *Sean* $B_{\mathcal{V}} = \{v_1, \ldots, v_n\}$ *y* $B_{\mathcal{W}} = \{w_1, \ldots, w_m\}$ *bases de $\mathcal{V}$ y $\mathcal{W}$. Dada $T \in L(\mathcal{V}, \mathcal{W})$ entonces $[T]_{B_{\mathcal{W}}, B_{\mathcal{V}}} \in K^{m\times n}$ (como en la Definición 4.1) es la* **única** *matriz $m \times n$ que verifica*

$$[Tv]_{B_{\mathcal{W}}} = [T]_{B_{\mathcal{W}}, B_{\mathcal{V}}} \cdot [v]_{B_{\mathcal{V}}} \quad \textit{para todo} \quad v \in \mathcal{V}\,. \tag{16}$$

Antes de comenzar formalmente con la prueba, notemos que en la identidad en Eq. (16), la expresión

$$[T]_{B_{\mathcal{W}}, B_{\mathcal{V}}} \cdot [v]_{B_{\mathcal{V}}}$$

corresponde al producto matriz-vector: notar que este producto está bien definido (por las dimensiones de la matriz y el vector). La identidad en Eq. (16) justifica la notación que hemos usado para el orden de las bases como subíndices: los subíndices que se encuentran en esta identidad son iguales ($B_{\mathcal{V}}$). Por otro lado, la utilidad de $[T]_{B_{\mathcal{W}}, B_{\mathcal{V}}}$ queda clara a partir de la identidad en Eq. 16: como el producto matriz-vector tiene una fórmula (sencilla) para su cálculo, $[T]_{B_{\mathcal{W}}, B_{\mathcal{V}}}$ nos permite calcular (mediante una fórmula) las coordenadas de $[Tv]_{B_{\mathcal{W}}}$ a partir de las de $[v]_{B_{\mathcal{V}}}$.

Demostración. Dado $v \in \mathcal{V}$, sea $[v]_{B_{\mathcal{V}}} = (\alpha_1, \ldots, \alpha_n) \in K^n = K^{n\times 1}$. En este caso,

$$v = \sum_{j=1}^{n} \alpha_j \, v_j\,.$$

Entonces, por la linealidad de T y de tomar coordenadas con respecto a una base

$$[Tv]_{B_{\mathcal{W}}} = \ [T(\sum_{j=1}^{n} \alpha_j \ v_j)]_{B_{\mathcal{W}}} = \sum_{j=1}^{n} \alpha_j \ [T(v_j)]_{B_{\mathcal{W}}} = [T]_{B_{\mathcal{W}}, B_{\mathcal{V}}} \, [v]_{B_{\mathcal{V}}}\,,$$

donde la última identidad es consecuencia de la Proposición 2.15 y el hecho de que, por construcción, la j-ésima columna de $[T]_{B_{\mathcal{W}}, B_{\mathcal{V}}}$ es $[T(v_j)]_{B_{\mathcal{W}}}$ (ver Eq. (15)).

Para verificar la unicidad de $[T]_{B_{\mathcal{W}}, B_{\mathcal{V}}}$, sea $A \in K^{m\times n}$ tal que

$$[Tv]_{B_{\mathcal{W}}} = A[v]_{B_{\mathcal{V}}} \quad \text{para todo} \quad v \in \mathcal{V}\,.$$

Sean $\vec{a}_1, \ldots, \vec{a}_n \in K^m$ las columnas de A (como en la Obs. 2.14). Sea $1 \le j \le n$ y consideremos $v_j \in B_{\mathcal{V}}$: entonces, $[v_j]_{B_{\mathcal{V}}} = e_j$, el j-ésimo vector de la bases canónica. Así, por hipótesis,

$$[T(v_j)]_{B_{\mathcal{W}}} = A\,[v_j]_{B_{\mathcal{V}}} = A\,e_j = \vec{a}_j$$

de forma que $\vec{a}_j = [T(v_j)]_{B_{\mathcal{W}}}$, que es la j-ésima columna de $[T]_{B_{\mathcal{W}}, B_{\mathcal{V}}}$. Así, columna a columna, las matrices A y $[T]_{B_{\mathcal{W}}, B_{\mathcal{V}}}$ son iguales, lo que implica que $A = [T]_{B_{\mathcal{W}}, B_{\mathcal{V}}}$. □

Obs 4.4. Sea $\mathcal{V}$ un K-ev, $\dim_K(\mathcal{V}) = n$ y sea $T \in L(\mathcal{V})$ un operador lineal sobre $\mathcal{V}$. Si $B = \{v_1, \ldots, v_n\}$ es base de $\mathcal{V}$ entonces podemos considerar la matriz de T con respecto a B tanto en el dominio como co-dominio, que en este caso notamos $[T]_B$, es decir

$$[T]_B = [T]_{B,B} \in K^{n\times n}\,.$$

En este caso, por el Teorema 4.3, tenemos que

$$[Tv]_B = [T]_B \, [v]_B \quad \text{para todo} \quad v \in \mathcal{V}\,.$$

△

Ejemplo 4.5. Sea $\mathcal{V} = \mathbb{R}^2$ como $\mathbb{R}$-espacio vectorial y sea $B = \{(1,0),\,(0,1)\} = \{e_1,\,e_2\}$ la base canónica. Supongamos que $T \in L(\mathbb{R}^2)$ es tal que

$$[T]_B = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}.$$

En este caso, la información que nos están dando es:

$$[Te_1]_B = \begin{pmatrix} 1 \\ -1 \end{pmatrix} \quad \text{y} \quad [Te_2]_B = \begin{pmatrix} 1 \\ -1 \end{pmatrix}.$$

Entonces

$$Te_1 = 1\,e_1 + 1\,e_2 = (1,1) \quad \text{y} \quad Te_2 = 1\,e_1 + \,(-1)\,e_2 = (1,-1)\,.$$

Así, si $(a,b) \in \mathbb{R}^2$ entonces, como $[(a,b)]_B = (a,b)$ es decir,

$$(a,b) = a\,e_1 + b\,e_2 \implies T(a,b) = a\,Te_1 + b\,Te_2 = a\,(1,1) + b\,(1,-1) = (a+b, a-b)\,.$$

De esta forma, podemos conocer la transformacion T (en todo $\mathbb{R}^2$). El hecho que hayamos tomado la base B como la base canónica no tiene nada de especial (solo permite hacer las cuentas de forma más sencilla).

Si por otra parte, consideramos la base $B' = \{(1,0),(1,-1)\} = \{u_1,\,u_2\}$ entonces, para calcular $[T]_{B'}$ debemos calcular $[Tu_1]_{B'}$ y $[Tu_2]_{B'}$. Cálculos sencillos muestran que

$$T(u_1) = (1,1) = 2\,(0,1) + (-1)\,(1,-1) = 2\,u_1 + (-1)\,u_2 \implies [T(u_1)]_{B'} = (2,-1)\,,$$

mientras que

$$T(u_2) = (0,2) = 2\,(0,1) + (-2)\,(1,-1) = 2\,u_1 + (-2)\,u_2 \implies [T(u_2)]_{B'} = (2,-2)\,.$$

Ahora podemos construir

$$[T]_{B'} = \begin{pmatrix} 2 & 2 \\ -1 & -2 \end{pmatrix}.$$

Notemos que $[T]_B$ y $[T]_{B'}$ son dos matrices distintas que representan la misma transformación, pero con respecto a bases distintas de $\mathbb{R}^2$. Más adelante veremos qué tipo de relación existe entre estas matrices (ver Sección 4.4 más abajo) . $\triangle$

4.2 Propiedades de la matriz asociada a una transformación lineal y un par de bases

Las siguientes son algunas propiedades muy útiles para trabjar con matrices asociadas a transformaciones lineales (y bases).

Proposición 4.6. *Sean $\mathcal{V}_1$, $\mathcal{V}_2$, $\mathcal{V}_3$ tres K-ev's de dimensión finita, dotados de bases B_1, B_2 y B_3 respectivamente. Entonces,*

1. *Si $S,\,T \in L(\mathcal{V}_1, \mathcal{V}_2)$ y $\alpha \in K$ entonces*

$$[\alpha S + T]_{B_2,B_1} = \alpha\,[S]_{B_2,B_1} + [T]_{B_2,B_1}\,.$$

2. *Si $T \in L(\mathcal{V}_1, \mathcal{V}_2)$ $U \in L(\mathcal{V}_2, \mathcal{V}_3)$ entonces*

$$[U\,T]_{B_3,B_1} = [U]_{B_3,B_2} \cdot [T]_{B_2,B_1}\,.$$

Antes de considerar la prueba, notemos que el ítem 1. del enunciado indica que, fijadas las bases B_1 y B_2, si conocemos la matriz de T (es decir $[T]_{B_2,B_1}$) y la matriz S ($[S]_{B_2,B_1}$) entonces podemos calcular la matriz de $\alpha\, S+T$ como la suma de las matrices $\alpha\,[S]_{B_2,B_1}$ y $[T]_{B_2,B_1}$.

De forma similar, si conocemos la matriz de U con respecto a las bases B_2, B_3 y la matriz de T con respecto a las bases B_1, B_2, entonces podemos calcular la matriz de la composición $U\,T$ como el producto de las matrices $[U]_{B_3,B_2}\cdot[T]_{B_2,B_1}$. Notemos que el producto está bien definido: en efecto, si $\dim_K \mathcal{V}_1=n_1$, $\dim_K \mathcal{V}_2=n_2$ y $\dim_K \mathcal{V}_3=n_3$ entonces

$$[T]_{B_2,B_1}\in K^{n_2\times n_1} \quad \text{y} \quad [U]_{B_3,B_2}\in K^{n_3\times n_2} \implies [U]_{B_3,B_2}\cdot[T]_{B_2,B_1}\in K^{n_3\times n_1}\,.$$

Demostración. Para verificar las identidades de los ítems 1 y 2 vamos a explotar la *unicidad* de la matrix de una transformación con respecto a un par de bases del Teorema 4.3. Concretamente, para verificar 1., sea $v\in\mathcal{V}_1$: entonces

$$\begin{aligned}[(\alpha\,S+T)\,v]_{B_2} &= [\alpha\,S\,v+T\,v]_{B_2}=\alpha\,[S\,v]_{B_2}+[T\,v]_{B_2}\\ &= \alpha\,[S]_{B_2,B_1}\,[v]_{B_1}+[T]_{B_2,B_1}\,[\,v]_{B_1}\\ &= (\alpha\,[S]_{B_2,B_1}+[T]_{B_2,B_1})\,[v]_{B_1}\,,\end{aligned}$$

donde hemos usado la definición de la suma y acción escalar en $L(\mathcal{V}_1,\mathcal{V}_2)$ para la evaluación en v, la linealidad de tomar coordenadas con respecto a la base B_2, la propiedad de la matriz de una transformación con respecto a las bases B_1, B_2 del Teorema 4.3 y las propiedades de la suma y multiplicación de matrices (distributividad). Así, la matriz $(\alpha\,[S]_{B_2,B_1}+[T]_{B_2,B_1})\in K^{n_2\times n_1}$ verifica: para todo $v\in\mathcal{V}_1$

$$[(\alpha\,S+T)\,v]_{B_2}=(\alpha\,[S]_{B_2,B_1}+[T]_{B_2,B_1})\,[v]_{B_1}\,.$$

Por la *unicidad* de $[\alpha S+T]_{B_2,B_1}$ verificada en el Teorema 4.3, concluimos que vale el ítem 1.

Para verificar el ítem 2, recordemos que dadas T y U como arriba entonces $U\,T\in L(\mathcal{V}_1,\mathcal{V}_3)$ está dada por

$$U\,T(v)=U(T(v)) \quad \text{para} \quad v\in\mathcal{V}_1\,.$$

Consideremos $v\in\mathcal{V}_1$: entonces,

$$\begin{aligned}[U\,T\,(v)]_{B_3} &= [U(T(v))]_{B_3}=[U]_{B_3,B_2}\,[Tv]_{B_2}\\ &= [U]_{B_3,B_2}\,(\,[T]_{B_2,B_1}\,[v]_{B_1}\,)=(\,[U]_{B_3,B_2}\,[T]_{B_2,B_1})\,[v]_{B_1}\end{aligned}$$

donde usamos la definición de composición, la propiedad de la matriz de U con respecto a B_2, B_3, la propiedad de la matriz de T con respecto a B_1, B_2 y la asociatividad del producto de matrices. Así, la matriz $[U]_{B_3,B_2}\,[T]_{B_2,B_1}\in K^{n_3\times n_1}$ verifica: para todo $v\in\mathcal{V}_1$

$$[U\,T\,(v)]_{B_3}=(\,[U]_{B_3,B_2}\,[T]_{B_2,B_1})\,[v]_{B_1}\,.$$

Por la *unicidad* de $[U\,T]_{B_3,B_1}$ verificada en el Teorema 4.3, concluimos que vale el ítem 2. □

Teorema 4.7. *Sean $\mathcal{V}$, $\mathcal{W}$ K-ev's con bases $B_{\mathcal{V}}=\{v_1,\dots,v_n\}$ y $B_{\mathcal{W}}=\{w_1,\dots,w_m\}$ respectivamente. Entonces la transformación*

$$[\,\cdot\,]_{B_{\mathcal{W}},\,B_{\mathcal{V}}}:L(\mathcal{V},\mathcal{W})\to K^{m\times n} \quad \textit{dada por} \quad T\mapsto[T]_{B_{\mathcal{W}},\,B_{\mathcal{V}}}$$

es un isomorfismo lineal.

Antes de entrar en la prueba, aclaremos un poco la notación: $[\,\cdot]_{B_\mathcal{W},\,B_\mathcal{V}}$ es una función cuyo dominio es el conjunto de las transformaciones lineales de $\mathcal{V}$ en $\mathcal{W}$. El puntito que aparece entre los corchetes (que funciona como una variable) indica que la función toma un elemento del dominio T y reemplaza el puntito por T es decir: si evaluamos $[\,\cdot]_{B_\mathcal{W},\,B_\mathcal{V}}$ en T el valor resultante es la matriz $[T]_{B_\mathcal{W},\,B_\mathcal{V}}$. Esto es en realidad equivalente a describir la función como se hizo en el enunciado.

Demostración. Notemos que el ítem 1. de Proposición 4.6 muestra, justamente, que la transformación

$$L(\mathcal{V},\mathcal{W}) \ni T \mapsto [T]_{B_\mathcal{W},\,B_\mathcal{V}} \in K^{m\times n}$$

es lineal, pues respeta sumas y acción por escalares: es decir, si $S,\,T \in L(\mathcal{V},\mathcal{W})$,

$$[S+T]_{B_\mathcal{W},\,B_\mathcal{V}} = [S]_{B_\mathcal{W},\,B_\mathcal{V}} + [T]_{B_\mathcal{W},\,B_\mathcal{V}} \quad \text{y} \quad [\alpha\, S]_{B_\mathcal{W},\,B_\mathcal{V}} = \alpha\, [S]_{B_\mathcal{W},\,B_\mathcal{V}}\,,$$

(a partir del ítem 1 de la Proposición 4.6, haciendo $\alpha = 1$ para la primer identidad, y $T = 0$ para la segunda).

Además, hemos calculado $\dim_K K^{m\times n} = m\,n = \dim_K L(\mathcal{V},\mathcal{W})$. Como los espacios tienen la misma dimensión entonces, por un resultado en la segunda parte de la teoría para el tp3, para probar que esta transformación lineal es un iso, basta ver que es mono!

Por un resultado ya visto, para verificar que la transformación es mono, basta ver que su núcleo es el subespacio nulo. Sea entonces $T \in L(\mathcal{V},\mathcal{W})$ tal que

$$[T]_{B_\mathcal{W},\,B_\mathcal{V}} = 0 \in K^{m\times n}$$

y veamos que entonces $T = 0$ (es decir, la única T que está en el núcleo de $[\,\cdot]_{B_\mathcal{W},\,B_\mathcal{V}}$ es la transformación $T = 0$). En efecto, por el Teorema 4.3, dado $v \in \mathcal{V}$ entonces

$$[Tv]_{B_\mathcal{W}} = [T]_{B_\mathcal{W},\,B_\mathcal{V}}\,[v]_{B_\mathcal{V}} = 0\,[v]_{B_\mathcal{V}} = \vec{0} \in K^m\,.$$

Es decir, las coordenadas de Tv son todas nulas: entonces

$$Tv = 0\,w_1 + \ldots + 0\,w_m = 0$$

es decir, $Tv = 0$. Como $v \in \mathcal{V}$ era un vector cualquiera, hemos probado que

$$Tv = 0 \quad \text{para todo} \quad v \in \mathcal{V} \implies T = 0 \in L(\mathcal{V},\mathcal{W})$$

es decir, $T = 0$ es la tranformación lineal nula. Esto muestra que el núcleo $\mathrm{N}([\,\cdot]_{B_\mathcal{W},B_\mathcal{V}}) = \{0\}$ y $[\,\cdot]_{B_\mathcal{W},B_\mathcal{V}} : L(\mathcal{V},\mathcal{W}) \to K^{m\times n}$ es un isomorfismo de K ev's. □

Obs 4.8. Sean $\mathcal{V}$, $\mathcal{W}$ K-ev's con bases $B_\mathcal{V} = \{v_1,\ldots,v_n\}$ y $B_\mathcal{W} = \{w_1,\ldots,w_m\}$ respectivamente. Como una consecuencia del Teorema 4.7, concluimos que:

1. Si $S,\,T \in L(\mathcal{V},\mathcal{W})$ son tales que $[S]_{B_\mathcal{W},B_\mathcal{V}} = [T]_{B_\mathcal{W},B_\mathcal{V}}$ entonces $S = T$. Esto es una consecuencia directa del hecho de que $[\,\cdot]_{B_\mathcal{W},B_\mathcal{V}}$ es una función inyectiva (pues es un mono).

2. Dada $A \in K^{m\times n}$ existe una (única) $T \in L(\mathcal{V},\,\mathcal{W})$ tal que $[T]_{B_\mathcal{W},B_\mathcal{V}} = A$. Esto es una consecuencia directa del hecho de que $[\,\cdot]_{B_\mathcal{W},B_\mathcal{V}}$ es una función suryectiva (pues también es un epi). La unicidad es consecuencia del primer ítem de esta observación!

△

Más adelante vamos a usar sistemáticamente (y muy frecuentemente) el siguiente resultado.

Corolario 4.9. *Sean $\mathcal{V}$, $\mathcal{W}$ K-ev's tales que $\dim_k \mathcal{V} = \dim_k \mathcal{W} = n \in \mathbb{N}$, con bases $B_\mathcal{V}$ y $B_\mathcal{W}$, respectivamente. Dada $T \in L(\mathcal{V},\mathcal{W})$, son equivalentes:*

1. T es isomorfismo lineal;

2. $[T]_{B_\mathcal{W},B_\mathcal{V}} \in K^{n\times n}$ es matriz inversible.

Más aún, si T es isomorfismo entonces

$$[T]_{B_\mathcal{W},B_\mathcal{V}}{}^{-1} = [T^{-1}]_{B_\mathcal{V},B_\mathcal{W}},.$$

Demostración. Supongamos que T es isomorfismo. Entonces, hemos probado que la función inversa $T^{-1} \in L(\mathcal{W},\mathcal{V})$ también es transformación lineal. En este caso, $I_\mathcal{V} = T^{-1}\,T \in L(\mathcal{V},\mathcal{V}) = L(\mathcal{V})$ es el operador identidad. Notemos que $[I_\mathcal{V}]_{B_\mathcal{V},B_\mathcal{V}} = I_n \in K^{n\times n}$ es decir, la matriz de $I_\mathcal{V}$ con respecto a la base $B_\mathcal{V}$ (tanto en el dominio como en el co-dominio) es la matriz identidad! En efecto, por definición de la transformación identidad, si $v \in \mathcal{V}$

$$[I_\mathcal{V}\, v]_{B_\mathcal{V}} = [v]_{B_\mathcal{V}} = I_n\ [v]_{B_\mathcal{V}}\,.$$

Por la unicidad de $[I_\mathcal{V}]_{B_\mathcal{V},B_\mathcal{V}}$ del Teorema 4.3, vemos que $[I_\mathcal{V}]_{B_\mathcal{V},B_\mathcal{V}} = I_n$.

Entonces, por el ítem 2 de la Proposición 4.6,

$$I_n = [I_\mathcal{V}]_{B_\mathcal{V},B_\mathcal{V}} = [T^{-1}\,T]_{B_\mathcal{V},B_\mathcal{V}} = [T^{-1}]_{B_\mathcal{V},B_\mathcal{W}}\ [T]_{B_\mathcal{W},B_\mathcal{V}}\,. \tag{17}$$

De forma similar (los detalles quedan como ejercicio),

$$I_n = [I_\mathcal{W}]_{B_\mathcal{W},B_\mathcal{W}} = [T\,T^{-1}]_{B_\mathcal{W},B_\mathcal{W}} = [T]_{B_\mathcal{W},B_\mathcal{V}}\ [T^{-1}]_{B_\mathcal{V},B_\mathcal{W}}\,. \tag{18}$$

Las ecuaciones (17) y (18) muestran que $[T]_{B_\mathcal{W},B_\mathcal{V}}$ es inversible y su matriz inversa está dada por

$$[T]^{-1}_{B_\mathcal{W},B_\mathcal{V}} = [T^{-1}]_{B_\mathcal{V},B_\mathcal{W}}\,.$$

Recíprocamente, supongamos que $[T]_{B_\mathcal{W},B_\mathcal{V}}$ es inversible. Sea $v \in \mathcal{V}$ tal que $T(v) = 0$. Entonces,

$$\vec{0} = [Tv]_{B_\mathcal{W}} = [T]_{B_\mathcal{W},B_\mathcal{V}}\,[v]_{B_\mathcal{V}} \implies [v]_{B_\mathcal{V}} = \vec{0}$$

pues el núcleo de la matriz $\mathrm{N}([T]_{B_\mathcal{W},B_\mathcal{V}}) = \{\vec{0}\}$, pues es inversible. Así, $v = 0$ y vemos que $\mathrm{N}(T) = \{0\}$. Como $\dim_K \mathcal{V} = \dim_K \mathcal{W}$ deducimos que T es isomorfismo lineal.

□

4.3 Matriz de una trasformación y cambios de base

El siguiente resultado muestra que relación existe entre las distintas representaciones matriciales de una misma transformación T cuando cambiamos las bases del espacio dominio y el espacio co-dominio de T.

Teorema 4.10. *Sean $\mathcal{V}$, $\mathcal{W}$ K-ev's de dimensión finita. Sean $B_\mathcal{V}$, $B'_\mathcal{V}$ bases de $\mathcal{V}$ y $B_\mathcal{W}$, $B'_\mathcal{W}$ bases de $\mathcal{W}$. Dada una transformación $T \in L(\mathcal{V},\mathcal{W})$ se verifica*

$$[T]_{B'_\mathcal{W},B'_\mathcal{V}} = M_{B'_\mathcal{W},B_\mathcal{W}} \cdot [T]_{B_\mathcal{W},B_\mathcal{V}} \cdot M_{B_\mathcal{V},B'_\mathcal{V}}\,, \tag{19}$$

donde $M_{B_\mathcal{V},B'_\mathcal{V}}$ denota la matriz de cambio de base de $B'_\mathcal{V}$ en la base $B_\mathcal{V}$ y donde $M_{B'_\mathcal{W},B_\mathcal{W}}$ denota la matriz de cambio de base de $B_\mathcal{W}$ en la base $B'_\mathcal{W}$.

Demostración. Para verificar la identidad propuesta en el enunciado usamos la unicidad de la matriz de T con respecto a un par de bases del Teorema 4.3. Para ello, recordemos la propiedades de las matrices cambio de base, definidas y consideradas en el apunto correspondiente al Tp2. De hecho, las matrices cambio de bases verifican: si $v \in \mathcal{V}$ y $w \in \mathcal{W}$ entonces,

$$[v]_{B_\mathcal{V}} = M_{B_\mathcal{V}, B'_\mathcal{V}}\, [v]_{B'_\mathcal{V}} \quad \text{y} \quad [w]_{B'_\mathcal{W}} = M_{B'_\mathcal{W}, B_\mathcal{W}}\, [w]_{B_\mathcal{W}} \tag{20}$$

Recordemos que nuestra notación está diseñada para que los sub-índices que se encuentran sean los mismos. En este sentido, notemos que la fórmula de la Eq. (19) tiene esta misma propiedad: los sub-índices correspondientes a las distintas matrices que se encuentran son iguales, tanto para el primer y segundo factor como para el segundo y tercer factor.

Siguiendo con la prueba, usando la Eq. (20), vemos que si $v \in \mathcal{V}$

$$\begin{aligned}
[Tv]_{B'_\mathcal{W}} &= M_{B'_\mathcal{W}, B_\mathcal{W}}\, [Tv]_{B_\mathcal{W}} = M_{B'_\mathcal{W}, B_\mathcal{W}}\, (\, [T]_{B_\mathcal{W}, B_\mathcal{V}}\, [v]_{B_\mathcal{V}}) \\
&= (\, M_{B'_\mathcal{W}, B_\mathcal{W}}\, [T]_{B_\mathcal{W}, B_\mathcal{V}}\,)\, [v]_{B_\mathcal{V}} = (\, M_{B'_\mathcal{W}, B_\mathcal{W}}\, [T]_{B_\mathcal{W}, B_\mathcal{V}}\,)\, M_{B_\mathcal{V}, B'_\mathcal{V}}\, [v]_{B'_\mathcal{V}} \\
&= (\, M_{B'_\mathcal{W}, B_\mathcal{W}}\, [T]_{B_\mathcal{W}, B_\mathcal{V}}\, M_{B_\mathcal{V}, B'_\mathcal{V}}\,)\, [v]_{B'_\mathcal{V}}
\end{aligned}$$

donde hemos usado la propiedad de la matriz cambio de base, la propiedad de la matriz $[T]_{B_\mathcal{W}, B_\mathcal{V}}$ para calcular coordenadas, la asociatividad del producto de matrices, la fórmula de cambio de coordenadas y nuevamente la asociatividad del producto de matrices.

Resumiendo, hemos verificado: para todo $v \in \mathcal{V}$,

$$[Tv]_{\mathcal{W}'} = (\, M_{B'_\mathcal{W}, B_\mathcal{W}}\, [T]_{B_\mathcal{W}, B_\mathcal{V}}\, M_{B_\mathcal{V}, B'_\mathcal{V}}\,)\, [v]_{B'_\mathcal{V}}\,.$$

Por la unicidad de la matriz de $[T]_{B'_\mathcal{W}, B'_\mathcal{V}}$ del Teorema 4.3, concluimos la identidad del enunciado. □

El siguiente caso particular del teorema anterior será de importancia en el estudio de operadores lineales.

Corolario 4.11. *Sea $\mathcal{V}$ un K-ev de dimensión finita y sean $B_\mathcal{V}$, $B'_\mathcal{V}$ bases de $\mathcal{V}$. Dado un operador $T \in L(\mathcal{V})$ se verifica*

$$[T]_{B'_\mathcal{V}} = M_{B'_\mathcal{V}, B_\mathcal{V}} \cdot [T]_{B_\mathcal{V}} \cdot M_{B_\mathcal{V}, B'_\mathcal{V}} = M^{-1}_{B_\mathcal{V}, B'_\mathcal{V}} \cdot [T]_{B_\mathcal{V}} \cdot M_{B_\mathcal{V}, B'_\mathcal{V}}\,,$$

donde $M_{B_\mathcal{V}, B'_\mathcal{V}}$ denota la matriz de cambio de base de $B'_\mathcal{V}$ en la base $B_\mathcal{V}$ y donde $M_{B'_\mathcal{V}, B_\mathcal{V}}$ denota la matriz de cambio de base de $B_\mathcal{V}$ en la base $B'_\mathcal{V}$.

Demostración. Recordemos que para el caso de operadores lineales sobre el espacio $\mathcal{V}$ hemos convenido en notar

$$[T]_{B_\mathcal{V}} = [T]_{B_\mathcal{V}, B_\mathcal{V}}$$

cuando consideramos la misma base $\mathcal{V}$ en el dominio como co-dominio de T, evitando el doble sub-índice. Por otro lado, en el contexto de matrices de cambio de bases, hemos probado que se verifica la identidad

$$M_{B'_\mathcal{V}, B_\mathcal{V}} = M^{-1}_{B_\mathcal{V}, B'_\mathcal{V}}$$

para un par de bases $B_\mathcal{V}$ y $B'_\mathcal{V}$ de $\mathcal{V}$. Una vez hechas estas alcaraciones, el enunciado del corolario es una consecuencia inmediata del Teorema 4.10.

□

4.4 Matrices semejantes

Sea $\mathcal{V}$ un K-ev de dimensión finita y sea $T \in L(\mathcal{V})$. Si B y B' son bases de $\mathcal{V}$ hemos visto (ver Ejemplo 4.5) que, en general $[T]_B$ y $[T]_{B'}$ son matrices distintas. Sin embargo, estas matrices están relacionadas entre sí (y comparten varias propiedades, como veremos más adelante). Para describir la relación entre $[T]_B$ y $[T]_{B'}$ consideramos la siguiente noción.

Definición 4.12. *Sea K un cuerpo y sean A, $B \in K^{n\times n}$. Decimos que A y B son semejantes, y notamos $A \sim B$, si existe una matriz inversible $P \in K^{n\times n}$ tal que*

$$B = P^{-1}\,A\,P\,.$$

□

Obs 4.13. Sea K un cuerpo. La relación de semejanza $\sim$ en $K^{n\times n}$ es una relación de equivalencia en $K^{n\times n}$. En efecto, verifica

1. Reflexiva: si $A \in K^{n\times n}$ entonces $A \sim A$ (tomando $P = I$)
2. Simétrica: si $A \sim B$ entonces $B \sim A$. En efecto, si $P \in K^{n\times n}$ es inversible y tal que

$$B = P^{-1}\,A\,P \implies A = P\,B\,P^{-1} = (P^{-1})^{-1}\,B\,P^{-1}$$

   donde hemos despejado a A de la primer ecuación usando que P es inversible y usado que $P^{-1} \in K^{n\times n}$ también es inversible con inversa $(P^{-1})^{-1} = P$.
3. Transitiva: si $A \sim B$ y $B \sim C$ entonces $A \sim C$: en efecto, si $P, Q \in K^{n\times n}$ son inversibles y tales que

$$B = P^{-1}\,A\,P \quad \text{y} \quad C = Q^{-1}\,B\,Q \implies C = Q^{-1}\,(P^{-1}\,A\,P)\,Q = (PQ)^{-1}\;A\;(PQ)$$

   donde hemos usado asociatividad del producto y el hecho de que si P y Q son inversibles entonces PQ es inversible y $(PQ)^{-1} = Q^{-1}P^{-1}$.

△

Proposición 4.14. *Sea $\mathcal{V}$ un K-ev, $\dim_K \mathcal{V} = n$ y sea $T \in L(\mathcal{V})$.*

1. *Si B y B' son bases de $\mathcal{V}$ entonces $[T]_B \sim [T]_{B'}$ son semejantes.*
2. *Recíprocamente, si B es base de $\mathcal{V}$ y $C \in K^{n\times n}$ es tal que $[T]_B \sim C$ entonces existe una base B' de $\mathcal{V}$ tal que $C = [T]_{B'}$.*

En particular, dada una base B_0 arbitraria (pero fija) de $\mathcal{V}$, se tiene la igualdad de conjuntos

$$\{\ [T]_B\ :\ B \ \ \textit{es base de}\ \ \mathcal{V}\ \} = \{\ A \in K^{n\times n}\ : A \sim [T]_{B_0}\ \}\,.$$

Demostración. El ítem 1 es una consecuencia del Corolario 4.11.

Para verificar el ítem 2, sea $C \in K^{n\times n}$ tal que $[T]_B \sim C$, donde $B = \{v_1, \ldots, v_n\}$ es base de $\mathcal{V}$. Por hipótesis existe $P \in K^{n\times n}$ inversible tal que $C = P^{-1}\,[T]_B\,P$. Por el Teorema 4.7 (ver también la Observación 4.8) existe $S \in L(\mathcal{V})$ tal que $[S]_B = P$. Por construcción de $[S]_B$ vemos que la j-ésima columna de P coincide con el vector de coordenadas $[S(v_j)]_B$, para $1 \le j \le n$. Es decir,

$$P = [S]_B = \begin{pmatrix} | & | & \cdots & | \\ [S(v_1)]_B & [S(v_2)]_B & \cdots & [S(v_n)]_B \\ | & | & \cdots & | \end{pmatrix} \in K^{m\times n}\,. \tag{21}$$

Por el Corolario 4.9, $S \in L(\mathcal{V})$ es un isomorfismo (de $\mathcal{V}$ en $\mathcal{V}$). En particular, S trasnforma bases de $\mathcal{V}$ en bases de $\mathcal{V}$: así, si

$$w_j = Sv_j \quad \text{para} \quad j = 1, \dots, n \implies B' = \{w_1, \dots, w_n\} \quad \text{es base de} \quad \mathcal{V}\,.$$

Calculemos ahora la matriz de cambio de base $M_{B,B'}$ de la base B' en la base B. Para ello, recordemos que la j-ésima columna de la matriz $M_{B,B'}$ coincide con el vector de coordenadas $[w_j]_B = [S(v_j)]_B$, para $1 \leq j \leq n$. Así, por la Eq. (21), vemos que columna a columna, las matrices $M_{B,B'}$ y P son iguales. Así, $M_{B,B'} = P$. Entonces

$$[T]_{B'} = M_{B,B'}^{-1}\,[T]_B\,M_{B,B'} = P^{-1}\,[T]_B\,P = C\,.$$

Finalmente, la igualdad de conjuntos es una consecuencia de los ítems 1. y 2. que permiten probar la doble inclusión.

□

4.5 Algunas consecuencias para $L(\mathcal{V})$

En este sección enunciamos algunas consecuencias (casi directas) de los resulados descriptos en la sección de repaso para el *álgebra de operadores* $L(\mathcal{V})$.

Teorema 4.15. *Sean $\mathcal{V}$ K-ev tal que $\dim_k \mathcal{V} = n \in \mathbb{N}$ y sea B una base de $\mathcal{V}$. En este caso podemos considerar las K-álgebras (con unidad) $L(\mathcal{V})$ y $K^{n\times n}$. Entonces la transformación*

$$[\,\cdot]_B : L(\mathcal{V}) \to K^{n\times n} \quad \text{dada por} \quad T \mapsto [T]_B$$

es un isomorfismo de K-álgebras con unidad: es decir, $[\cdot]_B$ es una función biyectiva que respeta la suma, la acción escalar, la unidad y el producto:

1. *Si $S, T \in L(\mathcal{V})$, $\alpha \in K$: $[\alpha\, S + T]_B = \alpha\,[S]_B + [T]_B$*
2. *$[I_{\mathcal{V}}]_B = I$*
3. *Si $S, T \in L(\mathcal{V})$: $[T\,S]_B = [T]_B\,[S]_B$.*

Proof. Antes que nada, recordemos la notación $[T]_B = [T]_{B,B}$, para $T \in L(\mathcal{V})$ (consideramos la misma base en $\mathcal{V}$ como dominio y co-dominio).

Por la Proposición 3.19 y la Observación 3.20 sabemos que $L(\mathcal{V})$ es una K-álgebra. En Álgebra I, han visto que $K^{n\times n}$ es también una K-álgebra.

Por el Teorema 4.7 sabemos que $[\cdot]_B$ es un isomorfismo lineal: en particular, es una biyección que verifica el item 1 del enunciado.

Por la Proposición 4.6 (haciendo $\mathcal{V}_1 = \mathcal{V}_2 = \mathcal{V}_3 = \mathcal{V}$), vemos que se verifica el ítem 3. El ítem 2 se de forma sencilla (ejercicio), teniendo en cuenta la Definición 4.1. □

Corolario 4.16. *Sean $\mathcal{V}$ K-ev tal que $\dim_k \mathcal{V} = n \in \mathbb{N}$, con base B. Dada $T \in L(\mathcal{V}, \mathcal{W})$, son equivalentes:*

1. *T es isomorfismo lineal;*
2. *$[T]_B \in K^{n\times n}$ es matriz inversible.*

Más aún, si T es isomorfismo entonces $[T]_B{}^{-1} = [T^{-1}]_{B,}$.

Proof. Es una consecuencia inmediata del Corolario 4.9 del repaso. □

5 Formas canónicas: operadores diagonalizables

5.1 Consideraciones generales

En lo que sigue:

- $\mathcal{V}$ denotará un K-ev, $\dim_K \mathcal{V} = n \in \mathbb{N}$
- $T \in L(\mathcal{V})$ denotará un operador que hemos fijado.

El *problema* que vamos a considerar es, a *grosso modo*, el siguiente: dado T como antes, *hallar* una base $B = \{v_1, \ldots, v_n\}$ de $\mathcal{V}$ tal que la matriz

$$[T]_B \qquad \text{tenga una estructura sencilla .}$$

Como veremos, el hecho de que podamos hallar una base en la que $[T]_B$ sea más sencilla ó menos sencilla dependerá de T (y en cierto punto de K también).

Claro, la primer pregunta es: qué significa estructura sencilla?! En esta primera parte de las notas sencillo significará *diagonal* (concepto que definimos más abajo) que sin dudas está intuitivamente relacionado con la noción de simplicidad. Más adelante, veremos algunas nociones un poco más técnicas... pero no nos apresuremos!

5.2 Matrices diagonales = matrices sencillas

Sea K un cuerpo y sea $A = (a_{ij}) \in K^{n\times n}$ una matriz. Recordemos que A es una *matriz diagonal* si las entradas de A fuera de la diagonal principal son nulas: en símbolos $a_{ij} = 0$ si $1 \le i \ne j \le n$: es decir

$$a_{ij} = \delta_{ij}\, a_{ii} = \delta_{ij}\, a_{jj} \tag{22}$$

con la notación de Kronecker. Gráficamente, tiene la forma

$$A = \begin{pmatrix} a_{11} & 0 & 0 & \ldots & 0 \\ 0 & a_{22} & 0 & \ldots & 0 \\ \vdots & \vdots & \ddots & \ldots & 0 \\ 0 & 0 & 0 & \ldots & a_{nn} \end{pmatrix} \in K^{n\times n} .$$

También escribimos $A = \operatorname{diag}(a_{11}, \ldots, a_{nn})$ para describir a la matriz diagonal A.

Si $A = \operatorname{diag}(a_{11}, \ldots, a_{nn}) \in K^{n\times n}$ entonces:

1. el producto matriz vector $A \cdot \vec{x}$ se puede realizar de forma sencilla: si $\vec{x} = (x_1, \ldots, x_n) \in K^n$ (recordemos nuestra convención: $K^n = K^{n\times 1}$ denotan vectores columna, que por una cuestión de espacio los notamos como fila)

$$A \cdot \vec{x} = \begin{pmatrix} a_{11}x_1, \\ a_{22}x_2 \\ \vdots \\ a_{nn}\, x_n \end{pmatrix} \in K^n \tag{23}$$

   En efecto, si calculamos la i-ésima coordenada

$$(A\vec{x})_i = \sum_{j=1}^{n} a_{ij}\, x_j = \sum_{j=1}^{n} \delta_{ij}\, a_{ii}\, x_j = a_{ii}\, x_i ,$$

donde usamos la notación de la Ec. (22). En particular, si $\vec{x} = e_1$ es el primer vector de la base canónica de K^n entonces

$$A \cdot e_1 = \begin{pmatrix} a_{11}\, 1 \\ a_{22}\, 0 \\ \vdots \\ a_{nn}\, 0 \end{pmatrix} = a_{11} \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} = a_{11}\, e_1 \in K^n \tag{24}$$

y más en general, si e_j es el j-ésimo vector de la base canónica

$$A \cdot e_j = \begin{pmatrix} a_{11}\, 0 \\ \vdots \\ a_{j-1,j-1}\, 0 \\ a_{jj}\, 1 \\ a_{j+1,j+1}\, 0 \\ \vdots \\ a_{nn}\, 0 \end{pmatrix} = a_{jj} \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} = a_{jj}\, e_j \in K^n \tag{25}$$

Lo anterior indica que la j-ésima columna de A es un múltiplo de e_j, para $1 \leq j \leq n$. Esto también se ve mirando la forma gráfica que tiene la matriz diagonal, en términos de sus columnas!

2. Notemos que si $\alpha \in K$ entonces

$$\alpha\, A = \mathrm{diag}(\alpha\, a_{11}, \ldots, \alpha\, a_{nn}) \begin{pmatrix} \alpha\, a_{11} & 0 & 0 & \ldots & 0 \\ 0 & \alpha\, a_{22} & 0 & \ldots & 0 \\ \vdots & \vdots & \ddots & \ldots & 0 \\ 0 & 0 & 0 & \ldots & \alpha\, a_{nn} \end{pmatrix}$$

Si $B = \mathrm{diag}(b_{11}, \ldots, b_{nn}) \in K^{n\times n}$ es otra matriz diagonal entonces

$$A + B = \mathrm{diag}(a_{11} + b_{11}, \ldots, a_{nn} + b_{nn}) = \begin{pmatrix} a_{11} + b_{11} & 0 & 0 & \ldots & 0 \\ 0 & a_{22} + b_{22} & 0 & \ldots & 0 \\ \vdots & \vdots & \ddots & \ldots & 0 \\ 0 & 0 & 0 & \ldots & a_{nn} + b_{nn} \end{pmatrix}$$

y por otro lado,

$$A \cdot B = \mathrm{diag}(a_{11}\, b_{11}, \ldots, a_{nn}\, b_{nn}) = \begin{pmatrix} a_{11}\, b_{11} & 0 & 0 & \ldots & 0 \\ 0 & a_{22}\, b_{22} & 0 & \ldots & 0 \\ \vdots & \vdots & \ddots & \ldots & 0 \\ 0 & 0 & 0 & \ldots & a_{nn}\, b_{nn} \end{pmatrix}$$

En efecto,

$$(AB)_{ij} = \sum_{k=1}^{n} a_{ik}\, b_{kj} = \sum_{k=1}^{n} \delta_{ik}\, a_{ii}\, \delta_{kj}\, b_{jj} = a_{ii}\, b_{jj} \sum_{k=1}^{n} \delta_{ik}\ \delta_{kj}\,.$$

Si $i \neq j$ entonces (verificar!)

$$\delta_{ik}\ \delta_{kj} = 0 \quad \text{para todo} \quad 1 \leq k \leq n \implies (AB)_{ij} = 0\,.$$

Si $i = j$ entonces

$$\delta_{ik}\ \delta_{kj} = 1 \quad \text{si y solo si} \quad k = i (= j) \implies (AB)_{ii} = a_{ii}\, b_{ii} \quad (i = j)\,.$$

Además, las potencias de A se pueden calcular de forma sencilla:

$$A^2 = \operatorname{diag}(a_{11}^2, \ldots, a_{nn}^2) \quad \text{y en general} \quad A^k = \operatorname{diag}(a_{11}^k, \ldots, a_{nn}^k)$$

(verificar por inducción, usando la fórmula para el producto de matrices diagonales).

Más aún, $AB = BA$ para A y B matrices diagonales.

3. El problema de determinar si una matriz diagonal $A = \operatorname{diag}(a_{11}, \ldots, a_{nn}) \in K^{n\times n}$ es inversible (cosa que en el caso de matrices arbitrarias es más bien complicado) se torna trivial: por un lado, se tiene la expresión (verificar por inducción en el tamaño de A)

$$\det A = \prod_{i=1}^n a_{ii} \implies \det A \neq 0 \quad \text{si y solo si} \quad a_{ii} \neq 0 \quad \text{para todo} \quad 1 \leq i \leq n\,.$$

Lo anterior muestra entonces que A es inversible si y solo si $a_{ii} \neq 0$ para todo $1 \leq i \leq n$; en ese caso está definido $a_{ii}^{-1} \in K$ y se verifica que

$$A^{-1} = \operatorname{diag}(a_{11}^{-1}, \ldots, a_{nn}^{-1}) \in K^{n\times n}$$

Para verificar lo anterior se puede usar la fórmula del producto de matrices diagonales:

$$\operatorname{diag}(a_{11}, \ldots, a_{nn}) \operatorname{diag}(a_{11}^{-1}, \ldots, a_{nn}^{-1}) = \operatorname{diag}(a_{11}\, a_{11}^{-1}, \ldots, a_{nn}\, a_{nn}^{-1}) = \operatorname{diag}(1, \ldots, 1) = I$$

donde hemos usado que la matriz identidad $I \in K^{n\times n}$ es diagonal, con diagonal principal formada por 1's.

Las observaciones anteriores muestran que la clase de matrices diagonales es una clase de matrices sencillas, tanto desde el punto de vista del producto matriz-vector, como desde el punto de vista de las operaciones algebraicas (sumas y productos por escalares y matrices) que se pueden hacer con esta clase de matrices.

5.3 Operadores diagonalizables

Definición 5.1. Sea $\mathcal{V}$ un K-ev, $\dim_K \mathcal{V} = n \in \mathbb{N}$ y sea $T \in L(\mathcal{V})$. Decimos que T es operador *diagonalizable* si existe una base B de $\mathcal{V}$ tal que $[T]_B \in K^{n\times n}$ es matriz diagonal. □

Las observaciones hechas en la sección anterior y el Teorema 4.15 indican que si $T \in L(\mathcal{V})$ es diagonalizable, entonces podemos representar de forma sencilla toda una serie de matrices asociadas con T. Además, podemos resolver algunos problemas relacionados con T de forma simple. Por ejemplo, si B es base de $\mathcal{V}$ tal que

$$[T]_B = \operatorname{diag}(a_{11}, \ldots, a_{nn})$$

entonces T es un isomorfismo lineal si y solo si $a_{ii} \neq 0$ para todo $1 \leq i \leq n$ (ver el ítem 3 de la sección anterior y el Corolario 4.16).

Ejemplo 5.2. Sea $\mathcal{V} = \mathbb{R}^2$ y $T \in L(\mathbb{R}^2)$ dado por

$$T(x, y) = (x + y, x + y) \quad \text{para} \quad (x, y) \in \mathbb{R}^2\,.$$

Si consideramos la base canónica de $\mathbb{R}^2$, $B = \{e_1, e_2\}$ entonces, cálculos sencillos muestran que

$$[T]_B = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}.$$

Claramente, $[T]_B$ no es matriz diagonal. Sin embargo, si consideramos la base $B' = \{u_1 = (1,1),\, u_2 = (1,-1)\}$ entonces, cálculos sencillos muestran que

$$[T]_{B'} = \begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix} .$$

En este caso, $[T]_{B'}$ es matriz diagonal. Deducimos que T es un operador diagonalizable, porque existe la base B' de $\mathbb{R}^2$ tal que $[T]_{B'}$ es matriz diagonal. △

Pero cómo hallar la base B' del ejemplo anterior?

Otra pregunta que surge es: será cierto que todo operador (en todo espacio vectorial) es diagonalizable? La respuesta a esta pregunta es: NO. Veamos el siguiente (no)-ejemplo:

Ejemplo 5.3. Sea

$$\mathcal{V} = \{p \in \mathbb{R}[x] \ : \ p = 0 \ \text{ ó } \operatorname{gr}(p) \leq 2\} = \{a_0 + a_1\, x + a_2 x^2 : a_0,\, a_1,\, a_2 \in \mathbb{R}\} \subset \mathbb{R}[x] \, .$$

En este contexto $B = \{1,\, x,\, x^2\}$ es una base de $\mathcal{V}$, de forma que $\dim_{\mathbb{R}} \mathcal{V} = 3$. Consideramos el operador derivación

$$D \in L(\mathcal{V}) \quad \text{ dado por } \quad D(a_0 + a_1\, x + a_2 x^2) = a_1 + 2\, a_2\, x \in \mathcal{V} \, .$$

Afirmamos que no existe base B' de $\mathcal{V}$ en la cual $[D]_{B'} \in \mathbb{R}^{3\times 3}$ sea diagonal. Notemos que si componemos a D con sí mismo tres veces, es decir, consideramos la potencia $D^3 = D\,D\,D \in L(\mathcal{V})$ se verifica:

$$D(a_0 + a_1\, x + a_2 x^2) = a_1 + 2\, a_2\, x \implies D^2(a_0 + a_1\, x + a_2 x^2) = D(a_1 + 2\, a_2\, x) = 2\, a_2$$

de forma que

$$D^3(a_0 + a_1\, x + a_2 x^2) = D(2\, a_2) = 0$$

En palabras: si derivamos tres veces cualquier polinomio de grado a lo sumo 2, obtenemos el polinomio nulo.

Lo anterior muestra que $D^3 = 0$ es el operador nulo en $\mathcal{V}$. Supongamos ahora que existe una base B' de $\mathcal{V}$ tal que $[D]_{B'} = \operatorname{diag}(a_{11},\, a_{22},\, a_{33}) \in \mathbb{R}^{3\times 3}$. Entonces, por el Teorema 4.15,

$$0 = [D^3]_{B'} = ([D]_{B'})^3 = \operatorname{diag}(a_{11}^3,\, a_{22}^3,\, a_{33}^3) \in \mathbb{R}^{3\times 3} \, .$$

Lo anterior muestra que $a_{ii}^3 = 0$ lo que a su vez implica que $a_{ii} = 0$, para $i = 1, 2, 3$. Pero entonces

$$[D]_{B'} = \operatorname{diag}(a_{11},\, a_{22},\, a_{33}) = \operatorname{diag}(0,\, 0,\, 0) = 0 \in \mathbb{R}^{3\times 3}$$

lo que implica que $D = 0$. Pero esto es absurdo, porque D no es el operador nulo: $D(x) = 1 \neq 0$.

Esta contradicción surge de suponer que D era diagonalizable. Entonces D no es diagonalizable! △

El ejemplo anterior es un tanto extenso, pero muy importante. No sólo porque muestra que hay operadores que no son diagonalizables, sino porque indica una clase que es la contra-cara de los operadores diagonalizables: los llamados operadores *nilpontentes* que estudiaremos más adelante (D del ejemplo anterior es nilpotente, porque $D^3 = 0 \in L(\mathcal{V})$).

Como un primer paso para entender cuando un operador es diagonalizable, consideramos el siguiente resultado.

Proposición 5.4. *Sea $\mathcal{V}$ un K-ev, sea $B = \{v_1, \ldots, v_n\}$ base de $\mathcal{V}$, sea $T \in L(\mathcal{V})$ y sean $\lambda_1, \ldots, \lambda_n \in K$. Son equivalentes:*

1. *$[T]_B = diag(\lambda_1, \ldots, \lambda_n) \in K^{n\times n}$;*

2. *$T v_j = \lambda_j \, v_j$, para $1 \le j \le n$.*

Demostración Supongamos que vale el ítem 1. Si $1 \le j \le n$ entonces podemos calcular las coordenadas de $[Tv_j]_B$ usando la propiedades de $[T]_B \in K^{n\times n}$:

$$[Tv_j]_B = [T]_B\,[v_j]_B = [T]_B\, e_j = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)\, e_j = \lambda_j\, e_j = \lambda_j(0, \ldots, 0, 1, 0, \ldots, 0)$$

(en único 1 está en la entrada j-ésima) donde hemos usando la fórmula de la Ec. (25) para e_j, el j-ésimo vector de la base canónica. Entonces, como $e_j = (\delta_{ij})_{i=1}^n$,

$$[Tv_j]_B = \lambda_j\, e_j = (\lambda_j\, \delta_{ij})_{i=1}^n \implies Tv_j = \sum_{i=1}^n (\lambda_j\, \delta_{ij})\ v_i = \lambda_j\ v_j\,.$$

Lo anterior prueba que $Tv_j = \lambda_j\, v_j$.

Recíprocamente, si $Tv_j = \lambda_j\, v_j$ entonces,

$$[Tv_j]_B = \lambda_j\ e_j \quad \text{para} \quad 1 \le j \le n\,.$$

Pero entonces

$$[T]_B = \begin{pmatrix} | & | & \cdots & | \\ \lambda_1\, e_1 & \lambda_2\, e_2 & \cdots & \lambda_n\, e_n \\ | & | & \cdots & | \end{pmatrix} = \begin{pmatrix} \lambda_1 & 0 & 0 & \ldots & 0 \\ 0 & \lambda_2 & 0 & \ldots & 0 \\ \vdots & \vdots & \ddots & \ldots & 0 \\ 0 & 0 & 0 & \ldots & \lambda_n \end{pmatrix} \in K^{n\times n}\,.$$

□

Como primer resultado, deducimos

Teorema 5.5. *Sea $\mathcal{V}$ un K-ev, $\dim_K \mathcal{V} = n \in \mathbb{N}$ y sea $T \in L(\mathcal{V})$. Son equivalentes:*

1. *T es operador diagonalizable*

2. *Existe una base $B = \{v_1, \ldots, v_n\}$ de $\mathcal{V}$ y $\lambda_1, \ldots, \lambda_n \in K$, tales que*

$$Tv_j = \lambda_j\, v_j \quad \textit{para} \quad 1 \le j \le n\,.$$

□

Obs 5.6. Con las notaciones del teorema anterior, cabe remarcar que hay cierta asimetría en los roles de la base B por un lado y los coeficientes $\lambda_1, \ldots, \lambda_n \in K$.

Concretamente, dada T diagonalizable, en general puede haber muchas bases $B' = \{v_1', \ldots, v_n'\}$ de $\mathcal{V}$ tales $[T]_{B'} = \operatorname{diag}(\mu_1, \ldots, \mu_n)$, es decir, tales que $Tv_j' = \mu_j\, v_j'$ para $1 \le j \le n$, para ciertos $\mu_1, \ldots, \mu_n$. Pero en este caso los escalares $\lambda_1, \ldots, \lambda_n$ coinciden con $\mu_1, \ldots, \mu_n$ salvo el órden (es decir, los μ's se obtienen de los λ's permutándolos). Así, los $\lambda_1, \ldots, \lambda_n$ son (esencialmente) únicos.

En lo que sigue vamos a estudiar estos escalares $\lambda_1, \ldots, \lambda_n$ asociados a T.

6 Autovalores y autoespacios

Definición 6.1. Sea $\mathcal{V}$ un K-ev tal que $\dim_K(\mathcal{V}) = n$ y sea $T \in L(\mathcal{V})$. Dado $\lambda \in K$

1. Decimos que λ es un *autovalor* de T si existe $v \in \mathcal{V}$, $v \neq 0$ tal que $Tv = \lambda v$.

2. En adelante $\sigma(T) = \{\lambda \in K \ : \ \lambda \text{ es autovalor de } T\}$ es llamado el *espectro* de T.

3. Si $\lambda \in \sigma(T)$ definimos el *autoespacio* de T asociado al autovalor λ, notado $E_\lambda(T) \subseteq \mathcal{V}$, dado por
$$E_\lambda(T) = \{w \in \mathcal{V} \ : \ Tw = \lambda\, w\} \neq \{0\}\,.$$
En este caso, si $w \in E_\lambda(T)$ (es decir, si $Tw = \lambda\, w$) decimos que w es un *autovector* de T asociado a λ.

□

Obs 6.2. Con las notaciones de la defínción anterior, si $\lambda \in \sigma(T)$ entonces $E_\lambda(T)$ es subespacio no nulo de $\mathcal{V}$: en efecto, notemos que

$$E_\lambda(T) = \{w \in \mathcal{V} \ : \ Tw = \lambda\, w\} = \{w \in \mathcal{V} \ : \ \lambda\, w - Tw = 0\} = \{w \in \mathcal{V} \ : \ (\lambda\, I - T)w = 0\}$$

donde I denota el operador identidad en $\mathcal{V}$, es decir $(\lambda\, I - T)w = \lambda I(w) - T(w) = \lambda\, w - Tw$; así

$$E_\lambda(T) = \{w \in \mathcal{V} \ : \ (\lambda\, I - T)w = 0\} = \mathrm{N}(\lambda\, I - T) \neq \{0\} \subseteq \mathcal{V}$$

es subespacio, pues es el núcleo del operador $\lambda\, I - T \in L(\mathcal{V})$. El subespacio no se reduce al subespacio nulo, porque por hipótesis existe $v \in \mathcal{V}$, $v \neq 0$ tal que $Tv = \lambda\, v \implies v \in E_T(\lambda)$. △

Obs 6.3. Sea $\mathcal{V}$ un K-ev, $\dim_K \mathcal{V} = n \in \mathbb{N}$ y sea $T \in L(\mathcal{V})$. Si T es operador diagonalizable entonces T debe tener autovalores! En efecto, por el Teorema 5.5 existe una base $B = \{v_1, \ldots, v_n\}$ de $\mathcal{V}$ y $\lambda_1, \ldots, \lambda_n \in K$, tales que $Tv_j = \lambda_j\, v_j$ para $1 \leq j \leq n$. Como v_j es un vector en una base, $v_j \neq 0$ (de otra forma B sería l.d), lo que muestra cada λ_j es autovalor de T. △

Definición 6.4. Sea $\mathcal{V}$ un K-ev tal que $\dim_K(\mathcal{V}) = n$ y sea $T \in L(\mathcal{V})$. Definimos el *determinante* de T, notado $\det T \in K$, como
$$\det T = \det([T]_B)$$
donde B denota una base (cualquiera) de $\mathcal{V}$. △

Con las notaciones de la definición anterior, parece haber un posible problema. Notemos que si B y B' son bases de $\mathcal{V}$ entonces sabemos que, en general las matrices $[T]_B$ y $[T]_{B'}$ son distintas; sin embargo, las matrices $[T]_B$ y $[T]_{B'}$ tienen el mismo determinante: en efecto, hemos probado que en este caso las matrices son semejantes $[T]_B \sim [T]_{B'}$, es decir, que existe una matriz $P \in K^{n\times n}$ tal que P es inversible y $P^{-1}\,[T]_B\, P = [T]_{B'}$ lo que implica

$$\det([T]_{B'}) = \det(P^{-1}\,[T]_B\, P) = \det(P^{-1})\, \det([T]_B)\, \det(P) = \det([T]_B) \in K$$

donde usamos que el producto en K es commutativo, que los determinantes son elementos de K, que $\det(AB) = \det(A)\det(B) \in K$ y que

$$\det(P)\det(P^{-1}) = \det(P\,P^{-1}) = \det(I) = 1 \in K\,.$$

Teorema 6.5. *Sea $\mathcal{V}$ un K-ev tal que $\dim_K(\mathcal{V}) = n$ y sea $T \in L(\mathcal{V})$. Dado $\lambda \in K$, son equivalentes:*

1. *λ es autovalor de T;*

2. $\lambda I - T \in L(\mathcal{V})$ *no es inversible;*

3. $\det(\lambda I - T) = 0$.

Demostración. 1. $\implies$ 2. Si λ es autovalor de T entonces (ver la Observación 6.2)

$$E_\lambda(T) = \mathrm{N}(\lambda I - T) \neq \{0\}\,,$$

porque como λ es autovalor de T existe $v \in \mathcal{V}$, $v \neq 0$ tal que $Tv = \lambda v \implies v \in E_T(\lambda)$. Entonces $\lambda I - T$ no es monomorfismo; así, por el Teorema 3.31 (tomando $\mathcal{W} = \mathcal{V}$ en ese teorema) T no es isomorfismo y vale 2.

2. $\implies$ 3. Supongamos que $\lambda I - T$ no es isomorfismo (es decir, $\lambda I - T$ no es inversible) y sea B una base de $\mathcal{V}$. Por el Corolario 4.16 concluimos que $[\lambda I - T]_B \in K^{n\times n}$ no es matriz inversible, de forma que su determinante es 0. Entonces, por la Definición 6.4,

$$\det(\lambda I - T) = \det([\lambda I - T]_B) = 0\,.$$

3. $\implies$ 1. Supongamos que $\det(\lambda I - T) = 0$. Si B es una base de $\mathcal{V}$ entonces, por la Definición 6.4, $\det([\lambda I - T]_B) = 0$. En particular, $[\lambda I - T]_B$ no es matriz inversible. Por el Corolario 4.16 vemos que $\lambda I - T \in L(\mathcal{V})$ no es isomorfismo. Ahora, por el Teorema 3.31 (tomando $\mathcal{W} = \mathcal{V}$ en ese teorema) concluimos que $\lambda I - T$ no es monomorfismo. Entonces $\mathrm{N}(\lambda I - T) \neq \{0\}$; así debe existir $v \in \mathrm{N}(\lambda I - T)$, $v \neq 0$: pero entonces

$$(\lambda I - T)(v) = 0 \implies Tv = \lambda v \quad \text{con} \quad v \neq 0\,.$$

Esto muestra que $\lambda \in K$ es autovalor de T. □

Ejemplo 6.6. Sea $\mathcal{V}$ un $\mathbb{R}$-ev con base $B = \{w_1, w_2\}$ y sea $T \in L(\mathcal{V})$ tal que

$$[T]_B = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \in \mathbb{R}^{2\times 2}\,.$$

Entonces $\sigma(T) = \emptyset$ (T no tiene autovalores). En efecto, si suponemos que $\lambda \in \mathbb{R}$ es autovalor de T entonces

$$0 = \det(\lambda I - T) = \det([\lambda I - T]_B) = \det(\lambda I_2 - [T]_B)$$

donde $I_2 \in \mathbb{R}^{2\times 2}$ denota la matriz identidad. Pero entonces

$$0 = \det(\lambda I_2 - [T]_B) = \det(\lambda \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} - \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}) = \det \begin{pmatrix} \lambda & -1 \\ 1 & \lambda \end{pmatrix} = \lambda^2 + 1 \geq 1$$

lo que es una contradicción ($0 \geq 1$) ya que $\lambda^2 \geq 0$ para todo $\lambda \in \mathbb{R}$.

En particular, teniendo en cuenta la Observación 6.3, vemos que T no es diagonalizable. □

Definición 6.7. Sea $\mathcal{V}$ un K-ev tal que $\dim_K(\mathcal{V}) = n$ y sea $T \in L(\mathcal{V})$. Definimos el polinomio característico de T, $p_T(x) \in K[x]$, dado por

$$p_T(x) = \det(x\, I - [T]_B) \in K[x]\,,$$

donde B es una base arbitraria de $\mathcal{V}$ y donde $I \in K^{n\times n}$ denota la matriz identidad. □

Ejemplo 6.8. Sea $\mathcal{V}$ un $\mathbb{C}$-ev tal que $\dim_{\mathbb{C}} \mathcal{V} = 3$ y sea B una base de $\mathcal{V}$. Sea $T \in L(\mathcal{V})$ tal que

$$[T]_B = \begin{pmatrix} 3 & 1 & -1 \\ 2 & 2 & -1 \\ 2 & 2 & 0 \end{pmatrix}$$

Calculemos $p_T(x) \in K[x]$: para ello

$$p_T(x) = \det(x\,I - [T]_B) = \det(\begin{pmatrix} x-3 & -1 & 1 \\ -2 & x-2 & 1 \\ -2 & -2 & x \end{pmatrix})$$

(donde I denota la matriz identidad 3×3). Cálculos sencillos muestran que

$$p_T(x) = x^3 - 5x^2 + 8x - 4 = (x-1)\,(x-2)^2$$

$\triangle$

Obs 6.9. Sea $\mathcal{V}$ un K-ev tal que $\dim_K(\mathcal{V}) = n$, sea $\mathcal{B}$ una baes de $\mathcal{V}$ y sea $T \in L(\mathcal{V})$. Si $p_T(x) = \det(x\,I - [T]_B) \in K[x]$ es el polinomio característico, entonces:

1. $p_T(x) \in K[x]$ está bien definido en el sentido que no depende de B (tan sólo depende de T): En efecto, si B' es otra base de $\mathcal{V}$ entonces hemos probado que en este caso las matrices son semejantes $[T]_B \sim [T]_{B'}$, es decir, que existe una matriz $P \in K^{n\times n}$ tal que P es inversible y $P^{-1}\,[T]_B\,P = [T]_{B'}$ lo que implica

$$\det(x\,I - [T]_{B'}) = \det(x\,I - P^{-1}\,[T]_B\,P) = \det(P^{-1}\,(x\,I - [T]_{B'})\,P) = \det([T]_B) \in K$$

   donde usamos que el producto en K es commutativo, que los determinantes son elementos de K, que $\det(AB) = \det(A)\det(B) \in K$ y que $\det(P)\det(P^{-1}) = \det(P\,P^{-1}) = \det(I) = 1 \in K$.

2. $p_T(x)$ es un polinomio mónico de grado n. Es decir,

$$p_T(x) = x^n + a_{n-1}\,x^{n-1} + \ldots + a_1\,x + a_0 \in K[x]\,.$$

3. $\lambda \in K$ es autovalor de T si y solo si $p_T(\lambda) = 0$, es decir, si $\lambda \in K$ es raíz de $p_T(x)$.

La afirmación del ítem 2 se verifica utilizando la fórmula del determinante en términos de las entradas de la matriz

$$p_T(x) = \det(x\,I - [T]_B)$$

donde B es una base de $\mathcal{V}$ y $I_n \in K^{n\times n}$ es la matriz identidad.

La afirmación del item 3 es una consecuencia directa del Teorema 6.5 y del hecho de que

$$p_T(\lambda) = \det(\lambda\,I - [T]_B) = \det([\lambda\,I - T]_B) = \det(\lambda\,I - T)\,.$$

$\square$

Usando la notación de la Observación anterior, notemos que T *tiene a lo sumo n autovalores distintos!* pues $p_T(x)$ tiene a lo sumo n raíces distintas (ya que es un polinomio de grado n).

Definición 6.10. Sea $\mathcal{V}$ un K-ev tal que $\dim_K(\mathcal{V}) = n$. Sea $T \in L(\mathcal{V})$ con polinomio característico $p_T(x)$. Si $\lambda \in K$ es autovalor de T:

1. Definimos la *multiplicidad algebraica* de λ en T, notada $m_a(\lambda)$ como el orden de λ como raíz de $p_T(x)$. Es decir, $m_a(\lambda) = r \geq 1$ si
$$p_T(x) = (x - \lambda)^r\, q(x) \quad \text{con} \quad q(\lambda) \neq 0\,.$$

2. Definimos la *multiplicidad geométrica* de λ en T, notada $m_g(\lambda)$ como la dimensión del autoespacio $E_T(\lambda)$
$$m_g(\lambda) = \dim_K(E_T(\lambda)) \geq 1\,.$$

□

6.1 Una estrategia para intentar diagonalizar un operador

A modo de motivación del contenido de esta sección, consideremos la siguiente *estrategia* para intentar diagonalizar un operador $T \in L(\mathcal{V})$, para $\mathcal{V}$ un K-ev tal que $\dim_K(\mathcal{V}) = n$:

1. Calcular $p_T(x)$ polinomio característico de T;

2. Hallar las raíces *distintas* de $p_T(x)$, $\lambda_1, \ldots, \lambda_k$ ($k \leq n$). Es decir, incluimos todas las raíces, pero sin multiplicidades algebraicas (sin repetir). En este caso, el espectro de T es (ver Definición 6.1)
$$\sigma(T) = \{\lambda_1, \ldots, \lambda_k\}\,.$$

3. Para cada autovalor λ_j de T, calcular una base $B_j = \{v_1^j, \ldots, v_{d_j}^j\}$ para el autoespacio $E_T(\lambda_j)$, $1 \leq j \leq k$. Notar que en este caso $\dim_K E_T(\lambda_j) = d_j$, la cantidad de vectores de B_j. Hemos usado un supra-índice para indicar en qué base están los vectores: v_i^j es el i-ésimo vector de la base B_j. Notemos que como $v_i^j \in E_T(\lambda_j)$ entonces
$$T\, v_i^j = \lambda_j\, v_i^j\,. \tag{26}$$

4. Unir (yuxtaponer) las bases de forma que $B = B_1 \cup B_2 \cup \ldots \cup B_k$ es decir
$$B = \{\underbrace{v_1^1, \ldots, v_{d_1}^1}_{B_1}, \underbrace{v_1^2, \ldots, v_{d_2}^2}_{B_2}, v_1^3, \ldots, \underbrace{v_1^k, \ldots, v_{d_k}^k}_{B_k}\}$$

5. Aquí surgen dos preguntas: la primera es si B es l.i. La segunda pregunta es si B es sistema de generadores. Notemos que si la respuesta a estas dos preguntas es *si*, entonces B es una base de $\mathcal{V}$ formada por autovectores de T, por la Eq. (26); por el Teorema 5.5, T sería diagonalizable en este caso.

Veremos que en general B es l.i. (ver Corolario 6.15 más abajo). Sin embargo, el hecho de que B sea un sistema de generadores dependerá de T (y en cierto punto de K).

Lema 6.11. *Sea $\mathcal{V}$ un K-ev tal que $\dim_K(\mathcal{V}) = n$ y sea $T \in L(\mathcal{V})$. Si $v_1, \ldots, v_k \in \mathcal{V}$ son autovectores* **no nulos** *de T asociados a los autovalores $\lambda_1, \ldots, \lambda_k$* **distintos** *de T entonces $\{v_1, \ldots, v_k\}$ es l.i.*

Demostración. Por hipótesis, $v_i \neq 0$ es tal que $T\, v_i = \lambda_i\, v_i$ para $1 \leq i \leq k$.

Vamos a probar el enunciado por inducción en k.

Si $k = 1$ es trivial porque v_1 es no nulo.

Hiótesis inductiva (H.I.): dados $v_1, \ldots, v_{k-1}$ ($k \geq 1$) autovectores no nulos asociados a autovalores distintos $\lambda_1, \ldots, \lambda_{k-1}$ se verifica que $\{v_1, \ldots, v_{k-1}\}$ es l.i.

Sean $v_1, \ldots, v_k$ autovectores no nulos asociados a autovalores distintos $\lambda_1, \ldots, \lambda_k$ y sean $\alpha_1, \ldots, \alpha_k \in K$ tales que

$$0 = \sum_{j=1}^{k} \alpha_j \, v_j \,. \tag{27}$$

Además, como cada v_j es autovector

$$0 = T(0) = T(\sum_{j=1}^{k} \alpha_j \, v_j) = \sum_{j=1}^{k} \alpha_j \, T(v_j) = \sum_{j=1}^{k} \alpha_j \, \lambda_j \, v_j \,. \tag{28}$$

Usando la Eq. (27) vemos que

$$\alpha_k \, v_k = \sum_{j=1}^{k-1} (-\alpha_j) \, v_j \,.$$

Reemplazando esta identidad en la Eq. (28) vemos que

$$\begin{aligned} 0 &= \sum_{j=1}^{k} \alpha_j \, \lambda_j \, v_j = \sum_{j=1}^{k-1} \lambda_j \, \alpha_j \, v_j + \lambda_k \, (\sum_{j=1}^{k-1} - \alpha_j \, v_j) \\ &= \sum_{j=1}^{k-1} \lambda_j \, \alpha_j \, v_j + \sum_{j=1}^{k-1} -(\lambda_k \, \alpha_j) \, v_j = \sum_{j=1}^{k-1} (\lambda_j - \lambda_k) \, \alpha_j \, v_j \,, \end{aligned}$$

donde hemos usado las propiedades de la acción escalar y la identidad $-(\alpha \, v) = (-\alpha) \, v$ (el opuesto de $\alpha \, v$ es el vector que se obtiene de actuar con el escalar $-\alpha$ sobre el vector v).

Por H.I. $\{v_1, \ldots, v_{k-1}\}$ es l.i. de forma que

$$0 = \sum_{j=1}^{k-1} (\lambda_j - \lambda_k) \, \alpha_j \, v_j \implies (\lambda_j - \lambda_k) \, \alpha_j = 0 \ , \ 1 \leq j \leq k-1 \implies \alpha_j = 0 \ , \ 1 \leq j \leq k-1$$

pues $(\lambda_j - \lambda_k) \neq 0$, dado que los autovalores $\lambda_1, \ldots, \lambda_k$ son *distintos*. Así, la Eq. (27) se convierte en $\alpha_k \, v_k = 0$ lo que implica que $\alpha_k = 0$, pues $v_k \neq 0$. $\square$

Definición 6.12. *Sea $\mathcal{V}$ un K-ev y sean $\mathcal{W}_1, \ldots, \mathcal{W}_k$ subespacios de $\mathcal{V}$.*

1. *Definimos el subespacio suma, notado $\mathcal{W}_1 + \ldots + \mathcal{W}_k \subset \mathcal{V}$, dado por*

$$\mathcal{W}_1 + \ldots + \mathcal{W}_k = \{w_1 + \ldots + w_k \ : \ w_i \in \mathcal{W}_i \ , \ 1 \leq i \leq k\} \,.$$

2. *Decimos que los subespacios son independientes si: siempre que $w_i \in \mathcal{W}_i$, $1 \leq i \leq k$ son tales que*

$$w_1 + \ldots + w_k = 0 \implies w_1 = \ldots = w_k = 0 \,.$$

3. *Si los subespacios $W_1, \ldots, W_k$ son independientes, se dice que la suma $W_1 + \ldots + W_k$ es directa y la suma se nota*

$$W_1 \dot{+} \ldots \dot{+} W_k \,.$$

$\square$

Proposición 6.13. *Sea $\mathcal{V}$ un K-ev tal que* $\dim_K(\mathcal{V}) = n$ *y sea $T \in L(\mathcal{V})$. Si $\lambda_1, \ldots, \lambda_k$ son los autovalores* **distintos** *de T entonces los subespacios $E_T(\lambda_1), \ldots, E_T(\lambda_k)$ son independientes (ver Definición 6.12).*

Demostración. Sean $v_j \in E_T(\lambda_j)$, $1 \leq j \leq k$, tales que

$$v_1 + \ldots + v_k = 0\,.$$

Supongamos que hay algún $v_j \neq 0$. En lo que sigue vamos a tener que discriminar entre los $v_j's$ nulos y los v_j's no nulos: formalizamos el argumento como sigue. Consideramos el conjunto $I = \{j \ : \ 1 \leq j \leq k \ , \ v_j \neq 0\} \neq \emptyset$. Si $[k] = \{1, \ldots, k\}$ entonces si consideramos la diferencia de conjuntos

$$j \in [k] \setminus I \implies v_j = 0$$

(en palabras: si $j \in [k]$ y $j \notin I$ entonces $v_j = 0$). Entonces $[k] = ([k] \setminus I) \cup I$ y

$$0 = v_1 + \ldots + v_k = \sum_{j \in [k] \setminus I} v_j + \sum_{j \in I} v_j = 0 + \sum_{j \in I} v_j = \sum_{j \in I} v_j\,.$$

Pero $\{v_j \ , \ j \in I\}$ son autovectores *no nulos* de T correspondientes a autovalores distintos λ_j con $j \in I$ (que son algunos de los autovalores distintos con los que empezamos): Por el Lema 6.11 concluimos que $\{v_j \ , \ j \in I\}$ es l.i., pero

$$\sum_{j \in I} 1\, v_1 = 0$$

con $0 \neq 1 \in K$, que contradice la independencia lineal de $\{v_j \ , \ j \in I\}$. La contradicción surge de suponer que existe algún $v_j \neq 0$. Así, se debe verificar que $v_j = 0$ para $1 \leq j \leq k$, y los subespacios son idenpendientes. □

Teorema 6.14. *Sea $\mathcal{V}$ un K-ev, $\dim_K \mathcal{V} = n \in \mathbb{N}$ y sean $\mathcal{W}_1, \ldots, \mathcal{W}_k$ subespacios de $\mathcal{V}$. Son equivalentes:*

1. *$\mathcal{W}_1, \ldots, \mathcal{W}_k$ son subespacios independientes;*
2. *Si $B_j = \{v_1^j, \ldots, v_{d_j}^j\}$ es base de $\mathcal{W}_j$, $1 \leq j \leq k$, entonces $B = B_1 \cup B_2 \cup \ldots \cup B_k$ (por yuxtaposición ó pegado) es decir*

   $$B = \{\underbrace{v_1^1, \ldots, v_{d_1}^1}_{B_1}, \underbrace{v_1^2, \ldots, v_{d_2}^2}_{B_2}, v_1^3, \ldots, \underbrace{v_1^k, \ldots, v_{d_k}^k}_{B_k}\}$$

   es base de $\mathcal{W} = \mathcal{W}_1 + \ldots + \mathcal{W}_k$.

En particular, si $\mathcal{W}_1, \ldots, \mathcal{W}_k$ son subespacios independientes entonces

$$\dim_K(\mathcal{W}_1 + \ldots + \mathcal{W}_k) = \dim_K(\mathcal{W}_1) + \ldots + \dim_K(\mathcal{W}_k)\,.$$

Demostración. $1 \implies 2$. Supongamos que $\mathcal{W}_1, \ldots, \mathcal{W}_k$ son subespacios independientes. Si B es como en el item 2, verificamos que B es l.i. y un sistema de generadores para $\mathcal{W} = \mathcal{W}_1 + \ldots + \mathcal{W}_k$.

B es l.i.: Es conveniente indicar los coeficientes que actúan sobre los vectores con el mismo tipo de índices (supra y sub-índices) que hemos elegido para los vectores: así, sean

$$\alpha_i^j \in K \quad , \quad 1 \leq i \leq d_j \qquad 1 \leq j \leq k$$

tales que

$$0 = \underbrace{\alpha_1^1\, v_1^1 + \ldots + \alpha_{d_1}^1\, v_{d_1}^1}_{d_1 \text{ términos}} + \underbrace{\alpha_1^2\, v_1^2 + \ldots + \alpha_{d_2}^2\, v_{d_2}^2}_{d_2 \text{ términos}} + \alpha_1^3\, v_1^3 + \ldots + \underbrace{\alpha_1^k\, v_1^k + \ldots + \alpha_{d_k}^k\, v_{d_k}^k}_{d_k \text{ términos}}$$

es decir, sumando cada sumatoria de d_j términos indicada arriba, $1 \leq j \leq k$

$$0 = \sum_{j=1}^{k} \sum_{i=1}^{d_j} \alpha_i^j \; v_i^j = 0\,.$$

Notemos que hemos elegido la notación de forma que el índice j sea un supra-índice en α_i^j, y no un exponente (es decir, no estamos considerando la j-ésima potencia de $\alpha_i \in K$). En este caso:

$$\sum_{j=1}^{k} \underbrace{\sum_{i=1}^{d_j} \alpha_i^j \; v_i^j}_{\in \mathcal{W}_j} = 0 \implies \sum_{i=1}^{d_j} \alpha_i^j \; v_i^j = 0 \quad \text{para} \quad 1 \leq j \leq k\,,$$

por la independencia de los subespacios (ver Definición 6.12). Para $1 \leq j \leq k$, como $B_j = \{v_1^j, \ldots, v_{d_j}^j\}$ es base de $\mathcal{W}_j$ se tiene que

$$\sum_{i=1}^{d_j} \alpha_i^j \; v_i^j = 0 \implies \alpha_i^j = 0 \quad \text{para} \quad 1 \leq i \leq d_j\,,$$

donde hemos usado la independencia lineal de B_j. Si miramos con cuidado, hemos verificado que todos los coeficientes

$$\alpha_i^j = 0 \quad , \quad 1 \leq i \leq d_j \qquad 1 \leq j \leq k\,,$$

que muestra que B es l.i.

B genera a $\mathcal{W}$: si $w \in \mathcal{W}$ entonces, por la Definición 6.12, existen $w_j \in \mathcal{W}_j$, $1 \leq j \leq k$ tales que

$$w = w_1 + \ldots + w_k\,.$$

Como B_j es base de $\mathcal{W}_j$ entonces existen coeficientes $\alpha_i^j \in K$, $1 \leq i \leq d_j$ tales que

$$w_j = \sum_{i=1}^{d_j} \alpha_i^j \; v_i^j \implies w = \sum_{j=1}^{k} w_j = \sum_{j=1}^{k} \sum_{i=1}^{d_j} \alpha_i^j \; v_i^j\,.$$

La identidad anterior muestra que B es un sistema de generadores de $\mathcal{W}$.

$2 \implies 1$. Ejercicio para el lector.

Finalmente, si suponemos que los subespacios son independientes, entonces la implicación $1 \implies 2$. probada más arriba muestra que B es base de $\mathcal{W}$. En particular,

$$\dim_K \mathcal{W} = \#(B) = \#(B_1) + \ldots + \#(B_k) = \dim_K \mathcal{W}_1 + \ldots + \dim_K(\mathcal{W}_k)$$

(donde $\#(A)$ denota la cantidad de elementos del conjunto A). □

Corolario 6.15. *Sea $\mathcal{V}$ un K-ev tal que $\dim_K(\mathcal{V}) = n$ y sea $T \in L(\mathcal{V})$. Sean $\lambda_1, \ldots, \lambda_k$ los autovalores* **distintos** *de T y sean $B_j = \{v_1^j, \ldots, v_{d_j}^j\}$ base del autoespacio $E_T(\lambda_j)$ para $1 \leq j \leq k$. Entonces*

1. *$B = B_1 \cup \ldots \cup B_k$ (por yuxtaposición) es linealmente independiente.*
2. *Si $\sum_{j=1}^{k} \dim_K E_T(\lambda_j) = n$ entonces T es diagonalizable.*

Demostración. Por la Proposición 6.13, los autoespacios $E_T(\lambda_j)$, $1 \le j \le k$, son independientes. Por el Teorema 6.14, B es una base de

$$\mathcal{W} = E_T(\lambda_1) \dot{+} \ldots \dot{+} E_T(\lambda_k) \subseteq \mathcal{V}\,.$$

En particular, B es conjunto l.i.

Finalmente, si $\sum_{j=1}^{k} \dim_K E_T(\lambda_j) = n$ entonces, por el Teorema 6.14, $\dim_K \mathcal{W} = n$ de forma $\mathcal{W} = \mathcal{V}$ y B es base de $\mathcal{V}$. Notemos que todo elemento de B es de la forma $v_i^j \in B_j$ para algún $1 \le i \le d_j$ y algún $1 \le j \le k$. En particular, como $B_j \subset E_T(\lambda_j)$ entonces

$$T\, v_i^j = \lambda_j\ v_i^j\,.$$

Así, B es base de $\mathcal{V}$ formada por autovectores de T. Entonces T es diagonalizable, por el Teorema 5.5. □

Ejemplo 6.16. Recordemos el Ejemplo 5.2. Así, sea $\mathcal{V} = \mathbb{R}^2$ y $T \in L(\mathbb{R}^2)$ dado por

$$T(x,y) = (x+y, x+y) \quad \text{para} \quad (x,y) \in \mathbb{R}^2\,.$$

Vamos a intentar diagonalizar a T: para esto, aplicamos la estrategia que describimos a comienzo de esta sección:

1. Para calcular $p_T(x)$ consideramos la base canónica de $\mathbb{R}^2$, $B = \{e_1, e_2\}$ entonces,

   $$[T]_B = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \implies p_T(x) = \det(\begin{pmatrix} x-1 & -1 \\ -1 & x-1 \end{pmatrix}) = (x-1)^2 - 1 = (x-2)\,x$$

   Las raíces distintas de p_T son $\lambda_1 = 2$ y $\lambda_2 = 0$.

2. Calculamos los autoespacios $E_T(\lambda_1)$:

   $$\begin{aligned} E_T(\lambda_1) &= \{(x,y) \in \mathbb{R}^2 \ :\ (2\,I - T)(x,y) = (0,0)\} \\ &= \{(x,y) \in \mathbb{R}^2 \ :\ 2(x,y) - (x+y, x+y) = (0,0)\} \\ &= \{(x,y) \ :\ x - y = 0,\, y - x = 0\} = \{(x,x) \ :\ x \in \mathbb{R}\}\,. \end{aligned}$$

   Notemos que $\{v_1^1 = (1,1)\}$ es una base de $E_T(\lambda_1)$.

   Calculamos los autoespacios $E_T(\lambda_2)$:

   $$\begin{aligned} E_T(\lambda_2) &= \{(x,y) \in \mathbb{R}^2 \ :\ (0\,I - T)(x,y) = (0,0)\} \\ &= \{(x,y) \in \mathbb{R}^2 \ :\ -(x+y, x+y) = (0,0)\} \\ &= \{(x,y) \ :\ x + y = 0\,\} = \{(x,-x) \ :\ x \in \mathbb{R}\}\,. \end{aligned}$$

   Notemos que $\{v_1^2 = (1,-1)\}$ es una base de $E_T(\lambda_2)$.

3. El Corolario 6.15 **garantiza** que el conjunto que se obtiene de pegar las bases de cada autoespacio (yuxtaposición) encontradas más arriba

   $$B' = \{v_1^1,\, v_1^2\} \quad \text{es l.i.}$$

   Más aún, como $\#(B') = 2$ (B' tiene dos elementos) y $\dim_{\mathbb{R}} \mathcal{V} = 2$ entonces B' es base de $\mathcal{V}$ formada por autovectores de T.

4. De hecho,

   $$[T]_{B'} = \begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix}\,.$$

Volver a mirar el Ejemplo 5.2!!

6.2 Caracterización de operadores diagonalizables

En las notas de la 1er parte del tp 5 hemos descrito una estrategia para diagonalizar un operador lineal $T \in L(\mathcal{V})$. Más aún, el Corolario 6.15 muestra que en el caso en que la suma de las dimensiones de los autoespacios correspondientes a los autovalores distintos de T coincida con la dimensión de $\mathcal{V}$, entonces la estrategia permite diagonalizar a T. Pero qué sucede si esta útlima condición falla? Veremos si esta condición falla entonces T no es diagonalizable. Como moraleja: si nuestra estrategia (particular) no diagonaliza a T entonces T no es diagonalizable!

De hecho, los siguientes resultados permiten mejorar notablemente nuestra estrategia para intentar diagonalizar un operador (ver la Observación 6.20 más abajo).

Teorema 6.17. *Sea $\mathcal{V}$ un K-ev tal que $\dim_K(\mathcal{V}) = n$ y sea $T \in L(\mathcal{V})$. Sean $\lambda_1, \ldots, \lambda_k$ (todos) los autovalores distintos de T y sean $d_j = \dim_K E_T(\lambda_j)$, para $1 \leq j \leq k$. Son equivalentes:*

1. *T es diagonalizable;*
2. *El polinomio característico de T se puede escribir como*
$$p_T(x) = (x - \lambda_1)^{d_1} \cdots (x - \lambda_k)^{d_k}$$
3. *$\sum_{j=1}^{k} \dim_K E_T(\lambda_j) = \sum_{j=1}^{k} d_j = n$.*

Demostración. 1. $\implies$ 2. Si T es diagonalizable, sea $B = \{v_1, \ldots, v_n\}$ una base de $\mathcal{V}$ tal que

$$[T]_B = \begin{pmatrix} \mu_1 & 0 & 0 & \ldots & 0 \\ 0 & \mu_2 & 0 & \ldots & 0 \\ 0 & 0 & \ddots & \ldots & 0 \\ \vdots & \vdots & 0 & \ddots & 0 \\ 0 & 0 & 0 & \ldots & \mu_n \end{pmatrix} \in K^{n \times n} .$$

En este caso, hemos visto que $Tv_i = \mu_i \, v_i$, $1 \leq i \leq n$. Así, cada μ_i es autovalor de T (pues $v_i \neq 0$): entonces cada μ_i coincide con algún λ_j para cierto $1 \leq j \leq k$ (que depende de i) porque $\lambda_1, \ldots, \lambda_k$ son los autovalores distintos de T. Por otro lado, dado $1 \leq j \leq k$ debe haber algún $1 \leq i \leq n$ tal que $\lambda_j = \mu_i$: de otra forma, $\lambda_j \neq \mu_i$ para $1 \leq i \leq n$, y la matriz

$$[\lambda_j \, I - T]_B = \operatorname{diag}(\lambda_j - \mu_1, \ldots, \lambda_j - \mu_n) \in K^{n \times n}$$

es una matriz inversible, pues es una matriz diagonal con entradas no nulas en su diagonal principal; este hecho implica que $\lambda_j \, I - T \in L(\mathcal{V})$ es operador inversible, que contradice el hecho de que λ_j sea autovalor de T (notar que $\{0\} \neq E_T(\lambda_j) = N(\lambda_j \, I - T)$ de forma que $\lambda_j \, I - T$ no es monomorfimos).

Podemos reordenar (permutar) los vectores en B de forma de ubicar los vectores v_i's tales que $Tv_i = \lambda_1 \, v_i$ al comienzo; luego ubicamos el grupo de vectores v_i's tales que $Tv_i = \lambda_2 \, v_i$, y continuamos de esta forma, agrupando los vectores tales que $Tv_i = \lambda_j \, v_i$, y ubicándolos en el j-ésimo grupo, $1 \leq j \leq k$. De esta forma, obtenemos la base (ordenada) B' que resulta de permutar los vectores de B con el criterio anterior. De hecho, podemos describir B' aprovechando los agrupamientos de vectores descritos antes, como

$$B' = \{ \underbrace{v_1^1, \ldots, v_{r_1}^1}_{\text{1er grupo}} , \underbrace{v_1^2, \ldots, v_{r_2}^2}_{\text{2do grupo}} , \ldots , \underbrace{v_1^k, \ldots, v_{r_k}^k}_{\text{k-ésimo grupo}} \}$$

donde

$$Tv_i^j = \lambda_j \, v_i^j \ , \ 1 \leq i \leq r_j \ , \ 1 \leq j \leq k \, .$$

En este caso, $[T]_{B'}$ es la matriz diagonal dada por

$$[T]_{B'} = \begin{pmatrix} \lambda_1 & 0 & 0 & 0 & \dots & 0 & \dots & 0 \\ 0 & \ddots & \vdots & 0 & \dots & 0 & \dots & 0 \\ \vdots & 0 & \lambda_1 & 0 & \dots & 0 & \dots & 0 \\ 0 & 0 & 0 & \lambda_2 & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & 0 & \ddots & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & 0 & \dots & \lambda_k & \dots & 0 \\ \vdots & \vdots & \vdots & 0 & \dots & 0 & \ddots & 0 \\ \underbrace{0 \quad 0 \quad 0}_{r_1} & & & 0 \dots 0 & & \underbrace{0 \quad \dots \quad \lambda_k}_{r_k} & & \end{pmatrix} \in K^{n\times n}\,.$$

En particular

$$x\,I - [T]_{B'} = \operatorname{diag}(\underbrace{x-\lambda_1, \dots, x-\lambda_1}_{r_1 \text{ veces}}, x-\lambda_2, \dots, x-\lambda_{k-1}, \underbrace{x-\lambda_k, \dots, x-\lambda_k}_{r_k \text{ veces}})\,.$$

Como consecuencia de la definición de polinomio caracteristico de T (y usando el hecho de que el determinante de una matriz diagonal es el producto de las entradas en su diagonal principal):

$$p_T(x) = \det(x\,I - [T]_{B'}) = (x-\lambda_1)^{r_1} \cdots (x-\lambda_k)^{r_k} \tag{29}$$

Como hemos partido a B' en k subconjuntos de r_j elementos cada uno $1 \leq j \leq k$, entonces $\sum_{j=1}^k r_j = n$. Otra forma de ver esto es usando que $\sum_{j=1}^k r_j$ es el grado del polinomio p_T, que siempre es $\dim_K \mathcal{V} = n$.

Además, si fijamos $1 \leq j \leq k$, entonces:

$$T v_i^j = \lambda_j\, v_i^j \;\;,\;\; 1 \leq i \leq r_j \implies \{v_1^j, \dots, v_{r_j}^j\} \subset E_T(\lambda_j) \quad \text{es conjunto l.i.} \quad .$$

Lo anterior muestra que $r_j \leq \dim_K E_T(\lambda_j) = d_j$. Como

$$\sum_{j=1}^k d_j = \dim_K(E_T(\lambda_1)\dot{+} \dots \dot{+} E_T(\lambda_k)) \leq n$$

pues los autoespacios son independientes por la Proposición 6.13 y el Teorema 6.14 (y recordemos que la dimensión de cualquier subespacio es menor ó igual que n). Entonces vemos que

$$n = \sum_{j=1}^k r_j \leq \sum_{j=1}^k d_j \leq n \implies r_j = d_j \quad , \quad 1 \leq j \leq k\,,$$

pues en las desigualdades $r_j \leq d_j$ no puede haber desigualdades estrictas! Así, usando la identidad en la Eq. (29)

$$p_T(x) = (x-\lambda_1)^{d_1} \cdots (x-\lambda_k)^{d_k}\,.$$

De esta forma, se verifica el item 2.

2. $\implies$ 3. Si vale 2, usando que $p_T(x)$ es polinomio mónico de grado $n = \dim_K \mathcal{V}$, concluimos que n=grado de $p_T(x) = \sum_{j=1}^k d_j$ y vale el item 3.

3. $\implies$ 1. Por el Corolario 6.15, si vale el ítem 3, T es diagonalizable. $\square$

Obs 6.18. Con las notaciones de la prueba del Teorema anterior, A veces resulta conveniente enfatizar la estructura de bloques de la matriz $[T]_{B'}$, y escribimos

$$[T]_{B'} = \begin{pmatrix} \lambda_1 I_{r_1} & 0 & 0 & \dots & 0 \\ 0 & \lambda_2 I_{r_2} & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \dots & 0 \\ 0 & 0 & 0 & \dots & \lambda_k I_{r_k} \end{pmatrix} \in K^{n\times n},$$

donde las entradas de la matriz anterior son matrices (si, matrices!) también llamadas bloques, de forma que $I_{r_j} \in K^{r_j\times r_j}$ denota la matriz identidad de tamaño r_j, para $1 \le j \le k$, y la matriz 0 que aparece en la entrada i, j $(i \ne j)$ es de tamaño $r_i \times r_j$. △

Obs 6.19 (Diagonalización y matrices de operadores)**.** Sea $\mathcal{V}$ un K-ev tal que $\dim_K(\mathcal{V}) = n$ y sea $T \in L(\mathcal{V})$. Sea B una base (arbitraria) de $\mathcal{V}$ y sea $A = [T]_B \in K^{n\times n}$. Recordemos que en este caso

$$[T\,v]_B = A\,[v]_B \quad \text{para todo} \quad v \in \mathcal{V}\,. \tag{30}$$

Además, recordemos que la transformación $C_B : \mathcal{V} \to K^n$, $C_B(v) = [v]_B$, es un isomorfismo de espacios vectoriales.

1. Entonces C_B identifica $\mathrm{N}(T) = \{v \in \mathcal{V} \ : \ Tv = 0\}$ con $\mathrm{N}(A) = \{\vec{x} \in K^n \ : \ A\vec{x} = \vec{0}\}$ como espacios vectoriales: concretamente

   $$C_B(\mathrm{N}(T)) = \{[v]_B \ : \ v \in \mathrm{N}(T)\} = \mathrm{N}(A)\,.$$

   En efecto, $Tv = 0$ si y solo si $\vec{0} = [Tv]_B = A\,[v]_B$ si y solo si $[v]_B \in \mathrm{N}(A)$. Así, la restricción $C_B : \mathrm{N}(T) \to \mathrm{N}(A)$ es un isomorfismo. Las observaciones anteriores muestran que $\{v_1, \dots, v_k\}$ es base de $\mathrm{N}(T)$ si y solo si $\{[v_1]_B, \dots, [v_k]_B\}$ es base de $\mathrm{N}(A)$. En particular

   $$\dim_K \mathrm{N}(T) = \dim_K \mathrm{N}(A)\,.$$

2. La observación anterior se aplica a los autoespacios de T: si λ es autovalor de T entonces

   $$E_T(\lambda) = \mathrm{N}(\lambda\, I - T) \text{ se identifica con } \mathrm{N}(\lambda\, I - A) \subseteq K^n\,.$$

   en el sentido del ítem anterior, usando una restricción del isomofismo C_B conveniente. En este contexto $E_A(\lambda) = \mathrm{N}(\lambda\, I - A)$ se llama el autoespacio de A asociado a λ. Las observaciones anteriores muestran que $\{v_1, \dots, v_k\}$ es base de $\mathrm{N}(\lambda\, I - T)$ si y solo si $\{[v_1]_B, \dots, [v_k]_B\}$ es base de $\mathrm{N}(\lambda\, I - A)$. En particular,

   $$\dim_K \mathrm{N}(\lambda\, I - T) = \dim_K \mathrm{N}(\lambda\, I - A)\,.$$

△

Obs 6.20. El Teorema 6.17, junto con la Observación 6.19 tienen la siguiente consecuencia importante para intentar diagonalizar un operador T.

Sea $\mathcal{V}$ un K-ev tal que $\dim_K(\mathcal{V}) = n$ y sea $T \in L(\mathcal{V})$. Sea B una base (arbitraria) de $\mathcal{V}$ y sea $A = [T]_B \in K^{n\times n}$.

1. Calculamos $p_T(x) = \det(x\,I - A)$, el polinomio característico de T;

2. Hallamos las raíces *distintas* de $p_T(x)$, $\lambda_1, \dots, \lambda_k$ $(k \le n)$. Si el polinomio $p_T(x) \in K[x]$ no se puede factorizar como producto de factores lineales, es decir, no se puede escribir como

   $$(x - \lambda_1)^{r_1} \cdots (x - \lambda_k)^{r_k}$$

   para ciertas potencias r_i, entonces T no es diagonalizable.

3. Si $p_T(x) = (x-\lambda_1)^{r_1}\cdots(x-\lambda_k)^{r_k}$: en este caso $\sum_{j=1}^k r_j = n$ (el grado de p_T es n).

   (a) Para cada autovalor λ_j de T, calculamos una base $B_j = \{\vec{x}_1^j, \ldots, \vec{x}_{d_j}^j\}$ para el núcleo
$$N(\lambda_j\, I - A)\ ,\quad 1 \le j \le k\,.$$
   Un hecho **fundamental** es que podemos calcular B_j resolviendo un sistema de ecuaciones lineales homogéneo con matriz dada por $\lambda\, I - A \in K^{n\times n}$. Recordemos que tenemos el método de reducción por filas para poder calcular este subespacio y una base para este subespacio (con la técnica de resolver las variables dependientes en función de las variables libres del sistema reducido, asociado a la matriz reducida y escalonada por filas equivalente a $\lambda\, I - A$).

   (b) Por la Observación 6.19, se tiene que $\dim_K E_T(\lambda_j) = d_j$, la cantidad de vectores de B_j. Si existe un $1 \le j \le k$ tal que
$$d_j = \dim_K E_T(\lambda_j) \neq r_j$$
   entonces T no es diagonalizable, por el Teorema 6.17.

   (c) Por otro lado, si para cada $1 \le j \le k$ se verifica: $d_j = \dim_K E_T(\lambda_j) = r_j$, entonces T es diagonalizable, porque $\sum_{j=1}^k d_j = \sum_{j=1}^k r_j = n$.

   Más aún, en este caso, si $B_j' = \{v_1^j, \ldots, v_{d_j}^j\}$ tales que $[v_i^j]_B = \vec{x}_i^j$, $1 \le i \le d_j$, entonces B_j' es base de $E_T(\lambda_j)$, $1 \le j \le k$. Finalmente, si $B' = B_1' \cup \ldots \cup B_k'$ entonces B' es base de $\mathcal{V}$ tal que $[T]_{B'}$ es matriz diagonal.

$\triangle$

Importante: con la notación de la Observación anterior, la condición $\dim_K E_T(\lambda_j) \neq r_j$ sólo se puede dar en casos en donde $r_j \ge 2$.

Por otro lado, una condición equivalente a que T sea diagonalizable es: $p_T(x) = (x-\lambda_1)^{r_1}\cdots(x-\lambda_k)^{r_k}$ y la multiplicidad algebraica de λ_j coincide con la multiplicidad geométrica de λ_j, para cada $1 \le j \le k$. (ver las notas de la 1er parte para las definiciones de las multiplicidades).

Ejemplo 6.21. Sea $T \in L(\mathbb{R}^3)$ un operador tal que su polinomio característico es
$$p_T(x) = (x-3)^2\, x\,.$$
Los autovalores de T son $\lambda_1 = 3$ y $\lambda_2 = 0$ (las raíces de $P_T(x)$). En este caso, la multiplicidad algebraica de $\lambda_1 = 3$ es $r_1 = 2$, mientras que la multiplicidad algebraica de $\lambda_2 = 0$ es $r_2 = 1$. Supongamos que $d_1 = \dim E_T(3) = \dim N(3I - T) = 1$: entonces T no es diagonalizable, pues en este caso la multiplicidad geométrica de $\lambda_1 = 1$ es $d_1 = 1 < 2 = r_1$ (es decir, $p_T(x) \neq (x-3)^{d_1}\, x = (x-3)\, x$. $\triangle$

Obs 6.22 (Extensión de varios conceptos a matrices)**.** Sea K un cuerpo, y sea $A \in K^{n\times n}$ matriz **cuadrada**. En este caso, podemos definir el operador lineal $T_A \in L(K^n)$, dado por
$$T_A\, \vec{x} = A\, \vec{x} \quad \text{ para } \quad \vec{x} \in K^n\,.$$
Definimos los autovalores de A, autovectores de A, los autoespacios de A y el polinomio característico de A como los autovalores, autovectores, autoespacios y polinomio característico de T_A. De hecho, si $B_c = \{e_1, \ldots, e_n\}$ denota la base canonica en K^n entonces $[T_A] = A$ y se tiene:

1. $\lambda \in K$ es autovalor de A si existe $\vec{x} \neq 0$ tal que $A\vec{x} = \lambda\, \vec{x}$;

2. Si $\lambda \in K$ es autovalor de A entonces definimos el autoespacio de A asociado a λ, notado $E_A(\lambda)$ dado por

$$E_A(\lambda) = \mathrm{N}(\lambda\, I - A) = \{\vec{y} \in K^n \ : \ (\lambda\, I - A)\, \vec{y} = \vec{0}\}\,.$$

Notemos que en este caso, podemos calcular el autoespacio resolviendo el sistema de ecuaciones lineales homogéneo con matriz dada por $\lambda\, I - A \in K^{n\times n}$. Recordemos que tenemos el método de reducción por filas para poder calcular este subespacio y una base para este subespacio.

3. El polinomio característico de A es $p_{T_A}(x) = \det(x\, I - [T_A]_{B_c}) = \det(x\, I - A) = p_A(x)$.

4. Decimos que A es diagonalizable si T_A es diagonalizable.

5. Notemos que A es diagonalizable si y solo si A es semejante a una matriz diagonal, es decir: A es diagonalizable si y solos si existe una matriz inversible $P \in K^{n\times n}$ tal que

$$A = P^{-1}\, \mathrm{diag}(\mu_1, \ldots, \mu_n)\, P$$

donde $\mu_1, \ldots, \mu_n$ son autovalores de A, pero que pueden aparecer repetidos (ejercicio. sugerencia: notar $A = [T_A]_{B_c}$ y si B es cualquier base de K^n entonces $[T_A]_{B_c}$ y $[T_A]_B$ son matrices semejantes, por la fórmula de cambio de base de la matriz de T).

Por supuesto, todos los resultados de diagonalización que hemos visto se aplican al operador $T_A \in L(K^n)$, lo que nos permite determinar si A es diagonalizable o no. $\triangle$

6.3 Sumas directas - sistemas de proyecciones

Comenzamos con una serie de nociones nuevas.

Definición 6.23. *Sea $\mathcal{V}$ un K-ev,* $\dim_K \mathcal{V} = n$.

1. *Una proyección E es un operador $E \in L(\mathcal{V})$ tal que $E \cdot E = E$ (es decir, $E^2 = E$).*

2. *Un sistema de proyecciones es un conjunto ordenado $\{E_1, \ldots, E_k\} \subset L(\mathcal{V})$ tal que*

   (a) $E_i\, E_j = \delta_{ij}\, E_i$, $1 \leq i, j \leq k$ (delta de Kronecker);
   (b) $E_1 + \ldots + E_k = I$.

$\square$

Obs 6.24. Con las notaciones de la definición anterior, si $\{E_1, \ldots, E_k\} \subset L(\mathcal{V})$ es sistema de proyecciones entonces

$$E_i\, E_i = \delta_{ii}\, E_i = E_i \implies E_i \ \text{ es proyección}$$

. $\square$

Obs 6.25. Con las notaciones de la definción anterior, sea $\{E_1, E_2\}$ un sistema de proyecciones (tomamos $k = 2$ porque este es un caso particularmente sencillo). Entonces, por la Observación anterior E_1 y E_2 son proyecciones, (usando las propiedades de la delta de Kronecker)

$$E_1\, E_2 = E_2\, E_1 = 0 \quad \text{y} \quad E_1 + E_2 = I\,.$$

Sea

$$\mathcal{W}_1 = \mathrm{Im}(E_1) \quad \text{y} \quad \mathcal{W}_2 = \mathrm{Im}(E_2)\,.$$

Así, $\mathcal{W}_1$ y $\mathcal{W}_2$ son subespacios de $\mathcal{V}$. Notemos que como $E_1 + E_2 = I$ entonces, si $v \in \mathcal{V}$

$$v = I\, v = (E_1 + E_2)\, v = E_1\, v + E_2\, v \in \mathcal{W}_1 + \mathcal{W}_2$$

pues $E_1\, v \in \mathcal{W}_1$ y $E_2\, v \in \mathcal{W}_2$: como $v \in \mathcal{V}$ era arbitrario, entonces $\mathcal{W}_1 + \mathcal{W}_2 = \mathcal{V}$.

Por otro lado, si $w_1 \in \mathcal{W}_1$ entonces existe $v_1 \in \mathcal{V}$ tal que $w_1 = E_1\, v_1$: así,

$$E_1 w_1 = E_1(E_1 v_1) = E_1 \cdot E_1(v_1) = E_1(v_1) = w_1$$

es decir $E_1 w_1 = w_1$; decimos que **el proyector E_1 fija los vectores de su imágen** (este es un hecho general, que vale para cualquier proyector: ejercicio).

De forma completamente análoga, si $w_2 \in \mathcal{W}_2$ entonces $E_2\, w_2 = w_2$. En particular

$$E_1\, w_2 = E_1\,(E_2 w_2) = (E_1\, E_2) w_2 = 0_{L(\mathcal{V})}\, w_2 = 0\,.$$

Lo anterior permite verificar que los subespacios $\mathcal{W}_1$ y $\mathcal{W}_2$ son independientes: en efecto, si $w_1 \in \mathcal{W}_1$, $w_2 \in \mathcal{W}_2$ son tales que $w_1 + w_2 = 0$ entonces

$$0 = w_1 + w_2 \implies 0 = E_1\, 0 = E_1(w_1 + w_2) = E_1(w_1) + E_1(w_2) = w_1 + 0 = w_1\,.$$

Así, $w_1 = 0$ lo que implica que $w_2 = 0$. En resumen hemos probado que

$$\mathcal{V} = \mathcal{W}_1 \dot{+} \mathcal{W}_2\,.$$

Recíprocamente, si partimos de dos subespacios $\mathcal{W}_1$, $\mathcal{W}_2$ tales que $\mathcal{V} = \mathcal{W}_1 \dot{+} \mathcal{W}_2$, entonces podemos construir un (único) sistema de proyecciones $\{E_1,\, E_2\}$ tal que $\mathrm{Im}(E_i) = \mathcal{W}_i$, $i = 1, 2$ (ver la Observación 6.28). De hecho, veremos que podemos extender estas observaciones al caso de sistemas con más de dos proyecciones (ver Teorema 6.29 más adelante). $\triangle$

Obs 6.26. Todo proyector $E \in L(\mathcal{V})$ es diagonalizable: en efecto, si E es proyector, sea $F = I - E$. Entonces, aplicando la propiedada distributiva:

$$F \cdot F = (I - E) \cdot (I - E) = I - E - E + E \cdot E = I - E - E + E = I - E = F\,,$$

donde usamos que $E^2 = E \cdot E = E$. Además, $E\,F = E\,(I - E) = E - E^2 = 0$, y $E + F = I$. Entonces $\{E, F\} = \{E, I - E\}$ es sistema de proyecciones. Si $\mathcal{W}_1 = \mathrm{Im}(E)$ y $\mathcal{W}_2 = \mathrm{Im}(F)$ entonces en la Observación anterior probamos: $\mathcal{V} = \mathcal{W}_1 \dot{+} \mathcal{W}_2$. Así, si B_1 es base de $\mathcal{W}_1$ y B_2 es base de $\mathcal{W}_2$ entonces $B = B_1 \cup B_2$ es base de $\mathcal{V}$, por el Teorema 6.13.

Si $v \in B$ está en B_1 entonces $Ev = v$, pues E fija los vectores de $\mathcal{W}_1$ (que es su imágen). Así v es autovector de E con autovalor 1.

Si $v \in B$ está en B_2 entonces $Ev = E(F(v)) = (EF)\, v = 0_{L(\mathcal{V})}\, v = 0 = 0_K\, v$. Entonces v es autovector de E con autovalor 0_K.

Así B es base de $\mathcal{V}$ formada por autovectores de E de forma que E es diagonalizable. En particular $[E]_B$ es matriz diagonal, que tiene en su diagonal principal 1's y 0's (hacer esquema). $\triangle$

Lema 6.27. *Sea $\mathcal{V}$ un K-ev y sean $\mathcal{W}_1, \ldots, \mathcal{W}_k$ subespacios independientes de $\mathcal{V}$. Si $w \in W_1 \dot{+} \ldots \dot{+} W_k$ entonces existen únicos $w_i \in \mathcal{W}_i$, $1 \le i \le k$, tales que $w = w_1 + \ldots + w_k$.*

Demostración. Por construcción de la suma de subespacios, sabemos que dado $w \in W_1 \dot{+} \ldots \dot{+} W_k$ entonces existen $w_i \in \mathcal{W}_i$, $1 \le i \le k$, tales que $w = w_1 + \ldots + w_k$. Veamos la unicidad: sean w_i, $w_i' \in \mathcal{W}_i$, $1 \le i \le k$ tales que

$$w_1 + \ldots + w_k = w = w_1' + \ldots + w_k' \implies (w_1 - w_1') + \ldots + (w_k - w_k') = 0\,.$$

Como cada diferencia $w_i - w_i' \in \mathcal{W}_i$, $1 \le i \le k$, y los subespacios son independientes, entonces concluimos que

$$w_i - w_i' = 0 \implies w_i = w_i' \ \ ,\ 1 \le i \le k\,.$$

$\square$

Obs 6.28. Realizamos una construcción recíproca a la realizada en la Observación 6.25. En este caso, partimos de dos subespacios independientes $\mathcal{W}_1$ y $\mathcal{W}_2$ tales que

$$\mathcal{W}_1 \dot{+} \mathcal{W}_2 = \mathcal{V}\,. \tag{31}$$

Construimos $E_1,\, E_2 \in L(\mathcal{V})$ como sigue: dado $v \in \mathcal{V}$, por el Lemma 6.27, existen únicos $w_1 \in \mathcal{W}_1$ y $w_2 \in \mathcal{W}_2$ tales que $v = w_1 + w_2$. En este caso, definimos

$$E_1(v) = w_1 \quad \text{ y } \quad E_2(v) = w_2\,.$$

Así, quedan definidas las funciones $E_1,\, E_2 : \mathcal{V} \to \mathcal{V}$. Veremos en la prueba del Teorema 6.29 que $\{E_1,\, E_2\}$ es un sistema de proyecciones tal que $\mathrm{Im}(E_i) = \mathcal{W}_i$, $i = 1, 2$.

Para finalizar, notemos que (asumiendo que $\{E_1,\, E_2\}$ tiene las propiedades mencionadas más arriba) en este caso $N(E_1) = \mathcal{W}_2$ y $N(E_2) = \mathcal{W}_1$. Verificamos $N(E_1) = \mathcal{W}_2$ y la otra identidad es ejercicio: si $w_2 \in \mathcal{W}_2$ entonces

$$E_1 w_2 = E_1\, E_2\, w_2 = 0\, w_2 = 0$$

donde usamos que el proyector E_2 fija los vectores de su imágen, de forma que $\mathcal{W}_2 \subset \mathrm{N}(E_1)$. Además, si $w \in \mathrm{N}(E_1)$ entonces $E_1 + E_2 = I$ implica que $E_2 = I - E_1$ y luego

$$E_2 w = (I - E_1)(w) = w - E_1 w = w - 0 = w$$

de forma que $w \in \mathcal{W}_2$. Entonces, $\mathrm{N}(E_1) \subseteq \mathcal{W}_2$. Los dos hechos anteriores muestran que $N(E_1) = \mathcal{W}_2$. En este caso, también notamos

$$E_1 = P_{\mathcal{W}_1 // \mathcal{W}_2}$$

para indicar que E_1 es un proyector cuya imágen es $\mathcal{W}_1$ y cuyo núcleo es $\mathcal{W}_2$ (y es uno de los dos proyectores del sistema asociado a la descomposición en Eq. (31)). Notemos que $E_2 = I - E_1 = P_{\mathcal{W}_2 // \mathcal{W}_1}$. △

El siguiente resultado extiende los hechos desarrollados en las Observaciones 6.25 y 6.28.

Teorema 6.29. *Sea $\mathcal{V}$ un K-ev.*

1. *Si asumimos que tenemos una descomposición en suma directa*

$$\mathcal{V} = \mathcal{W}_1 \dot{+} \ldots \dot{+} \mathcal{W}_k \tag{32}$$

entonces existe un único sistema de proyecciones $\{E_1, \ldots, E_k\}$ tal que $Im(E_i) = \mathcal{W}_i$, $1 \le i \le k$.

2. *Recíprocamente, dado un sistema de proyecciones $\{E_1, \ldots, E_k\}$ si definimos $\mathcal{W}_i = Im(E_i)$, $1 \le i \le k$, entonces se verifica la Eq.* (32)

Demostración. Supongamos que tenemos una descomposición de $\mathcal{V}$ en suma directa como en la Eq. (32). Definimos las funciones $E_1, \ldots,\, E_k : \mathcal{V} \to \mathcal{V}$ como sigue: dado $v \in \mathcal{V}$, por el Lema 6.27 existen únicos $w_i \in \mathcal{W}_i$, $1 \le i \le k$ tales que $v = w_1 + \ldots + w_k$: en este caso definimos

$$E_i(v) = w_i \in \mathcal{W}_i \quad , \; 1 \le i \le k\,.$$

Así, quedan bien definidas las funciones $E_i : \mathcal{V} \to \mathcal{V}$, $1 \le i \le k$. Por construcción, $\mathrm{Im}(E_i) \subseteq \mathcal{W}_i$. Notemos que de hecho vale la igualdad de conjuntos: si fijamos $1 \le i \le k$ y elegimos $w_i \in \mathcal{W}_i \subset \mathcal{V}$ entonces

$$w_i = 0 + \ldots + 0 + w_i + 0 + \ldots + 0\,,$$

donde cada términos distinto del i-ésimo es $0 \in \mathcal{W}_j$, $1 \le j \le k$, $j \neq i$, lo que representa a w_i como suma de vectores de cada $\mathcal{W}_h$, $1 \le h \le k$. Así, por definición,

$$E_i(w_i) = w_i \quad \text{y} \quad E_j(w_i) = 0 \quad \text{para} \quad 1 \le j \le k \ , \ \ j \neq i \, . \tag{33}$$

En particular, $\mathcal{W}_i \subseteq \operatorname{Im}(E_i)$ lo que muestra la igualdad $\operatorname{Im}(E_i) = \mathcal{W}_i$, $1 \le i \le k$. Más aún, si $v \in \mathcal{V}$ es arbitrario, tomando $w_i = E_i\, v \in \mathcal{W}_i$ vemos que

$$E_i(E_i\, v) = E_i\, v \implies E_i\, E_i = E_i \ , \ 1 \le i \le k \, .$$

Veamos que $E_i \in L(\mathcal{V})$, $1 \le i \le k$: en efecto, si u, $v \in \mathcal{V}$, $\alpha \in K$ entonces existen únicos w_i, $w_i' \in \mathcal{W}_i$, $1 \le i \le k$, tales

$$v = w_1 + \ldots + w_k \quad \text{y} \quad u = w_1' + \ldots + w_k' \implies E_i(v) = w_i \ \ , \ \ E_i(u) = w_i'$$
$$\text{y} \quad \alpha\, v + u = (\alpha\, w_1 + w_1') + \ldots + (\alpha\, w_k + w_k') \, .$$

Como cada $\alpha\, w_i + w_i' \in \mathcal{W}_i$, por la unicidad de la descomposición garantizada por la independencia de los subespacios, se tiene que

$$E_i(\alpha\, v + u) = \alpha\, w_i + w_i' = \alpha \ E_i(v) + E_i(u) \ \ , \ 1 \le i \le k \, .$$

Así, $E_i \in L(\mathcal{V})$, $\operatorname{Im}(E_i) = \mathcal{W}_i$ y $E_i\, E_i = E_i$. Por construcción $E_1 + \ldots + E_k = I$: en efecto, dado $v \in \mathcal{V}$ si $w_i \in \mathcal{W}_i$, $1 \le i \le k$ son los únicos tales que $v = w_1 + \ldots + w_k$ entonces $E_j(v) = w_j$, pero entonces

$$(E_1 + \ldots + E_k)(v) = E_1(v) + \ldots + E_k(v) = w_1 + \ldots + w_k = v \implies$$
$$(E_1 + \ldots + E_k)(v) = v = I(v) \qquad \text{para todo} \quad v \in \mathcal{V} \, .$$

Finalmente, si $v \in \mathcal{V}$ luego $w_i = E_i\, v \in \mathcal{W}_i$; como ya mostramos en la Eq. (33), si $j \neq i$, se tiene

$$E_j(w_i) = 0 \implies E_j\, E_i v = 0 \quad \text{para todo} \quad v \in \mathcal{V} \, .$$

Este último hecho muestra que $E_j\, E_i = 0$ si $1 \le i \neq j \le k$. Así, vemos también que $E_j\, E_i = \delta_{ij}\, E_i$ para $1 \le i\, , j \le k$. En resumen, $\{E_1, \ldots, E_k\}$ es un sistema de proyecciones que verifica las propiedades requeridas.

Además, es único: si $\{E_1' \, , \ldots, E_k' \, \}$ es otro sistema de proyecciones tales que $\operatorname{Im}(E_j' \,) = \mathcal{W}_j$, $1 \le j \le k$: entonces, si $v \in \mathcal{V}$ se verifica:

$$v = E_1' \ v + \ldots + E_k' \ v \quad \text{y} \quad v = E_1\, v + \ldots + E_k\, v \implies E_j' \ v = E_j\, v \in \mathcal{W}_j \ \ , \ \ 1 \le j \le k \, ,$$

por la unicidad de representación dada en el Lema 6.27. Así, $E_j v = E_j' \ v$ para $v \in \mathcal{V}$, lo que muestra que $E_j = E_j'$ para $1 \le j \le k$.

Para verificar el ítem 2, notemos que $\mathcal{W}_i = \operatorname{Im}(E_i)$ es subespacio $1 \le i \le k$. Dado $v \in \mathcal{V}$, entonces $w_i = E_i\, v \in \mathcal{W}_i$ y

$$v = (E_1 + \ldots + E_k)(v) = E_1(v) + \ldots + E_k(v) = w_1 + \ldots + w_k \, .$$

Como $v \in \mathcal{V}$ era arbitrario,

$$\mathcal{W}_1 + \ldots + \mathcal{W}_k = \mathcal{V} \, .$$

Veamos que los subespacios son independientes: si $w_i \in \mathcal{W}_i$, $1 \le i \le k$, son tales que $w_1 + \ldots + w_k = 0$ entonces, como $E_j(w_j) = w_j$ para $1 \le j \le k$ (cada proyector fija los vectores en su imágen) entonces, si $1 \le i \le k$:

$$\begin{aligned} 0 &= E_i(0) = E_i(w_1 + \ldots + w_k) = E_i(E_1\, w_1 + \ldots + E_k\, w_k) \\ &= E_i\, E_1 w_1 + \ldots + E_i\, E_k\, w_k = E_i\, E_i\, w_i = E_i\, w_i = w_i \end{aligned}$$

y $w_i = 0$, donde usamos que $E_i\, E_j = 0$ si $j \neq i$. Entonces $w_i = 0$, $1 \le i \le k$, que muestra que los subespacios son idenpendientes. Así, vale la descomposición de la Eq. (32). $\square$

6.4 Subespacios y sumas directas invariantes

En esta sección vamos a ver un concepto fundamental, que nos permitirá descomponer a un operador, de forma que cada parte en esta descomposición es *más chica* que el operador inicial (y por lo tanto más sencilla de entender).

Definición 6.30. *Sea $\mathcal{V}$ K-ev, $\dim_K \mathcal{V} = n \in \mathbb{N}$ y sea $T \in L(\mathcal{V})$.*

1. *Decimos que un subespacio $\mathcal{W} \subset \mathcal{V}$ es T-invariante (ó invariante por T) si*

$$T\,w \in \mathcal{W} \quad \textit{para todo} \quad w \in \mathcal{W}\,.$$

2. *Si $\mathcal{W}$ es T-invariante definimos la restricción $T|_{\mathcal{W}} : \mathcal{W} \to \mathcal{W}$ dada por*

$$T|_{\mathcal{W}}\,w = Tw \quad \textit{para todo} \quad w \in \mathcal{W}\,.$$

□

Obs 6.31. Con las notaciones de la definición anterior: si $\mathcal{W}$ es T invariante entonces

1. la restricción $T|_{\mathcal{W}} : \mathcal{W} \to \mathcal{W}$ está bien definida!: si $w \in \mathcal{W}$ entonces $T|_{\mathcal{W}}\,w = Tw \in \mathcal{W}$ está definida y es un vector de $\mathcal{W}$ (porque $\mathcal{W}$ es T-invariante).

2. $T|_{\mathcal{W}} \in L(\mathcal{W})$ es un **operador lineal en** $\mathcal{W}$: en efecto, si $u,\, v \in \mathcal{W}$ y $\alpha \in k$

$$T|_{\mathcal{W}}\,(\alpha\,u + v) = T(\alpha\,u + v) = \alpha\,T(u) + T(v) = \alpha\,T|_{\mathcal{W}}(u) + T|_{\mathcal{W}}(v)\,.$$

△

Obs 6.32 (La técnica de reducción - sumas directas invariantes)**.** Sea $\mathcal{V}$ un K-ev, $\dim_K \mathcal{V} = n \in \mathbb{N}$ y sea $T \in L(\mathcal{V})$. Supongamos que tenemos una descomposición en suma directa

$$\mathcal{V} = \mathcal{W}_1 \dot{+} \mathcal{W}_2 \quad \text{con} \quad \mathcal{W}_1\ ,\ \mathcal{W}_2 \quad T\text{-invariantes}\,.$$

Podemos considerar las restricciones $T_1 = T|_{\mathcal{W}_1} \in L(\mathcal{W}_1)$ y $T_2 = T|_{\mathcal{W}_2} \in L(\mathcal{W}_2)$:

$$T_j(w) = T(w) \quad \text{para} \quad w \in \mathcal{W}_j\ ,\ \ j = 1,\, 2\,.$$

1. Por un lado, T se puede describir en términos de T_1 y T_2: en efecto, si $\{E_1,\, E_2\}$ es el (único) sistema de proyecciones asociado a la descomposición $\mathcal{V} = \mathcal{W}_1 \dot{+} \mathcal{W}_2$: entonces $v = E_1 v + E_2 v$ con $E_j\,v \in \mathcal{W}_j$, $j = 1, 2$. Así, tenemos que

$$Tv = T(E_1 v + E_2 v) = T(E_1 v) + T(E_2 v) = T_1(E_1 v) + T_2(E_2 v)$$

pues $T(E_j\,v) = T_j(E_j\,v)$ por definición de la restricción T_j, $j = 1, 2$.

2. Sean $B_1 = \{v_1, \ldots, v_r\}$ y $B_2 = \{v_{r+1}, \ldots, v_n\}$ bases de $\mathcal{W}_1$ y $\mathcal{W}_2$, respectivamente. En este caso, podemos construir

$$[T_1]_{B_1} = \begin{pmatrix} | & \cdots & | \\ [T_1(v_1)]_{B_1} & \cdots & [T_1(v_r)]_{B_1} \\ | & \cdots & | \end{pmatrix} \in K^{r\times r} \quad \text{y}$$

$$[T_2]_{B_2} = \begin{pmatrix} | & \cdots & | \\ [T_2(v_{r+1})]_{B_2} & \cdots & [T_2(v_n)]_{B_2} \\ | & \cdots & | \end{pmatrix} \in K^{(n-r)\times(n-r)}\,.$$

Además, $B = B_1 \cup B_2$ es base de $\mathcal{V}$. Veamos que $[T]_B$ tiene una estructura **simplificada**: En efecto, si $1 \leq j \leq r$; entonces $v_j \in \mathcal{W}_1$ y $Tv_j = T_1\, v_j \in \mathcal{W}_1$, porque $\mathcal{W}_1$ es T-invariante: entonces Tv_j es combinación de los vectores de B_1 (que es base de $\mathcal{W}_1$)

$$Tv_j = T_1\, v_j = a_{1j}\, v_1 + \ldots + a_{rj}\, v_r = a_{1j}\, v_1 + \ldots + a_{rj}\, v_r + 0\, v_{r+1} + \ldots + 0\, v_n\,. \tag{34}$$

Por otro lado, si $r+1 \leq j \leq n$ entonces $v_j \in \mathcal{W}_2$ y $Tv_j = T_2\, v_j \in \mathcal{W}_2$ y Tv_j es combinación de los vectores de B_2

$$Tv_j = T_2\, v_j = a_{(r+1)j}\, v_{r+1} + \ldots + a_{nj}\, v_n = 0\, v_1 + \ldots + 0\, v_r + a_{(r+1)j}\, v_{r+1} + \ldots + a_{nj}\, v_n\,. \tag{35}$$

Así, las Eqs. (34) y (35) permiten calcular $[T]_B$ como

$$[T]_B = \begin{pmatrix} a_{11} & \ldots & a_{1r} & 0 & \ldots & 0 \\ \vdots & \ddots & \vdots & 0 & \ldots & 0 \\ a_{r1} & \ldots & a_{rr} & 0 & \ldots & 0 \\ 0 & 0 & 0 & a_{r+1,\,r+1} & \ldots & a_{r+1,n} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & a_{n,\,r+1} & \ldots & a_{nn} \end{pmatrix} \in K^{n\times n}\,.$$

Más aún, estas mismas ecuaciones indican que

$$[T_1]_{B_1} = \begin{pmatrix} a_{11} & \ldots & a_{1r} \\ \vdots & \ddots & \vdots \\ a_{r1} & \ldots & a_{rr} \end{pmatrix} \quad \text{y} \quad [T_2]_{B_2} = \begin{pmatrix} a_{r+1,\,r+1} & \ldots & a_{r+1,\,n} \\ \vdots & \ddots & \vdots \\ a_{n,\,r+1} & \ldots & a_{nn} \end{pmatrix}$$

Las identidades anteriores muestran que $[T]_B$ tiene una estructura de bloques que se puede describir como

$$[T]_B = \begin{pmatrix} [T_1]_{B_1} & 0_{r,\,(n-r)} \\ 0_{(n-r),\,r} & [T_2]_{B_2} \end{pmatrix}\,.$$

Cada entrada en la matriz por bloques 2×2 anterior es una matriz!: las matrices nulas que aparecen tienen tamaño indicado por los subíndices (ejemplo $0_{r,(n-r)} \in K^{r\times(n-r)}$).

Estas observaciones justifican la afirmación del comienzo de la sección: hemos descompuesto T en partes más sencillas (en este caso T_1 y T_2), de forma que si entendemos cada parte, podemos entender a T completo! pero cada parte es más sencilla de entender, porque es más chica. Por ejemplo, si podemos diagonalizar a T_1 con la base B_1 y podemos diagonalizar a T_2 con la base B_2 entonces la base B diagonaliza a T!!(hacer un esquema de esta situación en la hoja)

3. Las observaciones de los ítems anteriores se extienden al caso de la descomposición en suma directa

$$\mathcal{V} = \mathcal{W}_1 \dot{+} \ldots \dot{+} \mathcal{W}_k \tag{36}$$

donde $\mathcal{W}_i$ es subespacio T-invariante, $1 \leq i \leq k$. En este caso podemos definir las restricciones $T_i = T|_{\mathcal{W}_i} \in L(\mathcal{W}_i)$, $1 \leq i \leq k$. Como antes, se tiene que

$$Tv = \sum_{i=1}^{k} T_i(E_i\, v)$$

donde $\{E_1, \ldots, E_k\}$ es el sistema de proyecciones asociado a la descomposición en la Eq. (36). Más aún, si B_i es base de $\mathcal{W}_i$, $1 \le i \le k$, entonces $B = B_1 \cup \ldots \cup B_k$ es base de B y se verifica:

$$[T]_B = \begin{pmatrix} [T_1]_{B_1} & 0 & 0 & \ldots & 0 \\ 0 & [T_2]_{B_2} & 0 & \ldots & 0 \\ 0 & 0 & \ddots & \ldots & 0 \\ \vdots & \vdots & 0 & \ddots & 0 \\ 0 & 0 & 0 & \ldots & [T_k]_{B_k} \end{pmatrix} \in K^{n \times n} .$$

donde la representación anterior es por *diagonal por bloques*, de forma que los bloques $[T_i]_{B_i} \in K^{d_i \times d_i}$, $d_i = \dim_K \mathcal{W}_i$, $1 \le i \le k$. Recordemos que $\sum_{i=1}^k d_i = n$ en este caso (porqué?).

$\triangle$

La Observación 6.32 indica la importancia de poder contar con descomposiciones en sumas directas formadas por subespacios que son invariantes para un operador T. En el próximo resultado caracterizamos tales descomposiciones.

Teorema 6.33. *Sea $\mathcal{V}$ un K-ev, $\dim_K \mathcal{V} = n \in \mathbb{N}$ y sea $T \in L(\mathcal{V})$. Supongamos que tenemos una descomposición en suma directa*

$$\mathcal{V} = \mathcal{W}_1 \dot{+} \ldots \dot{+} \mathcal{W}_k$$

y sea $\{E_1, \ldots, E_k\}$ el sistema de proyecciones asociado a la descomposición. Son equivalentes:

1. $E_i\, T = T\, E_i$, $1 \le i \le k$.
2. $\mathcal{W}_i$ *es T-invariante,* $1 \le i \le k$.

Demostración. 1. $\implies$ 2. Supongamos que se verifican las relaciones de conmutación: $E_i\, T = T\, E_i$, $1 \le i \le k$. Sea $1 \le i \le k$: si $w_i \in \mathcal{W}_i = \mathrm{Im}(E_i)$ entonces

$$T\, w_i = T(E_i\, w_i) = (T\, E_i)\, w_i = (E_i\, T)\, w_i = E_i\, (T\, w_i) \in \mathrm{Im}(E_i) = \mathcal{W}_i \, ,$$

donde usamos que $E_i\, w_i = w_i$, porque todo proyector fija los vectores en su imágen y la hipótesis de conmutación. Así, si $w_i \in \mathcal{W}_i$ entonces $Tw_i \in \mathcal{W}_i$, y $\mathcal{W}_i$ es T-invariante, $1 \le i \le k$.

La recíproca es ejercicio. $\square$

7 Polinomio minimal y subespacios invariantes

En las próximas notas (para el tp6) veremos cómo los polinomios nos permiten obtener subespacios invariantes. Pero para poder hacer uso de los polinomios de forma adecuada, debemos desarrollar antes algunos aspectos de su teoría básica.

7.1 Ideales de polinomios

Comenzamos con el siguiente concepto y resultado, que serán muy importantes en el curso.

Definición 7.1. *Sea K un cuerpo y sea $\mathcal{I} \subset K[x]$ un subconjunto del anillo (comutativo, con unidad) de polinomios. Decimos que $\mathcal{I}$ es un ideal de $K[x]$ si:*

1. $0 \in \mathcal{I}$;
2. *Si $p(x)$, $q(x) \in \mathcal{I}$ entonces $p(x) + q(x) \in \mathcal{I}$;*

3. Si $p(x) \in \mathcal{I}$ y $r(x) \in K[x]$ ($r(x)$ es cualquier polinomio) entonces $p(x) \cdot r(x) \in \mathcal{I}$.

$\square$

Ejemplo 7.2 (Ejemplo fundamental de ideal)**.** Sea $\mathcal{V}$ un K-ev, $\dim_K \mathcal{V} = n$, y sea $T \in L(\mathcal{V})$. Consideramos $K[x]$ el anillo de polinomios sobre K. En este caso, dado

$$p(x) = a_0 + a_1\, x + a_2\, x^2 + \ldots + a_m\, x^m \in K[x]$$

podemos evaluar $p(x)$ en T, mediante la expresión

$$p(T) = a_0\, I + a_1\, T + a_2\, T^2 + \ldots + a_m\, T^m \in L(\mathcal{V})$$

donde las potencias de T se definen inductivamente: $T^0 = I$, $T^1 = T$ y en general, si $j \geq 1$ definimos $T^j = T^{j-1}\, T$. Notemos que en la evaluación, consideramos $a_0 = a_0\, x^0 \in K[x]$.

Usando el hecho de que las potencias satisfacen

$$T^k\, T^j = T^{j+k} = T^j\, T^k \qquad (\text{producto de } k+j \text{ factores iguales a } T)\,.$$

(que es una consecuencia inmediata de la asociatividad del producto de operadores) y las propiedades usuales de la suma de operadores se verifica que: Si $p(x)$, $q(x) \in K[x]$ entonces podemos sumar y multiplicar $(p+q)(x)$, $(p \cdot q)(x) \in K[x]$ y valen las identidades

1. $(p+q)(T) = p(T) + q(T) \in L(\mathcal{V})$;
2. $(p \cdot q)(T) = p(T) \cdot q(T) \in L(\mathcal{V})$.

Notemos que en particular, los operadores que son polinomio evaluados en T conmutan entre sí:

$$p(T) \cdot q(T) = (p \cdot q)(T) = (q \cdot p)(T) = q(T) \cdot p(T) \in L(\mathcal{V})\,,$$

porque $(p \cdot q)(x) = (q \cdot p)(x)$, ya que $K[x]$ es anillo conmutativo. Las propiedades 1. y 2. valen siempre que evaluemos todos los polinomios en cuestión en el *mismo* operador T.

Ahora podemos definir

$$\mathcal{I}_T = \{p(x) \in K[x] \ : \ p(T) = 0 \in L(\mathcal{V})\}\,.$$

Los elementos $p(x) \in \mathcal{I}_T$ son los *polinomios que anulan al operador* T.

$\mathcal{I}_T \subset K[x]$ es ideal: en efecto,

1. Si $0(x) \in K[x]$ es el polinimio nulo, entonces $0(T) = 0 \in L(\mathcal{V})$ es la transformación nula: entonces $0(x) \in \mathcal{I}_T$.
2. Si $p(x)$, $q(x) \in \mathcal{I}_T$ entonces $p(T) = q(T) = 0 \in L(\mathcal{V})$. Entonces, usando las propiedad 1. de la evaluación de polinomios: $(p+q)(T) = p(T) + q(T) = 0 + 0 = 0 \in L(\mathcal{V})$. Así, $(p+q)(x) \in \mathcal{I}_T$.
3. Si $p(x) \in \mathcal{I}_T$ y $r(x) \in K[x]$ entonces, usando las propiedad 2. de la evaluación de polinomios: $(p \cdot r)(T) = p(T) \cdot r(T) = 0 \cdot r(T) = 0 \in L(\mathcal{V})$. Así, $(p \cdot r)(x) \in \mathcal{I}_T$

De esta forma, hemos verificado que $\mathcal{I}_T \subset K[x]$ es un ideal.

Además, se verifica que $\mathcal{I}_T \neq \{0\}$, es decir, existe al menos un polinomio $p(x) \in \mathcal{I}_T$ tal que $p(x) \neq 0$: en efecto, notemos que

$$T^0 (= I),\, T,\, T^2,\, \ldots,\, T^{n^2} \in L(\mathcal{V})$$

son $n^2 + 1$ vectores en $L(\mathcal{V})$. Como $\dim_K L(\mathcal{V}) = n^2$ (recordemos $L(\mathcal{V}) = L(\mathcal{V}, \mathcal{V})$) entonces estos vectores forman un conjunto l.d. (de otra forma, contradecimos que $\dim_K L(\mathcal{V}) = n^2$, porqué?) Así, existen coeficientes *no todos nulos* $a_0, a_1, \ldots, a_{n^2} \in K$ tales que

$$a_0 I + a_1\, T + a_2\, T^2 + \ldots + a_{n^2} T^{n^2} = 0 \in L(\mathcal{V})$$

Entonces, si definimos $p(x) = a_0 + a_1\, x + a_2\, x^2 + \ldots + a_{n^2}\, x^{n^2} \in K[x]$, resulta que $p(x) \neq 0$ (porque no todos los coeficientes de $p(x)$ son nulos) y $p(T) = 0 \in L(\mathcal{V})$, por como construimos $p(x)$. Así, $p(x) \in \mathcal{I}_T$, y $p(x) \neq 0$: entonces $\mathcal{I}_T \neq \{0\}$. $\triangle$

El siguiente resultado juega un papel fundamental en la defnción del polinomio minimal de un operador. Para poder probar este resultado vamos a hacer uso del algoritmo de la división para polinomios en $K[x]$ tal cual lo han visto en Álgebra I.

Teorema 7.3. *Sea K un cuerpo y sea $\mathcal{I} \subset K[x]$ un ideal, tal que $\mathcal{I} \neq \{0\}$. Entonces existe un único polinomio mónico $d(x) \in \mathcal{I}$ tal que dado $p(x) \in K[x]$ se verifica:*

$$p(x) \in \mathcal{I} \quad \textit{si solo si} \quad d(x)|p(x) \tag{37}$$

donde $d(x)|p(x)$ indica que $d(x)$ divide a $p(x)$ en $K[x]$. En este caso $d(x)$ es llamado el generador de $\mathcal{I}$ y notamos $\mathcal{I} =< d(x) >$.

Demostración. Consideremos el conjunto

$$\mathcal{M}_{\mathcal{I}} = \{\mathrm{gr}(q(x)) : \ q(x) \in \mathcal{I} \text{ no nulo}\} \subseteq \mathbb{N}\,.$$

Notemos que $\mathcal{M}_{\mathcal{I}} \neq \emptyset$, pues $\mathcal{I} \neq \{0\}$. Por el principio de la buena ordenación en $\mathbb{N}$, podemos tomar $c(x) \in \mathcal{I}$ de grado mínimo entre los polinomios no nulos de $\mathcal{I}$ ($c(x) \in \mathcal{I}$ es un polinomio cuyo grado es el primer elemento de $\mathcal{M}_{\mathcal{I}}$). Así, $c(x) \in \mathcal{M}_{\mathcal{I}}$ y si $r(x) \in \mathcal{M}_{\mathcal{I}}$ entonces $\mathrm{gr}(c(x)) \leq \mathrm{gr}(r(x))$. Como $c(x) \neq 0$ entonces $c(x) = a_0 + a_1\,x + \ldots + a_m\,x^m$ con $a_m \neq 0$: Como $a_m^{-1} \in K \subset K[x]$ es polinomio de grado 0, entonces por el ítem 3 de la Definición 7.1, vemos que

$$d(x) := c(x)\,a_m^{-1} \in \mathcal{I} \quad , \quad \mathrm{gr}(c(x)) = \mathrm{gr}(d(x)) \quad , \quad d(x) \ \text{ es mónico}\,.$$

Así, $d(x) \in \mathcal{I}$ es polinomio mónico de grado mínimo (entre los grados de los polinomios no nulos de $\mathcal{I}$).

Sea $p(x) \in \mathcal{I}$: podemos aplicar el algoritmo de la división en $K[x]$ y concluir que existen $q(x)$, $r(x) \in K[x]$ tales que

$$p(x) = q(x)\,d(x) + r(x) \quad \text{ con } \quad r(x) = 0 \ \text{ ó } \ \mathrm{gr}(r(x)) < \mathrm{gr}(d(x))\,.$$

Si $r(x) \neq 0$ entonces $gr(r(x)) < \mathrm{gr}(d(x))$: además, como $d(x) \in \mathcal{I}$ entonces $q(x)\,d(x) \in \mathcal{I}$ y $(-1)d(x)\,q(x) \in \mathcal{I}$ (por el ítem 3 de la Definición 7.1). Pero entonces

$$r(x) = p(x) - d(x)\,q(x) = p(x) + (-1)\,d(x)\,q(x) \in \mathcal{I}$$

por el ítem 2. de la Definición 7.1, pues $p(x) \in \mathcal{I}$ y $(-1)d(x)\,q(x) \in \mathcal{I}$. Pero entonces $r(x) \in \mathcal{I}$ y $gr(r(x)) < \mathrm{gr}(d(x))$ contradicen la construcción de $d(x)$ (porque $d(x)$ tiene grado mínimo).

El argumento anterior muestra que se debe tener que $r(x) = 0$ de forma que $p(x) = q(x)\,d(x)$, lo que muestra que $d(x)|p(x)$, siempre que $p(x) \in \mathcal{I}$.

Recíprocamente, si $d(x)|p(x)$ entonces existe $q(x) \in K[x]$ tal que $p(x) = q(x)\,d(x) \in \mathcal{I}$ (por el ítem 3 de la Definición 7.1).

Los argumentos anteriores muestran que $d(x)$ construido como antes satisface la Eq. (37).

Finalmente, verificamos que existe un sólo polinomio que satisface las condiciones del Teorema: supongamos que $e(x) \in \mathcal{I}$ es otro polinomio mónico que verifica:

$$p(x) \in \mathcal{I} \quad \text{si solo si} \quad e(x)|p(x)\,.$$

Como $d(x) \in \mathcal{I}$ entonces $e(x)|d(x)$, por la condición anterior (hipótesis sobre $e(x)$). Así, existe $q_1(x) \in K[x]$ tal que $d(x) = q_1(x)\,e(x)$. Además, como $e(x) \in \mathcal{I}$ entonces, por la propiedad verificada para $d(x)$, $d(x)|e(x)$. Así, existe $q_2(x) \in K[x]$ tal que $e(x) = q_2(x)\,d(x)$. Así

$$d(x) = q_1(x)\,e(x) = q_1(x)\,(q_2(x) \cdot d(x)) = (q_1(x) \cdot q_2(x))\,d(x)\,.$$

Usando que, en general, gr$(r(x)\,s(x))$ =gr$(r(x))$+gr$(s(x))$, si $r(x)$ y $s(x)$ son no nulos, vemos que

$$\text{gr}(d(x)) = \text{gr}(q_1(x)\cdot q_2(x)) + \text{gr}(d(x)) \implies \text{gr}(q_1(x)\cdot q_2(x)) = \text{gr}(q_1(x)) + \text{gr}(q_2(x)) = 0\,.$$

Así, concluimos que gr$(q_1(x))$ = gr$(q_2(x)) = 0$, es decir, $q_1(x) = a \in K$ y $q_2(x) = b \in K$.

Entonces la identidad $d(x) = q_1(x)\,e(x)$ se transforma en $d(x) = a\,e(x)$. En particular, gr$(d(x))$ =gr$(e(x))$: así,

$$d(x) = d_0 + d_1\,x + \ldots + d_{m-1}\,x^{m-1} + x^m \quad , \quad e(x) = e_0 + e_1\,x + \ldots + e_{m-1}\,x^{m-1} + x^m$$

$$d_0 + d_1\,x + \ldots + d_{m-1}\,x^{m-1} + x^m = d(x) = a\,e(x) = a\,e_0 + a\,e_1\,x + \ldots + a\,e_{m-1}\,x^{m-1} + a\,x^m$$

Como los polinomios son iguales, entonces los coeficientes deben ser iguales; si miramos los coeficientes que acompañan a x^m deducimos que $1 = a$. Así $d(x) = a\,e(x) = 1\,e(x) = e(x)$ y $d(x) = e(x)$. □

7.2 El polinomio minimal

Ahora podemos definir otro concepto fundamental asociado a todo operador que actúe en un espacio vectorial de dimensión finita.

Definición 7.4. Sea $\mathcal{V}$ un K-ev, $\dim_{\mathcal{V}} = n$ y sea $T \in L(\mathcal{V})$. Consideremos el ideal de polinomios

$$\mathcal{I}_T = \{p(x) \in K[x] \ : \ p(T) = 0 \in L(\mathcal{V})\} \subset K[x]$$

definido en el Ejemplo 7.2. Definimos el *polinomio minimal* de T, notado $m_T(x) \in K[x]$, como el generador de $\mathcal{I}_T$ (ver Teorema 7.3). □

Obs 7.5. Con las notaciones de la definición anterior, el Teorema 7.3 garantiza que $m_T(x)$ está bien definido (existe) y es el único polinomio mónico tal que:

1. $m_T(T) = 0$ (m_T anula a T), pues $m_T(x) \in \mathcal{I}_T$.

2. Si $p(x)$ es tal que $p(T) = 0$ (es decir, $p(x) \in \mathcal{I}_T$) entonces $m_T(x)|p(x)$: es decir, existe

   $$q(x) \in K[x] \ \text{ tal que } \ p(x) = m_T(x)\,q(x)\,.$$

   En particular, si $p(x) \neq 0$ entonces gr$p(x)$ =gr$(m_T(x))$+gr$(q(x)) \geq$ gr$(m_T(x))$. Lo anterior indica que $m_T(x)$ es el polinomio (no nulo) de grado más chico que anula a T (de ahí su nombre).

△

Obs 7.6. Sea K un cuerpo y sea $A \in K^{n\times n}$ matriz **cuadrada**. Si $p(x) \in K[x]$ está dado por $p(x) = a_0 + a_1\,x + \ldots + a_m\,x^m$ definimos

$$p(A) = a_0\,I + a_1\,A + \ldots + a_m\,A^m \in K^{n\times n}$$

Podemos definir el ideal (verificar, ejercicio)

$$\mathcal{I}_A = \{p(x) \in K[x] \ : \ p(A) = 0\} \neq \{0\}$$

y el *polinomio minimal* de A como el generador del ideal $\mathcal{I}_A$.

Sea $\mathcal{V}$ es K-ev, $\dim_K \mathcal{V} = n$, $T \in L(\mathcal{V})$ y B una base de $\mathcal{V}$. Consideremos $[T]_B = A \in K^{n\times n}$: notemos que como la transformación $L(\mathcal{V}) \ni S \mapsto [S]_B$ es un isomorfimos de K-álgebras con unidad (respeta sumas, acción escalar, productos, unidad) entonces

$$\begin{aligned}
[p(T)]_B &= [a_0\, I + a_1\, T + \ldots + a_m\, T^m]_B = a_0\, [I]_B + a_1\, [T]_B + \ldots + a_m\, [T]_B^m \\
&= a_0\, I + a_1\, A + \ldots + a_m\, A^m = p(A) \in K^{n\times n}
\end{aligned}$$

donde hemos usado que $[T^2]_B = [T\, T]_B = [T]_B\, [T]_B = [T]_B^2$ y en general que $[T^j]_B = [T]_B^j$, $j \geq 0$. Así, como tomar matriz con respecto a una base es isomorfismo:

$$p(T) = 0 \in L(\mathcal{V}) \;\Leftrightarrow [p(T)]_B = 0 \;\Leftrightarrow\; p(A) = 0 \in K^{n\times n}\,.$$

Así,

$$\mathcal{I}_T = \{p(x) \in K[x] \;:\; p(T) = 0\} = \{p(x) \in K[x] \;:\; p(A) = 0\} = \mathcal{I}_A$$

de forma que $m_T(x) = m_A(x)$, para $A = [T]_B$ con respecto a cualquier base B de $\mathcal{V}$. △

Los siguientes resultados son una herramienta importante para poder determinar el polinomio minimal de un operador lineal T. Comenzamos con un lema técnico.

Lema 7.7. *Sea $\mathcal{V}$ es K-ev,* $\dim_K \mathcal{V} = n$*, $T \in L(\mathcal{V})$. Supongamos que $v \in \mathcal{V}$ y $\lambda \in K$ satisfacen $T\, v = \lambda\, v$. Si $p(x) \in K[x]$ entonces*

$$p(T) \in L(\mathcal{V}) \;\; \textit{verifica} \;\; p(T)\, v = p(\lambda)\, v\,,$$

donde $p(\lambda) \in K$ es la evaluación de $p(x)$ en $\lambda \in K$.

Demostración. Verificamos por inducción: $T^j\, v = \lambda^j\, v$ para $j \geq 0$: aquí $\lambda^j \in K$denota la j-ésima potencia de λ, es decir $\lambda^0 = 1$, $\lambda^1 = \lambda$, y en general $\lambda^j = \lambda^{j-1}\, \lambda$, para $j \geq 1$.

Si $j = 0$: $T^0 = I$, entonces $T^0\, v = I\, v = v = \lambda^0\, v$, pues $\lambda^0 = 1$.

Hipótesis inductiva: suponemos que vale para $j - 1 \geq 0$. Sea $j \geq 1$: entonces

$$T^j\, v = (T^{j-1}\, T)\, v = T^{j-1}(T\, v) = T^{j-1}\, \lambda\, v = \lambda\, T^{j-1}\, v = \lambda\, \lambda^{j-1}\, v = \lambda^j\, v\,,$$

donde usamos la propiedad $T^j = T^{j-1}\, T$, la definición del producto (que es la composición de transformaciones), la hipótesis, el hecho de que T^{j-1} es lineal, la hipótesis inductiva y las propiedades de las potencias en K.

Si $p(x) = a_0 + a_1, x + \ldots + a_m, x^m$ entonces

$$\begin{aligned}
p(T)\, v &= (a_0\, I + a_1\, T + \ldots + a_m\, T^m)(v) = a_0\, v + a_1\, T\, v + \ldots + a_m\, T^m\, v \\
&= a_0\, v + a_1\, \lambda\, v + \ldots + a_m\, \lambda^m\, v = (a_0\, + a_1\, \lambda + \ldots + a_m\, \lambda^m)\, v = p(\lambda)\, v\,.
\end{aligned}$$

□

Teorema 7.8. *Sea $\mathcal{V}$ es K-ev,* $\dim_K \mathcal{V} = n$*, $T \in L(\mathcal{V})$. Sean $p_T(x), m_T(x) \in K[x]$ el polinomio característico y minimal de T. Entonces $p_T(x)$ y $m_T(x)$ tienen las mismas raíces (pero, en general, no las mismas multiplicidades): es decir, dado $\lambda \in K$,*

$$m_T(\lambda) = 0 \;\Leftrightarrow\; p_T(\lambda) = 0\,.$$

Demostración. Sea $\lambda \in K$ tal que $m_T(\lambda) = 0$. Entonces $m_T(x) = (x - \lambda)\, q(x)$: entonces $\text{gr}(q(x)) < \text{gr}(m_T(x))$ lo que muestra que $q(T) \neq 0 \in L(\mathcal{V})$ (ver Observación 7.5). Así, existe $v \in \mathcal{V}$ tal que $w = q(T)\, v \neq 0$: entonces

$$0 = m_T(T) = ((x - \lambda) \cdot q(x))(T) = (T - \lambda\, I)\, q(T)$$

donde usamos la propiedad de la evaluación de polinomios en un mismo operador T, $(p(x) \cdot r(x))(T) = p(T) \cdot r(T)$ y que $(x - \lambda)(T) = (T - \lambda\, I)$. En particular

$$0 = 0\, v = (\, (T - \lambda\, I)\, q(T)\,)(v) = (T - \lambda\, I)(q(T)\, v) = (T - \lambda\, I)(w) \implies T\, w = \lambda\, w$$

con $w = q(T)\, v \neq 0$. Entonces λ es autovalor (con autovector w) y $p_T(\lambda) = 0$.

Recíprocamente, si $p_T(\lambda) = 0$ entonces λ es autovalor y existe $v \in \mathcal{V}$, $v \neq 0$ tal que $T\, v = \lambda\, v$. Por el Lema 7.7 (tomando $p(x) = m_T(x)$) tenemos que $m_T(\lambda) \in K$ y

$$m_T(\lambda)\, v = m_T(T)\, v = 0\, v = 0\,,$$

pues $m_T(T) = 0 \in L(\mathcal{V})$. Como $v \neq 0$ y $m_T(\lambda)\, v = 0$ deducimos que el escalar $m_T(\lambda) = 0 \in K$ y λ es raíz de $m_T(x)$. □

Teorema 7.9 (de Cayley-Hamilton)**.** *Sea $\mathcal{V}$ es K-ev, $\dim_K \mathcal{V} = n$, $T \in L(\mathcal{V})$. Sean $p_T(x)$, $m_T(x) \in K[x]$ el polinomio característico y minimal de T. Entonces $p_T(T) = 0$ ($p_T(x)$ anula a T). En particular, $m_T(x) | p_T(x)$ (el minimal divide al característico).* □

Por ahora no vamos a desarrollar una prueba del Teorema de Cayley-Hamilton. Sin embargo, vamos a aceptar este resultado como cierto y lo vamos a usar repetidamente.

Obs 7.10. Como consecuencia de los Teoremas 7.8 y 7.9 concluimos lo siguiente: si T es tal que $p_T(x) = (x - \lambda_1)^{r_1} \cdots (x - \lambda_k)^{r_k} \in K[x]$ entonces

$$m_T(x) = (x - \lambda_1)^{s_1} \cdots (x - \lambda_k)^{s_k} \quad \text{con} \quad 1 \leq s_j \leq r_j \ , \ 1 \leq j \leq k\,.$$

La afirmación anterior se deduce del hecho de que $m_T(x)$ tiene las mismas raíces $\lambda_1, \ldots, \lambda_k$ que $p_T(x)$ y además, $m_T(x) | p_T(x)$ de forma que la multiplicidad s_j de la raíz λ_j en $m_T(x)$ es menor ó igual que la multiplicidad r_j de la raíz λ_j en $p_T(x)$, $1 \leq j \leq k$. Este último hecho se deduce del teorema de la factorización prima (única salvo orden y asociados) de los polinomios y el hecho de que cada factor $(x - \lambda)$ es primo en $K[x]$. △

Ejemplo 7.11. Sea $S \in L(\mathbb{R}^3)$ el operador dado por

$$s(x, y, z) = (y, x, z) \quad \text{para} \quad (x, y, z) \in \mathbb{R}^3\,.$$

Si consideramos $B_c = \{e_1,\, e_2,\, e_3\}$ la base canónica de $\mathbb{R}^3$ entonces

$$[S]_{B_c} = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \implies x\, I - [S]_{B_c} = \begin{pmatrix} x & -1 & 0 \\ -1 & x & 0 \\ 0 & 0 & x-1 \end{pmatrix}.$$

En particular, desarrollando el determinante por la tercer columna:

$$p_S(x) = \det(x\, I - [S]_{B_c}) = (x - 1)\, (x^2 - 1) = (x - 1)^2\, (x + 1)\,,$$

donde factorizamos la diferencia de cuadrados $(x^2 - 1) = (x - 1)\, (x + 1)$.

Así, los autovalores de S son $\{1, -1\}$. Más aún, el polinomio minimal debe verificar (ver la Observación 7.10):

$$m_T(x) = (x - 1)^{s_1}(x + 1) \quad \text{con} \quad 1 \leq s_1 \leq 2\,.$$

Notemos que la potencia s_2 de la raíz $\lambda = 1$ en $m_T(x)$ debe ser $s_2 = 1$ (como aparece en la ecuación de arriba), pues $1 \leq s_2 \leq r_2 = 1$ (ver la Observación 7.10).

Así, tenemos dos posibilidades

$$m_T(x) = (x-1)\,(x+1) \quad \text{ó} \quad m_T(x) = (x-1)^2\,(x+1) = p_T(x)\,.$$

Comenzamos evaluando el polinomio $(x-1)\,(x+1)$ en S, es decir considerando $(S-I)\,(S+I) \in L(\mathbb{R}^3)$ y verificando si $(S-I)\,(S+I) = 0$. Para eso, consideramos la matriz

$$[(S-I)\,(S+I)]_{B_c} = ([S]_{B_c} - I)\,([S]_{B_c} + I) = \begin{pmatrix} -1 & 1 & 0 \\ 1 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \tag{38}$$

donde la última identidad resulta de hacer el producto de las matrices (ojo que al final aparece un número de ecuación).

Entonces, el polinomio minimal de T es $m_T(x) = (x-1)(x+1) = x^2 - 1$. En efecto, hemos visto que este polinomio *mónico* satisface:

1. Comparte las raíces con $p_T(x)$;
2. Divide a $p_T(x)$;
3. Anula a T (por la Eq. (38)).

En particular, $m_T(x)|(x-1)(x+1)$.

Además, el polinomio minimal $m_T(x)$ debe satisfacer estas propiedades (por los Teoremas 7.8 y 7.9); pero hemos visto que $(x-1)(x+1) = x^2 - 1$ es el polinomio *de grado más chico* que satisface estas tres propiedades (el otro candidato era $(x-1)^2\,(x+1)$ que tiene grado más grande!!). Así,

$$m_T(x)|(x-1)(x+1) \ \text{ y } \ \mathrm{gr}((x-1)(x+1)) \leq \ \mathrm{gr}(m_T(x)) \implies m_T(x) = (x-1)(x+1)$$

porque $(x-1)(x+1) = q(x)\,m_T(x)$ y $\mathrm{gr}((x-1)(x+1)) \leq \ \mathrm{gr}(m_T(x))$ implican que:

$$\mathrm{gr}((x-1)(x+1)) = \ \mathrm{gr}(m_T(x)) \quad \text{de forma que} \quad q(x) = 1 \quad \text{y} \quad m_T(x) = (x-1)(x+1)\,.$$

$\triangle$

Ejemplo 7.12. Sea $A \in \mathbb{C}^{4\times 4}$ dada por

$$A = \begin{pmatrix} 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \end{pmatrix}.$$

Nos piden calcular el minimal de la matriz A (ver Observación 7.6): para eso, nos sugieren verificar la identidad:

$$A^3 = \begin{pmatrix} 0 & 4 & 0 & 4 \\ 4 & 0 & 4 & 0 \\ 0 & 4 & 0 & 4 \\ 4 & 0 & 4 & 0 \end{pmatrix}.$$

para lo que calculamos A^2 y $A^3 = A^2\,A$ (hacer las cuentas). Una vez verificada la identidad propuesta, notemos que $A^3 = 4\,A$!! Es decir $A^3 - 4A = 0$. De esta forma, el polinomio $p(x) = x^3 - 4\,x$ anula a A.

Así, $m_A(x)|x^3 - 4x$; como $p(x) = x^3 - 4x = x(x^2-4) = x\,(x-2)\,(x+2)$, por el Teorema de factorización prima en $\mathbb{R}[x]$, los factores primos de $m_A(x)$ deben ser algunos de los factores primos de $p(x)$: en este caso, se tienen las siguientes posibilidades para $m_A(x)$:

1. $m_1 = x$;

2. $m_2 = x - 2$;

3. $m_3 = x + 2$;

4. $m_4 = x(x - 2)$;

5. $m_5 = x(x + 2)$;

6. $m_6 = (x - 2)\,(x + 2)$;

7. $m_7 = x(x - 2)\,(x + 2)$.

Notemos que hemos listado los candidatos a polinomio minimal de menor a mayor grado; de esta forma, si evaluamos estos polinomios en A, respetando el orden propuesto, el primer polinomio que anula a A será el polinomio mónico de grado mínimo que anule a A: es decir, el minimal de A.

Cuentas sencillas (pero tediosas) muestran que si $1 \leq j \leq 6$ entonces $m_j(A) \neq 0 \in \mathbb{R}^{4\times 4}$. Así, el minimal de A es $m_A(x) = m_7(x) = x\,(x - 2)\,(x + 2)$.

Qué sucede con el polinomio característico de A? (ver Observación 6.22). Notemos que $p_A(x)$ es un polinomio mónico de grado 4, que comparte las raíces con $m_A(x)$. Notemos entonces que los autovalores distintos de A son $\{0, 2, -2\}$.

Como $\dim_{\mathbb{C}} \mathrm{Im}(A) = \mathrm{rg}(A) = 2$ (la imágen de $T_A \in L(\mathbb{C}^4)$ es el subespacio generado por las columnas de A: pero la tercer y cuarta columnas de A coinciden con la primer y segunda columna de A). Entonces $\dim_{\mathbb{C}} \mathrm{N}(A) = \dim_{\mathbb{C}} \mathbb{C}^4 - \dim_{\mathbb{C}} \mathrm{Im}(A) = 2$. Así la multiplicidad geométrica de $\lambda = 0$ es 2 pues $\dim_{\mathbb{C}} E_A(0) = \dim_{\mathbb{C}} \mathrm{N}(0\,I - A) = \dim_{\mathbb{C}} \mathrm{N}(A) = 4 - 2 = 2$.

Así $\dim_{\mathbb{C}} E_A(0) = 2$, $\dim_{\mathbb{C}} E_A(2) \geq 1$ y $\dim_{\mathbb{C}} E_A(-2) \geq 1$ (pues los autoespacios son subespacios no nulos, y tienen dimensión por lo menos 1) implica que A es diagonalizable!! y luego su polinomio característico es $p_A(x) = x^2\,(x - 2)\,(x + 2)$. △

Como comentario final, notemos el siguiente hecho general que hemos usado en el ejemplo anterior: $\lambda = 0 \in K$ es autovalor de T si y solo si $N(T) \neq 0$. Así, si $N(T) \neq \{0\}$ entonces $E_T(0) = \mathrm{N}(T)$ (el autoespacio del autovalor 0 es el núcleo de T) y la multiplicidad geométrica de 0 como autovalor de T es la nulidad de T, es decir $\mathrm{Nul}(T) = \dim_K \mathrm{N}(T) = \dim_K E_T(0)$.

8 Primeros teoremas de estructura de operadores

En lo que sigue, vamos a aplicar técnicas y resultados de la teoría de polinomios para encontrar descomposiciones en suma directa
$$\mathcal{V} = \mathcal{W}_1 \dot{+} \ldots \dot{+} \mathcal{W}_k$$
tales que cada $\mathcal{W}_i$ es subespacio T-invariante, $1 \leq i \leq k$, y de forma que no sólo $T_i = T|_{\mathcal{W}_i} \in L(\mathcal{W}_i)$ sea *más chico* (como habíamos destacado antes) sino que en realidad T_i sea más *sencillo ó simple*. La complejidad de T_i será medida a través del polinomio minimal de T_i, $1 \leq i \leq k$.

8.1 Preliminares sobre polinomios

En lo que sigue, vamos a utilizar resultados vistos en Álgebra I sobre existencia y unicidad de factorización prima de polinomios. Para recordar estos hechos y concluir algunas consecuencias, consideramos las siguiente observaciones:

Obs 8.1. Sea K un cuerpo y consideremos $K[x]$ el anillo (conmutativo con unidad) de polinomios con coeficientes en K.

1. Recordemos que dado $p(x) \in K[x]$ tal que $\mathrm{gr}(p(x)) \geq 1$, decimos que es *primo* si siempre que $q(x) \in K[x]$ es tal que $q(x)|p(x)$, entonces $\mathrm{gr}(q(x)) = 0$ ($q(x) \in K$ es escalar) ó $\mathrm{gr}(q(x)) = \mathrm{gr}(p(x))$.

2. Así, todo polinomio lineal $(x - \lambda) \in K[x]$ es primo. Pero no todo polinomio primo debe ser lineal! Por ejemplo: $x^2 + 1$ es primo en $\mathbb{R}[x]$.

3. Notemos que si $p_1(x)$, $p_2(x) \in K[x]$ son primos, mónicos y **distintos** entonces

$$p_1(x) \not| p_2(x) \quad \text{y} \quad p_2(x) \not| p_1(x)\,.$$

   Por ejemplo, si suponemos que $p_1(x)|p_2(x)$ entonces $p_2(x) = q(x)\,p_1(x)$ con $\mathrm{gr}(q(x)) = 0$ es decir, $a = q(x) \in K$ (porqué?). Pero como $p_1(x)$ y $p_2(x)$ son mónicos, entonces

$$p_2(x) = a\,p_1(x) \implies a = 1 \quad \text{y} \quad p_2(x) = p_1(x)\,,$$

   que contradice la hipótesis.

4. Teorema de la factorización prima en $K[x]$: dado $p(x) \in K[x]$, $\mathrm{gr}(p(x)) \geq 1$, existen polinomios primos mónicos distintos $p_1(x), \ldots, p_k(x) \in K[x]$, naturales $1 \leq r_j$ para $1 \leq j \leq k$ (únicos salvo orden) y $a \in K$, tales que

$$p(x) = a\,p_1^{r_1}(x) \cdots p_k^{r_k}(x)\,,$$

   donde a coincide con el coeficiente principal de $p(x)$.

5. En adelante, vamos a usar muchas veces la siguiente observación (de Álgebra I): con las notaciones del ítem 4. (anterior), si $q(x) \in K[x]$ es polinomio primo, mónico y tal que $q(x)|p(x)$ entonces existe $1 \leq j \leq k$ tal que $q(x) = p_j(x)$. Es decir, si $q(x)$ es primo, mónico y divide a $p(x)$ debe ser uno de los factores en al descomposición prima de $p(x)$!! por la unicidad de tal factorización.

$\triangle$

Además, vamos a necesitar las siguientes observaciones sobre ideales.

Obs 8.2. Sea K un cuerpo y consideremos $K[x]$ el anillo (conmutativo con unidad) de polinomios con coeficientes en K.

Dados $f_1(x), \ldots, f_k(x) \in K[x]$ entonces el conjunto

$$\mathcal{I} = \{\sum_{j=1}^{k} f_j(x)\,g_j(x) \ : \ g_j(x) \in K[x]\ ,\ 1 \leq j \leq k\}$$

es un ideal de $K[x]$. En efecto:

1. $0 = \sum_{j=1}^{k} f_j(x)\,0 \in \mathcal{I}$, con $g_j(x) = 0 \in K[x]$;

2. Si $p(x)$, $q(x) \in \mathcal{I}$ entonces

$$p(x) = \sum_{j=1}^{k} f_j(x)\,g_j(x) \quad \text{y} \quad q(x) = \sum_{j=1}^{k} f_j(x)\,h_j(x)$$

   para ciertos $g_j(x)$, $h_j(x) \in K[x]$, $1 \leq j \leq k$: entonces

$$p(x) + q(x) = \sum_{j=1}^{k} f_j(x)\,g_j(x) + \sum_{j=1}^{k} f_j(x)\,h_j(x) = \sum_{j=1}^{k} f_j(x)(g_j(x) + h_j(x))$$

   con $g_j(x) + h_j(x) \in K[x]$, $1 \leq j \leq k$. Entonces $p(x) + q(x) \in \mathcal{I}$, por definición.

3. Si $p(x) \in \mathcal{I}$ es como arriba y $r(x) \in K[x]$ es arbitrario

$$p(x) \cdot r(x) = (\sum_{j=1}^{k} f_j(x)\, g_j(x)) \cdot r(x) = \sum_{j=1}^{k} f_j(x)\, (g_j(x)\, r(x))$$

   con $g_j(x)\, r(x) \in K[x]$, $1 \leq j \leq k$. Entonces $p(x) \cdot r(x) \in \mathcal{I}$.

El ideal $\mathcal{I}$ también es notado $\mathcal{I} =< \{f_j(x) \,:\, 1 \leq j \leq k\} >$ y llamado el ideal generado por los polinomios $\{f_j(x) \,:\, 1 \leq j \leq k\}$.

Si $f_j(x) \neq 0$ para algún $1 \leq j \leq k$ entonces $\mathcal{I} \neq \{0\}$. En este caso, por el resultado sobre ideales de las notas para el tp5 (2da parte) existe un único polinomio mónico $d(x) \in \mathcal{I}$ tal que dado $p(x) \in K[x]$ se verifica:

$$p(x) \in \mathcal{I} \;\Leftrightarrow\; d(x)|p(x)\,.$$

Afirmamos que $d(x)$ =mcd$(f_1(x), \ldots, f_k(x))$ (el máximo común divisor de los polinomios $\{f_j :\ 1 \leq j \leq k\}$). En efecto:

1. Si $1 \leq i \leq k$ fijo, sea $g_j(x) = \delta_{ij} \in K[x]$ ($g_j(x) = 0$ si $j \neq i$, $g_i(x) = 1$). Entonces

$$\sum_{j=1} f_j(x)\, g_j(x) = f_i(x) \in \mathcal{I} \implies d(x)|f_i(x)\,.$$

2. Además, como $d(x) \in \mathcal{I}$, entonces existen polinomios $h_1(x), \ldots, h_k(x) \in K[x]$ tales que

$$d(x) = \sum_{j=1}^{k} f_j(x)\, h_j(x)\,,$$

   por definición de $\mathcal{I}$.

Así, $d(x)$ es divisor de $f_i(x)$, $1 \leq i \leq k$, y si $r(x)$ es divisor de $f_i(x)$, $1 \leq i \leq k$, entonces

$$f_j(x) = r(x)\, q_j(x) \implies d(x) = \sum_{j=1}^{k} (r(x)\, q_j(x))\, h_j(x) = r(x) \sum_{j=1}^{k} q_j(x)\, h_j(x)$$

lo que muestra que $r(x)|d(x)$. $\triangle$

Vamos a aplicar las observaciones anteriores en la siguiente construcción

Obs 8.3. Sea K un cuerpo y consideremos $K[x]$ el anillo (conmutativo con unidad) de polinomios con coeficientes en K. Dado $p(x) \in K[x]$ polinomio mónico, gr$(p(x)) \geq 1$, sean $p_1(x), \ldots, p_k(x) \in K[x]$ polinomios primos mónicos distintos y $1 \leq r_j$, para $1 \leq j \leq k$ (únicos salvo orden) tales que

$$p(x) = p_1^{r_1}(x) \cdots p_k^{r_k}(x)\,,$$

(donde 1 es el coeficiente principal de $p(x)$). En lo que sigue, suponemos que k (la cantidad de factores) verifica $k \geq 2$.

Definimos, para $1 \leq j \leq k$:

$$f_j = \prod_{1 \leq i \neq j \leq k} p_i^{r_i}(x) \in K[x]$$

Ejemplo: si $p(x) = p_1^{r_1}(x)\, p_2^{r_2}(x)\, p_3^{r_3}(x)$ entonces

$$f_1(x) = p_2^{r_2}(x)\, p_3^{r_3}(x)\ ,\ \ f_2(x) = p_1^{r_1}(x)\, p_3^{r_3}(x)\ ,\ \ f_3(x) = p_1^{r_1}(x)\, p_2^{r_2}(x)$$

es decir, $f_j(x)$ tiene todos los factores primos de $p(x)$ con sus multiplicidades, salvo el j-ésimo factor primo (que está excluído de los índices de la productoria, indicando que el índice $i \neq j$).

Notemos que la construcción de $f_j(x)$ nos indica la descomposición prima de $f_j(x)$!!

Utilizando el argumento de la Observación 8.2, si $\mathcal{I}$ denota el ideal generado por $\{f_j \ : \ 1 \leq j \leq k\}$ entonces existe $d(x) \in \mathcal{I}$ polinomio mónico tal que

1. $d(x)|f_j(x)$, $1 \leq j \leq k$;

2. Existen polinomios $h_1(x), \ldots, h_k(x) \in K[x]$ tales que $d(x) = \sum_{j=1}^{k} f_j(x)\, h_j(x)$.

Afirmamos que en este caso $d(x) = 1$: en efecto, de otra forma $\mathrm{gr}(d(x)) \geq 1$; en este caso, $d(x)$ admite al menos un factor primo mónico: digamos que existe un polinomio primo mónico $q(x)$ tal que $q(x)|d(x)$: entonces, $q(x)|d(x)$ y $d(x)|f_1(x)$ entonces $q(x)|f_1(x)$.

Como $q(x)$ es polinomio primo mónico, entonces $q(x)$ coincide con alguno de los factores que aparecen en la definción de $f_1(x)$: es decir, existe $2 \leq \ell \leq k$ tal que $q(x) = p_\ell(x)$ (recordemos que $\ell \neq 1$ en la definición de $f_1(x)$), por unicidad de la factorización prima en $K[x]$ (ver item 5 de la Observación 8.1).

Ahora, usamos este índice $2 \leq \ell \leq k$ para hacer la siguiente observación (attenti!): $q(x)|d(x)$ y $d(x)|f_\ell(x)$ entonces $q(x)|f_\ell(x)$!! Pero esto último no es posible, porque $q(x) = p_\ell(x)$, y por construcción, $f_\ell(x)$ no tiene a $p_\ell(x)$ como factor primo en su descomposición prima (ver item 5 de la Observación 8.1).

La contradicción anterior surge de suponer que $\mathrm{gr}(d(x)) \geq 1$: entonces $\mathrm{gr}(d(x)) = 0$ y $d(x) = 1$. Volviendo al ítem 2. de arriba, recordemos que existían $h_1(x), \ldots, h_k(x) \in K[x]$ tales que

$$1 = \sum_{j=1}^{k} f_j(x)\, h_j(x)\,. \tag{39}$$

Para concluir esta observación, consideramos el siguiente hecho: si $1 \leq j \neq \ell \leq k$ entonces

$$p(x) \,|\, f_j(x)\, f_\ell(x)\,. \tag{40}$$

En efecto, por definición, $f_j(x)$ tiene a los factores $p_i(x)^{r_i}$ en su factorización con $1 \leq i \neq j \leq k$: así, para verificar que $f_j(x)\, f_\ell(x)$ es divisible por $p(x)$, basta ver que el factor $p_j(x)^{r_j}$ aparece en la factorización de $f_j(x)\, f_h(x)$; esto es cierto porque el factor $p_j(x)^{r_j}$ aparece en la factorización de $f_\ell(x)$!! (pues por definición, $f_\ell(x)$ tiene a los factores $p_i(x)^{r_i}$ en su factorización con $1 \leq i \neq \ell \leq k$ y hemos supuesto que $j \neq \ell$). $\triangle$

8.2 Polinomios de Lagrange

Vamos a considerar una construcción relacionada en algún sentido con la construcción realizada en la Observación 8.3.

Definición 8.4. *Sea K un cuerpo y sean $c_1, \ldots, c_k \in K$ elementos distintos de K, con $k \geq 2$. Para cada $1 \leq j \leq k$ construimos el polinomio*

$$q_j(x) = a_j^{-1} \prod_{1 \leq i \neq j \leq k} (x - c_i) \quad \text{donde} \quad a_j = \prod_{1 \leq i \neq j \leq k} (c_j - c_i) \neq 0\,.$$

Los polinomios $\{q_1(x), \ldots, q_k(x)\}$ son llamados los polinomios interpoladores de Lagrange, asociados a $c_1, \ldots, c_k \in K$. $\square$

Obs 8.5. Consideremos las notaciones de la definición anterior. Comencemos con algunas aclaraciones:

1. En la fórmula
$$q_j(x) = a_j^{-1} \prod_{1\leq i\neq j\leq k} (x - c_i) \in K[x]$$
se tiene el producto de $k-1$ factores; en efecto, el índice (variable) i varia entre $1 \leq i \leq k$, pero además, se exige que $i \neq j$ (es decir, el valor j queda excluído del conjunto en donde varía i). En particular $\mathrm{gr}(q_j(x)) = k-1$, $1 \leq j \leq k$. Además, el coeficiente
$$a_j = \prod_{1\leq i\neq j\leq k} (c_j - c_i) \neq 0$$
es no nulo, pues es el producto de $(k-1)$ factores no nulos, pues $(c_j - c_i) \neq 0$ si $i \neq j$, dado que $c_1, \ldots, c_k$ son **distintos** (así se tomaron). En particular, se puede considerar el inverso multiplicativo (en K) de a_j, es decir $a_j^{-1} \in K$.

2. Relaciones de dualidad: sea $1 \leq j \leq k$: consideramos las evaluaciones
$$q_j(c_j) = a_j^{-1} \prod_{1\leq i\neq j\leq k} (c_j - c_i) = a_j^{-1}\, a_j = 1\,,$$
donde hemos reemplazado a x por c_j y hemos usado la definición de a_j. Por otro lado, si $1 \leq h \leq k$ es tal que $h \neq j$ entonces:
$$q_j(c_h) = a_j^{-1} \prod_{1\leq i\neq j\leq k} (c_h - c_i) = 0\,.$$
En efecto, cuando reemplazamos x por c_h, aparece la productoria de los factores $(c_h - c_i)$, con $1 \leq i \neq j \leq k$. Como $1 \leq h \leq k$ y $h \neq j$ entonces el indice i debe tomar el valor $i = h$ en algún factor! En ese caso, aparece el factor $(c_h - c_i) = (c_h - c_h) = 0$ (si $i = h$) de forma que todo el producto se anula.

   A modo de resumen: con la notación de la delta de Kronecker,
$$q_j(c_h) = \delta_{j\,h} \quad \text{para cualesquiera} \quad 1 \leq j,\, h \leq k\,. \tag{41}$$

3. Sean $b_1, \ldots, b_k \in K$ escalares arbitrarios: entonces los polinomios de Lagrange nos permite construir de forma sencilla un polinomio $p(x) \in K[x]$ que verifica: $\mathrm{gr}(p(x)) \leq k-1$ y $p(x)$ *interpola* los pares (c_j, b_j), $1 \leq j \leq k$: es decir
$$p(c_j) = b_j\ ,\ 1 \leq j \leq k\,.$$
En efecto, sea
$$p(x) = \sum_{j=1}^{k} b_j\, q_j(x) \in K[x]\,.$$
Es sencillo verificar que $p(x)$ tiene las propiedades deseadas, utilizando las relaciones de dualidad del ítem 2 (anterior).

4. Otra propiedad que vamos a usar de esta familia de polinomios es la siguiente: si $1 \leq j \neq h \leq k$, entonces
$$\prod_{i=1}^{k} (x - c_i) \mid q_j(x)\, q_h(x)\,.$$

En efecto, recordemos que por definición $q_j(x)$ tiene a $(x-c_i)$ como factores primo, para $1 \leq i \neq j \leq k$. Para que el producto $q_j(x)\, q_h(x)$ sea divisible por $\prod_{i=1}^k (x-c_i)$, basta que el $q_h(x)$ tenga a $(x-c_j)$ como factor primo en su descomposición: pero, por definición, $q_h(x)$ tiene a $(x-c_i)$ como factores primo, para $1 \leq i \neq h \leq k$: en particular, el factor $(x-c_j)$ aparece en la factorización de $q_h(x)$, pues $j \neq h$!! El argumento anterior muestra que se verifica la relación de divisibilidad propuesta al comienzo de este ítem. $\triangle$

Teorema 8.6. *Sea K un cuerpo y sean $c_1, \ldots, c_k \in K$ elementos distintos de K, con $k \geq 2$, y sean $\mathcal{L} = \{q_1(x), \ldots, q_k(x)\}$ los polinomios de Lagrange asociados a $c_1, \ldots, c_k$. Sea*

$$K_{k-1}[x] = \{p(x) \in K[x] \ : \ p(x) = 0 \quad ó \quad gr(p(x)) \leq k-1\}\,.$$

Entonces $\mathcal{L}$ es una base de $K_{k-1}[x]$ (como K-ev). Más aún, si $p(x) \in K_{k-1}[x]$ entonces

$$[p(x)]_{\mathcal{L}} = (p(c_1), \ldots, p(c_k)) \in K^k\,.$$

Demostración. Notemos primero que $K_{k-1}[x]$ es un K-ev; además, $\{1,\, x,\, \ldots,\, x^{k-1}\}$ es una base de $K_{k-1}[x]$, de forma que $\dim_K K_{k-1}[x] = k$.

Veamos que $\mathcal{L}$ es conjunto linealmente independiente de $K_{k-1}[x]$. Primero, es claro que que $\mathcal{L}$ es subconjunto $K_{k-1}[x]$, pues $\mathrm{gr}(q_j(x)) = k-1$ para $1 \leq j \leq k$ (por construcción, ver la Observación 8.5). Sean $\alpha_1, \ldots, \alpha_k \in K$ tales que

$$\alpha_1\, q_1(x) + \ldots + \alpha_k\, q_k(x) = 0 \in K_{k-1}[x]\,.$$

Utilizamos el hecho de que la evaluación en un escalar respeta la suma de polinomios: concretamente, si $1 \leq h \leq k$, evaluamos la combinación lineal en c_h:

$$0 = 0(c_h) = (\alpha_1\, q_1(x) + \ldots + \alpha_k\, q_k(x))(c_h) = \alpha_1\, q_1(c_h) + \ldots + \alpha_k\, q_k(c_h) = \alpha_h\,,$$

de forma que $\alpha_h = 0$, donde hemos usado las relaciones de dualidad de la Eq. (41), es decir,

$$\alpha_j\, q_j(c_h) = \alpha_j\, \delta_{j\,h}$$

en cada término (así, la suma $\alpha_1\, q_1(c_h) + \ldots + \alpha_k\, q_k(c_h)$ de arriba, el único término posiblemente no nulo es el h-ésimo, que se transforma en $\alpha_h\, q_h(c_h) = \alpha_h$). El argumento anterior muestra que $\alpha_h = 0$, $1 \leq h \leq k$; entonces $\mathcal{L}$ es l.i. Como $\mathcal{L}$ tiene k elementos, entonces $\mathcal{L}$ es base de $K_{k-1}[x]$.

Ahora que sabemos que $\mathcal{L}$ es base de $K_{k-1}[x]$ entonces, dado $p(x) \in K_{k-1}[x]$ exiten únicos coeficientes $\beta_1, \ldots, \beta_k \in K$ tales

$$p(x) = \sum_{j=1}^{k} \beta_j\ q_j(x) \in K[x]\,.$$

Si $1 \leq h \leq k$ y evaluamos a ambos miembros en el coeficiente $c_h \in K$ entonces

$$p(c_h) = (\sum_{j=1}^{k} \beta_j\ q_j(x))\,(c_h) = \sum_{j=1}^{k} \beta_j\ q_j(c_h) = \beta_h$$

donde hemos usado nuevamente las relaciones de dualidad de la Eq. (41). En conclusión $\beta_h = p(c_h)$, $1 \leq h \leq k$. Así, el vector de coordenadas de $p(x)$ con respecto a la base $\mathcal{L}$ es

$$[p(x)]_{\mathcal{L}} = (\beta_1, \ldots, \beta_k) = (p(c_1), \ldots, p(c_k)) \in K^k\,.$$

$\square$

Obs 8.7. Sea K un cuerpo, y sean $c_1, \ldots, c_k \in K$ distintos, con $k \geq 2$. Consideramos los polinomios de Lagrange asociados $\mathcal{L} = \{q_1(x), \ldots, q_k(x)\}$. Por el teorema anterior, $\mathcal{L}$ es base $K_{k-1}[x]$. Además, también podemos considerar la base canónica de este espacio, dada por $B = \{1,\, x,\, x^2,\, \ldots,\, x^{k-1}\}$. Sea $M = M_{\mathcal{L},B} \in K^{k\times k}$ la matriz de cambio de base de la base B en la base $\mathcal{L}$. Por el Teorema 8.6, podemos calcular los vectores columna de M (vector de coordenadas)

$$[x^j]_{\mathcal{L}} = (c_1^j, \ldots, c_k^j)\ , \quad 0 \leq j \leq k-1\,.$$

Entonces

$$M = \begin{pmatrix} 1 & c_1 & c_1^2 & \ldots & c_1^{k-1} \\ 1 & c_2 & c_2^2 & \ldots & c_2^{k-1} \\ \vdots & \vdots & \vdots & \ldots & \vdots \\ 1 & c_k & c_k^2 & \ldots & c_k^{k-1} \end{pmatrix} \in K^{k\times k}\,.$$

La matriz M de arriba también es llamada matriz de Vandermonde y notada $M = V(c_1, \ldots, c_k)$. Notemos que por construcción, $V(c_1, \ldots, c_k)$ es matriz inversible (ya que se trata de una matriz de cambio de base). $\triangle$

8.3 Primeras aplicaciones de polinomios: el caso diagonalizable

En esta sección vamos a dar dos nuevas caracterizaciones de los operadores que son diagonalizables.

Teorema 8.8. *Sea $\mathcal{V}$ un K-ev, $\dim_K \mathcal{V} = n$. Sea $T \in L(\mathcal{V})$ y sean $\lambda_1, \ldots, \lambda_k \in K$ distintos dos a dos. Son equivalentes:*

1. *T es diagonalizable y $\lambda_1, \ldots, \lambda_k$ son los autovalores distintos de T;*
2. *Si $m_T(x) \in K[x]$ denota el polinomio minimal de T entonces*

$$m_T(x) = \prod_{j=1}^{k} (x - \lambda_j)\,.$$

3. *Existe un sistema de proyecciones $\{E_1, \ldots, E_k\}$ tal que $E_j \neq 0$ para $1 \leq i \leq k$ que satisface*

$$T = \sum_{j=1}^{k} \lambda_j\, E_j\,.$$

Más aún, si alguna de las condiciones de arriba se verifica, entonces el sistema $\{E_1, \ldots, E_k\}$ del ítem 3. está unívocamente determinado por la condición $Im(E_j) = E_T(\lambda_j)$, $1 \leq j \leq k$.

Demostración. $1 \implies 2$. Si T es diagonalizable, por el Teorema 6.17

$$p_T(x) = \prod_{j=1}^{k} (x - \lambda_j)^{d_j}$$

para $d_j = \dim_K E_T(\lambda_j) \geq 1$, $1 \leq j \leq k$. Como $m_T(x) | p_T(x)$ y $m_T(x)$ y $p_T(x)$ comparten raíces, por los Teoremas 7.8 y 7.9, entonces debemos tener que

$$m_T(x) = \prod_{j=1}^{k} (x - \lambda_j)^{s_j} \qquad \text{para ciertas potencias} \qquad 1 \leq s_j \leq d_j\ ,\ 1 \leq j \leq k\,. \tag{42}$$

Además, sabemos que existe una base $B = \{v_1, \ldots, v_n\}$ tal que $T\, v_i = \mu_i\; v_i$ donde $\mu_i \in K$ coincide con algún $\lambda_1, \ldots, \lambda_k$, para $1 \leq i \leq n$. Definamos

$$m(x) = \prod_{j=1}^{k} (x - \lambda_j)\,. \tag{43}$$

Notemos que si verificamos que $m(T) = 0 \in L(\mathcal{V})$ entonces $m(x)$ debe ser el polinomio minimal, por la Eq. (42) (pues $m(x)$ sería el polinomio de grado más chico entre todos los posibles candidatos para $m_T(x)$ dados por la Eq. (42), ya que estaríamos considerando $s_j = 1$, $1 \leq j \leq k$).

Con la notación de párrafos anteriores: si $T\, v_i = \mu_i\, v_i$ entonces $m(\mu_i) = 0$, pues $\mu_i \in K$ coincide con algún $\lambda_1, \ldots, \lambda_k$ y estos últimos escalares son todos raíces de $m(x)$(ver Eq. (43)). Si aplicamos el Lemma 7.7 (tomando $p(x) = m(x)$) entonces

$$m(T)\, v_i = m(\mu_i)\, v_i = 0\, v_i = 0 \quad \text{ para } \quad 1 \leq i \leq n\,.$$

Como $B = \{v_1, \ldots, v_n\}$ es base de $\mathcal{V}$ y $m(T)\, v_i = 0$, $1 \leq i \leq n$, concluimos que $m(T) = 0 \in L(\mathcal{V})$ y entonces $m_T(x) = m(x) = \prod_{j=1}^{k}(x - \lambda_j)$; así, vale 2.

2. $\implies$ 3. Si $k = 1$, entonces $m_T(x) = (x - \lambda_1)$ lo que implica que $T - \lambda_1\, I = 0$ es decir, $T = \lambda_1\, I$. En particular, el ítem 3 vale tomando el sistema de proyecciones $\{I\}$ formado solo por la identidad!

Si $k \geq 2$ entonces consideramos $\mathcal{L} = \{q_1(x), \ldots, q_k(x)\}$ los polinomios de Lagrange asociados a los escalares (dos a dos distintos) $\lambda_1, \ldots, \lambda_k \in K$. Entonces, podemos aplicar el Teorema 8.6. En particular, representamos al humilde polinomio $p(x) = x \in K_{k-1}[x]$ (pues $1 \leq k - 1$). Así, como $p(\lambda_h) = \lambda_h$ (por definición de $p(x) = x$) concluimos que

$$x = \sum_{j=1}^{k} \lambda_j\, q_j(x) \implies T = \sum_{j=1}^{k} \lambda_j\, q_j(T)$$

donde la segunda identidad se obtiene de evaluar en $T \in L(\mathcal{V})$ a ambos lados de la identidad de polinomios. Esto sugiere definir

$$E_j = q_j(T) \in L(\mathcal{V}) \quad \text{ para } \quad 1 \leq j \leq k \implies T = \sum_{j=1}^{k} \lambda_j\, E_j\,.$$

Veamos que $\{E_1, \ldots, E_k\}$ así definidos forman un sistemas de proyecciones no nulas. Por un lado, si aplicamos el Teorema 8.6 al todavía más humilde polinomio $p(x) = 1$, como $1(\lambda_h) = 1$ (es polinomio constante!) entonces

$$1 = \sum_{j=1}^{k} 1\, q_j(x) \implies I = \sum_{j=1}^{k} q_j(T) = \sum_{j=1}^{k} E_j$$

donde la segunda identidad de operadores se obtiene de indentidad de polinomios, evaluando a ambos miembros en T. Esta identidad verifica una de las propiedades de un sistema de proyecciones (ver Definición 6.23).

Si $1 \leq h \neq j \leq k$ entonces por el ítem 3 de la Observación 8.5, sabemos que

$$m_T(x) = \prod_{i=1}^{k} (x - \lambda_i) \mid q_j(x)\, q_h(x) \implies q_j(x)\, q_h(x) = m_T(x)\, r(x)$$

para algún $r(x) \in K[x]$. Si evaluamos a ambos lados de la identidad en el operador T:

$$E_j\, E_h = q_j(T)\, q_h(T) = m_T(T)\, r(T) = 0\; r(T) = 0$$

donde usamos que $m_T(x)$ anula a T, las propiedades del producto en $L(\mathcal{V})$ ($0\,S = 0$ para $S \in L(\mathcal{V})$) y la definición de E_j y E_h. Entonces, $E_j\,E_h = 0$ si $1 \leq j \neq h \leq k$. En particular,

$$E_j = E_j\,I = E_j\,(E_1 + \ldots + E_k) = E_j\,E_1 + \ldots + E_j\,E_k = E_j\,E_j$$

es decir, $E_j = E_j\,E_j$ lo que entonces muestra que

$$E_j\,E_h = \delta_{j\,h} E_j \quad \text{para todos} \quad 1 \leq j,\, h \leq k\,.$$

Así, $\{E_1, \ldots, E_k\}$ es sistema de proyecciones y $E_j = q_j(T) \neq 0$, pues $\mathrm{gr}(q_j(x)) = k-1 < \mathrm{gr}(m_T(x))$, de forma que $q_j(x)$ no puede anular a T (recordemos que $m_T(x)$ es el polinomio de grado más chico que anula a T). Todo lo anterior garantiza que vale 3.

3. $\implies$ 1. Definamos $\mathcal{W}_i = \mathrm{Im}(E_i) \neq \{0\}$ (pues $E_i \neq 0$), $1 \leq i \leq k$. Entonces, por el Teorema 6.29 se tiene la descomposición en suma directa

$$\mathcal{V} = \mathcal{W}_1 \dot{+} \ldots \dot{+} \mathcal{W}_k\,.$$

Sea $1 \leq j \leq k$ y sea $w_j \in \mathcal{W}_j$ un vector arbitrario. Entonces $E_j(w_j) = w_j$ pues E_j fija los vectores de su imágen (porque E_j es proyector). Así, si $h \neq j$ entonces

$$E_h(w_j) = E_h(E_j(w_j)) = (E_h\,E_j)w_j = 0\,w_j = 0$$

pues $E_h\,E_j = 0 \in L(\mathcal{V})$. Así,

$$T\,w_j = (\sum_{h=1}^{k} \lambda_h\,E_h)(w_j) = \sum_{h=1}^{k} \lambda_h\,E_h(w_j) = \lambda_j\,E_j(w_j) = \lambda_j\,w_j\,.$$

Lo anterior muestra que $T\,w_j = \lambda_j\,w_j$ para todo $w_j \in \mathcal{W}_j$: en particular,

$$\mathcal{W}_j \subseteq E_T(\lambda_j) \quad \text{para} \quad 1 \leq j \leq k\,.$$

Lo anterior muestra que $\dim_K \mathcal{W}_j \leq \dim_K E_T(\lambda_j)$. Pero entonces

$$n = \sum_{j=1}^{k} \dim_K \mathcal{W}_j \leq \sum_{j=1}^{k} \dim_K E_T(\lambda_j) = \dim_K(\sum_{j=1}^{k} E_T(\lambda_j)) \leq n$$

donde usamos el Teorema 6.14 dos veces: una vez para la primer igualdad, y otra vez para la última desigualdad (ver también el Corolario 6.15). Las desigualdades anteriores y la última igualdad prueban que la desigualdad $\dim_K \mathcal{W}_j \leq \dim_K E_T(\lambda_j)$ es en realidad una igualdad $\dim_K \mathcal{W}_j = \dim_K E_T(\lambda_j)$; entonces, $\mathcal{W}_j = E_T(\lambda_j)$, para $1 \leq j \leq k$ (porque $\mathcal{W}_j \subseteq E_T(\lambda_j)$ y tienen la misma dimensión).

En conclusión, $\sum_{j=1}^{k} \dim_K E_T(\lambda_j) = n$ lo que dice que T es diagonalizable (por el Corolario 6.15) y $\mathcal{W}_j = E_T(\lambda_j)$, para $1 \leq j \leq k$. Así, vale el ítem 1. y la última afirmación del enunciado. $\square$

8.4 Más aplicaciones: Teorema de la descomposición prima

Comenzamos con el siguiente lema auxiliar

Lema 8.9. *Sea $\mathcal{V}$ un K-ev, sea $T \in L(\mathcal{V})$ y sea $\mathcal{W}$ subespacio T-invariante. Entonces:*

1. *Si $p(x) \in K[x]$ entonces $\mathcal{W}$ es $p(T)$-invariante.*
2. *Además $p(T)|_{\mathcal{W}} = p(T|_{\mathcal{W}}) \in L(\mathcal{W})$.*

Demostración. Para ver el ítem 1, primero verificamos que $\mathcal{W}$ es T^j-invariante, $j \geq 0$, por inducción en j.

Si $j = 0$ entonces $T^0 = I$ y es claro que $\mathcal{W}$ es I-invariante: si $w \in \mathcal{W}$ entonces $Iw = w \in \mathcal{W}$.

Supongamos que $\mathcal{W}$ es T^{j-1}-invariante, es decir: $T^{j-1}\, z \in \mathcal{W}$, para todo $z \in \mathcal{W}$ (hipótesis inductiva).

Sea $w \in \mathcal{W}$: entonces, por hipótesis, $Tw = z \in \mathcal{W}$ (pues $\mathcal{W}$ es T-invariante). Así,

$$T^j w = T^{j-1}\, T\, w = T^{j-1} z \in \mathcal{W}$$

pues $z = Tw \in \mathcal{W}$ y por hip. inductiva.

Lo anterior muestra que $\mathcal{W}$ es T^j-invariante, $j \geq 0$. Si $p(x) = a_0 + a_1\, x + \ldots + a_m\, x^m \in K[x]$ con $a_0, \ldots, a_m \in K$, entonces $p(T) = a_0\, I + a_1\, T + \ldots + a_m\, T^m$. Así, si $w \in \mathcal{W}$ entonces

$$p(T)\, w = (a_0\, I + a_1\, T + \ldots + a_m\, T^m)\, w = a_0\, w + a_1\, T\, w + \ldots + a_m\, T^m\, w \in \mathcal{W}$$

pues cada término $a_j\, T^j w \in \mathcal{W}$, por la primer parte de la prueba. Así, $\mathcal{W}$ es $p(T)$-invariante.

Para ver el ítem 2, primero miremos con cuidado qué es lo que dice: por un lado, $\mathcal{W}$ es T-invariante, de forma que podemos considerar $T|_{\mathcal{W}} \in L(\mathcal{W})$. Así, dado $p(x) \in K[x]$ podemos evaluar el operador $T|_{\mathcal{W}}$ en $p(x)$, obteniendo $p(T|_{\mathcal{W}}) \in L(\mathcal{W})$. Por otro lado, como $\mathcal{W}$ es $p(T)$-invariante (por lo probado en el ítem 1) entonces podemos considerar la restricción del operador $p(T)$ a $\mathcal{W}$, que notamos $p(T)|_{\mathcal{W}}$. El ítem 2 asegura que estos dos operadores son en realidad el mismo. Ahora que hemos aclarado qué es lo que queremos probar, consideramos los detalles de la prueba.

Notar que, por definición, siempre que $w \in \mathcal{W}$ entonces $T|_{\mathcal{W}}\, w = Tw$. Como antes, comenzamos probando que $T^j|_{\mathcal{W}} = (T|_{\mathcal{W}})^j$ para $j \geq 0$, por inducción en j.

Si $j = 0$, $T^0 = I_{\mathcal{V}}$ (el operador identidad en $\mathcal{V}$) y $(T|_{\mathcal{W}})^0 = I_{\mathcal{W}}$ (el operador identidad en $\mathcal{W}$) y es claro que $(I_{\mathcal{V}})|_{\mathcal{W}} = I_{\mathcal{W}}$ (ejercicio).

Supongamos que $T^{j-1}|_{\mathcal{W}} = (T|_{\mathcal{W}})^{j-1}$ para $j - 1 \geq 0$. Consideramos $w \in \mathcal{W}$: entonces $T|_{\mathcal{W}}\, w = T\, w = z \in \mathcal{W}$ y

$$\begin{aligned} (T^j)|_{\mathcal{W}}\, w &= T^j\, w = T^{j-1}\, (T\, w) = T^{j-1} z = (T^{j-1})|_{\mathcal{W}} z = (T|_{\mathcal{W}})^{j-1} z \\ &= (T|_{\mathcal{W}})^{j-1} T|_{\mathcal{W}}\, w = (T|_{\mathcal{W}})^j\, w \end{aligned}$$

donde usamos la definición de la restricción, definición de T^j, definición de restricción a $\mathcal{W}$, hipótesis inductiva (en la quinta igualdad), y definición de $(T|_{\mathcal{W}})^j$. Si $p(T) = a_0\, I + a_1\, T + \ldots + a_m\, T^m$ como antes y $w \in \mathcal{W}$ entonces:

$$p(T)|_{\mathcal{W}}\, w = (a_0\, I + a_1\, T + \ldots + a_m\, T^m)\, w = a_0\, w + a_1\, T\, w + \ldots + a_m\, T^m\, w =$$

$$a_0\, I|_{\mathcal{W}}\, w + a_1\, T|_{\mathcal{W}}\, w + \ldots + a_m\, T^m|_{\mathcal{W}}\, w = a_0\, I|_{\mathcal{W}}\, w + a_1\, T|_{\mathcal{W}}\, w + \ldots + a_m\, (T|_{\mathcal{W}})^m\, w$$

$$= p(T|_{\mathcal{W}})\, w\, .$$

□

El lema anterior parece una especie de trabalenguas matemático (sobre todo la prueba). En realidad es una observación técnica que nos permite simplificar muchas pruebas. Por ejemplo, el siguiente resultado importante es una consecuencia inmediata.

Proposición 8.10. *Sea $\mathcal{V}$ un K-ev, $\dim_K \mathcal{V} = n$. Sea $T \in L(\mathcal{V})$ y sea $m_T(x) \in K[x]$ su polinomio minimal. Si $\mathcal{W} \subseteq \mathcal{V}$ es subespacio T-invariante y si $m_{T|_{\mathcal{W}}}(x) \in K[x]$ denota el polinomio minimal de $T|_{\mathcal{W}} \in L(\mathcal{W})$ entonces $m_{T|_{\mathcal{W}}}(x) \mid m_T(x)$.*

Demostración. Por el lema anterior, tenemos que

$$m_T(T|_{\mathcal{W}}) = m_T(T)|_{\mathcal{W}} = 0|_{\mathcal{W}} = 0 \in L(\mathcal{W}),$$

pues $m_T(T) = 0$ y la restricción $0|_{\mathcal{W}}$ es el operador nulo actuando en $\mathcal{W}$. Por la propiedad del polinomio minimal de $T|_{\mathcal{W}} \in L(\mathcal{W})$ entonces $m_{T|_{\mathcal{W}}}(x) \mid m_T(x)$.

□

Teorema 8.11 (Teorema de la descomposición prima)**.** *Sea $\mathcal{V}$ un K-ev, $\dim_K \mathcal{V} = n$. Sea $T \in L(\mathcal{V})$ y sea $m_T(x) \in K[x]$ su polinomio minimal. Consideremos la factorización prima*

$$m_T(x) = p_1(x)^{r_1} \cdots p_k^{r_k}(x)$$

donde $p_1(x), \ldots, p_k(x) \in K[x]$ son polinomios primos mónicos distintos y $1 \leq r_j$, $1 \leq j \leq k$ (únicos salvo órden). Si definimos

$$\mathcal{W}_j = N(p_j^{r_j}(T)) \subseteq \mathcal{V}\,, \quad 1 \leq j \leq k\,,$$

entonces:

1. *$\mathcal{V} = \mathcal{W}_1 \dot{+} \ldots \dot{+} \mathcal{W}_k$ (descomposición en suma directa de subespacios independientes).*
2. *$\mathcal{W}_j$ es T-invariante, $1 \leq j \leq k$.*
3. *Si $T_j = T|_{\mathcal{W}_j} \in L(\mathcal{W}_j)$ entonces su polinomio minimal es $m_{T_j}(x) = p_j^{r_j}(x)$, $1 \leq j \leq k$.*

Demostración. Comenzamos definiendo polinomios mediante la construcción desarrollada en la Observación 8.3. Así, para $1 \leq j \leq k$, sea:

$$f_j = \prod_{1 \leq i \neq j \leq k} p_i^{r_i}(x) \in K[x]\,. \tag{44}$$

Como se justificó en esa Observación, existen polinomios $h_1(x), \ldots, h_k(x) \in K[x]$ tales que

$$1 = \sum_{j=1}^{k} f_j(x)\, h_j(x)\,.$$

Además, por los comentarios al final de la Observación 8.3, si $1 \leq j \neq \ell \leq k$ entonces

$$m_T(x) \,|\, f_j(x)\, f_\ell(x)\,.$$

Así, si definimos $q_j(x) = f_j(x)\, h_j(x)$ para $1 \leq j \leq k$ entonces:

$$1 = \sum_{j=1}^{k} q_j(x) \quad \text{y} \quad m_T(x) \,|\, q_j(x)\, q_\ell(x) \quad \text{para} \quad 1 \leq j \neq \ell \leq k\,. \tag{45}$$

Definimos $E_j = q_j(T) \in L(\mathcal{V})$, para $1 \leq j \leq k$: entonces, evaluando en T en la Eq. (45) tenemos que

$$\sum_{j=1}^{k} E_j = 1(T) = I \quad \text{y} \quad E_j\, E_\ell = 0 \quad \text{si} \quad j \neq \ell\,. \tag{46}$$

Con respecto a la segunda identidad de más arriba: $E_j\, E_\ell = q_j(T)\, q_\ell(T) = 0$, pues el polinomio minimal $m_T(x)$ divide al producto $q_j(x)\, q_\ell(x)$, (de forma que $q_j(x)\, q_\ell(x) = m_T(x)\, r_{j,\ell}(x)$ y $m_T(T) = 0$). Usando las identidades de la Eq. (46), ahora podemos ver que

$$E_j = E_j\, I = E_j\,(E_1 + \ldots + E_k) = E_j\, E_1 + \ldots + E_j\, E_k = E_j\, E_j$$

lo que muestra que $E_j = E_j^2$. En resumen, $E_j\, E_\ell = \delta_{j\,\ell}\, E_j$, $1 \leq j,\, \ell \leq k$; este hecho, junto con la Eq. (46) muestran que $\{E_1, \ldots, E_k\}$ es sistema de proyecciones. Más aún, como cada $E_j = q_j(T)$ entonces,

$$T\, E_j = T\, q_j(T) = (x\, q_j(x))(T) = (q_j(x)\, x)(T) = q_j(T)\, T = E_j\, T\ ,\ 1 \leq j \leq k\,,$$

donde hemos usado que el producto de polinomios es conmutativo y las propiedades de la evaluación de polinomios en T. Por los Teoremas 6.29 y 6.33, vemos que los subespacios $\mathcal{W}_j$ =Im(E_j), $1 \leq j \leq k$, verifican:

$$\mathcal{V} = \mathcal{W}_1 \dot{+} \ldots \dot{+} \mathcal{W}_k \quad \text{y} \quad \mathcal{W}_j \ \text{ es } T\text{-invariante }\ ,\ 1 \leq j \leq k\,.$$

Sea $1 \leq j \leq k$ fijo: veamos ahora que $\mathrm{Im}(E_j) = \mathcal{W}_j$ coincide con el núcleo $\mathrm{N}(p_j^{r_j}(T))$:

Si $w \in \mathrm{Im}(E_j) = \mathcal{W}_j$ entonces $w = E_j\, w = q_j(T)\, w$. Así

$$p_j^{r_j}(T)\, w = p_j^{r_j}(T)\, q_j(T)\, w = m_T(T)\, h_j(T)\, w = 0\, h_j(T)\, w = 0\, w = 0$$

donde hemos usado que $p_j^{r_j}(x)\, f_j(x) = m_T(x)$, por construcción de $f_j(x)$ (ver la Eq. (44)), de forma que:

$$p_j^{r_j}(x)\, q_j(x) = p_j^{r_j}(x)\, f_j(x)\, h_j(x) = m_T(x)\, h_j(x) \implies p_j^{r_j}(T)\, q_j(T) = m_T(T)\, h_j(T) = 0 \in L(\mathcal{V})$$

pues $m_T(T) = 0$. En resumen, si $w \in \mathcal{W}_j$ entonces $p_j^{r_j}(T)\, w = 0$; entonces $\mathcal{W}_j \subset \mathrm{N}(p_j^{r_j}(T))$.

Si $w \in \mathrm{N}(p_j^{r_j}(T))$ entonces $p_j^{r_j}(T)\, w = 0$. Así, si $1 \leq j \neq \ell \leq k$, el único índice excluído en la productoria que define $f_\ell(x)$ es ℓ y $j \neq \ell$: en particular,

$$f_\ell(T)\, w = \left(\prod_{1 \leq i \leq k,\, i \neq \ell,\, i \neq j} p_i^{r_i}(T) \right) p_j^{r_j}(T)\, w = \left(\prod_{1 \leq i \leq k,\, i \neq \ell,\, i \neq j} p_i^{r_i}(T) \right) 0 = 0 \in \mathcal{V} \tag{47}$$

pues la productoria entre los paréntesis es un operador lineal (y manda el vector nulo en el vector nulo de $\mathcal{V}$). Así,

$$E_\ell\, w = q_\ell(T)\, w = h_\ell(T)\, f_\ell(T)\, w = h_\ell(T)\, 0 = 0 \quad \text{para} \quad 1 \leq j \neq \ell \leq k\,, \tag{48}$$

pues $h_\ell(T) \in L(\mathcal{V})$ es operador lineal. Usando que $\{E_1, \ldots, E_k\}$ es sistema de proyecciones y la Eq. (48)

$$w = (E_1 + \ldots + E_k)\, w = E_1\, w + \ldots + E_k\, w = E_j\, w \in \mathrm{Im}(E_j) = \mathcal{W}_j\,,$$

pues el término $E_\ell w = 0$, para $1 \leq j \neq \ell \leq k$. Así, $N(p_j^{r_j}(T)) \subset \mathcal{W}_j$ lo que implica que $N(p_j^{r_j}(T)) = \mathcal{W}_j$.

Los argumentos anteriores muestran que valen los ítems 1. y 2. del enunciado.

Finalizamos la prueba con la verificación del ítem 3: por un lado, hemos visto que si $w \in \mathcal{W}_j$ entonces $p_j^{r_j}(T)\, w = 0$. Entonces, si $w \in \mathcal{W}_j$

$$p_j^{r_j}(T|_{\mathcal{W}_j})\ w = p_j^{r_j}(T)|_{\mathcal{W}_j}\ w = p_j^{r_j}(T)\ w = 0$$

donde usamos el Lema 8.9 en la primer igualdad, y la definición de la evaluación de cualquier restricción en la segunda igualdad. Como $w \in \mathcal{W}_j$ era arbitrario, entonces

$$p_j^{r_j}(T_j) = p_j^{r_j}(T|_{\mathcal{W}_j}) = 0 \in L(\mathcal{W}_j) \implies m_{T_j}(x)\,|\,p_j^{r_j}(x)\,, \tag{49}$$

por la propiedad del polinomio minimal de $T_j = T|_{\mathcal{W}_j} \in L(\mathcal{W}_j)$.

Recíprocamente, sea $g(x) \in K[x]$ es tal que $g(T_j) = 0 \in L(\mathcal{W}_j)$. Afirmamos que $g(x)\, f_j(x) \in K[x]$ anula a T, es decir $g(T)\, f_j(T) = 0 \in L(\mathcal{V})$. En efecto, si $1 \leq j \neq \ell \leq k$ y $w_\ell \in \mathcal{W}_\ell$ entonces (ver Eq. (47), pero tener en mente que estamos cambiando los roles de j y ℓ)

$$f_j(T)\, w_\ell = 0 \implies g(T)\, f_j(T)\, w_\ell = 0 \quad \text{para} \quad w_\ell \in \mathcal{W}_\ell \ , \ \ 1 \leq j \neq \ell \leq k\,. \tag{50}$$

Además, por hipótesis $g(T_j) = 0$: así, si $w_j \in \mathcal{W}_j$ entonces (usando nuevamente el Lema 8.9)

$$g(T)\, w_j = g(T)|_{\mathcal{W}_j}\, w_j = g(T_j)\, w_j = 0\, w_j = 0 \in \mathcal{V}$$

$$\implies g(T)\, f_j(T)\, w_j = f_j(T)\, g(T)\, w_j = f_j(T)\, 0 = 0 \in \mathcal{V}\,. \tag{51}$$

Si combinamos las Eqs. (50) y (51): dado $v \in \mathcal{V}$ existen únicos $w_j \in \mathcal{W}_j$, $1 \leq j \leq k$, tales que $v = w_1 + \ldots + w_k$: entonces

$$g(T)\, f_j(T)\, v = g(T)\, f_j(T)\, (w_1 + \ldots + w_k) = g(T)\, f_j(T)\, w_1 + \ldots + g(T)\, f_j(T)\, w_k = 0 + \ldots + 0 = 0\,.$$

Como $v \in \mathcal{V}$ era arbitrario, entonces

$$g(T)\, f_j(T) = 0\,.$$

En particular, $m_T(x) \,|\, g(x)\, f_j(x)$, por la propiedad del minimal $m_T(x)$. Como $p_j^{r_j}(x) \,|\, m_T(x)$ (por construcción) y $p_j^{r_j}(x)$ y $f_j(x)$ son co-primos, pues no tienen factores primos en común, entonces

$$p_j^{r_j}(x) \,|\, g(x)\, f_j(x) \quad \text{y} \quad \text{m.c.d.}(p_j^{r_j}(x),\, f_j(x)) = 1 \implies p_j^{r_j}(x) \,|\, g(x)\,,$$

donde m.c.d. significa máximo común divisor, y donde hemos usado un resultado de divisibilidad de polinomios de Álgebra I : si $a(x)|b(x)\, c(x)$ y m.c.d.$(a(x), b(x)) = 1$ entonces $a(x)|c(x)$.

En resumen, hemos partido de $g(x) \in K[x]$ tal que $g(T_j) = 0 \in L(\mathcal{W}_j)$ y hemos probado que $p_j^{r_j}(x) \,|\, g(x)$. Si aplicamos este argumento a $g(x) = m_{T_j}(x)$ entonces $p_j^{r_j}(x) \,|\, m_{T_j}(x)$. Esta relación de divisibilidad junto con la relación en la Eq. (49) muestran que $p_j^{r_j}(x) = m_{T_j}(x)$ y luego, vale también el ítem 3. □

El teorema de la descomposición prima es una herramienta fundamental en el estudio de operadores. En particular, nos muestra una estrategia para descomponer a operadores en términos de operadores más chicos y más sencillos, como planteamos al comienzo de la Sección 8. En las notas de la segunda parte de la teoría para el tp 6, vamos a usar el teorema de la descomposición prima y otras nuevas herramientas para poder obtener la existencia y características de la forma de Jordan asociada a ciertos operadores T.

Obs 8.12. Como primer ejemplo del uso del Teorema de la descomposición prima, veamos como podemos deducir si $T \in L(\mathcal{V})$ es tal que su polinomio minimal es $m_T(x) = \prod_{j=1}^k (x - \lambda_j)$, $\lambda_1, \ldots, \lambda_k \in K$ distintos dos a dos, entonces T es diagonalizable.

Aclaramos que esto ya ha sido probado en Teorema 8.8, pero nuestra idea es ver como el Teorema 8.11 nos permite re-obtener este hecho.

Suponemos entonces que el polinomio minimal de T se factoriza como el producto de factores lineales mónicos distintos. Con la notación del Teorema 8.11,

$$m_T(x) = p_1(x) \cdots p_k(x) \quad \text{con} \quad p_j(x) = (x - \lambda_j) \ , \ 1 \leq j \leq k\,.$$

Es decir, las potencias de los factores primos son $r_j = 1$, $1 \leq j \leq k$.

Entonces definimos $\mathcal{W}_j = \mathrm{N}(p_j(T)) = \mathrm{N}(T - \lambda_j\, I)$. Notemos que en este caso resulta que $\mathcal{W}_j = E_T(\lambda_j)$, $1 \leq j \leq k$. En particular, si $w \in \mathcal{W}_j$ entonces $Tw = \lambda_j\, w$.

El Teorema 8.11 garantiza, entre otras cosas, que

$$\mathcal{V} = \mathcal{W}_1 \dot{+} \ldots \dot{+} \mathcal{W}_k \,.$$

Así, si B_j es base de $\mathcal{W}_j$, $1 \leq j \leq k$, entonces $B = B_1 \cup \ldots \cup B_k$ es base de $\mathcal{V}$. Más aún, si v denota un vector arbitrario de B, entonces $v \in B_j \subset \mathcal{W}_j$ para algún, $1 \leq j \leq k$. En particular, $Tv = \lambda_j \, v$. Lo anterior muestra que todo vector de la base B es autovector de T (correspondiente a algún autovalor). En particular, $[T]_B \in K^{n\times n}$ es matriz diagonal. △

8.5 Diagonalización simultánea

Concluimos estas notas con el estudio de la llamada diagonalización simultánea de operadores. Si bien el resultado vale con más generalidad, vamos a considerar una versión sencilla que involucre sólo dos operadores. Comenzamos con la siguiente observación.

Obs 8.13. Sea $\mathcal{V}$ un K-ev, $\dim_K \mathcal{V} = n$. Sean $S, T \in L(\mathcal{V})$ operadores diagonalizables Supongamos que existe una base B de $\mathcal{V}$ tal que $[T]_B$ y $[S]_B$ son matrices diagonales; en este caso, decimos que la base B diagonaliza *simultáneamente* a S y T.

En este caso, existen $s_1, \ldots,\, s_n \in K$ y $t_1, \ldots, t_n \in K$ tales que

$$[S]_B = \operatorname{diag}(s_1, \ldots, s_n) \quad \text{ y } \quad [T]_B = \operatorname{diag}(t_1, \ldots, t_n)\,.$$

Hemos visto que entonces

$$[S]_B\,[T]_B = \operatorname{diag}(s_1\, t_1, \ldots, s_n\, t_n) = [T]_B\,[S]_B\,.$$

Las identidades anteriores muestran que

$$[ST]_B = [S]_B\,[T]_B = [T]_B\,[S]_B = [TS]_B \implies ST = TS$$

pues tomar matriz con respecto a una base es un isomorfismo (en particular, es inyectiva). △

Una consecuencia de la Proposición 8.10 y del Teorema 8.8 es el siguiente resultado que muestra que vale la recíproca de la Observación 8.13.

Teorema 8.14. *Sea $\mathcal{V}$ un K-ev, $\dim_K \mathcal{V} = n$. Sean $S, T \in L(\mathcal{V})$ operadores diagonalizables tales que $SA = TS$. Entonces existe una base B de $\mathcal{V}$ tal que $[T]_B$ y $[S]_B$ son matrices diagonales (la base B diagonaliza simultáneamente a S y T).*

Demostración. Sea T diagonalizable, con autovalores distintos $\lambda_1, \ldots, \lambda_k \in K$. Sea $\{E_1, \ldots, E_k\}$ el sistema de proyecciones como en el Teorema 8.8. En este caso, $E_T(\lambda_j) = \operatorname{Im}(E_j)$, de forma que

$$\mathcal{V} = E_T(\lambda_1) \dot{+} \ldots \dot{+} E_T(\lambda_k) \tag{52}$$

donde la descomposición en suma directa es consecuencia del Teorema 6.29. Notemos que cada autoespacio $E_T(\lambda_j)$ es S-invariante, $1 \leq j \leq k$: en efecto, recordemos que existen polinomios $q_j(x)$ tales que $E_j = q_j(T)$, $1 \leq j \leq k$ (ver la prueba de 2. $\implies$ 3. del Teorema 8.8). En particular,

$$E_j\, S = q_j(T)\, S = S\, q_j(T) = S\, E_j \quad \text{ para } \quad 1 \leq j \leq k\,. \tag{53}$$

En efecto, como $T\,S = S\,T$ entonces se verifica que

$$T^2\, S = T\,(TS) = T\,(ST) = (TS)\,T = (ST)T = ST^2$$

y en general $T^j\, S = S\, T^j$ para $j \geq 0$ (que verificamos haciendo inducción en j), y luego $p(T)\, S = S\, p(T)$ para todo polinomio $p(x) \in K[x]$ (hacer detalles). En particular, tomando $p(x) = q_j(x)$ justificamos las identidades de la Eq. (53).

Así, las relaciones de conmutación de la Eq. (53) junto con el Teorema 6.33 muestran que $E_T(\lambda_j) = \mathrm{Im}(E_j)$ es subespacio S-invariante, $1 \leq j \leq k$. Sea

$$m_S(x) = \prod_{i=1}^{h} (x - \mu_i) \in K[x] \quad \text{y} \quad S_j = S|_{E_T(\lambda_j)} \in L(E_T(\lambda_j))\ ,\ 1 \leq j \leq k \tag{54}$$

donde $\mu_1, \ldots, \mu_h$ son los autovalores distintos de S (aquí usamos el Teorema 8.8 aplicado a S, que es diagonalizable por hipótesis).

Sea $m_{S_j}(x)$ el polinomio minimal del operador S_j. Por la Proposición 8.10, $m_{S_j}(x) \,|\, m_S(x)$, lo que indica que $m_{S_j}(x)$ es producto de factores lineales (mónicos) distintos: de hecho, $m_{S_j}(x)$ tiene algunos de los factores que aparecen en la factorización de $m_S(x)$ de la Eq. (54). Por el Teorema 8.8, $S_j \in L(E_T(\lambda_j))$ es diagonalizable !! (este es un hecho general: la restricción de cualquier operador diagonalizable a un subespacio invariante resulta diagonalizable, que se prueba con el argumento anterior).

Sea $B_j = \{v_1^j, \ldots, v_{d_j}^j\}$ una base de $E_T(\lambda_j)$ tal que $S_j\, v_i^j = \rho_{i,j}\, v_i^j$, $1 \leq i \leq d_j$ (que diagonaliza a S_j). Entonces

$$S\, v_i^j = S|_{E_T(\lambda_j)}\, v_i^j = S_j\, v_i^j = \rho_{i,j}\, v_i^j \quad \text{para} \quad 1 \leq i \leq d_j\ .$$

Además, por definición de $E_T(\lambda_j)$: si $v \in E_T(\lambda_j)$ entonces $T\, v = \lambda_j\, v$. En particular,

$$T\, v_i^j = \lambda_j\, v_i^j \quad \text{para} \quad 1 \leq i \leq d_j\ .$$

Así, los vectores de B_j son autovectores de S y T, para $1 \leq j \leq k$. Entonces, si definimos $B = B_1 \cup \ldots \cup B_k$, resulta que B es base de $\mathcal{V}$ (por la Eq. (52)) y cada vector de B (que es vector de alguna base B_j) resulta autovector de S y T. Lo anterior inmediatamente indica que $[T]_B$ y $[S]_B$ son matrices diagonales. $\square$

9 Forma de Jordan

Comenzamos la sección con las siguientes observaciones, que indican el rol destacado de los operadores nilpotentes (ver Definición 9.3 más abajo).

Obs 9.1. Sea $\mathcal{V}$ un K-ev, $\dim_K \mathcal{V} = n$. Sea $T \in L(\mathcal{V})$ y supongamos que el polinomio minimal de T satisface

$$m_T(x) = \prod_{j=1}^{k} (x - \lambda_j)^{r_j} \in K[x] \tag{55}$$

donde $\lambda_1, \ldots, \lambda_k \in K$ son los autovalores de T *distintos* y $1 \leq r_j$, $1 \leq j \leq k$. Recordemos que $\mathrm{gr}(m_T(x)) \leq \mathrm{gr}(p_T(x)) = n$, pues $m_T(x) \,|\, p_T(x)$, donde $p_T(x)$ denota al polinomio característico de T.

Si aplicamos el Teorema 8.11 de la descomposición prima en este caso (notemos que cada polinomio $(x - \lambda_j)$ es primo), entonces los subespacios $\mathcal{W}_j = N((T - \lambda_j\, I)^{r_j})$ son T-invariantes, $1 \leq j \leq k$, y se verifica:

1. $\mathcal{V} = \mathcal{W}_1 \dot{+} \ldots \dot{+} \mathcal{W}_k$;

2. Si $T_j = T|_{\mathcal{W}_j} \in L(\mathcal{W}_j)$ tiene polinomio minimal $m_{T_j}(x) = (x - \lambda_j)^{r_j}$.

En este caso, los operadores T_j, $1 \leq j \leq k$, nos permiten describir al operador completo T (ver Observación 6.32). Entonces, si podemos entender la estructura interna de cada T_j, $1 \leq j \leq k$, podemos describir la estructura de T. $\triangle$

Obs 9.2. La Observación anterior sugiere concentrar nuestra atención en operadores $S \in L(\mathcal{W})$ tales que su polinomio minimal $m_S(x) = (x - \lambda)^r$, para cierto $\lambda \in K$, $1 \leq r$, donde $\mathcal{W}$ denota un K-ev de dimensión finita. (Notemos que todos los operadores $T_j \in L(\mathcal{W}_j)$, $1 \leq j \leq k$, definidos en la Observación 9.1, tienen esta propiedad).

En este caso, podemos escribir

$$S = \lambda\, I + (S - \lambda\, I) = \lambda\, I + N \in L(\mathcal{W})$$

donde $N = (S - \lambda\, I) \in L(\mathcal{W})$ verifica

$$N^r = (S - \lambda\, I)^r = m_S(S) = 0 \in L(\mathcal{W})\,,$$

pues $m_S(x) = (x - \lambda)^r$ anula a S. Así, S es la suma de un múltiplo del operador identidad (es decir, el operador más sencillo de todos!) y un operador N tal que $N^r = 0$. Si podemos entender la estructura de N, podemos entender entonces la estructura de S! $\triangle$

Definición 9.3. *Sea $\mathcal{V}$ un K-ev, $\dim_K \mathcal{V} = n$ y sea $T \in L(\mathcal{V})$.*

1. *Decimos que T es nilpotente si existe $1 \leq m$ tal que $T^m = 0$;*
2. *Si T es nilpotente definimos el índice de nilpotencia de T al mínimo natural $m \geq 1$ tal que $T^m = 0$.*

$\square$

Ejemplo 9.4. Sea

$$\mathcal{V} = \{p \in \mathbb{R}[x] \ : \ p = 0 \ \text{ ó } \operatorname{gr}(p) \leq 2\} = \{a_0 + a_1\, x + a_2\, x^2 : a_0,\, a_1,\, a_2 \in \mathbb{R}\} \subset \mathbb{R}[x]\,.$$

En este contexto $B = \{1,\, x,\, x^2\}$ es una base de $\mathcal{V}$, de forma que $\dim_{\mathbb{R}} \mathcal{V} = 3$. Consideramos el operador derivación

$$D \in L(\mathcal{V}) \quad \text{ dado por } \quad D(a_0 + a_1\, x + a_2 x^2) = a_1 + 2\, a_2\, x \in \mathcal{V}\,.$$

Notemos que si componemos a D con sí mismo tres veces, es decir, consideramos la potencia $D^3 = D\, D\ D \in L(\mathcal{V})$ se verifica:

$$D(a_0 + a_1\, x + a_2 x^2) = a_1 + 2\, a_2\, x \implies D^2(a_0 + a_1\, x + a_2 x^2) = D(a_1 + 2\, a_2\, x) = 2\, a_2$$

de forma que

$$D^3(a_0 + a_1\, x + a_2 x^2) = D(2\, a_2) = 0$$

En palabras: si derivamos tres veces cualquier polinomio de grado a lo sumo 2, obtenemos el polinomio nulo.

Lo anterior muestra que $D^3 = 0$ es el operador nulo en $\mathcal{V}$. En este caso D es un operador nilpotente. Más aún, el índice de nilpotencia de D es 3 (verificar, ejercicio). $\triangle$

Obs 9.5. Sea $\mathcal{V}$ un K-ev, $\dim_K \mathcal{V} = n$ y sea $T \in L(\mathcal{V})$ nilpotente. Entonces existe $m \geq 1$ tal que $T^m = 0$. Así, si $q(x) = x^m \in K[x]$ entonces $q(T) = 0$.

En particular, $m_T(x) | q(x)$ que a su vez implica que

$$m_T(x) = x^r\ , \ \ 1 \leq r \leq n\,.$$

En efecto, como $x \in K[x]$ es polinomio primo, entonces $q(x) = x^m$ está factorizado en factores primos. Como $m_T(x)|q(x)$ los factores primos de la factorización prima de $m_T(x)$ son algunos de los factores primos de la factorización prima de $q(x)$. Esto muestra que $m_T(x) = x^r$ para algún $1 \leq r$. Ademas, como $\text{gr}(m_T(x)) \leq n$ entonces $1 \leq r \leq n$.

Como consecuencia del párrafo anterior, concluimos que el único autovalor de T es $0 \in K$. En efecto, la única raíz de $m_T(x)$ es $\lambda = 0$, y el minimal y característico tienen las mismas raíces.

Para finalizar esta observación, notemos que $T^{r-1} \neq 0$: en efecto, si suponemos que $T^{r-1} = 0$ entonces $p(x) = x^{r-1} \in K[x]$ es tal que $p(T) = 0$, con $\text{gr}(p(x)) < \text{gr}(m_T(x))$, que contradice la propiedad del polinomio minimal $m_T(x)$. Así, $T^{r-1} \neq 0$ y $T^r = 0$; entonces, el mínimo natural $m \geq 1$ tal que $T^m = 0$ es $r = \text{gr}(m_T(x))$, es decir el índice de nilpotencia de T (ver Definición 9.3) coincide con el grado del polinomio minimal $m_T(x)$. △

9.1 Descomposición cíclica de operadores nilpotentes

Para obtener una representación importante de los operadores nilpotentes vamos a introducir los siguientes conceptos nuevos.

Definición 9.6. *Sea $\mathcal{V}$ un K-ev, $\dim_K \mathcal{V} = n$ y sea $T \in L(\mathcal{V})$. Dado $v \in \mathcal{V}$ definimos el subespacio T-cíclico generado por $v \in \mathcal{V}$, notado $Z(v,T) \subseteq \mathcal{V}$ dado por*

$$Z(v,T) = \{p(T)\, v \ :\ p(x) \in K[x]\}\,.$$

□

Con la notación de la definición anterior, notemos que si $p(x) \in K[x]$ entonces $p(T) \in L(\mathcal{V})$ es un operador: en este caso tiene sentido evaluar $p(T)\, v \in \mathcal{V}$; el conjunto $Z(v,T)$ está formado por estos vectores.

En el siguiente resultado describimos las primeras propiedades de $Z(v,T)$. En la prueba vamos a usar las siguientes propiedades de la evaluación de polinomios en un operador lineal T: si $p(x),\, q(x) \in K[x]$, $\alpha \in K$ entonces $\alpha\, p(x) + q(x) \in K[x]$, $p(x)\, q(x) \in K[x]$ y

$$(\alpha\, p(x) + q(x))(T) = \alpha\, p(T) + q(T) \in L(\mathcal{V}) \quad \text{ y } \quad (p(x)\, q(x))(T) = p(T)\, q(T) \in L(\mathcal{V})\,.$$

Proposición 9.7. *Sea $\mathcal{V}$ un K-ev, $\dim_K \mathcal{V} = n$ y sea $T \in L(\mathcal{V})$. Dado $v \in \mathcal{V}$,*

1. *$Z(v,T) \subseteq \mathcal{V}$ es subespacio;*
2. *$v \in Z(v,T)$ y $Z(v,T)$ es T-invariante;*
3. *Si $\mathcal{W} \subseteq \mathcal{V}$ es subespacio T-invariante tal que $v \in \mathcal{W}$ entonces $Z(v,T) \subseteq \mathcal{W}$.*

Demostración. Para ver 1, notemos que $p(x) = 0 \in K[x]$ es tal que $p(T) = 0 \in L(\mathcal{V})$ y entonces $p(T)\, v = 0 \in Z(v,T)$. Además, si $u = q(T)\, v$, $w = s(T)\, v \in Z(v,T)$, con $q(x),\, s(x) \in K[x]$ y si $\alpha \in K$ entonces

$$\alpha\, u + v = \alpha\, q(T)\, v + s(T)\, v = (\alpha\, q(T) + s(T))\, v = h(T)\, v \in Z(v,T)$$

donde $h(x) = \alpha\, q(x) + s(x) \in K[x]$, de forma que $h(T) = \alpha\, q(T) + s(T)$. La afirmación $h(T)\, v \in Z(v,T)$ es por la definición de $Z(v,T)$.

Para verificar el ítem 2: notar que $1 \in K[x]$, $1(T) = I$, de forma que $1(T)\, v = I\, v = v \in Z(v,T)$. Además, si $u = q(T)\, v \in Z(v,T)$, para $q(x) \in K[x]$ entonces:

$$T\, u = T\, q(T)\, v = (x\, q(x))(T)\, v \in Z(v,T)$$

donde hemos usado que $x\,q(x) \in K[x]$ es tal que $(x\,q(x))(T) = T\,q(T) \in L(\mathcal{V})$. La afirmación $(x\,q(x))(T)\,v \in Z(v,T)$ es por la definición de $Z(v,T)$.

El ítem 3 es ejercicio (sugerencia: usar el Lema 8.9). □

Definición 9.8. *Sea $\mathcal{V}$ un K-ev, $\dim_K \mathcal{V} = n$, sea $T \in L(\mathcal{V})$ y sea $v \in \mathcal{V}$. Definimos*

1. *El ideal*
$$\mathcal{I}(v,T) = \{p(x) \in K[x] \ : \ p(T)\,v = 0\} \subset K[x]$$
(verificar que $\mathcal{I}(v,T) \neq \{0\}$ es ideal no nulo, ejercicio).

2. *El polinomio T-anulador de v, notado $m_{v,T}(x) \in K[x]$, como el único polinomio mónico que satisface: $m_{v,T}(T)\,v = 0$ y dado $q(x) \in K[x]$,*
$$q(x) \in \mathcal{I}(v,T) \quad \textit{si y solo si} \quad m_{v,T}(x)\,|\,q(x)\,.$$

□

Obs 9.9. Con las notaciones de las definición anterior (notar que T es un operador arbitrario): si $m_T(x) \in K[x]$ denota el operador minimal de T, entonces $m_T(T) = 0 \in L(\mathcal{V})$. En particular, $m_T(T)\,v = 0\,v = 0$ de forma que
$$m_{v,T}(x)\,|\,m_T(x)\,.$$
Si además suponemos que T es nilpotente, entonces existe $1 \leq r \leq n$ tal que
$$m_T(x) = x^r \quad \text{ y } \quad m_{v,T}(x)\,|\,m_T(x) \implies m_{v,T}(x) = x^k \quad \text{con} \quad 0 \leq k \leq r \ (\leq n)\,.$$
Más aún, si $v \neq 0$ entonces $k \geq 1$; en efecto, $T^0\,v = I\,v = v \neq 0$, de forma que $1 = x^0 \in K[x]$ es tal que $1(T)\,v \neq 0$. △

Teorema 9.10. *Sea $\mathcal{V}$ un K-ev, $\dim_K \mathcal{V} = n$ y sea $T \in L(\mathcal{V})$. Sea $v \in \mathcal{V}$ tal que $v \neq 0$, sea $m_{v,T}(x) \in K[x]$ el polinomio T-anulador de v, y sea $1 \leq k \leq n$ tal que $gr(m_{v,t}(x) = k$. Entonces*

1. *$B_v = \{v,\, T\,v,\, \ldots,\, T^{k-1}\,v\}$ es base de $Z(v,T)$ y $\dim_K Z(v,T) = k$.*

2. *Si $m_{T|_{Z(v,T)}}(x)$ el polinomio minimal de la restricción $T|_{Z(v,T)} \in L(Z(v,T))$ entonces*
$$m_{v,T}(x) = m_{T|_{Z(v,T)}}(x)\,.$$

Demostración. Para verificar el ítem 1, vamos a mostrar que B_v es l.i. y que es un sistema de generadores de $Z(v,T)$.

B_v es l.i: supongamos que existen $\alpha_0, \ldots, \alpha_{k-1} \in K$ tales que
$$\sum_{j=0}^{k-1} \alpha_j\, T^j\, v = \alpha_0\, I\, v + \alpha_1\, T\, v + \ldots + \alpha_{k-1}\, T^{k-1}\, v = 0\,.$$

En este caso, definimos $p(x) = \alpha_0 + \alpha_1\, x + \ldots + \alpha_{k-1}\, x^{k-1} \in K[x]$. Si existe algún $\alpha_j \neq 0$, para $0 \leq j \leq k-1$, entonces $p(x) \neq 0$ y $\mathrm{gr}(p(x)) \leq k-1$. Por construcción
$$p(T)\,v = (\alpha_0\, I + \alpha_1\, T + \ldots + \alpha_{k-1}\, T^{k-1})\,v = \alpha_0\, I\, v + \alpha_1\, T\, v + \ldots + \alpha_{k-1}\, T^{k-1}\, v = 0\,.$$

Entonces, por definición de polinomio T-anulador, $m_{v,T}(x)\,|\,p(x)$, que contradice el hecho de que $\mathrm{gr}(p(x)) \leq k-1$. Esta contradicción surge de suponer que existe algún $\alpha_j \neq 0$, para $0 \leq j \leq k-1$. En conclusión, $\alpha_0 = \ldots = \alpha_{k-1} = 0$ y B_v es l.i.

B_v es sistema de generadores de $Z(v,T)$. Sea $u \in Z(v,T)$; entonces existe $q(x) \in K[x]$ tal que $u = q(T)\,v$. Si aplicamos el algoritmo de la división en $K[x]$ concluimos que existen $s(x)$, $r(x) \in K[x]$ tales que $q(x) = s(x)\,m_{v,T}(x) + r(x)$, con $r(x) = 0$ ó $\mathrm{gr}(r(x)) < k = \mathrm{gr}(m_{v,T}(x))$. En cualquier caso, existen $\beta_0, \ldots, \beta_{k-1} \in K$ tales que

$$r(x) = \beta_0 + \beta_1\,x + \ldots + \beta_{k-1}\,x^{k-1}$$

donde los beta's podrían ser todos nulos (si $r(x) = 0$). En particular, $q(T) = s(T)\,m_{v,T}(T) + r(T) \in L(\mathcal{V})$ y

$$u = q(T)\,v = (s(T)\,m_{v,T}(T) + r(T))(v) = s(T)(m_{v,T}(T)v) + r(T)(v) = s(T)(0_{\mathcal{V}}) + r(T)\,v = r(T)\,v$$

pues $m_{v,T}(T)\,v = 0$ (por la propiedad del polinomio T-anulador de v) y $S(T) \in L(\mathcal{V})$ de forma que $S(T)(0_{\mathcal{V}}) = 0_{\mathcal{V}}$. Usando la representación de $r(x)$ de más arriba

$$u = r(T)\,v = (\beta_0\,I + \beta_1\,T + \ldots + \beta_{k-1}\,T^{k-1})\,v = \beta_0\,v + \beta_1\,T\,v + \ldots + \beta_{k-1}\,T^{k-1}\,v$$

que muestra que B_v es un sistema de generadores de $Z(v,T)$.

Para verificar el ítem 2, notemos que $m_{v,T}(x) \mid m_{T|_{Z(v,T)}}(x)$, pues

$$m_{T|_{Z(v,T)}}(T)\,v = m_{T|_{Z(v,T)}}(T)|_{Z(v,T)}\,v = m_{T|_{Z(v,T)}}(T|_{Z(v,T)})\,v = 0\,v = 0\,,$$

donde hemos usado que $v \in Z(v,T)$, la definición de restricción a un subespacio T-invariante y el Lema 8.9. Recíprocamente, si $u \in Z(v,T)$ es arbitrario, existe $q(x) \in K[x]$ tal que $q(T)\,v = u$; en este caso

$$m_{v,T}(T)\,u = m_{v,T}(T)\,q(T)\,v = q(T)\,m_{v,T}(T)\,v = q(T)0 = 0\,,$$

donde hemos usado la relación de conmutación $q(T)\,m_{v,T}(T) = m_{v,T}(T)\,q(T)$ (válida para cualquier par de polinomios en T) y la propiedad de del polinomio T-anulador de v, $m_{v,T}(T)\,v = 0$. Finalmente, recordemos que $q(T) \in L(\mathcal{V})$ de forma que $q(T)0 = 0$.

□

Teorema 9.11. *Sea $\mathcal{V}$ un K-ev, $\dim_K \mathcal{V} = n$ y sea $T \in L(\mathcal{V})$ operador nilpotente. Sea $v \in \mathcal{V}$ tal que $v \neq 0$ y sea $1 \leq k \leq n$ tal que $m_{v,T}(x) = x^k$. Entonces*

1. *$B_v = \{v,\, T\,v,\, \ldots,\, T^{k-1}\,v\}$ es base de $Z(v,T)$ y $\dim_K Z(v,T) = k$.*
2. *Se tiene que*

$$[T|_{Z(v,T)}]_{B_v} = \begin{pmatrix} | & | & \ldots & | & | \\ e_2 & e_3 & \ldots & e_k & 0 \\ | & | & \ldots & | & | \end{pmatrix} = \begin{pmatrix} 0 & 0 & \ldots & 0 & 0 & 0 \\ 1 & 0 & \ldots & \vdots & \vdots & 0 \\ 0 & 1 & \ddots & \vdots & \vdots & 0 \\ \vdots & \vdots & \ddots & 0 & 0 & 0 \\ \vdots & \vdots & & 1 & 0 & 0 \\ 0 & 0 & \ldots & 0 & 1 & 0 \end{pmatrix} \in K^{k\times k}\,. \tag{56}$$

   que es una matriz triangular inferior, con diagonal nula y subdiagonal (que corre debajo de la diagonal principal) formada por entradas iguales a 1.

3. *Si $m_{T|_{Z(v,T)}}(x)$ y $p_{T|_{Z(v,T)}}(x)$ denotan los polinomios minimal y característico de la restricción $T|_{Z(v,T)} \in L(Z(v,T))$ entonces se tienen las igualdades*

$$x^k = m_{v,T}(x) = m_{T|_{Z(v,T)}}(x) = p_{T|_{Z(v,T)}}(x)\,.$$

Demostración. El ítem 1 es consecuencia del ítem 1 del Teorema 9.10.

Para verificar el ítem 2, denotemos por $v_1 = v$, $v_2 = T\,v, \ldots, v_k = T^{k-1}\,v$: notemos que esto equivale a $v_j = T^{j-1}\,v$, $1 \leq j \leq k$. Así,

$$B_v = \{v_1, \ldots, v_k\}$$

y por construcción

$$T|_{Z(v,T)}\, v_j = T\,(T^{j-1}\,v) = T^j\,v = \begin{cases} v_{j+1} & \text{si } 1 \leq j \leq k-1\,; \\ 0 & \text{si } j = k\,. \end{cases}$$

pues $T^k\,v = m_{v,T}(T)\,v = 0$. Es decir,

$$T|_{Z(v,T)}\, v_1 = v_2,\ T|_{Z(v,T)}\, v_2 = v_3, \ldots, T|_{Z(v,T)}\, v_{k-1} = v_k \quad \text{y} \quad T|_{Z(v,T)}\, v_k = 0\,.$$

Entonces, por definición de la matriz de $T|_{Z(v,T)}$

$$[T|_{Z(v,T)}]_{B_v} = \begin{pmatrix} | & | & \ldots & | & | \\ [v_2]_{B_v} & [v_3]_{B_v} & \ldots & [v_k]_{B_v} & [0]_{B_v} \\ | & | & \ldots & | & | \end{pmatrix} = \begin{pmatrix} | & | & \ldots & | & | \\ e_2 & e_3 & \ldots & e_k & 0 \\ | & | & \ldots & | & | \end{pmatrix}$$

Finalmente, verificamos el ítem 3. Por el ítem 2 del Teorema 9.10, sabemos que $m_{v,T}(x) = m_{T|_{Z(v,T)}}(x) = x^k$. Calculamos $p_{T|_{Z(v,T)}}(x)$: recordemos que

$$p_{T|_{Z(v,T)}}(x) = \det(x\,I - [\,T|_{Z(v,T)}\,]_{B_v}) = x^k \tag{57}$$

pues, usando el ítem 2 verificado más arriba, tenemos que

$$x\,I - [\,T|_{Z(v,T)}\,]_{B_v} = \begin{pmatrix} x & 0 & \ldots & 0 & 0 & 0 \\ -1 & x & \ldots & \vdots & \vdots & 0 \\ 0 & -1 & \ddots & \vdots & \vdots & 0 \\ \vdots & \vdots & \ddots & x & 0 & 0 \\ \vdots & \vdots & & -1 & x & 0 \\ 0 & 0 & \ldots & 0 & -1 & x \end{pmatrix} \in K^{k\times k}$$

que es una matriz triangular inferior, cuya diagonal principal tiene por entradas a x. En este caso, el determinante de esta matriz triangular es el producto de las entradas en su diagonal principal, que prueba la identidad de la Eq. (57). □

En la última parte de la desmostración anterior hemos hecho uso del siguiente resultado de determinantes (probado en álgebra I): si $A = (a_{ij})$ es matriz $n \times n$ triangular (inferior o superior) entonces el determinante de A es el producto de las entradas en su diagonal principal, es decir

$$\det(A) = \prod_{j=1}^{n} a_{jj}\,.$$

Para simplificar la notación de los siguientes resultados, vamos a introducir la noción de bloque elemental de Jordan.

Definición 9.12. *Sea K un cuerpo y sean $\lambda \in K$ y $k \in \mathbb{N}$, $k \geq 1$. Definimos el bloque elemental de Jordan asociado a (λ, k), notado $J(\lambda, k)$ como la matriz: $J(\lambda, 1) = \lambda \in K^{1\times 1}$ y si $k \geq 2$,*

$$J(\lambda, k) = \begin{pmatrix} \lambda & 0 & \dots & 0 & 0 & 0 \\ 1 & \lambda & \dots & \vdots & \vdots & 0 \\ 0 & 1 & \ddots & \vdots & \vdots & 0 \\ \vdots & \vdots & \ddots & \lambda & 0 & 0 \\ \vdots & \vdots & & 1 & \lambda & 0 \\ 0 & 0 & \dots & 0 & 1 & \lambda \end{pmatrix} \in K^{k\times k}$$

Así, si $k \geq 2$, $J(\lambda, k)$ es matriz triangular inferior, tal que las entradas en la diagonal principal coinciden con λ, y las entradas en la subdiagonal (inferior) coinciden con 1. El resto de las entradas, son nulas. □

Obs 9.13. Sea K un cuerpo. Con la notación del Teorema 9.11, notemos que la matriz que aparece en la representación de la restricción de T en la Eq. (56) se puede describir como

$$[T|_{Z(v,T)}]_{B_v} = J(0, k) \in K^{k\times k}$$

es decir, como el bloque elemental de Jordan asociado a $0 \in K$ y $k \geq 1$.

Por otro lado si $\lambda \in K$ es arbitrario entonces, por construcción, se verifica:

$$J(\lambda, k) = \lambda\, I_k + J(0, k) \tag{58}$$

donde $I_k \in K^{k\times k}$ denota la matriz identidad (hacer ejemplos en la hoja). △

Corolario 9.14. *Sea $\mathcal{V}$ un K-ev, $\dim_K \mathcal{V} = n$ y sea $T \in L(\mathcal{V})$ operador nilpotente, con polinomio minimal $m_T(x) = x^n$. Entonces, si $v \in \mathcal{V}$ es tal que $T^{n-1}v \neq 0$ entonces $B_v = \{v,\, Tv,\, T^2\, v, \dots,\, T^{n-1}\, v\}$ es base de $\mathcal{V}$. En particular, $[T]_{B_v} = J(0, n)$.*

Demostración. Antes que nada, notemos que $T^{n-1} \neq 0$, de forma que existen vectores v como en el enunciado. Por otro lado, si $T^{n-1}\, v \neq 0$ entonces $v \neq 0$ y luego, el polinomio T-anulador de v verifica $m_{v,T}(x) = x^k$, con $1 \leq k \leq n$. En particular, $T^k\, v = 0$; además, si $s \geq k$ entonces $T^s\, v = 0$: la prueba es por inducción sobre s. El caso base, $s = k$ vale por lo anterior. Si suponemos que vale para un $s - 1 \geq k$ entonces $T^s\, v = T\,(T^{s-1}\, v) = T\,0 = 0$.

El hecho probado en el párrafo anterior indica que $k \nleq n - 1$ (de otra forma $T^{n-1}\, v = 0$, en contra de la hipótesis del enunciado). Así, $n - 1 < k \leq n$ que muestra que $k = n$ (ver Observación 9.9). Entonces hemos probado que $m_{v,T}(x) = x^n$. Si consideramos $Z(v, T) \subseteq \mathcal{V}$ entonces, el teorema anterior muestra que $\dim_K Z(v, T) = n$ y $Z(v, T) = \mathcal{V}$. Nuevamente, el teorema anterior muestra que B_v es base de $\mathcal{V}$ y que $[T]_{B_v} = J(0, n)$.

□

Obs 9.15. Con las notaciones del Corolario anterior, cabe aclarar que hay *muchos* vectores $v \in \mathcal{V}$ tales que $T^{n-1}\, v \neq 0$. Cualquiera de estos vectores permite desarrollar una base de $\mathcal{V}$, descrita en el enunciado del corolario.

Si consideramos el Ejemplo 9.4, entonces cualquier polinomio $p(x) \in \mathcal{V}$ tal que $\mathrm{gr}(p(x)) = 2$ es tal que $D^2\, p(x) \neq 0$ y entonces $\{p(x),\, D\, p(x),\, D^2\, p(x)\}$ es base de $\mathcal{V}$ (ejercicio).

En este sentido, el vector particular v no juega un papel esencial; lo esencial es su existencia, y la representación matricial de T derivada de su existencia. △

En lo que sigue enunciamos el Teorema de la descomposición cíclica en el caso particular de operadores nilpotentes.

Teorema 9.16 (Descomposición cíclica de operadores nilpotentes)**.** *Sea $\mathcal{V}$ un K-ev, $dim_K\mathcal{V} = n$ y sea $T \in L(\mathcal{V})$ operador nilpotente, con polinomio minimal $m_T(x) = x^r$. Entonces existen únicos $p \in \mathbb{N}$, y naturales $k_1 = r \geq k_2 \geq \ldots \geq k_p \geq 1$ tales que: existen vectores $v_1, \ldots, v_p \in \mathcal{V}$ que verifican*

1. *$\mathcal{V} = Z(v_1, T)\dot{+} \ldots \dot{+}Z(v_p, T)$ (suma directa);*
2. *$m_{v_j,T}(x) = m_{T|_{Z(v_j,T)}}(x) = x^{k_j}$, $1 \leq j \leq p$.*

En particular, si consideramos la base $B_j = \{v_j, T\,v_j, \ldots, T^{k_j-1}\,v_j\}$, para $1 \leq j \leq p$, entonces $B = B_1 \cup \ldots \cup B_p$ es base de $\mathcal{V}$ y

$$[T]_B = \begin{pmatrix} J(0,k_1) & 0 & 0 & \ldots & 0 \\ 0 & J(0,k_2) & 0 & \ldots & 0 \\ 0 & 0 & \ddots & \ldots & 0 \\ \vdots & \vdots & 0 & \ddots & 0 \\ 0 & 0 & 0 & \ldots & J(0,k_p) \end{pmatrix} \in K^{n\times n} . \tag{59}$$

Más aún, $\sum_{j=1}^{p} k_j = n$ y el polinomio característico $p_T(x) = x^n$.

Demostración. En estas notas no vamos a desarrollar la prueba en detalle. En particular, no vamos a desarrollar el argumento que muestra la existencia de los vectores $v_1, \ldots, v_p$.

Sin embargo, si *admitimos la existencia* de los vectores $v_1, \ldots, v_p$ que verifican el ítem 1 del enunciado (que es la parte crucial de la prueba del teorema) y del ítem 2, entonces la última afirmación del teorema se puede justificar como sigue.

Asumiendo los vectores $v_1, \ldots, v_p$ verifican los ítem 1 y 2, entonces, recordemos que cada espacio $\mathcal{W}_j = Z(v_j, T)$ es T-invariante, por la Proposición 9.7. Por el Teorema 9.11, B_j (definido en el enunciado) resulta una base de $\mathcal{W}_j$, $1 \leq j \leq p$. Como los subespacios están en suma directa y suman todo $\mathcal{V}$, entonces $B = B_1 \cup \ldots \cup B_p$ es base de $\mathcal{V}$. En particular,

$$n = \dim_K \mathcal{V} = \sum_{j=1}^{p} \dim_K Z(v_j, T) = \sum_{j=1}^{p} k_j .$$

Además, por la Observación 6.32 y el Teorema 9.11, concluimos la validez de la representación matricial $[T]_B$ (ver también la Observación 9.13). Notemos que $[T]_B$ es una matriz triangular inferior con diagonal principal nula (sugerencia: hacer algunos ejemplos en la hoja para ver el aspecto de este tipo de matrices, eligiendo cantidad de bloques y tamaños). Entonces $x\,I - [T]_B$ es matriz triangular inferior con diagonal principal con entradas todas iguales a x: esto muestra que

$$p_T(x) = \det(x\,I - [T]_B) = x^n .$$

□

Finalizamos esta sección con un comentario sobre la unicidad de los parámetros del Teorema de la descomposición cíclica que juega un papel importante en la unicidad de la forma de Jordan.

Obs 9.17. Sea $\mathcal{V}$ un K-ev, $dim_K\mathcal{V} = n$ y sea $T \in L(\mathcal{V})$ operador nilpotente, con polinomio minimal $m_T(x) = x^r$. Sean $p \in \mathbb{N}$, y $k_1 = r \geq k_2 \geq \ldots \geq k_p \geq 1$ únicos, como en el Teorema 9.16, para los cuales vale la Eq. (59).

Supongamos que $B' = \{w_1, \ldots, w_n\}$ es base de $\mathcal{V}$ tal que existen $q \in \mathbb{N}$ y $h_1 \geq \ldots \geq h_q \geq 1$ de forma que

$$[T]_{B'} = \begin{pmatrix} J(0,h_1) & 0 & 0 & \ldots & 0 \\ 0 & J(0,h_2) & 0 & \ldots & 0 \\ 0 & 0 & \ddots & \ldots & 0 \\ \vdots & \vdots & 0 & \ddots & 0 \\ 0 & 0 & 0 & \ldots & J(0,h_q) \end{pmatrix} \in K^{n\times n}\,. \tag{60}$$

Entonces

$$q = p \quad \text{y} \quad k_j = h_j \ , \ \ 1 \leq j \leq p\,.$$

Veamos un bosquejo (informal) de la prueba (el lector puede omitir este argumento en una primer lectura): en lo que sigue $\{e_1, \ldots, e_n\}$ denota la base canónica de K^n.

Notemos que la j-ésima columna de $[T]_{B'}$ es e_{j+1} para $1 \leq j \leq h_1 - 1$ y la columna h_1 es nula (esto se ve mirando con cuidado la matriz $[T]_{B'}$, y escribiendo el primer bloque $J(0,h_1)$; como sugerencia, fijar los valores de q y h_j's y hacer un bosquejo de la matriz en una hoja - ej: $q = 3$, $h_1 = 3$, $h_2 = 2$, $h_3 = 2$).

Renombremos $w_1 = z_1$; Como $[Tw_1]_{B'} = [T]_{B'}\, e_1 = e_2$ (pues $[T]_{B'}\, e_1$ coincide con la primer columna de $[T]_{B'}$, que es e_2) entonces $Tw_1 = w_2$! De forma similar, $T^2 w_1 = Tw_2 = w_3$ (porque $[T]_{B'}\, e_2 = e_3$) y en general: $T^j w_1 = w_{j+1} \neq 0$, $1 \leq j \leq h_1 - 1$, y $T^{h_1} w_1 = 0$. Esto muestra (usando el nombre $z_1 = w_1$) que $m_{z_1,T}(x) = x^{h_1} \in K[x]$ y que el subespacio generado por el conjunto l.i.

$$B_1 = \{w_1, \ldots, w_{h_1}\} = \{z_1,\, T\, z_1, \ldots, T^{h_1-1}\, z_1\}$$

es $Z(z_1,T)$, por el Teorema 9.11. En particular, B_1 es base de $Z(z_1,T)$.

Si consideramos el vector siguiente al grupo anterior y lo renombramos $z_2 = w_{h_1+1}$, podemos hacer el mismo análisis basado en las columnas de la matriz $[T]_{B'}$: de hecho, la columna $h_1 + j$ de $[T]_{B'}$ coincide con el vector e_{h_1+j+1}, $1 \leq j \leq h_2 - 1$ y coincide con el vector nulo si $j = h_2$. Entonces, concluimos (de forma similar al párrafo anterior) que $T^j\, w_{h_1+1} = w_{h_1+1+j} \neq 0$, $1 \leq j \leq h_2 - 1$ y $T^{h_2}\, w_{h_1+1} = 0$; Como antes, lo anterior muestra que $m_{z_2,T}(x) = x^{h_2}$ y que el subespacio generado por el conjunto l.i. $B_2 = \{w_{h_1+1}, \ldots, w_{h_1+h_2}\}$ coincide con $Z(z_2,T)$ (recordemos que $z_2 = w_{h_1+1}$), por el Teorema 9.11. En particular, B_2 es base de $Z(z_2,T)$.

Si continuamos de esta forma nuestro análisis (que se divide en q pasos, un paso por cada bloque), vemos que podemos encontrar vectores $z_1, \ldots, z_q$ en $\mathcal{V}$ tales que $m_{z_j,T}(x) = x^{h_j}$ y tales que podemos construir bases B_j de $Z(z_j,T)$, $1 \leq j \leq q$, tales que $B_1, \ldots, B_q$ es una partición de B' (es una familia de subconjuntos no vacíos de B', dos a dos disjuntos y cuya unión es B'). En particular,

$$\mathcal{V} = Z(z_1,T) \dot{+} \ldots \dot{+} Z(z_q,T)$$

donde la suma es directa y suma todo el espacio porque si yuxtaponemos (pegamos) las bases $B_1, \ldots, B_q$ obtenemos la base B' de $\mathcal{V}$ (aquí estamos usando una caracterización de subespacios independientes vista en notas anteriores).

Entonces, la unicidad garantizada por el Teorema 9.16 muestra que $q = p$ y $k_j = h_j$, $1 \leq j \leq p$, como queríamos verificar. $\triangle$

Obs 9.18. Sea $\mathcal{V}$ un K-ev, $dim_K \mathcal{V} = n$ y sea $T \in L(\mathcal{V})$ operador nilpotente. Si el polinomio minimal $m_T(x) = x^n$ entonces, con las notaciones del Teorema 9.16 de la descomposición cíclica, en este caso tenemos que $p = 1$ y $k_1 = n$. Más aún, si $v \in \mathcal{V}$ es tal que $T^{n-1}\, v \neq 0$ entonces, con las notaciones del Teorema 9.16, podemos tomar $v_1 = v$: en este caso, $\mathcal{V} = Z(v_1,T)$ y si $B = \{v_1,\, T\, v_1, \ldots, T^{n-1}\, v_1\}$ entonces B es base de $\mathcal{V}$ tal que $[T]_B = J(0,n)$; notemos que estas afirmaciones son consecuencia del Corolario 9.14. $\triangle$

9.2 Descomposición cíclica de operadores nilpotentes: + detalles

Consideremos la notación del Teorema 9.16 de la descomposición cíclica para operadores nilpotentes. Cabe remarcar que la construcción de los vectores $v_1, \ldots, v_p$ como en el enunciado implica hacer uso de ciertas nociones técnicas que no hemos introducido en estas notas.

Sin embargo, es posible hacer un estudio sobre los parámetros $p \in \mathbb{N}$ (que describe la cantidad de bloques elementales de Jordan que aparecen en la representación matricial de $[T]_B$) y $r = k_1 \geq k_2 \ldots \geq k_p \geq 1$ (que describen los tamaños de los bloques elementales de Jordan aparecen en la representación matricial de $[T]_B$) del enunciado, que son únicos, según afirma el teorema (ver también la Observación 9.17). En lo que sigue, hacemos un análisis relacionado con estos parámetros. Este análisis va a tener aplicaciones de interés en la determinación de los parámetros correspondientes a la forma de Jordan para ciertos operadores T más generales.

Comenzamos con el siguiente resultado que vale para operadores arbitrarios.

Proposición 9.19. *Sea $\mathcal{V}$ un K-ev, $\dim_K \mathcal{V} = n$ y sea $T \in L(\mathcal{V})$ arbitrario. Sean $\mathcal{W}_1 \ldots, \mathcal{W}_p$ subespacios T-invariantes, tales que*

$$\mathcal{V} = \mathcal{W}_1 \dot{+} \ldots \dot{+} \mathcal{W}_p .$$

Si $T_j = T|_{\mathcal{W}_j} \in L(\mathcal{W}_j)$, $1 \leq j \leq p$. Entonces

$$N(T) = N(T_1) \dot{+} \ldots \dot{+} N(T_p) \quad \text{ y } \quad Nul(T) = \sum_{j=1}^{p} Nul(T_j) .$$

Demostración. Notemos que como $\mathrm{N}(T_j) \subseteq \mathcal{W}_j$, $1 \leq j \leq p$, y los subespacios $\mathcal{W}_1, \ldots, \mathcal{W}_p$ son independientes, entonces $\mathrm{N}(T_1), \ldots, \mathrm{N}(T_p)$ resultan ser independientes (ejercicio, usar la definición de independencia de subespacios).

Sea $v \in N(T)$; entonces existen únicos $w_j \in \mathcal{W}_j$, $1 \leq j \leq p$, tales que $v = w_1 + \ldots + w_p$. En particular,

$$0 = T\,v = T(w_1 + \ldots + w_p) = T(w_1) + \ldots + T(w_p) .$$

Como cada $T(w_j) \in \mathcal{W}_j$, pues $\mathcal{W}_j$ es T-invariante, $1 \leq j \leq p$, concluimos que $T(w_j) = 0$, $1 \leq j \leq p$ (por independencia de los subespacios). Así, $0 = T(w_j) = T_j(w_j)$ lo que muestra que $w_j \in \mathrm{N}(T_j)$, $1 \leq j \leq p$ y finalmente, que

$$v = w_1 + \ldots + w_p \in \mathrm{N}(T_1) \dot{+} \ldots \dot{+} \mathrm{N}(T_p) .$$

Recíprocamente, si $v = w_1 + \ldots + w_p \in \mathrm{N}(T_1) \dot{+} \ldots \dot{+} \mathrm{N}(T_p)$ entonces

$$T\,v = T(w_1 + \ldots + w_p) = T(w_1) + \ldots + T(w_p) = T_1(w_1) + \ldots + T_p(w_p) = 0 + \ldots + 0 = 0$$

lo que muestra que $v \in N(T)$. Los párrafos anteriores muestran la primer identidad propuesta en el enunciado del resultado.

La segunda identidad (de dimensiones) es una consecuencia de la primer identidad (de subespacios) y de la independencia de los subespacios $\mathrm{N}(T_1), \ldots, \mathrm{N}(T_p)$ (junto con un resultado sobre pegado de bases de subespacios independientes que ya hemos probado!). $\square$

El próximo resultado establece las nulidades de las potencias de operadores nilpotentes especiales (con un solo bloque elemental de Jordan).

Proposición 9.20. *Sea $\mathcal{V}$ un K-ev, $\dim_K \mathcal{V} = n$ y sea $T \in L(\mathcal{V})$ operador nilpotente, con polinomio minimal $m_T(x) = x^n$. Si $m \geq 1$ entonces se verifica:*

$$Nul\ (T^m) = \begin{cases} m & si\ \ 1 \leq m \leq n\,; \\ & \\ n & si\ m \geq n+1\,. \end{cases} \tag{61}$$

Es decir, Nul $(T^m) = \min\{n, m\}$, para $m \geq 1$.

Demostración. Recordemos que $T^n = m_T(T) = 0$. Por otro lado, si $m \geq n$, entonces $T^m = 0$ (la prueba es por inducción en $m \geq n$ y se deja como ejercicio). Estos hechos muestran entonces que

$$\mathrm{Nul}(T^m) = \dim_K \mathrm{N}(T^m) = \dim_K \mathcal{V} = n\ ,\quad m \geq n\,.$$

Para estudiar los casos restantes, $1 \leq m \leq n-1$, consideramos $v \in \mathcal{V}$ tal que $T^{n-1}\, v \neq 0$. Construimos la base $B = \{v,\, T\, v, \ldots,\, T^{n-1}\, v\}$ de $\mathcal{V}$ como en el Corolario 9.14: de hecho, vamos a indicar a los elmentos de B como $B = \{v_1, \ldots, v_n\}$, de forma que

$$v_j = T^{j-1}\, v\ ,\quad 1 \leq j \leq n\,. \tag{62}$$

Para fijar ideas, consideremos el caso $m = 1$ con algún detalle. En este caso,

$$T(v_1) = T(v) = v_2\ ,\ T(v_2) = T(T\,v) = T^2\,v = v_3\ ,\ \ldots,\ T(v_{n-1}) = T(T^{n-2}v) = T^{n-1}\,v = v_n$$
$$\text{y}\quad T(v_n) = T(T^{n-1}\,v) = T^n\,v = 0\,v = 0\,.$$

Los cálculos anteriores muestran que

$$\{v_2, \ldots, v_n\} \subset \mathrm{Im}(T) \implies n-1 \leq \dim_K \mathrm{Im}(T)$$
$$\{v_n\} \subset \mathrm{N}(T) \implies 1 \leq \dim_K \mathrm{N}(T)$$

pues $\{v_2, \ldots, v_n\}$ y $\{v_n\}$ son conjuntos l.i. (porque son subconjuntos de una base). Como hemos probado que

$$n = \dim_K \mathrm{Im}(T) + \dim_K \mathrm{N}(T) \implies \dim_K \mathrm{Im}(T) = n-1 \quad \text{y} \quad \dim_K \mathrm{N}(T) = 1$$

Para ver el caso general, fijemos $1 \leq m \leq n-1$ y notemos que : si $1 \leq j \leq n$,

$$T^m\, v_j = \begin{cases} v_{j+m} & \text{si}\ \ 1 \leq j \leq n-m\,; \\ & \\ 0 & \text{si}\ \ n-m < j \leq n\,. \end{cases} \tag{63}$$

Justificamos los casos anteriores: notemos que

$$T^m\, v_j = T^m\,(T^{j-1}\,v) = T^{m+j-1}\,v$$

Si $1 \leq m+j-1 \leq n-1$ (ó equivalentemente $1 \leq j \leq n-m$) entonces $T^m\, v_j = v_{j+m}$ (ver la Eq. (62) con cuidado; en particular donde varía j allí).

Si $m+j-1 \geq n$ (ó equivalentemente $n-m < j \leq n$) entonces $T^{m+j-1} = 0 \in L(\mathcal{V})$ y $T^m\, v_j = T^{m+j+1}\, v = 0$.

Los dos casos anteriores justifican la Eq. (63). Estos hechos implican que

$$\{v_{1+m}, \ldots, v_n\} \subset \mathrm{Im}(T^m) \implies n-m \leq \dim_K \mathrm{Im}(T^m)$$

$$\{v_{n-m+1},\ldots,v_n\} \subset \mathrm{N}(T^m) \implies m \le \dim_K \mathrm{N}(T^m)$$

pues $\{v_{1+m},\ldots,v_n\}$ es conjunto l.i. con $n-m$ elementos y $\{v_{n-m+1},\ldots,v_n\}$ es conjunto l.i. con m elementos (Ojo: contar la cantidad de elementos con cuidado!). Como antes, usamos las desigualdades de dimensiones anteriores y el hecho de que

$$n = \dim_K \mathrm{Im}(T^m) + \dim_K \mathrm{N}(T^m) \implies \dim_K \mathrm{Im}(T^m) = n-m \quad \text{y} \quad \dim_K \mathrm{N}(T^m) = m\,.$$

$\square$

Ya estamos en condiciones de combinar los resultados de las proposiciones previas para estudiar los parámetros de la descomposición cíclica de operadores nilpotentes. Si bien el enunciado no parece decir mucho, vamos a ver que en realidad este resultado nos da información muy útil en observaciones posteriores.

Teorema 9.21. *Sea $\mathcal{V}$ un K-ev, $\dim_K\mathcal{V} = n$ y sea $T \in L(\mathcal{V})$ operador nilpotente, con polinomio minimal $m_T(x) = x^r$. Sean $p \in \mathbb{N}$, y $k_1 = r \ge k_2 \ge \ldots \ge k_p \ge 1$ únicos, como en el Teorema 9.16, para los cuales vale la Eq.* (59). *Entonces, si $m \ge 1$ tenemos que*

$$\mathit{Nul}\ (T^m) = \sum_{j=1}^{p} \min\{m,\, k_j\}\,.$$

Demostración. Con la notación del Teorema 9.16, sean $v_1,\ldots,v_p \in \mathcal{V}$ tales que verifican:

1. $\mathcal{V} = Z(v_1,T)\dot{+}\ldots\dot{+}Z(v_p,T)$ (suma directa);
2. $m_{v_j,T}(x) = m_{T|_{Z(v_j,T)}}(x) = x^{k_j}$, $1 \le j \le p$.

El ítem 2 de más arriba, junto con el Teorema 9.11, indican que $\dim_K Z(v_j,T) = k_j$, $1 \le j \le p$.

Definimos $T_j = T|_{Z(v_j,T)} \in L(Z(v_j,T))$, $1 \le j \le p$. Sea $1 \le m$ y consideremos $q(x) = x^m$. Entonces $q(T) = T^m$ y por el Lema 8.9 tenemos que

$$T^m\,|_{Z(v_j,T)} = q(T)|_{Z(v_j,T)} = q(T|_{Z(v_j,T)}) = q(T_j) = (T_j)^m\,. \tag{64}$$

Si aplicamos la Proposición 9.19 al operador $T^m \in L(\mathcal{V})$, usando la descomposición del ítem 1 (de arriba) y la Eq. (64), concluimos que:

$$\mathrm{Nul}\ (T^m) = \sum_{j=1}^{p} \mathrm{Nul}\ (T^m\,|_{Z(v_j,T)}) = \sum_{j=1}^{p} \mathrm{Nul}\ ((T_j)^m)\,. \tag{65}$$

Como $T_j \in L(Z(v_j,T))$ es nilpotente, con minimal $m_{T_j}(x) = x^{k_j}$ y $\dim_K Z(v_j,T) = k_j$, podemos aplicar la Proposición 9.20 a T_j y concluir

$$\mathrm{Nul}\ ((T_j)^m) = \min\{m,\, k_j\}\,. \tag{66}$$

Reemplazando la identidad de la Eq. (66) en la Eq. (65) concluimos la identidad del enunciado. $\square$

Obs 9.22. Consideremos la notación del Teorema 9.21; una consecuencia de este resultado es que los parámetros $p \in \mathbb{N}$, y $k_1 = r \ge k_2 \ge \ldots k_p \ge 1$ únicos, como en el Teorema 9.16, están codificados (se pueden recuperar) a partir de las dimensiones

$$\mathrm{Nul}\ (T^m) \quad \text{para} \quad 1 \le m \le r\ (\le n)\,.$$

Si bien no vamos a probar este hecho formalmente, vamos a ver cómo podemos utilizar los valores dados por las nulidades de arriba (junto con alguna información adicional) para determinar los parámetros $p \in \mathbb{N}$, y $k_1 = r \ge k_2 \ge \ldots k_p \ge 1$. $\triangle$

Ejemplo 9.23. Sea $\mathcal{V}$ un K-ev, $\dim_K \mathcal{V} = 6$ y sea $T \in L(\mathcal{V})$ operador nilpotente, con polinomio minimal $m_T(x) = x^3$. Con esta información no podemos determinar de forma exacta la estructura de T es decir, los parámetros p y $k_1 = 3 \geq k_2 \geq \ldots \geq k_p \geq 1$ (únicos) del Teorema 9.16.

Sin embargo, el Teorema 9.16 garantiza que $k_1 = 3$ (= al grado del polinomio minimal de T en este caso), $3 \geq k_2 \geq \ldots \geq k_p \geq 1$ y que $6 = \sum_{j=1}^{p} k_j$. Estos hechos reducen las posibilidades de los parámetros a los siguientes casos:

1. $p = 2$, $k_1 = 3$ y $k_2 = 3$;
2. $p = 3$, $k_1 = 3$, $k_2 = 2$ y $k_3 = 1$;
3. $p = 4$, $k_1 = 3$, $k_2 = 1$, $k_3 = 1$ y $k_4 = 1$.

En cada uno de los casos anteriores, existe B base de $\mathcal{V}$ tal que

$$\text{Si vale 1:} \quad [T]_B = \begin{pmatrix} J(0,3) & 0 \\ 0 & J(0,3) \end{pmatrix} \in K^{6\times 6} .$$

$$\text{Si vale 2:} \quad [T]_B = \begin{pmatrix} J(0,3) & 0 & 0 \\ 0 & J(0,2) & 0 \\ 0 & 0 & J(0,1) \end{pmatrix} \in K^{6\times 6} .$$

$$\text{Si vale 3:} \quad [T]_B = \begin{pmatrix} J(0,3) & 0 & 0 & 0 \\ 0 & J(0,1) & 0 & 0 \\ 0 & 0 & J(0,1) & 0 \\ 0 & 0 & 0 & J(0,1) \end{pmatrix} \in K^{6\times 6} .$$

Ejercicio: escribir cada una de las matrices de arriba como matriz 6×6, entrada por entrada y observar las diferencias entre ellas (que estará en cómo aparecen los 1's en la subdiagonal).

Notemos que, en general (tomando $m = 1$ en el Teorema 9.21)

$$\text{Nul }(T) = \sum_{j=1}^{p} \min\{1,\, k_j\} = \sum_{j=1}^{p} 1 = p$$

es decir, Nul (T) coincide con la cantidad de bloques elementales de Jordan de $[T]_B$.

En este caso particular, la cantidad de bloques elementales de Jordan nos permite discriminar los tres posibles casos de estructura de T. Así, si nos dan el dato

$$\text{Nul }(T) = \dim_K \mathrm{N}(T) = p \quad \text{ con } \quad 2 \leq p \leq 4$$

entonces podemos determinar de forma exacta la estructura de T (en términos de los parámetros p y k_j's).

Si el dato que nos dan es $\text{Nul}(T^2) = d$ entonces (tomando $m = 2$ en el Teorema 9.21)

1. Si vale 1: $\text{Nul}(T^2) = \min\{2,3\} + \min\{2,3\} = 2 + 2 = 4$;
2. Si vale 2: $\text{Nul}(T^2) = \min\{2,3\} + \min\{2,2\} + \min\{2,1\} = 2 + 2 + 1 = 5$;
3. Si vale 3: $\text{Nul}(T^2) = \min\{2,3\} + \min\{2,1\} + \min\{2,1\} + \min\{2,1\} = 2 + 1 + 1 + 1 = 5$.

Si por ejemplo nos indican que $\text{Nul}(T^2) = 5$ entonces nos quedan dos casos posibles (casos 2 y 3) y todavía necesitamos más información para determinar de forma exacta la estructura de T. $\triangle$

9.3 La forma de Jordan

Vamos a aplicar todos los resultados vistos en las secciones previas para poder determinar la existencia y características de la forma de Jordan de cierta clase de operadores lineales. Comenzamos con el caso más sencillo y luego derivamos el caso general. Pero primero, planteamos el siguiente ejercicio en forma de lema

Lema 9.24. *Sea $\mathcal{V}$ un K-ev y sea $T \in L(\mathcal{V})$. Dados $\lambda \in K$ y $v \in \mathcal{V}$ se verifica que*

$$Z(v, T) = Z(v, T - \lambda\, I) \subseteq \mathcal{V}\,.$$

Demostración. Ejercicio. sugerencia: probar una doble inclusión usando la definición de estos espacios; para ello, usar el ítem 3 de la Proposición 9.7 (que deben probar primero! porque lo dejamos como ejercicio). □

Proposición 9.25. *Sea $\mathcal{V}$ un K-ev, $dim_K\mathcal{V} = n$ y sea $T \in L(\mathcal{V})$ con polinomio minimal $m_T(x) = (x - \lambda)^r$, $1 \leq r$. Entonces existen únicos $p \in \mathbb{N}$, y naturales $k_1 = r \geq k_2 \geq \ldots \geq k_p \geq 1$ tales que: existe una base B de $\mathcal{V}$ tal que*

$$[T]_B = \begin{pmatrix} J(\lambda, k_1) & 0 & 0 & \ldots & 0 \\ 0 & J(\lambda, k_2) & 0 & \ldots & 0 \\ 0 & 0 & \ddots & \ldots & 0 \\ \vdots & \vdots & 0 & \ddots & 0 \\ 0 & 0 & 0 & \ldots & J(\lambda, k_p) \end{pmatrix} \in K^{n\times n}\,. \tag{67}$$

Más aún,

1. *$\sum_{j=1}^{p} k_j = n$ y el polinomio característico $p_T(x) = (x - \lambda)^n$.*
2. *Si $m \geq 1$ tenemos que*
$$Nul\ ((T - \lambda\, I)^m) = \sum_{j=1}^{p} \min\{m,\, k_j\}\,.$$
En particular, la cantidad de bloques elementales de Jordan es $p = Nul(T - \lambda)$.

Demostración. Para probar este resultado vamos a seguir la estrategia propuesta en la Observación 9.2 al comienzo de la Sección 9.

Sea $N = T - \lambda\, I \in L(\mathcal{V})$; entonces

$$T = \lambda\, I + N \quad \text{ y } \quad N^r = (T - \lambda\, I)^r = m_T(T) = 0$$

Lo anterior muestra que N es nilpotente; además, $m_N(x) = x^r$ pues $N^{r-1} = (T - \lambda\, I)^{r-1} \neq 0$, porque $\mathrm{gr}((x - \lambda)^{r-1}) < \mathrm{gr}(m_T(x))$ con lo que $(x - \lambda)^{r-1}$ no anula a T.

Así, podemos aplicar los Teoremas 9.16 y 9.21 al operador nilpotente $N \in L(\mathcal{V})$. En particular, existen únicos $p \in N$ y $k_1 = r \geq k_2 \geq \ldots \geq k_p \geq 1$, tales que existe una base B de $\mathcal{V}$ de forma que

$$[N]_B = \begin{pmatrix} J(0, k_1) & 0 & 0 & \ldots & 0 \\ 0 & J(0, k_2) & 0 & \ldots & 0 \\ 0 & 0 & \ddots & \ldots & 0 \\ \vdots & \vdots & 0 & \ddots & 0 \\ 0 & 0 & 0 & \ldots & J(0, k_p) \end{pmatrix} \in K^{n\times n}\,.$$

Entonces

$$[T]_B = [\lambda I + N]_B = \lambda I_n + [N]_B =$$

$$= \begin{pmatrix} \lambda I_{k_1} & 0 & 0 & \dots & 0 \\ 0 & \lambda I_{k_2} & 0 & \dots & 0 \\ 0 & 0 & \ddots & \dots & 0 \\ \vdots & \vdots & 0 & \ddots & 0 \\ 0 & 0 & 0 & \dots & \lambda I_{k_p} \end{pmatrix} + \begin{pmatrix} J(0,k_1) & 0 & 0 & \dots & 0 \\ 0 & J(0,k_2) & 0 & \dots & 0 \\ 0 & 0 & \ddots & \dots & 0 \\ \vdots & \vdots & 0 & \ddots & 0 \\ 0 & 0 & 0 & \dots & J(0,k_p) \end{pmatrix}.$$

donde hemos escrito a λI_n como una matriz (diagonal) por bloques, por conveniencia! La expresión anterior para $[T]_B$ muestra la validez de la Eq. (67), ya que la suma de matrices de bloques se realiza *bloque a bloque* y además, en cada bloque diagonal se tiene la suma $\lambda I_{k_j} + J(0,k_j) = J(\lambda, k_j)$ (ver Eq. (58) de la Observación 9.13), $1 \le j \le p$.

En particular se verifica que $\sum_{j=1}^{p} k_j = n$. Además $[T]_B$ resulta, por construcción, una matriz triangular inferior, cuyas entradas en la diagonal principal coinciden con λ. Entonces $x\, I_n - [T]_B \in K^{n\times n}$ es una matriz triangular inferior, cuyas entradas en la diagonal principal coinciden con $(x-\lambda)$. Como el determinante de una matriz triangular (inferior) es el producto de las entradas de su diagonal principal vemos que el polinomio característico

$$p_T(x) = \det(x\, I_n - [T]_B) = (x-\lambda)^n\,.$$

Por otro lado, el ítem 2 del enunciado es una consecuencia inmediata del Teorema 9.21 aplicado a $N = T - \lambda I \in L(\mathcal{V})$.

Finalmente, notemos que la unicidad de $p \in \mathbb{N}$ y $k_1 = r \ge k_2 \ge \ldots \ge k_p \ge 1$ es una consecuencia de la unicidad de los parámetros en el Teorema de la descomposición cíclica 9.16. Hacemos un bosquejo de la prueba, y dejamos los detalles al lector: si suponemos que $q \in N$ y $h_1 = r \ge h_2 \ge \ldots \ge h_q \ge 1$ son tales que existe una base $B' = \{w_1, \ldots, w_n\}$ de $\mathcal{V}$ tal que

$$[T]_{B'} = \begin{pmatrix} J(\lambda,h_1) & 0 & 0 & \dots & 0 \\ 0 & J(\lambda,h_2) & 0 & \dots & 0 \\ 0 & 0 & \ddots & \dots & 0 \\ \vdots & \vdots & 0 & \ddots & 0 \\ 0 & 0 & 0 & \dots & J(\lambda,h_q) \end{pmatrix} \in K^{n\times n}\,,$$

entonces

$$[N]_{B'} = [T]_{B'} - \lambda I_n = \begin{pmatrix} J(0,h_1) & 0 & 0 & \dots & 0 \\ 0 & J(0,h_2) & 0 & \dots & 0 \\ 0 & 0 & \ddots & \dots & 0 \\ \vdots & \vdots & 0 & \ddots & 0 \\ 0 & 0 & 0 & \dots & J(0,h_q) \end{pmatrix} \in K^{n\times n}\,.$$

Así, la Observación 9.17 muestra que $p = q$ y $k_j = h_j$, $1 \le j \le p$. □

Obs 9.26. Consideremos la notación del teorema anterior: notemos que la descomposición $T = \lambda I + N$ representa a T como la suma de un operador diagonalizable (en cualquier base! pues es un múltiplo de la identidad) y del operador nilpotente $N = T - \lambda I$. La información más compleja sobre T se encuentra codificada en N. En particular, los parámetros de la descomposición de Jordan $p \in \mathbb{N}$ y $k_1 = r \ge k_2 \ge \ldots \ge k_p \ge 1$ coinciden con los parámetros de la descomposición del Teorema 9.16 para el operador nilpotente N. Esto justifica el análisis detallado que hemos hecho del caso nilpotente, incluyendo el análisis de los parámetros de la descomposición cíclica en la sección 9.2. △

Obs 9.27. Sea $\mathcal{V}$ un K-ev, $dim_K\mathcal{V} = n$ y sea $T \in L(\mathcal{V})$ operador nilpotente. Si el polinomio minimal $m_T(x) = (x-\lambda)^n$ para $\lambda \in K$. Entonces, con las notaciones de la Proposición 9.25 en este caso tenemos que $p = 1$ y $k_1 = n$. Más aún, si definimos $N = T - \lambda\, I \in L(\mathcal{V})$ entonces N es operador nilpotente, con $m_N(x) = x^n$. Así, si $v \in \mathcal{V}$ es tal que $N^{n-1}\, v \neq 0$ (es decir, $(T-\lambda\, I)^{n-1}\, v \neq 0$) y si definimos $B = \{v,\, N\, v, \ldots, N^{n-1}\, v\}$, entonces B es base de $\mathcal{V}$ tal que

$$[N]_B = J(0,n) \implies [T]_B = [\lambda\, I + N]_B = \lambda\, I + J(0,n) = J(\lambda, n)\,.$$

Estas afirmaciones son consecuencia de la Observación 9.18 y la Eq. (58). $\triangle$

El siguiente resultado considera la existencia y unicidad de la forma de Jordan de operadores más generales. Si bien el enunciado es extenso, la información que aparece en él ha sido (esencialmente) adelantada y probada en los resultados previos.

Teorema 9.28. *Sea $\mathcal{V}$ un K-ev, $dim_K\mathcal{V} = n$ y sea $T \in L(\mathcal{V})$ con polinomio minimal*

$$m_T(x) = \prod_{j=1}^{q} (x-\lambda_j)^{r_j} \in K[x]\,.$$

donde $\lambda_1, \ldots, \lambda_q \in K$ son distintos. Entonces

1. *Si definimos $W_j = N((T-\lambda_j\, I)^{r_j})$ entonces $\mathcal{W}_j$ es T-invariante, $1 \leq j \leq q$, y*

$$\mathcal{V} = \mathcal{W}_1 \dot{+} \ldots \dot{+} \mathcal{W}_q\,.$$

2. *Para cada $1 \leq j \leq q$, existen únicos $p_j \in \mathbb{N}$, y naturales $k_{1,j} = r_j \geq k_{2,j} \geq \ldots \geq k_{p_j,j} \geq 1$ tales que: existe una base B_j de $\mathcal{W}_j$ para la cual*

$$[T|_{\mathcal{W}_j}]_{B_j} = \begin{pmatrix} J(\lambda_j, k_{1,j}) & 0 & 0 & \ldots & 0 \\ 0 & J(\lambda_j, k_{2,j}) & 0 & \ldots & 0 \\ 0 & 0 & \ddots & \ldots & 0 \\ \vdots & \vdots & 0 & \ddots & 0 \\ 0 & 0 & 0 & \ldots & J(\lambda_j, k_{p_j,j}) \end{pmatrix} \in K^{d_j \times d_j} \qquad (68)$$

   donde $d_j = \dim_K \mathcal{W}_j$. En este caso, la suma de los tamaños de los bloques elementales satisface:

$$\sum_{i=1}^{p_j} k_{i,j} = d_j\,.$$

3. *Para cada $m \geq 1$ se tiene*

$$\dim N((T-\lambda_j\, I)^m) = \dim N((T|_{\mathcal{W}_j} - \lambda_j\, I_{\mathcal{W}_j})^m) = \sum_{i=1}^{p_j} \min\{m\,,\, k_{i,j}\}\,.$$

En particular, si $B = B_1 \cup \ldots \cup B_q$ entonces B es base de $\mathcal{V}$ y tenemos que

$$[T]_B = \begin{pmatrix} [T|_{\mathcal{W}_1}]_{B_1} & 0 & 0 & \ldots & 0 \\ 0 & [T|_{\mathcal{W}_2}]_{B_2} & 0 & \ldots & 0 \\ 0 & 0 & \ddots & \ldots & 0 \\ \vdots & \vdots & 0 & \ddots & 0 \\ 0 & 0 & 0 & \ldots & [T|_{\mathcal{W}_q}]_{B_q} \end{pmatrix} \in K^{n\times n} \qquad (69)$$

donde cada bloque diagonal $[T|_{\mathcal{W}_j}]_{B_j}$ *está dado por la Eq.* (68)*. En este caso, el polinomio característico de* T *está dado por*

$$p_T(x) = \prod_{j=1}^{q} (x - \lambda_j)^{d_j} \in K[x] \,. \tag{70}$$

Demostración. La estrategia para probar este resultado es utilizar el argumento de descomposición de T dado en la Observación 9.1 al comienzo de la Sección 9 y luego la Proposición 9.25 en cada bloque de esta descomposición.

Como primer paso, aplicamos el Teorema 8.11 (de la descomposición prima) al operador T. En este caso, el Teorema 8.11 garantiza que si definimos $\mathcal{W}_j = \mathrm{N}((T - \lambda_j)^{r_j})$, $1 \leq j \leq q$, entonces vale el ítem 1 del enunciado (ver los comentarios en la Observación 9.1 sobre esta aplicación del teorema de la descomposición prima). Además, también por el Teorema de la descomposición prima, si definimos $T_j = T|_{\mathcal{W}_j} \in L(\mathcal{W}_j)$ entonces el polinomio minimal $m_{T_j}(x) = (x - \lambda_j)^{r_j}$.

Sea $1 \leq j \leq q$: las observaciones anteriores permiten aplicar la Proposición 9.25 al operador $T_j \in L(\mathcal{W}_j)$. En particular, existen únicos $p_j \in N$ y $k_{1,j} = r_j \geq k_{2,j} \geq \ldots \geq k_{p_j,j} \geq 1$ tales que existe una base B_j de $\mathcal{W}_j$ para la que vale la representación de la Eq. (68), donde hemos convenido en notar $d_j = \dim_K \mathcal{W}_j$. De esta forma se verifica el ítem 2.

Fijemos $1 \leq j \leq q$. Sea $1 \leq i \neq j \leq q$ (pero con j fijo!): entonces $T_i - \lambda_j \, I_{\mathcal{W}_i} \in L(\mathcal{W}_i)$ es operador inversible: en efecto, como $p_{T_i}(x) = (x - \lambda_i)^{r_i}$ por la Proposición 9.25, entonces el único autovalor de T_i es λ_i. Así, como $\lambda_j \neq \lambda_i$ (pues $i \neq j$) entonces $\lambda_j \, I_{\mathcal{W}_i} - T_i \in L(\mathcal{W}_i)$ es inversible, ó equivalentemente $T_i - \lambda_j \, I_{\mathcal{W}_i} \in L(\mathcal{W}_i)$ es inversible (de otra forma, λ_j sería autovalor de T_i, que contradice las observaciones anteriores). En particular, si $m \geq 1$ entonces $(T_i - \lambda_j \, I_{\mathcal{W}_i})^m \in L(\mathcal{W}_i)$ es inversible (producto de operadores inversibles es inversible!) y

$$\mathrm{Nul}((T_i - \lambda_j \, I_{\mathcal{W}_i})^m) = 0 \quad \text{para} \quad 1 \leq i \neq j \leq q \,. \tag{71}$$

Si consideramos ahora $1 \leq i \leq q$ arbitario, notemos que $\mathcal{W}_i$ es $(T - \lambda_j \, I)$-invariante: en efecto, $T - \lambda_j \, I = (x - \lambda_j)(T)$ y el Lema 8.9 ; de forma que

$$(T - \lambda_j \, I)|_{\mathcal{W}_i} = T_i - \lambda_j \, I_{\mathcal{W}_i} \implies (T - \lambda_j \, I)^m \,|_{\mathcal{W}_i} = (T_i - \lambda_j \, I_{\mathcal{W}_i})^m \;, \;\; 1 \leq i \leq q \tag{72}$$

por el Lema 8.9.

Si usamos la Eq. (72) en el caso $1 \leq i \neq j \leq q$, la Eq. (71) ahora muestra que

$$\mathrm{Nul}((T - \lambda_j \, I)^m \,|_{\mathcal{W}_i}) = 0 \quad \text{para} \quad 1 \leq i \neq j \leq q \,. \tag{73}$$

Entonces, si aplicamos la Proposición 9.19 con la descomposición dada en el ítem 1. y usamos la Eq. (73) tenemos que

$$\mathrm{Nul}(T - \lambda_j \, I)^m = \sum_{i=1}^{q} \mathrm{Nul}((T - \lambda_j \, I)^m \,|_{\mathcal{W}_i}) = \mathrm{Nul}((T - \lambda_j \, I)^m \,|_{\mathcal{W}_j}) \tag{74}$$

porque todos los términos para $i \neq j$ son nulos. Así, por la Proposición 9.25 aplicada al operador $T_j \in L(\mathcal{W}_j)$ (notemos que T_j cumple las hipótesis de ese resultado pues $m_{T_j}(x) = (x - \lambda_j)^{r_j}$) y las Eqs. (72) y (74) concluímos que

$$\mathrm{Nul}(T - \lambda_j \, I)^m = \mathrm{Nul}((T - \lambda_j \, I)^m \,|_{\mathcal{W}_j}) = (T_j - \lambda_j \, I_{\mathcal{W}_j})^m = \sum_{i=1}^{p_j} \min\{m, \, k_{i,j}\} \,.$$

y vale el ítem 3.

Para finalizar la prueba, es claro que si $B = B_1 \cup \ldots \cup B_q$ entonces B es base de $\mathcal{V}$ y que vale la representación en la Eq. (69). Notemos que la matriz $[T]_B$ es triangular inferior y en su diagonal principal tiene las entradas λ_j repetida d_j veces, $1 \leq j \leq q$: esto se aprecia mejor si notamos que la diagonal principal de $[T]_B$ es la yuxtaposición (pegado) de las diagonales principales de los bloques de Jordan de la Eq. (68), $1 \leq j \leq q$. De esta forma, $x\, I_n - [T]_B$ es una matriz triangular inferior y en su diagonal principal tiene las entradas $x - \lambda_j$ repetida d_j veces, $1 \leq j \leq q$: entonces

$$p_T(x) = \det(x\, I_n - [T]_B) = \prod_{j=1}^{q} (x - \lambda_j)^{d_j} \,.$$

□

Obs 9.29. Consideremos la notación del Teorema 9.28:

1. Algo de terminología: el bloque que aparece en la Eq. 68 se llama bloque de Jordan asociado a λ_j; recordemos que los bloques más pequenos $J(\lambda_j, k_{i,j})$ se llaman bloques elementales de Jordan (asociados a λ_j y $k_{i,j}$).

2. Para poder aplicar el teorema, el polinomio minimal de T se debe poder factorizar como producto de factores lineales en $K[x]$. Recordemos que el cuerpo K es *algebraicamente cerrado* cuando todo polinomio $p(x) \in K[x]$ con $\mathrm{gr}(p(x)) \geq 1$, se factoriza como producto de factores lineales en $K[x]$. Por ejemplo, el teorema fundamental del álgebra establece que el cuerpo de números complejos $\mathbb{C}$ es algebraicamente cerrado.

3. La parte de la prueba del resultado que nos demandó más trabajo corresponde al ítem 3. Notemos que la información del ítem 3 es similar a la información que aparece en el Teorema 9.21. Como hemos visto en el Ejemplo 9.23 (ver también Observación 9.22), podemos usar estas dimensiones para determinar los parámetros $p_j \in N$ y $k_{1,j} = r_j \geq k_{2,j} \geq \ldots \geq k_{p_j,j} \geq 1$ del bloque de Jordan correspondiente a λ_j.

4. El tamaño d_j del bloque de Jordan correspondiente a λ_j coincide con la multiplicidad algebraica de λ_j como raíz del polinomio característico $p_T(x)$ (ver Eq. (70)).

5. La cantidad de bloques $p_j = \mathrm{Nul}(T - \lambda_j\, I)$: en efecto (tomando $m = 1$)

$$\dim \mathrm{N}(T - \lambda_j\, I)^m = \sum_{i=1}^{p_j} \min\{1\,,\, k_{i,j}\} = \sum_{i=1}^{p_j} 1 = p_j \,.$$

6. Hay unicidad de los parámetros de la forma de Jordan en el siguiente sentido: si para cada $1 \leq j \leq q$ existen $m_j \in \mathbb{N}$ y $h_{1,j} \geq \ldots \geq h_{m_j,j} \geq 1$ tales que : existe una base B' de $\mathcal{V}$ tal que

$$[T]_{B'} = \begin{pmatrix} A_{11} & 0 & 0 & \ldots & 0 \\ 0 & A_{22} & 0 & \ldots & 0 \\ 0 & 0 & \ddots & \ldots & 0 \\ \vdots & \vdots & 0 & \ddots & 0 \\ 0 & 0 & 0 & \ldots & A_{qq} \end{pmatrix} \in K^{n\times n} \tag{75}$$

de forma que para cada $1 \leq j \leq q$ se tiene que

$$A_{jj} = \begin{pmatrix} J(\lambda_j, h_{1,j}) & 0 & 0 & \ldots & 0 \\ 0 & J(\lambda_j, h_{2,j}) & 0 & \ldots & 0 \\ 0 & 0 & \ddots & \ldots & 0 \\ \vdots & \vdots & 0 & \ddots & 0 \\ 0 & 0 & 0 & \ldots & J(\lambda_j, h_{m_j,j}) \end{pmatrix} \tag{76}$$

entonces

$$p_j = m_j \quad \text{y} \quad k_{i,j} = h_{i,j} \quad \text{para} \quad 1 \le i \le p_j \ \ , \ \ 1 \le j \le q\,.$$

Este es un hecho fundamental: no vamos a dar detalles de la prueba, pero mencionamos que es posible reducir la unicidad planteada en este ítem al tipo de análisis considerado en la prueba de la Proposición 9.25 (ver también Observación 9.17).

△

Corolario 9.30. *Sea K un cuerpo y sea $B \in K^{n\times n}$. Supongamos que el polinomio minimal de la matriz B se puede factorizar como $m_B(x) = \prod_{j=1}^{q}(x-\lambda_j)^{r_j}$ para ciertos $\lambda_1, \ldots, \lambda_q \in K$ distintos y $1 \le r_j$, $1 \le j \le q$. Entonces B es semejante a una matriz A como en la Eq. (75), cuyos bloques diagonales son de la forma dada en la Eq. (76), para $d_j = \dim_K N((B-\lambda_j)^{r_j})$, $1 \le j \le n$.* □

Ejemplo 9.31. Sea $T \in L(\mathbb{C}^{11})$ tal que

$$p_T(x) = (x-2)^7\,(x-1)^2\,(x+1)^2 \quad \text{y} \quad m_T(x) = (x-2)^4\,(x-1)^2\,(x+1)$$

Nos piden hallar la forma de Jordan de T con estos datos. En estos casos, es posible que los datos no alcancen para determinar exactamente la forma de Jordan y nuestra respuesta sea una serie de posibles casos:

Con los datos que nos han dado, entonces es posible ver que : si $\lambda_1 = 2$, $\lambda_2 = 1$, $\lambda_3 = -1$ son los autovalores distintos de T (raíces de $p_T(x)$), con multiplicidad algebraica $d_1 = 7$, $d_2 = 2$ y $d_3 = 2$ (como raíces de $p_T(x)$) y si $\mathcal{W}_1 = \mathrm{N}(T-\lambda_1\,I)^4$, $\mathcal{W}_2 = \mathrm{N}(T-\lambda_2\,I)^2$ y $\mathcal{W}_3 = \mathrm{N}(T-\lambda_3\,I)$ (donde las potencias "r_j" que consideramos vienen dadas por $m_T(x)$, como en el Teorema 9.28) entonces:

1. $\dim_k \mathcal{W}_j = d_j$, $1 \le j \le 3$ esto nos indica los tamaños de los bloques completos de Jordan correspondientes a cada autovalor de T.

2. Existen bases B_j de $\mathcal{W}_j$, $1 \le j \le 3$, tal que si $B = B_1 \cup B_2 \cup B_3$ entonces

$$[T]_B = \begin{pmatrix} [T_1]_{B_1} & 0 & 0 \\ 0 & [T_2]_{B_2} & 0 \\ 0 & 0 & [T_3]_{B_3} \end{pmatrix} \in \mathbb{C}^{11\times 11}$$

donde $T_j = T|_{\mathcal{W}_j} \in L(\mathcal{W}_j)$, $1 \le j \le 3$.

3. En este caso $[T_j]_{B_j} \in \mathbb{C}^{d_j\times d_j}$, $1 \le j \le 3$ (por el primer ítem de estas observaciones).

4. Con respecto a $[T_1]_{B_1} \in \mathbb{C}^{7\times 7}$, por el Teorema 9.28 sabemos que el tamaño del primer bloque elemental de Jordan de $\lambda_1 = 2$ es $k_{1,1} = 4$, es decir:

$$[T_1]_{B_1} = \begin{pmatrix} J(2,4) & 0 \\ 0 & * \end{pmatrix}$$

donde $* \in \mathbb{C}^{3\times 3}$ denota una matriz 3×3 que corresponde a uno de los siguientes casos:

(a)

$$* = J(2,3) \implies [T_1]_{B_1} = \begin{pmatrix} J(2,4) & 0 \\ 0 & J(2,3) \end{pmatrix}$$

con $p_1 = 2$ (dos bloques elementales) y tamaños: $k_{2,1} = 3$ $(k_{1,1} + k_{2,1} = 7)$

(b)

$$* = \begin{pmatrix} J(2,2) & 0 \\ 0 & J(2,1) \end{pmatrix} \implies [T_1]_{B_1} = \begin{pmatrix} J(2,4) & 0 & 0 \\ 0 & J(2,2) & 0 \\ 0 & 0 & J(2,1) \end{pmatrix}$$

con $p_1 = 3$ (tres bloques elementales) y tamaños: $k_{2,1} = 2$, $k_{3,1} = 1$ $(k_{1,1}+k_{2,1}+k_{3,1} = 7)$

(c)

$$* = \begin{pmatrix} J(2,1) & 0 & 0 \\ 0 & J(2,1) & 0 \\ 0 & 0 & J(2,1) \end{pmatrix} \implies [T_1]_{B_1} = \begin{pmatrix} J(2,4) & 0 & 0 & 0 \\ 0 & J(2,1) & 0 & 0 \\ 0 & 0 & J(2,1) & 0 \\ 0 & 0 & 0 & J(2,1) \end{pmatrix}$$

con $p_1 = 4$ (cuatro bloques elementales) y tamaños: $k_{2,1} = 1$, $k_{3,1} = 1$, $k_{4,1} = 1$ $(k_{1,1} + k_{2,1} + k_{3,1} + k_{4,1} = 7)$.

Notemos que siempre debemos elegir tamaños tales que satisfagan $k_{1,1} = 4 \geq k_{2,1} \geq \ldots \geq k_{p_1,1} \geq 1$.

Si además, nos dan el dato $\text{Nul}(T - 2\,I) = p_1$ (que es la cantidad de bloques elementales) entonces, en este ejemplo particular, podemos determinar completamente la estructura de $[T_1]_{B_1}$. Por ejemplo, si nos dicen que $p_1 = 3$ entonces

$$[T_1]_{B_1} = \begin{pmatrix} J(2,4) & 0 & 0 \\ 0 & J(2,2) & 0 \\ 0 & 0 & J(2,1) \end{pmatrix} = \begin{pmatrix} 2 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 2 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 2 \end{pmatrix}$$

Notemos el aspecto que tiene la matriz, cuando pegamos los bloques: si miramos la matriz con cuidado, se pueden apreciar los bloques, en función de cómo aparecen los 1's en la subdiagonal inferior. Con ese criterio, se pueden diferenciar los tres bloques (de 4×4, 2×2 y 1×1 en este caso).

5. Con respecto a $[T_2]_{B_2} \in \mathbb{C}^{2\times 2}$, sabemos que el tamaño del bloque elemental de Jordan más grande es $k_{2,1} = 2$ (ya que 2 es la potencia de $(x - \lambda_2)$ en $m_T(x)$). Pero entonces, hay un solo bloque elemental, pues $k_{2,1} = 2 = d_2$, que es el tamaño del bloque completo de Jordan correspondiente a $\lambda_2 = 1$. Entonces

$$[T_2]_{B_2} = J(1,2) = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$$

6. Con respecto a $[T_3]_{B_3} \in \mathbb{C}^{2\times 2}$, sabemos que el tamaño del bloque elemental de Jordan más grande es $k_{1,3} = 1$ (ya que 1 es la potencia de $(x - \lambda_3)$ en $m_T(x)$). Pero entonces, los demás bloques son, necesariamente, de tamaño 1 también : es decir, $k_{1,3} = 1$ implica que $k_{2,3} = 1$ porque hemos convenido en disponer los bloques elementales de forma que sus tamaños van decreciendo ó se mantienen igual, y todos los bloques son de tamaño mayor ó igual que 1. Así, $p_3 = 2$ (dos bloques elementales de forma que $k_{1,3} + k_{2,3} = d_3 = 2$) y como $\lambda_3 = -1$ tenemos

$$[T_3]_{B_3} = \begin{pmatrix} J(-1,1) & 0 \\ 0 & J(-1,1) \end{pmatrix} = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$$

7. La respuesta del ejercicio es: con la información inicial dada, las posibles formas de Jordan de T son:

$$\begin{pmatrix} J(2,4) & 0 & 0 & 0 & 0 \\ 0 & * & 0 & 0 & 0 \\ 0 & 0 & J(1,2) & 0 & 0 \\ 0 & 0 & 0 & J(-1,1) & 0 \\ 0 & 0 & 0 & 0 & J(-1,1) \end{pmatrix}$$

donde $* \in \mathbb{C}^{3\times 3}$ es alguna de las matrices correspondientes a los casos (a), (b) ó (c) del ítem 4. Como hemos mencionado antes, si nos dan el dato adicional de $\mathrm{Nul}(T - 2\,I) = p_1$, entonces podemos determinar (en este ejemplo) la forma completamente.

Para finalizar, si por ejemplo $p_1 = 3$ entonces la forma de Jordan de T resulta

$$= \begin{pmatrix} 2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 2 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 2 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & J(1,2) & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 \end{pmatrix}$$

donde hemos incluído el bloque $J(1,2)$ de tamaño 2×2. Noten que si miramos la matriz con cuidado podemos ver, por ejemplo, los tres bloques elemetales de Jordan de tamaños 4, 2 y 1 asociados al autovalor $\lambda = 2$. Como ejercicio, desarrollar también el bloque $J(1,2)$ (de tamaño 2×2) y escribir la matriz 11×11 completa! $\triangle$

10 El espacio dual

Sea K un cuerpo. Entonces podemos identificar a $K = K^1$, el K-espacio vectorial de los vectores columna de una entrada. En este sentido, notemos que $\dim_K K = 1$.

10.1 Definición y primeras propiedades

Comenzamos describiendo la noción de espacio dual.

Definición 10.1. *Sea $\mathcal{V}$ un K-ev. Un funcional lineal sobre $\mathcal{V}$ es una transformación lineal $f : \mathcal{V} \to K$, considerando a K como K-ev. Así, f satisaface:*

$$f(\alpha\, u + v) = \alpha\, f(u) + f(v) \quad \textit{para} \quad u,\, v \in \mathcal{V} \quad \textit{y} \quad \alpha \in K\,.$$

$\square$

Ejemplo 10.2. Consideremos los siguientes ejemplos:

1. Sea K un cuerpo y sea $\vec{x} = (x_1, \ldots, x_n) \in K^n$. Entonces la función $f_{\vec{x}} : K^n \to K$ dada por

   $$f_{\vec{x}}(\vec{y}) = \sum_{j=1}^{n} x_j\, y_j \quad \text{donde} \quad \vec{y} = (y_1, \ldots, y_n) \in K^n\,,$$

   es un funcional lineal (ejercicio).

   Recíprocamente, si $f : K^n \to K$ es funcional, definamos $x_j = f(e_j)$, $1 \le j \le n$, donde $B_c = \{e_1, \ldots, e_n\}$ denota la base canónica de K^n. Entonces $\vec{x} = (x_1, \ldots, x_n) \in K^n$ y si $\vec{y} = (y_1, \ldots, y_n) \in K^n$,

   $$f(\vec{y}) = f(\sum_{j=1}^{n} y_j\ e_j) = \sum_{j=1}^{n} y_j\ f(e_j) = \sum_{j=1}^{n} y_j\, x_j = f_{\vec{x}}(\vec{y})$$

   de forma que $f = f_{\vec{x}}$. Lo anterior muestra que todo funcional es de la forma $f_{\vec{x}}$, para algún $\vec{x} \in K^n$.

2. Sea K un cuerpo y sea $\operatorname{tr}(\cdot) : K^{n\times n} \to K$ dada por:

$$\operatorname{tr}(A) = \sum_{j=1}^{n} A_{jj} \quad \text{para} \quad A = (A_{ij}) \in K^{n\times n} .$$

   Entonces $\operatorname{tr}(\cdot)$ es un funcional lineal, llamado *traza*.

3. Si consideramos el $\mathbb{R}$-ev $C([a,b])$ de las funciones continuas definidas en un intervalo $[a,b] \subset \mathbb{R}$ (a valores reales) entonces la integral definida $I : C([a,b]) \to \mathbb{R}$ dada por

$$I(f) = \int_a^b f(t)\ dt \quad \text{para} \quad f \in C([a,b])$$

   es un funcional. De hecho, las llamadas propiedades de linealidad de la integral vistas en Análisis I muestran que:

$$I(\alpha\, f + g) = \int_a^b (\alpha\, f(t) + g(t))\ dt = \alpha \int_a^b f(t)\ dt + \int_a^b g(t)\ dt = \alpha\, I(f) + I(g)\,.$$

$\triangle$

Definición 10.3. *Sea $\mathcal{V}$ un K-ev, $\dim_K \mathcal{V} = n$. El espacio dual, notado $\mathcal{V}^*$, es el K-ev formado por todos los funcionales lineales $\mathcal{V}^* = L(\mathcal{V}\,,\, K)$, con la estructura de K-ev que hemos considerado en este espacio de transformaciones.*

Con la notación de la definición anterior, se tiene que

$$\dim_K \mathcal{V}^* = \dim_K \mathcal{V} \cdot \dim_K K = \dim_K \mathcal{V}$$

pues $\dim_K K = 1$.

Obs 10.4. En muchos contextos, los vectores de un espacio vectorial $\mathcal{V}$ describen fenómenos complejos (ó procesos, estados, etc). En estos casos, para entender el fenómeno, tomamos mediciones (que dan lugar a valores escalares, por ejemplo números reales, que son más sencillos) de distintos aspectos del proceso: en algunos casos, es posible modelar las mediciones como evaluar los vectores que describen el proceso en funcionales lineales. En este sentido, el valor $f(v) \in K$ puede pensarse como una medición escalar de un fenómeno complicado; así, es de esperar que si tomamos suficientes mediciones de distintos aspectos (independientes) del proceso podamos describir completamente al mismo. El siguiente resultado está orientado a este concepto.

Teorema 10.5. *Sea $\mathcal{V}$ un K-ev, $\dim_K \mathcal{V} = n$. Si $B = \{v_1, \ldots, v_n\}$ es base de $\mathcal{V}$ sea $B^* = \{f_1, \ldots, f_n\} \subset \mathcal{V}^*$ el conjunto cuyos elementos (funcionales) satisfacen las siguientes relaciones con los vectores de B:*

$$f_i(v_j) = \delta_{ij} \quad \textit{para} \quad 1 \le i,\, j \le n\,. \tag{77}$$

Entonces B^* queda determinada de forma única por las condiciones en la Eq. (77) *y es una base del espacio dual $\mathcal{V}^*$.*

Demostración. Notemos que dada la base B como antes y $1 \le i \le n$, existe un único funcional $f_i \in \mathcal{V}^* = L(\mathcal{V}, K)$ tal que $f_i(v_j) = \delta_{ij}$, para $1 \le j \le n$ (aquí estamos usando el resultado que indica la existencia de transformaciones lineales que transforman elementos de una base en vectores cualesquiera de K (pensado como K-ev): en este caso los vectores cualesquiera de K son 0 y 1).

De esta forma podemos construir $B^* = \{f_1, \ldots, f_n\}$ que verifica las identidades en la Eq. (77); estas identidades son llamadas *relaciones de dualidad* entre B y B^*.

Veamos que $B^* \subset \mathcal{V}^*$ es conjunto l.i. (aquí usamos la estructura de K-ev de $\mathcal{V}^* = L(\mathcal{V}, K)$). En efecto, si $\alpha_1, \ldots, \alpha_n \in K$ son tales que

$$0 = \sum_{i=1}^{n} \alpha_i \ f_i \in \mathcal{V}^*$$

entonces, como la igualdad anterior es una igualdad de funciones, podemos evaluar a ambos lados en v_j, $1 \leq j \leq n$:

$$0_K = 0(v_j) = (\sum_{i=1}^{n} \alpha_i \ f_i)\,(v_j) = \sum_{i=1}^{n} \alpha_i \ f_i(v_j) = \sum_{i=1}^{n} \alpha_i \ \delta_{ij} = \alpha_j$$

que muestra que $\alpha_j = 0$, $1 \leq j \leq n$.

Como $B^* \subset \mathcal{V}^*$ es conjunto l.i. con n elementos y $\dim_K \mathcal{V}^* = n$, hemos probado que B^* es necesariamente una base de $\mathcal{V}^*$ (resulta ser un sistema de generadores también).

Finalmente, verificamos la unicidad: supongamos que $B' = \{g_1, \ldots, g_n\}$ satisface las relaciones de dualidad dadas en la Eq. (77). En particular, si $1 \leq i \leq n$ entonces

$$f_i(v_j) = \delta_{ij} = g_i(v_j) \quad \text{para} \quad 1 \leq j \leq n\,.$$

Como f_i, g_i son transformaciones lineales (a valores escalares) que coinciden en una base del dominio, entonces $f_i = g_i$, $1 \leq i \leq n$. □

Definición 10.6. *Sea $\mathcal{V}$ un K-ev y sea $B = \{v_1, \ldots, v_n\}$ base de $\mathcal{V}$. Si $B^* = \{f_1, \ldots, f_n\}$ es como en el Teorema anterior (de forma que verifica la Eq. (77)), decimos que B^* es la base dual de la base B.*

Ejemplo 10.7. Sea K un cuerpo y consideremos K^n como K-ev. Si $B_c = \{e_1, \ldots, e_n\}$ denota la base canónica de K^n entonces su base dual $B_c^* = \{\pi_1, \ldots, \pi_n\}$ está formada por los funcionales

$$\pi_i : K^n \to K \quad \text{dado por} \quad \pi_i(x_1, \ldots, x_n) = x_i \ \ , \ \ (x_1, \ldots, x_n) \in K^n\,.$$

En efecto, basta notar que las $\pi_i \in (K^n)^*$ satisfacen las relaciones de dualidad (ejercicio)

$$\pi_i(e_j) = \delta_{ij} \quad \text{para} \quad 1 \leq i,\, j \leq n\,.$$

El funcional π_i es llamado *la proyección sobre la i-ésima coordenada.* △

Obs 10.8. Sea K un cuerpo y consideramos $\mathcal{V} = K^n$. Definimos $F : K^n \to (K^n)^*$ dada por

$$F(\vec{x}) = f_{\vec{x}} \quad \text{para} \quad \vec{x} \in K^n$$

donde $f_{\vec{x}}$ está definido como en el ítem 1 del Ejemplo 10.2. Entonces F es un isomorfismo lineal entre K^n y $(K^n)^*$ (verificar que F es lineal y que F transforma la base canónica $B_c = \{e_1, \ldots, e_n\}$ de K^n en la base dual de la base canónica $B_c^* = \{\pi_1, \ldots, \pi_n\}$). △

Ejemplo 10.9. Sea K un cuerpo y consideremos $\mathcal{V} = K_{n-1}[x]$ el espacio de los polinomios de grado menor ó igual que $n - 1$. Este es un K-ev tal que $\dim_K \mathcal{V} = n$. Dado $\alpha \in K$ entonces podemos considerar la evaluación en α, notada $e_\alpha : \mathcal{V} \to K$ dada por

$$e_\alpha(p(x)) = p(\alpha) \quad \text{para} \quad p(x) \in \mathcal{V}\,.$$

Entonces, gracias a las propiedades que tiene la evaluación en un escalar fijo ($\alpha \in K$), $e_\alpha \in \mathcal{V}^*$ (ejercicio), para cada $\alpha \in K$.

Sean $c_1, \dots, c_n \in K$ escalares distintos, y consideremos la base $\mathcal{L} = \{q_1(x), \dots, q_n(x)\}$de $\mathcal{V}$ dada por los polinomios de Lagrange asociados a $c_1, \dots, c_n$, es decir: si $1 \le j \le n$,

$$q_j(x) = a_j^{-1} \prod_{1\le i\ne j\le k} (x - c_i) \quad \text{donde} \quad a_j = \prod_{1\le i\ne j\le k} (c_j - c_i) \neq 0\,.$$

Recordemos las relaciones de dualidad que verifican estos polinomios:

$$e_{c_i}(q_j(x)) = q_j(c_i) = \delta_{ij} \quad \text{para} \quad 1 \le i,\, j \le n\,.$$

Las relaciones anteriores junto con el Teorema 10.5 indican que $\mathcal{L}^* = \{e_{c_1}, \dots, e_{c_n}\}$ es una base del espacio dual $\mathcal{V}^*$, que es la base dual de la base $\mathcal{L}$.

Finalmente, recordemos que si $p(x) \in \mathcal{V}$ entonces hemos probado que

$$[p(x)]_{\mathcal{L}} = (p(c_1), \dots, p(c_n)) = (e_{c_1}(p(x)), \dots, e_{c_n}(p(x))) \in K^n\,.$$

En realidad, este hecho es un caso particular del siguiente resultado. △

Proposición 10.10. *Sea $\mathcal{V}$ un K-ev, $B = \{v_1, \dots, v_n\}$ base de $\mathcal{V}$ y sea $B^*\{f_1, \dots, f_n\}$ la base dual de B. Entonces, si $v \in \mathcal{V}$ y $f \in \mathcal{V}^*$:*

$$\sum_{j=1}^n f_j(v)\, v_j \implies [v]_B = (f_1(v), \dots, f_n(v)) \in K^n\,,$$

$$f = \sum_{j=1}^n f(v_j)\, f_j \implies [f]_{B^*} = (f(v_1), \dots, f(v_n)) \in K^n\,.$$

Demostración. Como B es base de $\mathcal{V}$, existen $\alpha_1, \dots, \alpha_n \in K$ tales que $v = \sum_{j=1}^n \alpha_j\ v_j$: entonces

$$f_i(v) = f_i(\sum_{j=1}^n \alpha_j\ v_j) = \sum_{j=1}^n \alpha_j\ f_i(v_j) = \sum_{j=1}^n \alpha_j\ \delta_{ij} = \alpha_i\ , \quad \text{para} \quad 1 \le i \le n\,.$$

Así, $\alpha_i = f_i(v)$, $1 \le i \le n$, que muestra la primer identidad.

De forma similar, como B^* es base de $\mathcal{V}^*$, existen $\beta_1, \dots, \beta_n \in K$ tales que $f = \sum_{i=1}^n \beta_i\, f_i$: entonces

$$f(v_j) = (\sum_{i=1}^n \beta_i\ f_i)(v_j) = \sum_{i=1}^n \beta_i\ f_i(v_j) = \sum_{i=1}^n \beta_i\ \delta_{ij} = \beta_j\ , \quad \text{para} \quad 1 \le j \le n\,.$$

Así, $\beta_j = f(v_j)$, $1 \le j \le n$, que muestra la segunda identidad. □

Ajustensé los cinturones porque en lo que sigue vamos a ver los...

10.2 Hiperespacios

Comenzamos con el concepto de hiperespacio.

Definición 10.11. *Sea $\mathcal{V}$ un K-ev, $\dim_K \mathcal{V} = n$. Un hiperespacio de $\mathcal{V}$ es un subespacio $\mathcal{W} \subset \mathcal{V}$ tal que $\dim_K \mathcal{W} = n - 1$.*

Lema 10.12. *Sea $\mathcal{V}$ un K-ev, $\dim_K \mathcal{V} = n$. Dado un subconjunto $\mathcal{W} \subseteq \mathcal{V}$ entonces: $\mathcal{W}$ es un hiperespacio si y solo si existe $f \in \mathcal{V}^*$, $f \neq 0$ tal que $\mathcal{W} = N(f)$.*

Demostración. Sea $\mathcal{W} \subseteq \mathcal{V}$ tal que existe $f \in \mathcal{V}^*$, $f \neq 0$ tal que $\mathcal{W} = \mathrm{N}(f)$. Por hipótesis, $f \neq 0$ implica que $\mathrm{Im}(f) \neq \{0\}$. Como $f \in \mathcal{V}^* = L(\mathcal{V}, K)$ es transformación lineal, entonces $\mathrm{Im}(f) \subseteq K$ es un subespacio no nulo, de forma que $1 \leq \dim_K \mathrm{Im}(f) \leq \dim_K K = 1$ lo que muestra que $\dim_K \mathrm{Im}(f) = \dim_K K = 1$ y luego que $\mathrm{Im}(f) = K$ (los únicos subespacios de K son K y $\{0\}$ - ejercicio). Entonces, por un resultado previo

$$n = \dim_K \mathcal{V} = \dim_K \mathrm{Im}(f) + \dim_K \mathrm{N}(f) \implies \dim_K \mathrm{N}(f) = n - 1$$

y $\mathcal{W} = \mathrm{N}(f)$ es hiperespacio.

Si $\mathcal{W}$ es hiperespacio, sea $B' = \{w_1, \ldots, w_{n-1}\}$ una base de $\mathcal{W}$ (pues $\dim_K \mathcal{W} = n - 1$). En particular, $B' \subset \mathcal{V}$ es conjunto l.i. y lo podemos extender a una base, agregando un vector $w_n \in \mathcal{V}$: así, $B = \{w_1, \ldots, w_n\}$ es base de $\mathcal{V}$. Sea $B^* = \{f_1, \ldots, f_n\} \subset \mathcal{V}^*$ la base dual de B. Veamos que $\mathcal{W} = \mathrm{N}(f_n)$: en efecto, si $w \in \mathcal{W}$ entonces w es combinación lineal de los elementos de B' : entonces existen $\alpha_1, \ldots, \alpha_{n-1} \in K$ tales que $w = \sum_{j=1}^{n-1} \alpha_j \; w_j$. En este caso, usando las relaciones de dualidad entre B y B^* (y notando que B' está formada por los primeros $n - 1$ vectores de B)

$$f_n(w) = f_n(\sum_{j=1}^{n-1} \alpha_j \; w_j) = \sum_{j=1}^{n-1} \alpha_j \; f_n(w_j) = \sum_{j=1}^{n-1} \alpha_j \; \delta_{j,n} = 0$$

pues $\delta_{j,n} = 0$, para $1 \leq j \leq n - 1$. Entonces $\mathcal{W} \subseteq \mathrm{N}(f_n)$. Además, como $f_n \neq 0$ entonces $n - 1 = \dim_K \mathcal{W} \leq \dim \mathrm{N}(f_n) = n - 1$, donde la última igualdad es por la primer parte de la prueba. Entonces debe valer la igualdad de dimensiones $\dim_K \mathcal{W} = \dim \mathrm{N}(f_n)$ y como $W \subseteq \mathrm{N}(f_n)$ vale la igualdad de espacios $W = \mathrm{N}(f_n)$. □

Definición 10.13. *Sea $\mathcal{V}$ un K-ev,* $\dim_K \mathcal{V} = n$ *y sea $S \subset \mathcal{V}$ subconjunto. Definimos el anulador de S, notado $S^o \subseteq \mathcal{V}^*$, como el conjunto dado por*

$$S^o = \{f \in \mathcal{V}^* \;:\; f(v) = 0 \quad \textit{para todo} \quad v \in S\}\,.$$

□

Ejemplo 10.14. Sea $\mathcal{V}$ un K-ev, $\dim_K \mathcal{V} = n$. Entonces

1. $\{0\}^o = \mathcal{V}^*$;
2. $\mathcal{V}^o = \{0\} \subset \mathcal{V}^*$.

Las identidades anteriores son una consecuencia de la Definición 10.13 y se dejan como ejercicio.△

Proposición 10.15. *Sea $\mathcal{V}$ un K-ev,* $\dim_K \mathcal{V} = n$ *y sea $S \subset \mathcal{V}$ subconjunto.*

1. *S^o es un subespacios de $\mathcal{V}^*$.*
2. *Si $S_1 \subset S_2 \subset \mathcal{V}$ entonces $S_2^o \subset S_1^o \subset \mathcal{V}^*$.*
3. *Si $\overline{S} \subset \mathcal{V}$ denota el subespacio generado por S entonces $(\overline{S})^o = S^o$.*

Demostración. Veamos el ítem 1: si $0 \in \mathcal{V}^*$ entonces $0(v) = 0$ para todo $v \in S$ (de hecho para todo $v \in \mathcal{V}$) de forma que $0 \in S^o$. Si $f,\, g \in S^o$ y $\alpha \in K$ entonces, por hipótesis: $f(v) = g(v) = 0$ para todo $v \in S$. Así, si $v \in S$ vale que

$$(\alpha \; f + g)(v) = \alpha \; f(v) + g(v) = 0 + 0 = 0$$

con lo que $\alpha \; f + g \in S^o$. Lo anterior muestra que S^o es subespacio del dual $\mathcal{V}^*$.

El ítem 2 es una consecuencia de la definición de anulador: si $S_1 \subset S_2$ y si $f \in S_2^o$ entonces $f(v) = 0$ para $v \in S_2$; en particular, $f(v) = 0$ para $v \in S_1 \subset S_2$ que muestra que $f \in S_1^o$.

Como $S \subset \overline{S}$ entonces vemos que $(\overline{S})^o \subset S^o$, por el ítem 2. Sea $f \in S^o$ y sea $v \in \overline{S}$: entonces existen $\alpha_1, \ldots, \alpha_m \in K$ y $v_1, \ldots, v_m \in S$ tales que $v = \sum_{j=1}^m \alpha_j\ v_j$: entonces

$$f(v) = f(\sum_{j=1}^m \alpha_j\ v_j) = \sum_{j=1}^m \alpha_j\ f(v_j) = \sum_{j=1}^m \alpha_j\ 0 = 0$$

donde usamos que $f(v_j) = 0$, pues $v_j \in S$ y $f \in S^o$, $1 \leq j \leq m$. Lo anterior muestra que si $f \in S^o$ entonces $f \in (\overline{S})^o$. Así, $S^o \subseteq (\overline{S})^o$ y vale la igualdad $S^o = (\overline{S})^o$. □

Teorema 10.16. *Sea $\mathcal{V}$ un K-ev, $\dim_K \mathcal{V} = n$ y sea $\mathcal{W} \subset \mathcal{V}$ subespacio. Entonces*

$$\dim_K \mathcal{V} = \dim_K \mathcal{W} + \dim_K \mathcal{W}^o\,.$$

Demostración. Si $\mathcal{W} = \{0\}$ ó $\mathcal{W} = \mathcal{V}$ entonces el resultado es una consecuencia del Ejemplo 10.14. Así, si $\dim_K \mathcal{W} = k$ suponemos que $1 \leq k \leq n-1$. Sea $B' = \{w_1, \ldots, w_k\}$ base de $\mathcal{W}$. Como $B' \subset \mathcal{V}$ es conjunto l.i., lo podemos extender a una base $B = \{w_1, \ldots, w_n\}$ de $\mathcal{V}$. Sea $B^* = \{f_1, \ldots, f_n\}$ la base dual de la base B (de forma que B^* es base de $\mathcal{V}^*$).

Sea $k+1 \leq j \leq n$: si consideramos los vectores w_i de la base B', $1 \leq i \leq k$, se verifica

$$f_j(w_i) = \delta_{ij} = 0$$

donde usamos las relaciones de dualidad entre B y B^*, el hecho de que los primeros k vectores de B son los vectores de B', de forma que: $1 \leq i \leq k$, $k+1 \leq j \leq n$ implica que $i \neq j$ y $\delta_{ij} = 0$. Entonces, por la Proposición 10.15 (ítem 3),

$$f_j \in (B')^o = (\overline{B'})^o = \mathcal{W}^o \quad \text{para} \quad k+1 \leq j \leq n\,.$$

Entonces $B_+ = \{f_{k+1}, \ldots, f_n\} \subset \mathcal{W}^o$. Veamos que B_+ es una base de $\mathcal{W}^o$. Por un lado, $B_+ \subset \mathcal{V}^*$ es conjunto l.i., pues es un subconjunto de la base B^*. Por otro lado, si $f \in \mathcal{W}^o$ entonces, utilizando la base B^* de $\mathcal{V}^*$ y la Proposición 10.10 vemos que

$$f = \sum_{j=1}^n f(w_j)\ f_j \implies f = \sum_{j=k+1}^n f(w_j)\ f_j$$

pues $f(w_1) = \ldots = f(w_k) = 0$, ya que $w_1, \ldots, w_k \in \mathcal{W}$ y $f \in \mathcal{W}^o$. Así, $B_+ \subset \mathcal{W}^o$ es un sistema de generadores de $\mathcal{W}^o$ que es l.i. De esta forma, B_+ es una base de $\mathcal{W}^o$ y ahora podemos calcular $\dim_K \mathcal{W}^o = n-k$. Así

$$\dim_K \mathcal{W} + \dim_K \mathcal{W}^o = k + (n-k) = n = \dim_K \mathcal{V}\,.$$

□

Corolario 10.17. *Sea $\mathcal{V}$ un K-ev, $\dim_K \mathcal{V} = n$. Sea $\mathcal{W} \subseteq \mathcal{V}$ subespacio, $\dim_K \mathcal{W} = k \leq n-1$. Si $\{g_1, \ldots, g_{n-k}\}$ es base de $\mathcal{W}^o$ entonces*

$$\mathcal{W} = \bigcap_{j=1}^{n-k} N(g_j) = \{v \in \mathcal{V}\ :\ g_j(v) = 0\ ,\ 1 \leq j \leq n-k\}\,.$$

Demostración. Si $\dim_K \mathcal{W} = k$ entonces, por el Teorema anterior existe una base $B_+ = \{g_1, \ldots, g_{n-k}\}$ de $\mathcal{W}^o$. Sea

$$\mathcal{W}' = \bigcap_{j=1}^{n-k} \mathrm{N}(g_j) = \{v \in \mathcal{V} \ : \ g_j(v) = 0 \ , \ 1 \leq j \leq n-k\} \,.$$

Recordemos que la intersección de subespacios (en este caso, hiperespacios) es subespacio; entonces $\mathcal{W}'$ es subespacio. Además, notemos que si $w \in \mathcal{W}$ entonces $g_j(w) = 0$, pues $g_j \in \mathcal{W}^o$, $1 \leq j \leq n-k$. Entonces $\mathcal{W} \subset \mathcal{W}'$, que implica que $\dim_K \mathcal{W} \leq \dim_K \mathcal{W}'$.

Por otro lado, $\{g_1, \ldots, g_{n-k}\} \subset (\mathcal{W}')^o$, por definición de $\mathcal{W}'$ (verificar!). Como $B_+ = \{g_1, \ldots, g_{n-k}\}$ es base de $\mathcal{W}^o$ entonces $\mathcal{W}^o = \overline{B_+} \subset (\mathcal{W}')^o$ (aquí usamos el hecho de que $(\mathcal{W}')^o$ es subespacio, y las propiedades del subespacio generado por un conjunto: si $\{g_1, \ldots, g_{n-k}\} \subset (\mathcal{W}')^o$ entonces el subespacio generado por $\{g_1, \ldots, g_{n-k}\}$ esta incluido en $(\mathcal{W}')^o$). Luego $\dim_K \mathcal{W}^o \leq \dim_K (\mathcal{W}')^o$.

Los hechos anteriores muestran que

$$n = \dim_K \mathcal{W} + \dim_K \mathcal{W}^o \leq \dim_K \mathcal{W}' + \dim_K (\mathcal{W}')^o = n$$

de forma que

$$k = \dim_K \mathcal{W} = \dim_K \mathcal{W}' \quad , \quad n - k = \dim_K (\mathcal{W}')^o$$

Como $\mathcal{W} \subseteq \mathcal{W}'$ entonces vemos que $\mathcal{W} = \mathcal{W}'$. □

Corolario 10.18. *Sea $\mathcal{V}$ un K-ev, $\dim_K \mathcal{V} = n$. Sea $\mathcal{W}_1, \mathcal{W}_2 \subseteq \mathcal{V}$ subespacios. Entonces*

$$\mathcal{W}_1 = \mathcal{W}_2 \quad \textit{si y solo si} \quad \mathcal{W}_1^o = \mathcal{W}_2^o \,.$$

Demostración. Es claro que si $\mathcal{W}_1 = \mathcal{W}_2$ entonces $\mathcal{W}_1^o = \mathcal{W}_2^o$. Recíprocamente, si $\mathcal{W}_1^o = \mathcal{W}_2^o$ notemos que $\dim_K \mathcal{W}_1 = \dim_K \mathcal{W}_2$ (donde usamos el Teorema 10.16 para la igualdad de dimensiones). Si $\dim_K \mathcal{W}_1 = \dim_K \mathcal{W}_2 = n$ entonces $\mathcal{W}_j = \mathcal{V}$, $j = 1, 2$ y vale el resultado. Si $\dim_K \mathcal{W}_1 = \dim_K \mathcal{W}_2 = k \leq n - 1$ y $\{g_1, \ldots, g_{n-k}\}$ es base de $\mathcal{W}_1^o = \mathcal{W}_2^o$ entonces, por el Corolario 10.17 concluimos que

$$\mathcal{W}_1 = \bigcap_{j=1}^{n-k} \mathrm{N}(g_j) = \mathcal{W}_2 \,.$$

□

Obs 10.19. Sea K un cuerpo. Si consideramos un sistema de ecuaciones lineales homogéneo de n variables y m ecuaciones en K, es decir un sistema modelado por $A\vec{x} = 0_m$ donde $A \in K^{m \times n}$ y $\vec{x} = (x_1, \ldots, x_n)$, entonces sabemos bien que el espacio solución de este sistema es un subespacio de K^n. La dimensión del espacio solución es una cantidad que requiere cierto cálculo, porque el sistema puede tener ecuaciones redundantes ó dependientes: el ejemplo más sencillo es pensar que las m ecuaciones son en realidad una ecuación repetida m veces: en estes caso, si la ecuación es no trivial (es decir, no idénticamente nula) el espacio solución tendrá dimensión $n - 1$. En general, el espacio solución tendrá dimensión k, donde $n - k$ es la cantidad de ecuaciones *independientes* del sistema; recordemos que $n - k$ coincide con la cantidad de filas no nulas de la matriz escalonada y reducida por filas $A_R \in K^{m \times n}$ que se obtiene de A por reducción por filas. De hecho, como hemos visto en Álgebra I, se verifica

$$N(A) = \{\vec{x} \in K^n \ : \ A\,\vec{x} = 0_m\} = \{\vec{x} \in K^n \ : \ A_R\,\vec{x} = 0_{n-k}\} = N(A_R) \,.$$

Una pregunta natural es la siguiente: si nos dan de antemano un subespacio $\mathcal{W}$ de K^n de dimensión $\dim_K \mathcal{W} = k$, es posible describir ese subespacio como el espacio solución de un sistema de ecuaciones lineales homogéneo de $n - k$ ecuaciones (independientes) ? es decir, como el espacio solución del sistema $B\,\vec{x} = 0_{n-k}$, donde $B \in K^{(n-k) \times n}$.

La respuesta es que siempre se puede describir a $\mathcal{W}$ como espacio solución, y de hecho es posible hacer esto mediante un cierto método: todos estos hechos están codificados en el (más bien abstracto) Corolario 10.17.

En efecto, sea $B_+ = \{g_1, \ldots, g_{n-k}\}$ una base de $\mathcal{W}^o \subset (K^n)^*$. En este caso, el Corolario 10.17 prueba que

$$\mathcal{W} = \{v \in K^n \ : \ g_i(v) = 0 \ , \ 1 \leq i \leq n-k\} \, .$$

Recordemos además, que cada funcional $f \in (K^n)^*$ se puede describir como $f_{\vec{b}}$ para cierto vector $\vec{b} = (b_j)_{j=1}^n$ (ver el Ejemplo 10.2) es decir, $f_{\vec{b}}(\vec{y}) = \sum_{j=1}^n b_j \ y_j$, para $\vec{y} = (y_j) \in K^n$.

En particular, existen vectores $\vec{b}_i = (b_{i1}, \ldots, b_{in}) \in K^n$ tales que $g_i = f_{\vec{b}_i}$, $1 \leq i \leq n-k$. Entonces, dado $\vec{x} \in K^n$ se verifica: $\vec{x} = (x_j) \in \mathcal{W}$ si y solo si

$$0 = g_i(\vec{x}) = \sum_{j=1}^n b_{ij} \, x_j \quad \text{para} \quad 1 \leq i \leq n-k \, . \tag{78}$$

De esta forma, si construimos la matriz $B \in K^{(n-k)\times n}$ tal que $B = (b_{ij})$ con los coeficientes dados por los vectores $\vec{b}_i$ que representan los funcionales g_i, para $1 \leq i \leq n-k$, entonces la Eq. (78) es equivalente a afirmar que $\vec{x}$ es solución del sistema $B\,\vec{x} = 0_{n-k}$. Estos hechos muestran que

$$\mathcal{W} = \{\vec{x} \in K^n \ : \ B\,\vec{x} = 0_{n-k}\} \, .$$

$\triangle$

En el siguiente ejemplo consideramos un caso particular del problema propuesto en la Observación anterior.

Ejemplo 10.20. Consideremos el subespacio $\mathcal{W} \subset \mathbb{R}^4$ generado por $\{v_1, \, v_2\}$, donde $v_1 = (1,1,0,0)$ y $v_2 = (0,1,1,0)$.

Comenzamos calculando el anulador de $\mathcal{W}$. Por el ítem 3 de la Proposición 10.15, $f \in \mathcal{W}^o$ si y solo si $f \in \{v_1, \, v_2\}^o$ es decir, $f(v_1) = 0$ y $f(v_2) = 0$.

Por el Ejemplo 10.2, sabemos que todo funcional $f \in (\mathbb{R}^4)^*$ se escribe como $f = f_{\vec{x}}$ para algún $x = (x_j)_{j=1}^4 \in \mathbb{R}^4$. Así, $f = f_{\vec{x}} \in \mathcal{W}^o$ si y solo si

$$f_{\vec{x}}(v_1) = x_1 + x_2 = 0 \quad \text{y} \quad f_{\vec{x}}(v_2) = x_2 + x_3 = 0 \, .$$

Las condiciones anteriores pueden escribirse como un sistema de ecuaciones lineales homogéneo de la forma $A\,\vec{x} = 0$ donde

$$A = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \end{pmatrix} .$$

Notemos que las incógnitas de este sistema corresponden a las entradas del vector $\vec{x} = (x_j)_{j=1}^4$ que está asociado al funcional $f_{\vec{x}} \in (\mathbb{R}^4)^*$. Si calculamos la matriz escalonada y reducida por filas asociada a A (en este caso, lo podemos hacer en un paso, cambiando la F_1 de A por $F_1 - F_2$) llegamos a la matriz A_R, dada por

$$A_R = \begin{pmatrix} 1 & 0 & -1 & 0 \\ 0 & 1 & 1 & 0 \end{pmatrix} .$$

La forma de A_R indica que las variables x_1 y x_2 son variables dependientes del sistema, mientras que las variables x_3 y x_4 son libres. Si resolvemos el sistema reducido (siempre expresando las variables dependientes en función de las variables libres) obtenemos que

$$x_1 = x_3 \quad , \quad x_2 = -x_3 \quad \text{con} \quad x_3 \quad \text{y} \quad x_4 \ \text{libres} \, .$$

Así, los vectores solución del sistema tienen la forma

$$\vec{x} = (x_3,\, -x_3,\, x_3,\, x_4) = x_3\ (1,\, -1,\, 1,\, 0) + x_4\ (0,0,0,1) \quad \text{con} \quad x_3,\, x_4 \in \mathbb{R}\,.$$

Entonces $f = f_{\vec{x}} \in \mathcal{W}^o$ si y solo si existen $x_3,\, x_4 \in \mathbb{R}$ tales que

$$f = f_{\vec{x}} = f_{x_3\,(1,\,-1,\,1,\,0)\,+\,x_4\,(0,0,0,1)} = x_3\ f_{(1,\,-1,\,1,\,0)} + x_4\ f_{(0,0,0,1)}\,.$$

donde el último miembro es una combinación lineal de funcionales (ver la Observación 10.8). Así, si

$$g_1 = f_{(1,\,-1,\,1,\,0)} \quad \text{y} \quad g_2 = f_{(0,0,0,1)}$$

los cálculos anteriores muestran que $\{g_1,\, g_2\}$ es un sistema de generadores de $\mathcal{W}^o$. Más aún, $B_+ = \{g_1,\, g_2\}$ es una base de $\mathcal{W}^o$ (este es un hecho general que se verifica usando la Observación 10.8: Ejercicio).

Ahora podemos concluir que $\mathcal{W}^o$ es el subespacio de $(\mathbb{R}^4)^*$ cuya base es $B_+ = \{g_1,\, g_2\}$. En particular, $\dim_{\mathbb{R}} \mathcal{W}^o = 2$.

Por ejemplo, ahora podemos concluir que $\dim_{\mathbb{R}} \mathcal{W} = 2$, por el Teorema 10.16. También podríamos haber calculado la dimensón de $\mathcal{W}$ de forma más sencilla, notando que $\{v_1,\, v_2\}$ es conjunto l.i. !

Finalmente, notemos que por la Observación 10.19 anterior, podemos escribir a $\mathcal{W}$ como el espacio de soluciones del sistema dado por $g_1(\vec{y}) = 0$ y $g_2(\vec{y}) = 0$ (asociado a la base $\{g_1,\, g_2\}$ de $\mathcal{W}^o$). Así, $\mathcal{W}$ es el espacio solución del sistema

$$\begin{aligned} 0 &= g_1(\vec{y}) = y_1 - y_2 + y_3 \\ 0 &= g_2(\vec{y}) = y_4 \end{aligned}$$

Verificar esta última afirmación. $\triangle$

10.3 El doble dual

Comenzamos definiendo un nuevo espacio asociado a un espacio vectorial.

Definición 10.21. *Sea $\mathcal{V}$ un K-ev, $\dim_K \mathcal{V} = n$. Definimos el doble dual de $\mathcal{V}$, notado $\mathcal{V}^{**}$, dado por $\mathcal{V}^{**} = (\mathcal{V}^*)^*$. Formalmente, $\mathcal{V}^{**} = L(\mathcal{V}^*, K)$ es el K-ev de las trasnformaciones lineales de $\mathcal{V}^*$ en K.* $\square$

Con las notaciones de la Definición anterior, $\mathcal{V}^{**}$ es un K-ev, tal que

$$\dim_K \mathcal{V}^{**} = \dim_K \mathcal{V}^* = n\,.$$

Obs 10.22. Sea $\mathcal{V}$ un K-ev, $\dim_K \mathcal{V} = n$. Si $v \in \mathcal{V}$ podemos definir $e_v : \mathcal{V}^* \to K$, dada por

$$e_v(f) = f(v) \quad \text{para} \quad f \in \mathcal{V}^*\,.$$

Atenti! en la definición anterior la variable es $f \in \mathcal{V}^*$. A modo de comentario notacional, a veces se escribe L_v para denotar a e_v (son sinónimos).

Notemos que e_v es la función *evaluar en v*. En este caso, $e_v \in L(\mathcal{V}^*, K) = \mathcal{V}^{**}$: en efecto, si $f,\, g \in \mathcal{V}$ y $\alpha \in K$ entonces

$$e_v(\alpha\, f + g) = (\alpha\, f + g)(v) = \alpha\, f(v) + g(v) = \alpha\, e_v(f) + e_v(g)$$

donde hemos usado la definición de la suma de $\mathcal{V}^* = L(\mathcal{V}, K)$ de transformaciones lineales en la evaluación (segunda igualdad) y la definición de e_v (en la tercer igualdad).

Así, dado $v \in \mathcal{V}$ hemos definido $e_v \in \mathcal{V}^{**}$. En el siguiente resultado estudiamos las propiedades de la función $e_{(\cdot)} : \mathcal{V} \to \mathcal{V}^{**}$. $\triangle$

Teorema 10.23. *Sea $\mathcal{V}$ un K-ev, $\dim_K \mathcal{V} = n$. Entonces la función $e_{(\cdot)} : \mathcal{V} \to \mathcal{V}^{**}$ es un isomorfismo entre $\mathcal{V}$ y $\mathcal{V}^{**}$.*

Demostración. Primero, veamos que $e_{(\cdot)}$ es lineal: para eso, sean u, $v \in \mathcal{V}$ y $\alpha \in K$. Si $f \in \mathcal{V}^*$, entonces

$$e_{\alpha\, u+v}(f) = f(\alpha\, u + v) = \alpha\, f(u) + f(v) = \alpha\, e_u(f) + e_v(f) = (\alpha\, e_u + e_v)(f)$$

donde usamos la definición de la evaluación (en $\alpha\, u + v \in \mathcal{V}$), la linealidad de f, y la definición de la evaluación de $\alpha\, e_u + e_v \in L(\mathcal{V}^*, K)$ en el vector $f \in \mathcal{V}^*$ (del dominio). Si miramos el primer y último miembro, vemos que $e_{\alpha\, u+v}(f) = (\alpha\, e_u + e_v)(f)$ para todo $f \in \mathcal{V}^*$. Esto indica la igualdad (entre funciones)

$$e_{\alpha\, u+v} = \alpha\, e_u + e_v$$

que muestra que $e_{(\cdot)}$ es transformación lineal.

Como $\dim_K \mathcal{V} = \dim_K \mathcal{V}^{**} = n$, para probar que $e_{(\cdot)}$ es isomorfismo, basta ver que es monomorfismo. Para verificar que $e_{(\cdot)}$ es monomorfismo, basta ver que $\mathrm{N}(e_{(\cdot)}) = \{0\}$. Así, sea $v \in \mathcal{V}$ tal que $e_v = 0 \in \mathcal{V}^{**} = L(\mathcal{V}^*, K)$. En este caso,

$$0 = e_v(f) = f(v) \quad \text{ para todo } \quad f \in \mathcal{V}^* \,.$$

Sea $B = \{v_1, \ldots, v_n\}$ base de $\mathcal{V}$ y sea $B^* = \{f_1, \ldots, f_n\}$ la base dual (base de $\mathcal{V}^*$). Por la Proposición 10.10, sabemos que

$$v = \sum_{j=1}^{n} f_j(v)\ v_j = \sum_{j=1}^{n} e_v(f_j)\ v_j = \sum_{j=1}^{n} 0\ v_j = 0$$

de forma que $v = 0$ lo que implica que $\mathrm{N}(e_{(\cdot)}) = \{0\}$. Por el argumento previo, vemos que $e_{(\cdot)}$ es isomorfismo. $\square$

Corolario 10.24. *Sea $\mathcal{V}$ un K-ev, $\dim_K \mathcal{V} = n$. Entonces*

1. *Si v, $w \in \mathcal{V}$ son tales que $f(v) = f(w)$ para todo $f \in \mathcal{V}^*$ entonces $v = w$.*

2. *Si $\varphi \in \mathcal{V}^{**}$ entonces existe un único $v \in \mathcal{V}$ tal que $e_v = \varphi$ es decir, $\varphi(f) = f(v)$ para todo $f \in \mathcal{V}^*$.*

Demostración. El ítem 1 es consecuencia de la inyectividad de $e_{(\cdot)}$: en efecto, si v, $w \in \mathcal{V}$ son tales que $f(v) = f(w)$ para todo $f \in \mathcal{V}^*$ entonces $e_v = e_w$ (verificar!). Pero como $e_{(\cdot)}$ es inyectiva, concluimos que $v = w$.

El ítem 2 es consecuencia directa de la suryectividad de $e_{(\cdot)}$ (ejercicio). $\square$

Corolario 10.25. *Sea $\mathcal{V}$ un K-ev, $\dim_K \mathcal{V} = n$. Si $B' = \{f_1, \ldots, f_n\}$ es base de $\mathcal{V}^*$ entonces existe $B = \{v_1, \ldots, v_n\}$ base de $\mathcal{V}$ tal que su base dual es $B^* = B'$.*

Más aún, en este caso se tiene que $(B')^* = \{e_{v_1}, \ldots, e_{v_n}\} \subset \mathcal{V}^{**}$ es la base dual de B'.

Demostración. Si consideramos a $\mathcal{V}^*$ como K-ev, entonces su espacio dual es $(\mathcal{V}^*)^* = \mathcal{V}^{**}$, por definición. Así, como B' es base de $\mathcal{V}^*$, entonces podemos considerar su base dual $(B')^* = \{\varphi_1, \ldots, \varphi_n\}$ que es base de $(\mathcal{V}^*)^* = \mathcal{V}^{**}$. En este caso, las relaciones de dualidad se escriben como

$$\varphi_j(f_i) = \delta_{ij} \quad \text{ para } \quad 1 \leq i,\, j \leq n \,.$$

Por el Corolario 10.24, para cada $1 \leq j \leq n$ existe un único $v_j \in \mathcal{V}$ tal que $\varphi_j = e_{v_j}$: notemos que en este caso $e_{(\cdot)}^{-1} = T : \mathcal{V}^{**} \to \mathcal{V}$ (la transformación inversa de $e_{(\cdot)}$) es un isomorfismo: además

$T(\varphi_j) = v_j$ (pues $e_{v_j} = \varphi_j$) para $1 \leq j \leq n$. Así, si definimos $B = \{v_1, \ldots, v_n\}$, entonces B es una base de $\mathcal{V}$, porque se obtiene de transformar una base de $\mathcal{V}^{**}$ mediante el isomorfismo T!

Finalmente, notemos que si $1 \leq i, j \leq n$ entonces

$$f_i(v_j) = e_{v_j}(f_i) = \varphi_j(f_i) = \delta_{ij}$$

lo que muestra que $\{f_1, \ldots, f_n\} = B^*$ es la base dual de B (ver el enunciado del Teorema 77). Además, por construcción $(B')^* = \{\varphi_1, \ldots, \varphi_n\} = \{e_{v_1}, \ldots, e_{v_n}\}$. □

Definición 10.26. *Sea $\mathcal{V}$ un K-ev, $\dim_K \mathcal{V} = n$. Si $S \subseteq \mathcal{V}$ es subconjunto, definimos el doble anulador de S, notado S^{oo}, dado por $(S^o)^o \subseteq \mathcal{V}^{**}$.* □

Con la notación de la definición anterior, notemos que $S^o \subseteq \mathcal{V}^*$ por definición. Para obtener el doble anulador de S, calculamos el anulador de $S^o \subseteq \mathcal{V}^*$, es decir $(S^o)^o \subseteq \mathcal{V}^{**}$. Notemos que por construcción $(S^o)^o \subseteq \mathcal{V}^{**}$ resulta siempre un subespacio.

Teorema 10.27. *Sea $\mathcal{V}$ un K-ev, $\dim_K \mathcal{V} = n$. Si $S \subseteq \mathcal{V}$ es subconjunto entonces*

$$(S^o)^o = \{e_v \ : \ v \in \overline{S}\} \subset \mathcal{V}^{**} .$$

Demostración. Recordemos que $S^o = (\overline{S})^o$: así, si $v \in \overline{S}$ y $f \in S^o$ entonces $e_v(f) = f(v) = 0$, pues $f \in (\overline{S})^o$. Así,

$$\{e_v \ : \ v \in \overline{S}\} \subseteq (S^o)^o = S^{oo} .$$

Sea $k = \dim_K \overline{S}$: entonces $\dim_K S^o = \dim_K (\overline{S})^o = n - k$, por el Teorema 10.16. Así, $\dim_K S^{oo} = n - (n - k) = k$, nuevamente por el Teorema 10.16.

Para concluir la prueba, notemos que

$$\{e_v \ : \ v \in \overline{S}\} \subset \mathcal{V}^{**}$$

es un subespacio de $\mathcal{V}^{**}$ (ejercicio). Más aún, si definimos $T : \overline{S} \to \{e_v \ : \ v \in \overline{S}\}$ dado por $T(v) = e_v$, $v \in \overline{S}$, entonces T es un isomorfimo entre $\overline{S}$ y $\{e_v \ : \ v \in \overline{S}\}$, ya que es la restricción de un isomorfismo a un subespacio de su dominio y a su imagen (ejercicio). Lo anterior muestra que $k = \dim_K \overline{S} = \dim_K \{e_v \ : \ v \in \overline{S}\}$. Como ya hemos visto que $\{e_v \ : \ v \in \overline{S}\} \subset S^{oo}$ y que $\dim_K S^{oo} = k$, concluimos que $\{e_v \ : \ v \in \overline{S}\} = S^{oo}$. □

Obs 10.28. Sea $\mathcal{V}$ un K-ev, $\dim_K \mathcal{V} = n$. Entonces hemos definido $\mathcal{V}^*$ y $\mathcal{V}^{**}$ que son K-ev tales que $\dim_K \mathcal{V} = \dim_K \mathcal{V}^* = \dim_K \mathcal{V}^{**}$.

Las igualdades de dimensiones anteriores muestran que los tres espacios son isomorfos: de hecho, para construir un isomorfismo entre cualesquiera dos de estos espacios, basta tomar una base de cada espacio y construir la tranformación lineal que transforma el j-ésimo elemento de una de estas bases en el j-ésimo elemento de la otra, para $1 \leq j \leq n$. Entonces: porqué estudiar el isomorfismo $e_{(\cdot)} : \mathcal{V} \to \mathcal{V}^{**}$ con tanto esmero??

Sucede que $e_{(\cdot)} : \mathcal{V} \to \mathcal{V}^{**}$ es un isomorfismo especial. El carácter especial de $e_{(\cdot)}$ se puede apreciar en el hecho de que lo hemos construido sin usar bases!! Por el contrario, hemos hecho uso de las relaciones existentes entre los objetos para poder definirlo, sin hacer uso de bases ni coordenadas. Si $v \in \mathcal{V}$ y $f \in \mathcal{V}^*$ podemos calcular $f(v)$, por la naturaleza de la relación entre v y f. Es por esto que $e_{(\cdot)}$ es llamado el isomorfismo canónico entre $\mathcal{V}$ y $\mathcal{V}^{**}$.

△

10.4 Traspuesta de una transformación lineal

Comenzamos definiendo el siguiente concepto nuevo.

Definición 10.29. *Sean $\mathcal{V}$, $\mathcal{W}$ K-ev's tales que* $\dim_K \mathcal{V} = n$*,* $\dim_K \mathcal{W} = m$*. Dada $T \in L(\mathcal{V}, \mathcal{W})$ definimos la traspuesta de T, notadada $T^t : \mathcal{W}^* \to \mathcal{V}^*$, definida como sigue: si $g \in \mathcal{W}^*$ entonces*

$$T^t(g) \in \mathcal{V}^* \quad \textit{dado} \quad T^t(g)(v) = g(Tv) \in K \ , \quad \textit{para} \quad v \in \mathcal{V}\,.$$

□

Obs 10.30. Consideremos la notación de la definición anterior: verifiquemos que $T^t(g) : \mathcal{V} \to K$ es de hecho lineal (es decir, que $T^t(g) \in \mathcal{V}^*$, como afirmamos previamente): para ello, si u, $v \in \mathcal{V}$ y $\alpha \in K$ entonces

$$T^t(g)(\alpha\, u+v) = g(T(\alpha\, u+v)) = g(\alpha\, T(u)+T(v)) = \alpha\, g(T(u))+g(T(v)) = \alpha\, T^t(g)(u)+T^t(g)(v)\,.$$

Otra forma (más sencilla) de verificar que $T^t(g)$ es lineal es notar que $T^t(g) = g \circ T$ es la composición de transformaciones lineales (recordemos que hemos probado que la composición de transformaciones lineales es transformación lineal en general). De esta forma, $T^t : \mathcal{W}^* \to \mathcal{V}^*$ está bien definida como función. △

Proposición 10.31. *Sean $\mathcal{V}$, $\mathcal{W}$ K-ev's tales que* $\dim_K \mathcal{V} = n$*,* $\dim_K \mathcal{W} = m$*. Dada $T \in L(\mathcal{V}, \mathcal{W})$ entonces $T^t \in L(\mathcal{W}^*, \mathcal{V}^*)$ es transformación lineal.*

Demostración. Verificamos que T^t es lineal; para ello, sean g_1, $g_2 \in \mathcal{W}^*$ y $\alpha \in K$: entonces, si $v \in \mathcal{V}$

$$\begin{aligned} T^t(\alpha\, g_1 + g_2)(v) &= (\alpha\, g_1 + g_2)(Tv) = \alpha\, g_1(Tv) + g_2(Tv) \\ &= \alpha\, T^t(g_1)(v) + T^t(g_2)(v) = (\alpha\, T^t(g_1) + T^t(g_2))(v)\,. \end{aligned}$$

Como $v \in \mathcal{V}$ es arbitrario, entonces concluimos la igualdad entre funciones

$$T^t(\alpha\, g_1 + g_2) = \alpha\, T^t(g_1) + T^t(g_2)\,,$$

que muestra que T^t es lineal.

Otra forma (más directa) de mostrar que T^t es lineal es notar que

$$T^t(\alpha\, g_1 + g_2) = (\alpha\, g_1 + g_2) \circ T = \alpha\, g_1 \circ T + g_2 \circ T = \alpha\, T^t(g_1) + T^t(g_2)$$

donde la segunda identidad (entre composiciones) ha sido probada en notas anteriores. □

Teorema 10.32. *Sean $\mathcal{V}$, $\mathcal{W}$ K-ev's tales que* $\dim_K \mathcal{V} = n$*,* $\dim_K \mathcal{W} = m$*. Dada $T \in L(\mathcal{V}, \mathcal{W})$ entonces*

1. *$N(T^t) = (Im(T))^o$.*
2. *$rg(T^t) = rg(T)$.*
3. *$Im(T^t) = (N(T))^o$.*

Demostración. Para verificar el ítem 1, notemos que dado $g \in \mathcal{W}^*$ entonces

$$\begin{aligned} g \in \mathrm{N}(T^t) \quad &\Leftrightarrow \quad T^t(g) = 0 \in \mathcal{V}^* \ \Leftrightarrow \ g \circ T = 0 \in \mathcal{V}^* \\ &\Leftrightarrow \quad g(Tv) = 0 \quad \text{para todo} \quad v \in \mathcal{V} \ \Leftrightarrow \ g \in (\mathrm{Im}(T))^o\,. \end{aligned}$$

Las equivalencias anteriores muestran la igualdad de conjuntos en 1.

Para verificar el ítem 2, definamos $r = \mathrm{rg}(T) = \dim_K \mathrm{Im}(T)$. Como $\mathrm{Im}(T)$ es un subespacio de $\mathcal{W}$, el Teorema 10.16 muestra que

$$\dim_K (\mathrm{Im}(T))^o = m - \dim_K \mathrm{Im}(T) = m - r\,.$$

Por el ítem 1 que probamos antes, vemos que

$$\dim_K \mathrm{N}(T^t) = \dim_K (\mathrm{Im}(T))^o = m - r\,.$$

Además, como $T^t \in L(\mathcal{W}^*, \mathcal{V}^*)$ es transformación lineal, sabemos que

$$m = \dim_K \mathcal{W}^* = \dim_K \mathrm{Im}(T^t) + \dim_K \mathrm{N}(T^t)$$

$$\implies \mathrm{rg}(T^t) = \dim_K \mathrm{Im}(T^t) = m - \dim_K \mathrm{N}(T^t) = m - (m - r) = r = \mathrm{rg}(T)\,.$$

Así, vemos que $\mathrm{rg}(T^t) = \mathrm{rg}(T)$ y vale el ítem 2.

Para verificar el ítem 3, primero verificamos que $\mathrm{Im}\ (T^t) \subseteq (\mathrm{N}(T))^o$. En efecto, si $f \in \mathrm{Im}\ (T^t)$, entonces existe $g \in \mathcal{W}^*$ tal que $f = T^t(g)$: así, si $v \in \mathrm{N}(T)$ entonces

$$f(v) = T^t(g)(v) = g(Tv) = g(0) = 0$$

lo que muestra que $f(v) = 0$, para $v \in \mathrm{N}(T)$. Así, si $f \in \mathrm{Im}\ (T^t)$ entonces $f \in (\mathrm{N}(T))^o$ y vale nuestra afirmación $\mathrm{Im}\ (T^t) \subseteq (\mathrm{N}(T))^o$.

Finalmente, notemos que como $\mathrm{N}(T) \subseteq \mathcal{V}$ entonces, por el Teorema 10.16

$$\dim_K (\mathrm{N}(T))^o = n - \dim_K \mathrm{N}(T) = \dim_K \mathrm{Im}(T) = \mathrm{rg}(T) = \mathrm{rg}(T^t) = \dim_K \mathrm{Im}(T^t)\,.$$

Así, se tiene la igualdad de dimensiones $\dim_K (\mathrm{N}(T))^o = \dim_K \mathrm{Im}(T^t)$, que junto con la inclusión $\mathrm{Im}\ (T^t) \subseteq (\mathrm{N}(T))^o$ muestran que $\mathrm{Im}\ (T^t) = (\mathrm{N}(T))^o$. □

El siguiente resultado justifica el nombre que le hemos dado a la transformación T^t.

Teorema 10.33. *Sean $\mathcal{V}$, $\mathcal{W}$ K-ev's tales que $\dim_K \mathcal{V} = n$, $\dim_K \mathcal{W} = m$. Sean $B_\mathcal{V}$ y $B_\mathcal{W}$ bases de $\mathcal{V}$ y $\mathcal{W}$ respectivamente. Consideramos también las bases duales $B_\mathcal{V}^*$ y $B_\mathcal{W}^*$ de $\mathcal{V}^*$ y $\mathcal{W}^*$, respectivamente. Dada $T \in L(\mathcal{V}, \mathcal{W})$ entonces*

$$[T^t]_{B_\mathcal{V}^*, B_\mathcal{W}^*} = ([T]_{B_\mathcal{W}, B_\mathcal{V}})^t \in K^{n\times m}\,,$$

donde la expresión $([T]_{B_\mathcal{W}, B_\mathcal{V}})^t$ denota tomar traspuesta de la matriz $[T]_{B_\mathcal{W}, B_\mathcal{V}} \in K^{m\times n}$.

Demostración. Sean

$$B_\mathcal{V} = \{v_1, \ldots, v_n\} \quad , \quad B_\mathcal{W} = \{w_1, \ldots, w_m\}$$
$$B_\mathcal{V}^* = \{f_1, \ldots, f_n\} \quad , \quad B_\mathcal{W}^* = \{g_1, \ldots, g_m\}\,.$$

Además, sean

$$[T]_{B_\mathcal{W}, B_\mathcal{V}} = (a_{ij}) \in K^{m\times n} \quad \text{y} \quad [T^t]_{B_\mathcal{V}^*, B_\mathcal{W}^*} = (b_{ij}) \in K^{n\times m}$$

Por construcción de las matrices de transformaciones lineales, sabemos que

$$Tv_j = \sum_{i=1}^{m} a_{ij}\ w_i \quad \text{para} \quad 1 \le j \le n \tag{79}$$

$$T^t g_j = \sum_{i=1}^{n} b_{ij}\ f_i \quad \text{para} \quad 1 \le j \le m, \tag{80}$$

que corresponde a afirmar que la j-ésima columna de la matriz A (respectivamente B) es el vector de coordenadas $[Tv_j]_{B_\mathcal{W}}$ (respectivamente $[T^t g_j]_{B^*_\mathcal{V}}$).

Por la Proposición 10.10 (aplicada a las bases $B_\mathcal{W}$ y $B^*_\mathcal{W}$), la Eq. (79) y la unicidad de las coordenadas con respecto a una base, concluimos que

$$a_{ij} = g_i(Tv_j) = T^t(g_i)(v_j) \quad \text{para} \quad 1 \le i \le m \ , \ 1 \le j \le n \, .$$

De forma análoga, la Proposición 10.10 (aplicada a las bases $B_\mathcal{V}$ y $B^*_\mathcal{V}$), la Eq. (80) y la unicidad de las coordenadas con respecto a una base muestran que

$$b_{ij} = T^t(g_j)(v_i) \quad \text{para} \quad 1 \le i \le n \ , \ 1 \le j \le m \, .$$

Las expresiones anteriores para a_{ij} y b_{ij} nos permiten ver que (si intercambiamos los roles de i y j para a_{ij})

$$a_{ji} = T^t(g_j)(v_i) = b_{ij} \quad \text{para} \quad 1 \le i \le n \ , \ 1 \le j \le m \, .$$

Estas identidades verifican el enunciado

$$([T]_{B_\mathcal{W},B_\mathcal{V}})^t_{ij} = ([T]_{B_\mathcal{W},B_\mathcal{V}})_{ji} = a_{ji} = b_{ij} = ([T^t]_{B^*_\mathcal{V},B^*_\mathcal{W}})_{ij} \quad \text{para} \quad 1 \le i \le n \ , \ 1 \le j \le m \, .$$

□

Ejemplo 10.34. Consideramos $\mathcal{V} = \mathbb{R}^2$, $\mathcal{W} = \mathbb{R}^3$ como $\mathbb{R}$-ev, y sea $T \in L(\mathbb{R}^2, \mathbb{R}^3)$ dada por

$$T(x, y) = (x + y, y - x, -2x + y) \quad \text{para} \quad (x, y) \in \mathbb{R}^2 \, .$$

Sea $B_1 = \{e_1, e_2\}$ la base canónica de $\mathbb{R}^2$ y sea $B_2 = \{h_1, h_2, h_3\}$ la base canónica de $\mathbb{R}^3$ (usamos la letra h en vez de e para los elementos de la base canónica de $\mathbb{R}^3$ para no confundir con la base canónica de $\mathbb{R}^2$).

En este caso, podemos calcular la matriz

$$[T]_{B_2,B_1} = \begin{pmatrix} 1 & 1 \\ -1 & 1 \\ -2 & 1 \end{pmatrix} \in \mathbb{R}^{3\times 2} \, .$$

Consideremos ahora las bases duales $B_1^* = \{\pi_1, \pi_2\}$ y $B_2^* = \{\rho_1, \rho_2, \rho_3\}$ que corresponden a los funcionales que son proyecciones sobre las correspondientes coordenadas (ver Ejemplo 10.7):

$$\pi_i(x_1, x_2) = x_i \ \ , \ \ \rho_j(y_1, y_2, y_3) = y_j \quad \text{para} \quad 1 \le i \le 2 \ , \ 1 \le j \le 3 \, .$$

El Teorema anterior prueba que en este caso

$$[T^t]_{B_1^*,B_2^*} = \begin{pmatrix} 1 & -1 & -2 \\ 1 & 1 & 1 \end{pmatrix} \in \mathbb{R}^{2\times 3} \, .$$

Así, si $g \in (\mathbb{R}^3)^*$ entonces, por la Proposición 10.10,

$$g = g(h_1)\ \rho_1 + g(h_2)\ \rho_2 + g(h_3)\ \rho_3 \implies [g]_{B_2^*} = (g(h_1), g(h_2), g(h_3)) \in \mathbb{R}^3 \, .$$

Entonces

$$[T^t g]_{B_1^*} = \begin{pmatrix} 1 & -1 & -2 \\ 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} g(h_1) \\ g(h_2) \\ g(h_3) \end{pmatrix} = \begin{pmatrix} g(h_1) - g(h_2) - 2\,g(h_3) \\ \\ g(h_1) + g(h_2) + g(h_3) \end{pmatrix} .$$

Entonces, ahora podemos ver que

$$T^t g = (g(h_1) - g(h_2) - 2\,g(h_3))\ \pi_1 + (g(h_1) + g(h_2) + g(h_3))\ \pi_2 \quad \text{para} \quad g \in (\mathbb{R}^3)^* \, ,$$

que es una descripción explícita de la transformación T^t. △

11 Espacios con producto interno

En lo que sigue vamos a considerar espacios vectoriales $\mathcal{V}$ sobre el cuerpo K, donde K será o bien $\mathbb{R}$ ó bien $\mathbb{C}$.

Obs 11.1. Consideremos $\mathbb{R}^n$ como $\mathbb{R}$-ev. Este espacio tiene una geometría natural: podemos considerar vectores ortogonales; podemos definir la distancia entre vectores de forma que esta distancia satisface el Teorema de Pitágoras (que tiene que ver con triángulos cuyos catetos son vectores ortogonales!). También podemos definir el ángulo entre direcciones, etc. Todas estas nociones geométricas se pueden desarrollar en función del llamado *producto escalar* : recordemos que si $\vec{x} = (x_1, \dots, x_n)$ e $\vec{y} = (y_1, \dots, y_n)$ son vectores de $\mathbb{R}^n$ entonces el producto escalar de $\vec{x}$ y $\vec{y}$, notado $\vec{x} \cdot \vec{y} \in \mathbb{R}$ está dado por

$$\vec{x} \cdot \vec{y} = \sum_{j=1}^{n} x_j \, y_j \, .$$

En algunos casos, es conveniente *deformar* la geometría usual de $\mathbb{R}^n$ y considerar la geometría asociada a un producto escalar deformado, que viene dado por

$$\langle \vec{x} \, , \, \vec{y} \rangle_{\vec{a}} = \sum_{j=1}^{n} a_j \, x_j \, y_j \, ,$$

donde $\vec{a} = (a_1, \dots, a_n) \in \mathbb{R}^n$ tiene todas sus entradas estrictamente positivas. Esta nueva geometría derivada del producto interno $\langle \, , \, \rangle_{\vec{a}}$ ya no es intuitiva (como lo es la geometría natural de $\mathbb{R}^n$). Sin embargo, podemos considerar los conceptos de ortogonalidad, distancia, y hasta definir ángulos, basados en este nuevo producto escarlar (deformado).

En lo que sigue, vamos a dar una definición general de producto interno (que tiene como casos particulares tanto al producto escalar natural como al producto escalar deformado de más arriba) en K-ev's $\mathcal{V}$ (donde $K = \mathbb{R}$ ó $K = \mathbb{C}$) y desarrollaremos las distintas nociones geométricas asociadas a este producto interno. $\triangle$

11.1 Espacios con producto interno

Comenzamos con la noción fundamental de producto interno.

Definición 11.2. *Sea $\mathcal{V}$ un K-ev ($K = \mathbb{R}$ ó $K = \mathbb{C}$). Un producto interno en $\mathcal{V}$ es una función $\langle \cdot \, , \cdot \rangle : \mathcal{V} \times \mathcal{V} \to K$ que verifica: para todos $u, v, w \in \mathcal{V}$, $\alpha \in K$,*

1. $\langle \alpha \, u + v \, , \, w \rangle = \alpha \, \langle u \, , \, w \rangle + \langle v \, , \, w \rangle$
2. $\langle u \, , \, w \rangle = \overline{\langle w \, , \, u \rangle}$ *(conjugación compleja)*
3. $\langle v \, , \, v \rangle > 0$, *siempre que* $v \neq 0$.

$\square$

Obs 11.3. Con las notaciones de la definición anterior:

a. Usando 1.: $\langle \vec{0} \, , \, w \rangle = \langle \vec{0} + \vec{0} \, , \, w \rangle = \langle \vec{0} \, , \, w \rangle + \langle \vec{0} \, , \, w \rangle$ lo que implica que $\langle \vec{0} \, , \, w \rangle = 0$.

b. Usando 2.: $\langle w \, , \, \vec{0} \rangle = \overline{\langle \vec{0} \, , \, w \rangle} = \overline{0} = 0$.

c. En particular, $\langle \vec{0} \, , \, \vec{0} \rangle = 0$.

d. Usando 1. y 2. (dos veces)

$$\langle w\,,\,\alpha\,u+v\rangle=\overline{\langle \alpha\,u+v\,,\,w\rangle}=\overline{\alpha\,\langle u\,,\,w\rangle+\langle v\,,\,w\rangle}=\overline{\alpha}\,\overline{\langle u\,,\,w\rangle}+\overline{\langle v\,,\,w\rangle}=\overline{\alpha}\,\langle w\,,\,u\rangle+\langle w\,,\,v\rangle$$

En resumen $\langle w\,,\,\alpha\,u+v\rangle=\overline{\alpha}\,\langle w\,,\,u\rangle+\langle w\,,\,v\rangle$.

En particular, notemos que los escalares que actúan en los vectores en la segunda coordenada del producto interno salen conjugados del producto interno (ojo con esta cuestión!), es decir

$$\langle w\,,\,\alpha\,v\rangle=\overline{\alpha}\,\langle w\,,\,v\rangle\,.$$

Recordemos que si $z=a+i\,b\in\mathbb{C}$, con $a,\,b\in\mathbb{R}$ (aquí $i^2=-1$ es la unidad imaginaria) entonces: $a=\mathrm{Re}(z)\in\mathbb{R}$ es llamada la *parte real* de z y $b=\mathrm{Im}(z)\in\mathbb{R}$ es llamada la *parte imaginaria* de z; además $\overline{z}=a-i\,b$ es el conjugado de z. Recordemos que la conjugación de números complejos satisface: si $x,\,z\in\mathbb{C}$: $\overline{x+z}=\overline{x}+\overline{z}$, $\overline{x\,z}=\overline{x}\,\overline{z}$, $a^2+b^2=|z|^2=z\,\overline{z}$. Finalmente, notemos que $z\in\mathbb{R}\subset\mathbb{C}$ si y solo si $\mathrm{Im}(z)=0$ ó equivalentemente si $\overline{z}=z$.

$\triangle$

Ejemplos 11.4. Consideramos los siguientes ejemplos de producto interno:

1. En $\mathcal{V}=K^n$ ($K=\mathbb{R}$ ó $K=\mathbb{C}$) definimos: dados $\vec{x}=(x_j)$, $\vec{y}=(y_j)$,

$$\langle \vec{x}\,,\,\vec{y}\rangle=\sum_{j=1}^{n}x_j\,\overline{y_j}\in K$$

que es el llamado producto escalar usual de K^n. Verificar que esta función satisface las propiedades de un producto interno (ejercicio).

2. En $\mathcal{V}=\mathbb{R}^2$ como $\mathbb{R}$-ev, definimos: dados $\vec{x}=(x_1,x_2)$ y $\vec{y}=(y_1,y_2)$,

$$\langle \vec{x}\,,\,\vec{y}\rangle=x_1\,y_1-x_2\,y_1-x_1\,y_2+4x_2\,y_2\,.$$

Entonces $\langle\,,\,\rangle$ es un producto interno. Verificamos el ítem 3 de la Definición 11.2 (las propiedades de los ítems 1 y 2 son ejercicio para el lector): si $\vec{x}=(x_1,x_2)\neq(0,0)$ entonces

$$\langle \vec{x}\,,\,\vec{x}\rangle=x_1^2-2\,x_1\,x_2+4x_2^2=(x_1-x_2)^2+3x_2^2>0$$

(pues, si $x_2\neq 0$ entonces $(x_1-x_2)^2+3x_2^2\geq x_2^2>0$; de otra forma, $x_2=0$ y $x_1\neq 0$ lo que muestra que $(x_1-x_2)^2+3x_2^2=x_1^2>0$).

3. En $\mathcal{V}=K^{m\times n}$ definimos: dada $A=(a_{jk})$ y $B=(b_{jk})$

$$\langle A\,,\,B\rangle=\sum_{j=1}^{m}\sum_{k=1}^{n}A_{jk}\,\overline{B_{jk}}\,.$$

Verificar que esta función satisface las propiedades de un producto interno (ejercicio).

$\triangle$

Obs 11.5. Consideremos el siguiente NO ejemplo. En $\mathcal{V}=\mathbb{R}^2$ como $\mathbb{R}$-ev: dados $\vec{x}=(x_1,x_2)$ y $\vec{y}=(y_1,y_2)$ definimos

$$\langle \vec{x}\,,\,\vec{y}\rangle=x_1\,y_1-x_2\,y_1-x_1\,y_2+x_2\,y_2\,.$$

Entonces $\langle\,,\,\rangle$ NO es un producto interno.

Esta función verifica las propiedades 1 y 2 de la definición (ejercicio). Sin embargo no verifica la propiedad 3: si $\vec{x}=(x_1,x_2)$ entonces

$$\langle \vec{x}\,,\,\vec{x}\rangle=x_1^2-2\,x_1\,x_2+x_2^2=(x_1-x_2)^2$$

Así, el vector $\vec{x}=(1,1)\neq\vec{0}$ es tal que $\langle(1,1)\,,\,(1,1)\rangle=0$, que contradice la propiedad 3 de la definición. $\triangle$

Definición 11.6. *Un espacio producto interno (abreviado EPI) es un par $(\mathcal{V}, \langle\,,\,\rangle)$ formado por un K-ev (donde $K = \mathbb{R}$ ó $K = \mathbb{C}$) $\mathcal{V}$ de dimensión finita y un producto interno $\langle\,,\,\rangle : \mathcal{V} \times \mathcal{V} \to K$ en $\mathcal{V}$.* □

En lo que sigue vamos a estudiar varios aspectos de un EPI. Comenzamos definiendo la noción de forma cuadrática asociada.

Definición 11.7. *Sea $(\mathcal{V}, \langle\,,\,\rangle)$ un EPI. La forma cuadrática asociada a $\langle\,,\,\rangle$ es la función $Q : \mathcal{V} \to [0, \infty)$ dada por $Q(v) = \langle v\,,\,v\rangle$, para $v \in \mathcal{V}$.*

Proposición 11.8. *Sea $(\mathcal{V}, \langle\,,\,\rangle)$ un EPI y sea $Q(\cdot) : \mathcal{V} \to [0, \infty)$ la forma cuadrática asociada. Entonces, si $u, v \in \mathcal{V}$ y $\alpha \in K$:*

1. $Q(\alpha\, v) = |\alpha|^2\; Q(v)$;
2. $\mathrm{Re}(\langle u\,,\,v\rangle) = \frac{1}{4}\,(Q(u+v) - Q(u-v))$ *(donde $\mathrm{Re}(\beta)$ denota la parte real de $\beta \in \mathbb{C}$).*
3. *Si $K = \mathbb{C}$ se tiene la identidad de polarización:*

$$\langle u\,,\,v\rangle = \frac{1}{4}(Q(u+v) - Q(u-v) + i\,(Q(u+iv) - Q(u-iv))) = \frac{1}{4}\sum_{k=0}^{3} i^k\, Q(u + i^k\, v) \tag{81}$$

donde $i^2 = -1$.

Demostración. La prueba del ítem 1 es ejercicio.

Para verificar el ítem 2, calculamos $Q(u+v)$ usando la definición de Q y las propiedades del producto interno (incluyendo los hechos en la Observación 11.3):

$$\begin{aligned} Q(u+v) &= \langle u+v\,,\,u+v\rangle = \langle u\,,\,u+v\rangle + \langle v\,,\,u+v\rangle \\ &= (\langle u\,,\,u\rangle + \langle u\,,\,v\rangle) + (\langle v\,,\,u\rangle + \langle v\,,\,v\rangle) \\ &= Q(u) + \langle u\,,\,v\rangle + \overline{\langle u\,,\,v\rangle} + Q(v) = Q(u) + 2\;\mathrm{Re}(\langle u\,,\,v\rangle) + Q(v)\,. \end{aligned}$$

donde hemos usado que si $\beta = a + i\,b \in \mathbb{C}$ entonces $\beta + \overline{\beta} = 2\,a = 2\,\mathrm{Re}(\beta)$. En resumen, hemos probado que

$$Q(u+v) = Q(u) + 2\;\mathrm{Re}(\langle u\,,\,v\rangle) + Q(v)\,. \tag{82}$$

Si reemplazamos v por $-v$ en la identidad anterior, y notamos que $\mathrm{Re}(\langle u\,,\,-v\rangle) = \mathrm{Re}(-\langle u\,,\,v\rangle) = -\mathrm{Re}(\langle u\,,\,v\rangle)$ y que $Q(-v) = |-1|^2 Q(v) = Q(v)$, entonces obtenemos la identidad:

$$Q(u-v) = Q(u) - 2\;\mathrm{Re}(\langle u\,,\,v\rangle) + Q(v)\,. \tag{83}$$

Para confirmar la identidad anterior el lector puede hacer un desarrollo similar al hecho más arriba, pero ahora para la expresión $Q(u-v)$ - usando las propiedades del producto interno (ejercicio). Si restamos las identidades de las Eqs. (82) y (83) miembro a miembro obtenemos

$$Q(u+v) - Q(u-v) = 4\;\mathrm{Re}(\langle u\,,\,v\rangle)\,.$$

La identidad anterior es esencialmente la identidad del ítem 2. Notemos que si $K = \mathbb{R}$ entonces la identidad anterior muestra que $\langle u\,,\,v\rangle = \mathrm{Re}(\langle u\,,\,v\rangle) = \frac{1}{4}(Q(u+v) - Q(u-v))$, que prueba que el producto interno queda codificado por la forma cuadrática Q.

Para verificar el ítem 3, aplicamos la identidad del ítem 2, reemplazando v por $i\,v$ (notemos que estamos asumiendo que $K = \mathbb{C}$ y podemos actuar con la unidad imaginaria $i \in \mathbb{C}$). En este caso, por el item 2, usando que $Q(i\,v) = |i|^2\, Q(v) = Q(v)$,

$$\begin{aligned} Q(u + i\,v) &= Q(u) + 2\;\mathrm{Re}(\langle u\,,\,i\,v\rangle) + Q(v) \\ &= Q(u) + 2\;\mathrm{Re}(-i\,\langle u\,,\,v\rangle) + Q(v) \end{aligned}$$

Notemos que si $\beta = a + i\,b$ entonces $-i\,\beta = -i\,a - i^2\,b = -i\,a + b$, de forma que $\mathrm{Re}(-i\,\beta) = \mathrm{Im}(\beta)$. Si aplicamos esta observación en la identidad anterior, vemos que

$$Q(u + i\,v) = Q(u) + 2\ \mathrm{Im}(\langle u\,,\,v\rangle)) + Q(v)\,. \tag{84}$$

Si ahora reemplazamos a v por $-v$ en la identidad anterior y notamos que: $u - i\,v = u + i\,(-v)$ y $\mathrm{Im}(\langle u\,,\,-v\rangle) = \mathrm{Im}(-\langle u\,,\,v\rangle) = -\,\mathrm{Im}(\langle u\,,\,v\rangle)$ entonces

$$Q(u - i\,v) = Q(u) - 2\ \mathrm{Im}(\langle u\,,\,v\rangle)) + Q(v)\,. \tag{85}$$

El lector puede confirmar las identidades anteriores, desarrollando la expresión $Q(u+i\,v)$, $Q(u-i\,v)$ y aplicando las propiedades del producto interno (en este caso, hay que ser cuidadoso con la aparición de conjugaciones de complejos).

Si restamos las identidades de las Eqs. (84) y (85) miembro a miembro obtenemos

$$Q(u + i\,v) - Q(u - i\,v) = 4\,\mathrm{Im}(\langle u\,,\,v\rangle) \implies \mathrm{Im}(\langle u\,,\,v\rangle) = \frac{1}{4}\,(Q(u + i\,v) - Q(u - i\,v))\,.$$

Así,

$$\langle u\,,\,v\rangle = \mathrm{Re}(\langle u\,,\,v\rangle) + i\ \mathrm{Im}(\langle u\,,\,v\rangle) = \frac{1}{4}(Q(u+v) - Q(u-v) + i\,(Q(u + i\,v) - Q(u - i\,v)))\,.$$

Finalmente, la identidad

$$\frac{1}{4}(Q(u+v) - Q(u-v) + i\,(Q(u + i\,v) - Q(u - i\,v))) = \frac{1}{4}\sum_{k=0}^{3} i^k Q(u + i^k\,v)$$

se verifica expandiendo la sumatoria y comparando los términos de ambos miembros (ejercicio). Notemos que la identidad de polarización muestra que el producto interno está codificado en su forma cuadrática (en el sentido que podemos recuperar el producto interno a través de la forma cuadrática) también en el caso complejo. $\square$

Corolario 11.9. *Sea $(\mathcal{V},\,\langle\;,\;\rangle)$ un EPI y sea $Q(\cdot) : \mathcal{V} \to [0,\infty)$ la forma cuadrática asociada. Dados $u,\,v \in \mathcal{V}$, se verifica la identidad del paralelogramo:*

$$Q(u+v) + Q(u-v) = 2\,(Q(u) + Q(v))\,.$$

Demostración. La identidad es una consecuencia de las Eqs. (82) y (83). $\square$

En lo que sigue vamos a considerar la noción de longitud de un vector inducida por un producto interno.

Definición 11.10. *Sea $(\mathcal{V},\,\langle\;,\;\rangle)$ un EPI. Definimos la función $\|\cdot\| : \mathcal{V} \to [0,\infty)$ dada por*

$$\|v\| = (\langle v\,,\,v\rangle)^{1/2} = Q(v)^{1/2} \geq 0 \quad \textit{para} \quad v \in \mathcal{V}\,.$$

En este caso, $\|v\|$ es llamada la norma del vector $v \in \mathcal{V}$ (inducida por el producto interno). $\square$

Con las notaciones de la definición anterior, notemos que la norma está bien definida, en el sentido que $Q(v) = \langle v\,,\,v\rangle \geq 0$ para $v \in \mathcal{V}$; en este caso, podemos hallar la única raíz cuadrada no negativa de $Q(v)$.

Obs 11.11. Sea $(\mathcal{V}, \langle\, , \,\rangle)$ un EPI y sea Q la forma cuadrática asociada al producto interno. Entonces, por la Definición 11.10, $Q(v) = \|v\|^2$, para $v \in \mathcal{V}$. Así, podemos re-escribir algunas de las identidades que hemos deducido para Q en términos de la norma. Por ejemplo: la identidad en la Eq. (82) puede escribirse como

$$\|u+v\|^2 = \|u\|^2 + 2\ \mathrm{Re}(\langle u\,,\,v\rangle) + \|v\|^2\,. \tag{86}$$

De forma similar, la identidad en la Eq. (83) puede escribirse como

$$\|u-v\|^2 = \|u\|^2 - 2\ \mathrm{Re}(\langle u\,,\,v\rangle) + \|v\|^2\,. \tag{87}$$

La identidad de polarización en la Eq. (81) puede escribirse como

$$\langle u\,,\,v\rangle = \frac{1}{4}\sum_{k=0}^{3} i^k\,\|u +\, i^k\,v\|^2\,. \tag{88}$$

Finalmente, notemos que la identidad del paralelogramo puede escribirse como

$$\|u+v\|^2 + \|u-v\|^2 = 2\,(\|u\|^2 + \|v\|^2)\,.$$

$\triangle$

Teorema 11.12. *Sea $(\mathcal{V}, \langle\, , \,\rangle)$ un EPI. Dados $u,\ v \in \mathcal{V}$ y $\alpha \in K$ se verifica*

N1. $\|\alpha\ v\| = |\alpha|\,\|v\|$ *;*

N2. $\|v\| > 0$ *siempre que* $v \neq 0$ *(* $\|0\| = 0$ *) ;*

N3. $\|u+v\| \leq \|u\| + \|v\|$ *;*

C-S. $|\,\langle u\,,\,v\rangle\,| \leq \|u\|\ \|v\|$.

Demostración. Los items 1 y 2 son consecuencias directas de las propiedades del producto interno y de la función raíz cuadrada y quedan como ejercicio.

Veamos el cuarto ítem, conocido como la desigualdad de Cauchy-Schwarz (que hemos abreviado C-S). Si $u = \vec{0}$ entonces $\langle u\,,\,v\rangle = \langle \vec{0}\,,\,v\rangle = 0$, de forma que la desigualdad vale.

Si $u \neq 0$ entonces definimos

$$w = v - \frac{\langle v, u\rangle}{\|u\|^2}\,u \in \mathcal{V}$$

(notemos que $\|u\|^2 \neq 0$ pues $u \neq \vec{0}$). Entonces, usando la Eq. (87)

$$\begin{aligned}
0 \ &\leq\ \langle w\,,\,w\rangle = \|w\|^2 = \|v - \frac{\langle v, u\rangle}{\|u\|^2}\,u\|^2 = \|v\|^2 - 2\ \mathrm{Re}(\langle v\,,\,\frac{\langle v, u\rangle}{\|u\|^2}\,u\rangle) + \|\frac{\langle v, u\rangle}{\|u\|^2}\,u\|^2 \\
&=\ \|v\|^2 - 2\ \mathrm{Re}(\frac{\overline{\langle v\,,\,u\rangle}}{\|u\|^2}\ \langle v\,,\,u\rangle) + |\frac{\langle v, u\rangle}{\|u\|^2}|^2\,\|u\|^2 \\
&=\ \|v\|^2 - 2\ \frac{|\langle v\,,\,u\rangle|^2}{\|u\|^2} + \frac{|\langle v\,,\,u\rangle|^2}{\|u\|^2} = \|v\|^2 - \ \frac{|\langle v\,,\,u\rangle|^2}{\|u\|^2}
\end{aligned}$$

donde hemos utilizado las propiedades del producto interno y de la norma. Así, hemos probado que

$$0 \leq \|v\|^2 - \ \frac{|\langle v\,,\,u\rangle|^2}{\|u\|^2} \implies |\langle v\,,\,u\rangle|^2 \leq \|u\|^2\,\|v\|^2\,.$$

Como la función raíz cuadrada es una función creciente en $[0, \infty)$, entonces tomando raíz cuadrada a ambos miembros se preserva la desigualdad:

$$|\langle v\,,\,u\rangle|^2 \leq \|u\|^2\,\|v\|^2 \implies |\langle v\,,\,u\rangle| \leq \|u\|\,\|v\|\,.$$

Para probar el ítem 3, que es llamada la *desigualdad triangular* para la norma, comenzamos utilizando la identidad en la Eq. (86):

$$\begin{aligned}\|u+v\|^2 &= \|u\|^2 + 2\,\mathrm{Re}(\langle u\,,\,v\rangle) + \|v\|^2 \\ &\leq \|u\|^2 + 2\,|\langle u\,,\,v\rangle| + \|v\|^2 \\ &\leq \|u\|^2 + 2\,\|u\|\,\|v\| + \|v\|^2 = (\|u\| + \|v\|)^2\,.\end{aligned}$$

donde hemos usado que si $\alpha = a + i\,b \in \mathbb{C}$ entonces $\mathrm{Re}(\alpha) = a \leq (a^2+b^2)^{1/2} = |\alpha|$, con $\alpha = \langle u\,,\,v\rangle$ para la primer desigualdad y la desigualdad de Cauchy-Schwarz (probada más arriba) para la segunda desigualdad. Así,

$$0 \leq \|u+v\|^2 \leq (\|u\|+\|v\|)^2 \implies \|u+v\| \leq \|u\| + \|v\|\,.$$

□

11.2 Ortogonalidad en espacios con producto interno

Comenzamos con la noción fundamental de ortogonalidad (ó perpendicularidad) en un espacio con producto interno.

Definición 11.13. *Sea $(\mathcal{V}, \langle\,,\,\rangle)$ un EPI. Sean $u, v \in \mathcal{V}$ y $S = \{v_1, \ldots, v_k\} \subset \mathcal{V}$. Decimos que:*

1. *u y v son ortogonales, notado $u \perp v$, si $\langle u\,,\,v\rangle = 0$.*
2. *S es sistema ortogonal si $\langle v_h\,,\,v_j\rangle = 0$, para $1 \leq h \neq j \leq k$.*
3. *S es sistema ortonormal si S es sistema ortogonal tal que $\|v_j\| = 1$, para $1 \leq j \leq k$.*

□

Obs 11.14. Sea $(\mathcal{V}, \langle\,,\,\rangle)$ un EPI. Entonces

1. Si $u \in \mathcal{V}$ entonces $u \perp \vec{0}$ (pues $\langle u\,,\,\vec{0}\rangle = 0$).
2. Si $v \in \mathcal{V}$ es tal que $v \perp u$ para todo $u \in \mathcal{V}$ entonces $v = \vec{0}$: en efecto, si elegimos $u = v$ entonces $0 = \langle v\,,\,v\rangle$ lo que implica que $v = \vec{0}$.

△

Ejemplos 11.15. Sea $K = \mathbb{R}$ ó $K = \mathbb{C}$. Entonces

1. La base canónica $B = \{e_1, \ldots, e_n\}$ es un sistema ortonormal con respecto al producto interno usual de K^n. En este caso, como B es base, decimos que B es una base ortonormal de K^n.
2. Sea $\mathcal{V} = F^{m\times n}$ y consideremos el producto interno de los Ejemplos 11.4. Si $1 \leq p \leq m$ y $1 \leq q \leq n$ entonces sea $E^{(p,q)} \in K^{m\times n}$ tal que $E^{(p,q)}_{h,k} = \delta_{ph}\,\delta_{qk}$ para $1 \leq h \leq m$ y $1 \leq k \leq n$. Entonces $B = \{E^{(p,q)}\ :\ 1 \leq p \leq m\ ,\ 1 \leq q \leq n\ \}$ es un sistema ortonormal. Como B también es base de $K^{m\times n}$ decimos que B es base ortonormal de $K^{m\times n}$.

△

Teorema 11.16. *Sea $(\mathcal{V}, \langle\, , \,\rangle)$ un EPI, $\dim_K \mathcal{V} = n$. Si $S = \{v_1, \ldots, v_k\}$ es un sistema ortogonal formado por vectores no nulos entonces S es l.i. En particular, $k \leq n$.*

Demostración. Sean $\alpha_1, \ldots, \alpha_k \in K$ tales que $\sum_{j=1}^{k} \alpha_j\ v_j = \vec{0}$. Si $1 \leq h \leq k$ entonces

$$0 = \langle \vec{0}\,,\, v_h \rangle = \langle \sum_{j=1}^{k} \alpha_j\ v_j\,,\, v_h \rangle = \sum_{j=1}^{k} \alpha_j\ \langle v_j\,,\, v_h \rangle = \alpha_h\, \langle v_h\,,\, v_h \rangle$$

donde hemos usado que $\langle v_j\,,\, v_h \rangle = 0$ si $1 \leq j \neq h \leq k$. En resumen $0 = \alpha_h\, \langle v_h\,,\, v_h \rangle$. Como $v_h \neq \vec{0}$ entonces $\langle v_h\,,\, v_h \rangle > 0$, lo que muestra que $\alpha_h = 0$. Como $1 \leq h \leq k$ era arbitrario, hemos probado que $\alpha_h = 0$ para $1 \leq h \leq k$ y S es l.i. □

Teorema 11.17. *Sea $(\mathcal{V}, \langle\, , \,\rangle)$ un EPI y sea $\{u_1, \ldots, u_m\} \subset \mathcal{V}$ conjunto l.i. Entonces existe un sistema ortogonal formado por vectores no nulos $\{w_1, \ldots, w_m\} \subset \mathcal{V}$ tales que: si $1 \leq j \leq m$ entonces los subespacios*

$$\overline{\{u_1, \ldots, u_j\}} = \overline{\{w_1, \ldots, w_j\}}\,.$$

Demostración. Vamos a construir los vectores $w_1, \ldots, w_m$ recursivamente, de la siguiente forma: definimos $w_1 =: u_1$. Notemos que $w_1 \neq \vec{0}$ y $\{w_1\}$ es sistema ortogonal tal que $\overline{\{u_1\}} = \overline{\{w_1\}}$.

Supongamos que para un cierto $1 \leq j \leq m-1$, hemos definidos $w_1, \ldots, w_j$, tales que

1. $\{w_1, \ldots, w_j\}$ es sistema ortogonal de vectores no nulos;
2. $\overline{\{u_1, \ldots, u_j\}} = \overline{\{w_1, \ldots, w_j\}}$.

Entonces definimos

$$w_{j+1} = u_{j+1} - \sum_{h=1}^{j} \frac{\langle u_{j+1}\,,\, w_h \rangle}{\|w_h\|^2}\ w_h\,.$$

Notemos que este procedimiento está bien planteado: de hecho, como $w_h \neq \vec{0}$ entonces $\|w_h\|^2 > 0$ y podemos dividir por este escalar, para $1 \leq h \leq j$.

Veamos que $w_{j+1} \neq \vec{0}$: en efecto, si $w_{j+1} = \vec{0}$ entonces

$$u_{j+1} = \sum_{h=1}^{j} \frac{\langle u_{j+1}\,,\, w_h \rangle}{\|w_h\|^2}\ w_h \in \overline{\{w_1, \ldots, w_j\}} = \overline{\{u_1, \ldots, u_j\}}\,,$$

donde la pertenencia es por definición de subespacio generado, y hemos usado la igualdad de subespacios de nuestra hipótesis (ítem 2.). Así, existen $\alpha_1, \ldots, \alpha_j \in K$ tales que

$$u_{j+1} = \sum_{h=1}^{j} \alpha_h\ u_h \implies \sum_{h=1}^{j} \alpha_h\ u_h - u_{j+1} = \vec{0}$$

que contradice la hipótesis de que $\{u_1, \ldots, u_m\}$ son l.i. del enunciado (dado que hemos encontrado escalares $\alpha_1, \ldots,\, \alpha_j,\, -1$ no todos nulos, pues el escalar que acompaña a u_{j+1} es $-1 \neq 0$, tales que la combinación lineal correspondiente es nula). Esta contradicción surje de suponer que $w_{j+1} = \vec{0}$: así vemos que $w_{j+1} \neq \vec{0}$.

Veamos que $\{w_1, \ldots, w_{j+1}\}$ es sistema ortogonal: en este caso, usando el ítem 1, basta ver que $\langle w_{j+1}\,,\, w_\ell \rangle = 0$, para $1 \leq \ell \leq j$ (porque los w_h's para $1 \leq h \leq j$, ya son ortogonales entre sí, por el ítem 1). Entonces, calculamos

$$\begin{aligned}
\langle w_{j+1}\,,\,w_\ell\rangle &= \langle u_{j+1}-\sum_{h=1}^{j}\frac{\langle u_{j+1}\,,\,w_h\rangle}{\|w_h\|^2}\;w_h\,,\,w_\ell\rangle = \langle u_{j+1}\,,\,w_\ell\rangle - \langle\sum_{h=1}^{j}\frac{\langle u_{j+1}\,,\,w_h\rangle}{\|w_h\|^2}\;w_h\,,\,w_\ell\rangle\\
&= \langle u_{j+1}\,,\,w_\ell\rangle - \sum_{h=1}^{j}\frac{\langle u_{j+1}\,,\,w_h\rangle}{\|w_h\|^2}\;\langle\,w_h\,,\,w_\ell\rangle = \langle u_{j+1}\,,\,w_\ell\rangle - \frac{\langle u_{j+1}\,,\,w_\ell\rangle}{\|w_\ell\|^2}\;\langle w_\ell\,,\,w_\ell\rangle\\
&= \langle u_{j+1}\,,\,w_\ell\rangle - \langle u_{j+1}\,,\,w_\ell\rangle = 0
\end{aligned}$$

donde hemos usado las propiedades del producto interno, que $\langle w_h\,,\,w_\ell\rangle = 0$ siempre que $1\leq h\neq \ell\leq j$, y que $\|w_\ell\|^2=\langle w_\ell\,,\,w_\ell\rangle$. Lo anterior muestra que $\{w_1,\ldots,w_{j+1}\}$ es un sistema ortogonal formado por vectores no nulos. En particular, el Teorema 11.16, concluimos que $\{w_1,\ldots,w_{j+1}\}$ es l.i.

Finalmente, veamos que

$$\overline{\{u_1,\ldots,u_{j+1}\}}=\overline{\{w_1,\ldots,w_{j+1}\}}$$

En efecto: por definición

$$w_{j+1}\in\overline{\{w_1,\ldots,\,w_j,\,u_{j+1}\}}=\overline{\{u_1,\ldots,\,u_j,\,u_{j+1}\}}$$

donde en la igualdad anterior hemos usado la hipótesis del ítem 2 (verificar). Entonces, por el ítem 2,

$$w_1,\ldots,\,w_j,\,w_{j+1}\subset\overline{\{u_1,\ldots,\,u_j,\,u_{j+1}\}}\implies\overline{\{w_1,\ldots,\,w_j,\,w_{j+1}\}}\subseteq\overline{\{u_1,\ldots,\,u_j,\,u_{j+1}\}}\,.$$

Finalmente, como hemos probado que $\{w_1,\ldots,w_{j+1}\}$ es l.i. (y $\{u_1,\ldots,u_{j+1}\}$ es l.i. por hipótesis) vemos que

$$\dim_K\overline{\{w_1,\ldots,\,w_j,\,w_{j+1}\}}=\dim_K\overline{\{u_1,\ldots,\,u_j,\,u_{j+1}\}}=j+1$$

implica que $\overline{\{w_1,\ldots,\,w_j,\,w_{j+1}\}}=\overline{\{u_1,\ldots,\,u_j,\,u_{j+1}\}}$.

Finalmente, notamos que $w_1,\ldots,w_{j+1}$ cumplen con los análogos de los ítems 1 y 2 de más arriba, de forma que si $j+1<m$ podemos repetir el proceso: en particular, construimos w_{j+2} de forma que $w_1,\ldots,w_{j+2}$ también cumplen con los análogos de los ítems 1 y 2 de más arriba. Continuando este proceso completamos la construcción de los vectores $w_1,\ldots,w_m$ que verifican el teorema. □

Obs 11.18. Consideremos la notación del teorema anterior. Notemos que la prueba de este resultado indica un procedimiento para construir los vectores $w_1,\ldots,\,w_m$: en efecto, hemos definido $w_1=u_1$; en este caso construimos

$$w_2=u_2-\frac{\langle u_2,w_1\rangle}{\|w_1\|^2}\;w_1\,,$$

y podemos entonces construir (usando w_1, w_2 construidos previamente)

$$w_3=u_3-\Big(\frac{\langle u_3,w_1\rangle}{\|w_1\|^2}\;w_1+\frac{\langle u_3,w_2\rangle}{\|w_2\|^2}\;w_2\Big)\,.$$

En general: si asumimos que hemos construido $w_1,\ldots,w_j$ entonces podemos construir

$$w_{j+1}=u_{j+1}-\sum_{h=1}^{j}\frac{\langle u_{j+1}\,,\,w_h\rangle}{\|w_h\|^2}\;w_h\,.$$

Debe quedar claro que es posible continuar con este proceso, construyendo $w_1,\ldots,w_m$, que verifican las condiciones del Teorema 11.17.

Este proceso es conocido como el *método de ortogonalización de Gram-Schmidt.* △

Ejemplo 11.19. Consideremos $\mathbb{R}^3$ con el producto interno usual, y sea $B = \{u_1, u_2, u_3\}$ la base formada por

$$u_1 = (3,0,4)\ ,\ u_2 = (-1,0,7)\ ,\ u_3 = (2,9,11)\,.$$

Aplicamos el proceso de ortogonalización de Gram-Schmidt: Definimos $w_1 = u_1 = (3,0,4)$;

$$w_2 = u_2 - \frac{\langle u_2\,,\,w_1\rangle}{\|w_1\|^2}\ w_1 = (-1,0,7) - \frac{25}{25}\ (3,0,4) = (-4,0,3)$$

y finalmente,

$$\begin{aligned} w_3 &= u_3 - (\frac{\langle u_3\,,\,w_1\rangle}{\|w_1\|^2}\ w_1 + \frac{\langle u_3\,,\,w_2\rangle}{\|w_2\|^2}\ w_2) \\ &= (2,9,11) - (\frac{50}{25}\,(3,0,4) + \frac{25}{25}\ (-4,0,3)\ ) = (0,9,0)\,. \end{aligned}$$

Así

$$w_1 = (3,0,4)\ ,\ w_2 = (-4,0,3)\ ,\ w_3 = (0,9,0)$$

y $\{w_1, w_2, w_3\}$ es el sistema ortogonal (formado por vectores no nulos: es decir $\langle w_j\,,\,w_k\rangle = 0$ si $1 \leq j \neq k \leq 3$: verificar, ejercicio) obtenido de $\{u_1, u_2, u_3\}$ por el método de ortogonalización de Gram-Schmidt. En particular

$$\overline{\{u_1,\ldots,u_j\}} = \overline{\{w_1,\ldots,w_j\}} \quad \text{con} \quad j = 1,2,3\,.$$

$\triangle$

Obs 11.20. Sea $\mathcal{V}$ un K-ev, sean $v_1,\ldots,v_m \in \mathcal{V}$ y sean $\alpha_1,\ldots,\alpha_m \in K$ escalares *no nulos*. Entonces, se puede verificar que

$$\overline{\{v_1,\ldots,v_m\}} = \overline{\{\alpha_1\,v_1,\ldots,\alpha_m\,v_m\}}\,.$$

Ejercicio (sugerencia: usar las propiedades de subespacio generado por un conjunto de vectores y el hecho de que $\alpha_j \neq 0$ para $1 \leq j \leq m$). $\triangle$

Teorema 11.21. *Sea $(\mathcal{V}, \langle\,,\,\rangle)$ un EPI, $\dim_K \mathcal{V} = n$, y sea $\mathcal{W} \subset \mathcal{V}$ un subespacio tal que $1 \leq \dim_K \mathcal{W} = k$. Entonces existe un base ortonormal de $\mathcal{V}$ (es decir, una base que es un conjunto ortonormal) $B = \{v_1,\ldots,v_n\}$ tal que $\mathcal{W} = \overline{\{v_1,\ldots,v_k\}}$.*

Demostración. Sea $\{u_1,\ldots,u_k\}$ una base de $\mathcal{W}$; en este caso, podemos econtrar vectores $u_{k+1},\ldots,u_n \in \mathcal{V}$ tales que $\{u_1,\ldots,u_n\}$ es base de $\mathcal{V}$. Sea $\{w_1,\ldots,w_n\}$ el sistema ortogonal de vectores no nulos que resulta de aplicar el proceso de ortogonalización de Gram-Schmidt a $\{u_1,\ldots,u_n\}$. En particular,

$$\mathcal{W} = \overline{\{u_1,\ldots,u_k\}} = \overline{\{w_1,\ldots,w_k\}}\,.$$

Finalmente, definimos $v_j = \frac{1}{\|w_j\|}\, w_j$ para $1 \leq j \leq n$. Entonces $\{v_1,\ldots,v_n\}$ es un sistema ortonormal: en efecto, usando las propiedades del producto interno:

$$\langle v_h\,,\,v_\ell\rangle = \frac{1}{\|w_h\|\ \|w_\ell\|}\langle w_h\,,\,w_\ell\rangle = \delta_{h\,\ell}$$

pues $\langle w_h\,,\,w_\ell\rangle = 0$, si $h \neq \ell$ y $\langle w_h\,,\,w_h\rangle = \|w_h\|^2$. En particular, usando la Observación 11.20, ahora vemos que

$$\mathcal{W} = \overline{\{w_1,\ldots,w_k\}} = \overline{\{v_1,\ldots,v_k\}}\,.$$

Finalmente, notemos que $B = \{v_1,\ldots,v_n\}$ es conjunto l.i. por el Teorema 11.16, y B tiene $n = \dim_K \mathcal{V}$ elementos, de forma que es una base de $\mathcal{V}$. $\square$

Corolario 11.22. *Sea* $(\mathcal{V}, \langle\, , \,\rangle)$ *un EPI,* $\dim_K \mathcal{V} = n$. *Entonces existe una base ortonormal* $\{v_1, \ldots, v_n\}$ *de* $\mathcal{V}$.

Demostración. Consideramos el caso $\mathcal{W} = \mathcal{V}$ del teorema anterior. □

Obs 11.23. Sea $(\mathcal{V}, \langle\, , \,\rangle)$ un EPI, $\dim_K \mathcal{V} = n$. Aparte de la noción *algebraica* de dimensión (es decir, la noción de dimensión que hemos considerado en el curso), podemos considerar la noción *geométrica* de dimensión en $\mathcal{V}$. En este caso, la dimensión geométrica corresponde a la máxima cantidad de direcciones mutuamente ortogonales que podemos encontrar en $\mathcal{V}$, es decir, la cantidad máxima de elementos que tiene un sistema ortonormal (en el sentido de la Definición 11.13).

Por un lado, como todo sistema ortonormal es un sistema ortogonal de vectores no nulos, entonces el Teorema 11.16 indica que la dimensión geométrica es menor ó igual que la dimensión algebraica. Por otro lado, el Corolario 11.22 muestra que en realidad estas nociones de dimensión coinciden. △

Los siguientes resultados muestran algunas de las ventajas que tienen las bases ortogonales (y más aún las bases ortonormales) para el cálculo de coordenadas y de normas de vectores.

Proposición 11.24. *Sea* $(\mathcal{V}\, , \langle\, , \,\rangle)$ *un EPI,* $\dim_K \mathcal{V} = n$. *Sea* $B = \{v_1, \ldots, v_n\}$ *una base ortogonal de* $\mathcal{V}$. *Si* $u \in \mathcal{V}$ *entonces*

$$u = \sum_{j=1}^{n} \frac{\langle u\, , \, v_j\rangle}{\|v_j\|^2}\, v_j \qquad y \qquad \|u\|^2 = \sum_{j=1}^{n} \frac{|\langle u\, , \, v_j\rangle|^2}{\|v_j\|^2}\, .$$

Demostración. Para verificar la primer identidad, notemos primero que deben existir $\alpha_1, \ldots, \alpha_n \in K$ tales que $u = \sum_{j=1}^{n} \alpha_j\ v_j$, pues B es base de $\mathcal{V}$. Si $1 \le h \le n$ entonces

$$\langle u\, , \, v_h\rangle = \langle \sum_{j=1}^{n} \alpha_j\ v_j\, , \, v_h\rangle = \sum_{j=1}^{n} \alpha_j\ \langle v_j\, , \, v_h\rangle = \alpha_h\, \langle v_h\, , \, v_h\rangle = \alpha_h\ \|v_h\|^2$$

lo que implica que $\alpha_h = \frac{\langle u\, , \, v_h\rangle}{\|v_h\|^2}$. Si reemplazamos a α_h por $\frac{\langle u\, , \, v_h\rangle}{\|v_h\|^2}$ en la representación lineal de u con respecto a B obtenemos la primer identidad.

Para verificar la segunda identidad, utilizamos la representación anterior de u y las propiedades del producto interno para calcular:

$$\|u\|^2 = \langle u\, , \, u\rangle = \langle \sum_{j=1}^{n} \frac{\langle u\, , \, v_j\rangle}{\|v_j\|^2}\, v_j\, , \sum_{h=1}^{n} \frac{\langle u\, , \, v_h\rangle}{\|v_h\|^2}\, v_h\rangle = \sum_{j=1}^{n} \frac{\langle u\, , \, v_j\rangle}{\|v_j\|^2} \sum_{h=1}^{n} \frac{\overline{\langle u\, , \, v_h\rangle}}{\|v_h\|^2}\, \underbrace{\langle v_j\, , \, v_h\rangle}_{\delta_{jh}\ \|v_j\|^2}$$

$$= \sum_{j=1}^{n} \frac{\langle u\, , \, v_j\rangle}{\|v_j\|^2}\, \overline{\langle u\, , \, v_j\rangle} = \sum_{j=1}^{n} \frac{|\langle u\, , \, v_j\rangle|^2}{\|v_j\|^2}$$

donde hemos usado que los escalares que actúan en los vectores de las segundas coordenadas en el producto interno salen conjugados, y que $\langle v_j\, , \, v_h\rangle = \delta_{jh}\ \|v_j\|^2$ para $1 \le j\, , \, k \le n$. □

Corolario 11.25. *Sea* $(\mathcal{V}\, , \langle\, , \,\rangle)$ *un EPI,* $\dim_K \mathcal{V} = n$. *Sea* $B = \{v_1, \ldots, v_n\}$ *una base ortonormal de* $\mathcal{V}$ *(de forma que* $\|v_j\| = 1$, $1 \le j \le n$*). Si* $u \in \mathcal{V}$ *entonces*

$$u = \sum_{j=1}^{n} \langle u\, , \, v_j\rangle\, v_j \qquad y \qquad \|u\|^2 = \sum_{j=1}^{n} |\langle u\, , \, v_j\rangle|^2\, .$$

□

Obs 11.26. Sea $(\mathcal{V}\,,\,\langle\,,\,\rangle)$ un EPI, $\dim_K \mathcal{V} = n$. Entonces la norma $\|\cdot\|$ inducida por el producto interno satisface las propiedades $N1$, $N2$, $N3$ y $C-S$ del Teorema 11.12. En particular, estas propiedades garantizan que si definimos la *distancia* entre los vectores u y v, notada $d(u,v)$, dada por

$$d(u,v) = \|u-v\| \quad \text{para} \quad u,\, v \in \mathcal{V}$$

entonces la distancia satisface las propiedades

D1. $d(u,v) \geq 0$ para $u,\, v \in \mathcal{V}$; $d(u,v) = 0$ si y solo si $u = v$.

D2. $d(u,v) = d(v,u)$ para $u,\, v \in \mathcal{V}$.

D3. $d(u,v) \leq d(u,w) + d(w,v)$ para $u,\, v,\, w \in \mathcal{V}$.

La propiedad $D1$. es de no degeneración; la propiedad $D2$. es de simetría y la propiedad $D3$ es la llamada desigualdad triangular para la distancia. Además, como la métrica proviene de una norma, se verifican ciertas propiedades de invarianza y homogeneidad:

I. $d(u+w, v+w) = d(u,v)$ para $u,\, v,\, w \in \mathcal{V}$.

H. $d(\alpha\, u, \alpha\, v) = |\alpha|\; d(u,v)$ para $u,\, v \in \mathcal{V}$.

Estas son propiedades fundamentales (son todas ejercicio!) que permiten hacer muchos argumentos de naturaleza métrica (y hasta cierto punto geométrica) muy importantes. El hecho de que nuestra métrica provenga de un producto interno agrega otra propiedad fundamental: la identidad de Pitágoras

$$d(u,v) = d(u,0) + d(v,0) \quad \text{siempre que} \quad u \perp v\,.$$

$\triangle$

Obs 11.27 (El problema de cuadrados mínimos (ó la distancia a un subespacio))**.** Sea $(\mathcal{V}\,,\,\langle\,,\,\rangle)$ un EPI (en principio no pedimos que $\mathcal{V}$ sea de dimensión finita). Sea $\mathcal{W} \subseteq \mathcal{V}$ un subespacio y $v \in \mathcal{V}$. El *problema de cuadrados mínimos* asociados a $(\mathcal{W}\,,\, v)$ consiste en encontrar (si es que existe) un $w_0 \in \mathcal{W}$ que minimice la distancia a v, entre todos los vectores de $\mathcal{W}$, es decir:

$$\|v - w_0\| \leq \|v - w\| \quad \text{para todo} \quad w \in \mathcal{W}\,. \tag{89}$$

En este caso, decimos que la distancia de v al subespacio $\mathcal{W}$ es $\|v - w_0\|$. Podemos pensar que $w_0 \in \mathcal{W}$ es la *mejor aproximación* de $v \in \mathcal{V}$ mediante vectores de $\mathcal{W}$. $\triangle$

Teorema 11.28. *Sea $(\mathcal{V}\,,\,\langle\,,\,\rangle)$ un EPI. Sea $\mathcal{W} \subseteq \mathcal{V}$ un subespacio y $v \in \mathcal{V}$. Entonces:*

1. $w_0 \in \mathcal{W}$ satisface la Eq. (89) *si y solo si $(v - w_0) \perp w$ para todo $w \in \mathcal{W}$.*

2. Existe a lo sumo una solución del problema de minimización en la Eq. (89)*.*

3. Si $\dim_K \mathcal{W} = n \in \mathbb{N}$ y $\{u_1, \ldots, u_n\}$ es una base ortonormal de $\mathcal{W}$ entonces

$$w_0 = \sum_{j=1}^{n} \langle v\,,\, u_j\rangle\; u_j \in \mathcal{W}$$

es la única solución al problema de minimización de la Eq. (89)*.*

Demostración. Sean $w_0,\, w \in \mathcal{W}$ arbitrarios; notemos primero que $v - w = v - w_0 + w_0 - w$ de forma que, por la Eq. (86)

$$\|v - w\|^2 = \|(v - w_0) + (w_0 - w)\|^2 = \|v - w_0\|^2 + 2\ \operatorname{Re}(\langle v - w_0, w_0 - w\rangle) + \|w_0 - w\|^2\,. \tag{90}$$

Veamos el ítem 1: supongamos que $(v - w_0) \perp w$ para todo $w \in \mathcal{W}$. Entonces, como $w_0 - w \in \mathcal{W}$ vemos que

$$\langle v - w_0, w_0 - w\rangle = 0\,.$$

Si usamos este hecho en la Eq. (90) concluimos que

$$\|v - w\|^2 = \|(v - w_0) + (w_0 - w)\|^2 = \|v - w_0\|^2 + \|w_0 - w\|^2 \geq \|v - w_0\|^2\,,$$

dado que $\|w_0 - w\|^2 \geq 0$. Esto muestra que w_0 minimiza la distancia a $v \in \mathcal{V}$ entre todos los vectores $w \in \mathcal{W}$ es decir, satisface la Eq. (89).

Recíprocamente, si w_0 minimiza la distancia a v entre todos los vectores de $\mathcal{W}$ entonces, la Eq. (90) muestra que en este caso se debe verificar que

$$2\ \mathrm{Re}(\langle v - w_0, w_0 - w\rangle) + \|w_0 - w\|^2 \geq 0 \quad \text{para todo} \quad w \in \mathcal{W}\,.$$

Notemos que si $u \in \mathcal{W}$ es arbitrario, entonces si definimos $w = w_0 - u \in \mathcal{W}$, se tiene que $w_0 - w = u$. Así, podemos reemplazar $w_0 - w$ (con $w \in \mathcal{W}$ arbitrario) por un vector arbitrario $u \in \mathcal{W}$ en la desigualdad anterior y concluir que bajo nuestras hipótesis se verifica:

$$2\ \mathrm{Re}(\langle v - w_0, u\rangle) + \|u\|^2 \geq 0 \quad \text{para todo} \quad u \in \mathcal{W}\,. \tag{91}$$

Sea $u \in \mathcal{W}$, $u \neq \vec{0}$, fijo: definimos la función $f : \mathbb{R} \to \mathbb{R}$ asociada a u, dada por

$$f(t) = 2\ \mathrm{Re}(\langle v - w_0, t\,u\rangle) + \|t\,u\|^2 = 2\ t\ \mathrm{Re}(\langle v - w_0, u\rangle) + t^2\|u\|^2 \geq 0 \quad \text{para} \quad t \in \mathbb{R}\,.$$

Notemos que hemos usado que: si $t \in \mathbb{R}$ y $\alpha \in \mathbb{C}$ entonces $\mathrm{Re}(t\,\alpha) = t\,\mathrm{Re}(\alpha)$, en la identidad anterior. La desigualdad que aparece en la ecuacion de arriba se deriva de la Eq. (91) dado que $t\,u \in \mathcal{W}$ para $t \in \mathbb{R}$. Podemos re-escribir

$$f(t) = A\,t^2 + B\,t \geq 0 \quad \text{con} \quad A = \|u\|^2 > 0 \quad \text{y} \quad B = 2\ \mathrm{Re}(\langle v - w_0, u\rangle) \in \mathbb{R}\,.$$

La función $f(t)$ es una parábola con ramas hacia arriba (pues $A > 0$). La condición $f(t) \geq 0$ para todo $t \in \mathbb{R}$ se verifica si y solo si el discriminante de la cuadrática (notar que en este caso $C = 0$) verifica $B^2 - 4A\,0 \leq 0$ (en este caso la cuadrática tiene a lo sumo una raíz en el eje real, y está siempre por arriba del eje x - hacer dibujo de la parábola correspondiente!). Entonces $B^2 \leq 0$ lo que implica que $B = \mathrm{Re}(\langle v - w_0, u\rangle) = 0$ para todo $u \in \mathcal{W}$, $u \neq \vec{0}$; por supuesto que $\mathrm{Re}(\langle v - w_0, u\rangle) = 0$ si $u = \vec{0}$, por propiedades del producto interno.

El párrafo anterior muestra que

$$\mathrm{Re}(\langle v - w_0, u\rangle) = 0 \quad \text{para todo} \quad u \in \mathcal{W}\,. \tag{92}$$

Si $K = \mathbb{R}$, es decir, $\mathcal{V}$ es un $\mathbb{R}$-ev, entonces $\mathrm{Re}(\langle x\,,\, z\rangle) = \langle x\,,\, z\rangle$ para todo $x,\, z \in \mathcal{V}$, porque el producto interno siempre toma valores en el cuerpo (en este caso $K = \mathbb{R}$). Así, la identidad anterior muestra que $\langle v - w_0, u\rangle = 0$ para todo $u \in \mathcal{W}$, como queríamos probar.

Si $K = \mathbb{C}$, es decir, $\mathcal{V}$ es un $\mathbb{C}$-ev, entonces dado $u \in \mathcal{W}$ podemos considerar $i\,u \in \mathcal{W}$ (donde $i^2 = -1$) y concluir usando la Eq. (92) que

$$0 = \mathrm{Re}(\langle v - w_0, i\,u\rangle) = \mathrm{Re}(-i\,\langle v - w_0, u\rangle) = \mathrm{Im}(\langle v - w_0, u\rangle)\,, \tag{93}$$

donde hemos usado que $\bar{i} = -i$ (el conjugado complejo de i es $-i$), las propiedades del producto interno (los escalares que actúan en la segunda coordenada salen conjugados) y que $\mathrm{Re}(-i\,\alpha) = \mathrm{Im}(\alpha)$ para todo $\alpha \in \mathbb{C}$. Así, las Eqs. (92) y (93) implican que

$$\langle v - w_0,\, u\rangle = \mathrm{Re}(\langle v - w_0,\, u\rangle) + i\ \mathrm{Im}(\langle v - w_0,\, u\rangle) = 0 \quad \text{para todo} \quad u \in \mathcal{W}$$

como queríamos demostrar. Los argumentos anteriores muestran la equivalencia afirmada en el ítem 1.

Para verificar el ítem 2, supongamos que w_0, $w_1 \in \mathcal{W}$ minimizan la distancia de v a $\mathcal{W}$, es decir:

$$\|v - w_0\| \leq \|v - w\| \ , \ \|v - w_1\| \leq \|v - w\| \quad \text{para todo} \quad w \in \mathcal{W} \, .$$

El ítem 1 que hemos probado nos permite concluir que

$$\langle v - w_0 \, , \, w \rangle = 0 \quad \text{y} \quad \langle v - w_1 \, , \, w \rangle = 0 \quad \text{para todo} \quad w \in \mathcal{W}$$

En particular, si $w \in \mathcal{W}$ es arbitrario:

$$\langle w_1 - w_0 \, , \, w \rangle = \langle (w_1 - v) + (v - w_0) \, , \, w \rangle = \langle (w_1 - v) \, , \, w \rangle + \langle (v - w_0) \, , \, w \rangle = 0 \tag{94}$$

donde hemos usado que

$$\langle (w_1 - v) \, , \, w \rangle = \langle -(v - w_1) \, , \, w \rangle = -\langle (v - w_1) \, , \, w \rangle = 0 \, .$$

Si elegimos $w = w_1 - w_0 \in \mathcal{W}$ en la Eq. (94), vemos que

$$0 = \langle w_1 - w_0 \, , \, w \rangle = \langle w_1 - w_0 \, , \, w_1 - w_0 \rangle \implies w_1 - w_0 = \vec{0}$$

es decir, $w_0 = w_1$ y vale 2.

Finalmente, supongamos que $\{u_1, \ldots, u_n\}$ es base ortonormal de $\mathcal{W}$ y definamos

$$w_0 = \sum_{j=1}^{n} \langle v \, , \, u_j \rangle \ u_j \in \mathcal{W} \, .$$

Para verificar que w_0 es solucion al problema de cuadrados mínimos de la Eq. (89), basta ver que $\langle v - w_0 \, , \, w \rangle = 0$ para todo $w \in \mathcal{W}$, por el ítem 1 (que probamos antes). Para verificar este hecho, comenzamos notando que si $1 \leq h \leq n$ entonces

$$\begin{aligned} \langle v - w_0 \, , \, u_h \rangle &= \langle v - (\sum_{j=1}^{n} \langle v \, , \, u_j \rangle \ u_j \,) \, , \, u_h \rangle = \langle v \, , \, u_h \rangle - \langle \sum_{j=1}^{n} \langle v \, , \, u_j \rangle \ u_j \, , \, u_h \rangle \\ &= \langle v \, , \, u_h \rangle - \sum_{j=1}^{n} \langle v \, , \, u_j \rangle \ \langle u_j \, , \, u_h \rangle = \langle v \, , \, u_h \rangle - \langle v \, , \, u_h \rangle \ \langle u_h \, , \, u_h \rangle = 0 \end{aligned}$$

donde hemos usado que $\langle u_j \, , \, u_h \rangle = 0$ si $1 \leq j \neq h \leq n$ y que $\langle u_h \, , \, u_h \rangle = \|u_h\|^2 = 1$. Finalmente, si $w \in \mathcal{W}$ entonces existen $\alpha_1, \ldots, \alpha_n \in K$ tales que $w = \sum_{h=1}^{n} \alpha_h \ u_h$; en este caso

$$\langle v - w_0 \, , \, w \rangle = \langle v - w_0 \, , \, \sum_{h=1}^{n} \alpha_h \ u_h \rangle = \sum_{h=1}^{n} \overline{\alpha_h} \ \langle v - w_0 \, , \, u_h \rangle = \sum_{h=1}^{n} \overline{\alpha_h} \ 0 = 0$$

donde hemos usado la identidad $\langle v - w_0 \, , \, u_h \rangle = 0$ para $1 \leq h \leq n$, probada más arriba. Así, w_0 definido como antes es la solución del problema de cuadrados mínimos asociado a $(v \, , \, \mathcal{W})$; por el ítem 2, w_0 es la única solucion de este problema. □

11.3 Complementos ortogonales

Hay un concepto geométrico fundamental detrás de la solución al problema de cuadrados mínimos así como detrás del método de ortogonalización de Gram-Schmidt que está relacionado con el llamado complemento ortogonal y las proyecciones ortogonales asociadas a ciertas descomposiciones de un espacio con producto interno. Para describir estas relaciones, comenzamos con el concepto de complemento ortogonal.

Definición 11.29. *Sea $(\mathcal{V}\,,\,\langle\,,\,\rangle)$ un EPI, $\dim_K \mathcal{V} = n$. Dado $S \subset \mathcal{V}$ un subconjunto definimos el complemento ortogonal de S, notado $S^\perp \subset \mathcal{V}$, dado por*

$$S^\perp = \{v \in \mathcal{V} \ : \ v \perp w \text{ para todo } w \in S\} \subseteq \mathcal{V}\,.$$

□

Obs 11.30. Sea $(\mathcal{V}\,,\,\langle\,,\,\rangle)$ un EPI, $\dim_K \mathcal{V} = n$. Entonces

$$\mathcal{V}^\perp = \{\vec{0}\} \quad \text{y} \quad \{\vec{0}\}^\perp = \mathcal{V}$$

Verificar (ejercicio). △

Proposición 11.31. *Sea $(\mathcal{V}\,,\,\langle\,,\,\rangle)$ un EPI, $\dim_K \mathcal{V} = n$. Dado $S \subseteq \mathcal{V}$ se verifica:*

1. *$S^\perp \subset \mathcal{V}$ es subespacio;*
2. *Si $S \subset T \subset \mathcal{V}$ entonces $T^\perp \subset S^\perp$.*
3. *$S^\perp = (\overline{S})^\perp$.*

Demostración. Para verificar el ítem 1, usamos las propiedades del producto interno. Por un lado, queda claro que $\vec{0} \in S^\perp$ (ejercicio). Sean $u,\, v \in S^\perp$ y $\alpha \in K$ entonces, si $w \in S$ vale que $\langle u\,,\,w\rangle = \langle v\,,\,w\rangle = 0$, por hipotesis (pues $u \perp w$ y $v \perp w$). Luego,

$$\langle \alpha\, u + v\,,\,w\rangle = \alpha\,\langle u\,,\,w\rangle + \langle v\,,\,w\rangle = 0\,.$$

Lo anterior muestra que $\alpha\ u + v \in S^\perp$ de forma que $S^\perp$ es subespacio.

El ítem 2 es ejercicio para el lector.

Para verificar el ítem 3, notemos que por el ítem 2 tenemos que

$$S \subset \overline{S} \implies (\overline{S})^\perp \subset S^\perp\,.$$

Veamos la otra inclusión: sea $v \in S^\perp$ y sea $w \in \overline{S}$; entonces existen $w_1, \ldots, w_m \in S$ y escalares $\alpha_1, \ldots, \alpha_m \in K$ tales que $w = \sum_{j=1}^m \alpha_j\, w_j$. Por hipótesis, $\langle w_j\,,\,v\rangle = 0$ para $1 \leq j \leq m$ (pues $v \in S^\perp$). Así, ahora vemos que

$$\langle w\,,\,v\rangle = \langle \sum_{j=1}^m \alpha_j\, w_j\,,\,v\rangle = \sum_{j=1}^m \alpha_j\ \underbrace{\langle w_j\,,\,v\rangle}_{=0} = 0\,.$$

Como $w \in \overline{S}$ era arbitrario, vemos que $\langle w\,,\,v\rangle = 0$ para todo $w \in \overline{S}$ lo que muestra que $v \in (\overline{S})^\perp$. Esto muestra $S^\perp \subseteq (\overline{S})^\perp$ y finalmente que $S^\perp = (\overline{S})^\perp$. □

Teorema 11.32. *Sea $(\mathcal{V}\,,\,\langle\,,\,\rangle)$ un EPI, $\dim_K \mathcal{V} = n$. Dado $\mathcal{W} \subseteq \mathcal{V}$ subespacio, se tiene que*

$$\mathcal{V} = \mathcal{W} \dot{+} \mathcal{W}^\perp\,.$$

En particular, $n = \dim_K \mathcal{W} + \dim_K \mathcal{W}^\perp$.

Demostración. Si $\mathcal{W} = \mathcal{V}$ entonces $\mathcal{W}^\perp = \{\vec{0}\}$ y el resultado vale trivialmente. De forma similar, si $\mathcal{W} = \{\vec{0}\}$ entonces $\mathcal{W}^\perp = \mathcal{V}$ y el resultado vale trivialmente también.

Así, podemos suponer que $\mathcal{W}$ es un subespacio propio, es decir $1 \leq k = \dim_K \mathcal{W} \leq n-1$. En este caso, consideramos el Teorema 11.21, que muestra que podemos contruir una base ortonormal $B = \{v_1, \ldots, v_n\}$ de $\mathcal{V}$ tal $B_\mathcal{W} = \{v_1, \ldots, v_k\}$ es base (ortonormal) de $\mathcal{W}$.

En lo que sigue verificamos que $\{v_{k+1}, \ldots, v_n\}$ es base de $\mathcal{W}^\perp$. Por un lado, si $1 \leq j \leq k$ y $k+1 \leq h \leq n$ entonces

$$\langle v_j \,,\, v_h \rangle = 0 \implies v_h \in \{v_1, \ldots, v_k\}^\perp = \left(\overline{\{v_1, \ldots, v_k\}}\right)^\perp = \mathcal{W}^\perp$$

donde hemos usado el ítem 3 de la Proposición 11.31 y el hecho de que $\{v_1, \ldots, v_k\}$ es base de $\mathcal{W}$ (en realidad, solo usamos que genera a $\mathcal{W}$). Como $\mathcal{W}^\perp \subseteq \mathcal{V}$ es subespacio (ver Proposición 11.31) concluimos que

$$\overline{\{v_{k+1}, \ldots, v_n\}} \subset \mathcal{W}^\perp \,. \tag{95}$$

Veamos que en realidad vale la identidad en la inclusión anterior, lo que muestra en particular que $B_{\mathcal{W}^\perp} = \{v_{k+1}, \ldots, v_n\}$ es base de $\mathcal{W}^\perp$. En efecto, sea $v \in \mathcal{W}^\perp$: como $v \in \mathcal{V}$ entonces, por el Corolario 11.25 sabemos que

$$v = \sum_{j=1}^{n} \langle v \,,\, v_j \rangle \, v_j \implies v = \sum_{j=k+1}^{n} \langle v \,,\, v_j \rangle \, v_j$$

donde hemos usado que $\langle v \,,\, v_j \rangle = 0$ para $1 \leq j \leq k$, ya que $v \in \mathcal{W}^\perp$ y $v_j \in \mathcal{W}$ para $1 \leq j \leq k$. Lo anterior muestra que todo vector $v \in \mathcal{W}^\perp$ es combinación lineal de $B_{\mathcal{W}^\perp} = \{v_{k+1}, \ldots, v_n\}$; este hecho, junto con la inclusión en la Eq. (95) muestran que vale la igualdad

$$\overline{\{v_{k+1}, \ldots, v_n\}} = \mathcal{W}^\perp \,.$$

Además, como $B_{\mathcal{W}^\perp}$ es l.i. (porque es un subconjunto de una base), ahora podemos afirmar que es una base de $\mathcal{W}^\perp$. Por un resultado previo, como el pegado de las bases de $\mathcal{W}$ y $\mathcal{W}^\perp$ que hemos hallado

$$B_\mathcal{W} \cup B_{\mathcal{W}^\perp} = B$$

es una base de $\mathcal{V}$ (que es la base con la que comenzamos!) concluimos que los subespacios verifican $\mathcal{V} = \mathcal{W} \dot{+} \mathcal{W}^\perp$. Finalmente, por resultados previos concluimos que en este caso se verifica $\dim_K \mathcal{W} + \dim_K \mathcal{W}^\perp = \dim_K \mathcal{V}$. □

Corolario 11.33. *Sea $(\mathcal{V} \,,\, \langle \,,\, \rangle)$ un EPI, $\dim_K \mathcal{V} = n$. Dado $S \subseteq \mathcal{V}$, se tiene que*

$$(S^\perp)^\perp = \overline{S} \,.$$

En particular, si $\mathcal{W} \subseteq \mathcal{V}$ es subespacio entonces $(\mathcal{W}^\perp)^\perp = \mathcal{W}$.

Demostración. Sea $\dim_K \overline{S} = k$; como $S^\perp = (\overline{S})^\perp$ por la Proposición 11.31, entonces $\dim_K S^\perp = n - k$, por el Teorema 11.32. De forma similar, concluimos que $\dim_K (S^\perp)^\perp = n - (n-k) = k$. Además, notemos que si $v \in \overline{S}$ y $u \in S^\perp = (\overline{S})^\perp$ entonces $\langle v \,,\, u \rangle = 0$; como $u \in S^\perp$ era arbitrario, concluimos que $v \in (S^\perp)^\perp$, por definición de complemento ortogonal. El argumento anterior muestra que si $v \in \overline{S}$ entonces $v \in (S^\perp)^\perp$: en resumen, hemos probado que

$$\overline{S} \subset (S^\perp)^\perp \quad \text{y} \quad \dim_K \overline{S} = k = \dim_K (S^\perp)^\perp \implies \overline{S} = (S^\perp)^\perp \,.$$

□

Obs 11.34. Sea $(\mathcal{V}\,,\,\langle\,,\,\rangle)$ un EPI, $\dim_K \mathcal{V} = n$. Sea $\mathcal{W} \subseteq \mathcal{V}$ subespacio.

1. Hemos probado que en este caso $\mathcal{V} = \mathcal{W} \dot{+} \mathcal{W}^\perp$. Para enfatizar el hecho de que $\mathcal{W}$ y $\mathcal{W}^\perp$ son subespacios ortogonales entre si, introducimos la notación
$$\mathcal{V} = \mathcal{W} \oplus \mathcal{W}^\perp\,.$$

2. Supongamos que $1 \leq k = \dim_K \mathcal{W} \leq n-1$; notemos que podemos restringir el producto interno a los vectores de $\mathcal{W}$ y considerar $(\mathcal{W}\,,\,\langle\,,\,\rangle)$ como EPI de dimensión k. En este caso, podemos construir una base ortonormal $B_\mathcal{W} = \{w_1,\,\ldots,w_k\}$ de $\mathcal{W}$.

   De forma similar, considerando el EPI $(\mathcal{W}^\perp\,,\,\langle\,,\,\rangle)$ de dimensión $n-k$, podemos construir una base ortonormal $B_{\mathcal{W}^\perp} = \{u_1,\,\ldots,u_{n-k}\}$ de $\mathcal{W}^\perp$.

   Entonces la base
$$B = B_\mathcal{W} \cup B_{\mathcal{W}^\perp} = \{w_1,\ldots,w_k,\,u_1,\ldots,u_{n-k}\}$$
   es base ortonormal de $\mathcal{V}$. En efecto, por un lado sabemos que el pegado de bases de subespacios que forman una descomposición en suma directa de $\mathcal{V}$ es una base de $\mathcal{V}$. Además, B es un sistema ortonormal: si elegimos dos vectores distintos de B y consideramos su producto interno, entonces se da alguno de los siguientes casos:

   (a) Los dos vectores están en $B_\mathcal{W}$: entonces son ortogonales, porque $B_\mathcal{W}$ es sistema ortogonal;
   (b) Los dos vectores están en $B_{\mathcal{W}^\perp}$: entonces son ortogonales, porque $B_{\mathcal{W}^\perp}$ es sistema ortogonal;
   (c) Uno de los vectores está en $B_\mathcal{W}$ y el otro en $B_{\mathcal{W}^\perp}$: entonces son ortogonales, porque $B_\mathcal{W} \subset \mathcal{W}$ y $B_{\mathcal{W}^\perp} \subset \mathcal{W}^\perp$!

   Además, todos los vectores de B tienen norma uno (porque $B_\mathcal{W}$ y $B_{\mathcal{W}^\perp}$ son sistemas ortonormales).

3. Recordemos que dada la descomposición $\mathcal{V} = \mathcal{W} \oplus \mathcal{W}^\perp$ entonces existe un único sistema de proyecciones $\{E_1,\,E_2\}$ asociado a esta descomposición, de forma que E_1 y E_2 son proyecciones ($E_j^2 = E_j$, $j = 1,2$) tales que
$$I = E_1 + E_2 \quad , \quad E_1\,E_2 = E_2\,E_1 = 0 \quad , \ \mathrm{Im}(E_1) = \mathcal{W} \quad , \quad \mathrm{Im}(E_2) = \mathcal{W}^\perp\,.$$
   De hecho, dado $v \in \mathcal{V}$ existen únicos $w \in \mathcal{W}$ y $u \in \mathcal{W}^\perp$ tales que $v = w + u$ y en este caso se tiene que $E_1(v) = w$ y $E_2(v) = u$.

   Sean $B_\mathcal{W} = \{w_1,\,\ldots,w_k\}$ y $B_{\mathcal{W}^\perp} = \{u_1,\,\ldots,u_{n-k}\}$ bases ortonormales de $\mathcal{W}$ y $\mathcal{W}^\perp$, respectivamente (como antes). Por el ítem 2 de arriba, $B = B_\mathcal{W} \cup B_{\mathcal{W}^\perp}$ es base ortonormal de $\mathcal{V}$. Así, dado $v \in \mathcal{V}$ el Corolario 11.25 muestra que
$$v = \sum_{j=1}^{k} \langle v\,,\,w_j\rangle\; w_j + \sum_{j=1}^{n-k} \langle v\,,\,u_j\rangle\; u_j\,.$$
   Si definimos
$$w = \sum_{j=1}^{k} \langle v\,,\,w_j\rangle\; w_j \in \mathcal{W} \quad , \quad u = \sum_{j=1}^{n-k} \langle v\,,\,u_j\rangle\; u_j \in \mathcal{W}^\perp \qquad \text{entonces} \qquad v = w + u$$

Lo anterior indica que entonces

$$E_1(v) = \sum_{j=1}^{k} \langle v \, , \, w_j \rangle \; w_j \in \mathcal{W} \quad \text{ y } \quad E_2(v) = \sum_{j=1}^{n-k} \langle v \, , \, u_j \rangle \; u_j \in \mathcal{W}^{\perp} \, .$$

En este caso, la proyección E_1 es llamada la *proyección ortogonal sobre* $\mathcal{W}$ y notada $P_{\mathcal{W}}$; por otro lado, la proyección E_2 es llamada la *proyección ortogonal sobre* $\mathcal{W}^{\perp}$ y notada $P_{\mathcal{W}^{\perp}}$.

Como en todo sistema de dos proyecciones, se verifica que $P_{\mathcal{W}} + P_{\mathcal{W}^{\perp}} = I$; en particular, $P_{\mathcal{W}^{\perp}} = I - P_{\mathcal{W}}$ y valen la fórmulas

$$P_{\mathcal{W}} v = \sum_{j=1}^{k} \langle v \, , \, w_j \rangle \, w_j \quad \text{ y } \quad P_{\mathcal{W}^{\perp}} v = v - \big(\sum_{j=1}^{k} \langle v \, , \, w_j \rangle \, w_j \, \big) \quad \text{para} \quad v \in \mathcal{V} \, .$$

4. Si suponemos que $B_{\mathcal{W}} = \{w_1, \ldots, w_k\}$ es una base ortogonal (es decir, que en principio no verifica que la norma de los vectores w_j sea 1) entonces

$$\{ \frac{1}{\|w_1\|} \, w_1, \, \ldots \, , \frac{1}{\|w_k\|} \, w_k \}$$

es base ortonormal de $\mathcal{W}$ (verificar!) y podemos aplicar las observaciones del ítem 3 para esta base ortonormal: en particular,

$$P_{\mathcal{W}} \, v = \sum_{j=1}^{k} \langle v \, , \, \frac{1}{\|w_j\|} \, w_j \rangle \, \frac{1}{\|w_j\|} \, w_j = \sum_{j=1}^{k} \frac{\langle v \, , \; w_j \rangle}{\|w_j\|^2} \, w_j \quad , \quad P_{\mathcal{W}^{\perp}} \, v = v - \big(\sum_{j=1}^{k} \frac{\langle v \, , \; w_j \rangle}{\|w_j\|^2} \, w_j \, \big) \, .$$

5. Sobre el proceso de ortogonalización de Gram-Schmidt (ver Teorema 11.17 y su prueba): vamos a utilizar la notación del Teorema 11.17 : en este caso, $\{u_1, \, \ldots \, , u_m\}$ denota un conjunto l.i. (arbitrario) de $\mathcal{V}$.

   A partir de este conjunto, aplicamos el método de ortogonalización de Gram-Schmidt como antes: construimos $w_1 = u_1$; y si asumimos construidos $w_1, \ldots, w_j$ (que forma un sistema ortogonal de vectores no nulos) entonces

$$w_{j+1} = u_{j+1} - \sum_{h=1}^{j} \frac{\langle u_{j+1} \, , \, w_h \rangle}{\|w_h\|^2} \; w_h = P_{\mathcal{W}_j^{\perp}} \, u_{j+1}$$

   donde $\mathcal{W}_j = \overline{\{w_1, \ldots, w_j\}}$; es decir, w_{j+1} se construye como la proyección ortogonal sobre el complemento ortogonal de $\mathcal{W}_j$ del vector u_{j+1} (en donde hemos usado la identidad

$$P_{\mathcal{W}_j^{\perp}} \, v = v - \sum_{h=1}^{j} \frac{\langle v \, , \, w_h \rangle}{\|w_h\|^2} \; w_h \quad \text{para} \quad v \in \mathcal{V}$$

   del ítem 4 para calcular $P_{\mathcal{W}_j^{\perp}} \, u_{j+1}$, usando la base ortogonal $\{w_1, \ldots, w_j\}$ de $\mathcal{W}_j$, construida con los vectores obtenidos por el método de ortogonalización, en los pasos anteriores).

6. Sobre el problema de mínimos cuadrados: notemos que con la notación del Teorema 11.28, el vector que minimiza la distancia de $v \in \mathcal{V}$ al subespacio $\mathcal{W}$ está dado por $w_0 = P_{\mathcal{W}} \, v$, la proyección ortogonal de v sobre el subespacio $\mathcal{W}$ (en donde hemos utilizado la identidad del ítem 3 para calcular $P_{\mathcal{W}} \, v$ en términos de una base ortonormal de $\mathcal{W}$).

$\triangle$

Corolario 11.35. *Sea* $(\mathcal{V}, \langle\, ,\, \rangle)$ *un EPI,* $\dim_K \mathcal{V} = n$*. Si* $\{v_1, \ldots, v_k\} \subset \mathcal{V}$ *es un sistema ortogonal de vectores no nulos (respectivamente sistema ortonormal) con* $1 \leq k < n$*, entonces existen vectores* $v_{k+1}, \ldots, v_n \in \mathcal{V}$ *tales que* $\{v_1, \ldots, v_n\}$ *es base ortogonal (respectivamente ortonormal) de* $\mathcal{V}$*.*

Demostración. Por el Teorema 11.16 el conjunto $\{v_1, \ldots, v_k\}$ es l.i. y entonces es una base del subespacio

$$\mathcal{W} = \overline{\{v_1, \ldots, v_k\}}\,.$$

Sea $\{v_{k+1}, \ldots, v_n\}$ una base ortonormal de $\mathcal{W}^\perp$. Por el Teorema 11.32 y resultados previos (sobre el pegado de bases de subespacios independientes) $B = \{v_1, \ldots, v_n\}$ es una base de $\mathcal{V}$. Argumentando como en el ítem 2 de la Observación 11.34, concluimos que B es sistema ortogonal, de forma que B es una base ortogonal de $\mathcal{V}$. Si $\{v_1, \ldots, v_k\}$ era sistema ortonormal, entonces B es base ortonormal (verificar, ejercicio). $\square$

11.4 Funcionales en espacios con producto interno

En lo que sigue, dado $\mathcal{V}$ un K-ev de dimensión finita (donde $K = \mathbb{R}$ ó $K = \mathbb{C}$) entonces $\mathcal{V}^* = L(\mathcal{V}\,,\,K)$ denota el espacio dual del espacio $\mathcal{V}$.

Obs 11.36. Sea $(\mathcal{V}, \langle\, ,\, \rangle)$ un EPI, $\dim_K \mathcal{V} = n$. Dado $u \in \mathcal{V}$ entonces u determina un funcional lineal $f_u \in \mathcal{V}^*$ dado por

$$f_u(v) = \langle v\,,\,u\rangle \quad \text{para} \quad v \in \mathcal{V}\,.$$

Notemos que las propiedades de linealidad de f_u son una consecuencia de las propiedades de linealidad del producto interno con respecto a la primer coordenada (notar que u - que está fijo - aparece en la segunda coordenada del producto interno en la definición de f_u). En lo que sigue verificamos que todo funcional es de esta forma. $\triangle$

Teorema 11.37 (de Representación de Riesz)**.** *Sea* $(\mathcal{V}, \langle\, ,\, \rangle)$ *un EPI,* $\dim_K \mathcal{V} = n$*. Dado* $f \in \mathcal{V}^*$ *existe un único vector* $u \in \mathcal{V}$ *tal que*

$$f(v) = \langle v\,,\,u\rangle \quad \textit{para} \quad v \in \mathcal{V}\,.$$

Demostración. Sea $B = \{v_1, \ldots, v_n\}$ una base ortonormal de $\mathcal{V}$. Dado $f \in \mathcal{V}^*$ definimos

$$u = \sum_{j=1}^n \overline{f(v_j)}\; v_j \in \mathcal{V}\,.$$

Dado $v \in \mathcal{V}$ se verifica: $v = \sum_{j=1}^n \langle v\,,\,v_j\rangle\, v_j$. Entonces, usando esta representación de $v \in \mathcal{V}$ y la linealidad de $f \in \mathcal{V}^*$, se tiene que

$$f(v) = \sum_{j=1}^n \langle v\,,\,v_j\rangle\, f(v_j) = \sum_{j=1}^n \langle v\,,\,\overline{f(v_j)}\, v_j\rangle = \langle v\,,\,\sum_{j=1}^n \overline{f(v_j)}\, v_j\rangle = \langle v\,,\,u\rangle\,,$$

donde hemos usado las propiedades del producto interno. Lo anterior muestra que $f(v) = \langle v\,,\,u\rangle$, para $v \in \mathcal{V}$.

Con respecto a la unicidad: si suponemos que $w \in \mathcal{V}$ también verifica que $f(v) = \langle v\,,\,w\rangle$ para $v \in \mathcal{V}$ entonces tenemos que

$$\langle v\,,\,w\rangle = f(v) = \langle v\,,\,u\rangle \quad \text{para} \quad v \in \mathcal{V}$$

En particular, $\langle v\,,\,w-u\rangle = \langle v\,,\,w\rangle - \langle v\,,\,u\rangle = 0$, para $v \in \mathcal{V}$: finalmente, eligiendo $v = w-u$ vemos que $\langle w-u\,,\,w-u\rangle = 0$ lo que implica que $w-u = \vec{0}$, por propiedad del producto interno. $\square$

Teorema 11.38. *Sea $(\mathcal{V}, \langle\, , \,\rangle)$ un EPI, $\dim_K \mathcal{V} = n$. Sea $F : \mathcal{V} \to \mathcal{V}^*$ dada por $F(u) = f_u \in \mathcal{V}^*$ (definido como en la Observación 11.36). Entonces F es una función biyectiva que satisface:*

1. *$F(u + v) = F(u) + F(v)$, para $u, v \in \mathcal{V}$.*
2. *$F(\alpha\, u) = \overline{\alpha}\, F(u)$, para $u \in \mathcal{V}$.*

En este caso decimos que F es un isomorfismo lineal conjugado.

Demostración. Por la Observación 11.36 la función F está bien definida; además, el Teorema 11.37 garantiza que F es suryectiva. Más aún, la unicidad que aparece en ese mismo resultado muestra que F es inyectiva: si $F(u) = F(v)$ entonces $f_u = f_v$ lo que implica que $u = v$. Así, F es una biyección.

Las identidades de los ítems 1 y 2 son consecuencias de las propiedades del producto interno: si $w \in \mathcal{V}$ entonces

$$f_{u+v}(w) = \langle w\, ,\, u + v\rangle = \langle w\, ,\, u\rangle + \langle w\, ,\, v\rangle = f_u(w) + f_v(w)$$

y

$$f_{\alpha\, u}(w) = \langle w\, ,\, \alpha\, u\rangle = \overline{\alpha}\, \langle w\, ,\, u\rangle = \overline{\alpha}\ f_u(w)$$

lo que muestra que

$$F(u + v) = f_{u+v} = f_u + f_v = F(u) + F(v)$$

y $F(\alpha\, u) = f_{\alpha\, u} = \overline{\alpha}\, f_u = \overline{\alpha}\, F(u)$. □

Obs 11.39. Con la notación del Teorema 11.38, podemos identificar (de forma *lineal conjugada*) a $\mathcal{V}$ y $\mathcal{V}^*$ en el caso de espacios con producto interno. Utilizando esta identificación se tiene que si $S \subset \mathcal{V}$ y $S^o \subseteq \mathcal{V}^*$ denota el anulador de S (en el sentido definido en notas previas) entonces

$$F(S^{\perp}) = S^o$$

es decir, F identifica el anulador S^o con el complemento ortogonal $S^{\perp}$. En efecto

$$F(u) = f_u \in S^o \;\Leftrightarrow\; \langle v\, ,\, u\rangle = f_u(v) = 0 \quad \text{para} \quad v \in S \;\Leftrightarrow u \in S^{\perp}\,.$$

El isomorfismo F ha sido definido utilizando la naturaleza de los objetos disponibles en el espacio producto interno $(\mathcal{V}\, , \langle\, , \,\rangle)$ (sin usar bases ni coordenadas) y en este sentido también es un isomorfimos (lineal conjugado) *canónico.*

Este isomorfismo establece una relación fundamental entre la noción de algebraica de anulador (disponible en todo espacio vectorial) con la noción geométrica de complemento ortogonal (disponible solo en EPI's). Finalmente, notemos que en el caso real $K = \mathbb{R}$, F es un isomorfismo lineal, en el sentido usual. △

12 Adjuntos de operadores lineales

Comenzamos con el siguiente resultado, que establece la existencia del operador adjunto, que notaremos T^*, de un operador que T que actúa en un EPI de dimensión finita.

Teorema 12.1. *Sea $(\mathcal{V}, \langle\, , \,\rangle)$ un EPI, $\dim_K \mathcal{V} = n$, y sea $T \in L(\mathcal{V})$. Entonces existe un único $T^* \in L(\mathcal{V})$ que verifica:*

$$\langle T\, v\, ,\, w\rangle = \langle v\, ,\, T^*\, w\rangle \quad \textit{para todos} \quad v, w \in \mathcal{V} \tag{96}$$

T^* es llamado el (operador) adjunto de T con respecto al producto interno $\langle\, , \,\rangle$. En este caso, también se verifica $\langle T^*\, v\, ,\, w\rangle = \langle v\, ,\, T\, w\rangle$ para todos $v, w \in \mathcal{V}$.

Demostración. Fijemos $w \in \mathcal{V}$ arbitrario; en este caso definimos $g_w \in \mathcal{V}^*$ dado por

$$g_w(v) = \langle T\,v\,,\,w\rangle \quad \text{para} \quad v \in \mathcal{V}\,.$$

El hecho de que g_w es un funcional lineal es consecuencia de la linealidad de T y de las propiedades de linealidad del producto escalar en su primer coordenada:

$$g_w(\alpha\,u+v) = \langle T\,(\alpha\,u+v)\,,\,w\rangle = \langle \alpha\,T\,u + T\,v\,,\,w\rangle = \alpha\,\langle T\,u\,,\,w\rangle + \langle T\,v\,,\,w\rangle = \alpha\,g_w(u) + g_w(v)\,.$$

Por el Teorema 11.37 (de representación de Riesz) existe un único vector $w^* \in \mathcal{V}$ que representa al funcional $g_w \in \mathcal{V}^*$ es decir, tal que

$$\langle T\,v\,,\,w\rangle = g_w(v) = \langle v\,,\,w^*\rangle \quad \text{para} \quad v \in \mathcal{V}\,.$$

Definimos el valor $T^*w = w^*$; de esta forma (como consecuencia de la unicidad del vector w^* afirmada en el Teorema 11.37) queda bien definida la función $T^* : \mathcal{V} \to \mathcal{V}$ tal que, con las notaciones previas, $T^*w = w^*$, para $w \in \mathcal{V}$ arbitrario. Así, por construcción se verifica que

$$\langle T\,v\,,\,w\rangle = \langle v\,,\,w^*\rangle = \langle v\,,\,T^*\,w\rangle \quad \text{para todos} \quad v,\,w \in \mathcal{V}\,.$$

Además, $T^* : \mathcal{V} \to \mathcal{V}$ es lineal: en efecto, si $\alpha \in K$ y $w_1,\,w_2 \in \mathcal{V}$ entonces, por definicion de T^*, se verifica

$$\langle T\,v\,,\,w_1\rangle = \langle v\,,\,T^*\,w_1\rangle \quad \text{y} \quad \langle T\,v\,,\,w_2\rangle = \langle v\,,\,T^*\,w_2\rangle \quad \text{para} \quad v \in \mathcal{V}\,.$$

Además, si consideramos

$$\begin{aligned} g_{\alpha\,w_1+w_2}(v) &= \langle T\,v\,,\,\alpha\,w_1+w_2\rangle = \overline{\alpha}\,\langle T\,v\,,\,w_1\rangle + \langle T\,v\,,\,w_2\rangle = \overline{\alpha}\,\langle v\,,\,T^*\,w_1\rangle + \langle v\,,\,T^*\,w_2\rangle \\ &= \langle v\,,\,\alpha\,T^*\,w_1\rangle + \langle v\,,\,T^*\,w_2\rangle = \langle v\,,\,\alpha\,T^*\,w_1 + T^*\,w_2\rangle\,. \end{aligned}$$

Por la unicidad del vector que representa el funcional $g_{\alpha\,w_1+w_2}$ vemos que (usando la notación previa) $(\alpha\,w_1+w_2)^* = \alpha\;T^*\,w_1 + T^*\,w_2$; y por la definición de T^*, concluimos que

$$T^*(\alpha\,w_1+w_2) = (\alpha\,w_1+w_2)^* = \alpha\;T^*\,w_1 + T^*\,w_2$$

lo que prueba la linealidad de T^*.

Verificamos la unicidad de T^*; si $S \in L(\mathcal{V})$ es tal que

$$\langle T\,v\,,\,w\rangle = \langle v\,,\,Sw\rangle \quad \text{para todos} \quad v,\,w \in \mathcal{V}$$

entonces, dado $w \in \mathcal{V}$,

$$\langle v\,,\,T^*\,w\rangle = \langle T\,v\,,\,w\rangle = \langle v\,,\,S\,w\rangle \quad \text{para todos} \quad v \in \mathcal{V}$$

Así,

$$\langle v\,,\,T^*\,w - S\,w\rangle = 0 \quad \text{para todos} \quad v \in \mathcal{V} \implies T^*\,w - S\,w = \vec{0}$$

y luego $T^*\,w = S\,w$, para todo $w \in \mathcal{V}$; así, $T^* = S$.

La afirmación: $\langle T^*\,v\,,\,w\rangle = \langle v\,,\,T\,w\rangle$ para todos $v,\,w \in \mathcal{V}$, del enunciado es una consecuencia de la identidad en la Eq. (96) y de las propiedades del producto interno, y se deja como ejercicio. □

El próximo resultado muestra que calcular matrices de operadores con respecto a bases ortonormales es sencillo.

Proposición 12.2. *Sea $(\mathcal{V}, \langle \, , \, \rangle)$ un EPI, con base ortonormal $B = \{v_1, \dots, v_n\}$. Si $T \in L(\mathcal{V})$ y $A = [T]_B \in K^{n\times n}$ entonces*

$$A_{jk} = \langle T\, v_k \, , \, v_j \rangle \quad \textit{para} \quad 1 \le j,\, k \le n\,.$$

Demostración. Recordemos que por construcción, la k-ésima columna de la matriz $A = [T]_B$ coincide con el vector de coordenadas de $[T\, v_k]_B$, es decir,

$$T\, v_k = \sum_{j=1}^{n} A_{jk}\; v_j \quad \text{para} \quad 1 \le k \le n\,.$$

Por otro lado, el Corolario 11.25 indica que si $1 \le k \le n$,

$$T\, v_k = \sum_{j=1}^{n} \langle T\, v_k \, , \, v_j \rangle \; v_j\,.$$

Así, por unicidad de las coordenadas con respecto a la base B, concluimos la identidad $A_{jk} = \langle T\, v_k \, , \, v_j \rangle$, para $1 \le j \le n$, y para $1 \le k \le n$. □

Definición 12.3. *Sea $K = \mathbb{R}$ ó $K = \mathbb{C}$ y sea $A \in K^{n\times n}$. Definimos la matriz adjunta de A, notada $A^* \in K^{n\times n}$, dada por*

$$(A^*)_{jk} = \overline{A_{kj}} \quad \textit{para} \quad 1 \le j,\, k \le n\,,$$

donde $\overline{A_{kj}} \in K$ denota el complejo conjugado de $A_{kj} \in K$. □

Ejemplo 12.4. Sea $A \in \mathbb{C}^{2\times 2}$ dada por

$$A = \begin{pmatrix} 1 & i \\ \sqrt{3} & 2+i \end{pmatrix} \implies A^* = \begin{pmatrix} 1 & \sqrt{3} \\ -i & 2-i \end{pmatrix}.$$

△

Obs 12.5. Sea $K = \mathbb{R}$ ó $K = \mathbb{C}$ y sea $A \in K^{n\times n}$. En álgebra I hemos definido una matriz, llamada la matriz adjunta de A y notada $\operatorname{adj}(A)$ de forma que

$$[\operatorname{adj}(A)]_{jk} = (-1)^{j+k}\; \det(A(k|j)) \quad \text{para} \quad 1 \le j,\, k \le n\,,$$

donde $A(k|j) \in K^{(n-1)\times(n-1)}$ es la matriz que se obtiene tachando la k-ésima fila y la j-ésima columna de A. Esta matriz será llamada la *matriz adjunta de A en el sentido de álgebra 1* (y en realidad no va a figurar en estas notas). Por otro lado, la matriz adjunta en el sentido de la Definición 12.3 será llamada simplemente *matriz adjunta* de A. △

Obs 12.6. Sea $K = \mathbb{R}$ ó $K = \mathbb{C}$ y sea $A \in K^{n\times n}$. Consideremos K^n dotado del producto interno usual, y sea $A^* \in K^{n\times n}$ la matriz adjunta de A. Entonces, se verifica la identidad (ejercicio)

$$\langle A\vec{x} \, , \, \vec{y} \rangle = \langle \vec{x} \, , \, A^*\, \vec{y} \rangle \quad \text{para} \quad \vec{x},\, \vec{y} \in K^n\,.$$

△

Proposición 12.7. *Sea $(\mathcal{V}, \langle \, , \, \rangle)$ un EPI, con* **base ortonormal** *$B = \{v_1, \dots, v_n\}$. Dado $T \in L(\mathcal{V})$, si $T^* \in L(\mathcal{V})$ denota el operador adjunto (en el sentido definido en el Teorema 12.1) entonces*

$$[T^*]_B = [T]_B^* \in K^{n\times n}\,,$$

donde $[T]_B^*$ denota la matriz adjunta de $[T]_B$, en el sentido de la Definición 12.3.

Demostración. Por la Proposición 12.2, dados $1 \leq j,\, k \leq n$ se tiene que

$$([T]_B)_{jk} = \langle T\, v_k\,,\, v_j\rangle \quad \text{y} \quad ([T^*]_B)_{jk} = \langle T^*\, v_k\,,\, v_j\rangle = \langle v_k\,,\, T\, v_j\rangle = \overline{\langle T\, v_j\,,\, v_k\rangle}$$

donde hemos usado las propiedades del operador adjunto (del Teorema 12.1) y del producto interno. La identidad anterior muestra que

$$([T^*]_B)_{jk} = \overline{([T]_B)_{kj}} \quad \text{para} \quad 1 \leq j,\, k \leq n\,.$$

Así, se verifica $[T^*]_B = [T]_B^*$, por la Definición 12.3. □

Obs 12.8. Sea $(\mathcal{V}, \langle\, ,\, \rangle)$ un EPI, sea $T \in L(\mathcal{V})$ y sea $T^* \in L(\mathcal{V})$ el operador adjunto de T. Consideremos el isomorfismo lineal conjugado $F : \mathcal{V} \to \mathcal{V}^*$ del Teorema 11.38. Si $T^t \in L(\mathcal{V}^*)$ denota el operador transpuesto de T (como ha sido definido en notas anteriores) entonces se verifica que

$$F^{-1} \circ T^t \circ F = T^*\,.$$

La identidad anterior, cuya prueba es un ejercicio, establece una relación fundamental entre la noción de transpuesta y de adjunto de un operador que actúa en un epi. △

Ejemplos 12.9. Consideremos $K = \mathbb{R}$ ó $K = \mathbb{C}$. Entonces

1. Sea $(\mathcal{V}, \langle\, ,\, \rangle)$ un EPI de dimensión finita, $\mathcal{W} \subset \mathcal{V}$ un subespacio y $P_{\mathcal{W}}$ el proyector ortogonal sobre $\mathcal{W}$ ($P_{\mathcal{W}} = P_{\mathcal{W}//\mathcal{W}^\perp}$ es el proyector cuyo rango es $\mathcal{W}$ y su núcleo es $\mathcal{W}^\perp$). Entonces $P_{\mathcal{W}}^* = P_{\mathcal{W}}$ (es decir, el operador adjunto de $P_{\mathcal{W}}$ es el mismo operador $P_{\mathcal{W}}$). En efecto, recordemos que si $P_{\mathcal{W}^\perp}$ denota el proyector ortogonal sobre $\mathcal{W}^\perp$ entonces $\{P_{\mathcal{W}}\,,\, P_{\mathcal{W}^\perp}\}$ es el sistema de proyecciones asociado a la descomposición en suma directa (ortogonal) $\mathcal{V} = \mathcal{W} \dot{+} \mathcal{W}^\perp$ (cosa que ya hemos probado). En particular, $I = P_{\mathcal{W}} + P_{\mathcal{W}^\perp}$; dado $v \in \mathcal{V}$ entonces $v = P_{\mathcal{W}}\, v + P_{\mathcal{W}^\perp}\, v$, donde $P_{\mathcal{W}}\, v \in \mathcal{W}$ y $P_{\mathcal{W}\perp}\, v \in \mathcal{W}^\perp$; Así, si $v,\, w \in \mathcal{V}$ entonces

   $$\langle P_{\mathcal{W}}\, v\,,\, w\rangle = \langle P_{\mathcal{W}}\, v\,,\, P_{\mathcal{W}}\, w + P_{\mathcal{W}^\perp}\, w\rangle = \langle P_{\mathcal{W}}\, v\,,\, P_{\mathcal{W}}\, w\rangle + \langle P_{\mathcal{W}}\, v\,,\, P_{\mathcal{W}^\perp}\, w\rangle = \langle P_{\mathcal{W}}\, v\,,\, P_{\mathcal{W}}\, w\rangle$$

   donde usamos que $\langle P_{\mathcal{W}}\, v\,,\, P_{\mathcal{W}^\perp}\, w\rangle = 0$ pues $P_{\mathcal{W}}\, v \in \mathcal{W}$ y $P_{\mathcal{W}^\perp}\, w \in \mathcal{W}^\perp$. De forma similar, se verifica (ejercicio)

   $$\langle v\,,\, P_{\mathcal{W}}\, w\rangle = \langle P_{\mathcal{W}}\, v\,,\, P_{\mathcal{W}}\, w\rangle \implies \langle P_{\mathcal{W}}\, v\,,\, w\rangle = \langle v\,,\, P_{\mathcal{W}}\, w\rangle \quad , \quad v,\, w \in \mathcal{V}\,.$$

   Así, por la unicidad del operador adjunto del Teorema 12.1, vemos que $P_{\mathcal{W}}^* = P_{\mathcal{W}}$.

2. Consideramos $K^{m\times n}$ con el producto interno dado por

   $$\langle A\,,\, B\rangle = \sum_{j=1}^{m} \sum_{k=1}^{n} A_{jk}\, \overline{B_{jk}}\,.$$

   Dada una matriz $M \in K^{m\times m}$ consideramos el operador $L_M \in L(K^{m\times n})$ dado por $L_M(A) = M \cdot A \in K^{m\times n}$, para $A \in K^{m\times n}$. Entonces se verifica que $L_M^* = L_{M^*}$ es decir, el adjunto del operador L_M es el operador L_{M^*} que está dado por $L_{M^*}(A) = M^* \cdot A$, $A \in K^{m\times n}$, donde $M^* \in K^{m\times m}$ denota la matriz adjunta de la matriz M (ejercicio). △

En el siguiente resultado veremos algunas propiedades básicas pero importantes de la adjunción de operadores.

Proposición 12.10. *Sea $(\mathcal{V}, \langle\, ,\, \rangle)$ un EPI de dimensión finita, sean $T, U \in L(\mathcal{V})$ y $\alpha \in K$. Entonces:*

1. $0^* = 0$ *y* $I^* = I$;

2. $(\alpha T + U)^* = \overline{\alpha} T^* + U^*$*;*

3. $(T U)^* = U^* T^*$*;*

4. $(T^*)^* = T$*.*

Demostración. Verificamos el ítem 2: para eso, sean v, $w \in \mathcal{V}$, y consideremos

$$\begin{aligned} \langle (\alpha T + U) v , w \rangle &= \alpha \langle T v , w \rangle + \langle U v , w \rangle = \alpha \langle v , T^* w \rangle + \langle v , U^* w \rangle = \langle v , \overline{\alpha} T^* w + U^* w \rangle \\ &= \langle v , (\overline{\alpha} T^* + U^*) w \rangle . \end{aligned}$$

Así, el operador $\overline{\alpha} T^* + U^*$ cumple el rol de operador $(\alpha T + U)^*$, en el sentido que vale la identidad entre el primer y último miembro de la cadena de igualdades anterior (ver Teorema 12.1). Luego, por unicidad del adjunto, vemos que $(\alpha T + U)^* = \overline{\alpha} T^* + U^*$ y vale el ítem 2.

Verificamos el ítem 3. Sean v, $w \in \mathcal{V}$, entonces

$$\langle (TU) v , w \rangle = \langle T(U v) , w \rangle = \langle U v , T^* w \rangle = \langle v , U^*(T^* w) \rangle = \langle v , (U^* T^*) w \rangle$$

donde hemos usado que el producto $T U$ corresponde a la composición de los operadores, la propiedad de T^* (aplicada al par de vectores Uv y w), la propiedad de U^* (aplicada al par de vectores v y $T^* w$) y finalmente, la definición de $U^* T^*$ como la composición de estos operadores.

Los ítems 1 y 4 quedan como ejercicio para el lector. □

Obs 12.11. Sea $K = \mathbb{R}$ ó $K = \mathbb{C}$. Las propiedades que tiene la función $L(\mathcal{V}) \ni T \mapsto T^* \in L(\mathcal{V})$ tienen análogos en el caso de la función $K^{n\times n} \ni A \mapsto A^* \in K^{n\times n}$. Explícitamente, se verifica: si $\alpha \in K$ y A, $B \in K^{n\times n}$ entonces

1. $I^* = I$, $0^* = 0$;
2. $(\alpha A + B)^* = \overline{\alpha}\ A^* + B^*$;
3. $(A B)^* = B^* A^*$;
4. $(A^*)^* = A$.

La verificación de estas propiedades es un ejercicio (sugerencia: aplicar la Definición 12.3).

△

12.1 Isomorfismos entre espacios con producto interno

En lo que sigue vamos a considerar una noción de comparación entre distintos EPI's.

Definición 12.12. *Sean* $(\mathcal{V}, \langle \, , \, \rangle_{\mathcal{V}})$ *y* $(\mathcal{W}, \langle \, , \, \rangle_{\mathcal{W}})$ *dos EPI's sobre el cuerpo* K*. Dado* $T \in L(\mathcal{V}, \mathcal{W})$ *decimos que*

1. T *preserva el producto interno (preserva el p.i.) si*

$$\langle T v , T w \rangle_{\mathcal{W}} = \langle v , w \rangle_{\mathcal{V}} \quad \textit{para} \quad v, w \in \mathcal{V} .$$

2. T *es un isomorfismo de EPI's si* T *es un isomorfismo entre* K*-ev's que preserva el producto interno.*

□

Proposición 12.13. *Sean* $(\mathcal{V}, \langle\, , \,\rangle_{\mathcal{V}})$ *y* $(\mathcal{W}, \langle\, , \,\rangle_{\mathcal{W}})$ *dos EPI's sobre el cuerpo* K*. Dado* $T \in L(\mathcal{V}, \mathcal{W})$ *son equivalentes:*

1. *T preserva el p.i.;*

2. *T preserva la norma, es decir:* $\|Tv\|_{\mathcal{W}} = \|v\|_{\mathcal{V}}$ *para* $v \in \mathcal{V}$*.*

Más aún, en caso de que T preserve el p.i. entonces T es monomorfismo.

Demostración. 1. $\implies$ 2. Si T preserva el p.i. entonces

$$\|T\,v\|_{\mathcal{W}}^2 = \langle T\,v\,,\,T\,v\rangle_{\mathcal{W}} = \langle v\,,\,v\rangle_{\mathcal{V}} = \|v\|_{\mathcal{V}}^2 \implies \|T\,v\|_{\mathcal{W}} = \|v\|_{\mathcal{V}}\,.$$

2. $\implies$ 1. Aplicamos la identidad de polarización (ver Eq. (88)) en los EPI's; si $K = \mathbb{C}$ entonces obtenemos:

$$\langle T\,v\,,\,T\,w\rangle = \frac{1}{4}\sum_{k=0}^{3} i^k\,\|T\,v + i^k\,T\,w\|_{\mathcal{W}}^2 = \frac{1}{4}\sum_{k=0}^{3} i^k\,\|T\,(v + i^k\,w)\|_{\mathcal{W}}^2 = \frac{1}{4}\sum_{k=0}^{3} i^k\,\|v + i^k\,w\|_{\mathcal{V}}^2 = \langle v\,,\,w\rangle\,,$$

donde hemos usado la linealidad de T y la hipótesis (que T preserva la norma). Si $K = \mathbb{R}$ argumentamos de forma similar (ejercicio. Sugerencia: usar las Ecs. (86) y (87)).

Finalmente, si T preserva el p.i. entonces T preserva la norma y en este caso $\mathrm{N}(T) = \{0\}$: en efecto, si $v \in \mathrm{N}(T)$ entonces $Tv = \vec{0}$ de forma que $0 = \|T\,v\|_{\mathcal{W}} = \|v\|_{\mathcal{V}}$ que muestra que $v = \vec{0}$. $\square$

Teorema 12.14. *Sean* $(\mathcal{V}, \langle\, , \,\rangle_{\mathcal{V}})$ *y* $(\mathcal{W}, \langle\, , \,\rangle_{\mathcal{W}})$ *EPI's sobre el cuerpo K tales que* $\dim_K \mathcal{V} = \dim_K \mathcal{W}$*. Dado* $T \in L(\mathcal{V}, \mathcal{W})$ *son equivalentes:*

1. *T preserva el p.i.;*

2. *T es un isomorfismo de EPI's;*

3. *T transforma las bases ortonormales de $\mathcal{V}$ en bases ortonormales de $\mathcal{W}$;*

4. *T transforma una base ortonormal de $\mathcal{V}$ en una base ortonormal de $\mathcal{W}$.*

Demostración. 1. $\implies$ 2. Si T preserva el p.i. entonces T es monomorfismo, por la Proposición 12.13. Como $\dim_K \mathcal{V} = \dim_K \mathcal{W}$ deducimos que T es un isomorfismo de K-ev's, que preserva el p.i. Así, por la Definición 12.12 vemos que T es un isomorfismo de EPI's.

2. $\implies$ 3. Sea $B = \{v_1, \ldots, v_n\}$ base ortonormal de $\mathcal{V}$ arbitraria; veamos que $B' = \{T\,v_1, \ldots, T\,v_n\}$ es base ortonormal de $\mathcal{W}$. En efecto, como T es isomorfismo de K-ev's entonces B' es base de $\mathcal{W}$; además, como T preserva el p.i. entonces, si $1 \le j,\, k \le n$

$$\langle T\,v_j\,,\,T\,v_k\rangle_{\mathcal{W}} = \langle v_j\,,\,v_k\rangle_{\mathcal{V}} = \delta_{jk}\,.$$

Las identidades anteriores muestran que B' es sistema ortonormal; entonces B' es base ortonormal de $\mathcal{W}$.

3. $\implies$ 4. Es inmediato!

4. $\implies$ 1. Supongamos que existe una base ortonormal de $\mathcal{V}$, notada $B = \{v_1, \ldots, v_n\}$, tal que $B' = \{T\,v_1, \ldots, T\,v_n\}$ es base ortonormal de $\mathcal{W}$. Si $v \in \mathcal{V}$ entonces, por el Corolario 11.25

$$v = \sum_{j=1}^{n} \langle v\,,\,v_j\rangle_{\mathcal{V}}\;v_j \quad \text{ y } \quad \|v\|_{\mathcal{V}}^2 = \sum_{j=1}^{n} |\langle v\,,\,v_j\rangle_{\mathcal{V}}|^2\,.$$

Usando la representación de v anterior, la linealidad de T y el hecho de que B' también es base ortonormal de $\mathcal{W}$

$$T\,v = \sum_{j=1}^{n} \langle v\,,\, v_j\rangle_{\mathcal{V}}\; T\,v_j \implies \langle T\,v\,,\, T\,v_j\rangle_{\mathcal{V}} = \langle v\,,\, v_j\rangle_{\mathcal{V}} \quad \text{para} \quad 1 \le j \le n$$

donde la identidad anterior es una consecuencia de la unicidad de coordenadas con respecto a B' y Corolario 11.25 (pero en $\mathcal{W}$). Finalmente, otra vez usando el Corolario 11.25,

$$\|T\,v\|^2_{\mathcal{W}} = \sum_{j=1}^{n} |\langle T\,v\,,\, T\,v_j\rangle_{\mathcal{W}}|^2 = \sum_{j=1}^{n} |\langle v\,,\, v_j\rangle_{\mathcal{V}}|^2 = \|v\|^2_{\mathcal{V}}\,.$$

Así, T preserva la norma y luego, por la Proposición 12.13, preserva el p.i. □

Corolario 12.15. *Sean $(\mathcal{V}, \langle\,,\,\rangle_{\mathcal{V}})$ y $(\mathcal{W}, \langle\,,\,\rangle_{\mathcal{W}})$ EPI's sobre el cuerpo K de dimensión finita. Entonces existe $T \in L(\mathcal{V}, \mathcal{W})$ isomorfismo de EPI's si y solo si* $\dim_K \mathcal{V} = \dim_K \mathcal{W}$.

Demostración. Supongamos que existe $T \in L(\mathcal{V}, \mathcal{W})$ isomorfismo de EPI's entonces; como T es un isomorfismo de K-ev's, entonces $\dim_K \mathcal{V} = \dim_K \mathcal{W}$.

Recíprocamente, si $\dim_K \mathcal{V} = \dim_K \mathcal{W} = n$ sean $B = \{v_1, \ldots, v_n\}$ y $B' = \{w_1, \ldots, w_n\}$ bases ortonormales de $\mathcal{V}$ y $\mathcal{W}$ respectivamente. Entonces, como B es en particular una base de $\mathcal{V}$ entonces existe una única $T \in L(\mathcal{V}, \mathcal{W})$ tal que $T\,v_j = w_j$ para $1 \le j \le n$, por resultados previos (tp3). Notemos que T transforma, por construcción, B en B'. Así, por el Teorema 12.14, T es un isomorfismo de EPI's pues T transforma una base ortonormal de $\mathcal{V}$ (en este caso B) en una base ortonormal de $\mathcal{W}$ (en este caso B'). □

Corolario 12.16. *Sean $(\mathcal{V}, \langle\,,\,\rangle_{\mathcal{V}})$ un EPI sobre el cuerpo K, y sea $B = \{v_1, \ldots, v_n\}$ una base ortonormal de $\mathcal{V}$. Si consideramos a K^n dotado del producto interno usual, entonces $C_B : \mathcal{V} \to K^n$ dada por $C_b(v) = [v]_B = (\langle v\,,\, v_j\rangle)_{j=1}^n \in K^n$ es un isomorfismo de EPI's. En particular, C_B preserva el p.i., es decir:*

$$\langle u\,,\, v\rangle = \langle [u]_B\,,\, [v]_B\rangle_{K^n} = \sum_{j=1}^{n} \langle u\,,\, v_j\rangle\; \overline{\langle v\,,\, v_j\rangle}\,.$$

Demostración. C_B transforma la base ortonormal B en $\{e_1, \ldots, e_n\}$ que es una base ortonormal de K^n con el p.i. usual. Ahora, basta aplicar el Teorema 12.14 (ver también el Corolario 11.25). □

12.2 Operadores unitarios

Definición 12.17. *Sea $(\mathcal{V}, \langle\,,\,\rangle)$ un EPI sobre K, de dimensión finita. Un operador $U \in L(\mathcal{V})$ se llama unitario si U es un isomorfismo de EPI de $\mathcal{V}$ en $\mathcal{V}$.* □

En algunos contextos, los operadores unitarios $U \in L(\mathcal{V})$ que actúan en un EPI real, $(\mathcal{V}, \langle\,,\,\rangle)$ $\mathbb{R}$-EPI, son llamados operadores (ó trasnformaciones) *ortogonales*.

Con las notaciones de la Definición 12.17, $U \in L(\mathcal{V})$ es unitario si y solo si U es un isomorfismo de K-ev's que además preserva el p.i. Como consecuencia de la Proposición 12.13 y el Teorema 12.14 deducimos el siguiente resultado.

Teorema 12.18. *Sea $(\mathcal{V}, \langle\,,\,\rangle)$ un EPI sobre K, de dimensión finita. Dado $U \in L(\mathcal{V})$ son equivalentes:*

1. *U preserva la norma;*

2. U preserva el p.i.;

3. U es unitario;

4. U transforma bases ortonormales en bases ortonormales;

5. U transforma una base ortonormal en base ortonormal.

$\square$

El siguiente resultado caracteriza los operadores unitarios en términos de relaciones entre el operador y su adjunto (definido como en el Teorema 12.1).

Teorema 12.19. *Sea $(\mathcal{V}, \langle\, , \,\rangle)$ un EPI sobre K, de dimensión finita. Dado $U \in L(\mathcal{V})$ entonces U es unitario si y solo si $U^* U = U U^* = I$. Es decir, U es unitario si y solo si $U^* = U^{-1}$.*

Demostración. Supongamos que U es un operador unitario. En particular, U es un isomorfismo y resulta un operador inversible; en este caso, existe $U^{-1} \in L(\mathcal{V})$ tal que $U U^{-1} = U^{-1} U = I$. Veamos que $U^* = U^{-1}$: en efecto, si $u,\, v \in \mathcal{V}$ entonces

$$\langle U\,u\,,\,v\rangle = \langle U\,u\,,\,I\,v\rangle = \langle U\,u\,,\,U\,U^{-1}v\rangle = \langle u\,,\,U^{-1}\,v\rangle\,,$$

donde la última igualdad se verifica porque U preserva el producto interno. La identidad anterior, junto con la unicidad del operador adjunto muestran que $U^* = U^{-1}$. Entonces $U\,U^* = U^*\,U = I$.

Recíprocamente, si $U^*\,U = I$ entonces veamos que U preserva el producto interno: en efecto, si $u,\, v \in \mathcal{V}$ entonces

$$\langle U\,u\,,\,U\,v\rangle = \langle u\,,\,U^*(U\,v)\rangle = \langle u\,,\,(U^*\,U)\,v\rangle = \langle u\,,\,v\rangle$$

donde hemos usado la propiedad del adjunto U^*, la definición del producto en $L(\mathcal{V})$ (como la composición de operadores) y la hipótesis. Por el Teorema 12.18 concluimos que U es unitario. $\square$

Definición 12.20. *Sea $K = \mathbb{R}$ ó $K = \mathbb{C}$. Dada $A \in K^{n\times n}$ decimos que A es matriz unitaria si $A^*\,A = A\,A^* = I$. Es decir, A es unitaria si y solo si $A^* = A^{-1}$.* $\square$

Obs 12.21. Sea $K = \mathbb{R}$ ó $K = \mathbb{C}$. Si $A \in K^{n\times n}$ satisface $A^*A = I$ entonces A es unitaria es decir, también verifica la identidad $A\,A^* = I$. En efecto, $A^*A = I$ garantiza que $\mathrm{N}(A) = \{0\}$ (ejercicio). Por resultados de Álgebra I, deducimos que A es matriz inversible; en particular, si multiplicamos a izquierda por A^{-1} en ambos lados de la identidad $A^*\,A = I$ tenemos que $A^* = A^{-1}$, lo que garantiza que $A\,A^* = A\,A^{-1} = I$. $\triangle$

El interés en considerar la clase de matrices unitarias está en el hecho de que describen las matrices de operadores unitarios con respecto a *bases ortonormales* del espacio.

Corolario 12.22. *Sea $(\mathcal{V}, \langle\, , \,\rangle)$ un EPI sobre K y sea $B = \{v_1, \ldots, v_n\}$ una* **base ortonormal** *de $\mathcal{V}$. Dado $U \in L(\mathcal{V})$ entonces U es unitario si y solo si $[U]_B \in K^{n\times n}$ es matriz unitaria.*

Demostración. Sea $B = \{v_1, \ldots, v_n\}$ una base ortonormal de $\mathcal{V}$ y sea $U \in L(\mathcal{V})$. Por la Proposición 12.7, tenemos que $[U^*]_B = [U]_B^* \in K^{n\times n}$. En este caso, como tomar matriz con respecto a una base es isomorfismo de álgebras

$$U^*\,U = I \;\Leftrightarrow\; [U^*\,U]_B = I_n \;\Leftrightarrow\; [U^*]_B\,[U]_B = I_n \;\Leftrightarrow\; [U]_B^*\,[U]_B = I_n\,,$$

donde I_n denota la matriz identidad. De forma similar, $U\,U^* = I$ si y solo si $[U]_B\,[U]_B^* = I_n$. Las dos afirmaciones anteriores muestran que U es unitario si y solo si $[U]_B$ es matriz unitaria. $\square$

Obs 12.23. Sea $(\mathcal{V}, \langle\, , \,\rangle)$ EPI sobre K y sean $B_1 = \{v_1, \ldots, v_n\}$ y $B_2 = \{w_1, \ldots, w_n\}$ dos bases ortonormales de $\mathcal{V}$. Entonces la matriz de cambio de base $M_{B_2, B_1} \in K^{n\times n}$ es una matriz unitaria: en efecto, consideramos $U \in L(\mathcal{V})$ el único operador (transformación lineal de $\mathcal{V}$ en $\mathcal{V}$) que transforma los vectores de la base B_2 en los vectores de la base B_1, es decir $U\, w_j = v_j$ para $1 \leq j \leq n$. Entonces el Teorema 12.18 (ítem 5.) indica que U es operador unitario. Así $[U]_{B_2} \in K^{n\times n}$ es matriz unitaria por el Corolario 12.22 (ojo!: acá también usamos que B_2 es base ortonormal). Finalmente, notemos que si $1 \leq j \leq n$ entonces la j-ésima columna de $[U]_{B_2}$ es $[U\, w_j]_{B_2} = [v_j]_{B_2}$ que también es la j-ésima columna de la matriz M_{B_2, B_1}. Entonces $M_{B_2, B_1} = [U]_{B_2}$ es unitaria. En particular

$$M_{B_1, B_2} = M_{B_2, B_1}^{-1} = M_{B_2, B_1}^{*}$$

$\triangle$

Proposición 12.24. *Sea $(\mathcal{V}, \langle\, , \,\rangle)$ EPI sobre K y sean $B_1 = \{v_1, \ldots, v_n\}$ y $B_2 = \{w_1, \ldots, w_n\}$ dos bases ortonormales de $\mathcal{V}$. Dado $T \in L(\mathcal{V})$ entonces existe una matriz unitaria $A = M_{B_1, B_2} \in K^{n\times n}$ tal que*

$$[T]_{B_2} = A^*\, [T]_{B_1}\, A\,.$$

Demostración. Es una consecuencia inmediata de la fórmula de cambio de base para un operador y la Observación 12.23 (los detalles son ejercicio). $\square$

El resultado anterior sugiere considerar la siguiente nueva noción (que permite describir la relación entre las matrices de un mismo operador con respecto a bases ortonormales).

Definición 12.25. *Sea $K = \mathbb{R}$ ó $K = \mathbb{C}$. Dadas $B, C \in K^{n\times n}$ decimos que B y C son unitariamente equivalentes, notado $B \approx C$, si existe una matriz unitaria $W \in K^{n\times n}$ tal que*

$$B = W^*\, C\, W\,.$$

$\square$

Obs 12.26. Sea $K = \mathbb{R}$ ó $K = \mathbb{C}$. Por un lado, la relación de equivalencia unitaria $\approx$ es una relación de equivalencia en $K^{n\times n}$ (ejercicio). Además, si $B, C \in K^{n\times n}$ son tales que $B \approx C$, entonces existe $W \in K^{n\times n}$ unitaria tal que $B = W^*\, C\, W$; como W es unitaria, entonces es matriz inversible y $W^* = W^{-1}$. Así, $B = W^*\, C\, W = W^{-1}\, C\, W$ y en particular B es semejante a C (cosa que notabamos $B \sim C$).

Sin embargo, si $D, E \in K^{n\times n}$ son tales que $D \sim E$ entonces, en general, no se verifica que $D \approx E$; por ejemplo, las matrices en $\mathbb{R}^{2\times 2}$

$$D = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \quad \text{y} \quad E = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}$$

verifican que $D \sim E$ pero $D \not\approx E$ (es decir, no existe matriz $W \in \mathbb{R}^{2\times 2}$ unitaria tal que $D = W^*\, E\, W$) Ejercicio! (no tan sencillo).

Es por eso que decimos que la equivalencia unitaria es una relación de equivalencia estrictamente más fina que la relación de semejanza. $\triangle$

El siguiente resultado es un criterio útil para caracterizar a las matrices unitarias.

Proposición 12.27. *Sea $K = \mathbb{R}$ ó $K = \mathbb{C}$ y consideremos K^n con el producto interno usual. Sea $A \in K^{n\times n}$ con columnas dadas por $\vec{a}_1, \ldots, \vec{a}_n \in K^n$. Entonces A es matriz unitaria si y solo si $B = \{\vec{a}_1, \ldots, \vec{a}_n\}$ es base ortonormal de K^n.*

Demostración. Sea $A \in K^{n\times n}$ con columnas dadas por $\vec{a}_1, \dots, \vec{a}_n \in K^n$. Si consideramos las entradas de $A = (a_{jk})_{j,k=1}^n$ entonces $\vec{a}_k = (a_{jk})_{j=1}^n$, para $1 \le k \le n$. En este caso $(A^*)_{jk} = \overline{a_{kj}}$ y se tiene que

$$(A^* A)_{h\,k} = \sum_{j=1}^n (A^*)_{hj}\, A_{jk} = \sum_{j=1}^n \overline{a_{jh}}\, a_{jk} = \sum_{j=1}^n a_{jk}\, \overline{a_{jh}} = \langle \vec{a}_k \,,\, \vec{a}_h \rangle \quad \text{para} \quad 1 \le h,\, k \le n\,.$$

Así, $A^*A = I$ si y solo si $(A^*A)_{hk} = \delta_{hk}$ para $1 \le h,\, k \le n$ que equivale a que $\langle \vec{a}_k \,,\, \vec{a}_h \rangle = \delta_{hk}$ para $1 \le h,\, k \le n$.

Notemos que si los vectores $\vec{a}_1, \dots, \vec{a}_n \in K^n$ satisfacen $\langle \vec{a}_k \,,\, \vec{a}_h \rangle = \delta_{hk}$ para $1 \le h,\, k \le n$, entonces $B = \{\vec{a}_1, \dots, \vec{a}_n\}$ es un sistema ortonormal (en particular, un sistema ortogonal de vectores no nulos) y en particular es l.i. Como $\dim_K K^n = n$ entonces B es base ortonormal de K^n. Así, si A es unitaria (es decir, verifica $A^*A = I$) entonces B es base ortonormal.

Recíprocamente, si B es base ortonormal, entonces verifica $\langle \vec{a}_k \,,\, \vec{a}_h \rangle = \delta_{hk}$ para $1 \le h,\, k \le n$; por observaciones hechas más arriba (en esta prueba) concluimos que $A^*A = I$; por la Observación 12.21, ahora podemos decir que A es unitaria. $\square$

12.3 Diagonalización en bases ortonormales

En lo que sigue, dado un K-EPI $(\mathcal{V}, \langle\,,\,\rangle)$ de dimensión finita, y dado un operador $T \in L(\mathcal{V})$, vamos a estar interesados en determinar cuando existe una *base ortonomal* $B = \{v_1, \dots, v_n\}$ de $\mathcal{V}$, tal que $[T]_B \in K^{n\times n}$ sea una matriz diagonal; en caso que exista tal B, decimos que T es *unitariamente diagonalizable.* Vamos a ver que podemos caracterizar a los operadores unitariamente diagonalizables en términos de relaciones entre T y su operador adjunto T^*. Para ello, consideramos el siguiente resultado (que será utilizado para poder hacer reducciones):

Lema 12.28. *Sea $(\mathcal{V}, \langle\,,\,\rangle)$ un K-EPI, $\dim_K \mathcal{V} = n$. Sea $T \in L(\mathcal{V})$ (operador arbitrario) y sea $\mathcal{W} \subseteq \mathcal{V}$ subespacio T-invariante. Entonces $\mathcal{W}^\perp$ es T^*-invariante.*

Demostración. Por hipótesis, si $w \in \mathcal{W}$ entonces $T\,w \in \mathcal{W}$. Sea $u \in \mathcal{W}^\perp$ y verifiquemos que $T^*\,u \in \mathcal{W}^\perp$: para ello, sea $w \in \mathcal{W}$ arbitrario, entonces

$$\langle T^*\,u \,,\, w\rangle = \langle u \,,\, T\,w\rangle = 0\,,$$

donde hemos usado la propiedad de T^* y las hipótesis's ($T\,w \in \mathcal{W}$ y $u \in \mathcal{W}^\perp$). Así, $\langle T^*\,u \,,\, w\rangle = 0$ para $w \in \mathcal{W}$ y luego $T^*\,u \in \mathcal{W}^\perp$. $\square$

En lo que sigue, nuestro análisis se divide en este contexto en el caso $K = \mathbb{R}$ y $K = \mathbb{C}$.

12.3.1 Diagonalización en bases ortonormales: $K = \mathbb{R}$

Comenzamos con la siguiente observación en los espacios con producto interno *reales.*

Obs 12.29. Sea $(\mathcal{V}, \langle\,,\,\rangle)$ un $\mathbb{R}$-EPI, $\dim_{\mathbb{R}} \mathcal{V} = n$. Sea $T \in L(\mathcal{V})$ tal que existe una base ortonormal $B = \{v_1, \dots, v_n\}$ de $\mathcal{V}$ de forma que $[T]_B = \operatorname{diag}(\lambda_1, \dots, \lambda_n) \in \mathbb{R}^{n\times n}$ es decir

$$([T]_B)_{jk} = \delta_{jk}\,\lambda_j \quad \text{para} \quad 1 \le j,\, k \le n\,.$$

Entonces, como B es base ortonormal podemos aplicar la Proposición 12.7 y concluir que

$$([T^*]_B)_{jk} = \overline{([T]_B)_{kj}} = ([T]_B)_{kj} = \delta_{kj}\;\lambda_j = ([T]_B)_{jk} \quad \text{para} \quad 1 \le j,\, k \le n\,,$$

donde hemos usado que las entradas de $[T]_B$ son números reales (que no cambian por la conjugación) y que $\delta_{jk}\,\lambda_j = \delta_{kj}\,\lambda_k$ (delta de Kronecker). En conclusión $[T^*]_B = [T]_B \in \mathbb{R}^{n\times n}$. Como tomar matriz con respecto a B es isomorfismo, entonces ahora vemos que $T^* = T$; esta es una condición necesaria para que T sea unitariamente diagonalizable! $\triangle$

Definición 12.30. *Sea* $(\mathcal{V}, \langle\, ,\, \rangle)$ *un* $\mathbb{R}$*-EPI,* $\dim_{\mathbb{R}} \mathcal{V} = n$. *Dado* $T \in L(\mathcal{V})$ *decimos que* T *es autoadjunto si* $T^* = T$. □

Con la notación de la Observación 12.29 y la Definición 12.30, vemos que en el caso de operadores que actúan en un EPI *real*, una condición necesaria para que T sea unitariamente diagonalizable es que T sea autoadjunto.

Definición 12.31. *Sea* $A \in \mathbb{C}^{n\times n}$. *Decimos que* A *es matriz autoadjunta si* $A^* = A$. □

Notemos que si $A \in \mathbb{R}^{n\times n}$ entonces $A^* = A^t$, de forma que A es autoadjunta si y solo si A es matriz simétrica.

Corolario 12.32. *Sea* $(\mathcal{V}, \langle\, ,\, \rangle)$ *un* $\mathbb{R}$*-EPI,* $\dim_{\mathbb{R}} \mathcal{V} = n$ *y sea* B *una* **base ortonormal** *de* $\mathcal{V}$. *Entonces* $T \in L(\mathcal{V})$ *operador autoadjunto si y solo si* $[T]_B \in \mathbb{R}^{n\times n}$ *es matriz autoadjunta.*

Demostración. Ejercicio □

En lo que sigue veremos que si un operador T que actúa en un $\mathbb{R}$-EPI de dimensión finita es autoadjunto entonces T es unitariamente diagonalizable. Pero antes, veamos las siguientes propiedades:

Teorema 12.33. *Sea* $(\mathcal{V}, \langle\, ,\, \rangle)$ *un* $\mathbb{R}$*-EPI,* $\dim_{\mathbb{R}} \mathcal{V} = n \geq 1$. *Sea* $T \in L(\mathcal{V})$ *tal que* $T^* = T$. *Entonces*

1. $\emptyset \neq \sigma(T) \subset \mathbb{R}$ *;*
2. *Sean* $\lambda,\, \mu \in \sigma(T)$, $\lambda \neq \mu$*: si* $v \in E_T(\lambda)$ *,* $w \in E_T(\mu)$ *entonces:* $v \perp w$.

Demostración. Supongamos que $n \geq 1$. Sea B una base ortonormal de $\mathcal{V}$ y sea $A = [T]_B \in \mathbb{R}^{n\times n}$. Como $T = T^*$ entonces, por la Proposición 12.7, tenemos que $A = A^*$.

Sea $p_T(x) = p_A(x) = \det(x\, I - A) \in \mathbb{R}[x]$ el polinomio característico de T (de A). En particular, $p_A(x) \in \mathbb{C}[x]$ y admite al menos una raíz *compleja* $\lambda \in \mathbb{C}$, pues $\mathrm{gr}(p_A(x)) = n \geq 1$. Entonces existe $\vec{z} \in \mathbb{C}^n$ tal que $\vec{z} \neq \vec{0}$ y $A\, \vec{z} = \lambda\, \vec{z}$. Veamos que $\lambda \in \mathbb{R}$: en efecto, considerando el producto interno usual en $\mathbb{C}^n$

$$\lambda\, \langle \vec{z}\,,\, \vec{z}\rangle_{\mathbb{C}^n} = \langle \lambda\, \vec{z}\,,\, \vec{z}\rangle_{\mathbb{C}^n} = \langle A\, \vec{z}\,,\, \vec{z}\rangle_{\mathbb{C}^n} = \langle \vec{z}\,,\, A^*\, \vec{z}\rangle_{\mathbb{C}^n} = \langle \vec{z}\,,\, A\, \vec{z}\rangle_{\mathbb{C}^n} = \langle \vec{z}\,,\, \lambda\, \vec{z}\rangle_{\mathbb{C}^n} = \overline{\lambda}\, \langle \vec{z}\,,\, \vec{z}\rangle_{\mathbb{C}^n}$$

donde hemos usado las propiedades del producto interno, la hipótesis sobre A y $\vec{z}$ y la Observación 12.6. Así, $\lambda\, \langle \vec{z}\,,\, \vec{z}\rangle_{\mathbb{C}^n} = \overline{\lambda}\, \langle \vec{z}\,,\, \vec{z}\rangle_{\mathbb{C}^n}$; como $\vec{z} \neq \vec{0}$ entonces $\langle \vec{z}\,,\, \vec{z}\rangle_{\mathbb{C}^n} > 0$ y podemos concluir que $\lambda = \overline{\lambda}$ de forma que $\lambda \in \mathbb{R}$, como queríamos verificar. Así, hemos hallado una raíz real $\lambda \in \mathbb{R}$ del polinomio característico $p_A(x) = p_T(x)$ de forma que $\lambda \in \sigma(T)$, por las propiedades del polinomio característico.

Finalmente, sean $\lambda,\, \mu \in \sigma(T)$, $\lambda \neq \mu$: si $v \in E_T(\lambda)$, $w \in E_T(\mu)$ entonces

$$\lambda\, \langle v\,,\, w\rangle = \langle \lambda\, v\,,\, w\rangle = \langle T\, v\,,\, w\rangle = \langle v\,,\, T^*\, w\rangle = \langle v\,,\, T\, w\rangle = \langle v\,,\, \mu\, w\rangle = \overline{\mu}\, \langle v\,,\, w\rangle = \mu\, \langle v\,,\, w\rangle$$

donde hemos usado las propiedades del producto interno, las hipótesis y finalmente que los autovalores de T son reales (de forma que $\overline{\mu} = \mu$). En resumen, hemos probado que $\lambda\, \langle v\,,\, w\rangle = \mu\langle v\,,\, w\rangle$ que implica que $(\lambda - \mu)\, \langle v\,,\, w\rangle = 0$; como $\lambda - \mu \neq 0$ entonces $\langle v\,,\, w\rangle = 0$ y $v \perp w$. □

Teorema 12.34. *Sea* $(\mathcal{V}, \langle\, ,\, \rangle)$ *un* $\mathbb{R}$*-EPI,* $\dim_{\mathbb{R}} \mathcal{V} = n \geq 1$ *y sea* $T \in L(\mathcal{V})$. *Entonces* T *es unitariamente diagonalizable si y solo si* T *es autoadjunto.*

Demostración. Hemos visto que si un operador T que actúa en un EPI real es unitariamente diagonalizable entonces $T = T^*$ (ver Observación 12.29).

Recíprocamente, supongamos que T es autoadjunto, es decir $T = T^*$. Vamos a probar que existe una base ortonormal $B = \{v_1, \ldots, v_n\}$ de $\mathcal{V}$ tal que $[T]_B \in \mathbb{R}^{n\times n}$ es matriz diagonal, por inducción en $n \geq 1$.

Si $n = 1$ entonces el resultado es trivial (en este caso $[T]_B \in \mathbb{R}$ es matriz diagonal para cualquier base B, y todo $T \in L(\mathcal{V})$ es autoadjunto!)

Supongamos que el resultado vale para todo $\mathbb{R}$-EPI de dimensión menor ó igual que $n - 1$. Sea $(\mathcal{V}, \langle\, , \,\rangle)$ un $\mathbb{R}$-EPI, $\dim_{\mathbb{R}} \mathcal{V} = n \geq 1$ y $T \in L(\mathcal{V})$ autoadjunto. Entonces, por el Teorema 12.33 el operador T tiene al menos un autovalor es decir, existe $\lambda_1 \in \mathbb{R}$ tal que $\lambda_1 \in \sigma(T) \neq \emptyset$.

En particular, el autoespacio $E_T(\lambda_1) = \{v \in \mathcal{V} \ : \ T\,v = \lambda_1\, v\} \subseteq \mathcal{V}$ es no nulo (es decir, $\dim_{\mathbb{R}} E_T(\lambda_1) \geq 1$). Si $E_T(\lambda_1) = \mathcal{V}$ entonces todo $v \in \mathcal{V}$ satisface que $T\,v = \lambda_1\, v$; en este caso, si B es cualquier base ortonormal de $\mathcal{V}$ se verifica que $[T]_B = \lambda_1\, I_n$ y T es unitariamente diagonalizable.

Supongamos entonces que $1 \leq \dim_{\mathbb{R}} E_T(\lambda_1) \leq n - 1$. Notemos que por construcción $E_T(\lambda_1)$ es un subespacio T-invariante; por el Lema 12.28, $\mathcal{W} = E_T(\lambda_1)^\perp$ es invariante para $T^* = T$, pues T es autoadjunto. Es decir, $\mathcal{W}$ es T invariante. Además, como $\mathcal{W}^\perp = (E_T(\lambda_1)^\perp)^\perp = E_T(\lambda_1)$ y $\dim_{\mathbb{R}} \mathcal{W} + \dim_{\mathbb{R}} \mathcal{W}^\perp = n$ (ver los apuntes correspondientes a la primer parte del tp8); entonces $1 \leq k = \dim_{\mathbb{R}} \mathcal{W} = n - \dim E_T(\lambda_1) \leq n - 1$.

Entonces podemos considerar la restricción del producto interno de $\mathcal{V}$ al subespacio $\mathcal{W}$, de forma que $(\mathcal{W}, \langle\, , \,\rangle)$ resulta un $\mathbb{R}$-EPI, $1 \leq \dim_{\mathbb{R}} \mathcal{W} \leq n - 1$. Además, podemos considerar el operador $T_{\mathcal{W}} \in L(\mathcal{W})$, que es la restricción de T al subespacio T-invariante $\mathcal{W}$. Afirmamos que $T|_{\mathcal{W}}$ es autoadjunto en $(\mathcal{W}, \langle\, , \,\rangle)$: en efecto, si $w_1,\, w_2 \in \mathcal{W}$ entonces

$$\langle T|_{\mathcal{W}}\, w_1\, ,\, w_2\rangle = \langle T\, w_1\, ,\, w_2\rangle = \langle\, w_1\, ,\, T\, w_2\rangle = \langle w_1\, ,\, T|_{\mathcal{W}}\, w_2\rangle\, ,$$

donde hemos usado la definición del operador $T|_{\mathcal{W}}$, el hecho de que el producto interno que hemos considerado en $\mathcal{W}$ es en realidad la restricción del producto interno de $\mathcal{V}$, la hipótesis $T = T^*$ y prodiedades del adjunto y finalmente, nuevamente la definición del operador $T|_{\mathcal{W}}$. Por unicidad de $(T|_{\mathcal{W}})^* \in L(\mathcal{W})$ concluimos que $(T|_{\mathcal{W}})^* = T|_{\mathcal{W}} \in L(\mathcal{W})$ (ver el Teorema 12.1 y su prueba). Así, por hipótesis inductiva, existe una base ortonormal $B' = \{w_1, \ldots, w_k\}$ tal que $[T|_{\mathcal{W}}]_{B'} \in \mathbb{R}^{k\times k} = \operatorname{diag}(\mu_1, \ldots, \mu_k)$, para ciertos $\mu_1, \ldots, \mu_k \in \mathbb{R}$ (recordemos que hemos definido $k = \dim_{\mathbb{R}} \mathcal{W}$). En particular, recordemos que bajo estas hipótesis se verifica $T\, w_j = \mu_j\, w_j$ para $1 \leq j \leq k$.

Sea $B'' = \{v_1, \ldots, v_{n-k}\}$ una base ortonormal de $E_T(\lambda_1)$, con respecto al producto interno de $\mathcal{V}$: en particular $T\, v_j = \lambda_1\, v_j$ para $1 \leq j \leq n - k$, pues $v_1, \ldots, v_{n-k} \in E_T(\lambda_1)$ y todo vector de este subespacio $v \in E_T(\lambda_1)$ verifica $T\, v = \lambda_1\, v$.

Finalmente, recordemos que si $B = B'' \cup B' = \{v_1, \ldots, v_{n-k}, w_1, \ldots, w_k\}$ es el pegado de bases ortonormales de $E_T(\lambda_1)$ y $E_T(\lambda_1)^\perp = \mathcal{W}$, entonces B es base ortonormal de $\mathcal{V}$ (ver las notas de la primer parte del tp6, para el caso de descomposiciones en sumas directas $\mathcal{V} = \mathcal{W} \oplus \mathcal{W}^\perp$). Además, cada vector u de B verifica una ecuación de la forma $Tu = \rho\, u$, para cierto $\rho \in \mathbb{R}$ que depende de u (por ejemplo: si $u = v_j$ para $1 \leq j \leq n - k$, entonces $\rho = \lambda_1$; si $u = w_j$ para $1 \leq j \leq k$ entonces $\rho = \mu_j$). Este último hecho muestra que $[T]_B$ es una matriz diagonal! (recordar los primeros teoremas sobre diagonalización de las notas para el tp 5) donde B es base ortonormal de $\mathcal{V}$ y estamos hechos. □

Obs 12.35. Sea $(\mathcal{V}, \langle\, , \,\rangle)$ un $\mathbb{R}$-EPI, $\dim_{\mathbb{R}} \mathcal{V} = n \geq 1$ y sea $T \in L(\mathcal{V})$ operador autoadjunto. El resultado anterior garantiza que existe una base ortonormal B de $\mathcal{V}$ tal que $[T]_B$ es matriz diagonal.

Pero cómo hallar esta base B? Consideremos (todos) los autovalores $\lambda_1, \ldots, \lambda_k \in \mathbb{R}$ distintos de T; entonces se debe verificar que

$$\sum_{j=1}^{k} \dim_{\mathbb{R}} E_T(\lambda_j) = n$$

pues, en particular, T es diagonalizable!! (de forma que, en particular, valen los resultados para operadores diagonalizables).

Pero además, los autoespacios $E_T(\lambda_1), \ldots, E_T(\lambda_k)$ son mutuamente ortogonales: es decir, si $1 \leq j \neq h \leq k$ y $v \in E_T(\lambda_j)$, $w \in E_T(\lambda_h)$ entonces $\langle v\,,\,w\rangle = 0$ (ver el ítem 3 Teorema 12.33).

Así, sea B_j una base (cualquiera) de $E_T(\lambda_j)$, $1 \leq j \leq k$; notar que $\#(B_j) = \dim_{\mathbb{R}} E_T(\lambda_j)$. Si aplicamos el proceso de ortogonalización de Gram-Schmidt al conjunto (l.i.) B_j obtenemos la base ortogonal B'_j de $E_T(\lambda_j)$; entonces normalizamos los vectores de B'_j para obtener finalmente B''_j que es una base ortonormal de $E_T(\lambda_j)$ (puesto que B''_j y B_j generan el mismo subespacio, por las propiedades del proceso de Gram-Schmidt). Los comentarios anteriores ahora muestran que el conjunto

$$B = B''_1 \cup \ldots \cup B''_k$$

es una base ortonormal de $\mathcal{V}$ (pues es el pegoteo de los sistemas ortonormales B''_j, $1 \leq j \leq k$, que además son mutuamente ortogonales entre sí!). Finalmente, $[T]_B$ es una matriz diagonal, pues cada vector de B proviene de algún B''_j, y todos los vectores de $B''_j \subset E_T(\lambda_j)$ son autovectores de T asociados a λ_j (con $1 \leq j \leq k$). △

Corolario 12.36. *Sea $A \in \mathbb{R}^{n\times n}$. Las siguientes afirmaciones son equivalentes:*

1. *A es ortogonalmente equivalente a una matriz diagonal de entradas reales;*
2. *A es matriz autoadjunta;*
3. *$T_A \in L(\mathbb{R}^n)$ es operador autoadjunto ($T_A\,\vec{x} = A\,\vec{x}$, para $\vec{x} \in \mathbb{R}^n$).*

Demostración. $1 \Rightarrow 2$. Supongamos que A es ortogonalmente equivalente a una matriz diagonal; entonces existe una matriz ortogonal $W \in \mathbb{R}^{n\times n}$ (es decir, unitaria y de entradas reales) y una matriz diagonal $B = \mathrm{diag}(b_1, \ldots, b_n) \in \mathbb{R}^{n\times n}$ tales que $A = W^*\,B\,W$. En este caso, por la Observación 12.11

$$A^* = (W^*\,B\,W)^* = W^*\,B^*\,(W^*)^* = W^*\,B\,W = A$$

donde hemos usado que la matriz diagonal $B \in \mathbb{R}^{n\times n}$ satisface $B^* = B$.

$2. \Rightarrow 3.$ Basta notar que la base canónica B_c de $\mathbb{R}^n$ es una base ortonormal para el producto interno usual y que $[T_A]_{B_c} = A$. Entonces podemos aplicar el Corolario 12.32.

$3. \Rightarrow 1.$ Supongamos que $T_A \in L(\mathbb{C}^n)$ es autoadjunto, en donde consideramos el producto interno usual de $\mathbb{R}^n$. El Teorema 12.34 implica que existe una base ortonormal $B = \{v_1, \ldots, v_n\}$ de $\mathbb{R}^n$ tal que $[T_A]_B$ es matriz diagonal. Sea $B_c = \{e_1, \ldots, e_n\}$ la base canónica de $\mathbb{R}^n$. Entonces, sabemos que

$$[T_A]_B = M_{B_c,B}^{-1}\,[T_A]_{B_c}\,M_{B_c,B} = M_{B_c,B}^{-1}\,A\,M_{B_c,B}\,,$$

donde usamos que $[T_A]_{B_c} = A$. Por Observación 12.23 vemos que $M_{B_c,B} \in \mathbb{R}^{n\times n}$ es matriz unitaria, y en particular que $M_{B_c,B}^{-1} = M_{B_c,B}^*$ de forma que la identidad anterior se puede escribir $[T_A]_B = W^*\,A\,W$, para $W = M_{B_c,B}$ matriz ortogonal (unitaria y de entradas reales) y $[T_A]_B$ matriz diagonal. □

Ejemplo 12.37. Consideremos a $\mathbb{R}^3$ con el producto interno usual. Sea $T \in L(\mathbb{R}^3)$ el operador cuya matriz con respecto a la base canónica es

$$A = \begin{pmatrix} 1 & 2 & 2 \\ 2 & 4 & 4 \\ 2 & 4 & 4 \end{pmatrix}.$$

Notemos que en este caso $A^* = A$; de hecho, como A es una matriz real, entonces $A^* = A^t$ (dado que la conjugación de las entradas de A no modifica las entradas). Como la base canónica B_c de $\mathbb{R}^3$ es una base ortonormal con respecto al producto interno usual, concluimos que T es operador autoadjunto. En este caso, sabemos que existe una base ortonormal B de $\mathbb{R}^3$ tal que $[T]_B$ es matriz diagonal. Para hallar una tal base B, primero calculamos los autovalores distintos de T: en este caso, $p_A(x) = x^3 - 9x^2 = x^2(x-9)$; los autovalores de T son $\lambda_1 = 9$ con multiplicidad algebraica 1 y $\lambda_2 = 0$ con multiplicidad algebraica 2.

En el segundo paso, hallamos bases (arbitrarias) de los autoespacios de T: en este caso

$$E_T(\lambda_1) = E_T(9) = N(A - 9\,I) = \overline{\{u_1 = (1,2,2)\}}$$

donde hemos obtenido la base $B_1 = \{u_1\}$ de $E_T(9)$ por reducción por filas de la matriz $A - 9I$; de forma similar, calculamos

$$E_T(\lambda_2) = E_T(0) = N(A - 0\,I) = N(A) = \overline{\{u_2 = (-2,1,0)\,,\ u_3 = (-2,0,1)\}}$$

donde hemos obtenido la base $B_2 = \{u_2, u_3\}$ de $E_T(0)$ por reducción por filas de la matriz A.

Aplicando el proceso de ortogonalización de Gram-Schmidt a B_1 obtenemos $B_1' = \{u_1\}$ (en este caso, el proceso devuelve el único vector de B_1 sin modificación!). Luego, normalizando los vectores de B_1' obtenemos $B_1'' = \{v_1 = \frac{1}{3}\,(1,2,2)\}$.

De forma similar, aplicando el proceso de ortogonalización de Gram-Schmidt a $B_2 = \{u_2, u_3\}$ obtenemos: (a los vectores obtenidos por GS a partir de B_2 los notamos w_2, w_3:)

$$w_2 = u_2 = (-2,1,0) \quad \text{y}$$

$$w_3 = u_3 - \frac{\langle u_3\,,\, w_2\rangle}{\|w_2\|^2} w_2 = (-2,0,1) - \frac{4}{\sqrt{5}}\,(-2,1,0) = (\frac{8}{\sqrt{5}} - 2, \frac{-4}{\sqrt{5}}, 1)$$

Finalmente, normalizando los vectores de $B_2' = \{w_2, w_3\}$ obtenemos

$$v_2 = \frac{1}{\|w_2\|}\, w_2 = \frac{1}{\sqrt{5}}(-2,1,0) \quad \text{y} \quad v_3 = \frac{1}{\|w_3\|}\, w_3 = 0.39\,(\frac{8}{\sqrt{5}} - 2, \frac{-4}{\sqrt{5}}, 1)$$

donde hemos aproximado la norma $\|w_3\|^{-1} \approx 0.39$. Así, $B_2'' = \{v_2, v_3\}$ es una base ortonormal de $E_T(0)$.

Entonces, $B = \{v_1,\, v_2,\, v_3\}$ es una base ortonormal de $\mathbb{R}^3$ tal que $[T]_B = \operatorname{diag}(9,0,0)$ (verificar esta afirmación). $\triangle$

Obs 12.38. Sea $A \in \mathbb{R}^{n\times n}$ matriz autoadjunta. Podemos aplicar las observaciones realizadas en la prueba del Corolario 12.36 junto con la Observación 12.35 para hallar una matriz ortogonal $W \in \mathbb{R}^{n\times n}$ tal que W^*AW sea matriz diagonal. De hecho, como $T_A \in L(\mathbb{R}^n)$ resulta autoadjunto, podemos aplicar el método descripto en la Observación 12.35 para hallar una base ortonormal B tal que $[T_A]_B$ sea matriz diagonal. La prueba del Corolario 12.36 indica que podemos tomar $W = M_{B_c,B}$, donde B_c denota la base canónica de $\mathbb{R}^n$; como observación final, notemos que esta matriz se calcula de forma sencilla (una vez que conocemos B): porqué? $\triangle$

12.3.2 Diagonalización en bases ortonormales: $K = \mathbb{C}$

En esta sección estudiamos el problema de determinar cuando un operador T que actúa en un $\mathbb{C}$-EPI de dimensión finita es unitariamente diagonalizable. Comenzamos con la siguiente observación (comparar con la observación 12.29)

Obs 12.39. Sea $(\mathcal{V}, \langle\,,\,\rangle)$ un $\mathbb{C}$-EPI, $\dim_{\mathbb{C}} \mathcal{V} = n$. Sea $T \in L(\mathcal{V})$ tal que existe una base ortonormal $B = \{v_1, \ldots, v_n\}$ de $\mathcal{V}$ de forma que $[T]_B = \mathrm{diag}(\lambda_1, \ldots, \lambda_n) \in \mathbb{C}^{n\times n}$ es decir

$$([T]_B)_{jk} = \delta_{jk}\,\lambda_j \in \mathbb{C} \quad \text{para} \quad 1 \leq j,\, k \leq n\,.$$

Entonces, como B es base ortonormal podemos aplicar la Proposición 12.7 y concluir que

$$([T^*]_B)_{jk} = \overline{([T]_B)_{kj}} = \overline{\delta_{kj}\ \lambda_j} = \delta_{kj}\ \overline{\lambda_j} \quad \text{para} \quad 1 \leq j,\, k \leq n\,,$$

donde hemos usado que $\delta_{jk}\,\lambda_j = \delta_{kj}\,\lambda_k$ (delta de Kronecker). En conclusión

$$[T^*]_B = \mathrm{diag}(\overline{\lambda_1}, \ldots, \overline{\lambda_n}) \in \mathbb{C}^{n\times n}$$

también resulta una matriz diagonal. En particular, como las matrice diagonales conmutan entre sí,

$$[T]_B\,[T^*]_B = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)\ \mathrm{diag}(\overline{\lambda_1}, \ldots, \overline{\lambda_n}) = \mathrm{diag}(\overline{\lambda_1}, \ldots, \overline{\lambda_n})\ \mathrm{diag}(\lambda_1, \ldots, \lambda_n) = [T^*]_B\,[T]_B\,.$$

Así,

$$[T\,T^*]_B = [T]_B\,[T^*]_B = [T^*]_B\,[T]_B = [T^*\,T]_B\,;$$

como tomar matriz con respecto a B es isomorfismo, entonces ahora vemos que $T\,T^* = T^*\,T$; esta es una condición necesaria para que T sea unitariamente diagonalizable en el caso complejo! $\triangle$

Definición 12.40. *Sea $(\mathcal{V}, \langle\,,\,\rangle)$ un $\mathbb{C}$-EPI,* $\dim_{\mathbb{C}} \mathcal{V} = n$. *Dado $T \in L(\mathcal{V})$ decimos que T es operador normal si T conmuta con su adjunto es decir, si $T\,T^* = T^*\,T$.* $\square$

Con la notación de la Observación 12.39 y la Definición 12.40, vemos que en el caso de operadores que actúan en un EPI *complejo*, una condición necesaria para que T sea unitariamente diagonalizable es que T sea normal.

En lo que sigue veremos que esta condición es también suficiente para que T sea unitariamente diagonalizable. Comenzamos con las siguientes propiedades:

Proposición 12.41. *Sea $(\mathcal{V}, \langle\,,\,\rangle)$ un $\mathbb{C}$-EPI,* $\dim_{\mathbb{C}} \mathcal{V} = n$ *y sea $T \in L(\mathcal{V})$ operador normal. Entonces*

1. *$N(T) = N(T^*)$;*
2. *Si $\lambda \in \mathbb{C}$ entonces $\lambda\, I - T \in L(\mathcal{V})$ es operador normal;*
3. *Si $\lambda \in \mathbb{C}$ y $v \in \mathcal{V}$ son tales que $T\,v = \lambda\,v$ entonces $T^*\,v = \overline{\lambda}\,v$.*

Demostración. Para verificar el ítem 1, sea $v \in \mathcal{V}$; entonces

$$\|T\,v\|^2 = \langle T\,v\,,\,T\,v\rangle = \langle v\,,\,T^*\,T\,v\rangle = \langle v\,,\,T\,T^*\,v\rangle = \langle T^*\,v\,,\,T^*\,v\rangle = \|T^*\,v\|^2$$

donde hemos usado la propiedades del adjunto y la relación de conmutación $T^*\,T = T\,T^*$. Tomando raíces cuadradas en cada miembro vemos que $\|T\,v\| = \|T^*\,v\|$ para $v \in \mathcal{V}$. En particular,

$$v \in \mathrm{N}(T)\ \Leftrightarrow Tv = 0\ \Leftrightarrow \|Tv\| = 0\ \Leftrightarrow \|T^*v\| = 0\ \Leftrightarrow T^*v = 0\ \Leftrightarrow v \in \mathrm{N}(T^*)\,.$$

Para verificar el ítem 2, notemos que por la Proposición 12.10,

$$(\lambda\, I - T)^* = \overline{\lambda}\, I - T^* \in L(\mathcal{V})\,.$$

Así, para verificar que $\lambda\, I - T$ es normal, verificamos la conmutación con su adjunto (que describimos en la ecuación anterior): aplicando la propiedad distribuiva,

$$(\lambda\, I - T)\ (\overline{\lambda}\, I - T^*) = |\lambda|^2\, I - \lambda\, T^* - \overline{\lambda}\, T + T\, T^*$$

$$(\overline{\lambda}\, I - T^*)\ (\lambda\, I - T) = |\lambda|^2\, I - \lambda\, T^* - \overline{\lambda}\, T + T^*\, T$$

como las expresiones de la derecha en ambas ecuaciones coinciden término a término (notemos que los últimos términos son $T\, T^* = T^*\, T$ pues T es normal!), concluimos que

$$(\lambda\, I - T)\ (\overline{\lambda}\, I - T^*) = (\overline{\lambda}\, I - T^*)\ (\lambda\, I - T)\,,$$

que muestra que $\lambda\, I - T$ es operador normal.

Para verificar el ítem 3, aplicamos el ítem 1. junto con el ítem 2: notemos que $\lambda\, I - T$ es operador normal; si $v \in \mathcal{V}$ es tal que $T\, v = \lambda\, v$ entonces $v \in \mathrm{N}(\lambda\, I - T)$; así, por el ítem 1, $v \in \mathrm{N}(\lambda\, I - T) = \mathrm{N}(\overline{\lambda}\, I - T^*)$ es decir, $\overline{\lambda}\, v - T^*\, v = 0$ lo que muestra que $T^*\, v = \overline{\lambda}\, v$. □

Teorema 12.42. *Sea $(\mathcal{V}, \langle\, ,\, \rangle)$ un $\mathbb{C}$-EPI, $\dim_{\mathbb{C}} \mathcal{V} = n \geq 1$ y sea $T \in L(\mathcal{V})$ operador normal. Entonces*

1. *$\sigma(T) \neq \emptyset$;*
2. *Sea λ, $\mu \in \sigma(T)$, $\lambda \neq \mu$: si $v \in E_T(\lambda)$, $w \in E_T(\mu)$ entonces: $v \perp w$.*

Demostración. Notemos que en el caso complejo, T tiene su espectro no vacío de forma trivial; en efecto, si consideramos el polinomio característico $p_T(x) \in \mathbb{C}[x]$ resulta un polinomio mónico tal que $\mathrm{gr}(p(x)) = n \geq 1$; por el teorema fundamental del álgebra, $p_T(x)$ admite una raíz $\lambda \in \mathbb{C}$; en este caso $\lambda \in \sigma(T)$ por construcción de $p_T(x)$.

Con la notación del ítem 2, calculamos

$$\lambda\langle v\,,\, w\rangle = \langle \lambda\, v\,,\, w\rangle = \langle T\, v\,,\, w\rangle = \langle v\,,\, T^*\, w\rangle = \langle v\,,\, \overline{\mu}\, w\rangle = \mu\, \langle v\,,\, w\rangle\,,$$

donde hemos usado las propiedades del producto escalar, la hipótesis $v \in E_T(\lambda)$, las propiedades del adjunto, el ítem 3 de la Proposición 12.41 y la hipótesis $w \in E_T(\mu)$. Lo anterior muestra que $(\lambda - \mu)\, \langle v\,,\, w\rangle = 0$ que a su vez prueba que $\langle v\,,\, w\rangle = 0$ (ya que $\lambda \neq \mu$). □

Definición 12.43. *Decimos que $A \in \mathbb{C}^{n\times n}$ es matriz normal si $A\, A^* = A^*\, A$.* □

Corolario 12.44. *Sea $(\mathcal{V}, \langle\, ,\, \rangle)$ un $\mathbb{C}$-EPI, $\dim_{\mathbb{C}} \mathcal{V} = n$ y sea B una* **base ortonormal** *de $\mathcal{V}$. Entonces $T \in L(\mathcal{V})$ operador normal si y solo si $[T]_B$ es matriz normal.*

Demostración. Ejercicio. □

En lo que sigue veremos que si una matriz es normal y triangular superior entonces debe ser matriz diagonal. Este hecho junto con un resultado general sobre la existencia de representaciones triangulares de cualquier operador permitirá probar que los operadores normales son unitariamente diagonalizables.

Proposición 12.45. *Sea $A \in \mathbb{C}^{n\times n}$ una matriz normal y triangular superior. Entonces A es matriz diagonal.*

Demostración. Notemos que, en general, para una matriz A arbitraria y $1 \leq j \leq n$,

$$(A^*A)_{jj} = \sum_{\ell=1}^{n} A^*_{j\ell}\, A_{\ell j} = \sum_{\ell=1}^{n} \overline{A_{\ell j}}\, A_{\ell j} = \sum_{\ell=1}^{n} |\, A_{\ell j}\,|^2 = \|c_j(A)\|^2$$

donde $c_j(A) \in \mathbb{C}^n$ denota la j-ésima columna de A y $\|v\|^2$ denota el cuadrado de la norma de $v \in \mathbb{C}^n$ inducida por el producto interno usual en $\mathbb{C}^n$. De forma forma similar

$$(A\,A^*)_{jj} = \sum_{\ell=1}^{n} A_{j\ell}\, A^*_{\ell j} = \sum_{\ell=1}^{n} A_{j\ell}\, \overline{A_{j\ell}} = \sum_{\ell=1}^{n} |\, A_{j\ell}\,|^2 = \|f_j(A)\|^2$$

donde $f_j(A) \in \mathbb{C}^n$ denota la j-ésima fila de A.

Así, si A es matriz normal, entonces $A^*A = A\,A^*$; en particular, concluimos que

$$\|c_j(A)\|^2 = (A^*\,A)_{jj} = (A\,A^*)_{jj} = \|f_j(A)\|^2 \quad \text{para} \quad 1 \leq j \leq n\,.$$

Si además A es matriz triangular superior, entonces se verifica que las entradas de A por debajo de la diagonal principal son nulas; es decir,

$$A_{jk} = 0 \quad \text{si} \quad 1 \leq k < j \leq n\,. \tag{97}$$

La condición anterior dice, por ejemplo, que si miramos la primer columna de A (con $k = 1$) entonces todas las entradas de la forma A_{j1} con $1 < j \leq n$ son nulas (hacer esquema de la matriz A y comparar con las condición anterior sobre las entradas de A). En particular, $\|c_1(A)\|^2 = |A_{11}|^2$; como $\|c_1(A)\|^2 = \|f_1(A)\|^2$ entonces

$$|A_{11}|^2 = \|c_1(A)\|^2 = \|f_1(A)\|^2 = \sum_{k=1}^{n} |A_{1k}|^2 = |A_{11}|^2 + \sum_{k=2}^{n} |A_{1k}|^2 \implies \sum_{k=2}^{n} |A_{1k}|^2 = 0$$

Como $|A_{1k}|^2 \geq 0$ y $\sum_{k=2}^{n} |A_{1k}|^2 = 0$ concluimos que $A_{1k} = 0$ para $2 \leq k \leq n$; es decir, en la primer fila, la segunda entrada en adelante son todas nulas! En términos de la matriz A, ahora sabemos que

$$A = \begin{pmatrix} A_{11} & 0 & 0 & \dots & 0 \\ 0 & A_{22} & A_{23} & \dots & A_{2n} \\ 0 & 0 & A_{33} & \dots & A_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 & A_{nn} \end{pmatrix}$$

(la novedad son los ceros de la primer fila!). La matriz A empieza a tener pinta diagonal! Seguimos con el argumento: si ahora miramos la columna 2, con la nueva información sabemos que

$$\|c_2(A)\|^2 = |A_{22}|^2$$

(mirar bien el aspecto de A arriba). Argumentamos como antes: como $\|c_2(A)\|^2 = \|f_2(A)\|^2$ entonces

$$|A_{22}|^2 = \|c_2(A)\|^2 = \|f_2(A)\|^2 = \sum_{k=2}^{n} |A_{2k}|^2 = |A_{22}|^2 + \sum_{k=3}^{n} |A_{2k}|^2 \implies \sum_{k=3}^{n} |A_{2k}|^2 = 0\,.$$

Como $|A_{2k}|^2 \geq 0$ y $\sum_{k=3}^{n} |A_{2k}|^2 = 0$ concluimos que $A_{2k} = 0$ para $3 \leq k \leq n$; es decir, en la segunda fila, desde la tercera entrada en adelante son todas nulas! Esto permite ahora actualizar el aspecto de A (como hicimos antes). Luego de haber actualizado el aspecto de A, si miramos la tercer columna vamos a ver que $\|c_3(A)\|^3 = |A_{33}|^2$ (hacer esquema de A con la nueva información); argumentamos como antes y concluimos que $A_{3k} = 0$ para $4 \leq k \leq n$. A esta altura debe quedar claro que es posible seguir con este argumento con las columnas y filas siguientes. Finalmente, podemos concluir que dada $1 \leq j \leq n$ entonces se verifica que

$$A_{jk} = 0 \quad \text{si} \quad j+1 \leq k \leq n\,.$$

Este último hecho, junto con la condición en la Eq. (97) ahora prueban que $A_{jk} = 0$ si $1 \leq j \neq k \leq n$, de forma que A es matriz diagonal. □

El siguiente resultado, conocido como triangulazación de Schur, es un resultado fundamental dentro del álgebra lineal (este resultado tiene otras versiones en cuerpos arbitrarios).

Teorema 12.46. *Sea $(\mathcal{V}, \langle\,,\,\rangle)$ un $\mathbb{C}$-EPI,* $\dim_{\mathbb{C}} \mathcal{V} = n \geq 1$ *y sea $T \in L(\mathcal{V})$ operador arbitrario. Entonces existe una* **base ortonormal** *B de $\mathcal{V}$ tal que $[T]_B \in \mathbb{C}^{n\times n}$ es una matriz triangular superior.*

Demostración. Vamos a probar el resultado por inducción en n.

Si $n = 1$ entonces el resultado es esencialmente trivial ($[T]_B \in \mathbb{C}$!)

Supongamos que el resultado es cierto para todo $\mathbb{C}$-EPI de dimensión $\leq n-1$ y todo operador que actúe en el espacio (hipótesis inductiva).

Sea $(\mathcal{V}, \langle\,,\,\rangle)$ un $\mathbb{C}$-EPI, $\dim_{\mathbb{C}} \mathcal{V} = n \geq 2$ y sea $T \in L(\mathcal{V})$: sea $\lambda \in \sigma(T^*) \neq \emptyset$ ($p_{T^*}(x) \in \mathbb{C}[x]$ es polinomio de grado al menos 1, de forma que admite al menos una raíz compleja que resulta ser autovalor de T^*). Sea $v \in \mathcal{V}$ tal que $\|v\| = 1$ y $T^*\, v = \lambda\, v$. Notemos que si definimos $\mathcal{W} = \overline{\{v\}}$ el subespacio generado por v, entonces $\mathcal{W}$ es T^*-invariante (ejercicio!).

Por el Lema 12.28, $\mathcal{W}^\perp$ es un subespacio $(T^*)^*$-invariante; como $(T^*)^* = T$ entonces podemos considerar la restricción $T|_{\mathcal{W}^\perp} \in L(\mathcal{W}^\perp)$. Más aún, como $\dim_{\mathbb{C}} \mathcal{W}^\perp = n - \dim_{\mathbb{C}} \mathcal{W} = n-1 \geq 1$ entonces podemos aplicar la hipótesis inductiva en el $\mathbb{C}$-EPI $\mathcal{W}^\perp$ dotado del producto interno de $\mathcal{V}$ restringido a $\mathcal{W}^\perp$, y al operador $T|_{\mathcal{W}^\perp} \in L(\mathcal{W}^\perp)$. En particular, existe $B' = \{v_1, \ldots, v_{n-1}\}$ una base ortonormal de $\mathcal{W}^\perp$ tal que $[T|_{\mathcal{W}\perp}]_{B'} \in \mathbb{C}^{(n-1)\times(n-1)}$ es triangular superior.

Para poder aprovechar la existencia de la base B', vamos a re-escribir qué significa que $[T|_{\mathcal{W}\perp}]_{B'} \in \mathbb{C}^{(n-1)\times(n-1)}$ sea una matriz triangular superior.

Si miramos la primer columna de $[T|_{\mathcal{W}^\perp}]_{B'} = (a_{jk})_{j,k=1}^{n-1}$, vemos que es un vector de la forma $(a_{11}, 0, \ldots, 0) \in \mathbb{C}^{n-1}$, pues la matriz es triangular superior. Como este vector es, por definición, el vector de coordenadas $[T|_{\mathcal{W}^\perp} v_1]_{B'}$ concluimos que

$$T\, v_1 = T|_{\mathcal{W}^\perp}\, v_1 = a_{11}\, v_1\,.$$

En general, si miramos la j-esima columna de $[T|_{\mathcal{W}^\perp} v_1]_{B'}$ encontramos un vector de la forma $(a_{1j}, \ldots, a_{jj}, 0, \ldots, 0) \in \mathbb{C}^{n-1}$, pues la matriz es triangular superior. Como este vector es, por definición, el vector de coordenadas $[T|_{\mathcal{W}^\perp} v_j]_{B'}$ concluimos que

$$T\, v_j = T|_{\mathcal{W}^\perp}\, v_j = \sum_{h=1}^{j} a_{hj}\, v_h \quad \text{para} \quad 1 \leq j \leq n-1\,. \tag{98}$$

Si definimos $v_n = v$ (el vector con el que comenzamos el agumento) entonces $B = \{v_1, \ldots, v_{n-1}, v_n\}$ es base ortonormal de $\mathcal{V}$ (recordemos que $v \in \mathcal{W}$ de forma que $v_n = v \perp v_j$ para $1 \leq j \leq n-1$). Con estas definiciones, la Eq. (98) ahora se puede escribir como

$$T\, v_j = \sum_{h=1}^{j} a_{hj}\, v_h + \sum_{h=j+1}^{n} 0\, v_h \quad \text{para} \quad 1 \leq j \leq n-1\,, \tag{99}$$

lo que muestra que la j-ésima columna de $[T]_B$ es $[Tv_j]_B = (a_{1j}, \ldots, a_{jj}, 0, \ldots, 0) \in \mathbb{C}^n$, para $1 \leq j \leq n-1$. Así, la matriz $[T]_B$ es una matriz triangular superior! (hacer esquema de la matriz de $[T]_B$). $\square$

Ahora tenemos todas las herramientas para deducir la siguiente caracterización de los operadores unitariamente diagonalizables en el caso complejo.

Teorema 12.47. *Sea $(\mathcal{V}, \langle\,,\,\rangle)$ un $\mathbb{C}$-EPI, $\dim_{\mathbb{C}} \mathcal{V} = n \geq 1$ y sea $T \in L(\mathcal{V})$. Entonces T es unitariamente diagonalizable si y solo si T es operador normal.*

Demostración. Hemos visto en la Observación 12.39 que una condición necesaria para que un operador T sea unitariamente diagonalizable es que sea normal. Recíprocamente, si T es un operador normal, el Teorema de triangulaón de Schur 12.46 implica que existe una base ortonormal B de $\mathcal{V}$ tal que $[T]_B$ es una matriz triangular superior. Por el Corolario 12.44 vemos que $[T]_B$ es una matriz normal; pero la Proposición 12.45 ahora prueba que en este caso $[T]_B$ es una matriz diagonal. Lo anterior muestra que T es unitariamente diagonalizable. $\square$

Obs 12.48. Sea $(\mathcal{V}, \langle\,,\,\rangle)$ un $\mathbb{C}$-EPI, $\dim_{\mathbb{C}} \mathcal{V} = n \geq 1$ y sea $T \in L(\mathcal{V})$ operador normal. El resultado anterior garantiza que existe una base ortonormal B de $\mathcal{V}$ tal que $[T]_B$ es matriz diagonal. Pero cómo hallar esta base B? Consideremos (todos) los autovalores $\lambda_1, \ldots, \lambda_k \in \mathbb{C}$ distintos de T; entonces se debe verificar que

$$\sum_{j=1}^{k} \dim_{\mathbb{C}} E_T(\lambda_j) = n$$

pues, en particular, T es diagonalizable!! (de forma que, en particular, valen los resultados para operadores diagonalizables).

Pero además, los autoespacios $E_T(\lambda_1), \ldots, E_T(\lambda_k)$ son mutuamente ortogonales: es decir, si $1 \leq j \neq h \leq k$ y $v \in E_T(\lambda_j)$, $w \in E_T(\lambda_h)$ entonces $\langle v\,,\, w\rangle = 0$ (ver el ítem 2 Teorema 12.42).

Así, sea B_j una base (cualquiera) de $E_T(\lambda_j)$, $1 \leq j \leq k$; notar que $\#(B_j) = \dim_{\mathbb{C}} E_T(\lambda_j)$. Si aplicamos el proceso de ortogonalización de Gram-Schmidt al conjunto (l.i.) B_j obtenemos el conjunto B_j'; luego normalizamos los vectores de B_j' para obtener finalmente B_j'' que es una base ortonormal de $E_T(\lambda_j)$ (puesto que B_j'' y B_j generan el mismo subespacio, por las propiedades del proceso de Gram-Schmidt). Los comentarios anteriores ahora muestran que el conjunto

$$B = B_1'' \cup \ldots \cup B_k''$$

es una base ortonormal de $\mathcal{V}$ (pues es el pegado ó yuxtaposición) de los sistemas ortonormales B_j'', $1 \leq j \leq k$, que además son mutuamente ortogonales entre sí! Finalmente, $[T]_B$ es una matriz diagonal, pues cada vector de B proviene de algún B_j'', y todos los vectores de $B_j'' \subset E_T(\lambda_j)$ son autovectores de T (asociados a λ_j). $\triangle$

Corolario 12.49. *Sea $A \in \mathbb{C}^{n\times n}$. Las siguientes afirmaciones son equivalentes:*

1. *A es unitariamente equivalente a una matriz diagonal;*
2. *A es matriz normal;*
3. *$T_A \in L(\mathbb{C}^n)$ es operador normal ($T_A\, \vec{x} = A\, \vec{x}$, para $\vec{x} \in \mathbb{C}^n$).*

Demostración. $1 \Rightarrow 2$. Supongamos que A es unitariamente equivalente a una matriz diagonal; entonces existe una matriz unitaria $W \in \mathbb{C}^{n\times n}$ y una matriz diagonal $B = \operatorname{diag}(b_1, \ldots, b_n) \in \mathbb{C}^{n\times n}$ tales que $A = W^*\, B\, W$. En este caso, por la Observación 12.11

$$A^* = (W^*\, B\, W)^* = W^*\, B^*\, (W^*)^* = W^*\, B^*\, W$$

Entonces

$$A^*A = (W^*\,B^*\,W)\,(W^*\,B\,W) = W^*\,B^*\,B\,W = W^*\,B\,B^*\,W$$

donde hemos usado que, como B es matriz diagonal entonces B^* también es matriz diagonal de forma que se verifica $B^*\,B = B\,B^*$ (verificar!) y que $W\,W^* = I$ ya que W es matriz unitaria. De forma similar

$$AA^* = (W^*\,B\,W)\,(W^*\,B^*\,W) = W^*\,B\,B^*\,W\,.$$

Las identidades anteriores muestran que $A^*A = A\,A^*$, y A es matriz normal.

2. $\Rightarrow$ 3. Basta notar que la base canónica B_c de $\mathbb{C}^n$ es una base ortonormal para el producto interno usual y que $[T_A]_{B_c} = A$. Entonces podemos aplicar el Corolario 12.44.

3. $\Rightarrow$ 1. Supongamos que $T_A \in L(\mathbb{C}^n)$ es normal, en donde consideramos el producto interno usual de $\mathbb{C}^n$. El Teorema 12.47 implica que existe una base ortonormal $B = \{v_1, \ldots, v_n\}$ de $\mathbb{C}^n$ tal que $[T_A]_B$ es matriz diagonal. Sea $B_c = \{e_1, \ldots, e_n\}$ la base canónica de $\mathbb{C}^n$. Entonces, sabemos que

$$[T_A]_B = M_{B_c,B}^{-1}\,[T_A]_{B_c}\,M_{B_c,B} = M_{B_c,B}^{-1}\,A\,M_{B_c,B}\,,$$

donde usamos que $[T_A]_{B_c} = A$. Por Observación 12.23 vemos que $M_{B_c,B} \in \mathbb{C}^{n\times n}$ es matriz unitaria, y en particular que $M_{B_c,B}^{-1} = M_{B_c,B}^*$ de forma que la identidad anterior se puede escribir $[T_A]_B = W^*\,A\,W$, para W matriz unitaria y $[T_A]_B$ matriz diagonal. $\square$

Obs 12.50. Sea $A \in \mathbb{C}^{n\times n}$ matriz normal. Podemos aplicar las observaciones realizadas en la prueba del Corolario 12.49 junto con la Observación 12.48 para hallar una matriz unitaria $W \in \mathbb{C}^{n\times n}$ tal que W^*AW sea matriz diagonal. De hecho, como $T_A \in L(\mathbb{C}^n)$ resulta normal, podemos aplicar el método descripto en la Observación 12.48 para hallar una base ortonormal B tal que $[T_A]_B$ sea matriz diagonal. La prueba del Corolario 12.49 indica que podemos tomar $W = M_{B_c,B}$; como observación final, notemos que esta matriz se calcula de forma sencilla (una vez que conocemos B): porqué? $\triangle$

Obs 12.51. Sea $(\mathcal{V}, \langle\,,\,\rangle)$ un $\mathbb{C}$-EPI, $\dim_{\mathbb{C}} \mathcal{V} = n \geq 1$ y sea $T \in L(\mathcal{V})$ un operador tal que $T = T^*$. En este caso decimos que T es *autoadjunto* (así como en el caso de espacios reales). Notemos que entonces $T^*T = T^2 = TT^*$, de forma que T resulta normal. En particular, T es unitariamente diagonalizable. En este caso $\sigma(T) \subset \mathbb{R}$: en efecto, si $\mu \in \sigma(T)$ entonces existe $v \in \mathcal{V}$ tal que $v \neq 0$ y $T\,v = \mu\,v$: en este caso,

$$\mu\,\langle v\,,\,v\rangle = \langle \mu\,v\,,\,v\rangle = \langle T\,v\,,\,v\rangle = \langle v\,,\,T^*\,v\rangle = \langle v\,,\,\mu\,v\rangle = \overline{\mu}\,\langle v\,,\,v\rangle\,.$$

Así, $\mu\,\langle v\,,\,v\rangle = \overline{\mu}\,\langle v\,,\,v\rangle$; como $v \neq 0$ entonces $\langle v\,,\,v\rangle > 0$ y deducimos que $\mu = \overline{\mu}$ de forma que $\mu \in \mathbb{R}$. $\triangle$

Ejemplo 12.52. Cabe mencionar que existen operadores normales U en espacios con producto interno **reales** (en el sentido que $U^*U = U\,U^*$), que no son diagonalizables mediante una base ortonormal! Por ejemplo, el operador $U \in L(\mathbb{R}^2)$ que actúa sobre el EPI real $\mathbb{R}^2$ con el producto interno usual, dado por $U(x,y) = (-y,x)$ verifica: $U^*(x,y) = (y,-x)$ de forma que $U^* = U^{-1}$; en particular $U^*U = U\,U^* = I$. En este caso, el polinomio característico de U está dado por $p_U(x) = x^2+1$, de forma que U no tiene autovalores (como operador en $\mathbb{R}^2$, porque el característico no tiene raíces reales). En particular, el operador normal U no es diagonalizable en este espacio real.

Es por esto que debemos considerar (y recordar con cuidado) los casos real y complejo separadamente, al momento de caracterizar los operadores que son unitariamente diagonalizables. $\triangle$

Consideremos dos consecuencias del teorema de triangulación de Schur.

Teorema 12.53. *Recordemos que dada* $A = (a_{jk})_{j,k=1}^n \in \mathbb{C}^{n\times n}$ *entonces* $\operatorname{tr}(A) = \sum_{j=1}^n A_{jj}$. *Se verifican las siguientes propiedades:*

1. *Si* $A,\, B \in \mathbb{C}^{n\times n}$ *y* $\alpha \in \mathbb{C}$ *entonces*

$$\operatorname{tr}(\alpha\, A + B) = \alpha \operatorname{tr}(A) + \operatorname{tr}(B) \quad y \quad \operatorname{tr}(AB) = \operatorname{tr}(BA)\,.$$

2. *Si los autovalores de* A*, contando multiplicidades, son* $\lambda_1, \ldots, \lambda_n \in \mathbb{C}$ *entonces*

$$\operatorname{tr}(A) = \sum_{j=1}^n \lambda_j \quad y \quad \det(A) = \prod_{j=1}^n \lambda_j\,.$$

Demostración. La primer identidad del ítem 1 (que es la linealidad de la traza) es ejercicio. Para verificar la segunda identidad, sean $A = (a_{jk})_{j,k=1}^n$, $B = (b_{jk})_{j,k=1}^n \in \mathbb{C}^{n\times n}$. Entonces

$$\operatorname{tr}(AB) = \sum_{j=1}^n (AB)_{jj} = \operatorname{tr}(AB) = \sum_{j=1}^n \sum_{k=1}^n A_{jk}\, B_{kj} = \sum_{k=1}^n \sum_{j=1}^n B_{kj}\, A_{jk} = \operatorname{tr}(BA)\,,$$

donde hemos intercambiado el orden de sumación y conmutado los factores A_{jk} y B_{kj} de $\mathbb{C}$.

Para verificar el ítem 2, dada $A \in \mathbb{C}^{n\times n}$ consideremos el operador $T_A \in L(\mathbb{C}^n)$ dado por $T_A\, \vec{x} = A\, \vec{x}$. Si $B_c = \{e_1, \ldots, e_n\}$ denota la base canónica de $\mathbb{C}^n$ entonces $[T_A]_{B_c} = A$. Por el Teorema 12.46 existe una base ortonormal $B = \{v_1, \ldots, v_n\}$ tal que $[T_A]_B = (t_{jk})_{j,\,k=1}^n$ es matriz triangular superior. En particular, si calculamos el polinomio característico de T usando la base B obtenemos

$$p_T(x) = \det(x\, I - [T_A]_B) = \prod_{j=1}^n (x - t_{jj})\,,$$

pues $x\, I - [T_A]_B$ es una matriz triangular superior, con diagonal principal dada por $(x - t_{11}, \ldots, x - t_{nn})$; así, el determinante de esta matriz triangular superior es el producto de las entradas en su diagonal principal, lo que muestra la identidad anterior.

En particular, (todas) las raíces de $p_T(x)$ son $\lambda_1 = t_{11}, \ldots, \lambda_n = t_{nn}$; en la lista anterior permitimos repeticiones: es decir, contamos las raíces de $p_T(x)$ (autovalores de A) con sus respectivas multiplicidades.

Finalmente, recordemos que $A = [T_A]_{B_c} = M_{B,B_c}^{-1}\, [T_A]_B\, M_{B,B_c}$ de forma que

$$\det(A) = \det(M_{B,B_c}^{-1}\, [T_A]_B\, M_{B,B_c}) = \det([T_A]_B) = \prod_{j=1}^n t_{jj} = \prod_{j=1}^n \lambda_j$$

pues $[T_A]_B$ es matriz triangular; además se verifica

$$\operatorname{tr}(A) = \operatorname{tr}(M_{B,B_c}^{-1}\, ([T_A]_B\, M_{B,B_c})\,) = \operatorname{tr}(\, ([T_A]_B\, M_{B,B_c})\, M_{B,B_c}^{-1}\,) = \operatorname{tr}([T_A]_B) = \sum_{j=1}^n t_{jj} = \sum_{j=1}^n \lambda_j\,,$$

donde hemos usado la segunda propiedad del ítem 1 en la segunda igualdad.

□

12.4 Sobre la diagonalización simultánea con respecto a bases ortonormales

En esta sección consideramos el problema de diagonalizar simultáneamente un par de operadores con respecto a la misma base ortonormal. Naturalmente, desarrollamos los casos real y complejo por separado.

Teorema 12.54. *Sea $(\mathcal{V}, \langle\,,\,\rangle)$ un $\mathbb{R}$-EPI, $\dim_{\mathbb{R}} \mathcal{V} = n$, y sean $S, T \in L(\mathcal{V})$ operadores autoadjuntos. Las siguientes condiciones son equivalentes:*

1. *existe una base ortonormal B de $\mathcal{V}$ tal que $[T]_B$ y $[S]_B$ son matrices diagonales.*
2. *$ST = TS$.*

Demostración. Si vale el ítem 1, entonces se verifica que

$$[ST]_B = [S]_B\,[T]_B = [T]_B\,[S]_B = [TS]_B$$

donde hemos usado que las matrices diagonales $[T]_B$ y $[S]_B$ conmutan entre sí; la identidad $[ST]_B = [TS]_B$ implica que $ST = TS$.

Recíprocamente, si vale el ítem 2, entonces como S y T son diagonalizables y conmutan entre sí, hemos probado que existe una base (no necesariamente ortonormal) $B' = \{w_1, \ldots, w_n\}$ de $\mathcal{V}$ tal que $[S]_{B'}$, $[T]_{B'} \in \mathbb{R}^{n\times n}$ son matrices diagonales. En particular, existen $t_1, \ldots, t_n \in \mathbb{R}$ y $s_1, \ldots, s_n \in \mathbb{R}$ tales que $T\,w_j = t_j\,w_j$ y $S\,w_j = s_j\,w_j$, $1 \leq j \leq n$.

Definimos la siguiente relación de equivalencia en el conjunto $\{1, \ldots, n\}$: si $1 \leq j, k \leq n$ entonces $j \approx k$ si $t_j = t_k$ y $s_j = s_k$. Esta relación de equivalencia parte al conjunto $\{1, \ldots, n\}$ en clases de equivalencia $\{I_1, \ldots, I_p\}$; aquí, $j \approx k$ si y solo si $j, k \in I_\ell$ para algún $1 \leq \ell \leq p$.

Con las notaciones del párrafo anterior, notemos que si $1 \leq \ell \leq p$ entonces existen λ_ℓ, $\mu_\ell \in \mathbb{R}$ tales que $T\,w_j = \lambda_\ell\,w_j$, $S\,w_j = \mu_\ell\,w_j$ para $j \in I_\ell$; de hecho, si elegimos un elemento cualquiera $j_0 \in I_\ell$ entonces $\lambda_\ell = t_{j_0}$ y $\mu_\ell = s_{j_0}$!! En efecto, por definición de nuestra relación de equivalencia: $T\,w_j = t_j\,w_j = t_{j_0}\,w_j$ y de forma similar $S\,w_j = s_j\,w_j = s_{j_0}\,w_j$ siempre que $j \in I_\ell$ (ó equivalentemente, siempre que $j \approx j_0$).

Para $1 \leq \ell \leq p$ definimos $\mathcal{W}_\ell = \overline{\{w_j :\ j \in I_\ell\}}$, el subespacio generado por los vectores $\{w_j :\ j \in I_\ell\}$ (cuyos sub-índices pertenecen a la clase I_ℓ). Notemos que como estos subespacios están generados por conjuntos de vectores que forman una partición de la base B' entonces

$$\mathcal{V} = \mathcal{W}_1 \dot{+} \ldots \dot{+} \mathcal{W}_p$$

que es una descomposición de $\mathcal{V}$ en suma directa de subespacios independientes. Además, si $v \in \mathcal{W}_\ell$ entonces

$$v = \sum_{j\in I_\ell} \alpha_j\,w_j \implies Tv = \sum_{j\in I_\ell} \alpha_j\,T\,w_j = \sum_{j\in I_\ell} \alpha_j\,\lambda_\ell\,w_j = \lambda_\ell \sum_{j\in I_\ell} \alpha_j\,w_j = \lambda_\ell\,v$$

donde $\lambda_\ell \in \mathbb{R}$ es tal que $Tw_j = \lambda_\ell\,w_j$ para $j \in I_\ell$ (y no depende de la variable de sumación, de forma que podemos *sacar factor común* a este escalar). De forma análoga, si $v \in \mathcal{W}_\ell$ entonces

$$S\,v = \mu_\ell\,v \quad (\text{ y } \ T\,v = \lambda_\ell\,v \text{ como probamos antes }) .$$

Afirmamos que si $1 \leq \ell \neq k \leq p$ entonces $\mathcal{W}_\ell \perp \mathcal{W}_k$: de hecho, como estos son subespacios generados por familias de vectores que describimos previamente, basta mostrar que: si $j \in I_\ell$ y $h \in I_k$ entonces $w_j \perp w_h$ (verificar que basta con probar este último hecho). Pero en este caso, como $j \in I_\ell$ y $h \in I_k$ con $\ell \neq k$ entonces se debe verificar que

$$t_j \neq t_h \quad \text{ó} \quad s_j \neq s_h\,.$$

Si $t_j \neq t_h$ entonces, como $w_j \in E_T(t_j)$ y $w_h \in E_T(t_h)$ concluimos que $w_j \perp w_h$, por el Teorema 12.33. De forma similar, si $s_j \neq s_h$ entonces, como $w_j \in E_S(s_j)$ y $w_h \in E_S(s_h)$ concluímos que $w_j \perp w_h$, por el Teorema 12.33.

Finalmente, si construimos una base ortonormal B_j para $\mathcal{W}_j$, entonces los elementos de B_j son autovectores de T (con autovalor λ_j) y autovectores de S (con autovalores μ_j), por un párrafo anterior. Así, el pegado $B = B_1 \cup \ldots \cup B_p$ es base ortonormal de $\mathcal{V}$ formada por autovectores simultáneos de T y S, de forma que $[T]_B$ y $[S]_B$ son matrices diagonales. $\square$

De forma similar (con un argumento completamente análogo para la prueba) obtenemos

Teorema 12.55. *Sea $(\mathcal{V}, \langle\,,\,\rangle)$ un $\mathbb{C}$-EPI, $\dim_{\mathbb{C}} \mathcal{V} = n$, y sean S, T operadores normales. Las siguientes condiciones son equivalentes:*

1. *existe una base ortonormal B de $\mathcal{V}$ tal que $[T]_B$ y $[S]_B$ son matrices diagonales.*

2. *$ST = TS$.*

$\square$

13 Formas bilineales

Comenzamos con la definción del concepto central de estas notas

Definición 13.1. *Sea $\mathcal{V}$ un K-ev. Una forma bilineal sobre $\mathcal{V}$ es una función $\mathcal{A} : \mathcal{V} \times \mathcal{V} \to K$ que verifica: si $\alpha \in K$, u, v, $w \in \mathcal{V}$:*

1. *$\mathcal{A}(\alpha\, u + v\,,\, w) = \alpha\, \mathcal{A}(u\,,\, w) + \mathcal{A}(v\,,\, w)$;*

2. *$\mathcal{A}(u\,,\, \alpha\, v + w) = \alpha\, \mathcal{A}(u\,,\, v) + \mathcal{A}(u\,,\, w)$.*

$\square$

Obs 13.2. Con las notaciones de la definición anterior, una función $\mathcal{A} : \mathcal{V} \times \mathcal{V} \to K$ es una forma bilineal si dado $u \in \mathcal{V}$ arbitrario, entonces las funciones

$$\mathcal{A}(\cdot\,,\, u) : \mathcal{V} \to K \quad \text{y} \quad \mathcal{A}(u\,,\, \cdot) : \mathcal{V} \to K$$

son funcionales lineales (que corresponde a las condiciones 1 y 2 de la Definición 13.1). En particular, si $\alpha_j \in K$, $v_j \in \mathcal{V}$ para $1 \leq j \leq m$ entonces

$$\mathcal{A}(u\,,\, \sum_{j=1}^{m} \alpha_j\, v_j) = \sum_{j=1}^{m} \alpha_j\, \mathcal{A}(u\,,\, v_j) \quad \text{y} \quad \mathcal{A}(\sum_{j=1}^{m} \alpha_j\, v_j\,,\, u) = \sum_{j=1}^{m} \alpha_j\, \mathcal{A}(v_j\,,\, u)\,.$$

$\triangle$

Ejemplos 13.3. Sea K un cuerpo.

1. Si $\mathcal{V} = K^2$ como K-ev, entonces las funciones dadas por

$$\mathcal{A}((x_1, x_2)\,,\, (y_1, y_2)) = (x_1 + y_1)^2 - (x_1 - y_1)^2 + x_2\, y_2$$

$$\mathcal{B}((x_1, x_2)\,,\, (y_1, y_2)) = x_1\, y_2 + x_2\, y_1$$

son formas bilineales sobre $\mathcal{V}$ (verificar, ejercicio).

2. Dado $\mathcal{V}$ un K-ev, $\dim_K \mathcal{V} = n$ y dados $f,\, g \in \mathcal{V}^*$ (espacio dual) entonces la función dada por
$$\mathcal{A}(u\,,\,v) = f(u)\; g(v)$$
es una forma bilineal sobre $\mathcal{V}$ (verificar, ejercicio).

3. Si $\mathcal{V} = K^n$ y $A \in K^{n\times n}$ entonces la función dada por
$$\mathcal{A}(\vec{x}\,,\,\vec{y}) = (\vec{x})^t\, A\, \vec{y}$$
donde $(\vec{y})^t \in K^{1\times n}$ denota la transpuesta de $\vec{y} \in K^n = K^{n\times 1}$, es una forma bilineal sobre $\mathcal{V}$ (ejercicio). Notar que el producto de las tres matrices está bien definido (en términos de los tamaños de los vectores) y es una matriz en $K^{1\times 1} = K$.

4. Sea $(\mathcal{V}, \langle\,,\,\rangle)$ un $\mathbb{R}$-EPI de dimensión finita. Entonces el producto interno $\langle\cdot\,,\,\cdot\rangle : \mathcal{V}\times\mathcal{V} \to \mathbb{R}$ es una forma bilineal (ejercicio).

$\triangle$

Definición 13.4. *Sea $\mathcal{V}$ un K-ev. El conjunto formado por todas las formas bilineales sobre $\mathcal{V}$ es notado Bil($\mathcal{V}$). Si $\mathcal{A}$, $\mathcal{B} \in$ Bil($\mathcal{V}$) y $\alpha \in K$ entonces definimos*
$$\alpha\,\mathcal{A} + \mathcal{B} : \mathcal{V}\times\mathcal{V} \to K \qquad \textit{dada por} \qquad \alpha\,\mathcal{A} + \mathcal{B}(u\,,\,v) = \alpha\,\mathcal{A}(u\,,\,v) + \mathcal{B}(u\,,\,v)\,.$$

$\square$

Proposición 13.5. *Sea $\mathcal{V}$ un K-ev. Entonces Bil($\mathcal{V}$) dotado de la suma y la acción escalar descriptas en la Definición 13.4 es un K-ev.*

Demostración. Comenzamos notando que si $\alpha \in K$ y $\mathcal{A}$, $\mathcal{B} \in \text{Bil}(\mathcal{V})$ entonces $\alpha\,\mathcal{A} + \mathcal{B} \in \text{Bil}(\mathcal{V})$. De hecho, si $\beta \in K$ y $u,\, v,\, w \in \mathcal{V}$ entonces
$$(\alpha\,\mathcal{A} + \mathcal{B})(\beta\, u + v\,,\,w) = \alpha\,\mathcal{A}(\beta\, u + v\,,\,w) + \mathcal{B}(\beta\, u + v\,,\,w) =$$
$$\alpha\,(\beta\,\mathcal{A}(u\,,\,w) + \mathcal{A}(v\,,\,w)) + \beta\,\mathcal{B}(u\,,\,w) + \mathcal{B}(v\,,\,w) =$$
$$\beta\;(\alpha\,\mathcal{A} + \mathcal{B})(u\,,\,w) + (\alpha\,\mathcal{A} + \mathcal{B})(v\,,\,w)$$
los detalles para obtener la última desigualdad quedan a cargo del lector. De forma completamente análoga se verifica que
$$(\alpha\,\mathcal{A} + \mathcal{B})(u\,,\,\beta\, v + w) = \beta\;(\alpha\,\mathcal{A} + \mathcal{B})(u\,,\,v) + (\alpha\,\mathcal{A} + \mathcal{B})(u\,,\,w)\,.$$
Las identidades anteriores muestran que $\alpha\,\mathcal{A} + \mathcal{B} \in \text{Bil}(\mathcal{V})$. En particular, queda definida la suma en Bil($\mathcal{V}$) y la acción escalar de K en Bil($\mathcal{V}$). Queda como ejercicio para el lector verificar que la suma y la acción escalar definidas en Bil($\mathcal{V}$) satisfacen las propiedades correspondientes a las de un K-ev. $\square$

Definición 13.6. *Sea $\mathcal{V}$ un K-ev y sea $B = \{v_1, \ldots, v_n\}$ una base de $\mathcal{V}$. Si $\mathcal{A} \in$ Bil($\mathcal{V}$) definimos la matriz de $\mathcal{A}$ con respecto a B, notada $[\mathcal{A}]_B \in K^{n\times n}$, dada por*
$$([\mathcal{A}]_B)_{j,k} = \mathcal{A}(v_j\,,\,v_k) \qquad \textit{para} \qquad 1 \le j,\, k \le n\,.$$

$\square$

Proposición 13.7. *Sea $\mathcal{V}$ un K-ev y sea $B = \{v_1, \ldots, v_n\}$ una base de $\mathcal{V}$. Dada $\mathcal{A} \in Bil(\mathcal{V})$ entonces $[\mathcal{A}]_B \in K^{n\times n}$ es la única matriz que verifica*

$$\mathcal{A}(u\,,\,v) = [u]_B^t\,[\mathcal{A}]_B\,[v]_B \quad \textit{para} \quad u,\,v \in \mathcal{V}\,.$$

Demostración. Sean $u,\,v \in \mathcal{V}$ y consideremos los vectores de coordenadas

$$[u]_B = (\alpha_1, \ldots, \alpha_n) \quad \text{y} \quad [v]_B = (\beta_1, \ldots, \beta_n) \in K^n = K^{n\times 1}\,.$$

Entonces podemos reconstruir a los vectores usando sus coordenadas y la base B, de forma que

$$u = \sum_{j=1}^n \alpha_j\;v_j \quad \text{y} \quad v = \sum_{k=1}^n \beta_k\;v_k\,.$$

Entonces, como $\mathcal{A}$ es forma bilineal, tenemos que

$$\begin{aligned}
\mathcal{A}(u\,,\,v) &= \mathcal{A}(\sum_{j=1}^n \alpha_j\;v_j\,,\,\sum_{k=1}^n \beta_k\;v_k) = \sum_{j=1}^n \alpha_j\;\mathcal{A}(v_j\,,\,\sum_{k=1}^n \beta_k\;v_k) = \sum_{j=1}^n \alpha_j\;\sum_{k=1}^n \beta_k\;\mathcal{A}(v_j\,,\,v_k)\\
&= \sum_{j=1}^n\sum_{k=1}^n \alpha_j\;\beta_k\;([\mathcal{A}]_B)_{jk} = \sum_{j=1}^n \alpha_j \sum_{k=1}^n([\mathcal{A}]_B)_{jk}\;\beta_k = \sum_{j=1}^n \alpha_j \sum_{k=1}^n([\mathcal{A}]_B\,[v]_B)_j\\
&= [u]_B^t\,([\mathcal{A}]_B\,[v]_B) = [u]_B^t\,[\mathcal{A}]_B\,[v]_B\,,
\end{aligned}$$

donde hemos usado la bilinealidad de $\mathcal{A}$ (ver la Observación 13.2), hemos reordenado el orden de sumación, y hemos usado las fórmulas para las entradas del producto de matrices (dos veces) y la asociatividad del producto de matrices (para omitir los últimos paréntesis). Para verificar la unicidad, supongamos que $C = (C_{jk})_{j,\,k=1}^n \in K^{n\times n}$ también verifica

$$\mathcal{A}(u\,,\,v) = [u]_B^t\,C\,[v]_B \quad \text{para} \quad u,\,v \in \mathcal{V}\,.$$

En particular, si elegimos $u = v_j$ y $v = v_k$ para $1 \le j,\,k \le n$ entonces, usando que $[u]_B = e_j = (\delta_{j\ell})_{\ell=1}^n$ y $[v]_B = e_k = (\delta_{k\ell})_{\ell=1}^n$ entonces

$$\mathcal{A}(v_j\,,\,v_k) = e_j^t\,C\,e_k = \sum_{h=1}^n\;\sum_{\ell=1}^n (e_j)_h\;C_{h\ell}\;(e_k)_\ell = \sum_{h=1}^n\;\sum_{\ell=1}^n\;\delta_{jh}\;C_{h\ell}\;\delta_{lk} = C_{jk}\,.$$

Lo anterior indica que $[\mathcal{A}]_{jk} = C_{jk}$ para $1 \le j,\,k \le n$, que prueba la unicidad de $[\mathcal{A}]_B$. □

Obs 13.8. Sea $\mathcal{V}$ un K-ev y sea $B = \{v_1, \ldots, v_n\}$ base de $\mathcal{V}$. Sea $\mathcal{A} \in \text{Bil}(\mathcal{V})$ y sea $[\mathcal{A}]_B = (a_{jk})_{j,k=1}^n \in K^{n\times n}$ la matriz de $\mathcal{A}$ con respecto a B. Si $[u]_B = (\alpha_1, \ldots, \alpha_n)$ y $[u]_B = (\beta_1, \ldots, \beta_n)$ entonces

$$\mathcal{A}(u\,,\,v) = [u]_B^t\,A\,[v]_B = \sum_{j=1}^n\;\sum_{k=1}^n \alpha_j\;a_{jk}\;\beta_k\,,$$

donde la última identidad se obtiene desarrollando el producto de las tres matrices (con la fórmula del producto). La verificación de esta identidad es ejercicio para el lector. △

Corolario 13.9. *Sea $\mathcal{V}$ un K-ev y sean B y B' dos bases de $\mathcal{V}$. Dada $\mathcal{A} \in Bil(\mathcal{V})$ entonces se verifica la siguiente relación entre las matrices de $\mathcal{A}$ asociadas a cada una de estas bases:*

$$[\mathcal{A}]_{B'} = (M_{B,B'})^t\,[\mathcal{A}]_B\,M_{B,B'}\,, \tag{100}$$

donde $M_{B,B'}$ denota la matriz de cambio de base de la base B' en la base B y $(M_{B,B'})^t$ denota la matriz transpuesta de la matriz $M_{B,B'}$.

Demostración. Vamos a verificar que la matriz que figura a la derecha de la Eq. (100) verifca formalmente la propiedad que determina a $[\mathcal{A}]_{B'}$ (ver la Proposición 13.7). En efecto, si $u,\, v \in \mathcal{V}$ entonces, recordemos las relaciones

$$[u]_B = M_{B,B'}\,[u]_{B'} \quad \text{y} \quad [v]_B = M_{B,B'}\,[v]_{B'}\,,$$

con lo que, en particular, usando las propiedades de transponer (como en Álgebra I)

$$[u]_B^t = (M_{B,B'}\,[u]_{B'})^t = [u]_{B'}^t\,(M_{B,B'})^t\,.$$

Entonces la matriz $(M_{B,B'})^t\,[\mathcal{A}]_B\,M_{B,B'}$ verifica

$$[u]_{B'}^t\,(M_{B,B'})^t\,[\mathcal{A}]_B\,M_{B,B'}\,[v]_{B'} = [u]_B^t\,[\mathcal{A}]_B\,[v]_B = \mathcal{A}(u\,,\,v)\,,$$

en donde hemos usado la propiedad de $[\mathcal{A}]_B$ en la última igualdad. Por la unicidad de la matriz $[\mathcal{A}]_{B'} \in K^{n\times n}$ concluimos que $[\mathcal{A}]_{B'} = (M_{B,B'})^t\,[\mathcal{A}]_B\,M_{B,B'}$.

□

Ejemplo 13.10. Consideremos $\mathcal{V} = \mathbb{R}^2$ como $\mathbb{R}$-ev, y sea $\mathcal{A} \in \mathrm{Bil}(\mathcal{V})$ la forma bilineal dada por

$$\mathcal{A}((x_1, x_2)\,,\,(y_1, y_2)) = x_1\,y_1 + x_1\,y_2 + x_2\,y_1 + x_2\,y_2$$

Si consideramos la base canónica $B_c = \{e_1,\, e_2\}$ de $\mathcal{V}$ entonces realizando las evaluaciones correspondientes $\mathcal{A}(e_j\,,\,e_k)$ obtenemos que

$$[\mathcal{A}]_{B_c} = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}\,.$$

Si ahora consideramos la base $B = \{(1,-1),(1,1)\}$ de $\mathcal{V}$ entonces es sencillo verificar que

$$M_{B_c,\,B} = \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}$$

Los cálculos anteriores junto con el Corolario 13.9 muestran que

$$[\mathcal{A}]_B = (M_{B_c,\,B})^t\,[\mathcal{A}]_{B_c}\,M_{B_c,\,B} = \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 4 \end{pmatrix}$$

donde la última identidad se obtiene desarrollando el producto de las matrices.

El hecho de que la matriz $[\mathcal{A}]_B$ sea diagonal sugiere que es conveniente considerar la representación de $\mathcal{A}$ en las coordenadas con respecto a la base B. En este caso, notemos que

$$M_{B,B_c} = (M_{B_c,B})^{-1} = \frac{1}{2}\begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}$$

en donde hemos invertido la matriz $M_{B_c,B}$ que calculamos previamente. Así,

$$[(x_1,\,x_2)]_B = \frac{1}{2}\begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} \frac{1}{2}(x_1 - x_2) \\ \frac{1}{2}(x_1 + x_2) \end{pmatrix}$$

Entonces, el cálculo anterior de $[\mathcal{A}]_B$ ahora indica que

$$\mathcal{A}((x_1, x_2)\,,\,(y_1, y_2)) = \begin{pmatrix} \frac{1}{2}(y_1 - y_2) \\ \frac{1}{2}(y_1 + y_2) \end{pmatrix}^t \begin{pmatrix} 0 & 0 \\ 0 & 4 \end{pmatrix} \begin{pmatrix} \frac{1}{2}(x_1 - x_2) \\ \frac{1}{2}(x_1 + x_2) \end{pmatrix} = (y_1 + y_2)\,(x_1 + x_2)$$

donde la última identidad se obtiene desarrollando el producto de las matrices. Finalmente, el lector puede ahora verificar (cosa que en este ejemplo es sencillo) que $\mathcal{A}((x_1, x_2)\,,\,(y_1, y_2)) = (y_1 + y_2)\,(x_1 + x_2)$ comparando esta expresión con la expresión que defina a $\mathcal{A}$ al comienzo del ejemplo. △

Más adelante vamos a hacer un estudio sistemático sobre el fenómeno detrás del ejemplo anterior.

Obs 13.11. Con las notaciones del Corolario 13.9, notemos la diferencia entre la fórmula para la matriz de una forma bilineal bajo un cambio de base del espacio con respecto a la fórmula para la matriz de un operador lineal bajo un cambio de base del espacio. Esta diferencia en las fórmulas indica una naturaleza distinta de los objetos que representan las matrices (aún cuando en ambos casos se usen matrices en $K^{n\times n}$). Más adelante vamos a considerar estas diferencias en un contexto general de tensores de tipo (p, q) sobre $\mathcal{V}$. △

Teorema 13.12. *Sea $\mathcal{V}$ un K-ev y sea $B = \{v_1, \ldots, v_n\}$ una base de $\mathcal{V}$. Entonces la función*

$$[\cdot]_B : Bil(\mathcal{V}) \to K^{n\times n} \quad \text{dada por} \quad \mathcal{A} \mapsto [\mathcal{A}]_B$$

es un isomorfismo lineal. En particular, $\dim_K Bil(\mathcal{V}) = n^2$.

Demostración. Verificamos primero que la función definida en el enunciado es lineal; para ello, sean decir $\alpha \in K$ y $\mathcal{A}$, $\mathcal{B} \in \mathrm{Bil}(\mathcal{V})$ entonces,

$$([\alpha\,\mathcal{A} + \mathcal{B}]_B)_{jk} = (\alpha\,\mathcal{A} + \mathcal{B})(v_j\,,\,v_k) = \alpha\,\mathcal{A}(v_j\,,\,v_k) + \mathcal{B}(v_j\,,\,v_k) = (\alpha\,[\mathcal{A}]_B + [\mathcal{B}]_B)_{jk}\,,$$

donde hemos usado la definición de la forma bilineal $\alpha\,\mathcal{A} + \mathcal{B}$ (ver Definición 13.4) y la definición de matriz asociada a una forma bilineal con respecto a un base (tres veces!). Así,

$$[\alpha\,\mathcal{A} + \mathcal{B}]_B = \alpha\,[\mathcal{A}]_B + [\mathcal{B}]_B$$

lo que muestra que la función definida en el enunciado es lineal. Si $\mathcal{A}$, $\mathcal{B} \in \mathrm{Bil}(\mathcal{V})$ son tales que $[\mathcal{A}]_B = [\mathcal{B}]_B$ entonces, por la Proposición 13.7, si u, $v \in \mathcal{V}$:

$$\mathcal{A}(u\,,\,v) = [u]_B^t\,[\mathcal{A}]_B\,[v]_B = [u]_B^t\,[\mathcal{B}]_B\,[v]_B = \mathcal{B}(u\,,\,v)\,,$$

de forma que $\mathcal{A} = \mathcal{B}$. Así, la función definida en el enunciado es monomorfismo. Finalmente, si $A \in K^{n\times n}$ podemos definir la función $\mathcal{A} : \mathcal{V} \times \mathcal{V} \to K$ dada por

$$\mathcal{A}(u\,,\,v) = [u]_B^t\,A\,[v]_B \in K \quad \text{para} \quad u,\,v \in \mathcal{V}\,.$$

Entonces $\mathcal{A}$ es forma bilineal tal que $[\mathcal{A}]_B = A$ (ejercicio. Sugerencia: recordar que tomar coordenadas con respecto a una base es transformación lineal y las propiedades de linealidad del producto de matrices). Los hechos anteriores indican que la función definida en el enunciado es isomorfismo. □

Definición 13.13. *Sea $\mathcal{V}$ un K-ev, $\dim_K \mathcal{V} = n$ y sea $\mathcal{A} \in Bil(\mathcal{V})$. Definimos el rango de $\mathcal{A}$ como el rango de la matriz $[\mathcal{A}]_B \in K^{n\times n}$, donde B denota una base (arbitraria de $\mathcal{V}$).* □

Obs 13.14. Con las notaciones de la definición anterior, notemos que el rango de $\mathcal{A}$ está bien definido, en el sentido que no depende de la base B de $\mathcal{V}$. En efecto, si B' es otra base de $\mathcal{V}$ entonces, por el Corolario 13.9, se tiene que

$$[\mathcal{A}]_{B'} = (M_{B,B'})^t\,[\mathcal{A}]_B\,M_{B,B'}\,.$$

La identidad anterior muestra que los rangos de las matrices $[\mathcal{A}]_{B'}$ y $[\mathcal{A}]_B$ coinciden (ejercicio. Sugerencia: probar que si $A \in K^{n\times n}$ es arbitraria y $B \in K^{n\times n}$ es matriz inversible entonces los rangos de las matrices A, AB y BA coinciden). △

Proposición 13.15. *Sea $\mathcal{V}$ un K-ev, $\dim_K \mathcal{V} = n$ y sea $\mathcal{A} \in Bil(\mathcal{V})$. Son equivalentes:*

1. *$\mathcal{A}$ tiene rango n;*

2. *Para todo $u \in \mathcal{V}$, $u \neq 0$, existe $v \in \mathcal{V}$ tal que $\mathcal{A}(u\,,\,v) \neq 0$;*

3. *Para todo $v \in \mathcal{V}$, $v \neq 0$, existe $u \in \mathcal{V}$ tal que $\mathcal{A}(u\,,\,v) \neq 0$.*

Demostración. 1. $\Longrightarrow$ 2: sea $u \in \mathcal{V}$ tal que $u \neq 0$; en particular, $[u]_B \neq \vec{0}$. Como el rango de $[\mathcal{A}]_B$ coincide con el rango de su transpuesta $[\mathcal{A}]_B^t$ entonces las dimensiones de los núcleos de $[\mathcal{A}]_B$ y $[\mathcal{A}]_B^t$ también coinciden; como $N([\mathcal{A}]_B) = \{0\}$ concluimos que $N([\mathcal{A}]_B^t) = \{0\}$. Así, $[\mathcal{A}]_B^t\,[u]_B \neq \vec{0}$ de forma que

$$\vec{0}^{\,t} \neq ([\mathcal{A}]_B^t\,[u]_B)^t = [u]_B^t\,[\mathcal{A}]_B\,.$$

A partir de este hecho, concluimos que existe $v \in \mathcal{V}$ tal que

$$0 \neq [u]_B^t\,[\mathcal{A}]_B\,[v]_B = \mathcal{A}(u\,,\,v)\,,$$

en donde la verificación de la existencia de $v \in \mathcal{V}$ como antes es ejercicio.

2. $\Longrightarrow$ 1. Supongamos que vale 2. Sea $B = \{v_1, \ldots, v_n\}$ una base de $\mathcal{V}$. Sea $\vec{0} \neq \vec{x} = (x_j)_{j=1}^n \in K^n$; si definimos $u = \sum_{j=1}^n x_j\; v_j \in \mathcal{V}$ entonces $[u]_B = \vec{x}$ de forma que $u \neq 0$. Por el ítem 2, existe $v \in \mathcal{V}$ tal que

$$\mathcal{A}(u\,,\,v) = [u]_B^t\,[\mathcal{A}]_B\,[v]_B \neq 0\,.$$

Lo anterior indica que $[u]_B^t\,[\mathcal{A}]_B \neq \vec{0}^{\,t}$ es decir, que $[\mathcal{A}]_B^t\,[u]_B = [\mathcal{A}]_B^t\,\vec{x} \neq \vec{0}$. Así, dado $\vec{x} \neq \vec{0}$ hemos verificado que $[A]_B^t\,\vec{x} \neq \vec{0}$; este hecho indica que el núcleo $\mathrm{N}([\mathcal{A}]_B^t) = \{\vec{0}\}$ y entonces el rango de $[\mathcal{A}]_B^t$ es n. Entonces, el rango de $[\mathcal{A}]_B$ también es n y vale el ítem 1.

La prueba de la equivalencia entre los ítems 1. y 3. es similar y se deja como ejercicio. $\square$

Definición 13.16. *Sea $\mathcal{V}$ un K-ev, $\dim_K \mathcal{V} = n$ y sea $\mathcal{A} \in Bil(\mathcal{V})$. Decimos que $\mathcal{A}$ es no degenerada si verifica alguna de las condiciones (equivalentes) de la Proposición 13.15.* $\square$

13.1 Formas simétricas, antisimétricas

En lo que sigue vamos a asumir, por simplicidad, que el cuerpo de escalares K verifica la condición: si $1_K \in K$ denota el nuetro del producto (unidad) entonces la suma $\sum_{j=1}^n 1_K \neq 0$, donde en la suma anterior hay n términos, con $n \geq 1$ (aquí n es un número natural); en este caso, convenimos en escribir $1_K + \ldots + 1_K = n_K$, donde ahora $n_K \in K$. Como $n_K \neq 0$ entonces siempre podemos considerar $n_K^{-1} \in K$: más adelante, escribimos simplemente $n_K = n \in K$. Por ejemplo, si $K = \mathbb{R}$ ó $K = \mathbb{C}$ entonces esta propiedad se verifica. Sin embargo, si $K = \mathbb{Z}_p$ (los enteros módulo $p \in \mathbb{N}$, con p primo) entonces $\sum_{j=1}^p \overline{1} = \overline{p} = \overline{0}$ (donde $\overline{a}$ denota la clase de a módulo p) de forma que $\mathbb{Z}_p$ es un cuerpo que no verifica nuestra hipótesis (y queda excluído de nuestro análisis).

Comenzamos con dos definiciones relacionadas con el concepto de simetría y antisimetría (para formas bilineales y matrices cuadradas).

Definición 13.17. *Sea $\mathcal{V}$ un K-ev, $\dim_K \mathcal{V} = n$ y sea $\mathcal{A} \in Bil(\mathcal{V})$. Decimos que*

1. *$\mathcal{A}$ es simétrica si $\mathcal{A}(u\,,\,v) = \mathcal{A}(v\,,\,u)$, para todos u, $v \in \mathcal{V}$.*

2. *$\mathcal{A}$ es antisimétrica si $\mathcal{A}(u\,,\,v) = -\mathcal{A}(v\,,\,u)$, para todos u, $v \in \mathcal{V}$.*

$\square$

Definición 13.18. *Sea K un cuerpo y $A \in K^{n\times n}$. Decimos que A es matriz*

1. *A es simétrica si $A_{kj} = A_{jk}$, para $1 \leq j,\, k \leq n$ (es decir, $A^t = A$).*

2. *A es antisimétrica si $A_{kj} = -A_{jk}$, para $1 \leq j,\, k \leq n$ (es decir $A^t = -A$).*

$\square$

Ejemplo 13.19. Sea $K = \mathbb{R}$: entonces las matrices

$$A = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix} \quad \text{y} \quad B = \begin{pmatrix} 0 & 3 \\ -3 & 0 \end{pmatrix}$$

son tales que $A^t = A$ (A es simétrica) y $B^t = -B$ (B es antisimétrica). $\triangle$

Obs 13.20. Sea K un cuerpo y $A \in K^{n\times n}$.

1. Si A es matriz simétrica entonces para determinar a A basta conocer las entradas de la forma $A_{jk} \in K$ con $1 \leq j \leq k \leq n$ (la parte "triangular superior" de A).
2. El conjunto de matrices $n \times n$ simétricas es un subespacio de $K^{n\times n}$ de dimensión $\frac{1}{2}\, n\,(n+1)$ (ejercicio. Sugerencia: considerar el ítem anterior para obtener una base de este espacio).
3. Si A es matriz antisimétrica entonces $A_{jj} = 0$ para $1 \leq j \leq n$ (pues $A_{jj} = -A_{jj}$); para determinar a A basta conocer las entradas de la forma $A_{jk} \in K$ con $1 \leq j < k \leq n$ (la parte "estrictamente por arriba" de la diagonal principal de A).
4. El conjunto de matrices $n\times n$ antisimétricas es un subespacio de $K^{n\times n}$ de dimensión $\frac{1}{2}\, n\,(n-1)$ (ejercicio. Sugerencia: considerar el ítem anterior para obtener una base de este espacio).

$\triangle$

Proposición 13.21. *Sea $\mathcal{V}$ un K-ev,* $\dim_K \mathcal{V} = n$ *y sea $\mathcal{A} \in Bil(\mathcal{V})$. Si B es una base arbitraria de $\mathcal{V}$ entonces*

1. *$\mathcal{A}$ es simétrica si y solo si $[\mathcal{A}]_B \in K^{n\times n}$ es matriz simétrica.*
2. *$\mathcal{A}$ es antisimétrica si y solo si $[\mathcal{A}]_B \in K^{n\times n}$ es matriz antisimétrica.*

Demostración. Probamos la equivalencia del ítem 1. y la prueba del ítem 2 queda como ejercicio.

Por un lado, es claro que si $\mathcal{A}$ es simétrica entonces $[\mathcal{A}]_B$ resulta ser una matriz simétrica, por la construcción de esta matriz (ver Definición 13.6). Recírocamente, si $[\mathcal{A}]_B$ es matriz simétrica entonces $[\mathcal{A}]_B^t = [\mathcal{A}]_B$. En particular, si u, $v \in \mathcal{V}$ entonces $\mathcal{A}(v\,,\,u) = [v]_B^t\,[\mathcal{A}]_B\,[u]_B$; usando que $K = K^{1\times 1}$ de forma que la transposición de elementos de K (pensados como matrices) satisface $\alpha^t = \alpha$ para $\alpha \in K$:

$$\mathcal{A}(v\,,\,u) = \mathcal{A}(v\,,\,u)^t = ([v]_B^t\,[\mathcal{A}]_B\,[u]_B)^t = [u]_B^t\,[\mathcal{A}]_B^t\,[u]_B = [u]_B^t\,[\mathcal{A}]_B\,[u]_B = \mathcal{A}(u\,,\,v)\,,$$

en donde hemos usado las propiedades de la transposición de matrices (vistas en Álgebra I). Las igualdades anteriores muestran que $\mathcal{A}$ es simétrica en este caso. $\square$

Como consecuencia del Teorema 13.12, la Observación 13.20 y la Proposición 13.21 deducimos el siguiente resultado (cuya prueba es un ejercicio).

Corolario 13.22. *Sea $\mathcal{V}$ un K-ev,* $\dim_K \mathcal{V} = n$. *Entonces*

1. *El conjunto formado por todas las formas bilineales simétricas es un subespacio de $Bil(\mathcal{V})$ de dimensión $\frac{1}{2}\, n\,(n+1)$.*
2. *El conjunto formado por todas las formas bilineales antisimétricas es un subespacio de $Bil(\mathcal{V})$ de dimensión $\frac{1}{2}\, n\,(n-1)$.*

$\square$

Obs 13.23. Los conceptos de simetría y antisimetría en el caso de formas bilineales son, en cierto sentido, contrapuestos. Concretamente, si $\mathcal{V}$ un K-ev, $\dim_K \mathcal{V} = n$ y $\mathcal{A} \in \mathrm{Bil}(\mathcal{V})$ es tal que $\mathcal{A}$ es simétrica y antisimétrica simultáneamente, entonces: si $u, v \in \mathcal{V}$ se tiene que

$$\mathcal{A}(u\,,\,v) = \mathcal{A}(v\,,\,u) = -\mathcal{A}(u\,,\,v) \implies \mathcal{A}(u\,,\,v) = 0$$

en donde hemos usado la simetría en la primer igualdad y la antisimetría en la segunda desigualdad y el hecho de que si $\alpha \in K$ es tal que $\alpha + \alpha = 0$ entonces $\alpha = 0$ (esto es una consecuencia de que $0 = \alpha + \alpha = \alpha(1_K + 1_K) = \alpha \cdot 2_K$ y $2_K \neq 0$: ver los comentarios al comienzo de la Sección 13.1). Así, la forma nula es la que es la única forma bilineal simétrica y antisimétrica simultáneamente. $\triangle$

Proposición 13.24. *Sea $\mathcal{V}$ un K-ev, $\dim_K \mathcal{V} = n$ y sea $\mathcal{A} \in Bil(\mathcal{V})$. Entonces existen únicas $\mathcal{A}_s, \mathcal{A}_a \in Bil(\mathcal{V})$ tales que:*

1. *$\mathcal{A}_s$ es simétrica y $\mathcal{A}_a$ es antisimétrica;*
2. *$\mathcal{A} = \mathcal{A}_s + \mathcal{A}_a$.*

Demostración. Dada $\mathcal{A} \in \mathrm{Bil}(\mathcal{A})$ definimos $\mathcal{A}^t : \mathcal{V} \times \mathcal{V} \to K$ dada por $\mathcal{A}^t(u\,,\,v) = \mathcal{A}(v\,,\,u)$, para $u, v \in \mathcal{V}$. Entonces se verifica que $\mathcal{A}^t \in \mathrm{Bil}(\mathcal{V})$ (ejercicio). Ahora podemos definir $\mathcal{A}_s, \mathcal{A}_a \in \mathrm{Bil}(\mathcal{V})$ dadas por

$$\mathcal{A}_s = \frac{1}{2_K}(\mathcal{A} + \mathcal{A}^t) \quad \text{y} \quad \mathcal{A}_s = \frac{1}{2_K}(\mathcal{A} - \mathcal{A}^t)\,.$$

Notemos que en la definicion de $\mathcal{A}_s$ y $\mathcal{A}_a$ hemos hecho uso de la estructura de K-ev de $\mathrm{Bil}(\mathcal{V})$ (y del hecho de que $\mathcal{A}^t \in \mathrm{Bil}(\mathcal{V})$). Además, la expresión $1/2_K$ corresponde a 2_K^{-1}, que está bien definido por la hipótesis general de esta sección (ver los comentarios al comienzo de la Sección 13.1), y se verifica (aquí notamos $1 = 1_K$ la unidad de K):

$$\frac{1}{2_K}(\alpha + \alpha) = 2_K^{-1} \cdot ((1_K + 1_K) \cdot \alpha) = 2_K^{-1} \cdot (2_K \cdot \alpha) = \alpha\,.$$

Por definición de combinación lineal en $\mathrm{Bil}(\mathcal{V})$ tenemos que, para $u, v \in \mathcal{V}$:

$$\mathcal{A}_s(u\,,\,v) = \frac{1}{2_K}(\mathcal{A}(u\,,\,v) + \mathcal{A}(v\,,\,u)) \quad \text{y} \quad \mathcal{A}_a(u\,,\,v) = \frac{1}{2_K}(\mathcal{A}(u\,,\,v) - \mathcal{A}(v\,,\,u))\,.$$

Por definición, se verifica que

$$\mathcal{A}_s + \mathcal{A}_s = \frac{1}{2_K}(\mathcal{A} + \mathcal{A}^t) + \frac{1}{2_K}(\mathcal{A} - \mathcal{A}^t) = \mathcal{A}\,.$$

Verifiquemos ahora la unicidad de la descomposición; si suponemos que $\mathcal{B}, \mathcal{C} \in \mathrm{Bil}(\mathcal{V})$ son tales que $\mathcal{B}$ es simétrica, $\mathcal{C}$ es antisimétrica y $\mathcal{A} = \mathcal{B} + \mathcal{C}$ entonces

$$\mathcal{A}_s + \mathcal{A}_a = \mathcal{A} = \mathcal{B} + \mathcal{C} \implies \mathcal{A}_s - \mathcal{B} = \mathcal{C} - \mathcal{A}_a\,. \tag{101}$$

En este caso, $\mathcal{A}_s - \mathcal{B} \in \mathrm{Bil}(\mathcal{V})$ es simétrica, mientras que $\mathcal{C} - \mathcal{A}_a \in \mathrm{Bil}(\mathcal{V})$ es antisimétrica (ver el Corolario 13.22). Estas observaciones, junto con la identidad de la Eq. (101) y la Observación 13.23 muestran que $\mathcal{A}_s - \mathcal{B} = 0 = \mathcal{C} - \mathcal{A}_a$, porque la única forma bilineal que es simultáneamente simétrica y anti-simétrica es la forma nula; así, $\mathcal{A}_s = \mathcal{B}$ y $\mathcal{C} = \mathcal{A}_a$. $\square$

13.2 Formas cuadráticas

Definición 13.25. *Sea $\mathcal{V}$ un K-ev, $\dim_K \mathcal{V} = n$. Una forma cuadrática sobre $\mathcal{V}$ es una función $Q : \mathcal{V} \to K$ tal que existe una forma bilineal $\mathcal{A} \in Bil(\mathcal{V})$, que verifica $Q(v) = \mathcal{A}(v\,,\,v)$, para $v \in \mathcal{V}$.* $\square$

Obs 13.26. Sea $\mathcal{V}$ un K-ev, $\dim_K \mathcal{V} = n$ y sea $Q : \mathcal{V} \to K$ una forma cuadrática sobre $\mathcal{V}$. Por hipótesis existe $\mathcal{A} \in \mathrm{Bil}(\mathcal{V})$ tal que $Q(v) = \mathcal{A}(v\,,\,v)$, para $v \in \mathcal{V}$. Sea B una base de $\mathcal{V}$ y sea $[\mathcal{A}]_B = (a_{jk})_{j,k=1}^n \in K^{n\times n}$. Entonces, si $v \in \mathcal{V}$ tiene coordenadas $[v]_B = (\alpha_j)_{j=1}^n \in K^n$,

$$Q(v) = \mathcal{A}(v\,,\,v) = [v]_B^t\,[\mathcal{A}]_B\,[v]_B = \sum_{j=1}^n \sum_{k=1}^n \alpha_j\, a_{jk}\, \alpha_k = \sum_{j=1}^n \sum_{k=1}^n a_{jk}\, \alpha_j\, \alpha_k\,.$$

Notemos que el factor $\alpha_j\,\alpha_k$ $(= \alpha_k\,\alpha_j)$ en la sumatoria anterior aparece dos veces: una acompanañado por a_{jk} y otra acompañado por a_{kj}. Entonces, podemos sacar factor común y re-escribir

$$Q(v) = \sum_{j=1}^n \sum_{k=1}^n a_{jk}\,\alpha_j\,\alpha_k = \sum_{j=1}^n \sum_{k=1}^{j-1} (a_{jk} + a_{kj})\,\alpha_j\,\alpha_k + \sum_{j=1} a_{jj}\ \alpha_j^2 \tag{102}$$

donde hemos restringido la sumatoria en k a los índices $1 \le k \le j-1$, para no repetir términos (dado que la expresión $(a_{jk}+a_{kj})\,\alpha_j\,\alpha_k$ corresponde a dos términos de la sumatoria inicial) y hemos sumado los términos *diagonales* (es decir, cuando $j = k$) en la última sumatoria.

Finalmente, si definimos b_{jk} para $1 \le k \le j \le n$ de forma que $b_{jk} = (a_{jk} + a_{kj})$ si $k < j$ y $b_{jj} = a_{jj}$ entonces la expresión en la Eq. (102) se transforma en

$$Q(v) = \sum_{j=1}^n \sum_{k=1}^{j} b_{jk}\,\alpha_j\,\alpha_k \quad \text{con} \quad [v]_B = (\alpha_j)_{j=1}^n\,, \tag{103}$$

que es la típica presentación de una forma cuadrática. $\triangle$

Ejemplo 13.27. Sea $U \subseteq \mathbb{R}^n$ un conjunto abierto y sea $f : U \to \mathbb{R}$ una función de clase $C^2(U)$ es decir, tal que existen las segundas derivadas $\frac{\partial^2 f}{\partial x_j \partial x_k}$ en U, y son funciones continuas. En este caso se define el Hessiano de f en el punto $u \in \mathbb{R}^n$ como

$$H_{f,u}(v) = \frac{d^2 f}{dt^2}(u + t\,v)|_{t=0} \quad \text{para} \quad v = (v_1, \ldots, v_n)\,.$$

Como $f \in C^2(U)$ entonces se verifica que

$$\frac{df}{dt}(u + t\,v)|_{t=0} = \sum_{j=1}^n \frac{\partial f}{\partial x_j}(u)\ v_j$$

de forma que

$$H_{f,u}(v) = \frac{d^2 f}{dt^2}(u + t\,v)|_{t=0} = \sum_{j=1}^n \sum_{k=1}^n \frac{\partial^2 f}{\partial x_j \partial x_k}(u)\ v_j\ v_k\,.$$

La expresión anterior permite verificar que $H_{f,u} : \mathbb{R}^n \to \mathbb{R}$ es una forma cuadrática. Esta forma cuadrática está relacionada con propiedades de *concavidad* de la gráfica de la función f; es por ello que esta forma cuadrática, llamada forma Hessiana, juega un papel importante al momento de determinar el comportamiento local de los llamados puntos críticos de f.

$\triangle$

Obs 13.28. Sea $\mathcal{V}$ un K-ev, $\dim_K \mathcal{V} = n$ y sea $Q : \mathcal{V} \to K$ una forma cuadrática sobre $\mathcal{V}$. En general existen muchas formas bilineales $\mathcal{A} \in \mathrm{Bil}(\mathcal{V})$ tales que $Q(v) = \mathcal{A}(v\,,\,v)$. Por ejemplo, si $Q : \mathcal{V} \to K$ es la forma cuadrática nula (que corresponde a la forma bilineal nula) entonces, si $\mathcal{A}$ es cualquier forma bilineal antisimétrica se verifica que $\mathcal{A}(v\,,\,v) = -\mathcal{A}(v\,,\,v)$ (por antisimetría) con lo que $\mathcal{A}(v\,,\,v) = 0 = Q(v)$, para $v \in \mathcal{V}$. Así toda forma bilineal antisimétrica representa a la forma cuadrática nula.

$\triangle$

Sin embargo, si nos restringimos a considerar formas bilineales simétricas para representar formas cuadráticas, entonces sí hay unicidad. Para verificar este hecho, primero probamos el siguiente lema: para simplificar la notación escribimos 2 y $\frac{1}{2}$ para denotar $2_K = 1_K + 1_K$ y $\frac{1}{2_K} = 2_K^{-1}$, donde $1_K = 1 \in K$ es la unidad (neutro para el producto) del cuerpo K.

Lema 13.29. *Sea $\mathcal{V}$ un K-ev, $\dim_K \mathcal{V} = n$ y sea $\mathcal{A}$ una forma bilineal* **simétrica** *sobre $\mathcal{V}$. Si $Q : \mathcal{V} \to K$ es la forma cuadrática sobre $\mathcal{V}$ dada por $Q(v) = \mathcal{A}(v\,,\,v)$, para $v \in \mathcal{V}$, entonces*

$$\mathcal{A}(u\,,\,v) = \frac{1}{2}(Q(u+v) - Q(u) - Q(v)) \quad \textit{para} \quad u,\, v \in \mathcal{V}\,.$$

Demostración. Sean $u,\, v \in \mathcal{V}$; entonces

$$\begin{aligned} Q(u+v) &= \mathcal{A}(u+v\,,\,u+v) = \mathcal{A}(u\,,\,u+v) + \mathcal{A}(v\,,\,u+v) \\ &= (\mathcal{A}(u\,,\,u) + \mathcal{A}(u\,,\,v)) + (\mathcal{A}(v\,,\,u) + \mathcal{A}(v\,,\,v)) = Q(u) + 2\,\mathcal{A}(u\,,\,v) + Q(v)\,, \end{aligned}$$

donde hemos usado la relación entre Q y $\mathcal{A}$ y el hecho de que $\mathcal{A}$ es simétrica. Despejando $\mathcal{A}(u\,,\,v)$ de la ecuación anterior, obtenemos la identidad del enunciado. $\square$

Teorema 13.30. *Sea $\mathcal{V}$ un K-ev, $\dim_K \mathcal{V} = n$ y sea $Q : \mathcal{V} \to K$ una forma cuadrática sobre $\mathcal{V}$. Entonces existe una única forma bilineal* **simétrica** *$\mathcal{A}$ sobre $\mathcal{V}$ tal que $Q(v) = \mathcal{A}(v\,,\,v)$, para $v \in \mathcal{V}$.*

Demostración. Por definición, existe una forma bilineal $\mathcal{B} \in \mathrm{Bil}(\mathcal{V})$ tal que $Q(v) = \mathcal{B}(v\,,\,v)$, para $v \in \mathcal{V}$. Definimos

$$\mathcal{A} = \frac{1}{2}(\mathcal{B} + \mathcal{B}^t) \in \mathrm{Bil}(\mathcal{V})\,.$$

(con la notación de la Proposición 13.24, $\mathcal{A} = \mathcal{B}_s$ (ver también la prueba de ese resultado)). Entonces, si $v \in \mathcal{V}$:

$$\mathcal{A}(v\,,\,v) = \frac{1}{2}(\mathcal{B} + \mathcal{B}^t)(v\,,\,v) = \frac{1}{2}(\mathcal{B}(v\,,\,v) + \mathcal{B}^t(v\,,\,v)) = \frac{1}{2}(Q(v) + Q(v)) = Q(v)\,.$$

En particular, por el Lema 13.29, como $\mathcal{A}$ es simétrica tenemos que

$$\mathcal{A}(u\,,\,v) = \frac{1}{2}(Q(u+v) - Q(u) - Q(v)) \quad \text{para} \quad u,\, v \in \mathcal{V}\,.$$

Si $\mathcal{C} \in \mathrm{Bil}(\mathcal{V})$ es simétrica y tal que $\mathcal{C}(v\,,\,v) = Q(v)$, para $v \in \mathcal{V}$ entonces, nuevamente por el Lema 13.29,

$$\mathcal{C}(u\,,\,v) = \frac{1}{2}(Q(u+v) - Q(u) - Q(v)) \quad \text{para} \quad u,\, v \in \mathcal{V}\,.$$

Las dos identidades anteriores muestran que $\mathcal{A} = \mathcal{C}$ lo que establece la unicidad de $\mathcal{A}$ del enunciado.

$\square$

Definición 13.31. *Sea $\mathcal{V}$ un K-ev, $\dim_K \mathcal{V} = n$ y sea $Q : \mathcal{V} \to K$ una forma cuadrática sobre $\mathcal{V}$. Sea $\mathcal{A} \in Bil(\mathcal{V})$ es la única forma bilineal* **simétrica** *tal que $\mathcal{A}(v\,,\,v) = Q(v)$, para $v \in \mathcal{V}$. Dada una base B de $\mathcal{V}$ definimos la matriz de Q con respecto a B, notada $[Q]_B$ como la matriz simétrica $[Q]_B = [\mathcal{A}]_B \in K^{n\times n}$.* $\square$

Obs 13.32. Sea $\mathcal{V}$ un K-ev, $\dim_K \mathcal{V} = n$ y sea $Q : \mathcal{V} \to K$ una forma cuadrática sobre $\mathcal{V}$. Si B es una base de $\mathcal{V}$, en la Observación 13.26 hemos mostrado que existen coeficientes $b_{jk} \in K$ para $1 \leq k \leq j \leq n$ tales que

$$Q(v) = \sum_{j=1}^{n} \sum_{k=1}^{j} b_{jk}\, \alpha_j\, \alpha_k \quad \text{con} \quad [v]_B = (\alpha_j)_{j=1}^n \, .$$

En este caso, si definimos, para $1 \leq j,\ k \leq n$:

$$c_{jk} = \begin{cases} \frac{1}{2}\, b_{jk} & \text{si } \ 1 \leq k \leq j-1 \\ \\ b_{jj} & \text{si } k = j \\ \\ \frac{1}{2}\, b_{kj} & \text{si } \ n \geq k > j \geq 1 \end{cases} \tag{104}$$

entonces $[Q]_B = (c_{jk})_{j,k=1}^n$ (verificar, ejercicio. Sugerencia: notar que la matriz $(c_{jk})_{j,k=1}^n$ es simétrica por construcción; si $\mathcal{A} \in \mathrm{Bil}(\mathcal{V})$ es la única forma bilineal tal que $[\mathcal{A}]_B = (c_{jk})_{j,k=1}^n$ entonces $\mathcal{A}$ es una forma simétrica. Finalmente, probar que $\mathcal{A}$ representa a Q es decir: $Q(v) = \mathcal{A}(v, v)$, para $v \in \mathcal{V}$). $\triangle$

13.3 Formas sobre un espacio euclídeo

En lo que sigue, vamos a llamar espacio euclídeo a un $\mathbb{R}$-EPI de dimensión finita. El arquetipo de estos espacios es $\mathbb{R}^n$ con el producto interno usual (que corresponde a la geometría euclíea).

Proposición 13.33. *Sea $(\mathcal{V}, \langle\ ,\ \rangle)$ un espacio euclídeo, $\dim_{\mathbb{R}} \mathcal{V} = n$. Si $\mathcal{A} \in Bil(\mathcal{V})$ entonces existe un único $T \in L(\mathcal{V})$ tal que*

$$\mathcal{A}(u\,,\, v) = \langle u\,,\, T\, v\rangle \quad \textit{para} \quad u,\, v \in \mathcal{V}\, .$$

Más aún, si B es base ortonomal de $\mathcal{V}$ entonces $[\mathcal{A}]_B = [T]_B \in \mathbb{R}^{n\times n}$.

Demostración. Sea $\mathcal{E} = \{v_1, \ldots, v_n\}$ una base ortonormal (arbitraria pero fija) de $\mathcal{V}$. Consideramos la matriz $[\mathcal{A}]_{\mathcal{E}} = (a_{jk})_{j,k=1}^n \in \mathbb{R}^{n\times n}$; recordemos que entonces $a_{jk} = \mathcal{A}(v_j\,,\, v_k)$ para $1 \leq j, n \leq n$. Sea $T \in L(\mathcal{V})$ el único operador tal que $[T]_{\mathcal{E}} = [\mathcal{A}]_{\mathcal{E}}$. Por construcción de $[T]_{\mathcal{E}}$, sabemos que si $1 \leq k \leq n$,

$$T v_k = \sum_{\ell=1}^{n} a_{\ell k}\ v_\ell \implies \langle v_j\,,\, T\, v_k\rangle = \overline{a_{jk}} = a_{jk} = \mathcal{A}(v_j\,,\, v_k) \in \mathbb{R}\,, \tag{105}$$

pues la k-esima columna de $[T]_{\mathcal{E}}$ es $[T\, v_k]_{\mathcal{E}}$ y $\mathcal{E}$ es base ortonormal de $\mathcal{V}$ (verificar).

Así, si $u,\, v \in \mathcal{V}$ entonces, como $\mathcal{E}$ es base ortonormal, tenemos que

$$u = \sum_{j=1}^{n} \langle u\,,\, v_j\rangle\ v_j \quad , \quad v = \sum_{k=1}^{n} \langle v\,,\, v_k\rangle\ v_k \implies [u]_{\mathcal{E}} = (\langle u\,,\, v_j\rangle)_{j=1}^n \quad , \quad [v]_{\mathcal{E}} = (\langle v\,,\, v_k\rangle)_{k=1}^n \, .$$

Entonces, por las propiedades del producto interno en espacios reales y la linealidad de T,

$$\begin{aligned} \langle u\,,\, T\, v\rangle &= \langle \sum_{j=1}^{n} \langle u\,,\, v_j\rangle\ v_j\,,\, T\,(\sum_{k=1}^{n} \langle v\,,\, v_k\rangle\ v_k)\rangle = \sum_{j=1}^{n} \sum_{k=1}^{n} \langle u\,,\, v_j\rangle\ \langle v\,,\, v_k\rangle\, \langle v_j\,,\, T\, v_k\rangle \\ &= \sum_{j=1}^{n} \sum_{k=1}^{n} \langle u\,,\, v_j\rangle\ \langle v\,,\, v_k\rangle\, \mathcal{A}(v_j\,,\, v_k) = [u]_{\mathcal{E}}^t\, [\mathcal{A}]_{\mathcal{E}}\, [v]_{\mathcal{E}} = \mathcal{A}(u\,,\, v)\,, \end{aligned}$$

donde hemos usado la identidad en la Eq. (105); la anteúltima igualdad es consecuencia de un cálculo directo (que en realidad ya hemos hecho, ver la Observación 13.8). Los hechos anteriores muestran que T representa la forma bilineal $\mathcal{A}$ en el sentido que vale la identidad del enunciado del teorema.

Consideremos una base ortonormal arbitraria B de $\mathcal{V}$. Entonces hemos verificado que la matriz de cambio de base $M_{B,\,E} \in \mathbb{R}^{n\times n}$ es una matriz unitaria (en algunos casos, las matrices unitarias de entradas reales son también llamadas matrices ortogonales). Notemos que como $M_{\mathcal{E},B}$ es unitaria y tiene entradas reales entonces (la conjugación de las entradas de $M_{\mathcal{E},B}$ no cambia las entradas porque son reales!) tenemos las identidades

$$M_{\mathcal{E},B}^{-1} = (M_{\mathcal{E},B})^* = (M_{\mathcal{E},B})^t\,. \tag{106}$$

En este caso, podemos calcular las matrices $[T]_B$ y $[\mathcal{A}]_B$ usando las correspondientes fórmulas (distintas!) para el cambio de base de estas matrices

$$[T]_B = M_{\mathcal{E},B}^{-1}\,[T]_{\mathcal{E}}\,M_{\mathcal{E},B} \quad \text{ y } \quad [\mathcal{A}]_B = (M_{\mathcal{E},B})^t\,[\mathcal{A}]_{\mathcal{E}}\,M_{\mathcal{E},B} = M_{\mathcal{E},B}^{-1}\,[\mathcal{A}]_{\mathcal{E}}\,M_{\mathcal{E},B}$$

en donde hemos usado la identidad de la Eq. (106). Como $[T]_{\mathcal{E}} = [\mathcal{A}]_{\mathcal{E}}$ (por construcción de T) las identidades anteriores nos permiten ver que $[T]_B = [\mathcal{A}]_B$, para la base ortonormal B también.

Finalmente, notemos que la unicidad de T es una consecuencia directa de la identidad $\mathcal{A}(u\,,\,v) = \langle u\,,\,T\,v\rangle$ para $u,\,v \in \mathcal{V}$. En efecto, si $S \in L(\mathcal{V})$ es tal que $\mathcal{A}(u\,,\,v) = \langle u\,,\,S\,v\rangle$ para $u,\,v \in \mathcal{V}$ entonces

$$\langle u\,,\,T\,v\rangle = \mathcal{A}(u\,,\,v) = \langle u\,,\,S\,v\rangle \quad \text{ para } \quad u,\,v \in \mathcal{V}\,.$$

Así, $\langle u\,,\,(S-T)\,v\rangle = 0$ para $u,\,v \in \mathcal{V}$; tomando $u = (S-T)v \in \mathcal{V}$ concluimos que $(S-T)v = 0$ para $v \in \mathcal{V}$. Este último hecho muestra que $S = T$. □

Corolario 13.34. *Sea $(\mathcal{V}, \langle\,,\,\rangle)$ un espacio euclídeo,* $\dim_{\mathbb{R}} \mathcal{V} = n$. *Sea $\mathcal{A} \in Bil(\mathcal{V})$ simétrica y $T \in L(\mathcal{V})$ es el único operador tal que $\mathcal{A}(u\,,\,v) = \langle u\,,\,T\,v\rangle$ para $u,\,v \in \mathcal{V}$. Entonces $T \in L(\mathcal{V})$ es autoadjunto.*

Demostración. Como $\mathcal{A}$ es simétrica entonces, dados $u,\,v \in \mathcal{V}$ tenemos que

$$\langle u\,,\,T\,v\rangle = \mathcal{A}(u\,,\,v) = \mathcal{A}(v\,,\,u) = \langle v\,,\,T\,u\rangle = \langle T\,u\,,\,v\rangle\,,$$

donde hemos usado que el producto interno es a valores reales en la última identidad para poder intercambiar los roles de la primer y segunda coordenada (sin modificar el valor del producto interno). La identidad $\langle T\,u\,,\,v\rangle = \langle u\,,\,T\,v\rangle$ para $u,\,v \in \mathcal{V}$, que se deduce de lo anterior, muestra que T cumple el rol de su adjunto. Por la unicidad del operador adjunto de T ahora vemos que $T^* = T$ y T es autoadjunto. □

Definición 13.35. *Sea $\mathcal{V}$ un K-ev,* $\dim_K \mathcal{V} = n$ *y sea $Q : \mathcal{V} \to K$ una forma cuadrática. Dada una base B de $\mathcal{V}$ decimos que B presenta a Q en forma diagonal si existen $a_1, \ldots, a_n \in K$ tales que*

$$Q(u) = \sum_{j=1}^{n} a_j\,\alpha_j^2 \quad \textit{para} \quad [u]_B = (\alpha_j)_{j=1}^n \in K^n\,.$$

□

Obs 13.36. Con las notaciones de la defınción anterior, es conveniente aclarar que los coeficientes $a_1, \ldots, a_n \in K$ no son únicos (si no que pueden depender de las distintas bases que presentan a Q

en forma diagonal). Por ejemplo, sea $B = \{v_1, \ldots, v_n\}$ una base de $\mathcal{V}$ que presenta a Q en forma diagonal; sean $a_1, \ldots, a_n \in K$ tales que

$$Q(u) = \sum_{j=1}^{n} a_j\, \alpha_j^2 \quad \text{para} \quad [u]_B = (\alpha_j)_{j=1}^n \in K^n .$$

Consideremos $B' = \{2\, v_1, \ldots, 2\, v_n\}$ que también es base de $\mathcal{V}$ (ejercicio). De hecho, es sencillo verificar que si $u \in \mathcal{V}$ es tal que

$$[u]_B = (\alpha_j)_{j=1}^n \implies [u]_{B'} = (\frac{\alpha_j}{2})_{j=1}^n .$$

Entonces, haciendo el viejo truco de multiplicar y dividir por 4:

$$Q(u) = \sum_{j=1}^{n} a_j\, \alpha_j^2 = Q(u) = \sum_{j=1}^{n} 4\, a_j\, (\frac{\alpha_j}{2})^2 = Q(u) = \sum_{j=1}^{n} 4\, a_j\, \beta_j^2 \quad \text{para} \quad [u]_{B'} = (\beta_j)_{j=1}^n \in K^n .$$

Así, B' también presenta a Q en forma diagonal, con coeficientes $4a_1, \ldots, 4a_n \in K$.

$\triangle$

Teorema 13.37 (Existencia de presentaciones diagonales de formas cuadráticas). *Sea $(\mathcal{V}, \langle\, ,\, \rangle)$ un espacio euclídeo, $\dim_{\mathbb{R}} \mathcal{V} = n$. Si $Q : \mathcal{V} \to \mathbb{R}$ es una forma cuadrática, entonces existe una base ortonormal $B = \{v_1, \ldots, v_n\}$ de $\mathcal{V}$ tal que B presenta a Q en forma diagonal es decir, existen $a_1, \ldots, a_n \in \mathbb{R}$ tales que*

$$Q(v) = \sum_{j=1}^{n} a_j\ \langle v\, ,\, v_j \rangle^2 \quad \textit{para} \quad v \in \mathcal{V} \quad (\ [v]_B = (\langle v\, ,\, v_j \rangle)_{j=1}^n\ ) .$$

Demostración. Dada $Q : \mathcal{V} \to \mathbb{R}$ forma cuadrática, entonces el Teorema 13.30 existe una (única) forma bilineal simétrica $\mathcal{A} \in \mathrm{Bil}(\mathcal{V})$ tal que $Q(u) = \mathcal{A}(u\, ,\, u)$ para $u \in \mathcal{V}$.

Por la Proposición 13.33 existe un (único) $T \in L(\mathcal{V})$ tal que $\mathcal{A}(u\, ,\, v) = \langle u\, ,\, T\, v \rangle$, para $u,\, v \in \mathcal{V}$. Más aún, como $\mathcal{A}$ es simétrica, el Corolario 13.34 garantiza que $T = T^*$ es operador autoadjunto.

Así, por resultados vistos en las notas correspondientes al tp8 (segunda parte) T es unitariamente diagonalizable (que corresponde al estudio que hemos hecho en el caso de EPI reales). Entonces, existe una base ortonormal $B = \{v_1, \ldots, v_n\}$ de $\mathcal{V}$ tal que $[T]_B = \mathrm{diag}(a_1, \ldots, a_n) \in \mathbb{R}^{n \times n}$ es una matriz diagonal, con entradas en su diagonal principal dadas por $a_1, \ldots, a_n \in \mathbb{R}$. En este caso, aplicando resultados sobre operadores cuya matriz en una base resulta diagonal, vemos que $T\, v_j = a_j\, v_j$, $1 \le j \le n$.

Sea $v \in \mathcal{V}$; entonces, como B es base ortonormal de $\mathcal{V}$ tenemos que

$$v = \sum_{j=1}^{n} \langle v\, ,\, v_j \rangle\ v_j \implies T\, v = T(\sum_{j=1}^{n} \langle v\, ,\, v_j \rangle\ v_j) = \sum_{j=1}^{n} \langle v\, ,\, v_j \rangle\ T\, v_j = \sum_{j=1}^{n} \langle v\, ,\, v_j \rangle\ a_j\ v_j$$

donde hemos usado la linealidad de T y las observaciones previas ($T\, v_j = a_j\, v_j$, $1 \le j \le n$). Finalmente, observemos que

$$\begin{aligned} Q(v) &= \mathcal{A}(v\, ,\, v) = \langle v\, ,\, T\, v \rangle = \langle \sum_{j=1}^{n} \langle v\, ,\, v_j \rangle\ v_j\, ,\, \sum_{k=1}^{n} \langle v\, ,\, v_k \rangle\ a_k\ v_k \rangle \\ &= \sum_{j=1}^{n} \langle v\, ,\, v_j \rangle\ \sum_{k=1}^{n} a_k\ \langle v\, ,\, v_k \rangle\ \underbrace{\langle v_j\, ,\, v_k \rangle}_{\delta_{jk}} = \sum_{j=1}^{n} \langle v\, ,\, v_j \rangle\ a_j\ \langle v\, ,\, v_j \rangle = \sum_{j=1}^{n}\ a_j\ \langle v\, ,\, v_j \rangle^2 . \end{aligned}$$

$\square$

Obs 13.38. Con las notaciones del teorema anterior y de su prueba, notemos que la base B así como los coeficientes $a_1, \ldots, a_n \in \mathbb{R}$ surgen de encontrar una base ortonormal de $\mathcal{V}$ que diagonalice al operador autoadjunto T (en este caso, $a_1, \ldots, a_n \in \mathbb{R}$ son los autovalores de T, contando multiplicidades); el proceso de hallar tal base B (asumiendo que hemos podido hallar todas las raíces del polinomio característico de T) ha sido descripto en notas anteriores (mediante la ortonornalización de las bases de los autoespacios de T). $\triangle$

Ejemplo 13.39 (Forma canónica de una forma cuadrática)**.** Consideremos $\mathcal{V} = \mathbb{R}^3$ como $\mathbb{R}$-ev con el producto interno usual (que es un espacio euclídeo), y sea $Q : \mathbb{R}^3 \to \mathbb{R}$ la forma cuadrática presentada como en la Eq. (103) de la Observación 13.26, con respecto a la base canónica $B_c = \{e_1,\, e_2,\, e_3\}$:

$$Q((x_1, x_2, x_3)) = \sum_{j=1}^{3} \sum_{k=1}^{j} b_{jk}\, x_j\, x_k$$

donde, en este caso concretamente Q está dada por:

$$Q((x_1, x_2, x_3)) = \underbrace{x_1^2}_{j=1} + \underbrace{4\, x_2\, x_1 + 4\, x_2^2}_{j=2,\ 1\le k\le 2} + \underbrace{4\, x_3\, x_1 + 8\, x_3\, x_2 + 4\, x_3^2}_{j=3,\ 1\le k\le 3} .$$

Así,

$$b_{11} = 1\,,\ b_{21} = 4\,,\ b_{22} = 4\,,\ b_{31} = 4\,,\ b_{32} = 8\,,\ b_{33} = 4\,.$$

Podemos aplicar la fórmula de la Eq. (104) de la Observación 13.32 para calcular la matriz de Q con respecto a la base canónica, es decir $[Q]_{B_c}$: de esta forma, si $[Q]_{B_c} = (c_{jk})_{j,k=1}^3 \in \mathbb{R}^{3\times 3}$ entonces

$$[Q]_{B_c} = \begin{pmatrix} 1 & 2 & 2 \\ 2 & 4 & 4 \\ 2 & 4 & 4 \end{pmatrix} .$$

Notemos que $[Q]_{B_c}$ resulta una matriz simétrica, pues por definición $[Q]_{B_c} = [\mathcal{A}]_{B_c}$, donde $\mathcal{A} \in \mathrm{Bil}(\mathbb{R}^3)$ es la única forma bilineal simétrica que representa a Q (de forma que $[\mathcal{A}]_{B_c}$) es una matriz simétrica).

Más aún, como B_c es una base ortonormal de $\mathbb{R}^3$ con respecto al producto interno usual de $\mathbb{R}^3$ entonces, la Proposición 13.33 garantiza que

$$[T]_{B_c} = [\mathcal{A}]_{B_c}$$

donde $T \in L(\mathbb{R}^3)$ es el único operador tal que

$$\mathcal{A}(\vec{x},\, \vec{y}) = \langle \vec{x},\, T\,\vec{y}\rangle \quad \text{para} \quad \vec{x},\, \vec{y} \in \mathbb{R}^3 .$$

Por el Corolario 13.34 sabemos que T es autoadjunto (también lo podemos deducir del hecho de que la matriz $[T]_{B_c}$ es una matriz simétrica real, y luego es una matriz autoadjunta (real), junto con el hecho de que B_c es una base ortonormal de $\mathbb{R}^3$).

Si ahora aplicamos los resultados de la teoría del tp8 (segunda parte), sabemos que existe una base ortonormal $B = \{v_1,\, v_2,\, v_3\}$ de $\mathbb{R}^3$ de autovectores de T: de hecho, para calcular la base B, primero hallamos bases (arbitrarias) de los autoespacios de T: en este caso, si usamos la descripción de la matriz $[T]_{B_c}$ anterior, concluimos que

$$p_T(x) = \det(x\, I - [T]_{B_c}) = x^3 - 9x^2 = x^2(x-9)$$

Así, los autovalores de T son $\lambda_1 = 9$ con multiplicidad algebraica 1 y $\lambda_2 = 0$ con multiplicidad algebraica 2.

Más aún, podemos hallar bases (arbitrarias) de los autoespacios:

$$E_T(\lambda_1) = E_T(9) = N(A - 9\,I) = \overline{\{u_1 = (1,2,2)\}}$$
$$E_T(\lambda_2) = E_T(0) = N(A - 0\,I) = N(A) = \overline{\{u_2 = (-2,1,0)\,,\ u_3 = (-2,0,1)\}}$$

Luego, aplicamos el algoritmo de ortogonalización de Gram-Schmidt a cada una de las bases de cada uno de los autoespacios:

$$B_1 = \{u_1\} \quad \text{y} \quad B_2 = \{u_2,\, u_3\} \quad \text{entonces obtenemos} \quad B_1' = \{w_1\} \quad \text{y} \quad B_2' = \{w_2,\, w_3\}$$

donde $w_1 = u_1$ (por GS a B_1) y $w_2 = u_2$,

$$w_3 = u_3 - \frac{\langle u_3\,,\, w_2\rangle}{\|w_2\|^2} w_2 = (-2,0,1) - \frac{4}{\sqrt{5}}\,(-2,1,0) = (\frac{8}{\sqrt{5}} - 2, \frac{-4}{\sqrt{5}}, 1)\,,$$

por GS a B_2. Finalmente, normalizando los vectores de B_1' obtenemos

$$v_1 = \frac{1}{\|w_1\|}\, w_1 = \frac{1}{3}\,(1,2,2)$$

y normalizando los vectores de $B_2' = \{w_2, w_3\}$ obtenemos

$$v_2 = \frac{1}{\|w_2\|}\, w_2 = \frac{1}{\sqrt{5}}(-2,1,0) \quad \text{y} \quad v_3 = \frac{1}{\|w_3\|}\, w_3 = 0.39\,(\frac{8}{\sqrt{5}} - 2, \frac{-4}{\sqrt{5}}, 1)$$

donde hemos aproximado la norma $\|w_3\|^{-1} \approx 0.39$. Así, $B_2'' = \{v_2, v_3\}$ es una base ortonormal de $E_T(0)$. Entonces,

$$B = \{v_1,\, v_2,\, v_3\}$$

es una base ortonormal de $\mathbb{R}^3$ tal que $[T]_B = \operatorname{diag}(9,0,0)$ (verificar esta afirmación). Finalmente, dado $v \in \mathbb{R}^3$ tal que

$$Q(v) = 9\,\langle v\,,\, v_1\rangle^2 + 0\,\langle v\,,\, v_2\rangle^2 + 0\,\langle v\,,\, v_3\rangle^2 = 9\,\langle v\,,\, v_1\rangle^2\,.$$

Verificar estos hechos (notar que la última identidad vale en general, ver la prueba del Teorema 13.37). △

13.4 Ley de inercia para formas cuadráticas

Hemos visto que dada una forma cuadrática en un espacio euclídeo $\mathcal{V}$, entonces es posible hallar una base (incluso una base ortonormal) B de $\mathcal{V}$, tal que B presenta a Q en forma diagonal, en términos de ciertos coeficientes reales. También hemos visto que estos coeficientes reales (los llamados $a_1, \ldots, a_n$ en la Definición 13.35) no son únicos. Sin embargo, como veremos a continuación, hay unicidad en términos de la cantidad de coeficientes positivos y negativos aparecen en las representaciones diagonales.

Teorema 13.40. *Sea $(\mathcal{V}, \langle\,,\,\rangle)$ un espacio euclídeo, $\dim_{\mathbb{R}} \mathcal{V} = n$ y sea $Q : \mathcal{V} \to \mathbb{R}$ una forma cuadrática. Sean $B = \{v_1, \ldots, v_n\}$ y $B' = \{w_1, \ldots, w_n\}$ bases de $\mathcal{V}$ que presentan a Q en forma diagonal; en este caso, sean $a_1, \ldots, a_n \in \mathbb{R}$ tales que, dado $u \in \mathcal{V}$ arbitrario*

$$Q(u) = \sum_{j=1}^{n} a_j\ \alpha_j^2 \quad \textit{para} \quad [u]_B = (\alpha_j)_{j=1}^n \in K^n\,,$$

y sean $b_1, \ldots, b_n \in \mathbb{R}$ tales que, dado $u \in \mathcal{V}$ arbitrario

$$Q(u) = \sum_{j=1}^{n} b_j\ \beta_j^2 \quad \textit{para} \quad [u]_{B'} = (\beta_j)_{j=1}^n \in K^n\,.$$

Entonces

$$\#(\{j:\ a_j > 0\}) = \#(\{j:\ b_j > 0\}) \quad \textit{y} \quad \#(\{j:\ a_j < 0\}) = \#(\{j:\ b_j < 0\})\,.$$

Demostración. Antes que nada, aclaremos que la notación $\#(A)$ hace referencia a la cantidad de elementos del conjunto A.

Consideremos el conjunto de vectores $P_B = \{v_j : \ a_j > 0\}$ y $N_B = \{v_j : \ a_j < 0\}$; más aún, llamemos
$$k = \#(P_B) = \#(\{j : \ a_j > 0\}) \quad \text{y} \quad m = \#(N_B) = \#(\{j : \ a_j < 0\})\,.$$
Podemos reordenar los elementos de la base B de forma que los elementos de P_B aparezcan listados primero y luego aparezcan listados los elementos de N_B. Finalmente, el grupo restante de vectores (es decir los vectores $\{v_j : \ a_j = 0\}$) los listamos al final. Llamando a este re-ordenamiento nuevamente $B = \{v_1, \ldots, v_n\}$, en este caso tenemos que

$$Q(u) = \sum_{j=1}^{k} c_j \ \alpha_j^2 + \sum_{j=k+1}^{k+m} -c_j \cdot \alpha_j^2 \quad \text{para} \quad [u]_B = (\alpha_j)_{j=1}^n \in K^n\,,$$

donde todos los $c_1, \ldots, c_{m+k} > 0$, $\{c_1, \ldots, c_k\} = \{a_j : \ a_j > 0\}$ y $\{-c_{k+1}, \ldots, -c_{k+m}\} = \{a_j : \ a_j < 0\}$ (aquí hemos usado que las coordenadas con respecto a este reordenamiento -de la base inicial- se obtienen aplicando el mismo reordenamiento a las coordenadas con respecto a la base inicial: por eso los coeficientes positivos aparecen primero, luego los negativos y finalmente los nulos). En la expresión para $Q(u)$ hemos evitado escribir los términos correspondientes a $a_j = 0$.

De forma análoga, re-ordenando los elementos de la base B' (como antes, al reordenamiento de B' lo seguimos llamando B'), podemos asumir que

$$Q(u) = \sum_{j=1}^{p} d_j \, \beta_j^2 + \sum_{j=p+1}^{p+r} -d_j \cdot \beta_j^2 \quad \text{para} \quad [u]_{B'} = (\beta_j)_{j=1}^n \in K^n\,,$$

donde $p = \#(\{j : \ b_j > 0\})$, $r = \#(\{j : \ b_j < 0\})$, todos los $d_1, \ldots, d_{p+r} > 0$ y tales que $\{d_1, \ldots, d_p\} = \{b_j : \ b_j > 0\}$ y $\{-d_{p+1}, \ldots, -d_{+r}\} = \{b_j : \ b_j < 0\}$. En la expresión para $Q(u)$ hemos evitado escribir los términos correspondientes a $b_j = 0$.

En resumen, en los párrafos anteriores hemos re-ordenado las bases para poder presentar a Q en forma diagonal con respecto a (los re-ordenamientos de) B y B', de forma tal que los coeficientes positivos en la representación diagonal aparecen primero en cada lista, y los coeficientes negativos en la representación diagonal aparecen en un segundo grupo en cada lista. Es importante notar que este procedimiento (la re-ordenación) no cambia la cantidad de coeficientes positivos y negativos que teníamos inicialmente, para cada una de las presentaciones diagonales de Q (porque justamente, queremos entender como son esas cantidades, de forma que no debemos modificarlas!!).

Ahora que hemos organizado de forma conveniente las presentaciones de Q, argumentamos como sigue: supongamos que las cantidades k y p de coeficientes positivos en cada representación diagonal no coinciden. En este caso se tiene, por ejemplo, que $k > p$ (el caso $k < p$ se trata de forma completamente análoga).

Definimos los subespacios $\mathcal{V}_1 = \overline{\{v_1, \ldots, v_k\}}$ y $\mathcal{V}_1 = \overline{\{w_{p+1}, \ldots, w_n\}}$. Entonces $\dim_{\mathbb{R}} \mathcal{V}_1 = k$, $\dim_{\mathbb{R}} \mathcal{V} = n - p$ (porqué?). Dos hechos fundamentales son:

1. si $u \in \mathcal{V}_1$, $u \neq 0$ entonces $Q(u) > 0$: en efecto, si $u \in \mathcal{V}_1$, $u \neq 0$ entonces $u = \sum_{j=1}^k \alpha_j \, v_j$ con algún $\alpha_j \neq 0$ para $1 \leq j \leq k$. Entonces

$$[u]_B = (\alpha_1, \ldots, \alpha_k, 0, \ldots, 0) \implies Q(u) = \sum_{j=1}^{k} c_j \, \alpha_j^2 > 0$$

pues: como $c_j > 0$ y $\alpha_j^2 \geq 0$ entonces $c_j \, \alpha_j^2 \geq 0$, para $1 \leq j \leq k$; y existe algún $\alpha_j \neq 0$ lo que implica que $c_j \, \alpha_j^2 > 0$ para algún $1 \leq j \leq k$.

2. si $u \in \mathcal{V}_2$ entonces $Q(u) \leq 0$: en efecto, si $u \in \mathcal{V}_2$, $u \neq 0$ entonces $u = \sum_{j=p+1}^{n} \beta_j \, w_j$. Entonces

$$[u]_{B'} = (0, \ldots, 0, \beta_{p+1}, \ldots, \beta_n) \implies Q(u) = -\sum_{j=p+1}^{p+r} d_j \, \beta_j^2 \leq 0$$

pues: como $d_j > 0$ y $\beta_j^2 \geq 0$ entonces $d_j \, \beta_j^2 \geq 0$, para $p + 1 \leq j \leq r + r$.

Pero como hemos supuesto que las cantidades (números naturales) $k > p$ entonces se tiene que

$$\dim_{\mathbb{R}} \mathcal{V}_1 + \dim_{\mathbb{R}} \mathcal{V}_2 = k + n - p = n + (k - p) > n \, ,$$

pues $k - p \geq 1$.

Si aplicamos un resultado visto en las primeras notas:

$$n \geq \dim_{\mathbb{R}}(\mathcal{V}_1 + \mathcal{V}_2) = \dim_{\mathbb{R}}(\mathcal{V}_1) + \dim_{\mathbb{R}}(\mathcal{V}_1) - \dim_{\mathbb{R}}(\mathcal{V}_1 \cap \mathcal{V}_2) > n - \dim_{\mathbb{R}}(\mathcal{V}_1 \cap \mathcal{V}_2)$$

concluimos que $\dim_{\mathbb{R}}(\mathcal{V}_1 \cap \mathcal{V}_2) \geq 1$; de otra forma $\dim_{\mathbb{R}}(\mathcal{V}_1 \cap \mathcal{V}_2) = 0$ y la desigualdad se transforma en $n > n$, que es una contradicción!. En particular, $\mathcal{V}_1 \cap \mathcal{V}_2 \neq \{0\}$ no se reduce al subespacio nulo.

Sea $u \in \mathcal{V}_1 \cap \mathcal{V}_2 \neq \{0\}$ tal que $u \neq 0$. Entonces, en tanto $u \in \mathcal{V}_1$ podemos aplicar el ítem 1 de las observaciones previas y concluir que $Q(u) > 0$. Por otro lado, en tanto $u \in \mathcal{V}_2$ podemos aplicar el ítem 2 de las observaciones previas y concluir que $Q(u) \leq 0$, que contradice la desigualdad $Q(u) > 0$ anterior. Esta contradicción surge de suponer que $k > p$. En consecuencia, $k \leq p$. Si suponemos que $k < p$ entonces, cambiando los roles de las bases B y B', llegamos también a una contradicción (en este caso, la contradicción depende de construir subespacios $\mathcal{W}_1$ y $\mathcal{W}_2$ adecuados, ejercicio). Así, concluimos que $k = p$.

Podemos aplicar un argumento similar al visto previamente para concluir que $m = r$ (como antes, si $m \neq r$ entonces se pueden construir subespacios adecuados $\mathcal{W}_1$ y $\mathcal{W}_2$ de $\mathcal{V}$, cuya existencia lleva a una contradicción). Los detalles quedan a cargo del lector.

Finalmente, completando los detalles dejados al lector, concluimos que $k = p$ y $m = r$, como queríamos. □

Definición 13.41. *Sea* $(\mathcal{V}, \langle\, , \,\rangle)$ *un espacio euclídeo,* $\dim_{\mathbb{R}} \mathcal{V} = n$ *y sea* $Q : \mathcal{V} \to \mathbb{R}$ *una forma cuadrática. Sea* B *una base de* $\mathcal{V}$ *que presenta a* Q *en forma diagonal; en este caso, sean* $a_1, \ldots, a_n \in \mathbb{R}$ *tales que, dado* $u \in \mathcal{V}$ *arbitrario*

$$Q(u) = \sum_{j=1}^{n} a_j \; \alpha_j^2 \quad \text{para} \quad [u]_B = (\alpha_j)_{j=1}^{n} \in K^n \, ,$$

Definimos:

1. *El índice de inercia positivo de* Q *como* $\#(\{j : \; a_j > 0\})$*;*
2. *El índice de inercia negativo de* Q *como* $\#(\{j : \; a_j < 0\})$*;*
3. *El índice de inercia de* Q *como* $\#(\{j : \; a_j \neq 0\})$*.*

□

Notemos que el Teorema 13.40 garantiza que los índices de inercia están bien definidos en el sentido que no dependen de la base B elegida.

13.5 Formas semidefinidas

Comenzamos con los siguientes conceptos:

Definición 13.42. *Sea $(\mathcal{V}, \langle\,,\,\rangle)$ un espacio euclídeo,* $\dim_{\mathbb{R}} \mathcal{V} = n$ *y sea* $Q : \mathcal{V} \to \mathbb{R}$ *una forma cuadrática. Decimos que Q es*

1. *semidefinida (resp. definida) positiva si $Q(v) \geq 0$ (resp. $Q(v) > 0$) si $v \neq 0$;*
2. *semidefinida (resp. definida) negativa si $Q(v) \leq 0$ (resp. $Q(v) < 0$) si $v \neq 0$;*
3. *indefinida si existen v, $w \in \mathcal{V}$ tales que $Q(v) \cdot Q(w) < 0$.*

□

Proposición 13.43. *Sea $(\mathcal{V}, \langle\,,\,\rangle)$ un espacio euclídeo,* $\dim_{\mathbb{R}} \mathcal{V} = n$. *Sea $Q : \mathcal{V} \to \mathbb{R}$ una forma cuadrática, sea $\mathcal{A} \in Bil(\mathcal{V})$ la única forma bilineal simétrica que representa a Q y sea $T \in L(\mathcal{V})$ el único operador autoadjunto que representa a $\mathcal{A}$ (ver Proposición 13.33 y el Corolario 13.34). Las siguientes condiciones son equivalentes:*

1. *Q es semidefinida positiva;*
2. *El índice de inercia negativo de Q es 0;*
3. *$\sigma(T) \subset [0, \infty)$.*

Demostración. La prueba de la equivalencia $1 \Leftrightarrow 2$ es una consecuencia de las Definiciones 13.41 y 13.42 (ejercicio).

Sea $T \in L(\mathcal{V})$ como en el enunciado. Utilizando el mismo argumento de la prueba del Teorema 13.37, conluimos que existe una base ortonormal $B = \{v_1, \ldots, v_n\}$ de $\mathcal{V}$ y escalares $a_j \in \mathbb{R}$ tales que $T\,v_j = a_j\,v_j$ para $1 \leq j \leq n$ tales que

$$Q(u) = \sum_{j=1}^{n} a_j\, \langle u\,,\, v_j\rangle^2 \quad \text{para} \quad u \in \mathcal{V}\,.$$

Notemos que por definición $[T]_B = \operatorname{diag}(a_1, \ldots, a_n)$ de forma que los autovalores de T son (posiblemente contando repeticiones) $a_1, \ldots, a_n$. Lo anterior indica que $p_T(x) = \prod_{j=1}^{n}(x - a_j)$ de forma que $\sigma(T) = \{a_1, \ldots, a_n\}$ (en donde hemos incluido posibles repeticiones de los autovalores).

Finalmente, notemos que Q es semidefinida positiva si y solo si $a_j \geq 0$ para $1 \leq j \leq n$, dado que $\langle u\,,\, v_j\rangle^2 \geq 0$ para todo $u \in \mathcal{V}$ y todo $1 \leq j \leq n$ (conviene notar aquí que $Q(v_j) = a_j$ para $1 \leq j \leq n$): ejercicio. La equivalencia entre los ítems 2 y 3 es consecuencia del los hechos anteriores junto con la observación de que $\sigma(T) = \{a_1, \ldots, a_n\}$. □

Para indicar el interés de la noción de forma semidefinida y para mostrar un posible uso de la proposición anterior, consideramos el siguiente ejemplo (que ustedes ya han visto en el contexto de extremos locales en Análisis 2).

Ejemplo 13.44. Consideramos nuevamente el Ejemplo 13.27. Así, consideramos $U \subseteq \mathbb{R}^n$ un conjunto abierto y sea $f : U \to \mathbb{R}$ una función de clase $C^2(U)$. En este el Hessiano de f en el punto $u \in \mathbb{R}^n$ está dado por

$$H_{f,u}(v) = \frac{d^2 f}{dt^2}(u + t\,v)|_{t=0} \quad \text{para} \quad v = (v_1, \ldots, v_n)\,,$$

de forma que

$$H_{f,u}(v) = \frac{d^2 f}{dt^2}(u + t\,v)|_{t=0} = \sum_{j=1}^{n}\sum_{k=1}^{n} \frac{\partial^2 f}{\partial x_j \partial x_k}(u)\; v_j\; v_k\,. \tag{107}$$

La expresión anterior permite verificar que $H_{f,u} : \mathbb{R}^n \to \mathbb{R}$ es una forma cuadrática. En este presentación de Q estamos usando las coordenadas de $v \in \mathbb{R}^n$ con respecto a la base canónica B_c: recordemos que se tiene que $[v]_{B_c} = v \in \mathbb{R}^n$.

Es un hecho conocido de Análisis 2 que en nuestro caso se verifica:

$$\frac{\partial^2 f}{\partial x_j \partial x_k}(u) = \frac{\partial^2 f}{\partial x_k \partial x_j}(u) \quad \text{para} \quad 1 \leq j, n \leq n\,. \tag{108}$$

En particular, si definimos la matriz Hessiana

$$\nabla^2 f(u) = (\frac{\partial^2 f}{\partial x_j \partial x_k}(u))_{j,k=1}^n \in \mathbb{R}^{n\times n}$$

entonces $\nabla^2 f(u)$ resulta una matriz simétrica real (por la Eq. (108)); así $\nabla^2 f(u) \in \mathbb{R}^{n\times n}$ es una matriz real autoadjunta.

Más aún, si consideramos la forma bilineal en $\mathbb{R}^n$ dada por

$$(v, w) \mapsto \langle v\,,\, \nabla^2 f(u) \cdot w\rangle$$

entonces esta forma simétrica (pues la matriz de esta forma bilineal con respecto a la base canónica B_c es $\nabla^2 f(u)$ que es una matriz simétrica) y, usando la Eq. (107), vemos que representa a Q. Si consideramos a $\mathbb{R}^n$ como espacio euclídeo, con el producto interno usual, entonces el (único) operador autoadjunto $T_{f,u} \in L(\mathbb{R}^n)$ que representa a esta forma bilineal simétrica es $T_{f,u}(v) = \nabla^2 f(u) \cdot v$, para $v \in \mathbb{R}^n$. De hecho, notemos que B_c es base ortonormal que verifica que $[T_{f,u}]_{B_c} = \nabla^2 f(u)$ (ver Proposición 13.33).

En Análisis 2 han probado que si $u \in U$ es un punto crítico (es decir $\nabla f(u) = 0$) entonces u es un mínimo local si y solo si la forma Hessiana es semidefinida positiva. La Proposición 13.43 permite determinar si la forma es semidefinida, mediante el estudio de los signos de los autovalores de la matriz Hessiana $\nabla^2 f(u)$ (que coinciden con los autovalores del operador $T_{f,u}$). Esto permite formalizar el estudio realizado en Análisis 2, mediante técnicas del Álgebra Lineal. $\triangle$

14 Transformaciones multilineales y producto tensorial

En lo que sigue, dados $\mathcal{V}_1, \ldots, \mathcal{V}_p$ K-ev's vamos a considerar el producto cartesiano

$$\mathcal{V}_1 \times \ldots \times \mathcal{V}_p$$

como conjuntos: en este caso, los elementos de $\mathcal{V}_1 \times \ldots \times \mathcal{V}_p$ son p-uplas de la forma $(v_1, \ldots, v_p)$, tal que $v_j \in \mathcal{V}_j$ para $1 \leq j \leq p$.

Definición 14.1. *Sean $\mathcal{V}_1, \ldots, V_p, \mathcal{U}$ K-ev's. Una función $\varphi : \mathcal{V}_1 \times \ldots \times \mathcal{V}_p \to \mathcal{U}$ (definida en el producto cartesiano de los espacios) es una transformación p-lineal (ó multilineal) si para todo $1 \leq i \leq p$ se verifica: para cualesquiera $v_j \in \mathcal{V}_j$ para $1 \leq j \neq i \leq p$ y todos u, $v \in \mathcal{V}_i$ y $\alpha \in K$*

$$\varphi(v_1, \ldots, v_{i-1}, \alpha\, u + v, v_{i+1}, \ldots, v_p) = \alpha\, \varphi(v_1, \ldots, v_{i-1}, u, v_{i+1}, \ldots, v_p) + \varphi(v_1, \ldots, v_{i-1}, v, v_{i+1}, \ldots, v_p)$$

$\square$

Obs 14.2. Con las notaciones de la definición anterior, una función $\varphi : \mathcal{V}_1 \times \ldots \times \mathcal{V}_p \to U$ es multilineal si para cada $1 \leq i \leq n$ se verifica: para cualesquiera $v_j \in \mathcal{V}_j$ la función

$$\varphi(v_1, \ldots, v_{i-1}, \cdot, v_{i+1}, \ldots, v_p) : \mathcal{V}_i \to \mathcal{U}$$

tal que a $v \in \mathcal{V}_i$ le asigna el valor

$$\varphi(v_1, \ldots, v_{i-1}, v, v_{i+1}, \ldots, v_p)$$

es transformación lineal: notemos que la condición escrita en la Eq. que aparece en la Definición 14.1 corresponde a la linealidad de la transformación $\varphi(v_1, \ldots, v_{i-1}, \cdot, v_{i+1}, \ldots, v_p)$.

En particular, si φ es multilineal entonces: si $u_1, \ldots, u_m \in \mathcal{V}_i$ y $\alpha_1, \ldots, \alpha_m \in K$ entonces

$$\varphi(v_1, \ldots, v_{i-1}, \sum_{j=1}^{m} \alpha_j\, u_j, v_{i+1}, \ldots, v_p) = \sum_{j=1}^{m} \alpha_j\, \varphi(v_1, \ldots, v_{i-1}, u_j, v_{i+1}, \ldots, v_p)\,.$$

$\triangle$

Ejemplos 14.3. Sea K un cuerpo.

1. Si $\mathcal{V} = K^m$, $\mathcal{W} = K^n$ y $U = K^{m\times n}$ como K-ev, entonces la función $\varphi : \mathcal{V} \times \mathcal{W} \to U$ dada por

   $$\varphi(\vec{x}\,,\,\vec{y}) = \vec{x} \cdot \vec{y}^{\,t} \quad \text{para} \quad \vec{x} \in \mathcal{V} \quad \text{y} \quad \vec{y} \in \mathcal{W}$$

   es una transformación 2-lineal (ó transformación bilineales): verificar, ejercicio.

2. Si $\mathcal{V} = K^{n\times n}$ entonces la $\varphi : \mathcal{V} \times \mathcal{V} \times \mathcal{V} \to \mathcal{V}$ dada por

   $$\varphi(A, B, C) = A \cdot B \cdot C \quad \text{para} \quad A,\, B,\, C \in \mathcal{V}$$

   es transformación 3-lineal : verificar, ejercicio..

3. Sea $\mathcal{V}$ un K-ev y sea $E : \mathcal{V} \times L(\mathcal{V}) \to \mathcal{V}$ dada por $E(v, T) = T(v)$ para $v \in \mathcal{V}$ y $T \in L(\mathcal{V})$ es una transformación 2-lineal : verificar, ejercicio.

$\triangle$

Definición 14.4. *Sean $\mathcal{V}_1, \ldots, V_p, \mathcal{U}$ K-ev's. El conjunto de todas las transformaciones p-lineales es notado Hom$(\mathcal{V}_1, \ldots, V_p\,; \mathcal{U})$.* $\square$

Proposición 14.5. *Sean $\mathcal{V}_1, \ldots, V_p, \mathcal{U}$ K-ev's. Dadas $\varphi,\, \psi \in$ Hom$(\mathcal{V}_1, \ldots, V_p\,; \mathcal{U})$ y $\alpha \in K$ entonces definimos, si $\alpha\,\varphi + \psi : \mathcal{V}_1 \times \ldots \times V_p \to \mathcal{U}$ dada por*

$$\alpha\,\varphi + \psi(v_1, \ldots, v_p) = \alpha\,\varphi(v_1, \ldots, v_p) + \psi(v_1, \ldots, v_p) \in \mathcal{U}$$

se verifica que $\alpha\,\varphi + \psi \in$ Hom$(\mathcal{V}_1, \ldots, V_p\,; \mathcal{U})$. Más aún, Hom$(\mathcal{V}_1, \ldots, V_p\,; \mathcal{U})$ dotado de la suma y la acción escalar definidas previamente resulta un K-ev. En particular el vector nulo de este espacio es la transformación p-lineal $0 : \mathcal{V}_1 \times \ldots \times V_p \to \mathcal{U}$ dada por $0(v_1, \ldots, v_p) = 0_{\mathcal{U}}$, para todo $(v_1, \ldots, v_p)$.

Demostración. Ejercicio. $\square$

Ejemplo 14.6. Sean $\mathcal{V}$, $\mathcal{W}$ K-ev's. Hemos considerado previamente al menos dos familias de espacios vectoriales de la siguiente forma:

1. Hom$(\mathcal{V}\,; \mathcal{W}) = L(\mathcal{V}, \mathcal{W})$: es decir, las transformaciones 1-lineales de $\mathcal{V}$ en $\mathcal{W}$ son, simplemente, las transformaciones lineales de $\mathcal{V}$ en $\mathcal{W}$. En este caso, también coinciden la estructura de espacio vectorial sobre K.

2. Si consideramos a K como K-ev, entonces Hom$(\mathcal{V}, \mathcal{W}\,; K)$ =Bil$(\mathcal{V}, \mathcal{W})$: es decir, las transformaciones 2-lineales a valores en K son las formas bilineales que hemos considerado en las notas anteriores.

En los dos casos anteriores, notar el uso del “punto y coma” en la notación con el “Hom”, para diferenciar el rol del último de los espacios vectoriales. $\triangle$

14.1 Producto tensorial de dos espacios vectoriales

Proposición 14.7. *Sean* $\mathcal{V}, \mathcal{W}, \mathcal{U}$ *K-ev's y sea* $\varphi \in Hom(\mathcal{V}, \mathcal{W}\,; \mathcal{U})$. *Supongamos que* $B_\mathcal{V} = \{v_1, \ldots, v_n\}$ *y* $B_\mathcal{W} = \{w_1, \ldots, w_m\}$ *son bases de* $\mathcal{V}$ *y* $\mathcal{W}$*, tales que*

$$B = \{\varphi(v_j, w_k) : 1 \le j \le n,\ 1 \le k \le m\}$$

es una base de $\mathcal{U}$. *Entonces dadas cualesquiera* $B'_\mathcal{V} = \{u_1, \ldots, u_n\}$ *y* $B'_\mathcal{W} = \{z_1, \ldots, z_m\}$ *bases de* $\mathcal{V}$ *y* $\mathcal{W}$ *respectivamente, entonces* $B' = \{\varphi(u_j, z_k) : 1 \le j \le n,\ 1 \le k \le m\}$ *es base de* $\mathcal{U}$.

Demostración. Por hipótesis $\dim_K \mathcal{U} = \#(B) = n \cdot m$, pues B es base de $\mathcal{U}$. Como $\#(B') = n \cdot m$, entonces para verificar que B' es base, basta ver que B' genera a $\mathcal{U}$. Así, verificamos que B' es un sistema de generadores de $\mathcal{U}$; sea entonces $u \in \mathcal{U}$: por hipótesis, existen coeficientes $\alpha_{j,k} \in K$ para $1 \le j \le n$ y $1 \le k \le m$ tales que

$$u = \sum_{j=1}^{n} \sum_{k=1}^{m} \alpha_{j,k}\ \varphi(v_j, w_k)\,. \tag{109}$$

Por otro lado, dados $1 \le j \le n$ y $1 \le k \le m$ existen coeficientes $\beta_{\ell,j}$, $\gamma_{h,k} \in K$, para $1 \le \ell \le n$ y $1 \le h \le m$ tales que

$$v_j = \sum_{\ell=1}^{n} \beta_{\ell,j}\, u_\ell \quad \text{y} \quad w_k = \sum_{h=1}^{m} \gamma_{h,k}\, z_h\,. \tag{110}$$

Usando las representaciones en las Eqs. 109 y 110 vemos que

$$\begin{aligned} u &= \sum_{j=1}^{n}\sum_{k=1}^{m} \alpha_{j,k}\ \varphi\Big(\sum_{\ell=1}^{n} \beta_{\ell,j}\, u_\ell\,,\ \sum_{h=1}^{m} \gamma_{h,k}\, z_h\Big) = \sum_{j=1}^{n}\sum_{k=1}^{m} \alpha_{j,k}\ \sum_{\ell=1}^{n} \beta_{\ell,j}\, \varphi\Big(u_\ell\,,\ \sum_{h=1}^{m} \gamma_{h,k}\, z_h\Big) \\ &= \sum_{j=1}^{n}\sum_{k=1}^{m} \alpha_{j,k}\ \sum_{\ell=1}^{n} \beta_{\ell,j}\ \sum_{h=1}^{m} \gamma_{h,k}\, \varphi(u_\ell\,,\ z_h) = \sum_{j=1}^{n}\sum_{k=1}^{m}\sum_{\ell=1}^{n}\sum_{h=1}^{m} \alpha_{j,k}\ \beta_{\ell,j}\ \gamma_{h,k}\ \varphi(u_\ell\,,\ z_h) \\ &= \sum_{\ell=1}^{n}\sum_{h=1}^{m} \left(\sum_{j=1}^{n}\sum_{k=1}^{m} \alpha_{j,k}\ \beta_{\ell,j}\ \gamma_{h,k} \right)\ \varphi(u_\ell\,,\ z_h) = \sum_{\ell=1}^{n}\sum_{h=1}^{m} \eta_{\ell,h}\ \varphi(u_\ell\,,\ z_h) \end{aligned}$$

donde hemos usado la 2-linealidad de φ, y hemos intercambiado el orden de sumanción (sin modificar la suma total), y donde hemos considerado

$$\eta_{\ell,h} = \sum_{j=1}^{n}\sum_{k=1}^{m} \alpha_{j,k}\ \beta_{\ell,j}\ \gamma_{h,k}\ \in\ K\,.$$

Las identidades anteriores muestran que B' es un sistema de generadores de $\mathcal{U}$; por el argumento previamente mencionado en este caso, deducimos que B' es base de $\mathcal{U}$. □

Ya estamos en condiciones de introducir el concepto de producto tensorial de dos K-ev's.

Definición 14.8. *Sean* $\mathcal{V}, \mathcal{W}$ *K-ev's de dimensión finita. Un producto tensorial de* $\mathcal{V}$ *y* $\mathcal{W}$ *es un par* $(\otimes\,, \mathcal{X})$ *formado por un K-ev* $\mathcal{X}$ *y una transformación 2-lineal* $\otimes : \mathcal{V} \times \mathcal{W} \to \mathcal{X}$ *tales que si* $B_\mathcal{V} = \{v_1, \ldots, v_n\}$ *y* $B_\mathcal{W} = \{w_1, \ldots, w_m\}$ *son bases de* $\mathcal{V}$ *y* $\mathcal{W}$ *respectivamente, entonces* $\{\otimes(v_j, w_k)\ :\ 1 \le j \le n\ ,\ 1 \le k \le m\}$ *es base de* $\mathcal{X}$. □

Obs 14.9. Con las notaciones de la definición anterior, notemos que la Proposición 14.7, indica que basta verificar que $\{\otimes(v_j, w_k)\ :\ 1 \le j \le n\ ,\ 1 \le k \le m\}$ es base de $\mathcal{X}$ para un par de bases

$B_{\mathcal{V}}$ y $B_{\mathcal{W}}$ particulares. Por otro lado, dado un producto tensorial $(\otimes\,,\,\mathcal{X})$ de $\mathcal{V}$ y $\mathcal{W}$ en adelante notaremos:

$$v\otimes w:=\otimes(v\,,\,w)\quad \text{ para }\quad (v\,,\,w)\in\mathcal{V}\times\mathcal{W}\,.$$

En este contexto $v\otimes w$ es llamado un *tensor elemental*.

Las propiedades de 2-linealidad de $\otimes$ se pueden re-escribir con nuestra nueva convención como sigue: si $\mathbf{v}_1,\ldots,\mathbf{v}_n\in\mathcal{V}$ y $w_1,\ldots,w_m\in\mathcal{W}$ y $\alpha_1,\ldots,\alpha_n\in K$ y $\beta_1,\ldots,\beta_m\in K$ entonces

$$\begin{aligned}(\sum_{j=1}^n\alpha_j\,v_j)\otimes(\sum_{k=1}^m\beta_k\,w_k) &= \sum_{j=1}^n\alpha_j\ (v_j\otimes(\sum_{k=1}^m\beta_k\,w_k))=\sum_{j=1}^n\alpha_j\ \sum_{k=1}^m\beta_k\ (v_j\otimes w_k)\\ &= \sum_{j=1}^n\sum_{k=1}^m\alpha_j\,\beta_k\ v_j\otimes w_k\end{aligned}$$

Hemos incluido el uso de parántesis, para clarificar el sentido de las expresiones anteriores. Sin embargo, en adelante, vamos a evitar el uso de tales paréntesis y el lector deberá inspeccionar con cuidado las expresiones y en particular, el uso de la notación $\otimes$. $\triangle$

Ejemplo 14.10. Sea K un cuerpo, y consideremos $K^n=K^{n\times 1}$, $K^m=K^{m\times 1}$ y $K^{n\times m}$ como K-ev's. Sea $\otimes:K^n\times K^m\to K^{n\times m}$ dada por

$$\otimes(\vec{x},\vec{y})=\vec{x}\,\vec{y}^{\,t}\in K^{n\times m}\,,$$

donde $y^t\in K^{1\times m}$ denota el transpuesto del vector $\vec{y}$. Entonces $\otimes$ es una transformación 2-lineal (ejercicio). Más aún, si $B_1=\{e_1,\ldots,e_n\}$ y $B_2=\{f_1,\ldots,f_m\}$ denotan las bases canónicas de K^n y K^m respectivamente, entonces

$$e_i\otimes f_j=\otimes(e_i,f_j)=e_i\,f_j^t=E^{i,j}\in K^{n\times m}$$

donde $E^{i,j}$ es la matriz que tiene una única entrada no nula, que es igual a 1 y se encuentra en la entrada (i,j): verificar, ejercicio. Hemos visto que $\{E^{i,j}\ :\ 1\leq i\leq n\ ,\ 1\leq j\leq m\}$ es una base de $K^{n\times m}$. De esta forma, el par $(\otimes,K^{n\times m})$ cumple con las propiedades de la Definición 14.8 y es un producto tensorial de K^n y K^m. $\triangle$

Obs 14.11 (Fundamental!!)**.** Sean $\mathcal{V},\mathcal{W}$ K-ev's de dimensión finita. Calculamos un producto tensorial de los espacios $\mathcal{V}$ y $\mathcal{W}$; para eso, consideramos $(v,w)\in\mathcal{V}\otimes W$: en este caso, podemos definir $v\otimes w\in L(\mathcal{V}^*,\mathcal{W})$ (ojo: aquí la expresión $v\otimes w$ es un nombre para cierta función!) dada por

$$v\otimes w(f)=f(v)\cdot w\quad \text{ para }\quad f\in\mathcal{V}^*\,,$$

(recordar que $f(v)\in K$ es un escalar). Es claro que $v\otimes w$ así definido es una transformación lineal de $\mathcal{V}^*$ en $\mathcal{W}$:

$$\begin{aligned}v\otimes w(\alpha\,f+g) &= (\alpha\,f+g)(v)\cdot w=(\alpha f(v)+g(v))\cdot w=\alpha f(v)\cdot w+g(v)\cdot w\\ &= \alpha\ v\otimes w(f)+v\otimes w(g)\,.\end{aligned}$$

Además, la función $\otimes(\cdot,\cdot):\mathcal{V}\times\mathcal{W}\to L(\mathcal{V}^*,\mathcal{W})$ dada por

$$\otimes(v,w)=v\otimes w\in L(\mathcal{V},\mathcal{W})$$

es 2-lineal: por ejemplo

$$\begin{aligned}\otimes(\alpha\,u+v,w)(f) &= (\alpha\cdot u+v)\otimes w(f)=f(\alpha\,u+v)\cdot w=(\alpha\,f(u)+f(v))\,w\\ &= (\alpha\,f(u))\cdot w+f(v)\cdot w=\alpha\ u\otimes w(f)+v\otimes w(f)\\ &= [\alpha\cdot\otimes(u,w)+\otimes(v,w)](f)\end{aligned}$$

donde hemos usado la definición del valor $(\alpha\, u+v)\otimes w(f)$, la definición de $\alpha\cdot u\otimes w$ y $v\otimes w$, y finalmente la definición de la suma en $L(\mathcal{V}^*,\mathcal{W})$. Como $f\in\mathcal{V}^*$ era arbitrario, entonces concluimos la identidad entre las transformaciones lineales

$$\otimes(\alpha\, u+v,w)=\alpha\cdot\otimes(u,w)+\otimes(v,w)\,.$$

De forma similar, se verifica

$$\otimes(v,\alpha\, w+z)=\alpha\cdot\otimes(v,w)+\otimes(v,z)\,.$$

Las dos ecuaciones anteriores muestran que la función $\otimes:\mathcal{V}\times\mathcal{W}\to L(\mathcal{V}^*,\mathcal{W})$ es 2-lineal. Por otro lado, sea $B_\mathcal{V}=\{v_1,\ldots,v_n\}$ base de $\mathcal{V}$, sea $B^*_\mathcal{V}=\{f_1,\ldots,f_n\}$ la base dual de la base $B_\mathcal{V}$ y sea $B_\mathcal{W}=\{w_1,\ldots,w_m\}$ base de $\mathcal{W}$.

Recordemos que hemos construido una base de $L(\mathcal{V}^*,\mathcal{W})$ como sigue: dado $1\le p\le m$ y $1\le q\le n$ entonces definimos

$$E^{(p,q)}\in L(\mathcal{V}^*,\mathcal{W})\quad\text{ tal que }\quad E^{(p,q)}(f_j)=\delta_{jq}\,w_p\ \ ,\ \ 1\le j\le n\,,$$

donde δ_{jp} denota la delta de Kroncker; en este caso, hemos probado que el conjunto

$$B=\{E^{(p,q)}\ :\ 1\le p\le m\ ,\ 1\le q\le n\}$$

es base de $L(\mathcal{V}^*,\mathcal{W})$ (ver notas anteriores). Notemos que si elegimos $1\le p\le m$, $1\le q\le n$ y $1\le j\le n$ entonces

$$v_q\otimes w_p(f_j)=f_j(v_q)\cdot w_p=\delta_{jp}\,w_p=E^{(p,q)}(f_j)$$

donde hemos usado la definición de la transformación $v_q\otimes w_p$ y las relaciones de dualidad entre la base $B_\mathcal{V}$ y $B^*_\mathcal{V}$. Como $1\le j\le n$ es arbitrario, y $B^*_\mathcal{V}=\{f_1,\ldots,f_n\}$ es base de $\mathcal{V}^*$ entonces se tiene la igualdad entre las transformaciones lineales

$$v_q\otimes w_p=E^{(p,q)}\implies B=\{v_q\otimes w_p\ :\ 1\le p\le m\ ,\ 1\le q\le n\}$$

es base de $L(\mathcal{V}^*,\mathcal{W})$. Las propiedades anteriores junto con la Definición 14.8 muestran que el par $(\otimes\,,\,L(\mathcal{V}^*,\mathcal{W})$ es un producto tensorial de los espacios $\mathcal{V}$ y $\mathcal{W}$. Para finalizar, notemos que según la Definición 14.4 se tiene que $L(\mathcal{V}^*,\mathcal{W})=\mathrm{Hom}(\mathcal{V}^*\,;\mathcal{W})$ (transformaciones 1-lineales = transformaciones lineales!!). $\triangle$

Como consecuencia del desarrollo de la Observación anterior, hemos probado:

Teorema 14.12. *Sean $\mathcal{V}$, $\mathcal{W}$ K-ev's. Entonces existe un producto tensorial de $\mathcal{V}$ y $\mathcal{W}$. En este caso,* $\dim_K\mathcal{X}=\dim_K\mathcal{V}\cdot\dim_K\mathcal{W}$.

$\square$

Obs 14.13. Sean $\mathcal{V}$, $\mathcal{W}$ K-ev's de dimensión finita y consideremos las notaciones del Ejemplo 14.11. Hay una serie de cuestiones que siempre conviene tener en mente en el caso de productos tensoriales:

1. Por un lado, no todo elemento del producto tensorial $\mathcal{X}$ de $\mathcal{V}$ y $\mathcal{W}$ (en nuestro caso $\mathcal{X}=L(\mathcal{V}^*,\mathcal{W})$) es un tensor elemental: de hecho, si $v\in\mathcal{V}$ y $w\in\mathcal{W}$ entonces $v\otimes w\in L(\mathcal{V}^*,\mathcal{W})$ es una transformación que verifica: $v\otimes w(f)=f(v)\,w$, de forma que $\mathrm{Im}(v\otimes w)\subset\overline{\{w\}}$: así, si $w\ne 0\ne v$ entonces $\dim_K\mathrm{Im}(v\otimes w)=1$. Si las dimensiones de $\mathcal{W}$ y $\mathcal{V}$ son (ambas) mayores ó iguales que 2, entonces podemos construir transformaciones $T\in L(\mathcal{V}^*,\mathcal{W})$ con $\dim_K\mathrm{Im}(T)\ge 2$, de forma que T no puede ser igual a $v\otimes w$ para vectores $v\in\mathcal{V}$ y $w\in\mathcal{W}$. Esto no contradice el hecho de que los tensores elementales generan linealmente a $\mathcal{X}$ (porque siempre hay bases formadas por tensores elementales!).

2. Si $h \in \mathcal{V} \otimes \mathcal{W}$ entonces, en general, hay distintas formas de expresar a h como combinación de tensores elementales: el ejemplo más sencillo tal vez sea notar que
$$2v \otimes w = 2(v \otimes w) = v \otimes 2w$$
donde hemos usado la 2-linealidad del producto tensorial: así $(2v) \otimes w = v \otimes (2w)$, pero en general, $2v \neq v$ y $w \neq 2w$. En consecuencia, hay que tener cuidado al momento de representar un elemento arbitrario del producto tensorial como combinación lineal de tensores elementales. Sin embargo, recordemos que siempre podemos tomar una *base* de tensores elementales: en este caso, como es el caso de cualquier base, sí tenemos unicidad de la representación.

$\triangle$

Más adelante, veremos que el producto tensorial de espacios es *esencialmente* único.

Teorema 14.14 (Propiedad universal del producto tensorial)**.** *Sean $\mathcal{V}$, $\mathcal{W}$, U K-ev's de dimensión finita y sea $(\otimes\,,\,\mathcal{X})$ un producto tensorial de $\mathcal{V}$ y $\mathcal{W}$. Dada $\varphi \in Hom(\mathcal{V},\mathcal{W}\,;\mathcal{U})$ entonces existe una única transformación lineal $T_\varphi \in L(\mathcal{X},\,U)$ tal que*
$$\varphi = T_\varphi \circ \otimes \qquad \textit{es decir :} \qquad \varphi(v,w) = T_\varphi(v \otimes w) \qquad \textit{para} \qquad (v,w) \in \mathcal{V} \times \mathcal{W}\,. \tag{111}$$

Demostración. Sean $B_\mathcal{V} = \{v_1, \ldots, v_n\}$ y $B_\mathcal{W} = \{w_1, \ldots, w_m\}$ bases de $\mathcal{V}$ y $\mathcal{W}$ respectivamente. Entonces, por la Definición 14.8 y nuestra hipótesis, el conjunto
$$B = \{v_j \otimes w_k \ : \ 1 \leq j \leq n\ ,\ 1 \leq k \leq m\}$$
es base de $\mathcal{X}$. Entonces, definimos $T_\varphi \in L(\mathcal{X},\mathcal{U})$ como la única transformación lineal que transforma los elementos de la base B según la fórmula
$$T_\varphi(v_j \otimes w_k) = \varphi(v_j, w_k) \in \mathcal{U} \quad \text{para} \quad 1 \leq j \leq n\ ,\ 1 \leq k \leq m\,.$$
Notemos que hemos usado un resultado sobre la existencia (y abundancia) de transformaciones lineales. En este caso, T_φ queda completamente determinada y de forma única. Veamos que T_φ satisface la Eq. (111) : dados $v \in \mathcal{V}$ y $w \in \mathcal{W}$ arbitrarios, entonces existen coeficientes $\alpha_1, \ldots, \alpha_n \in K$ y $\beta_1, \ldots, \beta_m \in K$ tales que
$$v = \sum_{j=1}^n \alpha_j\, v_j \quad \text{y} \quad w = \sum_{k=1}^m \beta_k\, w_k\,.$$
Entonces, usando las propiedades de 2-linealidad de $\otimes$ y φ (ver Observaciones 14.2 y 14.9), y las propiedades de linealidad de T_φ tenemos que
$$\begin{aligned}
T_\varphi(v \otimes w) &= T_\varphi(\sum_{j=1}^n \alpha_j\, v_j \otimes \sum_{k=1}^m \beta_k\, w_k) = T_\varphi(\sum_{j=1}^n \alpha_j\ \sum_{k=1}^m \beta_k\ v_j \otimes w_k) \\
&= \sum_{j=1}^n \alpha_j\ \sum_{k=1}^m \beta_k\ T_\varphi(v_j \otimes w_k) = \sum_{j=1}^n \alpha_j\ \sum_{k=1}^m \beta_k\ \varphi(v_j, w_k) \\
&= \sum_{j=1}^n \alpha_j\ \ \varphi(v_j\,,\, \sum_{k=1}^m \beta_k\, w_k) = \varphi(\sum_{j=1}^n \alpha_j\, v_j\ ,\ w) = \varphi(v,w)
\end{aligned}$$
Finalmente, verificamos la unicidad de T_φ: si $R \in L(\mathcal{X},\mathcal{U})$ es una transformación lineal tal que $U(v \otimes w) = \varphi(v,w)$ para $(v,w) \in \mathcal{V} \times \mathcal{W}$ entonces
$$R(v_j \otimes w_k) = \varphi(v_j, w_k) = T_\varphi(v_j \otimes w_k) \quad \text{para} \quad 1 \leq j \leq n\ ,\ 1 \leq k \leq m\,.$$
De forma que T_φ, $R \in L(\mathcal{X},\mathcal{U})$ son transformaciones que toman los mismos valores en cada uno de los vectores de la base B de $\mathcal{X}$; pero esto último muestra que $T_\varphi = R$. $\square$

Obs 14.15. Con las notaciones del Teorema 14.14:

1. Decimos que T_φ es la linealización de φ ;
2. Podemos considerar que en la representación $\varphi = T_\varphi \circ \otimes$, la 2-linealidad de φ quedó absorvida por la 2-linealidad de $\otimes$.

$\triangle$

Corolario 14.16 (Unicidad del producto tensorial)**.** *Sean $\mathcal{V}$, $\mathcal{W}$ dos K-ev's y sean $(\otimes_1, \mathcal{X}_1)$ y $(\otimes_2, \mathcal{X}_2)$ dos productos tensoriales de $\mathcal{V}$ y $\mathcal{W}$, según la Definición 14.8. Entonces existe un único isomorfismo $\Phi : \mathcal{X}_1 \to \mathcal{X}_2$ que verifica*

$$\Phi(v \otimes_1 w) = v \otimes_2 w \quad \textit{para} \quad (v, w) \in \mathcal{V} \times \mathcal{W} .$$

Demostración. Para probar este resultado, vamos a usar la propiedad universal del producto tensorial enunciada en el Teorema 14.14. En efecto, notemos que por hipótesis, $\otimes_2 \in \operatorname{Hom}(\mathcal{V}, \mathcal{W}\,; \mathcal{X}_2)$ es una transformación 2-lineal. Por el Teorema 14.14 aplicada al producto tensorial $(\otimes_1,\, \mathcal{X}_1)$ y $\varphi = \otimes_2$, existe una única transformación $T_{\otimes_2} \in L(\mathcal{X}_1, \mathcal{X}_2)$ tal que

$$T_{\otimes_2}(v \otimes_1 w) = \otimes_2(v, w) = v \otimes_2 w \quad \text{para} \quad (v, w) \in \mathcal{V} \times \mathcal{W} .$$

Llamemos $\Phi = T_{\otimes_2}$; entonces, usando la identidad anterior y la definición de Φ:

$$\Phi(v \otimes_1 w) = v \otimes_2 w \quad \text{para} \quad (v, w) \in \mathcal{V} \times \mathcal{W} .$$

Veamos que Φ es isomorfismo: si $B_\mathcal{V} = \{v_1, \ldots, v_n\}$ y $B_\mathcal{W} = \{w_1, \ldots, w_m\}$ son bases de $\mathcal{V}$ y $\mathcal{W}$ respectivamente, entonces sabemos que

$$B_1 = \{v_j \otimes_1 w_k \ : \ 1 \le j \le n \ , \ 1 \le k \le m\} \quad \text{y} \quad B_2 = \{v_j \otimes_2 w_k \ : \ 1 \le j \le n \ , \ 1 \le k \le m\}$$

son bases de $\mathcal{X}_1$ y $\mathcal{X}_2$ respectivamente. Pero notemos que, en particular,

$$\Phi(v_j \otimes_1 w_k) = v_j \otimes_2 w_k \quad \text{para} \quad 1 \le j \le n \ , \ 1 \le k \le m \, ,$$

de forma que Φ transforma la base B_1 en la base B_2. Por resultados vistos previamente, Φ es un isomorfismo lineal. $\square$

Obs 14.17. Con las notaciones del corolario anterior, notemos que sabíamos de antemano que $\dim_K \mathcal{X}_1 = \dim_K \mathcal{X}_2$ (ver el Teorema 14.12). De forma que los espacios son isomorfos! Pero el isomorfismo Φ del corolario anterior es especial: este isomorfismo además identifica cada tensor elemental $v \otimes_1 w$ con el correspondiente $v \otimes_2 w$, de forma que en realidad Φ permite identificar toda la información contenida en los pares $(\otimes_1, \mathcal{X}_1)$ y $(\otimes_2, \mathcal{X}_2)$. (Cabe aclarar que es posible construir isomorfismos lineales entre $\mathcal{X}_1$ y $\mathcal{X}_2$ que no respetan a los tensores elementales, en el sentido anterior). Es por esto que Φ es un isomorfismo entre (toda la estructura) de productos tensoriales. Por otro lado, notemos que la construcción de este isomorfismo no hizo uso de bases de los espacios; es por esto que es llamado un isomorfismo natural.

Es por esto que en adelante, usaremos la notación

$$(\otimes, \mathcal{V} \otimes \mathcal{W})$$

para el (esencialmente) único producto tensorial de los espacios $\mathcal{V}$ y $\mathcal{W}$ (incluso, a veces simplemente escribimos $\mathcal{V} \otimes \mathcal{W}$ para el producto tensorial, donde la transformación 2-lineal $\otimes : \mathcal{V} \times \mathcal{W} \to \mathcal{V} \otimes \mathcal{W}$ queda implícita). $\triangle$

Ejemplo 14.18. Sean $\mathcal{V}$, $\mathcal{W}$ K-ev's de dimensión finita. Entonces

$$\mathcal{V}^* \otimes \mathcal{W} = L(\mathcal{V}\,,\,\mathcal{W})\,. \tag{112}$$

En efecto, dados $f \in \mathcal{V}^*$, $w \in \mathcal{W}$ definimos $f\otimes w \in L(\mathcal{V}\,,\,\mathcal{W})$ dada por $f\otimes w(v) = f(v)\cdot w$, para $v \in \mathcal{V}$ (verificar que $f\otimes w$ es transformación lineal). De esta forma, queda definida $\otimes : \mathcal{V}^*\times\mathcal{W} \to L(\mathcal{V}\,,\,\mathcal{W})$ dada por

$$\mathcal{V}^* \times \mathcal{W} \ni (f\,,\,w) \mapsto f\otimes w \in L(\mathcal{V}\,,\,\mathcal{W})$$

que es una transformación 2-lineal (verificar). Finalmente, si $B_{\mathcal{V}} = \{v_1,\ldots,v_n\}$ es base de $\mathcal{V}$ con base dual $B_{\mathcal{V}}^* = \{f_1,\ldots,f_n\}$ y $B_{\mathcal{W}} = \{w_1,\ldots,w_m\}$ es base de $\mathcal{W}$: si $1\le p\le m$ y $1\le q\le n$ entonces

$$f_q\otimes w_p(v_j) = f_q(v_j)\,w_p = \delta_{jq}\,w_p = E^{(p,q)}(v_j) \quad \text{para} \quad 1\le j\le n$$

donde $\{E^{(p,q)} = f_q\otimes w_p \ :\ 1\le p\le m\ ,\ 1\le q\le n\}$ es la base de $L(\mathcal{V}\,,\,\mathcal{W})$ asociada a las bases $B_{\mathcal{V}}$ y $B_{\mathcal{W}}$, como en las notas para el tp4. Por el Corolario 14.16 y la Observación 14.17 concluimos que vale la Eq. (112). △

Ejemplo 14.19. Sean $\mathcal{V}$, $\mathcal{W}$ K-ev's de dimensión finita. Entonces

$$\mathcal{V}^* \otimes \mathcal{W}^* = \mathrm{Bil}(\mathcal{V}\,,\,\mathcal{W})\,. \tag{113}$$

En efecto, dados $f \in \mathcal{V}^*$, $g \in \mathcal{W}^*$ definimos $f\otimes g \in \mathrm{Bil}(\mathcal{V}\,,\,\mathcal{W})$ dada por $f\otimes g(v,w) = f(v)\cdot g(w)$, para $(v,w)\in\mathcal{V}\times\mathcal{W}$ (verificar que $f\otimes g$ es bilineal). De esta forma, queda definida $\otimes : \mathcal{V}^*\times\mathcal{W}^* \to \mathrm{Bil}(\mathcal{V}\,,\,\mathcal{W})$ dada por

$$\mathcal{V}^* \times \mathcal{W}^* \ni (f\,,\,g) \mapsto f\otimes g \in \mathrm{Bil}(\mathcal{V}\,,\,\mathcal{W})$$

que es una transformación 2-lineal (verificar). Finalmente, si $B_{\mathcal{V}} = \{v_1,\ldots,v_n\}$ es base de $\mathcal{V}$ con base dual $B_{\mathcal{V}}^* = \{f_1,\ldots,f_n\}$ y $B_{\mathcal{W}} = \{w_1,\ldots,w_m\}$ es base de $\mathcal{W}$ con base dual $B_{\mathcal{W}}^* = \{g_1,\ldots,g_m\}$ entonces $B = \{f_i\otimes g_j \ :\ 1\le i\le n\ ,\ 1\le j\le m\}$ es base de $\mathrm{Bil}(\mathcal{V}\,,\,\mathcal{W})$: de hecho, si $\mathcal{A}\in\mathrm{Bil}(\mathcal{V}\,,\,\mathcal{W})$ entonces

$$\mathcal{A} = \sum_{i=1}^{n}\sum_{j=1}^{m} \mathcal{A}(v_i,w_j)\cdot f_i\otimes g_j \in \mathrm{Bil}(\mathcal{V}\,,\,\mathcal{W})\,.$$

(verificar la identidad anterior). Entonces B es un sistema de generadores de $\mathrm{Bil}(\mathcal{V}\,,\,\mathcal{W})$ con $n\cdot m$ elementos, con lo cual B es una base de este espacio. Por el Corolario 14.16 y la Observación 14.17 concluimos que vale la Eq. (113). △

Corolario 14.20. *Sean $\mathcal{V}$, $\mathcal{W}$, $\mathcal{U}$ K-ev's de dimensión finita. Entonces la función*

$$\Psi : \mathit{Hom}(\mathcal{V},\mathcal{W}\,;\mathcal{U}) \to L(\mathcal{V}\otimes\mathcal{W},\mathcal{U}) = \mathit{Hom}(\mathcal{V}\otimes\mathcal{W}\,;\mathcal{U})$$

dada por $\Psi(\varphi) = T_\varphi$, para $\varphi \in \mathit{Hom}(\mathcal{V},\mathcal{W}\,;\mathcal{U})$ (donde T_φ es como en el Teorema 14.14) es un isomorfismo de K-ev's. En particular,

$$\dim_K \mathit{Hom}(\mathcal{V},\mathcal{W}\,;\mathcal{U}) = \dim_K L(\mathcal{V}\otimes\mathcal{W},\mathcal{U}) = \dim_K \mathcal{V}\otimes\mathcal{W}\cdot\dim_K\mathcal{U} = \dim_K\mathcal{V}\cdot\dim_K\mathcal{W}\cdot\dim_K\mathcal{U}\,.$$

Demostración. Como fue indicado en la Proposición 14.5, $\mathrm{Hom}(\mathcal{V},\mathcal{W}\,;\mathcal{U})$ es un K-ev (de hecho, en el enunciado de este resultado hemos descrito la suma y la acción por escalares de este espacio vectorial). Recordemos además que dada $\varphi \in \mathrm{Hom}(\mathcal{V},\mathcal{W}\,;\mathcal{U})$ entonces $T_\varphi \in L(\mathcal{V}\otimes\mathcal{W},\mathcal{U})$ esta determinada de forma única por la identidad $T_\varphi(v\otimes w) = \varphi(v,w)$ para $(v,w)\in\mathcal{V}\times\mathcal{W}$ (ver Teorema 14.14). Veamos que Ψ es lineal: en efecto, si φ, $\psi \in \mathrm{Hom}(\mathcal{V},\mathcal{W}\,;\mathcal{U})$ y $\alpha\in K$ entonces:

$$(\alpha\,\varphi+\psi)(v,w) = \alpha\,\varphi(v,w)+\psi(v,w) = \alpha\,T_\varphi(v\otimes w)+T_\psi(v\otimes w) = (\alpha\,T_\varphi+T_\psi)(v\otimes w)\,.$$

Así, $(\alpha\,\varphi+\psi)(v,w)=(\alpha\,T_\varphi+T_\psi)(v\otimes w)$ para $(v,w)\in\mathcal{V}\times\mathcal{W}$. Por la unicidad mencionada más arriba,

$$\Psi(\alpha\,\varphi+\psi)=\alpha\,T_\varphi+T_\psi=\alpha\,\Psi(\varphi)+\Psi(\psi)\,,$$

de forma que Ψ es lineal.

Por un lado, Ψ es inyectiva: si $\varphi\in\mathrm{Hom}(\mathcal{V},\mathcal{W}\,;\mathcal{U})$ es tal que $\Psi(\varphi)=0$ entonces $\varphi(v,w)=0(v\otimes w)=0_{\mathcal{U}}$ para $(v,w)\in\mathcal{V}\times\mathcal{W}$, de forma que $\varphi=0$ (es el vector nulo de $\varphi\in\mathrm{Hom}(\mathcal{V},\mathcal{W}\,;\mathcal{U})$). Lo anterior indica que $\mathrm{N}(\Psi)=\{0\}$ y Ψ es inyectiva.

Finalmente, Ψ es suryectiva: dada $T\in L(\mathcal{V}\otimes\mathcal{W},\mathcal{U})$ entonces definimos $\varphi:\mathcal{V}\times\mathcal{W}\to\mathcal{U}$ dada por $\varphi(v,w)=T(v\otimes w)$, para $(v,w)\in\mathcal{V}\times\mathcal{W}$: entonces φ es 2-lineal: en efecto

$$\begin{aligned}\varphi(\alpha\,u+v,w)&=T((\alpha\,u+v)\otimes w)=T(\alpha\,u\otimes w+v\otimes w)=\alpha\,T(u\otimes w)+T(v\otimes w)\\&=\alpha\,\varphi(u,w)+\varphi(v,w)\,,\end{aligned}$$

donde hemos usado que el producto tensorial es 2-lineal y la linealidad de la transformación T definida en $\mathcal{V}\otimes\mathcal{W}$. De forma similar, se verifica $\varphi(v,\alpha w+z)=\alpha\varphi(v,w)+\varphi(v,z)$. Es claro que en este caso, $T_\varphi=T$; es decir, $\Psi(\varphi)=T$ y Ψ es suryectiva.

Notemos que la afirmación sobre las dimensiones es consecuencia de que Ψ es isomorfismo, del cálculo de la dimensión del espacio de transformaciones lineales (resultado previo) y del hecho de que el producto tensorial verifica $\dim_K\mathcal{V}\otimes\mathcal{W}=\dim_K\mathcal{V}\cdot\dim_K\mathcal{W}$. □

Teorema 14.21. *Sean $\mathcal{V}$, $\mathcal{W}$, $\mathcal{Y}$ K-ev's de dimensión finita. Entonces*

1. *La función $\mathcal{V}\otimes\mathcal{W}\ni v\otimes w\mapsto w\otimes v\in\mathcal{W}\otimes\mathcal{V}$ se extiende (de forma única) a un isomorfismo $C:\mathcal{V}\otimes\mathcal{W}\to\mathcal{W}\otimes\mathcal{V}$ de K-ev's.*

2. *La función $(\mathcal{V}\otimes\mathcal{W})\otimes\mathcal{Y}\ni(v\otimes w)\otimes y\mapsto v\otimes(w\otimes y)\in\mathcal{V}\otimes(\mathcal{W}\otimes\mathcal{Y})$ se extiende (de forma única) a un isomorfismo $A:(\mathcal{V}\otimes\mathcal{W})\otimes\mathcal{Y}\to\mathcal{V}\otimes(\mathcal{W}\otimes\mathcal{Y})$ de K-ev's.*

Demostración. Comenzamos verificando 1: en este caso, consideramos $\varphi:\mathcal{V}\times\mathcal{W}\to\mathcal{W}\otimes\mathcal{V}$ dada por $\varphi(v,w)=w\otimes v$, para $(v,w)\in\mathcal{V}\times\mathcal{W}$. Usando las propiedades de 2-linealidad del producto tensorial $\otimes:\mathcal{W}\times\mathcal{V}\to\mathcal{W}\otimes\mathcal{V}$ concluimos que φ es una transformación 2-lineal. Así, por el Teorema 14.14 existe un única $C:=T_\varphi\in L(\mathcal{V}\otimes\mathcal{W},\mathcal{W}\otimes\mathcal{V})$ tal que $C(v\otimes w)=w\otimes v$, para $(v,w)\in\mathcal{V}\times\mathcal{W}$. Si $B_{\mathcal{V}}=\{v_1,\ldots,v_n\}$ y $B_{\mathcal{W}}=\{w_1,\ldots,w_m\}$ son bases de $\mathcal{V}$ y $\mathcal{W}$ entonces $C(v_j\otimes w_k)=w_k\otimes v_j$ para $1\le j\le n$, $1\le k\le m$; de esta forma, C transforma una base $\mathcal{V}\otimes\mathcal{W}$ en una base $\mathcal{W}\otimes\mathcal{V}$ lo que indica que C es un isomorfismo lineal.

Para verificar 2, sea $y\in\mathcal{Y}$ fijo: entonces, la función $\varphi_y:\mathcal{V}\times\mathcal{W}\to\mathcal{V}\otimes(\mathcal{W}\otimes\mathcal{Y})$ dada por $\varphi_y(v,w)=v\otimes(w\otimes y)$, para $(v,w)\in\mathcal{V}\times\mathcal{W}$ es 2-lineal (ejercicio). Entonces, existe un única $T_y\in L(\mathcal{V}\otimes\mathcal{W},\mathcal{V}\otimes(\mathcal{W}\otimes\mathcal{Y}))$ tal que

$$T_y(v\otimes w)=\varphi_y(v,w)=v\otimes(w\otimes y)\,.$$

Más aún, por la unicidad de T_y (para cada $y\in\mathcal{Y}$), si $y,z\in\mathcal{Y}$ y $\alpha\in K$ entonces

$$\begin{aligned}T_{\alpha\,y+z}(v\otimes w)&=v\otimes(w\otimes\alpha\,y+z)=v\otimes(\alpha\cdot w\otimes y+w\otimes z)=v\otimes(\alpha\,w\otimes y)+v\otimes(w\otimes z)\\&=\alpha\cdot[v\otimes(w\otimes y)]+v\otimes(w\otimes z)=(\alpha\,T_y+T_z)(v\otimes w)\,.\end{aligned}$$

Como los tensores elementales son un sistema de generadores para $\mathcal{V}\otimes\mathcal{W}$, las identidades anteriores muestran que la igualdad entre las transformaciones lineales $T_{\alpha\,y+z}=\alpha\,T_y+T_z$.

Finalmente, definimos $\psi:(\mathcal{V}\otimes\mathcal{W})\times\mathcal{Y}\to\mathcal{V}\otimes(\mathcal{W}\otimes\mathcal{Y})$ dada por $\psi(v\otimes w,y)=T_y(v\otimes w)$; entonces ψ es una transformación 2-lineal !! (la linealidad en la primer variable - es decir, en el espacio

$\mathcal{V}\otimes\mathcal{W}$ - es consecuencia de que T_y es transformación lineal; la linealidad en la segunda variable es consecuencia de la identidad $T_{\alpha\, y+z}=\alpha\, T_y+T_z$): ejercicio. Una nueva aplicación del Teorema 14.14 nos permite considerar la transformación lineal $A:=T_\psi\in L((\mathcal{V}\otimes\mathcal{W})\times\mathcal{Y},\mathcal{V}\otimes(\mathcal{W}\times\mathcal{Y}))$ tal que

$$A((v\otimes w)\otimes y)=\psi(v\otimes w,y)=v\otimes(w\otimes y)\,.$$

Si $B_\mathcal{V}=\{v_1,\ldots,v_n\}$, $B_\mathcal{W}=\{w_1,\ldots,w_m\}$ y $B_\mathcal{Y}=\{y_1,\ldots,y_p\}$ son bases de $\mathcal{V}$, $\mathcal{W}$ e $\mathcal{Y}$ entonces: $\{v_j\otimes w_k\ :\ 1\le j\le n,\, 1\le k\le m\}$ es base de $\mathcal{V}\otimes\mathcal{W}$; entonces

$$B_1=\{(v_j\otimes w_k)\otimes y_\ell\ :\ 1\le j\le n,\, 1\le k\le m,\, 1\le \ell\le p\}$$

es base de $(\mathcal{V}\otimes\mathcal{W})\times\mathcal{Y}$. De forma análoga

$$B_2=\{v_j\otimes(w_k\otimes y_\ell)\ :\ 1\le j\le n,\, 1\le k\le m,\, 1\le \ell\le p\}$$

es base de $\mathcal{V}\otimes(\mathcal{W}\times\mathcal{Y})$. Notemos que A transforma los elementos de la base B_1 en los elementos de la base B_2: $A(v_j\otimes w_k)\otimes y_\ell)=v_j\otimes(w_k\otimes y_\ell)$; Así, A es un isomorfismo de K-ev's. □

Obs 14.22. Con las notaciones del Teorema anterior y de su prueba, la propiedad del ítem 1 es la conmutatividad del producto tensorial de espacios vectoriales. Por otro lado, la propiedad del ítem 2 es la asociatividad del producto tensorial de espacios vectoriales. Notemos que los isomorfismos C y A se han construido sin hacer uso de bases del espacio, y además permiten identificar tensores elementales de cada espacio de forma natural.

Es por esto que en adelante identificaremos

$$\mathcal{V}\otimes\mathcal{W}=\mathcal{W}\otimes\mathcal{V}\quad\text{ y}$$

$$\mathcal{V}\otimes\mathcal{W}\otimes\mathcal{Y}:=(\mathcal{V}\otimes\mathcal{W})\otimes\mathcal{Y}=\mathcal{V}\otimes(\mathcal{W}\otimes\mathcal{Y})\,.$$

Además, definimos $\otimes:\mathcal{V}\times\mathcal{W}\times\mathcal{Y}\to\mathcal{V}\otimes\mathcal{W}\otimes\mathcal{Y}$ tal que

$$\otimes(v,w,y)=v\otimes w\otimes y$$

que es una transformación 3-lineal. En el último caso, escribimos

$$B=\{v_j\otimes w_k\otimes y_\ell\ :\ 1\le j\le n,\, 1\le k\le m,\, 1\le \ell\le p\}$$

para una base de $\mathcal{V}\otimes\mathcal{W}\otimes\mathcal{Y}$. En particular, $\dim_K\mathcal{V}\otimes\mathcal{W}\otimes\mathcal{Y}=\dim_K\mathcal{V}\dim_K\mathcal{W}\dim_K\mathcal{Y}$.

△

14.2 Producto tensorial de espacios vectoriales (caso general)

Basados en la propiedad asociativa del producto tensorial, ahora podemos definir el producto tensorial de varios espacio, como sigue:

Definición 14.23. *Sea $\mathcal{V}_j$ un K-ev con base $B_j=\{v_1^j,\ldots,v_{d_j}^j\}$, para $1\le j\le p$. Definimos el espacio producto tensorial como el par $(\otimes_{j=1}^p,\mathcal{V}_1\otimes\ldots\otimes\mathcal{V}_p)$ tal que:*

1. *Si $p=2$, entonces la construcción coincide con la construcción previa $(\otimes,\mathcal{V}_1\otimes\mathcal{V}_2)$ del producto tensorial;*
2. *Si $p\ge 3$ (asumiendo que hemos definido el producto de $p-1$ espacios): entonces*

$$\mathcal{V}_1\otimes\ldots\otimes\mathcal{V}_p:=(\mathcal{V}_1\otimes\ldots\otimes\mathcal{V}_{p-1})\otimes\mathcal{V}_p$$

y la transformación p-lineal $\otimes_{j=1}^p : \mathcal{V}_1 \times \ldots \times \mathcal{V}_p \to \mathcal{V}_1 \otimes \ldots \otimes \mathcal{V}_p$ *dada por*

$$\otimes_{j=1}^p (v_1, \ldots, v_p) = \otimes_{j=1}^{p-1}(v_1, \ldots, v_{p-1}) \otimes v_p \in \mathcal{V}_1 \otimes \ldots \otimes \mathcal{V}_p$$

y notamos $\otimes_{j=1}^p (v_1, \ldots, v_p) = v_1 \otimes \ldots \otimes v_p$. *En este caso, el conjunto*

$$B = \{v_{i_1}^1 \otimes \ldots \otimes v_{i_p}^p \ : \ 1 \leq i_j \leq d_j \ , \ 1 \leq j \leq p\}$$

es base de $\mathcal{V}_1 \otimes \ldots \otimes \mathcal{V}_p$. □

Obs 14.24. Sea $\mathcal{V}_j$ un K-ev de dimensión finita, para $1 \leq j \leq p$. El producto tensorial de los espacios también es notado $\otimes_{j=1}^p \mathcal{V}_j$. Notemos que por construicción

$$\dim_K \otimes_{j=1}^p \mathcal{V}_j = \prod_{j=1}^p \dim_K \mathcal{V}_j \, .$$

Como consecuencia de la propiedad asociativa y conmutativa, se verifica que

$$\otimes_{j=1}^p \mathcal{V}_j = (\otimes_{j=1}^k \mathcal{V}_{\sigma(j)}) \otimes (\otimes_{j=k+1}^p \mathcal{V}_{\sigma(j)})$$

donde $\sigma : \{1, \ldots, p\} \to \{1, \ldots, p\}$ es una permutación de los índices (función biyectiva). Por ejemplo

$$\mathcal{V}_1 \otimes \mathcal{V}_2 \otimes \mathcal{V}_3 \otimes \mathcal{V}_4 = (\mathcal{V}_2 \otimes \mathcal{V}_3) \otimes (\mathcal{V}_1 \otimes \mathcal{V}_4) \, .$$

(aquí usamos $\sigma(1) = 2$, $\sigma(2) = 3$, $\sigma(3) = 1$, $\sigma(4) = 4$). △

Teorema 14.25 (Propiedad universal del producto tensorial)**.** *Sea* $\mathcal{V}_j$ *un* K*-ev para* $1 \leq j \leq p$ *y* $\mathcal{U}$ *un* K*-ev, todos de dimensión finita. Si* $\varphi \in Hom(\mathcal{V}_1, \ldots, \mathcal{V}_p ; \mathcal{U})$ *es transformación p-lineal, entonces existe una única* $T_\varphi \in L(\otimes_{j=1}^p \mathcal{V}_j \, , \mathcal{U})$ *tal que*

$$T_\varphi(v_1 \otimes \ldots \otimes v_p) = \varphi(v_1, \ldots, v_p) \quad \textit{para} \quad (v_1, \ldots, v_p) \in \mathcal{V}_1 \times \ldots \times \mathcal{V}_p \, .$$

Demostración. Sea $B_j = \{v_1^j, \ldots, v_{d_j}^j\}$ base de $\mathcal{V}_j$, para $1 \leq j \leq p$. En este caso, el conjunto

$$B = \{v_{i_1}^1 \otimes \ldots \otimes v_{i_p}^p \ : \ 1 \leq i_j \leq d_j \ , \ 1 \leq j \leq p\}$$

es base de $\otimes_{j=1}^p \mathcal{V}_j$. Definimos $T_\varphi \in L(\otimes_{j=1}^p \mathcal{V}_j, \mathcal{U})$ como la única transformación lineal que transforma los vectores de la base B de la siguiente forma:

$$T_\varphi(v_{i_1}^1 \otimes \ldots \otimes v_{i_p}^p) = \varphi(v_{i_1}^1, \ldots, v_{i_p}^p) \in \mathcal{U} \quad \text{para} \quad 1 \leq i_j \leq d_j \ , \ 1 \leq j \leq p \, .$$

Para cada $1 \leq j \leq p$, si $v_j \in \mathcal{V}_j$ es arbitrario, entonces se puede representar como

$$v_j = \sum_{i_j=1}^{d_j} \alpha_{i_j \, j} \ v_{i_j}^j$$

para ciertos coeficiente $\alpha_{i_j \, j} \in K$: aquí, nuestra variable de sumación es i_j, que varía $1 \leq i_j \leq d_j$ (con este subíndice en la variable i, podemos considerar las variables de sumación distintas $i_1, \ldots, i_p$:

vamos a necesitar un nombre distinto para cada variable de sumación en la cuenta de más abajo!). Entonces

$$\begin{aligned}
\varphi(v_1,\ldots,v_p) &= \varphi(\sum_{i_1=1}^{d_1}\alpha_{i_1\,1}\;v^1_{i_1},\ldots,\sum_{i_p=1}^{d_p}\alpha_{i_p\,p}\;v^p_{i_p}) = \sum_{i_1=1}^{d_1}\alpha_{i_1 1}\;\varphi(v^1_{i_1},\ldots,\sum_{i_p=1}^{d_p}\alpha_{i_p\,p}\;v^p_{i_p}) \\
&= \ldots = \sum_{i_1=1}^{d_1}\alpha_{i_1 1}\ldots\sum_{i_p=1}^{d_p}\alpha_{i_p\,p}\;\varphi(v^1_{i_1},\ldots,v^p_{i_p}) \\
&= \sum_{i_1=1}^{d_1}\alpha_{i_1 1}\ldots\sum_{i_p=1}^{d_p}\alpha_{i_p\,p}\;T_\varphi(v^1_{i_1}\otimes\ldots\otimes v^p_{i_p}) \\
&= \sum_{i_1=1}^{d_1}\alpha_{i_1 1}\ldots\sum_{i_{p-1}=1}^{d_{p-1}}\alpha_{i_{p-1}\,p-1}\;T_\varphi(v^1_{i_1}\otimes\ldots\otimes\sum_{i_p=1}^{d_p}\alpha_{i_p\,p}v^p_{i_p}) \\
&= T_\varphi(\sum_{i_1=1}^{d_1}\alpha_{i_1\,1}\;v^1_{i_1}\otimes\ldots\otimes\sum_{i_p=1}^{d_p}\alpha_{i_p\,p}\;v^p_{i_p}) = T_\varphi(v_1\otimes\ldots\otimes v_p)\,,
\end{aligned}$$

donde hemos usado la p-linealidad de φ (es decir, la dependencia lineal en cada una de sus coordenadas - dejando las otras fijas); de esta forma, salen p signos de sumación acompañados de los coeficientes del cuerpo correspondientes; hemos usado la definición de T_φ, la linealidad de T_φ (las sumatorias y coeficientes entran dentro de T_φ) y la p-linealidad del producto tensorial (es decir, la dependencia lineal en cada una de sus coordenadas - dejando las otras fijas). La unicidad es ejercicio (sencillo). $\square$

Corolario 14.26 (Unicidad del producto tensorial)**.** *Sea $\mathcal{V}_j$ un K-ev para $1\le j\le p$ y $\mathcal{U}$ un K-ev, todos de dimensión finita. Supongamos que existe $\otimes_\mathcal{U}\in Hom(\mathcal{V}_1,\ldots,\mathcal{V}_p\,;\mathcal{U})$ con la siguiente propiedad: para $1\le j\le p$ existe una base $B_j=\{v^j_1,\ldots,v^j_{d_j}\}$ de $\mathcal{V}_j$, tales que*

$$B_\mathcal{U}=\{\otimes_\mathcal{U}(v^1_{i_1},\ldots,v^p_{i_p})\ :\ 1\le i_j\le d_j\ ,\ 1\le j\le p\}$$

es base de $\mathcal{U}$. Entonces existe un (único) isomorfismo de K-ev's $\Phi:\otimes^p_{j=1}\mathcal{V}_1\to\mathcal{U}$ tal que

$$\Phi(v_1\otimes\ldots\otimes v_p)=\otimes_\mathcal{U}(v_1,\ldots,v_p)\quad \textit{para}\quad (v_1,\ldots,v_p)\in\mathcal{V}_1\times\ldots\times\mathcal{V}_p\,. \tag{114}$$

Demostración. Por el Teorema 14.25, existe $\Phi=T_{\otimes_\mathcal{U}}\in L(\otimes^p_{j=1}\mathcal{V}_j\,,\mathcal{U})$ que verifica la Eq. (114). Además, Φ resulta un isomorfismo: en efecto, si consideramos las bases B_j como en el enunciado, entonces

$$B=\{v^1_{i_1}\otimes\ldots\otimes v^p_{i_p}\ :\ 1\le i_j\le d_j\ ,\ 1\le j\le p\}$$

es base de $\mathcal{V}_1\otimes\ldots\otimes\mathcal{V}_p$, por construcción (ver la Definición 14.23). Notemos que Φ transforma los elementos de B en los elementos de $B_\mathcal{U}$ (como en el enunciado). Así, Φ transforma una base de su dominio en una base de su co-dominio lo que muestra que Φ es un isomorfismo, por resultados previos. $\square$

Obs 14.27. Con las notaciones del Corolario anterior, el isomorfismo Φ identifica los espacios vectoriales de forma tal que también identifica los *tensores elementales*. En este caso, vamos a escribir $\mathcal{U}=\otimes^p_{j=1}\mathcal{V}_j$ y usar la notación $\otimes_\mathcal{U}(v_1,\ldots,v_p)=v_1\otimes\ldots\otimes v_p$, que es la unicidad del producto tensorial. $\triangle$

Corolario 14.28. *Sea $\mathcal{V}_j$ un K-ev para $1 \leq j \leq p$ y $\mathcal{U}$ un K-ev, todos de dimensión finita. Entonces, la función $\Psi : Hom(\mathcal{V}_1, \ldots, \mathcal{V}_p\,; \mathcal{U}) \to L(\otimes_{j=1}^p \mathcal{V}_j\,, \mathcal{U})$ dada por $\Psi(\varphi) = T_\varphi$ (como en el Teorema 14.25) para $\varphi \in Hom(\mathcal{V}_1, \ldots, \mathcal{V}_p\,; \mathcal{U})$, es un isomorfismo de K-ev's. En particular*

$$\dim_K Hom(\mathcal{V}_1, \ldots, \mathcal{V}_p\,; \mathcal{U}) = \prod_{j=1}^p \dim_K \mathcal{V}_j \cdot \dim_K \mathcal{U}\,.$$

Demostración Ejercicio para el lector. $\square$

Obs 14.29. Sea $\mathcal{V}_j$ un K-ev de dimensión finita, para $1 \leq j \leq p$. En este caso, tenemos que

$$(\otimes_{j=1}^p \mathcal{V}_j)^* = \otimes_{j=1}^p \mathcal{V}_j^*\,. \tag{115}$$

Notemos primero que si $f_j \in \mathcal{V}_j$, para $1 \leq j \leq p$, entonces la función

$$\varphi_{(f_1,\ldots,f_p)}(v_1, \ldots, v_p) = \prod_{j=1}^p f_j(v_j) \in K \quad \text{para} \quad (v_1, \ldots, v_p) \in \mathcal{V}_1 \times \ldots \times \mathcal{V}_p$$

es tal que $\varphi_{(f_1,\ldots,f_p)} \in \operatorname{Hom}(\mathcal{V}_1, \ldots, \mathcal{V}_p\,; K)$ (verificar). Entonces, existe una única transformación lineal $\otimes_*(f_1, \ldots, f_p) \in L(\otimes_{j=1}^p \mathcal{V}_j\,, K) = (\otimes_{j=1}^p \mathcal{V}_j)^*$ tal que

$$\otimes_*(f_1, \ldots, f_p)(v_1 \otimes \ldots \otimes v_p) = \prod_{j=1}^k f_j(v_j) \quad \text{para} \quad v_1 \otimes \ldots \otimes v_p \in \otimes_{j=1}^p \mathcal{V}_j\,.$$

De esta forma, hemos definido la función $\otimes_* : \mathcal{V}_1^* \times \ldots \times \mathcal{V}_p^* \to (\otimes_{j=1}^p \mathcal{V}_j)^*$. Notemos que $\otimes_*$ es una función p-lineal: en efecto, si $1 \leq j \leq p$, $g_j \in \mathcal{V}^*$ y $\alpha \in K$ entonces

$$\otimes_*(f_1, \ldots, f_{j-1}\,, \alpha\, f_j + g_j\,, f_{j+1}, \ldots, f_p)(v_1 \otimes \ldots \otimes v_p) = \prod_{1 \leq i \neq j \leq p} f_i(v_i) \cdot (\alpha\, f_j + g_j)(v_j)$$

$$= \prod_{1 \leq i \neq j \leq p} f_i(v_i) \cdot (\alpha\, f_j(v_j) + g_j(v_j)) = \alpha \prod_{1 \leq i \neq j \leq p} f_i(v_i) \cdot f_j(v_j) + \prod_{1 \leq i \neq j \leq p} f_i(v_i) \cdot g_j(v_j) =$$

$$= \alpha \otimes_* (f_1, \ldots, f_{j-1}, f_j, f_{j+1}, \ldots, f_p)(v_1 \otimes \ldots \otimes v_p) + \otimes_*(f_1, \ldots, f_{j-1}, g_j, f_{j+1}, \ldots, f_p)(v_1 \otimes \ldots \otimes v_p)$$

$$= (\alpha \otimes_* (f_1, \ldots, f_{j-1}, f_j, f_{j+1}, \ldots, f_p) + \otimes_*(f_1, \ldots, f_{j-1}, g_j, f_{j+1}, \ldots, f_p))(v_1 \otimes \ldots \otimes v_p)\,.$$

Así, los funcionales

$$\otimes_*(f_1, \ldots, f_{j-1}\,, \alpha\, f_j + g_j\,, f_{j+1}, \ldots, f_p) \quad \text{y}$$

$$\alpha \otimes_* (f_1, \ldots, f_{j-1}, f_j, f_{j+1}, \ldots, f_p) + \otimes_*(f_1, \ldots, f_{j-1}, g_j, f_{j+1}, \ldots, f_p)$$

coinciden en los vectores $v_1 \otimes \ldots \otimes v_p$, para $v_j \in \mathcal{V}_j$, $1 \leq j \leq p$: como este conjunto de vectores es un sistema de generadores, entonces los funcionales son iguales! La igualdad anterior muestra que $\otimes_*$ es lineal como función de su j-ésima entrada (fijando las demás entradas), para cada $1 \leq j \leq p$. Así, $\otimes_*$ es p-lineal, es decir, $\otimes_* \in \operatorname{Hom}(\mathcal{V}_1^*, \ldots, \mathcal{V}_p^*\,; (\otimes_{j=1}^p \mathcal{V}_j)^*)$.

Para $1 \leq j \leq p$ sea $B_j = \{v_1^j, \ldots, v_{d_j}^j\}$ una base de $\mathcal{V}_j$ y $V_j^* = \{f_1^j, \ldots, f_{d_j}^j\}$ su base dual (que es base de $\mathcal{V}_j^*$). Entonces

$$\otimes_*(f_{i_1}^1, \ldots, f_{i_p}^p)(v_{j_1}^1 \otimes \ldots \otimes v_{j_p}^p) = f_{i_1}^1(v_{j_1}^1) \cdots f_{i_p}^p(v_{j_p}^p) = \prod_{h=1}^p f_{i_h}^h(v_{j_h}^h) = \prod_{h=1}^p \delta_{i_h\, j_h}$$

donde hemos considerado las relaciones de dualidad entre los vectores de las bases B_h y B_h^*; lo anterior indica que

$$\otimes_*(f^1_{i_1},\ldots,f^p_{i_p})(v^1_{j_1}\otimes\ldots\otimes v^p_{j_p})=0 \quad \text{si} \quad (i_1,\ldots,i_p)\neq(j_1,\ldots,j_p) \quad \text{y}$$

$$\otimes_*(f^1_{i_1},\ldots,f^p_{i_p})(v^1_{j_1}\otimes\ldots\otimes v^p_{j_p})=1 \quad \text{si} \quad (i_1,\ldots,i_p)=(j_1,\ldots,j_p)$$

Así, si consideramos la base de $\otimes_{j=1}^p\mathcal{V}_j$ dada por

$$B=\{v^1_{i_1}\otimes\ldots\otimes v^p_{i_p}\ :\ 1\le i_j\le d_j\ ,\ 1\le j\le p\}$$

entonces el conjunto

$$B_*=\{\otimes_*(f^1_{i_1},\ldots,f^p_{i_p})\ :\ 1\le i_j\le d_j\ ,\ 1\le j\le p\}\subset(\otimes_{j=1}^p\mathcal{V}_j)^*$$

verifica las relaciones de dualidad con respecto a la base B. Así, $B_*=B^*$ es la base dual de la base B y en particular, es base de $(\otimes_{j=1}^p\mathcal{V}_j)^*$. El Corolario 14.26 (junto con la Observación 14.27) ahora muestran que vale la identidad en la Eq. (115). $\triangle$

Como consecuencia de la Observación 14.29 concluimos las siguientes observaciones

Obs 14.30. Sea $\mathcal{V}_j$ un K-ev para $1\le j\le p$ y $\mathcal{U}$ un K-ev, todos de dimensión finita. Por el Corolario 14.28 podemos identificar

$$\operatorname{Hom}(\mathcal{V}_1,\ldots,\mathcal{V}_p\,;\mathcal{U})=L(\otimes_{j=1}^p\mathcal{V}_j\,,\mathcal{U})$$

a través de un isomorfismo lineal (natural, en el sentido que fue definido sin usar bases). Por otro lado, el Ejemplo 14.18 y la Observación 14.29 indican que

$$\operatorname{Hom}(\mathcal{V}_1,\ldots,\mathcal{V}_p\,;\mathcal{U})=L(\otimes_{j=1}^p\mathcal{V}_j\,,\mathcal{U})=(\otimes_{j=1}^p\mathcal{V}_j)^*\otimes\mathcal{U}=\otimes_{j=1}^p\mathcal{V}_j^*\otimes\mathcal{U}\,.$$

$\triangle$

Obs 14.31. Sean $\mathcal{V},\mathcal{W}$ K-ev's. Entonces el Ejemplo 14.18 indica que $\mathcal{V}^*\otimes\mathcal{W}=L(\mathcal{V}\,,\mathcal{W})$. Aplicando las propiedades de asociatividad y conmutatividad del producto tensorial (ver Teorema 14.21) concluimos que

$$\begin{aligned} L(\mathcal{V})\otimes L(\mathcal{W}) &= (\mathcal{V}^*\otimes\mathcal{V})\otimes(\mathcal{W}^*\otimes\mathcal{W})=\mathcal{V}^*\otimes\mathcal{W}^*\otimes\mathcal{V}\otimes\mathcal{W}\\ &= (\mathcal{V}\otimes\mathcal{W})^*\otimes(\mathcal{V}\otimes\mathcal{W})=L(\mathcal{V}\otimes\mathcal{W})\,, \end{aligned}$$

en donde hemos usado nuevamente el Ejemplo 14.18 en la última igualdad. La identificación anterior permite considerar a los elementos del producto tensorial $L(\mathcal{V})\otimes L(cW)$ como operadores lineales en el espacio $\mathcal{V}\otimes\mathcal{W}$. De hecho, dadas $(S,T)\in L(\mathcal{V})\times L(\mathcal{W})$ definimos $\varphi_{(S,T)}:\mathcal{V}\times\mathcal{W}\to\mathcal{V}\otimes\mathcal{W}$ dada por

$$\varphi_{(S,T)}(v,w)=Sv\otimes Tw\in\mathcal{W}\otimes\mathcal{W}\,.$$

Entonces $\varphi_{(S,T)}$ es 2-lineal (verificar). Por el Teorema 14.14 existe una única transformación $S\otimes T:=T_{\varphi_{(S,T)}}\in L(\mathcal{V}\otimes\mathcal{W})$ tal que

$$S\otimes T(v\otimes w)=Sv\otimes Tw \quad \text{para} \quad (v,w)\in\mathcal{V}\times\mathcal{W}\,.$$

Sean $B_\mathcal{V}=\{v_1,\ldots,v_n\}$ y $B_\mathcal{W}=\{w_1,\ldots,w_m\}$ bases de $\mathcal{V}$ y $\mathcal{W}$; Entonces $B=\{v_i\otimes w_j\ :\ 1\le i\le n\ ,\ 1\le j\le m\}$ es base de $\mathcal{V}\otimes\mathcal{W}$: por un lado, cada operador tiene su matriz

$$[S]_{B_\mathcal{W},B_\mathcal{V}}=A=(a_{ij})_{i,j=1}^n\in K^{n\times n} \quad \text{y} \quad [T]_{B_\mathcal{W},B_\mathcal{V}}=B=(b_{ij})_{i,j=1}^m\in K^{m\times m}\,,$$

construidas como es habitual. Notemos que para poder construir la matriz de $S \otimes T$ con respecto a la base B, debemos ordenar los elementos de la base B : como por construcción los elementos de B están indicados por los pares de la forma (i, j) para $1 \leq i \leq n$, $1 \leq j \leq m$, entonces podemos considerar el orden *lexicográfico* entre ellos. Es decir, dados $1 \leq i,\, h \leq n$, $1 \leq j,\, k \leq m$ entonces

$$(i, j) \leq_{\text{lex}} (h, k) \quad \text{si y solo si} \quad i < h \quad \text{ó} \quad i = h \quad \text{y} \quad j \leq k\,.$$

En este caso, la base B ordenada por el orden lexicográfico es:

$$B = \{v_1 \otimes w_1, \ldots, v_1 \otimes w_m,\, v_2 \otimes w_1, \ldots, v_2 \otimes w_m, v_3 \otimes w_1, \ldots, v_n \otimes w_1, \ldots, v_n \otimes w_m\}\,.$$

Así, para construir la columna h-ésima de $[S \otimes T]_B$, para $1 \leq h \leq m$, debemos calcular:

$$\begin{aligned} S \otimes T(v_1 \otimes w_h) &= Sv_1 \otimes Tw_h = (\sum_{i=1}^{n} a_{i1}\, v_i) \otimes (\sum_{j=1}^{m} b_{jh}\, w_j) \\ &= \sum_{i=1}^{n} \sum_{j=1}^{m} a_{i1}\, b_{jh}\ v_i \otimes w_j\,, \end{aligned}$$

que expresa al valores $S \otimes T(v_1 \otimes w_h)$ como combinación lineal de los elementos de B. Teniendo en cuenta que hemos ordenados los elementos de B con el orden lexicográfico, concluimos que la primer columna de $[S \otimes T]_B$ es

$$(a_{11}\, b_{1h}, \ldots, a_{11}\, b_{mh}, a_{21}\, b_{1h}, \ldots, a_{21}\, b_{mh}, \ldots, a_{n1}\, b_{1h}, \ldots, a_{n1}\, b_{mh})$$

donde tenemos que recordar que si bien hemos escrito al vector como una fila, en realidad es un vector columna. Podemos calcular el segundo grupo de columnas $[S \otimes T]_B$, es decir, la columna h-ésima para $m + 1 \leq h \leq 2m$: en este caso, notemos que el h-ésimo vector de B es $v_2 \otimes w_k$, donde $1 \leq k \leq m$ es tal que $h = m + k$ (verificar!!) Cálculos similares muestran que la h-ésima columna (con $m + 1 \leq h \leq 2m$, es decir, para $1 \leq k \leq m$) es

$$(a_{12}\, b_{1k}, \ldots, a_{12}\, b_{mk}, a_{22}\, b_{1k}, \ldots, a_{22}\, b_{mk}, \ldots, a_{n2}\, b_{1k}, \ldots, a_{n2}\, b_{mk})\,.$$

Podemos continuar calculando el r-ésimo grupo de columnas $(r - 1)m + 1 \leq h \leq rm$ de $[S \otimes T]_B$, para $1 \leq r \leq n$ (notemos que hemos hecho esto en los casos $r = 1$ y $r = 2$) : realizar los cálculos! Si miramos con cuidado la matriz resultante, concluimos que $[S \otimes T]_B$ tiene la siguiente estructura por bloques:

$$[S \otimes T]_B = \begin{pmatrix} a_{11}\, B & \ldots & a_{1n}\, B \\ \vdots & \ldots & \vdots \\ a_{n1}\, B & \ldots & a_{nn}\, B \end{pmatrix} \in K^{nm \times nm} \tag{116}$$

donde el bloque (i, j) es: $a_{ij} \cdot B \in K^{m \times m}$, para $1 \leq i,\, j \leq n$. Esta identidad se verifica notando que las columnas de la matriz a la derecha en la igualdad anterior coinciden con las columnas de $[S{\otimes}T]_B$ que hemos calculado previamente (ejercicio). La matriz por bloques que figura a la derecha de la identidad de la Eq. (116) es llamada *producto de Kronecker* de las matrices A y B , y es notado $A \otimes B$.

A modo de aplicación de la identidad de la Eq. (116), concluimos que si los polinomios de característicos de S y T son

$$p_S(x) = \prod_{h=1}^{n} (x - \lambda_h) \quad \text{y} \quad p_T(x) = \prod_{k=1}^{m} (x - \mu_k) \implies p_{S\otimes T}(x) = \prod_{h=1}^{n} \prod_{k=1}^{m} (x - \lambda_h \cdot \mu_k)\,.$$

En particular, los autovalores de $S \otimes T$ (contando multiplicidades) son $\lambda_h \cdot \mu_k$, para $1 \leq h \leq n$ y $1 \leq k \leq m$. Para verificar la afirmación anterior notemos que, por el Teorema de triangulación de Schur, podemos considerar bases para las cuales A y B son matrices triangulares superiores. En este caso, el producto de Kronecker $A \otimes B$ resulta una matriz triangular superior (verificar); más aún, podemos calcular el polinomio característico de $S \otimes T$ con respecto a esta representación matricial (ejercicio. Sugerencia: calcular con cuidado la diagonal principal de $A \otimes B$ en este caso). $\triangle$

14.3 Álgebra tensorial de un espacio vectorial

Definición 14.32. *Sea $\mathcal{V}$ un K-ev de dimensión finita. Si $p, q \in \mathbb{N}_0$ entonces definimos el espacio de los tensores de tipo (p, q) de $\mathcal{V}$, notado $T^p_q(\mathcal{V})$, dado por:*

1. *Si $p = q = 0$ entonces $T^0_0(\mathcal{V}) = K$;*
2. *Si $q = 0$, ≥ 1:*
$$T^p_0(\mathcal{V}) = \otimes^p_{j=1}\mathcal{V} = \underbrace{\mathcal{V} \otimes \ldots \otimes \mathcal{V}}_{p \text{ veces}}$$
3. *Si $p = 0$ y $q \geq 1$:*
$$T^0_q(\mathcal{V}) = \otimes^q_{j=1}\mathcal{V}^* = \underbrace{\mathcal{V}^* \otimes \ldots \otimes \mathcal{V}^*}_{q \text{ veces}}$$
4. *Si $p \geq 1$ y $q \geq 1$ entonces*
$$T^p_q(\mathcal{V}) = \otimes^p_{j=1}\mathcal{V} \otimes \otimes^q_{k=1}\mathcal{V}^* = \underbrace{\mathcal{V} \otimes \ldots \otimes \mathcal{V}}_{p \text{ veces}} \otimes \underbrace{\mathcal{V}^* \otimes \ldots \otimes \mathcal{V}^*}_{q \text{ veces}}\,.$$

$\square$

Obs 14.33. Sea $\mathcal{V}$ un K-ev de dimensión finita. Entonces, aplicando las propiedades de conmutatividad y asociatividad del producto tensorial, tenemos que: si p, $q \geq 1$ entonces, por ejemplo:
$$T^p_q(\mathcal{V}) = \underbrace{\mathcal{V}^* \otimes \ldots \otimes \mathcal{V}^*}_{\text{q veces}} \otimes \underbrace{\mathcal{V} \otimes \ldots \otimes \mathcal{V}}_{\text{p veces}} \tag{117}$$
ó también
$$T^p_q(\mathcal{V}) = (\otimes^h_{j=1}\mathcal{V}) \otimes (\otimes^k_{j=1}\mathcal{V}^*) \otimes (\otimes^{(p-h)}_{j=1}\mathcal{V}) \otimes (\otimes^{(q-k)}_{j=1}\mathcal{V}^*)$$
donde $1 \leq h < p$ y $1 \leq k < q$. En particular, lo anterior indica que
$$T^p_q(\mathcal{V}) = T^h_k(\mathcal{V}) \otimes T^{p-h}_{q-k}(\mathcal{V})\,.$$
En algunos casos es conveniente recordar estas identificaciones. $\triangle$

Ejemplo 14.34. Sea $\mathcal{V}$ un K-ev de dimensión finita. Entonces:

1. $T^1_0(\mathcal{V}) = \mathcal{V}$; $T^0_1(\mathcal{V}) = \mathcal{V}^*$.
2. En el Ejemplo 14.18 hemos visto que
$$T^1_1(\mathcal{V}) = \mathcal{V} \otimes \mathcal{V}^* = \mathcal{V}^* \otimes \mathcal{V} = L(\mathcal{V})\,.$$
3. En el Ejemplo 14.19 hemos visto que
$$T^0_2(\mathcal{V}) = \mathcal{V}^* \otimes \mathcal{V}^* = \mathrm{Bil}(\mathcal{V}\,,\,\mathcal{V})\,.$$
4. En la Observación 14.31 (ver también la Observación 14.33) hemos visto que
$$T^2_2(\mathcal{V}) = \mathcal{V} \otimes \mathcal{V} \otimes \mathcal{V}^* \otimes \mathcal{V}^* = (\mathcal{V}^* \otimes \mathcal{V}) \otimes (\mathcal{V}^* \otimes \mathcal{V}) = L(\mathcal{V}) \otimes L(\mathcal{V}) = L(\mathcal{V} \otimes \mathcal{V})\,.$$

$\triangle$

Obs 14.35. Sea $\mathcal{V}$ un K-ev, $\dim_K \mathcal{V} = n$. Entonces:

1. $\dim_K T^p_q(\mathcal{V}) = n^{p+q}$;

2. Usando la identificación en la Eq. (117) tenemos que

$$T^p_q(\mathcal{V})^* = (\underbrace{\mathcal{V}^* \otimes \ldots \otimes \mathcal{V}^*}_{\text{q veces}} \otimes \underbrace{\mathcal{V} \otimes \ldots \otimes \mathcal{V}}_{\text{p veces}})^* = (\underbrace{\mathcal{V} \otimes \ldots \otimes \mathcal{V}}_{\text{q veces}} \otimes \underbrace{\mathcal{V}^* \otimes \ldots \otimes \mathcal{V}^*}_{\text{p veces}}) = T^q_p(\mathcal{V})\,.$$

3. Si $B = \{v_j\}_{j=1}^n$ es base de $\mathcal{V}$. En lo que sigue vamos a denotar los elementos de la base dual $B^* = \{f^j\}_{j=1}^n$ con supra-índices. En este caso definimos

$$B^p_q = \{v_{i_1} \otimes \ldots \otimes v_{i_p} \otimes f^{j_1} \otimes \ldots \otimes f^{j_q} : 1 \leq i_\ell,\, j_k \leq n\,,\ 1 \leq \ell \leq p,\, 1 \leq k \leq q\}\,.$$

Como consecuencias de la construcción del espacio producto tensorial, sabemos que B^p_q es base de $T^p_q(\mathcal{V})$

En este caso, dado $T \in T^p_q(\mathcal{V})$ entonces T admite un desarrollo único como combinación lineal de los elementos de la base B^p_q; los coeficientes de la combinación lineal (que son únicos) corresponden a las coordenadas de T con respecto a la base B^p_q. Notemos que para escribir la correspondiente combinación lineal, hacen falta $(p+q)$ símbolos de sumatoria!! Para evitar la escritura (innecesaria) de esos símbolos seguimos *la convención de Albert Einstein*: denotamos las coordenadas de T mediante las expresiones

$$T^{i_1,\ldots,i_p}_{j_1,\ldots,j_q}(B) \in K \quad \text{para} \quad 1 \leq i_\ell,\, j_k \leq n\,,\ 1 \leq \ell \leq p,\, 1 \leq k \leq q$$

y la combinación lineal correspondiente queda escrita como

$$T^{i_1,\ldots,i_p}_{j_1,\ldots,j_q}(B)\ v_{i_1} \otimes \ldots \otimes v_{i_p} \otimes f^{j_1} \otimes \ldots \otimes f^{j_q}$$

donde los símbolos de sumatoria están implícitos, porque hemos seguido la siguiente convención: siempre que en una expresión como la anterior aparezca una variable de sumación (por ejemplo i_1) tanto como super-índice como sub-índice, entonces convenimos que se está sumando sobre el índice en cuestión: en el caso de i_1, aparece como super-índice en $T^{\mathbf{i_1},\ldots,i_q}_{j_1,\ldots,j_p}$ y aparece como sub-índice en la expresión $v_{\mathbf{i_1}} \otimes \ldots \otimes v_{i_p} \otimes f^{j_1} \otimes \ldots \otimes f^{j_q}$. De hecho, lo mismo sucede con cada uno de los índices $i_1, \ldots i_p, j_1, \ldots, j_q$ de forma que en la expresión anterior hay implícitas $(p+q)$ símbolos de sumanción.

Finalmente, notemos que la base dual $(B^p_q)^* = B^q_p$, en términos de la identificación de $T^p_q(\mathcal{V})^* = T^q_p(\mathcal{V})$ dada en el ítem 2.

$\triangle$

14.3.1 Operaciones entre tensores de tipo (p, q) y sus coordenadas.

En esta sección introducimos una serie de operaciones que se pueden realizar entre los espacios de tensores del tipo $T^p_q(\mathcal{V})$ para un K-ev de dimensión finita. Además, veremos como calcular las coordendas correspondientes a cada operación.

Obs 14.36. Sea $\mathcal{V}$ un K-ev de dimensión finita. Entonces $T^p_q(\mathcal{V})$ es un K-ev, donde la estructura de este espacio corresponde a la estructura vectorial del producto tensorial de K-ev's, como ha sido descrita en la Definición 14.23. En particular, si T_1, $T_2 \in T^p_q(\mathcal{V})$ y $\alpha \in K$ entonces podemos formar la combinación lineal: $\alpha\, T_1 + T_2 \in T^p_q(\mathcal{V})$.

Si $B=\{v_1,\dots,v_n\}$ es base de $\mathcal{V}$ con base dual $B^*=\{f_1,\dots,f_n\}$ de $\mathcal{V}^*$ entonces, si las coordenadas de T_h con respecto a B_q^p son

$$(T_h)^{i_1,\dots,i_p}_{j_1,\dots,j_q}\ :\ 1\le i_\ell,\, j_k\le n\,,\ 1\le \ell\le p,\, 1\le k\le q \quad \text{y} \quad h=1,\,2$$

entonces las coordenadas de $\alpha\, T_1+T_2$ con respecto a B_q^p son

$$(\alpha\, T_1+T_2)^{i_1,\dots,i_p}_{j_1,\dots,j_q}=\alpha\,(T_1)^{i_1,\dots,i_p}_{j_1,\dots,j_q}+(T_2)^{i_1,\dots,i_p}_{j_1,\dots,j_q}\ :\ 1\le i_\ell,\, j_k\le n\,,\ 1\le \ell\le p,\, 1\le k\le q$$

porque tomar coordenadas con respecto a una base es una transformación lineal!

△

Definición 14.37 (Producto tensorial entre tensores de tipo (p,q))**.** *Sea $\mathcal{V}$ un K-ev de dimensión finita. Si $p,q,r,s\in\mathbb{N}_0$ entonces podemos construir: $\otimes\in Hom(T_q^p(\mathcal{V})\,,\,T_s^r(\mathcal{V});\,T_{q+s}^{p+r}(\mathcal{V}))$ como sigue: comenzamos definiendo*

$$\Phi\in Hom(T_q^p(\mathcal{V})\,,\,T_s^r(\mathcal{V});\,T_q^p(\mathcal{V})\otimes T_s^r(\mathcal{V}))$$

como la transformación 2-lineal dada por

$$\Phi(\otimes_{i=1}^p v_j\ \otimes\ \otimes_{j=1}^q f_j\,,\,\otimes_{h=1}^r u_h\ \otimes\ \otimes_{k=1}^s g_k)=(\otimes_{i=1}^p v_j\ \otimes\ \otimes_{j=1}^q f_j)\ \otimes\ (\otimes_{h=1}^r u_h\ \otimes\ \otimes_{k=1}^s g_k)\,.$$

En términos de la identificación (ver Observación 14.33)

$$T_q^p(\mathcal{V})\otimes T_s^r(\mathcal{V})=T_{q+s}^{p+r}(\mathcal{V})$$

definimos $\otimes\in Hom(T_q^p(\mathcal{V})\,,\,T_s^r(\mathcal{V})\,;\,T_{q+s}^{p+r}(\mathcal{V}))$ determinado en los tensores elementales por

$$\otimes(\otimes_{i=1}^p v_j\ \otimes\ \otimes_{j=1}^q f_j\,,\,\otimes_{h=1}^r u_h\ \otimes\ \otimes_{k=1}^s g_k)=(\otimes_{i=1}^p v_j\ \otimes\ \otimes_{h=1}^r u_h)\ \otimes\ (\otimes_{j=1}^q f_j\ \otimes\ \otimes_{k=1}^s g_k)\in T_{q+s}^{p+r}(\mathcal{V})$$

y en este caso al valor $\otimes((\otimes_{i=1}^p v_j\ \otimes\ \otimes_{j=1}^q f_j)\,,\,(\otimes_{h=1}^r u_h\ \otimes\ \otimes_{k=1}^s g_k))$ lo notamos

$$\underbrace{(\otimes_{i=1}^p v_j\ \otimes\ \otimes_{j=1}^q f_j)}_{\in T_q^p(\mathcal{V})}\otimes\underbrace{(\otimes_{h=1}^r u_h\ \otimes\ \otimes_{k=1}^s g_k)}_{\in T_s^r(\mathcal{V})}\in T_{q+s}^{p+r}(\mathcal{V})$$

y lo llamamos el **producto tensorial** *de $\otimes_{i=1}^p v_j\ \otimes\ \otimes_{j=1}^q f_j$ y $\otimes_{h=1}^r u_h\ \otimes\ \otimes_{k=1}^s g_k$.*

□

Obs 14.38. Sea $\mathcal{V}$ un K-ev de dimensión finita. Consideremos el producto tensorial de los espacios $T_q^p(\mathcal{V})$ y $T_s^r(\mathcal{V})$ como en la Definición 14.37. En este caso, consideramos la transformación $\otimes\in \mathrm{Hom}(T_q^p(\mathcal{V})\,,\,T_s^r(\mathcal{V})\,;\,T_{q+s}^{p+r}(\mathcal{V}))$ determinado en los tensores elementales por

$$(\otimes_{i=1}^p v_j)\ \otimes\ (\otimes_{j=1}^q f_j)\otimes(\otimes_{h=1}^r u_h)\ \otimes\ (\otimes_{k=1}^s g_k)=(\otimes_{i=1}^p v_j\ \otimes\ \otimes_{h=1}^r u_h)\ \otimes\ (\otimes_{j=1}^q f_j\ \otimes\ \otimes_{k=1}^s g_k)$$

donde la expresión de arriba está en $T_{q+s}^{p+r}(\mathcal{V})$.

Notemos que el producto tensorial se extiende al producto de tensores arbitrarios (es decir, posiblemente no elementales) $T_1\in T_q^p(\mathcal{V})$ y $T_2\in T_s^r(\mathcal{V})$ de forma que

$$T_1\otimes T_2\in T_{q+s}^{p+r}(\mathcal{V})\,.$$

Sea $B=\{v_1,\dots,v_n\}$ una base $\mathcal{V}$ con base dual $B^*=\{f_1,\dots,f_n\}$. Supongamos que tenemos las representaciones lineales

$$T_1=(T_1)^{i_1,\,\dots,i_p}_{j_1,\dots,j_q}\ v_{i_1}\otimes\ldots\otimes v_{i_p}\otimes f^{j_1}\otimes\ldots\otimes f^{j_q}$$

$$T_2 = (T_2)^{k_1,\ldots,k_r}_{\ell_1,\ldots,\ell_s}\ v_{k_1}\otimes\ldots\otimes v_{k_r}\otimes f^{\ell_1}\otimes\ldots\otimes f^{\ell_s}$$

Entonces, usando la 2-linealidad del producto tensorial $\otimes : T^p_q(\mathcal{V})\times T^r_s(\mathcal{V})\to T^{p+r}_{q+s}(\mathcal{V})$ tenemos que el producto tensorial $T_1\otimes T_2$ se puede representar como:

$$\Big((T_1)^{i_1,\ldots,i_p}_{j_1,\ldots,j_q}\ v_{i_1}\otimes\ldots\otimes v_{i_p}\otimes f^{j_1}\otimes\ldots\otimes f^{j_q}\Big)\otimes\Big((T_2)^{k_1,\ldots,k_r}_{\ell_1,\ldots,\ell_s}\ v_{k_1}\otimes\ldots\otimes v_{k_r}\otimes f^{\ell_1}\otimes\ldots\otimes f^{\ell_s}\Big)$$

$$= (T_1)^{i_1,\ldots,i_p}_{j_1,\ldots,j_q}\ (T_2)^{k_1,\ldots,k_r}_{\ell_1,\ldots,\ell_s}\ v_{i_1}\otimes\ldots\otimes v_{i_p}\otimes v_{k_1}\otimes\ldots\otimes v_{k_r}\ \otimes\ f^{j_1}\otimes\ldots\otimes f^{j_q}\otimes f^{\ell_1}\otimes\ldots\otimes f^{\ell_s}$$

Pero notemos que los elementos de la base B^{p+r}_{q+s} de T^{p+r}_{q+s} son justamente los tensores elementales de la forma:

$$v_{i_1}\otimes\ldots\otimes v_{i_p}\otimes v_{k_1}\otimes\ldots\otimes v_{k_r}\ \otimes\ f^{j_1}\otimes\ldots\otimes f^{j_q}\otimes f^{\ell_1}\otimes\ldots\otimes f^{\ell_s}$$

para $1\leq i_1,\ldots,i_p,k_1,\ldots,k_r\leq n$ y $1\leq j_1,\ldots,j_q,\ell_1,\ldots,\ \ell_s\leq n$. Así, las identidades anteriores prueban que las coordenadas de $T_1\otimes T_2$ con respecto a B^{p+r}_{q+s} son

$$(T_1\otimes T_2)^{i_1,\ldots,i_p,k_1,\ldots,k_r}_{j_1,\ldots,j_q,\ell_1,\ldots,\ell_s}(B) = (T_1)^{i_1,\ldots,i_p}_{j_1,\ldots,j_q}(B)\cdot(T_2)^{k_1,\ldots,k_r}_{\ell_1,\ldots,\ell_s}(B)\in K$$

para $1\leq i_1,\ldots,i_p,k_1,\ldots,k_r\leq n$ y $1\leq j_1,\ldots,j_q,\ell_1,\ldots,\ \ell_s\leq n$. △

La siguiente operación entre tensores permite modelar las distintas operaciones de evaluación que hemos considerado (y otro tipo de construcciones más generales).

Definición 14.39 (Contracción de tensores de tipo (p,q))**.** *Sea $\mathcal{V}$ un K-ev de dimensión finita y sean $1\leq i\leq p$ y $1\leq j\leq q$. Consideramos* $\Psi^i_j : \underbrace{\mathcal{V}\times\ldots\times\mathcal{V}}_{p\ veces}\otimes\underbrace{\mathcal{V}^*\times\ldots\times\mathcal{V}^*}_{q\ veces}\to T^{p-1}_{q-1}(\mathcal{V})$

$$\Psi^i_j(v_1,\ldots,v_p,f_1,\ldots,f_q) = \underbrace{f_j(v_i)}_{\in K}\ v_1\otimes\ldots\otimes v_{i-1}\otimes v_{i+1}\otimes\ldots\otimes v_p\otimes f_1\otimes\ldots\otimes f_{j-1}\otimes f_{j+1}\otimes\ldots\otimes f_q\,.$$

Entonces Ψ^i_j es una transformación $(p{+}q)$-multilineal, de forma que existe una única transformación lineal $C^i_j : T^p_q(\mathcal{V})\to T^{p-1}_{q-1}(\mathcal{V})$ determinada por su acción sobre los tensores elmentales:

$$C^i_j(v_1\otimes\ldots\otimes v_p\otimes f_1\otimes\ldots\otimes f_q) = f_j(v_i)\ v_1\otimes\ldots\otimes v_{i-1}\otimes v_{i+1}\otimes\ldots\otimes v_p\otimes f_1\otimes\ldots\otimes f_{j-1}\otimes f_{j+1}\otimes\ldots\otimes f_q$$

llamada **contracción por los índices** (i,j). □

Ejemplo 14.40. Sea $\mathcal{V}$ un K-ev de dimensión finita. En este caso podemos considerar $C^1_1 : T^1_1(\mathcal{V}) = \mathcal{V}\otimes\mathcal{V}^*\to T^0_0(\mathcal{V}) = K$ dada por $C^1_1(v\otimes f) = f(v)\in K$, para $v\otimes f\in\mathcal{V}\otimes\mathcal{V}^*$. Esta propiedad determina de forma única la contracción por los índices $(1,1)$ en todo el espacio $\mathcal{V}\otimes\mathcal{V}^*$.

Consideremos la identificación $T^1_1(\mathcal{V}) = L(\mathcal{V})$ desarrollada en el Ejemplo 14.18. En este caso, consiramos al tensor elemental como una transformación lineal $v\otimes f(w) = f(w)\,v$ para $w\in\mathcal{V}$. Consideramos una base $B = \{v_1,\ldots,v_n\}$ de $\mathcal{V}$; para calcular la matriz de la transformación $v\otimes f$ consideramos

$$v = \sum_{i=1}^n \alpha_i\,v_i \implies v\otimes f(v_j) = \sum_{i=1}^n f(v_j)\alpha_i\,v_i$$

que muestra que

$$[v\otimes f]_B = \begin{pmatrix} f(v_1)\,\alpha_1 & \ldots & f(v_n)\,\alpha_1\\ \vdots & \vdots & \vdots\\ f(v_1)\,\alpha_n & \ldots & f(v_n)\,\alpha_n\end{pmatrix}$$

que implica que

$$\operatorname{tr}(v \otimes f) = \operatorname{tr}([v \otimes f]_B) = \sum_{j=1}^{n} f(v_j)\,\alpha_j = f(\sum_{j=1}^{n} \alpha_j\, v_j) = f(v) = C_1^1(v \otimes f)\,.$$

Es decir, bajo la identificación $T_1^1(\mathcal{V}) = L(\mathcal{V})$ entonces la contracción $(1,1)$ coincide con tomar la traza del operador en los tensores elementales. Notemos que como la contracción y la traza son funcionales lineales (en este caso) y como los tensores elementales general a $T_1^1(\mathcal{V})$ entonces, la contracción y la traza deben coincidir en todo tensor de $T_1^1(\mathcal{V})$ (no solo en los tensores elementales). $\triangle$

Ejemplo 14.41. Sea $\mathcal{V}$ un K-ev de dimensión finita. En este caso podemos considerar $C_1^2 : T_1^2(\mathcal{V}) \to T_0^1(\mathcal{V}) = \mathcal{V}$ determinada por $C_1^2(v \otimes w \otimes f) = f(w)\,v$ en los tensores elementales.

Identificamos $\mathcal{V}^* \otimes \mathcal{V}$ con $L(\mathcal{V})$ de forma que podemos identificar $T_1^2(\mathcal{V}) = \mathcal{V}^* \otimes \mathcal{V} \otimes \mathcal{V} = L(\mathcal{V}) \otimes \mathcal{V}$, tal que $T_1^2(\mathcal{V}) \ni v \otimes w \otimes f \approx (f \otimes v) \otimes w \in L(\mathcal{V}) \otimes \mathcal{V}$ donde $f \otimes v(u) = f(u)\,v$, para $u \in \mathcal{V}$.

Así, el valor $C_1^2(v \otimes w \otimes f) = C_1^2((f \otimes v) \otimes w) = f(w)\,v$ coincide con el valor de la transformación $f \otimes v$ en el vector $w \in \mathcal{V}$, donde hemos utilizado las identificaciones descriptas en el párrafo anterior. Así, en este caso la contracción por los índices $(2,1)$ corresponde a la evaluación de una transformación lineal en un vector en los tensores elementales. Notemos que como la contracción y la evaluación son transformaciones lineales, y como los tensores elementales general a $T_2^2(\mathcal{V})$ entonces, la contracción y la evaluación deben coincidir en todo tensor de $T_2^2(\mathcal{V})$ (no solo en los tensores elementales). $\triangle$

Ejemplo 14.42. Sea $\mathcal{V}$ un K-ev de dimensión finita. En este caso podemos considerar $C_1^2 : T_2^2(\mathcal{V}) \to T_1^1(\mathcal{V}) = \mathcal{V} \otimes \mathcal{V}^*$ determinada por $C_1^2(v \otimes w \otimes f \otimes g) = f(w)\,v \otimes g$ en los tensores elementales.

Identificamos $T_2^2(\mathcal{V}) = \mathcal{V}^* \otimes \mathcal{V} \otimes \mathcal{V}^* \otimes \mathcal{V} = L(\mathcal{V}) \otimes L(\mathcal{V})$ de forma que $T_2^2(\mathcal{V}) \ni v \otimes w \otimes f \otimes g \approx (f \otimes v) \otimes g \otimes w \in L(\mathcal{V}) \otimes L(\mathcal{V})$ donde $f \otimes v(u) = f(u)\,v$ y $g \otimes w(u) = g(u)\,w$, para $u \in \mathcal{V}$. Bajo esta identificación se tiene que el producto (la composición) de los operadores

$$(f \otimes v) \cdot (g \otimes w)(u) = f \otimes v(g(u)\,w) = f(g(u)\,w)\,v = f(w)\,g(u)\,v = (f(w)\ g \otimes v)(u)$$

que muestra la identidad: $(f \otimes v) \cdot (g \otimes w) = f(w)\ g \otimes v$.

Así, el valor $C_1^2(v \otimes w \otimes f \otimes g) = C_1^2((f \otimes v) \otimes (g \otimes w)) = f(w)\,v$ coincide con el producto de los operadores $f \otimes v$ y $g \otimes w$, donde hemos utilizado las identificaciones descriptas en el párrafo anterior. Así, en este caso la contracción por los índices $(2,1)$ corresponde al producto de operadores lineales, en los tensores elementales. Consideraciones similares a las hechas en los ejemplos anteriores permiten ver que esta identificación se extiende a todos tensores en $T_2^2(\mathcal{V})$. $\triangle$

Obs 14.43. Sea $\mathcal{V}$ un K-ev de dimensión finita y sean $1 \le h \le p$ y $1 \le k \le q$. Consideramos la contracción de los índices (h,k) es decir, la transformación lineal $C_k^h : T_q^p(\mathcal{V}) \to T_{q-1}^{p-1}(\mathcal{V})$ tal que su acción en los tensores elementales está dada por:

$$C_k^h(v_1 \otimes \ldots \otimes v_p \otimes f_1 \otimes \ldots \otimes f_q) = f_k(v_h)\; v_1 \otimes \ldots \otimes v_{h-1} \otimes v_{h+1} \otimes \ldots \otimes v_p \otimes f_1 \otimes \ldots \otimes f_{k-1} \otimes f_{k+1} \otimes \ldots \otimes f_q\,.$$

En este caso, C_j^i se extiende (por linealidad) a todo el espacio de tensores elementales. Sea $T \in T_q^p(\mathcal{V})$; si consideramos una base $B = \{v_1, \ldots, v_n\}$ de $\mathcal{V}$ con base dual $B^* = \{f_1, \ldots, f_n\}$ de $\mathcal{V}^*$ entonces se tiene una representación única de T con respecto a la base B_q^p de la forma

$$T = T_{j_1,\ldots,j_q}^{i_1,\ldots,i_p}\ v_{i_1} \otimes \ldots \otimes v_{i_p} \otimes f^{j_1} \otimes \ldots \otimes f^{j_q}$$

En este caso, usando la linealidad de C^h_k y la representación anterior tenemos que

$$C^h_k(T) = T^{i_1,\ldots,i_p}_{j_1,\ldots,j_q}\ C^h_k(v_{i_1}\otimes\ldots\otimes v_{i_p}\otimes f^{j_1}\otimes\ldots\otimes f^{j_q}) =$$

$$\sum_{i_h=1}^n\sum_{j_k=1}^n T^{i_1,\ldots,i_p}_{j_1,\ldots,j_q}\ f_{j_k}(v_{i_h})\ v_{i_1}\otimes\ldots\otimes v_{i_{h-1}}\otimes v_{i_{h+1}}\otimes\ldots\otimes v_{i_p}\otimes f^{j_1}\otimes\ldots\otimes f^{j_{k-1}}\otimes f^{j_{k+1}}\otimes\ldots\otimes f^{j_q}$$

$$=\sum_{\ell=1}^n T^{i_1,\ldots,i_{h_1},\ell,i_{h+1},\ldots,i_p}_{j_1,\ldots,j_{k_1},\ell,j_{k+1},\ldots,j_q}\ v_{i_1}\otimes\ldots\otimes v_{i_{h-1}}\otimes v_{i_{h+1}}\otimes\ldots\otimes v_{i_p}\otimes f^{j_1}\otimes\ldots\otimes f^{j_{k-1}}\otimes f^{j_{k+1}}\otimes\ldots\otimes f^{j_q}$$

$$= T^{i_1,\ldots,i_{h_1},\ell,i_{h+1},\ldots,i_p}_{j_1,\ldots,j_{k_1},\ell,j_{k+1},\ldots,j_q}\ v_{i_1}\otimes\ldots\otimes v_{i_{h-1}}\otimes v_{i_{h+1}}\otimes\ldots\otimes v_{i_p}\otimes f^{j_1}\otimes\ldots\otimes f^{j_{k-1}}\otimes f^{j_{k+1}}\otimes\ldots\otimes f^{j_q}$$

donde hemos seguido la convención de Einstein, salvo para los índices de sumación de i_h y j_k (para los cuales hemos escrito explícitamente la sumatoria); además, hemos usado las relaciones de dualidad $f_{j_k}(v_{i_h}) = \delta_{i_h,j_k}$: esto permite reemplazar la doble sumatoria por una sumatoria simple, correspondiente a los índices $i_h = j_k$. Más aún, hemos renombrado en este caso $i_h = j_k = \ell$ introduciendo un nuevo índice de sumanción (para difereciarlo del resto); finalmente, hemos hecho uso de la convención de Einstein en la última identidad.

Finalmente, notemos que la base $B^{p-1}_{q_1}$ de $T^{p-1}_{q_1}(\mathcal{V})$ se puede describir como el conjunto de los vectores de la forma :

$$v_{i_1}\otimes\ldots\otimes v_{i_{h-1}}\otimes v_{i_{h+1}}\otimes\ldots\otimes v_{i_p}\otimes f^{j_1}\otimes\ldots\otimes f^{j_{k-1}}\otimes f^{j_{k+1}}\otimes\ldots\otimes f^{j_q})$$

con $1\le i_1,\ldots,i_{h-1},i_{h+1},\ldots,i_p\le n$ y $1\le j_1,\ldots,j_{k-1},j_{k+1},\ldots,j_q\le n$. Entonces las identidades anteriores nos permiten calcular las coordenadas de $C^h_k(T)$ con respecto a la base $B^{p-1}_{q_1}$ como:

$$(C^h_k(T))^{i_1,\ldots,i_{h-1},i_{h+1},\ldots,i_p}_{j_1,\ldots,j_{k-1},j_{k+1},\ldots,j_q}(B) = T^{i_1,\ldots,i_{h-1},\ell,i_{h+1},\ldots,i_p}_{j_1,\ldots,j_{k-1},\ell,j_{k+1},\ldots,j_q}(B)\,,$$

donde estamos usando la convención de Einstein (en este caso, para omitir la sumatoria en el índice ℓ). △

14.3.2 Cambio de coordenadas de tensores de tipo (p, q)

En esta sección desarrollamos la fórmula del cambio de base para tensores de tipo (p, q) con respecto a bases del tipo descripto en el ítem 3. de la Observación 14.35. Para ello, comenzamos con la siguiente:

Obs 14.44. Sea $\mathcal{V}$ un K-ev, $\dim_K\mathcal{V} = n$. Sean $B_1 = \{v_1,\ldots,v_n\}$ y $B_2 = \{w_1,\ldots,w_n\}$ bases de $\mathcal{V}$. Consideramos las respectivas bases duales $B_1^* = \{f^1,\ldots,f^n\}$ y $B_2^* = \{g^1,\ldots,g^n\}$ de $\mathcal{V}^*$. Sea M_{B_2,B_1} la matriz de cambio de base de la base B_1 en la base B_2. En lo que sigue notamos las entradas de M como a^h_k, $1\le h,\, k\le n$: es decir, $M_{B_2,B_1} = (a^h_k)^n_{h,k=1}$. Por construcción de M_{B_2,B_1}, sabemos que

$$v_i = \sum_{h=1}^n a^h_i\,w_h = a^h_i\,w_h \quad \text{para} \quad 1\le j\le n\,, \tag{118}$$

pues la i-ésima columna de M es el vector $[v_i]_{B_2}\in K^n$. Notemos que hemos introducido la nueva notación para los coeficientes de M para poder aplicar la convención de Einstein (en la identidad $\sum_{h=1}^n a^h_i\,w_h = a^h_i\,w_h$) ver el ítem 3 de la Observación 14.35.

De forma similar, si $M_{B_1,B_2} = (b^i_j)^n_{i,j=1}$ entonces

$$w_j = \sum_{i=1}^n b^i_j\,v_i = b^i_j\,v_i \quad \text{para} \quad 1\le j\le n\,,$$

Además, por las propiedades de la base dual B_2^* se tiene que

$$f^j = \sum_{i=1}^{n} f^j(w_i)\, g^i = \sum_{i=1}^{n} f^j(\sum_{\ell=1}^{n} b_i^\ell v_\ell)\, g^i = \sum_{i=1}^{n}\sum_{\ell=1}^{n} b_i^\ell\, f^j(v_\ell)\, g^i = \sum_{i=1}^{n} b_i^j\, g^i = b_i^j\, g^i \tag{119}$$

donde hemos usado las relaciones de dualidad $f^j(v_\ell) = \delta_{j\,\ell}$ y la convención de Einstein en la última identidad. Así, $[f^j]_{B_2^*} = (b_1^j, \ldots, b_n^j)$ que corresponde a la j-ésima **fila** de la matriz cambio de base $M_{B_1,B_2} = (b_j^i)_{i,j=1}^n$. Pero por otro lado, $[f^j]_{B_2^*} = (b_1^j, \ldots, b_n^j)$ corresponde a la j-ésima columna de la matriz de cambio de base $M_{B_2^*,B_1^*}$ por construcción. Así, se tiene la relación

$$(b_j^i)_{i,j=1}^n = M_{B_2^*,B_1^*} = M_{B_1,B_2}^t\,.$$

△

Proposición 14.45. *Sea $\mathcal{V}$ un K-ev y sean $B_1 = \{v_1, \ldots, v_n\}$ y $B_2 = \{w_1, \ldots, w_n\}$ bases de $\mathcal{V}$, con bases duales $B_1^* = \{f^1, \ldots, f^n\}$ y $B_2^* = \{g^1, \ldots, g^n\}$ de $\mathcal{V}^*$. Si $M_{B_2,B_1} = (a_j^i)_{i,j=1}^n$ y $M_{B_1,B_2}^t = (b_j^i)_{i,j=1}^n$ entonces: dado un tensor $T \in T_q^p(\mathcal{V})$ con coordenadas con respecto a la base $(B_1)_q^p$ dadas por*

$$\{T_{j_1,\ldots,j_q}^{i_1,\ldots,i_p}(B_1) : 1 \le i_\ell,\, j_k \le n\,,\ 1 \le \ell \le p,\, 1 \le k \le q\}$$

y coordenadas con respecto a la base $(B_2)_q^p$ dadas por

$$\{T_{j_1,\ldots,j_q}^{i_1,\ldots,i_p}(B_2) : 1 \le i_\ell,\, j_k \le n\,,\ 1 \le \ell \le p,\, 1 \le k \le q\}$$

entonces vale la fórmula de cambio de base para las coordenadas:

$$T_{k_1,\ldots,k_q}^{h_1,\ldots,h_p}(B_2) = b_{i_1}^{h_1} \cdots b_{i_p}^{h_p} \cdot T_{j_1,\ldots,j_q}^{i_1,\ldots,i_p}(B_1) \cdot a_{k_1}^{j_1} \cdots a_{k_q}^{j_p}\,,$$

donde estamos usando la convención de Einstein (y hemos omitido los $(p+q)$ símbolos de sumación correspondientes a las variables $i_1, \ldots, i_p, j_1, \ldots, j_q$).

Demostración. La prueba de este resultado es una consecuencia de la unicidad de las coordenadas con respecto a una base (en este caso $(B_2)_q^p$ de $T_q^p(\mathcal{V})$) y de las Eqs. (118) y (119). En efecto, basta notar que, por la multilinealidad del producto tensorial:

$$\begin{aligned}
T &= T_{j_1,\ldots,j_q}^{i_1,\ldots,i_p}(B)\ v_{i_1} \otimes \ldots \otimes v_{i_p} \otimes f^{j_1} \otimes \ldots \otimes f^{j_q} \\
&= T_{j_1,\ldots,j_q}^{i_1,\ldots,i_p}(B)\ a_{i_1}^{h_1} w_{h_1} \otimes \ldots \otimes a_{i_p}^{h_p} w_{h_p} \otimes b_{k_1}^{j_1} g^{k_1} \otimes \ldots \otimes b_{k_q}^{j_q} g^{k_q} \\
&= b_{k_1}^{j_1} \cdots b_{k_q}^{j_q} \cdot T_{j_1,\ldots,j_q}^{i_1,\ldots,i_p}(B) \cdot a_{i_1}^{h_1} \cdots a_{i_p}^{h_p}\ w_{h_1} \otimes \ldots \otimes w_{h_p} \otimes g^{k_1} \otimes \ldots \otimes g^{k_q}
\end{aligned}$$

Finalmente, los coeficientes que acompañan a los vectores de la base $(B_2)_q^p$ en la expresión anterior deben ser las coordenadas de T con respecto a esa base; estas observaciones prueban el enunciado.

□