1. Espacios vectoriales

Def. Un **espacio vectorial** sobre un campo F consiste de un conjunto V junto con dos operaciones una interna denominada suma y otra externa denominada multiplicación escalar de manera que se cumplen las siguientes condiciones:

1. Para todo u,v en V, $u + v = v + u$.
2. Para todo u,v,w en V $(u + v) + w = u + (v + w)$.
3. Existe un elemento en V, denotado por *0*, tal que para todo u en V, $u + 0 = u$.
4. Para todo elemento u en V existe otro elemento v también en V tal que $u + v = 0$.
5. Para todo elemento u en V, $1u = u$.
6. Para cualquier par de elementos a, b en F y cualquier elemento u en V, $(ab)u = a(bu)$.
7. Para cualquier elemento a en F y cualquier par de elementos u, v en V, $a(u + v) = au + av$.
8. Para cualquier par de elementos a, b en F y cualquier elemento u en V, $(a + b)u = au + bu$.

Los elementos de un espacio vectorial se denominan vectores y los elementos de un campo escalares.

Propiedades Adicionales

Teorema 1.1 (Ley de cancelación)**.** *Si u, v y w son elementos de un espacio vectorial V tal que $u + w = v + w$, entonces $u = v$.*

Demostración. Existe un vector x en V tal que $w + x = 0$ (propiedad 4). Entonces,

$$u = u + 0 = u + (w + x) = (u + w) + x = (v + w) + x = v + (w + x) = v + 0 = v$$

□

Corolario 1.1.1. *El vector 0 descrito en la propiedad 3 es único.*

Demostración. Supongamos que existen dos vectores *0* y *0'* como el descrito en la propiedad 3. Entonces, tenemos que $v + 0 = v$ y que $v + 0' = v$ De modo que, $v + 0 = v + 0'$. Por la ley de cancelación tenemos entonces que $0 = 0'$. □

Corolario 1.1.2. *El vector v descrito en propiedad 4 es único*

Demostración. Análoga a la demostración del corolario anterior. □

Por este corolario, podemos utilizar, sin riesgo de ambigüedad, $-v$ para denotar el vector descrito en la propiedad 4.

Teorema 1.2. *En todo espacio vectorial V, lo siguiente siempre es verdadero:*

(a) $0v = 0$ *para todo* $v \in V$

(b) $(-a)v = -(av) = a(-v)$ *para todo* $a \in F$ *y todo* $v \in V$.

(c) $a0 = 0$ *para todo* $a \in F$

1.1. Subespacios Vectoriales

Def. Un subconjunto W de un espacio vectorial V sobre un campo F se llama **subespacio** de V si W es un espacio vectorial sobre F con las operaciones de suma y multiplicación escalar definidas en V.

Para probar que un subconjunto es un subespacio no hace falta demostrar que se cumplen todas las propiedades de un espacio vectorial. Basta con probar tres condiciones.

Teorema 1.3. *Sean V un espacio vectorial y W un subconjunto de V. Entonces W es subespacio de V si y solo si las siguientes condiciones se cumplen para las operaciones definidas en V:*

(*a*) *$u + v \in W$ para cualesquiera dos elementos $u, v \in W$ (Clausura bajo sumas).*

(*b*) *$cv \in W$ para cualesquiera $c \in F$ y $v \in W$ (Clausura bajo multiplicación escalar).*

(*c*) *W tiene vector 0.*

Teorema 1.4. *La intersección de subespacios de V es un subespacio de V.*

Demostración. Sea C una colección de subespacios de V y sea W la intersección de los subespacios en C. Dado que todo subespacio contiene al vector *0*, $0 \in W$. Sean $a \in F$ y $u, v \in W$. Entonces u, v están contenidos en todo subespacio en C. Como cada subespacio es cerrado bajo sumas y bajo múltiplos, se cumple que $u + v$ y au pertenecen a todo subespacio en C de modo que au y $u+v$ también pertenecen a W. Esto verifica las tres condiciones enunciadas en el teorema anterior, luego W es subespacio es subespacio de V. □

Por otra parte, la unión de subespacios no siempre es subespacio. De hecho, la unión de subespacios sólo es subespacio cuando uno está contenido dentro del otro. De lo contrario, no se cumple que la unión sea cerrada bajo sumas.

Teorema 1.5. *Sean W_1 y W_2 subespacios de un espacio vectorial V. $W_1 \cup W_2$ es subespacio vectorial si y solo si $W_1 \subseteq W_2$ ó $W_2 \subseteq W_1$.*

Demostración. Por doble implicación:

($\rightarrow$) Supongamos que $W_1 \nsubseteq W_2$ y $W_2 \nsubseteq W_1$, entonces existe un u_1 que pertenece a W_1 pero no a W_2, y existe un u_2 que pertenece a W_2 pero no a W_1. Sea $v = u_1 + u_2$. Ahora supongamos que $W_1 \cup W_2$ es subespacio. Como $u_1, u_2 \in W_1 \cup W_2$, $v \in W_1 \cup W_2$. Esto implica que $v \in W_1$ ó $v \in W_2$ (o las dos al tiempo).

- Si $v \in W_1$ entonces $v - u_1 \in W_1$, pero $v - u_1 = u_2$, de modo que $u_2 \in W_1$. Contradicción.
- Si $v \in W_2$ entonces $v - u_2 \in W_2$, pero $v - u_2 = u_1$, de modo que $u_1 \in W_2$. Contradicción.

Como v no pertenece a W_1 ni a W_2, entonces si no se cumple que $W_1 \subseteq W_2$ ó $W_2 \subseteq W_1$, $W_1 \cup W_2$ no es subespacio. Por contradicción esto es, si $W_1 \cup W_2$ es subespacio, entonces $W_1 \subseteq W_2$ ó $W_2 \subseteq W_1$.

($\leftarrow$) Si $W_1 \subseteq W_2$ entonces para todo $u \in W_1$ se cumple que $u \in W_2$. Ahora, sea $v \in W_1 \cup W_2$. Al ser W_1 subconjunto de W_2, tenemos automáticamente que $v \in W_2$, de modo que $u + v \in W_2$ o lo que es equivalente $u + v \in W_1 \cup W_2$. Por otra parte, para todo $a \in F$, $au \in W_1$ y $au \in W_1 \cup W_2$. Igualmente al ser W_1 y W_2 subespacios, ambos contienen el vector *0*, de modo que su unión contiene el vector *0*. Esto verifica que si $W_1 \subseteq W_2$ entonces $W_1 \cup W_2$ es subespacio. El mismo razonamiento aplica si $W_2 \subseteq W_1$.

□

Si la unión de subespacios no genera subespacios, tenemos que considerar una operación entre conjuntos que sí lo haga. Tal operación es la suma de subespacios.

Def. Si S_1 y S_2 son dos conjuntos no vacíos de un espacio vectorial V, entonces la **suma de** S_1 **y** S_2, denotada $S_1 + S_2$, es el conjunto $\{x + y : x \in S_1, y \in S_2\}$.

Teorema 1.6. *Sean W_1 y W_2 subespacios de un espacio vectorial V sobre F. Entonces,*

(*a*) *$W_1 + W_2$ es un subespacio de V;*

(*b*) *todo subespacio que contiene a W_1 y a W_2 contiene también a $W_1 + W_2$.*

Demostración.

(*a*) Sean $u_1, v_1 \in W_1$ y $u_2, v_2 \in W_2$. Entonces $u_1 + u_2$, $v_1 + v_2 \in W_1 + W_2$. Tenemos entonces la suma $(u_1 + u_2) + (v_1 + v_2) = (u_1 + v_1) + (u_2 + v_2)$. Nótese que el primer sumando es elemento de W_1 y el segundo es elemento de W_2, de modo que esta suma pertenece a $W_1 + W_2$. Ahora, sean $a \in F$ y $u + v \in W_1 + W_2$, donde $u \in W_1$ y $v \in W_2$. Entonces se cumple lo siguiente: $a(u+v) = au + uv$. Como $au \in W_1$ y $av \in W_2$, por ser W_1 y W_2 espacios vectoriales, tenemos que $a(u + v) \in W_1 + W_2$. Al ser W_1 y W_2 subespacios, ambos contienen el vector *0*, tenemos que *0* + *0* = *0* pertenece a $W_1 + W_2$. Esto prueba que $W_1 + W_2$ es subespacio.

(*b*) Sea W un subespacio que contiene a W_1 y a W_2 y sea u un vector en $W_1 + W_2$, tal que $u = u_1 + u_2$, donde $u_1 \in W_1$ y $u_2 \in W_2$. Como $W_1 \subset W$ y $W_2 \subset W$, $u_1, u_2 \in W$, de modo que $W_1 + W_2 \subset W$.

□

A continuación definiremos una operación entre subespacios vectoriales que produce (sub)espacios vectoriales con ciertas propiedades. Esta será de gran utilidad para el futuro desarrollo del curso.

Def. Un espacio vectorial V se llama **suma directa** de W_1 y W_2 si W_1 y W_2 son subespacios de V tales que $W_1 \cap W_2 = \{0\}$ y $W_1 + W_2 = V$. Denotamos la suma directa de W_1 y W_2 escribiendo $W_1 \oplus W_2$.

1.2. Combinaciones lineales

Def. Sea V un espacio vectorial sobre F. Dada una colección finita de vectores $v_1, v_2, \ldots, v_r$ en V, una expresión de la forma

$$a_1 v_1 + a_2 v_2 + \cdots + a_r v_r,$$

para una colección de escalares $a_1, a_2, \ldots, a_r$ en F, se llama **combinación lineal** (de los vectores $v_1, v_2, \ldots, v_r$). La combinación lineal

$$0v_1 + 0v_2 + \cdots + 0v_r$$

se llama combinación lineal trivial (de los vectores $v_1, v_2, \ldots, v_r$).

Def. Sea S un subconjunto del espacio vectorial V sobre un campo F. El subconjunto de V que consta de todas las posibles combinaciones lineales de los elementos en S se llama **span** de (o **espacio generado** por) S, denotado por $\operatorname{span}(S)$. Por convención, $\operatorname{span}(\varnothing) = \{0\}$.

Teorema 1.7. *Sea V un espacio vectorial sobre F. Entonces para cualquier subconjunto S de V, $\operatorname{span}(S)$ es subespacio de V y todo subespacio de V que contiene a S también debe contener al $\operatorname{span}(S)$.*

Demostración. Sea $S = \{v_1, v_2, \ldots, v_r\}$ un subconjunto finito de V. Si $a_1, a_2, \ldots, a_r \in F$ y $b_1, b_2, \ldots, b_r \in F$ son dos colecciones de escalares, entonces $a_1v_1 + a_2v_2 + \cdots + a_rv_r + b_1v_1 + b_2v_2 + \cdots + b_rv_r = (a_1 + b_1)v_1 + (a_2 + b_2)v_2 + \cdots + (a_r + b_r)v_r$ Esto muestra que $\mathrm{span}(S)$ es cerrado bajo adición.

Si $c \in F$ y $a_1, a_2, \ldots, a_r \in F$ es una colección de escalares, entonces

$$c(a_1v_1 + a_2v_2 + \cdots + a_rv_r) = (ca_1)v_1 + (ca_2)v_2 + \cdots + (ca_r)v_r$$

Esto muestra que $\mathrm{span}(S)$ es cerrado bajo multiplicación escalar. Por otra parte, es fácil ver que la combinación lineal trivial produce al vector *0*. Esto prueba que $\mathrm{span}(S)$ es un subespacio vectorial.

Que todo subespacio que contiene a S contiene al span de S es consecuencia del teorema 1.7. □

Def. Sea S un subconjunto de un espacio vectorial V sobre un campo F. El subconjunto S es **linealmente independiente** si la única combinación lineal de vectores en S igual al vector cero de V es la combinación lineal trivial. Un subconjunto de un espacio vectorial que no es linealmente independiente se dice **linealmente dependiente**. Un subconjunto finito $S = v_1, v_2, \ldots, v_r$ de un espacio vectorial V sobre un campo F es linealmente independiente si:

$$a_1v_1 + a_2v_2 + \cdots + a_rv_r = \mathit{0} \Rightarrow a_1 = a_2 = \cdots = a_r = 0$$

Si S es infinito, entonces S es linealmente independiente si y sólo si todo subconjunto finito de S es linealmente independiente. Claramente, un subconjunto de un conjunto linealmente independiente es linealmente independiente.

Teorema 1.8 (NKB)**.** *Sea S un subconjunto linealmente independiente de un espacio vectorial sobre un campo F. Sea v un vector en V, $v \notin S$. Entonces $S \cup v$ es linealmente dependiente si y sólo si $v \in \mathrm{span}(S)$.*

Demostración. Sea v un vector en V tal que $S \cup \{v\}$ es linealmente dependiente. Esto significa que existe una colección finita $\{v_1, v_2, \ldots, v_r\}$ en S tal que $\{v_1, v_2, \ldots, v_r\} \cup \{v\}$ es linealmente dependiente. Entonces, existen escales $a_1, a_2, \ldots, a_r, b \in F$, no todos cero, tales que

$$a_1v_1 + a_2v_2 + \cdots + a_rv_r + bv = \mathit{0}$$

Como $\{v_1, v_2, \ldots, v_r\}$ es linealmente independiente por hipótesis, sabemos que $b \neq 0$, lo cual implica $v = -(1/b)(a_1v_1 + a_2v_2 + \cdots + a_rv_r) \in \mathrm{span}(S)$.

Recíprocamente, sea $v \in \mathrm{span}(S)$. Esto implica que existe una colección finita $\{v_1, v_2, \ldots, v_r\}$ en S tal que $v = a_1v_1 + a_2v_2 + \cdots + a_rv_r$, donde los a_i son distintos de cero. Entonces, $a_1v_1 + a_2a_2 + \cdots + a_rv_r - v = \mathit{0}$, lo cual implica que $\{v_1, v_2, \ldots, v_r\} \cup \{v\}$, o lo que es equivalente $S \cup \{v\}$, es linealmente dependiente, dado que la última es una combinación lineal no trivial. □

Def. Sea S un subconjunto linealmente independiente de un espacio vectorial V sobre un campo F. S se denomina **linealmente independiente maximal** si es maximal del orden parcial $(L, \subseteq)$, donde L es la familia de todas las colecciones linealmente independientes de vectores en V. Un tal subconjunto se llama **base** de V.

Teorema 1.9. *Sea V un espacio vectorial sobre un campo F y sea S un subconjunto de V. Entonces S es una base de V si y sólo si S es linealmente independiente y $\mathrm{span}(S) = V$.*

Demostración. Si S es una base de V, en particular, S es linealmente independiente. Para mostrar que $\mathrm{span}(S) = V$, sea v un vector en V tal que $v \notin S$. Esto implica que $S \cup \{v\}$ es linealmente dependiente, puesto que S es maximal. En consecuencia, por 1.8, $v \cup \mathrm{span}(S)$. Por el contrario, supongamos que S es linealmente independiente y que $\mathrm{span}(S) = V$. Esto último implica que todo $v \notin V$ pertenece al $\mathrm{span}(S)$, lo cual a su vez implica por 1.8 que $S \cup \{v\}$ es linealmente dependiente para todo $v \notin S$. Entonces, S debe ser maximal. □

Si $\text{span}(S) = V$ decimos que S genera a V o que S es un conjunto de generadores de V. Tener un conjunto de generadores S de un espacio vectorial V implica que todo vector en V puede escribirse como combinación lineal de vectores en S. Si este conjunto resulta ser también base de V entonces, por su independencia lineal, todo vector en V se escribe como combinación lineal de vectores en V de forma única.

Lema 1.10 (Lema de Zorn). *Sea $(S, \leq)$ un orden parcial. Si toda cadena en S (i.e. todo subconjunto linealmente ordenado por $\leq$) tiene cota superior, entonces S tiene por lo menos un elemento maximal.*

Teorema 1.11. *Sea V un espacio vectorial no trivial sobre un campo F. Entonces V tiene una base.*

Demostración. Consideremos la familia L de todas las colecciones linealmente independientes en V junto al relación de orden parcial $\subseteq$. La familia L es no vacío puesto que V es no trivial. Como una base es un elemento máximo de L, el 1.10 proveerá uno si logramos probar que cada cadena en L tiene cota superior. Sea K una cadena en L. Decimos que la unión S de todos los elementos de K es una cota superior para K. Claramente S contiene todos los elementos en K, de modo que basta demostrar que S es linealmente independiente. Puesto que K es una cadena y S es la unión de todos los elementos en K, todo subconjunto finito de S está en algún elemento de K, luego es linealmente independiente. $\square$

Def. Sea V un espacio vectorial sobre un campo F. V es **finitamente generado** o de **dimensión finita** si V tiene un conjunto finito de generadores.

Teorema 1.12 (Replacement Theorem). *Sea V un espacio vectorial finitamente generado sobre un campo F. Supongamos que $L = \{v_1, v_2, \ldots, v_k\}$ es un subconjunto linealmente independiente de V y $G = \{u_1, u_2, \ldots, u_n\}$ es un conjunto de generadores de V. Entonces,*

(*a*) *$k \leq m$;*

(*b*) *existe una subconjunto H de G con exactamente $m - k$ vectores tal que $L \cup H$ también es un conjunto de generadores de V.*

Demostración. Por inducción sobre m. Si $m = 0$, $L = \varnothing$, tomando $H = G$ obtenemos el resultado deseado. Ahora supongamos que el teorema es verdadero para algún entero $m \geq 0$. Verificamos que el teorema sea verdadero para $m + 1$. Sea $L = \{v_1, v_2, \ldots, v_(m + 1)\}$ un subconjunto linealmente independiente de V. Sabiendo que $\{v_1, v_2, \ldots, v_m\}$ es linealmente independiente también, aplicamos la hipótesis de inducción para concluir que $m \leq n$ y que existe un subconjunto $\{u_1, u_2, \ldots, u_(n-m)\}$ de G tal que $\{v_1, v_2, \ldots, v_m\} \cup \{u_1, u_2, \ldots, u_(n - m)\}$ genera V. De modo que existen escalares $a_1, a_2, \ldots, a_m, b_1, b_2, \ldots, b_{n-m}$ tales que

$$a_1 v_1 + a_2 v_2 + \cdots + a_m v_m + b_1 u_1 + b_2 u_2 + \cdots + b_(n - m) u_(n - m) = v_(m + 1)$$

Nótese que $n - m > 0$, no sea que $v_(m + 1)$ sea una combinación lineal de $v_1, v_2, \ldots, v_m$, lo cual por 1.8 contradice la suposición que L es linealmente independiente. Entonces $n > m$; esto es, $n \geq m + 1$. Más aún, algún b_i, digamos b_1, es diferente de cero, pues de otro modo obtendríamos la misma contradicción. Despejando u_1 obtenemos

$$u_1 = (-b_1^{-1} a_1) v_1 + (-b_1^{-1} a_2) v_2 + \cdots + (-b_1^{-1} a_m) v_m + (b_1^{-1}) v_{m+1} + (-b_1^{-1} b_2) u_2 + \cdots + (-b_1^{-1} b_{n-m}) u_{n-m}$$

Sea $H = \{u_2, \ldots, u_{n-m}\}$. Entonces $u_1 \in \text{span}(L \cup H)$, y puesto que $v_1, v_2, \ldots, v_m, u_2, \ldots, u_{n-m}$ claramente están en $\text{span}(L \cup H)$, se cumple que $\{v_1, v_2, \ldots, v_m, u_2, \ldots, u_{n-m}\} \subseteq \text{span}(L \cup H)$. Como $\{v_1, v_2, \ldots, v_m, u_1, u_2, \ldots, u_{n-m}\}$ genera a V, el primer teorema de esta entrada implica que $\text{span}(L \cup H) = V$. Luego H es un subconjunto de G que contiene $(n - m) - 1 = n - (m + 1)$ vectores, el teorema es válido para $m + 1$. $\square$

Corolario 1.12.1. *Sea V un espacio vectorial finitamente generado. Entonces toda base de V contiene el mismo número de vectores.*

Def. Sea V un espacio vectorial finitamente generado. Se denomina **dimensión** de V al número de vectores en una base de V.

Observación. La dimensión del espacio vectorial trivial es 0.

Anteriormente, definimos la suma directa de subespacios y afirmamos que esta operación será importante en el curso. El siguiente resultado establece una propiedad fundamental de la suma directa.

Teorema 1.13. *Sea V un espacio vectorial sobre F y W_1, W_2 subespacios de dimensión finita de V. Entonces,*

$$\dim(W_1 + W_2) = \dim(W_1) + \dim(W_2) - \dim(W_1 \cap W_2).$$

Demostración. Sea $\{u_1, u_2, \ldots, u_k\}$ una base de $W_1 \cap W_2$. Esta base la extendemos a $\{u_1, u_2, \ldots, u_k, v_1, v_2, \ldots, v_m\}$ para formar una base de W_1 y a $\{u_1, u_2, \ldots, u_k, w_1, w_2, \ldots, w_n\}$ para formar una base de W_2. Basta con demostrar que

$$\beta = \{u_1, u_2, \ldots, u_k, v_1, v_2, \ldots, v_m, w_1, w_2, \ldots, w_n\}$$

es una base de $W_1 + W_2$. Primero, veamos que es linealmente independiente. Sean $a_1, a_2, \ldots, a_k, b_1, \ldots, b_m, c_1, \ldots, c_n \in F$ escalares tales que

$$a_1u_1 + a_2u_2 + \cdots + a_ku_k + b_1v_1 + \cdots + b_mv_m + c_1w_1 + \cdots + c_nw_n = 0.$$

Ahora, sea

$$v = \sum_{i=1}^{k} a_iu_i + \sum_{i=1}^{m} b_iv_i = -\sum_{i=1}^{n} c_iw_i.$$

Nótese que $v \in W_1 \cap W_2$, puesto que el término del medio está en W_1 y el término de la derecha está en W_2. Por lo tanto, existen $d_1, d_2, \ldots, d_k \in F$ tales que

$$v = \sum_{i=1}^{k} d_iu_i,$$

entonces,

$$0 = v - v = \sum_{i=1}^{k} (a_i - d_i)u_i + \sum_{i=1}^{m} b_iv_i.$$

Por independencia lineal de $\{u_1, u_2, \ldots, u_k, v_1, v_2, \ldots, v_m\}$, los b_i son todos iguales a cero y los a_i son iguales a los d_i. Luego,

$$0 = \sum_{i=1}^{k} a_iu_i + \sum_{i=1}^{n} c_iw_i.$$

Por independencia lineal de $\{v_1, v_2, \ldots, v_k, w_1, w_2, \ldots, w_n\}$ concluimos que $a_1 = a_2 = \cdots = a_k = c_1 = \cdots = c_n = 0$. Esto establece la independencia lineal de β.

Ahora, para verificar que β genera a $W_1 + W_2$ basta con notar que si $u \in W_1 + W_2$, existen v y w en W_1 y W_2 respectivamente, tales que $u = v + w$. Por hipótesis, existen $a_1, a_2, \ldots, a_k, b_1, \ldots, b_m, c_1, \ldots, c_k, d_1, \ldots, d_n \in F$ tales que

$$v = \sum_{i=1}^{k} a_iu_i + \sum_{i=1}^{m} b_iv_i \quad \text{y} \quad w = \sum_{i=1}^{k} c_iu_i + \sum_{i=1}^{n} d_iw_i,$$

de modo que u se puede escribir como

$$u = \sum_{i=1}^{k}(a_i + c_i)u_i + \sum_{i=1}^{m} b_i v_i + \sum_{i=1}^{n} d_i w_i.$$

Esto establece que β genera a $W_1 + W_2$.

Nótese que β tiene $k + n + m = (k + n) + (k + m) - k = \dim(W_1) + \dim(W_2) - \dim(W_1 \cap W_2)$ vectores, lo cual establece el teorema. □

Corolario 1.13.1. *Sean W_1 y W_2 subespacios de dimensión finita de un espacio V y sea $V = W_1 + W_2$. V es la suma directa de W_1 y W_2 si y solo si $\dim(V) = \dim(W_1) + \dim(W_2)$.*

Demostración. Si $V = W_1 \oplus W_2$, $W_1 \cap W_2 = \{0\}$, luego $\dim(W_1 \cap W_2) = 0$. Por el teorema, $\dim(V) = \dim(W_1) + \dim(W_2) - 0 = \dim(W_1) + \dim(W_2)$.

Recíprocamente, si $\dim(V) = \dim(W_1 + W_2) = \dim(W_1) + \dim(W_2)$, $\dim(W_1 \cap W_2) = 0$, por el teorema anterior. Luego, $W_1 \cap W_2 = \{0\}$. Esto implica que la suma de W_1 y W_2 es directa. □

2. Transformaciones lineales

Def. Sean V y W espacios vectoriales sobre F. Decimos que $T : V \to W$ es una **transformación lineal** de V en W si, para todo $u, v \in V$ y $c \in F$, tenemos que:

1. $T(u + v) = T(u) + T(v)$
2. $T(cu) = cT(u)$

Teorema 2.1. *Sean V y W espacios vectoriales y $T : V \to W$ una función. Entonces,*

(*a*) *si T es lineal, $T(0) = 0$*

(*b*) *T es lineal si y sólo si $T(cu + v) = cT(u) + T(v)$ para todo $u, v \in V$ y $c \in F$.*

(*c*) *si T es lineal, entonces $T(u - v) = T(u) - T(v)$ para todo $u, v \in V$.*

(*d*) *T es lineal si y sólo si, para $v_1, v_2, \ldots, v_n \in V$ y $a_1, a_2, \ldots, a_n \in F$, tenemos que $T(\sum_{i=1}^{n} a_i v_i) = \sum_{i=1}^{n} a_i T(v_i)$*

Demostración. Las pruebas de *a, b* y *c* son elementales.

La prueba de d es por inducción sobre n. El caso base es la parte 2 de la definición anterior. Supongamos que d se cumple para $n = k - 1$. Tomemos $v = \sum_{i=1}^{k} a_i v_i$ donde $a_i \in F$ y $v_i \in V$ para $i = 1, 2, \ldots, k$. Entonces,

$$T(v) = T\left(\sum_{i=1}^{k} a_i v_i\right) = T\left(\sum_{i=1}^{k-1} a_i v_i + a_k v_k\right) = T\left(\sum_{i=1}^{k-1} a_i v_i\right) + T(a_k v_k)$$

Por hipótesis de inducción, el primer término es $\sum_{i=1}^{k-1} a_i T(v_i)$ y así:

$$T(v) = \sum_{i=1}^{k-1} a_i T(v_i) + a_k T(v_k) = \sum_{i=1}^{k} a_i T(v_i).$$

□

Existen dos ejemplos importantes de transformaciones lineales.

Def. Sean V y W dos espacios vectoriales sobre F. Definimos a la transformación identidad, denotada por $\mathrm{id}_V : V \to V$ como $\mathrm{id}_V(v) = v$ para todo $v \in V$. Además, definimos la transformación cero, denotada por $T_0 : V \to W$, como $T_0(v) = 0$ para todo $v \in V$.

Def. Sean V y W espacios vectoriales y sea $T : V \to W$ una transformación lineal. Definimos el **núcleo** (espacio nulo o kernel) de T, denotado como $\ker(T)$, como el conjunto de todos los vectores $v \in V$ tales que $T(v) = 0$; esto es, $\ker(T) = \{v \in V : T(v) = 0\}$. Definimos también la **imagen** (o rango) de T, denotado $\mathrm{im}(T)$, como el subconjunto de W que consiste en todas las imágenes bajo T de vectores en V; esto es $\mathrm{im}(T) = \{T(v) : v \in V\}$.

El núcleo de T es subconjunto de V, mientras la imagen de T es subconjunto de W. Más aún, estos dos conjuntos resultan ser subespacios de V y de W, respectivamente.

Teorema 2.2. *Sean V y W espacios vectoriales y $T : V \to W$ una transformación lineal. Entonces $\ker(T)$ y $\mathrm{im}(T)$ son subespacios de V y W respectivamente.*

Demostración. Con respecto a la notación, utilizamos 0_V y 0_W para denotar los vectores cero de V y W respectivamente. Como $T(0_V) = 0_W$, tenemos que $0_V \in \ker(T)$. Sean $u, v \in \ker(T)$ y $c \in F$. Entonces $T(u + v) = T(u) + T(v) = 0_W + 0_W = 0_W$ y $T(cu) = cT(u) = c0_W = 0_W$. Entonces $u + v \in \ker(T)$ y $cu \in \ker(T)$, de modo que $\ker(T)$ es subespacio de V.

Como $T(0_V) = 0_W$, tenemos que $0_W \in \mathrm{im}(T)$. Ahora sean $w, x \in \mathrm{im}(T)$ y $c \in F$. Entonces existen u y v en V tales que $T(u) = w$ y $T(v) = x$. Entonces $T(u + v) = T(u) + T(v) = w + x$, y $T(cu) = cT(u) = cw$. Luego $w + x \in \mathrm{im}(T)$ y $cw \in \mathrm{im}(T)$, por ende $\mathrm{im}(T)$ es subespacio de W. □

Teorema 2.3. *Sean V y W espacios vectoriales y sea $T : V \to W$ una transformación lineal. Si $\beta = \{v_1, v_2, \ldots, v_n\}$ es una base de V, entonces $\mathrm{im}(T) = \mathrm{span}(T(\beta)) = \mathrm{span}(\{T(v_i)\}_{i=1}^n)$.*

Demostración. Es evidente que $T(v_i) \in \mathrm{im}(T)$ para $i = 1, 2, \ldots, n$. Dado que $\mathrm{im}(T)$ es un subespacio, por teorema 1.7 $\mathrm{im}(T)$ contiene a $\mathrm{span}(\{T(v_1), T(v_2), \ldots, T(v_n)\}) = \mathrm{span}(T(\beta))$. Ahora supongamos que $w \in \mathrm{im}(T)$. Entonces $w = T(v)$ para algún $v \in V$. Como β es una base de V, tenemos que

$$v = \sum_{i=1}^{n} a_i v_i \quad \text{para algunos } a_1, a_2, \ldots, a_n \in F$$

Como T es lineal, se cumple que

$$w = T(v) = \sum_{i=1}^{n} a_i T(v_i) \in \mathrm{span}(T(\beta))$$

Entonces $\mathrm{im}(T)$ está contenido en $\mathrm{span}(T(\beta))$. □

Este último teorema nos permite encontrar una base de $\mathrm{im}(T)$ y determinar su dimensión. A continuación le damos un nombre a la dimensión de $\mathrm{im}(T)$ y de $\ker(T)$.

Def. Sean V y W espacios vectoriales y $T : V \to W$ una transformación lineal. Si $\ker(T)$ y $\mathrm{im}(T)$ son de dimensión finita, definimos la **nulidad** de T, denotado $\mathrm{null}(T)$, y el **rango** de T, denotado $\mathrm{rank}(T)$, como las dimensiones de $\ker(T)$ y $\mathrm{im}(T)$, respectivamente.

Teorema 2.4 (Teorema de la dimensión)**.** *Sean V y W espacios vectoriales y $T : V \to W$ una transformación lineal. Si V es finitamente generado, entonces $\mathrm{null}(T) + \mathrm{rank}(T) = \dim(V)$.*

Demostración. Supongamos que $\dim(V) = n$, $\dim(N(T)) = k$ y $\{v_1, v_2, \ldots, v_k\}$ es una base de $\ker(T)$. Podemos extender $\{v_1, v_2, \ldots, v_k\}$ a una base $\beta = \{v_1, v_2, \ldots, v_k, v_{k+1}, \ldots, v_n\}$ de V. Afirmamos que $S = \{T(v_{k+1}), T(v_{k+2}), \ldots, T(v_n)\}$ es una base de $\mathrm{im}(T)$. Primero, verificamos que S

genera a $\text{im}(T)$. Utilizando el teorema 2.3, y el hecho de que $T(v_i) = 0$ para $1 \leq i \leq k$, tenemos que

$$\begin{aligned}\text{im}(T) &= \text{span}(\{T(v_1), T(v_2), \ldots, T(v_k), T(v_{k+1}, \ldots, T(v_n)\}) \\ &= \text{span}(\{T(v_{k+1}), T(v_{k+2}), \ldots, T(v_n)\}) = \text{span}(S).\end{aligned}$$

Ahora probemos que S es linealmente independiente. Supongamos que

$$\sum_{i=k+1}^{n} b_i T(v_i) = \mathit{0} \quad \text{para } b_{k+1}, b_{k+2}, \ldots, b_n \in F$$

Utilizando el hecho de que T es lineal, tenemos que

$$T\left(\sum_{i=k+1}^{n} b_i v_i\right) = \mathit{0}$$

De modo que,

$$\sum_{i=k+1}^{n} b_i v_i \in \ker(T)$$

Entonces existen $c_1, c_2, \ldots, c_k \in F$ tales que

$$\sum_{i=k+1}^{n} b_i v_i = \sum_{i=1}^{k} c_i v_i$$

$$\sum_{i=k+1}^{n} b_i v_i + \sum_{i=1}^{k} (-c_i) v_i = \mathit{0}$$

Como β es una base de V, tenemos que $b_i = 0$ para todos los i. Así que S es linealmente independiente. Nótese que este argumento también muestra que $T(v_{k+1}), T(v_{k+1}), \ldots, T(v_n)$ son distintos; entonces $\text{rank}(T) = n - k$. □

Teorema 2.5. *Sean V y W espacios vectoriales y $T : V \to W$ una transformación lineal. Entonces T es inyectiva si y sólo si* $\ker(T) = \{\mathit{0}\}$

Demostración. Supongamos que T es inyectiva y $v \in \ker(T)$. Entonces $T(v) = \mathit{0} = T(\mathit{0})$, pero como T es inyectiva, obtenemos que $v = \mathit{0}$. Luego, $\ker(T) = \{\mathit{0}\}$.

Ahora asumamos que $\ker(T) = \{\mathit{0}\}$, y supongamos que $T(u) = T(v)$, donde $u, v \in V$. Entonces $\mathit{0} = T(u) - T(v) = T(u - v)$ (por parte tres de teorema 2.1). En consecuencia $u - v \in \ker(T) = \{\mathit{0}\}$. Entonces $u - v = \mathit{0}$ y $u = v$. Esto último implica que T es inyectiva. □

Teorema 2.6. *Sean V y W espacios vectoriales finitamente generados con* $\dim(V) = \dim(W)$. *Sea $T : V \to W$ una transformación lineal. Entonces las siguientes afirmaciones son equivalentes:*

(a) *T es inyectiva*

(b) *T es sobreyectiva*

(c) $\text{rank}(T) = \dim(V)$

Demostración. Del teorema de la dimensión (2.4) tenemos que

$$\text{null}(T) + \text{rank}(T) = \dim(V)$$

Ahora, aplicando el teorema 2.5 obtenemos que T es inyectiva si y sólo si $\ker(T) = \{0\}$, si y sólo si $\text{null}(T) = 0$, si y sólo si $\text{rank}(T) = \dim(V)$, si y sólo si $\text{rank}(T) = \dim(W)$, si y sólo si $\dim(\text{im}(T)) = \dim(W)$. Esta última igualdad sumada al hecho de que $\text{im}(T)$ es subespacio de W implica (por Teorema 1.11 de Friedberg, Insel & Spence) que $\text{im}(T) = W$, la definición de ser sobreyectiva. □

2.1. Representación matricial de una transformación lineal

Def. Sea V un espacio vectorial de dimensión finita. Una **base ordenada** de V es una base de V dotada con un orden específico; esto es, una base ordenada de V es una secuencia finita de vectores linealmente independientes en V que generan a V.

Para un espacio vectorial F^n, e_i es una n-tupla donde la única entrada distinta de cero es la i-ésima, cuyo valor es 1. El conjunto $\{e_1, e_2, \dots, e_n\}$ es la base ordenada estándar de F^n. Similarmente para $P_n(F)$, llamamos a $\{1, x, \dots, x^n\}$ la base ordenada estándar de $P_n(F)$.
Las bases ordenadas nos permiten hablar de vectores de coordenadas.

Def. Sea $\beta = \{u_1, u_2, \dots, u_3\}$ una base ordenada para un espacio vectorial finitamente generado V. Para $v \in V$, sean $a_1, a_2, \dots, a_n \in F$ los únicos escalares tales que

$$v = \sum_{i=1}^{n} a_i u_i$$

Definimos el **vector de coordenadas** de v relativo a β, denotado $[v]_\beta$, como

$$[v]_\beta = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix}$$

A partir de esta definición, supongamos que V y W son espacios vectoriales de dimensión finita con las bases ordenadas $\beta = \{v_1, v_2, \dots, v_n\}$ y $\gamma = \{w_1, w_2, \dots, w_m\}$, respectivamente. Sea además $T : V \to W$ una transformación lineal. Entonces, para cada j, tal que $1 \leq j \leq n$, existen escalares únicos $a_{ij} \in F$, $1 \leq i \leq m$, tales que:

$$T(v_j) = \sum_{i=1}^{m} a_{ij} w_i$$

Def. Utilizando la notación anterior, llamamos a la matriz A de tamaño $m \times n$ definida por $A_{ij} = a_{ij}$ la **representación matricial** de T en las bases ordenadas β y γ. Escribimos $A = [T]_\beta^\gamma$. Si $V = W$ y $\beta = \gamma$, escribimos entonces $A = [T]_\beta$.

La i-ésima columna de la matriz $[T]_\beta^\gamma$ corresponde al vector de coordenadas de $T(v_i)$ relativo a γ, donde v_i es el i-ésimo vector en β.
Nótese que si $U : V \to W$ es una transformación lineal tal que $[U]_\beta^\gamma = [T]_\beta^\gamma$, entonces $U = T$ por corolario al teorema 2.6 de Friedberg, Insel & Spence.

Def. Sean $T, U : V \to W$ funciones arbitrarias, donde V y W son espacios vectoriales sobre F, y sea $a \in F$. Definimos $T + U : V \to W$ como $(T + U)(v) = T(v) + U(v)$ para todo $v \in V$ y $aT : V \to W$ como $(aT)(v) = aT(v)$ para todo $v \in V$.

Teorema 2.7. *Sean V y W espacios vectoriales sobre un campo F y sean $T, U : V \to W$ transformaciones lineales.*

(a) *Para todo $a \in F$, $aT + U$ es lineal.*

(b) *Con base en las operaciones de suma y multiplicación escalar de la definición precedente, la colección de todas las transformaciones lineales de V a W es un espacio vectorial sobre F.*

Demostración.

(a) Sean $u, v \in F$ y $c \in F$. Entonces

$$\begin{aligned}(aT + U)(cu + v) &= aT(cu + v) + U(cu + v)\\ &= a[cT(u) + T(v)] + cU(u) + U(v)\\ &= acT(u) + cU(v) + aT(u) + U(v)\\ &= c(aT + U)(u) + (aT + U)(v)\end{aligned}$$

Entonces $aT + U$ es lineal.

(b) Teniendo en cuenta que T_0, la transformación cero, sería el vector cero, es fácil verificar que los axiomas de los espacios vectoriales se satisfacen, luego el conjunto de todas las transformaciones de V en W, junto a la definición de suma y multiplicación escalar anterior, es un espacio vectorial.

□

Def. Sean V y W espacios vectoriales sobre F. Denotamos el espacio vectorial de todas las transformaciones de V en W por $\mathcal{L}(V, W)$. En caso que $V = W$, escribimos $\mathcal{L}(V)$.

Teorema 2.8. *Sean V y W espacios vectoriales finitamente generados dotados con las bases ordenadas β y γ respectivamente. Sean $T, U : V \to W$ transformaciones lineales. Entonces,*

(a) $[T + U]_\beta^\gamma = [T]_\beta^\gamma + [U]_\beta^\gamma$

(b) $[kT]_\beta^\gamma = k[T]_\beta^\gamma$

Demostración. Sean $\beta = \{v_1, v_2, \ldots, v_n\}$ y $\gamma = \{w_1, w_2 \ldots, w_m\}$. Existen escalares únicos a_{ij} y b_{ij} $(1 \leq i \leq m, 1 \leq j \leq n)$ tales que

$$T(v_j) = \sum_{i=1}^{n} a_{ij} w_i \quad \text{y} \quad U(v_j) = \sum_{i=1}^{m} b_{ij} w_i \quad \text{para} \quad 1 \leq j \leq n$$

Luego

$$(T + U)(v_j) = \sum_{i=1}^{m} (a_{ij} + b_{ij}) w_i$$

Entonces

$$([T + U]_\beta^\gamma)_{ij} = a_{ij} + b_{ij} = ([T]_\beta^\gamma + [U]_\beta^\gamma)_{ij}$$

Esto prueba a. Para la prueba de b., nuevamente

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

Luego

$$\begin{aligned}(kT)(v_j) &= \sum_{i=1}^{n} ka_{ij}w_i \\ &= k\sum_{i=1}^{n} a_{ij}w_i\end{aligned}$$

De modo que

$$([kT]_\beta^\gamma)_{ij} = (k[T]_\beta^\gamma)_{ij}$$

□

2.2. Composición de transformaciones lineales y multiplicación de matrices

Anteriormente, hemos visto cómo asociar una matriz con una transformación lineal de tal forma que la suma y la multiplicación escalar de las matrices está asociada a las sumas y multiplicaciones escalares correspondientes de transformaciones. Ahora pretendemos establecer como se relacionan las representaciones matriciales con la composición de transformaciones lineales.
Utilizaremos la notación UT para representar la compuesta de U y T.

Teorema 2.9. *Sean V, W y Z espacios vectoriales sobre el mismo campo F y sean $T : V \to W$ y $U : W \to Z$ transformaciones lineales. Entonces $UT : V \to Z$ es lineal.*

Demostración. Sean $u, v \in V$ y $a \in F$. Entonces

$$\begin{aligned}UT(au+v) &= U(T(au+v)) = U(aT(u)+T(v)) \\ &= aU(T(u)) + U(T(v)) = a(UT)(u) + UT(v)\end{aligned}$$

□

Teorema 2.10. *Sea V un espacio vectorial. Sean $T, U_1, U_2 \in \mathcal{L}(V)$, entonces:*

(*a*) $T(U_1+U_2) = TU_1 + TU_2$ *y* $(U_1+U_2)T = U_1T + U_2T$

(*b*) $T(U_1U_2) = (TU_1)U_2$

(*c*) $T\,\mathrm{id}_V = \mathrm{id}_V T = T$

(*d*) $a(U_1U_2) = (aU_1)U_2 = U_1(aU_2)$ *para todo escalar a.*

Demostración. Sea v un vector en V.

(*a*)

$$T(U_1+U_2)(v) = T(U_1(v)+U_2(v))$$

Por linealidad de T,

$$T(U_1(v)) + T(U_2(v)) = TU_1(v) + TU_2(v) = (TU_1+TU_2)(v)$$

Por otra parte, $(U_1+U_2)T(v) = (U_1+U_2)(T(v)) = U_1(T(v)) + U_2(T(v))$ por definición. Entonces

$$(U_1+U_2)T(v) = U_1T(v) + U_2T(v)$$

(*b*) $T(U_1U_2)(v) = T(U_1U_2(v)) = T(U_1(U_2(v))) = TU_1(U_2(v))$

(c) $T\,\mathrm{id}_V(v) = T(\mathrm{id}_V(v)) = T(v) = \mathrm{id}_V(T(v))$

(d) $a(U_1U_2)(v) = a(U_1(U_2(v))) = (aU_1)(U_2(v)) = (aU_1)U_2(v)$
Partiendo de la primera igualdad, $a(U_1U_2)(v) = a(U_1(U_2(v)))$, por linealidad de U_1,

$$a(U_1(U_2(v))) = U_1(a(U_2(v))) = U_1(aU_2(v))$$

□

Un resultado más general vale para transformaciones lineales cuyos dominios no coinciden con sus codominios.

Def. Sean A una matriz $m \times n$ y B una matriz $n \times p$. Definimos el producto de A y B (en ese orden), denotado AB como la matriz $m \times p$ tal que

$$(AB)_{ij} = \sum_{k=1}^{n} A_{ik}B_{kj} \qquad \text{para } 1 \leq i \leq m, \quad 1 \leq j \leq p$$

Para que el producto matricial esté definido se requiere que el número de columnas del primer factor sea igual al número de filas del segundo factor.

Teorema 2.11. *Sean V, W, y Z espacios vectoriales de dimensión finita con las bases ordenadas α, β y γ respectivamente. Sean $T : V \to W$ y $U : W \to Z$ transformaciones lineales. Entonces,*

$$[UT]_\alpha^\gamma = [U]_\beta^\gamma [T]_\alpha^\beta$$

Demostración. Si $\alpha = \{u_1, u_2, \ldots, u_n\}$, $\beta = \{v_1, v_2, \ldots, v_m\}$ y $\gamma = \{w_1, w_2, \ldots, w_p\}$, la imagen de u_i a través de T es

$$T(u_i) = \sum_{j=1}^{m} a_{ij}v_j$$

Luego, la imagen de u_i a través de UT será

$$UT(u_i) = U(T(u_i)) = U\left(\sum_{j=1}^{m} a_{ij}v_j\right)$$

Por linealidad de U, tenemos que

$$U\left(\sum_{j=1}^{m} a_{ij}v_j\right) = \sum_{j=1}^{m} a_{ij}U\left(v_j\right)$$

Los $U(v_j)$ pueden expresarse a su vez como una combinación lineal de los vectores en γ. Entonces,

$$UT(u_i) = \sum_{j=1}^{m} a_{ij} \sum_{k=1}^{p} b_{jk}w_k$$

De aquí vemos que la entrada en la columna i y en la fila k de $[UT]_\alpha^\gamma$ está dada por $\sum_{j=1}^{m} a_{ij}b_{jk}$, lo cual concuerda con que esta matriz sea el producto entre una matriz con entradas a_{ij} y otra matriz con entradas b_{jk}. Tales matrices son justamente $[T]_\alpha^\beta$ y $[U]_\beta^\gamma$, respectivamente. □

Def. Definimos el **delta de Kronecker** como

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

La **matriz identidad** de tamaño $n \times n$ se define por $(I_n)_{ij} = \delta_{ij}$.

Teorema 2.12. *Sean A una matriz $m \times n$, B y C matrices $n \times p$ y D y E matrices $q \times m$. Entonces*

(a) $A(B+C) = AB + AC$ *y* $(D+E)A = DA + EA$

(b) $a(AB) = (aA)B = A(aB)$ *para cualquier escalar* a

(c) $I_m A = A = AI_n$

(d) *Si V es un espacio vectorial de dimensión n con una base ordenada β,* $[\mathrm{id}_V]_\beta = I_n$

Demostración.

(a) Tenemos que

$$\begin{aligned}[][A(B+C)]_{ij} &= \sum_{k=1}^{n} A_{ik}(B+C)_{kj} \sum_{k=1}^{n} A_{ik}(B_{kj} + C_{kj}) = \\ &= \sum_{k=1}^{n} (A_{ik}B_{kj} + A_{ik}C_{kj}) = \sum_{k=1}^{n} A_{ik}B_{kj} + \sum_{k=1}^{n} A_{ik}C_{kj} \\ &= (AB)_{ij} + (AC)_{ij} = [AB + AC]_{ij}\end{aligned}$$

Entonces $A(B+C) = AB + AC$. Por otra parte,

$$\begin{aligned}[][(D+E)A]_{ij} &= \sum_{k=1}^{m} (D+E)_{ik}A_{kj} = \sum_{k=1}^{m} (D_{ik} + E_{ik})A_{kj} \\ &= \sum_{k=1}^{m} (D_{ik}A_{kj} + E_{ik}A_{kj}) = \sum_{k=1}^{m} D_{ik}A_{kj} + \sum_{k=1}^{m} E_{ik}A_{kj} \\ &= (DA)_{ij} + (EA)_{ij} = [DA + EA]_{ij}\end{aligned}$$

(b)

$$\begin{aligned}[][a(AB)]_{ij} = a\sum_{k=1}^{n} A_{ik}B_{kj} = \sum_{k=1}^{n} aA_{ik}B_{kj} &= \sum_{k=1}^{n} (aA_{ik})B_{kj} = [(aA)B]_{ij} \\ &= \sum_{k=1}^{n} A_{ik}(aB_{kj}) = [A(aB)]_{ij}\end{aligned}$$

(c)

$$(I_m A)_{ij} = \sum_{k=1}^{m} (I_m)_{ik}A_{kj} = \sum_{k=1}^{m} \delta_{ik}A_{kj} = A_{ij} = \sum_{k=1}^{n} A_{ik}\delta_{kj} = \sum_{k=1}^{n} A_{ik}(I_n)_{kj} = (AI_n)_{ij}$$

(d) Si $\beta = \{v_1, v_2, \ldots, v_n\}$, $I_V(v_i) = v_i$. Entonces, $I_V(v_i) = \sum_{k=1}^{n} \delta_{ik}v_k$, de modo que

$$\left([\mathrm{id}_V]_\beta\right)_{ij} = \delta_{ij} = (I_n)_{ij}$$

$\square$

Teorema 2.13. *Sean V y W espacios vectoriales finitamente generados con las bases ordenadas β y γ, respectivamente, y sea $T : V \to W$ una transformación lineal. Entonces para cada $u \in V$,*

$$[T(u)]_\gamma = [T]_\beta^\gamma [u]_\beta$$

Demostración. Sea u un vector cualquiera en V y sea $u = k_1u_1 + k_2u_2 + \cdots + k_nu_n$ su única representación como combinación lineal de los vectores en la base ordenada β. Entonces, $T(u) = T(k_1u_1 + k_2u_2 + \cdots + k_nu_n)$. Como T es lineal, tenemos que

$$\begin{aligned}
T(u) &= k_1T(u_1) + k_2T(u_2) + \cdots + k_nT(u_n) \\
&= k_1(a_{11}v_1 + \cdots + a_{i1}v_i + \cdots + a_{m1}v_m) + \cdots + k_j(a_{1j}v_1 + \cdots + a_{ij}v_i + \cdots + a_{mj}v_m) \\
&+ \cdots + k_n(a_{1n}v_1 + \cdots + a_{in}v_i + \cdots + a_{mn}v_m) \\
&= (k_1a_{11} + \cdots + k_ja_{1j} + \cdots + k_na_{1n})v_1 + \cdots + (k_1a_{i1} + \cdots + k_ja_{ij} + \cdots + k_na_{in})v_i \\
&+ \cdots + (k_1a_{m1} + \cdots + k_ja_{mj} + \cdots + k_na_{mn})v_m
\end{aligned}$$

En formal matricial esto es

$$\begin{aligned}
[T(u)]_\beta &= \begin{bmatrix} k_1a_{11} + \cdots + k_ja_{1j} + \cdots + k_na_{1n} \\ \vdots \\ k_1a_{i1} + \cdots + k_ja_{ij} + \cdots + k_na_{in} \\ \vdots \\ k_1a_{m1} + \cdots + k_ja_{mj} + \cdots + k_na_{mn} \end{bmatrix} \\
&= \begin{bmatrix} a_{11} & \dots & a_{1j} & \dots & a_{1n} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ a_{i1} & \dots & a_{ij} & \dots & a_{in} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ a_{m1} & \dots & a_{mj} & \dots & a_{mn} \end{bmatrix} \begin{bmatrix} k_1 \\ \vdots \\ k_j \\ \vdots \\ k_n \end{bmatrix} = [T]_\beta^\gamma [u]_\beta
\end{aligned}$$

$\square$

Def. Sea A una matriz $m \times n$ con entradas de un campo F. Denotamos por L_A a la función $L_A : F^n \to F^m$ definida por $L_A(x) = Ax$ para todo vector columna $x \in F^n$. Llamamos a L_A una transformación de multiplicación por izquierda.

Teorema 2.14. *Sea A una matriz $m \times n$ con entradas de F. Entonces la transformación de multiplicación por izquierda $L_A : F_n \to F_m$ es lineal.*

Demostración. Es consecuencia inmediata del teorema 2.12. $\square$

Teorema 2.15. *Sea A una matriz $m \times n$ con entradas de F. Adicionalmente, si B es cualquier otra matriz $m \times n$ con entradas de F y β y γ son bases ordenadas estándar de F^n y F^m, respectivamente, entonces se cumple lo siguiente:*

(a) $[L_A]_\beta^\gamma = A$.

(b) $L_A = L_B$ *si y solo si* $A = B$.

(c) $L_{A+B} = L_A + L_B$ *y* $L_{aA} = aL_A$ *para todo* $a \in F$.

(d) *Si* $T : F^n \to F^m$ *es lineal, entonces exista una matriz única* C *tal que* $T = L_C$. *De hecho,* $C = [T]_\beta^\gamma$.

(e) *Si* E *es una matriz* $n \times p$, *entonces* $L_{AE} = L_A L_E$.

(f) *Si* $m = n$, *entonces* $L_{I_n} = \mathrm{id}_{F^n}$

Demostración.

(a) La j-ésima columna de $[L_A]_\beta^\gamma$ es igual a $L_A(e_j) = Ae_j$ por definición. Luego,

$$Ae_j = \begin{bmatrix} \sum_{k=1}^n A_{1k}(e_j)_k \\ \sum_{k=1}^n A_{2k}(e_j)_k \\ \vdots \\ \sum_{k=1}^n A_{mk}(e_j)_k \end{bmatrix}$$

Como $(e_j)_k = \delta_{jk}$, se obtiene que

$$Ae_j = \begin{bmatrix} A_{1j} \\ A_{2j} \\ \vdots \\ A_{mj} \end{bmatrix} = A_j$$

(b) Si $L_A = L_B$, entonces podemos usar el numeral anterior para escribir $A = [L_A]_\beta^\gamma = [L_B]_\beta^\gamma = B$. La prueba del converso es trivial.

(c) Sea x un vector de F^n. Por definición, $L_{A+B}(x) = (A+B)x$. De acuerdo al teorema 2.12, esto es es igual a $Ax + Bx$, o lo que es igual, $L_A(x) + L_B(x)$. Como x es un elemento genérico de F^n, se cumple que en general $L_{A+B} = L_A + L_B$. Por otra parte, $L_{aA}(x) = (aA)x$. Por el teorema 2.12 esto es igual a $A(ax) = L_A(ax)$ y esto por linealidad de L_A es igual a $aL_A(x)$.

(d) Sea $C = [T]_\beta^\gamma$. Por el teorema 2.13, $[T(x)]_\gamma = [T]_\beta^\gamma [x]_\beta$ o, alternativamente, $T(x) = Cx = L_C(x)$ para todo $x \in F^n$. Entonces, $T = L_C$. La unicidad de C es consecuencia de *2*.

(e) Tenemos que

$$L_{AE}(e_j) = (AE)e_j = \begin{bmatrix} \sum_{k=1}^p (AE)_{1k}(e_j)_k \\ \sum_{k=1}^p (AE)_{2k}(e_j)_k \\ \vdots \\ \sum_{k=1}^p (AE)_{mk}(e_j)_k \end{bmatrix}$$

Sin embargo, $(AE)_{ij} = \sum_{k=1}^n A_{ik}E_{kj}$, de modo que

$$L_{AE}(e_j) = \begin{bmatrix} \sum_{k=1}^p \left(\sum_{l=1}^n A_{1l}E_{lk}\right)(e_j)_k \\ \sum_{k=1}^p \left(\sum_{l=1}^n A_{2l}E_{lk}\right)(e_j)_k \\ \vdots \\ \sum_{k=1}^p \left(\sum_{l=1}^n A_{ml}E_{lk}\right)(e_j)_k \end{bmatrix} = \begin{bmatrix} \sum_{l=1}^n A_{1l}\left(\sum_{k=1}^p E_{lk}(e_j)_k\right) \\ \sum_{l=1}^n A_{2l}\left(\sum_{k=1}^p E_{lk}(e_j)_k\right) \\ \vdots \\ \sum_{l=1}^n A_{ml}\left(\sum_{k=1}^p E_{lk}(e_j)_k\right) \end{bmatrix}$$

$$= \begin{bmatrix} \sum_{l=1}^n A_{1l}(Ee_j)_l \\ \sum_{l=1}^n A_{2l}(Ee_j)_l \\ \vdots \\ \sum_{l=1}^n A_{ml}(Ee_j)_l \end{bmatrix} = A(Ee_j) = AL_E(e_j) = L_A L_E(e_j)$$

(f) Sea v un vector columna en F^n. Entonces,

$$L_{I_n}(v) = I_n v = \begin{bmatrix} \sum_{k=1}^n (I_n)_{1k} v_k \\ \sum_{k=1}^n (I_n)_{2k} v_k \\ \vdots \\ \sum_{k=1}^n (I_n)_{nk} b_k \end{bmatrix} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} = v$$

Como v es un elemento genérico de F^n, se cumple que en general $L_{I_n}(x) = x$ para todo $x \in F^n$, de modo que L_{I_n} coincide con id_{F^n}.

□

Teorema 2.16. *Sean A,B y C matrices tales que $A(BC)$ está definido. Entonces $(AB)C$ también está definido y $A(BC) = (AB)C$; esto es, la multiplicación de matrices es asociativa.*

Demostración. Si $(AB)C$ está definido es porque el número de columnas de AB es igual al número de filas de C. Adicionalmente, el número de columnas de AB viene dado por el número de columnas de B. Por transitividad, el número de columnas de B y el número de filas de C concuerdan, luego el producto BC está definido. Similarmente, como el producto AB está definido, el número de columnas de A coincide con el número de filas de B, el cual fija a su vez el número de filas del producto BC. Vemos entonces que el número de columnas de A es igual al número de filas de BC, por lo cual el producto $A(BC)$ está definido.
Por otra parte, usando la parte *5.* del teorema 2.15 y la asociatividad de la composición de funciones, tenemos que

$$L_{A(BC)} = L_A L_{BC} = L_A(L_B L_C) = (L_A L_B)L_C = L_{AB} L_C = L_{(AB)C}$$

Luego, por la parte *2.* del teorema 2.15, $A(BC) = (AB)C$. □

2.3. Invertibilidad e isomorfismos

Esta sección aplica los conceptos de invertibilidad de funciones a las transformaciones lineales. A su vez algunos resultados sobre invertibilidad serán aplicados al concepto de isomorfismo.

Def. Sean V y W espacios vectoriales y sea $T : V \to W$ una transformación lineal. Una función $U : W \to V$ se dice **inversa** de T si $TU : I_W$ y $UT = I_V$. Si T tiene una inversa, entonces T es **invertible**. Si T es invertible, su inversa es única y se denota por T^{-1}.

Una vez establecida la noción de invertibilidad para transformaciones lineales, podemos reescribir el teorema 2.6 en términos de invertibilidad.

Teorema 2.17 (2.6 revisado)**.** *Sea $T : V \to W$ una transformación lineal, donde V y W son espacios finitamente generados de dimensión igual. Entonces T es invertible si y solo si* $\mathrm{rank}(T) = \dim(V)$.

Teorema 2.18. *Sean V y W espacios vectoriales y sea $T : V \to W$ lineal e invertible. Entonces $T^{-1} : W \to V$ es lineal.*

Demostración. Sean $w_1, w_2 \in W$ y $c \in F$. Como T es sobreyectiva e inyectiva, existen vectores únicos v_1 y v_2 en V tales que $T(v_1) = w_1$ y $T(v_2) = w_2$. En consecuencia, $v_1 = T^{-1}(w_1)$ y $v_2 = T^{-1}(w_2)$; entonces

$$\begin{aligned} T^{-1}(cw_1 + w_2) &= T^{-1}[cT(v_1) + T(v_2)] = T^{-1}[T(cv_1 + v_2)] \\ &= cv_1 + v_2 = cT^{-1}(w_1) + T^{-1}(w_2) \end{aligned}$$

□

Ahora procedemos a definir la inversa de una matriz. Existe una clara analogía entre la inversa de una matriz y la de una transformación lineal.

Def. Sea A una matriz $n \times n$. Entonces A es **invertible** si existe una matriz B de tamaño $n \times n$ tal que $AB = BA = I$. De existir, B se conoce como matriz inversa de A.

Lema 2.19. *Sea T una transformación lineal invertible de V a W. Entonces V es finitamente generado si y solo si W es finitamente generado. En tal caso,* $\dim(V) = \dim(W)$.

Demostración. Supongamos que V es finitamente generado. Sea $\beta = \{x_1, x_2, \ldots, x_n\}$ una base de V. Por el teorema 2.3, $T(\beta)$ genera $\operatorname{im}(T) = W$; luego W es finitamente generado pues $T(\beta)$ es un conjunto finito y por el teorema 1.9 de Friedberg, Insel & Spence (p. 44) un espacio generado por un conjunto finito es de dimensión finita. Por otro lado, si W es finitamente generado entonces V también lo es por un argumento similar, usando T^{-1}.

Ahora supongamos que V y W son de dimensión finita. Como T es inyectiva y sobre, tenemos que

$$\operatorname{null}(T) = 0 \quad \text{y} \quad \operatorname{rank}(T) = \dim(\operatorname{im}(T)) = \dim(W)$$

Entonces por el teorema 2.4 (teorema de la dimensión), tenemos que $\dim(V) = \dim(W)$. □

Teorema 2.20. *Sean V y W espacios vectoriales finitamente generados con las bases ordenadas β y γ, respectivamente. Sea $T : V \to W$ lineal. Entonces T es invertible si y solo si $[T]_\beta^\gamma$ es invertible. Adicionalmente,* $[T^{-1}]_\gamma^\beta = \left([T]_\beta^\gamma\right)^{-1}$

Demostración. Supongamos que T es invertible. Entonces, por el lema anterior tenemos que $\dim(V) = \dim(W)$. Sea $n = \dim(V)$. Entonces $[T]_\beta^\gamma$ es una matriz $n \times n$. Ahora bien $T^{-1} : W \to V$ satisface que $TT^{-1} = \mathrm{id}_W$ y $T^{-1}T = \mathrm{id}_V$. Entonces

$$I_n = [\mathrm{id}_V]_\beta = [T^{-1}T]_\beta = [T^{-1}]_\gamma^\beta [T]_\beta^\gamma$$

Similarmente, $[T]_\beta^\gamma [T^{-1}]_\gamma^\beta = I_n$. Entonces $[T]_\beta^\gamma$ es invertible y $\left([T]_\beta^\gamma\right)^{-1} = [T^{-1}]_\gamma^\beta$.

Ahora, supongamos que $A = [T]_\beta^\gamma$ es invertible, de modo que existe una matriz de tamaño $n \times n$ B tal que $AB = BA = I_n$. Por el teorema 2.6 de Friedberg, Insel & Spence (p. 72), existe una transformación $U \in \mathcal{L}(W, V)$ tal que

$$U(w_j) = \sum_{i=1}^{n} B_{ij} v_i \qquad \text{para } j = 1, 2, \ldots, n$$

donde $\gamma = \{w_1, w_2, \ldots, w_n\}$ y $\beta = \{v_1, v_2, \ldots, v_n\}$. De aquí se obtiene que $[U]_\beta^\gamma = B$. Para mostrar que $U = T^{-1}$, es necesario notar que

$$[UT]_\beta = [U]_\gamma^\beta [T]_\beta^\gamma = BA = I_n = [\mathrm{id}_V]_\beta$$

por el teorema 2.11. Por lo tanto, $UT = \mathrm{id}_V$ y $TU = \mathrm{id}_W$. □

Def. Sean V y W espacios vectoriales. Decimos que V es **isomorfo** a W si existe una transformación lineal $T : V \to W$ que es invertible. Tal transformación lineal se denomina **isomorfismo** de V sobre W.

Teorema 2.21. *Sean V y W espacios vectoriales de dimensión finita (sobre el mismo campo). Entonces V es isomorfo a W si y solo si* $\dim(V) = \dim(W)$.

Demostración. Supongamos que V es isomorfo a W y que $T : V \to W$ es un isomorfismo de V a W. Por el lema 2.19, tenemos que $\dim(V) = \dim(W)$.

Ahora supongamos que $\dim(V) = \dim(W)$ y sean $\beta = \{v_1, v_2, \ldots, v_n\}$ y $\gamma = \{w_1, w_2, \ldots, w_n\}$ bases para V y W, respectivamente. Por el teorema 2.6 de Friedberg, Insel & Spence (p. 72), existe $T : V \to W$ tal que T es lineal y $T(v_i) = w_i$ para $i = 1, 2, \ldots, n$. A partir del teorema 2.3, tenemos que

$$\operatorname{im}(T) = \operatorname{span}(T(\beta)) = \operatorname{span}(\gamma) = W$$

Por lo tanto T es sobreyectiva. Por el teorema 2.6, tenemos que T también es inyectiva. Luego, T es un isomorfismo. □

Teorema 2.22. *Sean V y W espacios vectoriales de dimensión finita n y m, respectivamente, y sean β y γ bases ordenadas de V y W, respectivamente. Entonces la función $\Phi : \mathcal{L}(V, W) \to M_{m\times n}(F)$, definida por $\Phi(T) = [T]_\beta^\gamma$ para $T \in \mathcal{L}(V, W)$, es un isomorfismo.*

Demostración. Por el teorema 2.8, Φ es lineal. Resta demostrar que Φ es sobreyectiva e inyectiva. Esto se logra al mostrar que para cada matriz $m \times n$ A, existe una única transformación lineal $T : V \to W$ tal que $\Phi(T) = A$. Sean $\beta = \{v_1, v_2, \ldots, v_n\}$, $\gamma = \{w_1, w_2, \ldots, w_m\}$ y A una matriz $m \times n$ dada. Supongamos que T y U satisfacen que $\Phi(T) = A$. En tal caso, $T(v_j) = U(v_j) = \sum_{i=1}^m A_{ij} w_i$ para $i = 1, 2, \ldots, n$, y por el corolario al teorema 2.6 de Friedberg, Insel & Spence (p. 73), $T = U$. Por lo que T es única. □

Corolario 2.22.1. *Sean V y W espacios vectoriales de dimensión finita n y m, respectivamente. Entonces $\mathcal{L}(V, W)$ es de dimensión finita mn.*

Def. Sea β una base ordenada para un espacio vectorial sobre el campo F de dimensión n. La **representación estándar** de V con respecto a β es la función $\phi_\beta : V \to F^n$ definida por $\phi_\beta(v) = [v]_\beta$ para cada $v \in V$.

Teorema 2.23. *Para cualquier espacio vectorial V de dimensión finita con una base ordenada β, ϕ_β es un isomorfismo.*

Demostración. Basta con probar que ϕ_β es lineal y que su kernel consta únicamente del vector *0*, puesto que $\dim(V) = \dim(F^n)$.

ϕ_β ***es lineal:*** Sean $\beta = \{v_1, v_2, \ldots, v_n\}$, $x = \sum_{i=1}^n a_i v_i$ y $y = \sum_{i=1}^n b_i v_i$ vectores en V y $k \in F$ un escalar. Entonces,

$$\begin{aligned}
\phi_\beta(kx + y) &= \phi_\beta\left(k\left(\sum_{i=1}^n a_i v_i\right) + \sum_{i=1}^n b_i v_i\right) = \phi_\beta\left(\sum_{i=1}^n (ka_i + b_i)\, v_i\right) \\
&= \begin{bmatrix} ka_1 + b_1 \\ ka_2 + b_2 \\ \vdots \\ ka_n + b_n \end{bmatrix} = \begin{bmatrix} ka_1 \\ ka_2 \\ \vdots \\ ka_n \end{bmatrix} + \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix} = k \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} + \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix} \\
&= k\phi_\beta(x) + \phi_\beta(y)
\end{aligned}$$

$\ker(\phi_\beta) = \{\mathit{0}\}$***:*** Si $\phi_\beta(x) = (0, 0, \ldots, 0) \in F^n$, $x = \sum_{i=1}^n 0\, v_i$, de modo que $x = \mathit{0} \in V$.

□

2.4. Matriz de cambio de coordenadas

Teorema 2.24. *Sean β y β' dos bases ordenadas para una espacio vectorial de dimensión finita y sea $Q = [\mathrm{id}_V]_{\beta'}^{\beta}$. Entonces,*

(*a*) *Q es invertible*

(*b*) *Para cualquier $v \in V$, $[v]_\beta = Q[v]_{\beta'}$*

Demostración.

(*a*) Como id_V es invertible, Q también es invertible por el teorema 2.20.

(*b*) Para cualquier $v \in V$,
$$[v]_\beta = [\mathrm{id}_V(v)]_\beta = [\mathrm{id}_V]_{\beta'}^{\beta}[v]_{\beta'} = Q[v]_{\beta'}$$
por el teorema 2.13.

□

La matriz Q descrita en el teorema anterior se denomina matriz de cambio de coordenadas de coordenadas en β' a coordenadas en β.

Def. Sea V un espacio vectorial y $T : V \to V$ una transformación lineal. Entonces T se conoce como un **operador lineal** sobre V.

Teorema 2.25. *Sea T un operador lineal sobre un espacio vectorial de dimensión finita V y sean β y β' bases ordenadas de V. Supongamos que Q es la matriz de cambio de coordenadas de β' a β. Entonces*
$$[T]_{\beta'} = Q^{-1}[T]_\beta Q$$

Demostración. Sea id la transformación identidad en V. Entonces $T = \mathrm{id}\, T = T\, \mathrm{id}$; luego por el teorema 2.11,
$$Q[T]_{\beta'} = [\mathrm{id}]_{\beta'}^{\beta}[T]_{\beta'} = [\mathrm{id}\, T]_{\beta'}^{\beta} = [T\, \mathrm{id}]_{\beta'}^{\beta} = [T]_\beta[\mathrm{id}]_{\beta'}^{\beta} = [T]_\beta Q$$
En consecuencia, multiplicando ambos lados por Q^{-1} a la izquierda, $Q^{-1}Q[T]_{\beta'} = Q^{-1}[T]_\beta Q$. Como $Q^{-1}Q$ corresponde a la matriz identidad del tamaño apropiado y una matriz A multiplicada por derecha por la matriz identidad del tamaño apropiado es A, obtenemos que
$$[T]_{\beta'} = Q^{-1}[T]_\beta Q.$$

□

Def. Sean A y B matrices en $M_{n\times n}(F)$. Decimos que B es **similar** a A si existe una matriz invertible Q tal que $B = Q^{-1}AQ$.

3. Matrices y sistemas de ecuaciones lineales

Def. Sea A una matriz $m\times n$. Cualquiera de las siguientes tres operaciones sobre las filas (columnas) de A se llama una **operación elemental** sobre las filas (columnas)

1. intercambiando dos filas (columnas) cualesquiera de A;
2. multiplicando cualquier fila (columna) de A por un escalar distinto de cero;
3. sumar cualquier múltiplo escalar de una fila (columna) de A a otra fila (columna).

Una **matriz elemental** $n \times n$ es una matriz obtenida al realizar una operación elemental sobre I_n. La matriz elemental es de tipo 1, 2 o 3 de acuerdo a si la operación elemental realizada sobre I_n es de tipo 1, 2 o 3, respectivamente.

Teorema 3.1. *Sea $A \in M_{m\times n}(F)$ y supongamos que B se obtiene a partir de A al realizar una operación elemental sobre las filas (columnas). Entonces existe una matriz elemental $m\times m$ ($n\times n$) E tal que $B = EA$ ($B = AE$). De hecho E se obtiene de I_m (I_n) al realizar la misma operación elemental que fue realizada sobre A para obtener B. Recíprocamente, si E es una matriz elemental $m\times m$ ($n\times n$), entonces EA (AE) es la matriz obtenida a partir de A al realizar la misma operación elemental que produce E a partir de I_m (I_n).*

Demostración. La prueba requiere verificar el teorema para cada tipo de operación elemental.

Tipo 1: Sea A una matriz de tamaño $n \times m$ y sea B la matriz que se obtiene al intercambiar la filas k y l de A. Supongamos que existe una matriz E tal que $B = EA$. En tal caso, $B_{ij} = \sum_{p=1}^{m} E_{ip}A_{pj} = \sum_{p=1}^{m} E_{ip}a_{pj}$. Las condiciones $B_{ij} = A_{ij}$ para $i \neq k, l$, $B_{kj} = A_{lj}$ y $B_{lj} = A_{kj}$ sugieren que $E_{ij} = \delta_{ij}$ para $i \neq i.j$, $E_{kj} = \delta_{lj}$ y $E_{lj} = \delta_{kj}$. Esta descripción de las entradas de E coinciden con la matriz elemental de tipo 1 que resulta de intercambiar las filas k y l de I_m.

Tipo 2: Sea A una matriz de tamaño $n \times m$ y sea B la matriz que se obtiene al multiplicar la k-ésima fila de A por un escalar distinto de cero, h. Nuevamente supongamos que existe una matriz E tal que $B = EA$. Las condiciones $B_{ij} = A_{ij}$ para $i \neq k$ y $B_{kj} = hA_{kj}$ sugieren que $E_{ij} = \delta_{ij}$ para $i \neq k$ y $E_{kj} = h\delta_{kj}$. Esta descripción de las entradas de E coinciden con la matriz elemental de tipo 2 que resulta de multiplicar la i-ésima fila de I_m por h.

Tipo 3: Sea A una matriz de tamaño $n \times m$ y sea B la matriz que se obtiene al sumar h veces la k-ésima fila a la l-ésima fila de A. Supongamos que existe una matriz E tal que $B = EA$. Las condiciones $B_{ij} = A_{ij}$ para $i \neq l$ y $B_{lj} = A_{lj} + hA_{kj}$ sugieren que $E_{ij} = \delta_{ij}$ para $i \neq l$ y $B_{lj} = \delta_{lj} + h\delta_{kj}$.

Para verificar el teorema sobre las columnas basta con aplicar el procedimiento anterior a matrices transpuestas. □

Teorema 3.2. *Las matrices elementales son invertibles y la matriz inversa de una matriz elemental es una matriz elemental del mismo tipo.*

Demostración. Sea E una matriz elemental $n \times n$. Entonces E puede obtenerse al realizar una operación elemental sobre I_n. Al deshacer la operación mediante la cual se obtuvo E, podemos transformar E de vuelta a I_n. El resultado es que I_n puede obtenerse a partir E al realizar una operación elemental del mismo tipo. Por el teorema anterior (3.1), existe una matriz elemental $\overline{E}$ tal que $\overline{E}E = I_n$, luego E es invertible y $\overline{E} = E^{-1}$. □

Def. Si $A \in M_{n\times m}(F)$, definimos el **rank** de A, denotado $\text{rank}(A)$, como el rank de la transformación lineal $L_A : F^n \to F^m$.

Lema 3.3. *Sean V y W espacios vectoriales de dimensión finita y $T : V \to W$ un isomorfismo. Sea V_0 un subespacio de V. $T(V_0)$ es subespacio de W y $\dim(V_0) = \dim(T(V_0))$.*

Demostración. Para probar que $T(V_0)$ es subespacio de W, antes probaremos una condición más fuerte (ignorando la restricción que impone la existencia de un isomorfismo entre V y W): sea $U : V \to W$ una transformación lineal, entonces $U(V_0)$ es subespacio de W. Para demostrarlo consideramos, la restricción de U a V_0, denotada U_{V_0}. De acuerdo al teorema 2.2, $\text{im}(U_{V_0})$ es subespacio de W. Por construcción, $U(V_0) = \text{im}(U_{V_0})$, luego $U(V_0)$ es subespacio de W.

Para probar que $\dim(V_0) = \dim(T(V_0))$, consideramos una base de V_0, digamos $\beta = \{v_1, v_2, \ldots, v_n\}$. Todo vector $v \in V_0$ puede expresarse de forma única como una combinación lineal de los vectores en β:

$$v = \sum_{i=1}^{n} a_i v_i.$$

Entonces,

$$T(v) = T\left(\sum_{i=1}^{n} a_i v_i\right) = \sum_{i=1}^{n} a_i T(v_i).$$

Esto muestra que $T(\beta)$ genera a $T(V_0)$. Al ser isomorfismo, T es 1:1, de modo que $|T(\beta)| = n$. Adicionalmente, $T(\beta)$ debe ser linealmente independiente porque de lo contrario existen escalares b_i $(1 \leq i \leq n)$ no todos iguales a cero tales que $\mathit{0} = \sum_{i=1}^{n} b_i T(v_i) = T\left(\sum_{i=1}^{n} b_i v_i\right)$. Pero como T es isomorfismo, el único vector $x \in V$ tal que $T(x) = \mathit{0}$ es el vector $\mathit{0} \in V$. Luego, $\sum_{i=1}^{n} b_i v_i = \mathit{0}$, de modo que β no es base de V_0. En resumen, hemos probado que $T(\beta)$ es linealmente independiente, genera a $T(V_0)$ y tiene n vectores. Esto comprueba que $\dim(V_0) = \dim(T(V_0))$. □

Teorema 3.4. *Sea $T : V \to W$ una transformación lineal entre espacios de dimensión finita y sean β y γ bases ordenadas de V y W, respectivamente. Entonces* $\operatorname{rank}(T) = \operatorname{rank}\left([T]_\beta^\gamma\right)$.

Demostración. Esto es equivalente a probar que si $A = [T]_\beta^\gamma$, $\operatorname{rank}(T) = \operatorname{rank}(L_A)$. Partimos de que $\operatorname{im}(L_A) = \phi_\gamma(R(T))$. Aplicando el lema anterior, al ser $\operatorname{im}(T)$ un subespacio de W y ϕ_β un isomorfismo:

$$\begin{aligned}\dim(R(L_A)) &= \dim(\phi_\gamma(R(T))) = \dim(R(T)) \\ \operatorname{rank}(L_A) &= \operatorname{rank}(T)\end{aligned}$$

□

Teorema 3.5. *Sea A una matriz $m \times n$. Si P y Q son invertibles $m \times m$ y $n \times n$, respectivamente, entonces,*

(*a*) $\operatorname{rank}(PA) = \operatorname{rank}(A)$,

(*b*) $\operatorname{rank}(AQ) = \operatorname{rank}(A)$,

y por lo tanto,

c. $\operatorname{rank}(PAQ) = \operatorname{rank}(A)$.

Demostración.

(*a*) Notamos que

$$\operatorname{im}(L_{PA}) = \operatorname{im}(L_P L_A) = L_P(\operatorname{im}(L_A))$$

Como L_P es isomorfismo, por el lema 3.3, $\dim(\operatorname{im}(L_A)) = \dim(L_P(\operatorname{im}(L_A)))$. Entonces, $\operatorname{rank}(PA) = \operatorname{rank}(L_{PA}) = \operatorname{rank}(L_A) = \operatorname{rank}(A)$.

(*b*) Primero, observemos que

$$\operatorname{im}(L_{AQ}) = \operatorname{im}(L_A L_Q) = L_A L_Q(F^n) = L_A(L_Q(F^n)).$$

Como L_Q es sobreyectiva, esto es igual a $L_A(F^n) = \operatorname{im}(L_A)$. Luego,

$$\operatorname{rank}(AQ) = \operatorname{rank}(L_{AQ}) = \operatorname{rank}(L_A) = \operatorname{rank}(A)$$

(c) Aplicando los dos numerales anteriores:

$$\text{rank}(PAQ) = \text{rank}(PA) = \text{rank}(A)$$

$\square$

Corolario 3.5.1. *Las operaciones elementales sobre las filas y columnas de una matriz preservan el rank.*

Demostración. Si B se obtiene de A a través de una operación elemental sobre las filas (columnas), entonces existe una matriz elemental E tal que $B = EA$ $(B = AE)$. Por el teorema 3.2, E es invertible, entonces $\text{rank}(B) = \text{rank}(EA) = \text{rank}(A)$ $(\text{rank}(B) = \text{rank}(AE) = \text{rank}(A))$. $\square$

Teorema 3.6. *El rank de cualquier matriz es igual al máximo número de sus columnas linealmente independientes; esto es, el rank de una matriz es la dimensión del subespacio generado por sus columnas.*

Demostración. Para cualquier $A \in M_{m\times n}(F)$,

$$\text{rank}(A) = \text{rank}(L_A) = \dim(R(L_A))$$

Sea β la base estándar ordenada de F^n. Entonces β genera a F^n, luego, por el teorema 2.2 de Friedberg, Insel & Spence (p. 68),

$$\text{im}(L_A) = \text{span}(L_A(\beta)) = \text{span}(\{L_A(e_1), L_A(e_2), \ldots, L_A(e_n)\}).$$

Por el teorema 2.13.b) de Friedberg, Insel & Spence (p. 90), para cualquier j, $L_A(e_j) = Ae_j = a_j$, donde a_j es la j-ésima columna de A. Entonces,

$$\text{im}(L_A) = \text{span}(\{a_1, a_2, \ldots, a_n\}).$$

Por lo tanto,

$$\text{rank}(A) = \dim(\text{im}(L_A)) = \dim(\text{span}(\{a_1, a_2, \ldots, a_n\})).$$

$\square$

Teorema 3.7. *Sea A una matriz $m \times n$ de rank r. Entonces $r \leq m$, $r \leq n$ y, a través de un número finito de operaciones elementales sobre las filas y las columnas, A puede transformarse en la matriz*

$$D = \begin{bmatrix} I_r & O_1 \\ O_2 & O_3 \end{bmatrix},$$

donde O_1, O_2 y O_3 son matrices de ceros. Por lo tanto, $D_{ii} = 1$ para $i \leq r$ y $D_{ij} = 0$ en cualquier otro caso.

Demostración. Si A es la matriz de ceros, $r = 0$, pues la dimensión de L_A es el espacio nulo. En este caso la conclusión se obtiene con $D = A$.

Ahora supongamos que $A \neq O$ y $r = \text{rank}(A)$; entonces $r > 0$. Procedemos a demostrar el teorema por inducción sobre m, el número de filas de A.

Supongamos que $m = 1$. A través de a lo sumo una operación de tipo 1 sobre las columnas y a lo sumo una operación de tipo 2, A puede transformarse en una matriz con 1 en la posición 1,1. Ahora, con a lo sumo $n - 1$ operaciones de tipo 3 sobre las columnas, esta matriz puede transformarse en la matriz

$$\begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \end{bmatrix}.$$

Nótese que hay una columna (y una fila) linealmente independiente en esta matriz. Por lo tanto su rank, que es igual al de A por el corolario 3.5.1, es 1 por el teorema 3.6. Esto establece el teorema para $m = 1$.

Ahora, asumimos que el teorema es válido para cualquier matriz con máximo $m-1$ filas (para algún $m > 1$). Veamos que el teorema vale para cualquier matriz con m filas. Supongamos que A es una matriz $m \times n$. Si $n = 1$, el teorema se establece de manera análoga a como se hizo para $m = 1$.

Ahora, si $n > 1$, como $A \neq O$, $A_{ij} \neq 0$ para algún para i, j. A través de a lo sumo una operación elemental sobre las filas y a lo sumo una operación elemental sobre las columnas, cada una de tipo 1, podemos mover la entrada distinta de cero a la posición 1,1. Luego, a través de a lo sumo una operación adicional de tipo 2, se obtiene un 1 en la posición 1,1. A través de a lo sumo $m-1$ operaciones elementales de tipo 3 sobre las filas y a lo sumo $n-1$ operaciones elementales de tipo 3 sobre las columnas, podemos hacer nulas todas las entradas de la primera fila y la primera columna, excepto la entrada 1,1. Así, A puede transformarse en una matriz

$$B = \left[\begin{array}{c|ccc} 1 & 0 & \cdots & 0 \\ \hline 0 & & & \\ \vdots & & B' & \\ 0 & & & \end{array}\right]$$

Por hipótesis de inducción, B' puede transformarse en una matriz D' tal que

$$D' = \begin{bmatrix} I_{r-1} & O_4 \\ O_5 & O_6 \end{bmatrix},$$

donde O_4, O_5 y O_6 son matrices de ceros. Sea

$$D = \left[\begin{array}{c|ccc} 1 & 0 & \cdots & 0 \\ \hline 0 & & & \\ \vdots & & D' & \\ 0 & & & \end{array}\right].$$

Podemos ver que D puede obtenerse de B a través de un número finito de operaciones elementales. De aquí, es fácil ver que D es una matriz como la descrita en el teorema y como A puede transformarse en B y B a su vez puede transformarse en D, cada una a través de un número finito de operaciones elementales, A puede transformarse en D a través de un número finito de operaciones elementales. $\square$

Corolario 3.7.1. *Sea A una matriz $m \times n$ de rank r. Entonces existen matrices invertibles B y C de tamaños $m \times m$ y $n \times n$, respectivamente, tales que $D = BAC$, donde*

$$D = \begin{bmatrix} I_r & O_1 \\ O_2 & O_3 \end{bmatrix}$$

es la matriz $m \times n$ en la que O_1, O_2 y O_3 son matrices de ceros.

Demostración. Por el teorema 3.7, A puede transformarse en D a través de un número finito de operaciones elementales sobre las filas o columnas de A. Podemos aplicar el teorema 3.1 cada vez que se realiza una operación elemental. Entonces, existen matrices elementales $m \times m$ $E_1, E_2, \ldots, E_p$ y matrices elementales $n \times n$ $G_1, G_2, \ldots, G_q$ tales que

$$D = E_p E_{p-1} \cdots E_2 E_1 A G_1 G_2 \cdots G_q.$$

Por el teorema 3.2, cada E_j y G_j es invertible. Sea $B = E_p E_{p-1} \cdots E_2 E_1$ y $C = G_1 G_2 \cdots G_q$. Entonces B y C son invertibles, puesto que el producto de matrices invertibles es invertible. $\square$

Corolario 3.7.2. *Sea A una matriz $m \times n$. Entonces,*

(a) $\operatorname{rank}(A) = \operatorname{rank}(A^t)$,

(b) *el rank de una matriz es el máximo número de sus filas linealmente independientes; esto es el rank de una matriz es la dimensión del espacio generado por sus filas,*

(c) *Las filas y las columnas de cualquier matriz A generan subespacios de la misma dimensión, igual al rank de la matriz A.*

Demostración.

(a) Por el corolario anterior, existen matrices invertibles B y C tales que $D = BAC$, donde D satisface las condiciones establecidas en el corolario. Al tomar transpuestas tenemos que

$$D^t = (BAC)^t = C^t A^t B^t.$$

Como B y C son invertibles, B^t y C^t también lo son. Luego, por el teorema 3.5,

$$\operatorname{rank}(A^t) = \operatorname{rank}(C^t A^t B^t) = \operatorname{rank}(D^t).$$

Supongamos que $r = \operatorname{rank}(A)$. Entonces D^t es una matriz $n \times m$ de la forma de la matriz D en el corolario anterior, por lo tanto $\operatorname{rank}(D^t) = r$. Por el teorema 3.6,

$$\operatorname{rank}(A^t) = \operatorname{rank}(D^t) = r = \operatorname{rank}(A).$$

(b) De a. sabemos que $\operatorname{rank}(A) = \operatorname{rank}(A^t)$. Por el teorema 3.6, $\operatorname{rank}(A^t)$ es la dimensión del subespacio generado por las columnas de A^t, pero las columnas de A^t son las filas A.

(c) A partir de b. sabemos que el rank de A es la dimensión del subespacio generado por sus filas, pero el rank de A también es la dimensión del subespacio generado por sus columnas por el teorema 3.6. En consecuencia, la dimensión de estos dos subespacios debe ser igual.

□

Corolario 3.7.3. *Toda matriz invertible es producto de matrices elementales.*

Demostración. Si A es una matriz invertible $n \times n$ entonces $\operatorname{rank}(A) = n$. Entonces, por el corolario 3.7.1, existen matrices invertibles B y C tales que $I_n = BAC$. Como en la prueba del corolario 3.7.1, B y C pueden expresarse como productos de matrices elementales, $E_1, E_2, \ldots, E_p, G_1, G_2, \ldots, G_q$, así

$$B = E_p E_{p-1} \cdots E_1,$$
$$C = G_1 G_2 \cdots G_q.$$

Luego, $A = B^{-1} I_n C^{-1} = B^{-1} C^{-1}$. Por lo tanto,

$$A = E_1^{-1} E_2^{-1} \cdots E_p^{-1} G_q^{-1} G_{q-1}^{-1} \cdots G_1.$$

Como las inversas de matrices elementales son matrices elementales, se verifica el corolario. □

Teorema 3.8. *Sean $T : V \to W$ y $U : W \to Z$ transformaciones lineales sobre espacios de dimensión finita V, W y Z. Sean además A y B matrices tales que el producto AB está definido. Entonces,*

(a) $\operatorname{rank}(UT) \leq \operatorname{rank}(U)$,

(b) $\operatorname{rank}(UT) \leq \operatorname{rank}(T)$,

(c) $\operatorname{rank}(AB) \leq \operatorname{rank}(A)$,

(d) $\operatorname{rank}(AB) \leq \operatorname{rank}(B)$

Demostración. Probaremos las afirmaciones en el siguiente orden: a, c, d y b.

(a) Es evidente que $\operatorname{im}(T) \subseteq W$. Luego,

$$\operatorname{im}(UT) = UT(V) = U(T(V)) = U(\operatorname{im}(T)) \subseteq U(W) = \operatorname{im}(U).$$

Por lo tanto,

$$\operatorname{rank}(UT) = \dim(\operatorname{im}(UT)) \leq \dim(\operatorname{im}(U)) = \operatorname{rank}(U).$$

(c) Por a,

$$\operatorname{rank}(AB) = \operatorname{rank}(L_{AB}) = \operatorname{rank}(L_A L_B) \leq \operatorname{rank}(L_A) = \operatorname{rank}(A)$$

(d) Por c y el corolario 3.7.2,

$$\operatorname{rank}(AB) = \operatorname{rank}((AB)^t) = \operatorname{rank}(B^t A^t) \leq \operatorname{rank}(B^t) = \operatorname{rank}(B).$$

(b) Sean α, β y γ bases ordenadas de V, W y Z, respectivamente, y sean $A' = [U]_\beta^\gamma$ y $B' = [T]_\alpha^\beta$. Entonces, $A'B' = [UT]_\alpha^\gamma$ por el teorema 2.11. Entonces por el teorema 3.4 y d,

$$\operatorname{rank}(UT) = \operatorname{rank}(A'B') \leq \operatorname{rank}(B') = \operatorname{rank}(T).$$

□

3.1. Sistemas de ecuaciones lineales - aspectos teóricos

Definiciones. El sistema de ecuaciones

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

donde a_{ij} y b_i ($1 \leq i \leq m$ y $1 \leq j \leq n$) son escalares en un campo F y $x_1, x_2, \ldots, x_n$ son n variables con valores en F, se llama un **sistema de m ecuaciones lineales en n incógnitas sobre el campo F.**

La matriz $m \times n$

$$A = \begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{bmatrix}$$

se denomina **matriz de coeficientes** del sistema (S).

Si tomamos

$$x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \qquad \text{y} \qquad b = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix},$$

el sistema (S) puede escribirse como una única ecuación matricial

$$Ax = b.$$

Una **solución** al sistema (S) es una n-tupla

$$s = \begin{pmatrix} s_1 \\ s_2 \\ \vdots \\ s_n \end{pmatrix} \in F^n$$

tal que $As = b$ (usamos paréntesis para s para denotando que es una tupla en F^n escrita como vector columna). El conjunto de todas las soluciones al sistema (S) se denomina **conjunto solución** del sistema. El sistema (S) se dice **consistente** si su conjunto solución es no vacío; de lo contrario se dice **inconsistente**.

Def. Un sistema $Ax = b$ de m ecuaciones lineales en n incógnitas es **homogéneo** si $b = \mathit{0}$. De lo contrario, es **no homogéneo**.

Teorema 3.9. *Sea $Ax = \mathit{0}$ un sistema homogéneo de m ecuaciones en n incógnitas sobre un campo F. Sea K el conjunto solución del sistema. Entonces, $K = \ker(L_A)$. Por lo tanto, K es un subespacio de F^n de dimensión $n - \operatorname{rank}(L_A) = n - \operatorname{rank}(A)$.*

Demostración. Claramente, $K = \{s \in F^n : As = \mathit{0}\} = \ker(L_A)$. La segunda parte es consecuencia del teorema de la dimensión (2.4). $\square$

Corolario 3.9.1. *Si $m < n$, el sistema $Ax = \mathit{0}$ tiene una solución distinta de cero.*

Demostración. Supongamos que $m < n$. Entonces $\operatorname{rank}(A) = \operatorname{rank}(L_A) \leq m$. Así,

$$\dim(K) = n - \dim(L_A) \geq n - m > 0,$$

donde $K = \ker(L_A)$. Como $\dim(K) > 0$, $K \neq \varnothing$ y en consecuencia, existe un vector no nulo $s \in K$, de modo que s es una solución distinta de cero a $Ax = \mathit{0}$. $\square$

Ahora, centramos nuestro estudio en sistemas no homogéneos.

Def. Sea $Ax = b$ un sistema no homogéneo de m ecuaciones en n incógnitas sobre un campo F. El sistema $Ax = \mathit{0}$ se denomina **sistema homogéneo correspondiente a** $Ax = b$.

Teorema 3.10. *Sea K el conjunto solución de un sistema de ecuaciones lineales $Ax = b$ y sea K_H es conjunto solución del sistema homogéneo correspondiente $Ax = \mathit{0}$. Entonces, para cualquier solución s de $Ax = b$*

$$K = \{s\} + K_H = \{s + k : k \in K_H\}.$$

Demostración. Sea s cualquier solución a $Ax = b$. Veamos que $K = \{s\} + K_H$. Sea $w \in K$, por ende $Aw = b$. Luego,

$$A(w - s) = Aw - As = b - b = \mathit{0}.$$

De aquí concluimos que $w - s \in K_H$. Por lo tanto, existe $k \in K_H$ tal que $w - s = k$. Se sigue que $w = s + k \in s + K_H$ y por lo tanto, $K \subseteq \{s\} + K_H$.

Recíprocamente, supongamos que $w \in \{s\} + K_H$, entonces $w = s + k$ para algún $k \in K_H$. Así, $Aw = A(s + k) = As + Ak = b + \mathit{0} = b$ y en consecuencia $w \in K$. Esto nos permite concluir que $\{s\} + K_H \subseteq K$ y en últimas que $K = \{s\} + K_H$. $\square$

Teorema 3.11. *Sea $Ax = b$ un sistema de n ecuaciones lineales en n incógnitas. El sistema tiene una única solución si y solo si A es invertible.*

Demostración. Supongamos que el sistema tiene exactamente una solución, digamos s. Sea K_H el conjunto solución para el sistema homogéneo correspondiente $Ax = \mathit{0}$. Por el teorema 3.10, $\{s\} = \{s\} + K_H$, pero esto solo se cumple si $K_H = \{\mathit{0}\}$ y esto por el teorema 2.6 implica que L_A es invertible y, por extensión, A también lo es.

Recíprocamente, supongamos que A es invertible y s es una solución arbitraria del sistema, entonces $As = b$. Multiplicando el sistema a izquierda por A^{-1} llegamos que $s = A^{-1}b$. Por lo tanto, el sistema tiene una única solución, a saber $A^{-1}b$. □

Def. Sean A y B matrices de tamaño $m \times n$ y $m \times p$ respectivamente. La matriz aumentada $[A|B]$ es la matriz de tamaño $m \times (n+p)$ cuyas n primeras columnas son las columnas de A y cuyas p últimas columnas son las de B.

Teorema 3.12. *Sea $Ax = b$ un sistema de ecuaciones lineales. Entonces el sistema es consistente si y solo si $\operatorname{rank}(A) = \operatorname{rank}(A|b)$.*

Demostración. Decir que $Ax = b$ tiene solución es equivalente a afirmar que $b \in \operatorname{im}(L_A)$. En la prueba del teorema 3.6 vimos que $\operatorname{im}(L_a) = \operatorname{span}(\{a_1, a_2, \ldots, a_n\})$, el espacio generado por las columnas de A. Por ende, $Ax = b$ tiene solución si y solo si $b \in \operatorname{span}(\{a_1, a_2, \ldots, a_n\})$. Pero $b \in \operatorname{span}(\{a_1, a_2, \ldots, a_n\})$ si y solo si $\operatorname{span}(\{a_1, a_2, \ldots, a_n\}) = \operatorname{span}(\{a_1, a_2, \ldots, a_n, b\})$. Esta última afirmación es equivalente a

$$\dim(\operatorname{span}(\{a_1, a_2, \ldots, a_n\})) = \dim(\operatorname{span}(\{a_1, a_2, \ldots, a_n, b\})).$$

Por el teorema 3.6, esto se reduce a

$$\operatorname{rank}(A) = \operatorname{rank}(A|b).$$

□

3.2. Sistemas de ecuaciones lineales - aspectos computacionales

El procedimiento para hallar la solución de un sistema de ecuaciones lineales es bastante directo e involucra usar operaciones elementales para obtener sistemas de ecuaciones equivalente pero más sencillos de resolver.

Def. Dos sistemas de ecuaciones lineales se dicen **equivalentes** si tienen el mismo conjunto solución.

Teorema 3.13. *Sea $Ax = b$ un sistema de m ecuaciones lineales en n incógnitas y sea C una matriz invertible $m \times m$. Entonces, el sistema $(CA)x = Cb$ es equivalente a $Ax = b$.*

Demostración. Sea K el conjunto solución para $Ax = b$ y sea K' el conjunto solución para $(CA)x = Cb$. Si $s \in K$, entonces $As = b$. Multiplicando a izquierda ambos lados de la igualdad por C, $(CA)s = Cb$, y así $s \in K'$. Por consiguiente, $K \subseteq K'$.

Por el contrario, si $s \in K'$, entonces $(CA)s = Cb$. Multiplicando a izquierda ambos lados de esta igualdad por C^{-1},

$$\begin{aligned} C^{-1}[(CA)s] &= C^{-1}(Cb) \\ As &= b, \end{aligned}$$

de modo que $s \in K$. Por ende $K' \subseteq K$ y así $K = K'$. □

Corolario 3.13.1. *Sea $Ax = b$ un sistema de m ecuaciones lineales en n incógnitas. Si $(A'|b')$ se obtiene de $(A|b)$ a través de un número finito de operaciones elementales sobre las filas, entonces el sistema $A'x = b'$ es equivalente al sistema original.*

Demostración. Supongamos que $(A'|b')$ se obtiene de $(A|b)$ a través de operaciones elementales sobre las filas. Estas se obtienen al multiplica $(A|b)$ por matrices elementales de tamaño $m \times m$ $E_1, E_2, \ldots, E_p$. Sea $C = E_p \cdots E_2 E_1$; entonces,

$$(A'|b') = C(A|b) = (CA|Cb).$$

Como cada matriz E_i para $i = 1, 2, \ldots, p$ es invertible, C también lo es. $A' = CA$ y $b' = Cb$. Así, por el teorema 3.13, el sistema $A'x = b'$ es equivalente al sistema $Ax = b$. □

Este teorema y su corolario nos permiten transformar cualquier matriz en una más simple, a partir de la cual se puede calcular la solución.

Def. Una matriz se encuentra en **forma escalonada reducida por filas** si se cumplen las siguientes condiciones.

i. Cualquier fila que contiene una entrada no nula precede a cualquier fila en la cual todas las entradas son cero.

ii. La primera entrada no nula en cada fila es 1 y está en una columna a la derecha de la primera entrada no nula en la fila precedente.

iii. La primera entrada no nula en cada fila es la única entrada no nula en su columna.

Si una matriz cumple las condiciones i y ii, se dice que está en **forma escalonada por filas**.

Hay dos métodos en los que se emplea la forma reducida para solucionar el sistema de ecuaciones.

Eliminación gaussiana

Si en algún punto del siguiente procedimiento se obtiene una fila nula, ubicarla como última fila usando operaciones elementales de tipo 1.

1. Se ubica un 1 en la posición 1,1 ya sea a través de una operación elemental de tipo 1 o una operación de tipo 2.

2. Se suman múltiplos de la primera fila a las demás para obtener ceros en las demás posiciones de la primera columna.

3. Se multiplica la siguiente fila no nula por una constante, tal que la primera entrada distinta de cero sea 1.

4. Se suman múltiplos de la fila sobre la cual se opero en el paso anterior a las filas debajo de ella, de modo que se obtengan ceros debajo de la primera entrada no nula de la fila operada.

5. Se repiten los pasos 3 y 4 hasta obtener una matriz triangular superior.

6. Sumar múltiplos apropiados de la última fila no nula a las filas superiores para ubicar ceros en todas las filas superiores sobre las entradas no nulas.

7. Repetir el paso 6 para cada fila precedente hasta obtener una matriz en forma escalonada reducida.

Reducción de Gauss-Jordan

1. Se ubica un 1 en la posición 1,1 ya sea a través de una operación elemental de tipo 1 o una operación de tipo 2.
2. Se suman múltiplos de la primera fila a las demás para obtener ceros en las demás posiciones de la primera columna.
3. Se multiplica la siguiente fila no nula por una constante, tal que la primera entrada distinta de cero sea 1.
4. Se suman múltiplos de la fila sobre la cual se opero en el paso anterior a las filas debajo de ella, de modo que se obtengan ceros en la columna de la primera entrada no nula de la fila operada.
5. Se repiten los pasos 3 y 4 hasta obtener una matriz en forma escalonada reducida.

Tanto la eliminación gaussiana como la reducción de Gauss-Jordan transforman cualquier matriz en su forma escalonada reducida por filas.

El siguiente ejemplo ilustra ambos métodos.

Ejemplo Consideremos el sistema de ecuaciones lineales

$$\begin{array}{rcrcrcrcr} 3x_1 & + & 2x_2 & + & 3x_3 & - & 2x_4 & = & 1 \\ x_1 & + & x_2 & + & x_3 & & & = & 3 \\ x_1 & + & 2x_2 & + & x_3 & - & x_4 & = & 2. \end{array}$$

A este sistema le corresponde la matriz aumentada

$$\left[\begin{array}{rrrr|r} 3 & 2 & 3 & -2 & 1 \\ 1 & 1 & 1 & 0 & 3 \\ 1 & 2 & 1 & -1 & 2 \end{array}\right]$$

Ambos métodos comienzan igual. Para el primer paso, intercambiamos las filas 1 y 3:

$$\left[\begin{array}{rrrr|r} 1 & 2 & 1 & -1 & 2 \\ 1 & 1 & 1 & 0 & 3 \\ 3 & 2 & 3 & -2 & 1 \end{array}\right].$$

Ahora, sumamos -1 veces la primera fila a la segunda y -3 veces la primera a la tercera.

$$\left[\begin{array}{rrrr|r} 1 & 2 & 1 & -1 & 2 \\ 0 & -1 & 0 & 1 & 1 \\ 0 & -4 & 0 & 1 & -5 \end{array}\right].$$

Ahora multiplico la segunda fila por -1 para obtener un 1 en su primera entrada no nula.

$$\left[\begin{array}{rrrr|r} 1 & 2 & 1 & -1 & 2 \\ 0 & 1 & 0 & -1 & -1 \\ 0 & -4 & 0 & 1 & -5 \end{array}\right].$$

Aquí es donde divergen los dos métodos. A la izquierda procedo por eliminación gaussiana y a la derecha por reducción de Gauss-Jordan.

Para eliminación gaussiana y para Gauss-Jordan, sumo 4 veces la segunda fila a la tercera. Además, para Gauss-Jordan sumo -2 veces la segunda fila a la primera.

$$\left[\begin{array}{rrrr|r} 1 & 2 & 1 & -1 & 2 \\ 0 & 1 & 0 & -1 & -1 \\ 0 & 0 & 0 & -3 & -9 \end{array}\right]; \qquad \left[\begin{array}{rrrr|r} 1 & 0 & 1 & 1 & 4 \\ 0 & 1 & 0 & -1 & -1 \\ 0 & 0 & 0 & -3 & -9 \end{array}\right].$$

Ahora multiplico la tercera fila por $-\frac{1}{3}$, en ambos casos.

$$\left[\begin{array}{cccr|r} 1 & 2 & 1 & -1 & 2 \\ 0 & 1 & 0 & -1 & -1 \\ 0 & 0 & 0 & 1 & 3 \end{array}\right]; \qquad \left[\begin{array}{cccr|r} 1 & 0 & 1 & 1 & 4 \\ 0 & 1 & 0 & -1 & -1 \\ 0 & 0 & 0 & 1 & 3 \end{array}\right].$$

Para Gauss-Jordan basta con sumar la tercera fila a la segunda y restar la tercera a la primera. Para el otro método en vez de restar, se suma la tercera a la primera y después es necesario sumar -2 veces la segunda fila a la primera.

$$\left[\begin{array}{cccc|c} 1 & 2 & 1 & 0 & 5 \\ 0 & 1 & 0 & 0 & 2 \\ 0 & 0 & 0 & 1 & 3 \end{array}\right]; \qquad \left[\begin{array}{cccc|c} 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 0 & 2 \\ 0 & 0 & 0 & 1 & 3 \end{array}\right];$$
$$\left[\begin{array}{cccc|c} 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 0 & 2 \\ 0 & 0 & 0 & 1 & 3 \end{array}\right].$$

Ambos métodos llegan a la misma forma escalonada reducida por filas. Aunque el procedimiento de Gauss-Jordan tiene menos pasos, este involucra más procesos aritméticos. Por lo tanto, es preferible usar el método de eliminación gaussiana.

Ya sabemos cómo calcular la forma escalonada reducida de una matriz. Ahora, corresponde interpretar la forma escalonada reducida en términos del conjunto solución al sistema de ecuaciones.

Def. Sea A una matriz en forma escalonada reducida por filas. La primera entrada distinta de cero en una fila es un pivote.

Para resolver un sistema de ecuaciones en el que la matriz aumentada está en forma escalonada reducida por filas, dividimos las variables en dos conjuntos. Un conjunto que corresponde a los pivotes de la forma escalonada reducida por filas y otro que corresponde a las demás variables. A cada variable en el segundo caso, le asignamos un valor paramétrico y resolvemos las variables en el primer conjunto en términos de las variables en el segundo. Para el ejemplo anterior, la forma escalonada reducida encontrada corresponde al sistema

$$\begin{array}{ccccccc} x_1 & & + & x_3 & & = & 1 \\ & x_2 & & & & = & 2 \\ & & & & x_4 & = & 3. \end{array}$$

Los pivotes corresponden a las variables x_1, x_2 y x_4. A la variable restante x_3 le asociamos un valor paramétrico t. Así, toda solución es de la forma

$$\begin{pmatrix} 1-t \\ 2 \\ t \\ 3 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \\ 0 \\ 3 \end{pmatrix} + t \begin{pmatrix} -1 \\ 0 \\ 1 \\ 0 \end{pmatrix}.$$

El primer vector de la suma corresponde a una solución particular y el segundo vector corresponde a un vector en el núcleo de L_A.

En general, el vector b' corresponde de algún modo a la solución particular. La correspondencia la dicta la posición de los pivotes en la forma escalonada reducida por filas. Por otra parte, la base del núcleo está relacionada con las columnas que no contienen pivotes.

4. Determinantes

Existen dos definiciones del determinante de una matriz. Ambas conducen a una función con las mismas propiedades y el mismo valor para una matriz dada.

Def. Sea S_n el conjunto de todas las biyecciones de $\{1, 2, \ldots, n\}$ a $\{1, 2, \ldots, n\}$. El **determinante** es la función $\det : M_{n\times n}(F) \to F$ definida por

$$\det(A) = \sum_{\sigma \in S_n} \operatorname{sign}(\sigma) A_{1\sigma(1)} A_{2\sigma(2)} \cdots A_{n\sigma(n)},$$

donde $\operatorname{sign}(\sigma)$ es la función signo de una permutación (Álgebra Abstracta 1), definida por

$$\operatorname{sign}(\sigma) = \begin{cases} 1 & \text{si } \sigma \text{ es producto de un número par de transposiciones} \\ -1 & \text{si } \sigma \text{ es producto de un número impar de transposiciones} \end{cases}$$

Observación. Esta definición garantiza que en cada término de la sumatoria no hay dos entradas en la misma fila ni en la misma columna. Por lo tanto, esta definición es equivalente a

$$\det(A) = \sum_{\sigma \in S_n} \operatorname{sign}(\sigma) A_{\sigma(1)1} A_{\sigma(2)2} \cdots A_{\sigma(n)n}.$$

Def (alternativa)**.** Definimos el determinante de una matriz 1×1 como la única entrada de la matriz. Inductivamente, definimos el **determinante** de una matriz de tamaño $n \times n$, A, así:

$$\det(A) = \sum_{j=1}^{n} (-1)^{i+j} A_{ij} \cdot \det(\tilde{A}_{ij}),$$

donde $\tilde{A}_{ij}$ es la matriz (llamada **matriz menor**) de tamaño $(n-1) \times (n-1)$ obtenida al eliminar la fila i y la columna j de A. El escalar $(-1)^{i+j} A_{ij} \cdot \det(\tilde{A}_{ij})$ se llama **cofactor** entrada i, j de A.

Observación. Nótese que el determinante de una matriz 2×2, B, es simplemente $\det(B) = B_{11}B_{22} - B_{12}B_{21}$. Además, en la definición anterior tenemos la libertad de fijar la fila (i) a partir de la cual calculamos el determinante.

Cualquiera de las dos definiciones anteriores es suficiente para desarrollar el estudio que pretendemos del determinante.

Observación. El determinante se puede ver también como una función de n vectores en F^n, a saber las n columnas de la matriz. Así, si $A \in M_{n\times n}(F)$ y $a_1, a_2, \ldots, a_n \in F^n$ son las columnas de A,

$$\det : \underbrace{F^n \times F^n \times \cdots \times F^n}_{n \text{ veces}} \to F$$
$$(a_1, a_2, \ldots, a_n) \mapsto \det(A)$$

Teorema 4.1. *Sea $A \in M_{n\times n}(F)$ una matriz con columnas $a_1, a_2, \ldots, a_n$ y b un vector en F^n. Sea $c \in F$. Entonces,*

(*a*) $\det(a_1, a_2, \ldots, a_j + b, \ldots, a_n) = \det(a_1, a_2, \ldots, a_j, \ldots, a_n) + \det(a_1, a_2, \ldots, b, \ldots, a_n)$,

(*b*) $\det(a_1, a_2, \ldots, ca_j, \ldots, a_n) = c \det(a_1, a_2, \ldots, a_j, \ldots, a_n)$,

(*c*) $\det(a_1, a_2, \ldots, a_i, \ldots, a_j, \ldots, a_n) = -\det(a_1, a_2, \ldots, a_j, \ldots, a_i, \ldots, a_n)$, *con $i < j$.*

Demostración. La prueba es directa.

(*a*)

$$\det(a_1, a_2, \ldots, a_j + b, \ldots, a_n) = \sum_{\sigma \in S_n} \operatorname{sign}(\sigma) A_{\sigma(1)1} A_{\sigma(2)2} \cdots (A_{\sigma(j)j} + b_{\sigma(j)}) \cdots A_{\sigma(n)n}$$
$$= \sum_{\sigma \in S_n} \operatorname{sign}(\sigma) A_{\sigma(1)1} A_{\sigma(2)2} \cdots A_{\sigma(j)j} \cdots A_{\sigma(n)n} + \sum_{\sigma \in S_n} \operatorname{sign}(\sigma) A_{\sigma(1)1} A_{\sigma(2)2} \cdots b_{\sigma(j)} \cdots A_{\sigma(n)n}$$
$$= \det(a_1, a_2, \ldots, a_j, \ldots, a_n) + \det(a_1, a_2, \ldots, b, \ldots, a_n).$$

(*b*)

$$\det(a_1, a_2, \ldots, ca_j, \ldots, a_n) = \sum_{\sigma \in S_n} \operatorname{sign}(\sigma) A_{\sigma(1)1} A_{\sigma(2)2} \cdots cA_{\sigma(j)j} \cdots A_{\sigma(n)n}$$
$$= c \sum_{\sigma \in S_n} \operatorname{sign}(\sigma) A_{\sigma(1)1} A_{\sigma(2)2} \cdots A_{\sigma(j)j} \cdots A_{\sigma(n)n} = c \det((a_1, a_2, \ldots, a_j, \ldots, a_n)$$

(*c*) Consideremos $\tau \in S_n$, definido por $\tau = (i\, j)$, de modo que

$$\delta := \det(a_1, a_2, \ldots, a_j, \ldots, a_i, \ldots, a_n) = \det(a_{\tau(1)}, a_{\tau(2)}, \ldots, a_{\tau(i)}, \ldots, a_{\tau(j)}, \ldots, a_{\tau(n)}).$$

Así,

$$\delta = \sum_{\sigma \in S_n} \operatorname{sign}(\sigma) A_{\sigma\tau(1)1} A_{\sigma\tau(2)2} \cdots A_{\sigma\tau(i)i} \cdots A_{\sigma\tau(j)j} \cdots A_{\sigma\tau(n)n}.$$

Pero como $\sigma\tau \in S_n$ y $\operatorname{sign}(\tau) = -1$, podemos definir una nuevo índice $\rho = \sigma\tau$ y escribir $\operatorname{sign}(\sigma) = -\operatorname{sign}(\rho)$, permitiéndonos expresar el determinante δ así:

$$\begin{aligned}\delta &= \sum_{\rho \in S_n} -\operatorname{sign}(\rho) A_{\rho(1)1} A_{\rho(2)2} \cdots A_{\rho(i)i} \cdots A_{\rho(j)j} \cdots A_{\rho(n)n} \\ &= -\sum_{\rho \in S_n} \operatorname{sign}(\rho) A_{\rho(1)1} A_{\rho(2)2} \cdots A_{\rho(i)i} \cdots A_{\rho(j)j} \cdots A_{\rho(n)n} \\ &= -\det(a_1, a_2, \ldots, a_i, \ldots, a_j, \ldots, a_n).\end{aligned}$$

□

Observación. Este teorema dice que la función determinante, vista como una función de las columnas de la matriz, es una función lineal cuando las demás componentes permanecen constantes. Más aún, la determinante conserva estas propiedades vista como una función de las filas de la matriz.

Corolario 4.1.1. *Sea A una matriz con dos (o más) columnas idénticas. Entonces,* $\det(A) = 0$.

Demostración. Supongamos que A es una matriz con dos columnas idénticas y B se obtiene de A intercambiando estas dos columnas. Por el teorema 4.1(c), $\det(A) = -\det(B)$, pero A y B son la misma matriz, de modo que $\det(A) = \det(B)$ y por lo tanto, $\det(A) = -\det(A)$. En un campo arbitrario esto solo es cierto si $\det(A) = 0$. □

Corolario 4.1.2. *Sea A una matriz con una columna compuesta por ceros únicamente. Entonces,* $\det(A) = 0$.

Demostración. Sea B una matriz arbitraria y supongamos que A se obtiene a partir de B al multiplicar una (o más) de sus columnas por 0. Entonces, por el teorema 4.1, $\det(A) = 0 \cdot \det(B) = 0$. $\square$

Corolario 4.1.3. *Sea $A \in M_{n\times n}(F)$ una matriz cualquiera y $\sigma \in S_n$. Supongamos que*

$$B = \begin{bmatrix} A_{;\sigma(1)} & A_{;\sigma(2)} & \dots & A_{;\sigma(n)} \end{bmatrix}.$$

Entonces, $\det(B) = \operatorname{sign}(\sigma) \det(A)$.

Demostración. Análoga a la prueba de la parte (c) del teorema 4.1, pero con $\tau \in S_n$ arbitrario. $\square$

Corolario 4.1.4. *Sea A una matriz cualquiera y supongamos que B se obtiene de A al sumar un múltiplo de una columna de A a otra columna de A. Entonces,* $\det(B) = \det(A)$.

Demostración. Supongamos que B se obtiene a partir de A al sumar c veces la j-ésima columna de A a la i-ésima columna de A. Usando el teorema 4.1,

$$\det(B) = \det(A) + c\det(C),$$

donde C se obtiene al reemplazar la i-ésima columna de A por la j-ésima. C tiene dos columna idénticas, de modo que $\det(C) = 0$ por el corolario 4.1.1. Esto establece el corolario. $\square$

Corolario 4.1.5. *Sea $A \in M_{n\times n}(F)$ tal que* $\operatorname{rank}(A) < n$. *Entonces,* $\det(A) = 0$.

Demostración. Si $\operatorname{rank}(A) < n$, el número de columnas linealmente independientes es menor que n. Por lo tanto, por lo menos una de las columnas de A es combinación lineal de las demás. Nos permitimos afirmar que la columna r es una combinación lineal de las demás, entonces existen $a_i \in F$ con $i \neq r$ no todos iguales a cero tales que

$$A_{;r} = \sum_{\substack{i=1 \\ i\neq r}}^{n} a_i A_{;i}.$$

Sea B la matriz que se obtiene a partir de A al sumar $-a_i$ veces la columna i a la columna r para cada $i \neq r$. Como no todos los a_i son iguales a cero, la r-ésima columna de B es nula y así $\det(B) = 0$ por el corolario 4.1.2. Además, por el corolario anterior, $\det(A) = \det(B) = 0$. $\square$

Observación. Lo que se ha dicho hasta sobre las columnas de una matriz también es válido para las filas. La demostración del teorema y sus corolarios para las filas de una matriz solo requieren cambiar el índice sobre el cual actúa la permutación.

Teorema 4.2. *Para dos matrices cualesquiera $A, B \in M_{n\times n}(F)$,* $\det(AB) = \det(A) \cdot \det(B)$.

Demostración. Directamente, notamos que las columnas del producto AB están dadas por

$$(AB)_{;j} = A(B_{;j})$$

de modo que el determinante será

$$\begin{aligned}
\det(AB) &= \det\begin{pmatrix} A(B_{;1}) & A(B_{;2}) & \dots & A(B_{;n}) \end{pmatrix} \\
&= \det\begin{pmatrix} A(\sum_{i=1}^n B_{i1}e_i) & A(\sum_{i=1}^n B_{i2}e_i) & \dots & A(\sum_{i=1}^n B_{in}e_i) \end{pmatrix} \\
&= \det\begin{pmatrix} \sum_{i=1}^n B_{i1}Ae_i & \sum_{i=1}^n B_{i2}Ae_i & \dots & \sum_{i=1}^n B_{in}Ae_i \end{pmatrix} \\
&= \sum_{i_1=1}^{n} \sum_{i_2=1}^{n} \cdots \sum_{i_n=1}^{n} B_{i_1 1}B_{i_2 2}\cdots B_{i_n n} \det\begin{pmatrix} Ae_{i_1} & Ae_{i_2} & \dots & Ae_{i_n} \end{pmatrix}
\end{aligned}$$

Este último resultado se obtiene al aplicar varias veces la multi-linealidad del determinante. Nótese que cuando dos o más de los i_j son iguales, el determinante se anula. Por lo tanto, las n sumatorias se pueden reemplazar sobre una única sumatoria sobre las permutaciones. Esto es,

$$\begin{aligned}\det(AB) &= \sum_{\sigma \in S_n} B_{\sigma(1)1} B_{\sigma(2)2} \cdots B_{\sigma(n)n} \det\begin{pmatrix} Ae_{\sigma(1)} & Ae_{\sigma(2)} & \ldots & Ae_{\sigma(n)} \end{pmatrix} \\ &= \sum_{\sigma \in S_n} B_{\sigma(1)1} B_{\sigma(2)2} \cdots B_{\sigma(n)n} \operatorname{sign}(\sigma) \det(A)\end{aligned}$$

Esta última igualdad se debe a que $Ae_{\sigma(i)}$ corresponde a la $\sigma(i)$-ésima columna de A, de modo que el determinante dentro de la sumatoria en la primera fila no es más que el determinante de una matriz obtenida al permutar las columnas de A bajo σ. Por lo tanto,

$$\begin{aligned}\det(AB) &= \det(A) \sum_{\sigma \in S_n} \operatorname{sign}(\sigma) B_{\sigma(1)1} B_{\sigma(2)2} \cdots B_{\sigma(n)n} \\ &= \det(A)\det(B).\end{aligned}$$

□

Observación. Nótese que como $\det(A), \det(B) \in F$, su producto conmuta y el determinante es independiente del orden del producto entre A y B; esto es, $\det(AB) = \det(BA)$.

Corolario 4.2.1. *Una matriz* $A \in M_{n\times n}(F)$ *es invertible si y solo si* $\det(A) \neq 0$. *Más aún, si* A *es invertible, entonces* $\det(A^{-1}) = \det(A)^{-1}$.

Demostración. Si $A \in M_{n\times n}$ no es invertible, $\operatorname{rank}(A) < n$ y por el corolario 4.1.5, esto implica que $\det(A) = 0$. Por otra parte, si A es invertible,

$$1 = \det(I_n) = \det(AA^{-1}) = \det(A)\det(A^{-1}).$$

La última igualdad es consecuencia del teorema 4.2. Por lo tanto, $\det(A) \neq 0$ y es evidente que

$$\det(A^{-1}) = \det(A)^{-1}.$$

□

Corolario 4.2.2. *Sean* $A, B \in M_{n\times n}(F)$ *matrices similares. Entonces,* $\det(A) = \det(B)$.

Demostración. Supongamos que B es similar a A. Entonces, existe una matriz invertible Q tal que $B = Q^{-1}AQ$, de modo que

$$\begin{aligned}\det(B) &= \det(Q^{-1}AQ) \\ &= \det(Q^{-1})\det(A)\det(Q) \\ &= \det(Q)^{-1}\det(Q)\det(A) \\ &= \det(A).\end{aligned}$$

□

Teorema 4.3. *Sea* $A \in M_{n\times n}(F)$ *una matriz triangular superior. Entonces, el determinante de* A *es el producto de los términos en la diagonal.*

Demostración. Primero, notemos que $A_{ij} = 0$ si $i > j$. Entonces, si $\sigma(i) > i$ para $i = 1, 2, \ldots, n$ el término de la sumatoria del determinante se hace 0. Entonces, las únicas permutaciones que afectan el determinante son aquellas para las que $\sigma(i) \geq i$ para $i = 1, 2, \ldots, n$. La única permutación que cumple esta restricción es la identidad, de modo que el determinante viene dado por

$$\det(A) = A_{11}A_{22}\cdots A_{nn}.$$

□

Lema 4.4. *Sea E una matriz elemental.* $\det(E) = \det(E^t)$.

Demostración. Supongamos que E_1 es una matriz elemental de tipo 1, su determinante es -1, por el teorema 4.1(c). Una matriz elemental de tipo 1 es simétrica, de modo que $E_1 = E_1^t$ y automáticamente, el determinante de la matriz y su transpuesta es igual.

Supongamos que E_2 es una matriz elemental de tipo 2. Entonces, E_2 es una matriz diagonal y por ende simétrica. Nuevamente, es automático que $\det(E_2) = \det(E_2^t)$.

Finalmente, supongamos que E_3 es una matriz elemental de tipo 3. Supongamos además que E_3 se obtiene de I_n al aplicar una operación elemental de tipo 3 sobre las columnas. Por el corolario 4.1.4, $\det(E_3) = \det(I_n)$. E_3^t se obtiene de I_n al aplicar la misma operación elemental, pero sobre las filas. Por lo tanto, $\det(E_3^t) = \det(I_n) = \det(E_3)$. Intercambiando los roles de E_3 y E_3^t, se establece la igualdad cuando la operación elemental es sobre las filas.

Habiendo probado la igualdad para matrices elementales de los tres tipos, la afirmación es cierta para toda matriz elemental. □

Teorema 4.5. *Sea $A \in M_{n\times n}(F)$.* $\det(A) = \det(A^t)$.

Demostración. Supongamos que A no es invertible, de modo que $\operatorname{rank}(A) < n$. Por el corolario 3.7.2, $\operatorname{rank}(A^t) < n$ y en consecuencia, $\operatorname{rank}(A^t) = 0 = \operatorname{rank}(A)$, por el corolario 4.1.5.

Por otra parte, supongamos que A es invertible. Entonces, A puede ser escrita como el producto de matrices elementales, digamos $A = E_m \cdots E_2E_1$, por el corolario 3.7.3. Así,

$$\begin{aligned}
\det(A^t) &= \det(E_1^tE_2^t\cdots E_m^t) \\
&= \det(E_1^t)\det(E_2^t)\cdots\det(E_m^t) \\
&= \det(E_1)\det(E_2)\cdots\det(E_m) \\
&= \det(E_m)\cdots\det(E_2)\det(E_1) \\
&= \det(E_m\cdots E_2E_1) \\
&= \det(A).
\end{aligned}$$

La tercera igualdad es consecuencia del lema anterior y la penúltima se debe al teorema 4.2. En cualquier caso, $\det(A) = \det(A^t)$. □

5. Diagonalización

Def. Sea T un operador lineal sobre un espacio vectorial de dimensión finita V. T se dice **diagonalizable** si existe una base ordenada β de V tal que $[T]_\beta$ es una matriz diagonal. Una matriz cuadrada A es diagonalizable si L_A es diagonalizable.

5.1. Valores y vectores propios

Def. Sea T un operador lineal sobre un espacio vectorial V. Un vector no nulo $v \in V$ se dice **vector propio** de T si existe un escalar λ tal que $T(v) = \lambda v$. El escalar λ se denomina **valor propio** correspondiente al vector propio v.

Teorema 5.1. *Un operador lineal T sobre un espacio vectorial de dimensión finita V es diagonalizable si y solo si existe una base ordenada β de V que consiste únicamente de vectores propios de T. Más aún, si T es diagonalizable, $\beta = \{v_1, v_2, \ldots, v_n\}$ es una base ordenada de vectores propios de T y $D = [T]_\beta$, entonces D es una matriz diagonal y D_{jj} es el valor propio correspondiente a v_j para $j = 1, 2, \ldots, n$.*

Demostración. Supongamos que T es diagonalizable, de modo que existe una base $\beta = \{v_1, v_2, \ldots, v_n\}$ tal que $D = [T]_\beta$ es una matriz diagonal. Para cada vector $v_j \in \beta$,

$$T(v_j) = \sum_{i=1}^{n} D_{ij} v_i = D_{jj} v_j = \lambda_j v_j,$$

donde $D_{jj} = \lambda_j$. Como los v_j forman una base, ninguno de ellos es nulo y por lo tanto son vectores propios.

Recíprocamente, si $\beta = \{v_1, v_2, \ldots, v_n\}$ es una base ordenada de V de vectores propios, existen $\lambda_1, \lambda_2, \ldots, \lambda_n \in F$ tales que $T(v_j) = \lambda_j v_j$. En consecuencia,

$$D = [T]_\beta = \begin{pmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & \ldots & 0 \\ \vdots & \vdots & \ddots & \ldots \\ 0 & 0 & \ldots & \lambda_n \end{pmatrix}$$

es una matriz diagonal. $\square$

Teorema 5.2. *Sea $A \in M_{n \times n}(F)$. Entonces, λ es un valor propio de A si y solo si* $\det(A - \lambda I_n) = 0$.

Demostración. Un escalar λ es valor propio de A si y solo si existe un vector no nulo $v \in F^n$ tal que $Av = \lambda v$; esto es, $(A - \lambda I_n)(v) = \mathit{0}$. Por el teorema 2.6, es se cumple si y solo si $A - \lambda I_n$ no es invertible. Esto a su vez es cierto si y solo si $\det(A - \lambda I_n) = 0$. $\square$

Def. Sea $A \in M_{n \times n}(F)$. El polinomio $p_A(t) = \det(A - tI_n)$ se denomina **polinomio característico** de A.

Observación. La matriz $A - tI_n$ no es una matriz sobre el campo F sino sobre el campo de polinomios en la variable t con coeficientes en F, denotado $F[t]$.

Teorema 5.3. *Sean $A, B \in M_{n \times n}(F)$ matrices similares. Entonces, sus polinomios característicos son iguales.*

Demostración. Sea Q una matriz tal que $B = Q^{-1}AQ$ y sean $p(t)$ y $q(t)$ los polinomios característicos de A y B respectivamente. Tenemos que

$$\begin{aligned} q(t) &= \det(B - tI_n) = \det(Q^{-1}AQ - tI_n) \\ &= \det(Q^{-1}AQ - tQ^{-1}I_nQ) \\ &= \det(Q^{-1}(A - tI_n)Q) \\ &= \det(Q^{-1})\det(A - tI_n)\det(Q) \\ &= \det(A - tI_n) = p(t). \end{aligned}$$

$\square$

Def. Sea T un operador lineal sobre espacio vectorial V de dimensión n con base ordenada β. Definimos el **polinomio característico** $p_T(t)$ de T como el polinomio característico de $A = [T]_\beta$.

Observación. El teorema 5.3 me permite definir independiente de la base escogida el polinomio característico de T.

Teorema 5.4. *Sea T un operador lineal sobre un espacio vectorial V y sea λ un valor propio de T. Un vector $v \in V$ es un vector propio de T correspondiente a λ si y solo si $v \neq 0$ y $v \in \ker(T - \lambda \operatorname{id}_V)$.*

Demostración. Sea v un vector propio de T. Entonces, $T(v) = \lambda v$, esto es,

$$\begin{aligned} T(v) - \lambda \operatorname{id}_V(v) &= 0 \\ (T - \lambda \operatorname{id}_V)(v) &= 0. \end{aligned}$$

Esto prueba que $v \in \ker(T - \lambda \operatorname{id}_V)$.

Recíprocamente, si $v \in \ker(T - \lambda \operatorname{id}_V)$, $(T - \lambda \operatorname{id}_V)(v) = 0$. Esto implica que

$$\begin{aligned} T(v) - \lambda \operatorname{id}_V(v) &= 0 \\ T(v) - \lambda v &= 0 \\ T(v) &= \lambda v. \end{aligned}$$

Esto establece que v es un vector propio de T con valor propio λ. □

5.2. Diagonalizabilidad

Teorema 5.5. *Sea T un operador lineal sobre un espacio vectorial V y sean λ_1, λ_2, ..., λ_k valores propios distintos de T. Si v_1, v_2, ..., v_k son vectores propios tales que λ_i corresponde a v_i (para $1 \leq i \leq k$), entonces $\{v_1, v_2, \ldots, v_k\}$ es linealmente independiente.*

Demostración. La prueba es por inducción sobre k. Supongamos que $k = 1$. Entonces $v_1 \neq 0$ pues v_1 es un vector propio y por ende $\{v_1\}$ es linealmente independiente.

Ahora supongamos que el teorema vale para $k-1$ valores propios distintos, donde $k - 1 \geq 1$, y tenemos k vectores propios v_1, v_2, ..., v_k correspondientes a los distintos valores propios λ_1, λ_2, ..., λ_k. Queremos mostrar que $\{v_1, v_2, \ldots, v_k\}$ es linealmente independiente. Supongamos que a_1, a_2, ..., a_k son escalares tales que

$$a_1 v_1 + a_2 v_2 + \cdots + a_k v_k = 0.$$

Aplicando $T - \lambda_k \operatorname{id}_V$ a ambos lados de la ecuación obtenemos

$$a_1(\lambda_1 - \lambda_k)v_1 + a_2(\lambda_2 - \lambda_k)v_2 + \cdots + a_{k-1}(\lambda_{k-1} - \lambda_k)v_{k-1} = 0.$$

Por la hipótesis de inducción $\{v_1, v_2, \ldots, v_{k-1}\}$ es linealmente independiente y por lo tanto,

$$a_1(\lambda_1 - \lambda_k) = a_2(\lambda_2 - \lambda_k) = \cdots = a_{k-1}(\lambda_{k-1} - \lambda_k) = 0.$$

Como λ_1, λ_2,..., λ_k son distintos, $\lambda_i - \lambda_k \neq 0$ para $1 \leq i \leq k-1$. Entonces $a_1 = a_2 = \cdots = a_{k-1}$ y así la primera ecuación de la demostración se reduce a $a_k v_k = 0$, pero $v_k \neq 0$ y en consecuencia $a_k = 0$. Así, $a_1 = a_2 = \cdots = a_k = 0$ y concluimos que $\{v_1, v_2, \ldots, v_k\}$ es linealmente independiente. □

Corolario 5.5.1. *Sea T un operador lineal sobre un espacio vectorial n-dimensional V. Si T tiene n valores propios distintos, entonces T es diagonalizable.*

Demostración. Supongamos que T tiene n valores propios distintos $\lambda_1, \ldots, \lambda_n$. Para cada i escogemos un vector propio v_i correspondiente a λ_i. Por el teorema 5.5, $\{v_1, \ldots, v_n\}$ es linealmente independiente y como $\dim(V) = n$, este conjunto es una base de V. Entonces, por el teorema 5.1, T es diagonalizable. □

Teorema 5.6. *Sea T un operador lineal diagonalizable sobre un espacio vectorial V. Entonces, su polinomio característico se parte.*

Demostración. Sea T un operador lineal diagonalizable sobre un espacio vectorial n-dimensional V y sea β una base ordenada de V tal que $[T]_\beta = D$ es una matriz diagonal. Supongamos que

$$D = \begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n \end{bmatrix},$$

y sea $f(t)$ el polinomio característico de T. Entonces,

$$\begin{aligned} f(t) = \det(D - tI_n) = \det \begin{pmatrix} \lambda_1 - t & 0 & \dots & 0 \\ 0 & \lambda_2 - t & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n - t \end{pmatrix} \\ = (\lambda_1 - t)(\lambda_2 - t) \cdots (\lambda_n - t) = (-1)^n (t - \lambda_1)(t - \lambda_2) \cdots (t - \lambda_n). \end{aligned}$$

□

Def. Sea T un operador lineal sobre un espacio vectorial V y sea λ un valor propio de T. Definimos $E_\lambda = \{v \in V \,|\, T(v) = \lambda v\} = \ker(T - \lambda \mathrm{id}_V)$. El conjunto E_λ se llama el **espacio propio** de T correspondiente al valor propio λ.

Observación. E_λ es un subespacio de V, pues es el núcleo de un operador sobre V. E_λ consiste en vector cero y los vectores propios asociados al valor propio λ.

Def. Sea λ un valor propio de un operador con polinomio característico $f(t)$. La **multiplicidad algebraica** de λ es el máximo entero positivo k para el cual $(t - \lambda)^k$ es factor de $f(t)$. Al entero $\dim(E_\lambda)$ lo llamamos **multiplicidad geométrica** de λ.

Teorema 5.7. *Sea T un operador lineal sobre un espacio vectorial de dimensión finita V y sea λ un valor propio con multiplicidad algebraica m. Entonces, $1 \leq \dim(E_\lambda) \leq m$.*

Demostración. Sea $d = \dim(E_\lambda)$. La desigualdad $1 \leq d$ es inmediata puesto que los espacios propios de un operador no son triviales.

Para la segunda desigualdad, sea $\{v_1, v_2, \dots, v_d\}$ una base ordenada de E_λ. Extendemos esta base a una base ordenada de V, $\beta = \{v_1, v_2, \dots, v_d, v_{d+1}, \dots, v_n\}$. Como v_i (para $1 \leq i \leq d$) es un vector propio de T con valor propio λ,

$$[T]_\beta = \begin{bmatrix} \lambda I_d & B \\ O & C \end{bmatrix}.$$

El polinomio característico de T es

$$\begin{aligned} f(t) &= \det \begin{pmatrix} (\lambda - t) I_d & B \\ O & C - tI_{n-d} \end{pmatrix} = \det((\lambda - t) I_d) \det(C - tI_{n-d}) \\ &= (\lambda - t)^d g(t), \end{aligned}$$

donde $g(t)$ es un polinomio. Así, $(\lambda - t)^d$ es un factor de $f(t)$ y la multiplicidad de λ es al menos d y por lo tanto $d \leq m$. □

Sumas directas

El concepto de suma directa permite descomponer un espacio vectorial en subespacios más sencillos que ofrecen una perspectiva del comportamiento de T, incluso si este no es diagonalizable.

A continuación ofrecemos una generalización de la suma directa a más de dos subespacios.

Def. Sean $W_1, W_2, \ldots, W_k$ subespacios de un espacio vectorial V. Definimos la **suma** de estos subespacios como el conjunto

$$\{v_1 + v_2 + \cdots + v_k \,|\, v_i \in W_i \text{ para } 1 \leq i \leq k\},$$

lo cual denotamos por $W_1 + W_2 + \cdots + W_k$ ó $\sum\limits_{i=1}^{k} W_i$.

Def. Sean $W_1, W_2, \ldots, W_k$ subespacios de un espacio vectorial V. Decimos que la suma de $W_1, W_2, \ldots, W_k$ es directa si para todo j $(1 \leq j \leq k)$

$$W_j \cap \sum_{i \neq j} W_i = \{0\}.$$

Esto lo denotamos $W_1 \oplus W_2 \oplus \cdots \oplus W_k$ ó $\bigoplus\limits_{i=1}^{k} W_i$.

Teorema 5.8. *Sean $W_1, W_2, \ldots, W_k$ subespacios de un espacio vectorial de dimensión finita V. Las siguientes condiciones son equivalentes.*

(*a*) *$V = W_1 \oplus W_2 \oplus \cdots \oplus W_k$.*

(*b*) *$V = W_1 + W_2 + \cdots + W_k$ y para cualesquiera vectores $v_1, v_2, \ldots, v_k$ tales que $v_i \in W_i$ $(1 \leq i \leq k)$, si $v_1 + v_2 + \cdots + v_k = 0$, entonces $v_i = 0$ para todo i.*

(*c*) *Cada vector $v \in V$ puede escribirse de forma unívoca como $v = v_1 + v_2 + \cdots + v_k$ donde $v_i \in W$.*

(*d*) *Si γ_i es una base ordenada de W_i $(1 \leq i \leq k)$, entonces $\gamma_1 \cup \gamma_2 \cup \cdots \cup \gamma_k$ es una base ordenada para V.*

(*e*) *Para cada $i = 1, 2, \ldots, k$, existe una base ordenada γ_i para W_i tal que $\gamma_1 \cup \gamma_2 \cup \cdots \cup \gamma_k$ es una base ordenada de V.*

Demostración. Primero asumimos (a) y probamos (b). Claramente,

$$V = \sum_{i=1}^{k} W_i.$$

Ahora supongamos que $v_1, v_2, \ldots, v_k$ son vectores tales que $v_i \in W_i$ $(1 \leq i \leq k)$ y $v_1 + v_2 + \cdots + v_k = 0$. Para cualquier j,

$$-v_j = \sum_{i \neq j} v_i \in \sum_{i \neq j} W_i.$$

Pero $-v_j \in W_j$ y

$$-v_j \in W_j \cap \sum_{i \neq j} W_i = \{0\}.$$

Esto prueba que $v_j = 0$, lo cual prueba (b).

Ahora supongamos (b) y probamos (c). Por (b), existen vectores $v_1, v_2, \ldots, v_k$ son vectores tales que $v_i \in W_i$ y $v = v_1 + v_2 + \cdots + v_k$. Queremos mostrar que esta representación es única. Supongamos que $v = w_1 + w_2 + \cdots + w_k$, donde $w_i \in W_i$ para todo i. Entonces

$$(v_1 - w_1) + (v_2 - w_2) + \cdots + (v_k - w_k) = 0.$$

Pero $v_i - w_i \in W_i$ para todo i y por (b) $v_1 - w_i = 0$. Así $v_i = w_i$ para todo i, lo cual prueba la unicidad de la representación.

Ahora asumimos (c) y probamos (d). Para cada i, sea γ_i una base ordenada de W_i. Como

$$V = \sum_{i=1}^{k} W_i$$

por (c), tenemos que $\gamma_1 \cup \gamma_2 \cup \cdots \cup \gamma_k$ genera a V. Para probar que este conjunto es linealmente independiente consideremos vectores $v_{ij} \in \gamma_i$ $(j = 1, 2, \ldots, m_1$ e $i = 1, 2, \ldots, k)$ y escalares a_{ij} tales que

$$\sum_{i,j} = a_{ij} v_{ij} = 0.$$

Para cada i, fijemos

$$w_i = \sum_{j=1}^{m_i} a_{ij} v_{ij}.$$

Entonces para cada i, $w_i \in \operatorname{span}(\gamma_i) = W_i$ y

$$w_1 + w_2 + \cdots + w_k = \sum_{i,j} a_{ij} v_{ij} = 0.$$

Como $0 \in W_i$ para cada i y $0 + 0 + \cdots + 0 = w_1 + w_2 + \cdots + w_k$, (c) implica que $w_i = 0$ para todo i. Así,

$$0 = w_i = \sum_{j=1}^{m_i} a_{ij} v_{ij}$$

para cada i. Pero cada γ_i es linealmente independiente y por ende $a_{ij} = 0$ para todo i y todo j. En consecuencia, $\gamma_1 \cup \gamma_2 \cup \cdots \cup \gamma_k$ es linealmente independiente y por lo tanto es una base de V.

Claramente (e) se sigue de (d).

Resta verificar que (e) implica (a). Para cada i, sea γ_i una base ordenada de W_i tal que $\gamma_1 \cup \gamma_2 \cup \cdots \cup \gamma_k$ es una base ordenada de V. Entonces

$$V = \operatorname{span}(\gamma_1 \cup \gamma_2 \cup \cdots \cup \gamma_k).$$

Es fácil probar que el espacio generado por una unión de conjuntos es la suma de los subespacios generados por cada conjunto, de modo que

$$V = \operatorname{span}(\gamma_1) + \operatorname{span}(\gamma_2) + \cdots + \operatorname{span}(\gamma_k) = \sum_{i=1}^{k} W_i.$$

Tomemos un vector no nulo $v \in \bigcap_{i=1}^{k} W_i$. Entonces v puede escribirse de forma única como combinación lineal de los vectores en γ_i para todo i. Esto quiere decir que v se puede escribir de varias maneras como combinación lineal de $\gamma_1 \cup \gamma_2 \cup \cdots \cup \gamma_k$, lo cual contradice la unicidad de la representación de un vector en una base. Por lo tanto, concluimos que $v = 0$; esto es, $\bigcap_{i=1}^{k} W_i = \{0\}$, lo cual prueba (a). $\square$

Teorema 5.9. *Sea T un operador lineal sobre un espacio vectorial V. Supongamos que $\lambda_1, \lambda_2, \ldots, \lambda_k$ son valores propios distintos de T. Entonces, la suma de los espacios propios $E_{\lambda_1}, E_{\lambda_2}, \ldots, E_{\lambda_k}$ es directa.*

Demostración. Tomemos $v_1, v_2, \ldots, v_k$ tales que $v_i \in E_{\lambda_i}$ para cada i y

$$v_1 + v_2 + \cdots + v_k = \mathit{0}.$$

Como los vectores propios correspondientes a distintos valores propios son linealmente independientes (teorema 5.5), esto implica que cada $v_i = \mathit{0}$. Usando el teorema anterior (5.8), esto implica que la suma es directa. □

Teorema 5.10. *Sea T un operador lineal sobre un espacio vectorial de dimensión finita V. Sean $\lambda_1, \lambda_2, \ldots, \lambda_k$ valores propios distintos de T. Las siguientes afirmaciones son equivalentes.*

(a) *T es diagonalizable.*

(b) *V tiene una base que consiste en vectores propios de T.*

(c) $V = \bigoplus_{i=1}^{k} E_{\lambda_i}$.

(d) *Para cada λ_i ($1 \leq i \leq k$), su multiplicidad algebraica es igual a su multiplicidad geométrica.*

Demostración. La equivalencia de (a) y (b) es el teorema 5.1.

Ahora procedemos a probar la equivalencia entre (b) y (c). Si V tiene una base que consiste en vectores propios de T, todo vector en v se puede escribir como una combinación lineal de vectores propios de T, de modo que

$$V = \sum_{i=1}^{k} E_{\lambda_k}.$$

Por el teorema 5.9, (c) se cumple. Recíprocamente, si (c) es verdadero, por el teorema 5.8 para cada i existe una base ordenada γ_i de E_{λ_i} tal que $\gamma_1 \cup \gamma_2 \cup \cdots \cup \gamma_k$ es una base ordenada de V. Como cada γ_i contiene vectores propios, $\gamma_1 \cup \gamma_2 \cup \cdots \cup \gamma_k$ es una base de V con vectores propios, lo cual prueba (b).

Para culminar la prueba, mostraremos que (b) y (d) son equivalentes. Para cada i, sea m_i la multiplicidad algebraica de λ_i y d_i la multiplicidad geométrica.

Supongamos que β es una base ordenada de V conformada por vectores propios de T. Para cada i, sea $\beta_i = \beta \cap E_{\lambda_i}$, el conjunto de vectores en β que son vectores propios correspondientes a λ_i y sea n_i el número de vectores en β_i. Entonces $n_i \leq d_i$ para cada i pues β_i es un subconjunto linealmente independiente de un subespacio de dimensión d_i y $d_i \leq m_i$ por el teorema 5.7. Los n_i suman n puesto que β contiene n vectores. Los m_i también suman n porque la suma de las multiplicidades algebraicas es igual al grado del polinomio característico de T. Así,

$$n = \sum_{i=1}^{k} n_i \leq \sum_{i=1}^{k} d_i \leq \sum_{i=1}^{k} m_i = n.$$

De esto sigue que

$$\sum_{i=1}^{k} (m_i - d_i) = 0,$$

pero como $(m_i - d_i) \geq 0$ para todo i, concluimos que $m_i = d_i$ para todo i.

Recíprocamente, supongamos que $m_i = d_i$ para todo i. Para cada i, sea β_i una base ordenada para E_{λ_i} y sea $\beta = \beta_1 \cup \beta_2 \cup \cdots \cup \beta_k$. Como dos vectores cualesquiera en un mismo β_i son linealmente independientes y dos vectores cualesquiera, uno en β_i y otro en β_j $(i \neq j)$, corresponden a valores propios distintos, β es linealmente independiente. Más aún, β contiene

$$\sum_{i=1}^{k} d_i = \sum_{i=1}^{k} m_i = n$$

vectores. En consecuencia, β es una ordenada de V que consiste en vectores propios de T. □

5.3. Espacios invariantes y el teorema de Cayley-Hamilton

Def. Sea T un operador lineal sobre un espacio vectorial V. Un subespacio W de V se dice un subespacio T-invariante de V si $T(W) \subseteq W$, esto es $T(v) \in W$ para todo $v \in W$.

Observación. Un espacio propio de T es un subespacio T-invariante.

Def. Sean $f(t) \in F[t]$ y T un operador lineal sobre un espacio vectorial V sobre el campo F. Supongamos que $f(t) = a_n t^n + a_{n-1} t^{n-1} + \cdots + a_1 t + a_0$. Definimos el operador $f(T)$ por

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

donde para todo entero positivo k,

$$T^k = \underbrace{T \circ T \circ \cdots \circ T}_{k \text{ veces}}.$$

Def. Sea T un operador lineal sobre un espacio vectorial V y sea x un vector no nulo en V. El subespacio

$$W = \mathrm{span}(\{x, T(x), T^2(x), \dots\})$$

se denomina el **subespacio T-cíclico generado por** x.

Teorema 5.11. *Sea T un operador lineal sobre un espacio vectorial V, sea $v \in V$ un vector no nulo y sea W el subespacio T-cíclico generado por v. W es T-invariante y cualquier subespacio T-invariante de V que contiene a v también contiene a W.*

Demostración. Sea $w \in W$. Como $\{v, T(v), T^2(v), \dots\}$ genera a W, w se puede escribir como combinación lineal de un subconjunto finito de vectores en $\{v, T(v), T^2(v), \dots\}$, esto es existe una colección finita de índices $I \subseteq \mathbb{N}$ y un escalar para cada índice tal que

$$w = \sum_{i \in I} a_i T^i(v),$$

para alguna colección de escalares $\{a_i\}_{i \in I}$ $(T^0 = \mathrm{id}_V)$. En consecuencia,

$$T(w) = T\left(\sum_{i \in I} a_i T^i(v)\right) = \sum_{i \in I} a_i T^{i+1}(v)$$

Como $T^k \in \{v, T(v), T^2(v), \dots\}$ para todo entero no negativo k, $T(w) \in \mathrm{span}(\{v, T(v), T^2(v), \dots\})$ y W es T-invariante.

Para probar la segunda afirmación supongamos que Z es un subespacio T-invariante de V que contiene a v. Como Z es T-invariante, $T(v) \in Z$. Similarmente, al ser Z T-invariante si $T^{k-1}(v) \in Z$, entonces $T^k(v) \in Z$. Vemos entonces que $\{v, T(v), T^2(v), \dots\} \subseteq Z$ y por el teorema 1.7 esto implica que $W \subseteq Z$. □

Teorema 5.12. *Sea T un operador lineal sobre un espacio vectorial de dimensión finita V y sea W un subespacio T-invariante de V. El polinomio característico de la restricción de T a W, T_W, divide el polinomio característico de T.*

Demostración. Escojamos una base ordenada de W $\gamma = \{v_1, v_2, \ldots, v_k\}$ y extendámosla a un base ordenada de V $\beta = \{v_1, v_2, \ldots, v_k, v_{k+1}, \ldots, v_n\}$. Sea $A = [T]_\beta$ y $B_1 = [T_W]_\gamma$. Como W es T-invariante,

$$T(v_i) = \sum_{j=1}^{n} a_{ij} v_j = \sum_{j=1}^{k} b_{ij} v_j,$$

para todo $i = 1, 2, \ldots, k$, de modo que

$$A = \begin{bmatrix} B_1 & B_2 \\ O & B_3 \end{bmatrix}.$$

Sea $f(t)$ el polinomio característico de T y $g(t)$ el polinomio característico de T_W. Entonces,

$$f(t) = \det(A - tI_n) = \det \begin{bmatrix} B_1 - tI_k & B_2 \\ O & B_3 - tI_{n-k} \end{bmatrix} = g(t) \cdot \det(B_3 - tI_{n-k}).$$

Así, $g(t)$ divide a $f(t)$. $\square$

Lema 5.13. *Sea A la matriz de tamaño $k \times k$*

$$\begin{bmatrix} 0 & 0 & \ldots & 0 & -a_0 \\ 1 & 0 & \ldots & 0 & -a_1 \\ 0 & 1 & \ldots & 0 & -a_2 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \ldots & 0 & -a_{k-2} \\ 0 & 0 & \ldots & 1 & -a_{k-1} \end{bmatrix},$$

donde $a_0, a_1, \ldots, a_{k-1}$ son escalares arbitrarios. Entonces, el polinomio característico de A es

$$(-1)^k (a_0 + a_1 t + \cdots + a_{k-1} t^{k-1} + t^k).$$

Demostración. Por inducción sobre k. Cuando $k = 1$, $A = [0]$ y su polinomio característico es $-t = (-1)^1 t^1$.

Supongamos ahora que la afirmación se cumple para $k - 1$. Sea $f(t)$ el polinomio característico de A. Entonces,

$$f(t) = \det(A - tI_k) = \det \begin{pmatrix} -t & 0 & \ldots & 0 & -a_0 \\ 1 & -t & \ldots & 0 & -a_1 \\ 0 & 1 & \ldots & 0 & -a_2 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \ldots & -t & -a_{k-2} \\ 0 & 0 & \ldots & 1 & -a_{k-1} - t \end{pmatrix}.$$

Calculando el determinante por cofactores a lo largo de la primera fila,

$$f(t) = -t \det \begin{pmatrix} -t & 0 & \ldots & 0 & -a_1 \\ 1 & -t & \ldots & 0 & -a_2 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \ldots & -t & -a_{k-2} \\ 0 & 0 & \ldots & 1 & -a_{k-1} - t \end{pmatrix} - a_0 (-1)^{k+1} \det \begin{pmatrix} 1 & -t & \ldots & 0 \\ 0 & 1 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & -t \\ 0 & 0 & \ldots & 1 \end{pmatrix}.$$

El primero de estos determinantes corresponde al polinomio característico de una matriz de tamaño $k-1$ con las condiciones del teorema. Por hipótesis de inducción, ese determinante es $(-1)^{k-1}(a_1+a_2t+\cdots+a_{k-1}t^{k-2}+t^{k-1})$. Por el teorema 4.3, el segundo determinante es 1. Así,

$$\begin{aligned} f(t) &= -t(-1)^{k-1}(a_1+a_2t+\cdots+a_{k-1}t^{k-2}+t^{k-1})-a_0(-1)^{k+1} \\ &= (-1)^k(a_1t+a_2t^2+\cdots+a_{k-1}t^{k-1}+t^k)+(-1)^{k+2}a_0 \end{aligned}$$

Como $(-1)^k=(-1)^{k+2}$, obtenemos finalmente que

$$f(t)=(-1)^k(a_0+a_1t+a_2t^2+\cdots+a_{k-1}t^{k-1}+t^k).$$

□

Teorema 5.14. *Sea T un operador lineal sobre un espacio vectorial de dimensión finita y sea W el subespacio T-cíclico generado por un vector $v\in V$. Sea $k=\dim(W)$. Entonces,*

(*a*) *$\{v,T(v),T^2(v),\ldots,T^{k-1}(v)\}$ es una base de W,*

(*b*) *si $a_0v+a_1T(v)+\cdots+a_{k-1}T^{k-1}(v)+T^k(v)=0$, entonces el polinomio característico de T_W es $f(t)=(-1)^k(a_0+a_1t+\cdots+a_{k-1}t^{k-1}+t^k)$.*

Demostración. (*a*) Como $v\neq 0$, el conjunto $\{v\}$ es linealmente independiente. Sea j el mayor entero positivo tal que

$$\beta=\{v,T(v),\ldots,T^{j-1}(v)\}$$

es linealmente independiente. Tal j debe existir puesto que V es un espacio vectorial de dimensión finita. Sea $Z=\mathrm{span}(\beta)$. Entonces β es una bae de Z. Más aún, $T^j(v)\in Z$ por el teorema 1.8. Usamos esta información para mostrar que Z es un subespacio T-invariante de V. Tomemos $w\in Z$. Como w es una combinación lineal de los vectores en β, existen escalares $b_0,b_1,\ldots,b_{j-1}$ tales que

$$w=b_0v+b_1T(v)+\cdots+b_{j-1}T^{j-1}(v),$$

y por ende

$$T(w)=b_0T(v)+b_1T^2(v)+\cdots+b_{j-1}T^j(v).$$

Entonces $T(w)$ es una combinación lineal de vectores en Z y por lo tanto pertenece a Z. Esto verifica que Z es T-invariante. Más aún, $v\in Z$ y por el teorema 5.11 $W\subseteq Z$. Claramente, $Z\subseteq W$, lo que nos permite concluir que $Z=W$. Se sigue que β es una base de W y por ende $\dim(W)=j$. Así, $j=k$. Esto prueba (*a*).

(*b*) Ahora tomemos β de la parte anterior como base de W. Sean $a_0,a_1,\ldots,a_{k-1}$ escalares tales que

$$a_0v+a_1T(v)+\cdots+a_{k-1}T^{k-1}(v)+T^k(v)=0.$$

Tenemos que

$$[T_W]_\beta=\begin{bmatrix} 0 & 0 & \ldots & 0 & -a_0 \\ 1 & 0 & \ldots & 0 & -a_1 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \ldots & 1 & -a_{k-1} \end{bmatrix},$$

que por el lema anterior tiene polinomio característico $f(t)=(-1)^k(a_0+a_1t+\cdots+a_{k-1}t^{k-1}+t^k)$, lo cual prueba (*b*).

□

Teorema 5.15 (Teorema de Cayley-Hamilton). *Sea T un operador lineal sobre un espacio de dimensión finita y sea $f(t)$ el polinomio característico de T. Entonces, $f(T) = T_0$.*

Demostración. Queremos mostrar que $f(T)(v) = 0$ para todo $v \in V$. Esto es evidente si $v = 0$ porque $f(T)$ es lineal. Supongamos entonces que $v \neq 0$. Sea W el subespacio T-cíclico generado por v y supongamos que $\dim(W) = k$. Por el teorema 5.14(a), existen escalares tales que

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

Entonces, el teorema 5.14 implica que

$$g(t) = (-1)^k (a_0 + a_1 t + \cdots + a_{k-1} t^{k-1} + t^k)$$

es el polinomio característico de T_W. Combinando estas dos ecuaciones tenemos que

$$g(T)(v) = (-1)^k (a_0 \mathrm{id}_V + a_1 T + \cdots + a_{k-1} T^{k-1} + T^k)(v) = 0.$$

Por el teorema 5.12, $g(t)$ divide a $f(t)$, de modo que existe un polinomio $q(t)$ tal que $f(t) = q(t)g(t)$ y

$$f(T)(v) = q(T)g(T)(v) = q(T)(g(T)(v)) = q(T)(0) = 0.$$

□

Teorema 5.16. *Sea T un operador lineal sobre un espacio vectorial de dimensión finita V y supongamos que $V = W_1 \oplus W_2 \oplus \cdots \oplus W_k$, donde W_i es un subespacio T-invariante de V para cada i ($1 \leq i \leq k$). Supongamos que $f_i(t)$ es el polinomio característico de T_{W_i} ($1 \leq i \leq k$). Entonces $f_1(t) \cdot f_2(t) \cdots f_k(t)$ es el polinomio característico de T.*

Demostración. Por inducción sobre k. Sea $f(t)$ el polinomio característico de T. Supongamos primero que $k = 2$. Sean β_1 y β_2 bases ordenadas de W_1 y W_2 respectivamente y sea $\beta = \beta_1 \cup \beta_2$. Entonces, β es una base ordenada de V por el teorema 5.8(d). Sean $A = [T]_\beta$, $B_1 = [T_{W_1}]_{\beta_1}$ y $B_1 = [T_{W_2}]_{\beta_2}$. Entonces,

$$A = \begin{bmatrix} B_1 & O \\ O' & B_2 \end{bmatrix},$$

donde O y O' son matrices nulas de los tamaños apropiados. Así,

$$f(t) = \det(A - tI) = \det(B_1 - tI) \cdot \det(B_2 - tI) = f_1(t) \cdot f_2(t).$$

Esto prueba el resultado para $k = 2$.

Ahora supongamos que el teorema vale para $k - 1$ y supongamos que V es la suma directa de k subespacios, digamos,

$$V = W_1 \oplus W_2 \oplus \cdots \oplus W_k.$$

Sea $W = W_1 + W_2 + \cdots + W_{k-1}$. Es fácil verificar que W es T-invariante y por lo tanto $V = W \oplus W_k$. Por el caso con $k = 2$, $f(t) = g(t) \cdot f(t)$ donde $g(t)$ es el polinomio característico de T_W. Claramente, $W = W_1 \oplus W_2 \oplus \cdots \oplus W_{k-1}$ y por ende $g(t) = f_1(t) \cdot f_2(t) \cdots f_{k-1}(t)$ por la hipótesis de inducción. Concluimos que

$$f(t) = g(t) \cdot f_k(t) = f_1(t) \cdot f_2(t) \cdots f_k(t).$$

□

Def. Sean $B_1 \in M_{n\times n}(F)$ y $B_2 \in M_{m\times m}$. Definimos la **suma directa** de B_1 y B_2, denotada $B_1 \oplus B_2$, como la matriz de tamaño $(m + n) \times (m + n)$ A tal que

$$A_{ij} = \begin{cases} (B_1)_{ij} & \text{para } 1 \leq i, j \leq n \\ (B_2)_{(i-n),(j-m)} & \text{para } n + 1 \leq i, j \leq n + m \\ 0 & \text{de lo contrario.} \end{cases}$$

Si $B_1, B_2, \ldots, B_k$ son matrices cuadradas con entradas en F definimos recursivamente su suma directa por

$$B_1 \oplus B_2 \oplus \cdots \oplus B_k = (B_1 \oplus B_2 \oplus \cdots \oplus) B_k.$$

Observación. La suma directa de una colección de matrices es una matriz diagonal por bloques.

Teorema 5.17. *Sea T un operador lineal sobre un espacio vectorial de dimensión finita V y sean $W_1, W_2, \ldots, W_k$ subespacios T-invariantes de V tales que $V = W_1 \oplus W_2 \oplus \cdots \oplus W_k$. Para cada i, sea β_i una base ordenada de W_i y definamos $\beta = \beta_1 \cup \beta_2 \cup \cdots \cup \beta_k$. Sea $A = [T]_\beta$ y $B_i = [T_{W_i}]_{\beta_i}$ para $i = 1, 2, \ldots, k$. Entonces,*

$$A = B_1 \oplus B_2 \oplus \cdots \oplus B_k.$$

Demostración. Por inducción sobre k, el número de subespacios. Veamos que el teorema vale para $k = 2$. Sea n la dimensión de W_1 y m la dimensión de W_2. Al ser W_1 T-invariante, la imagen mediante T de los vectores en β_1 será una combinación lineal de los vectores en β_1 donde los coeficientes se encuentran consignados en la matriz B_1. Por lo tanto, para cada $v_j \in \beta_1$,

$$T(v_j) = \sum_{i=1}^{n} (B_1)_{ij} v_i,$$

donde cada $v_i \in \beta_1$. En consecuencia,

$$A_{ij} = (B_1)_{ij} \qquad \text{para } 1 \leq i, j \leq n.$$

Además, $A_{ij} = 0$ si $j < n$ e $i > n$. Un argumento similar prueba que $A_{ij} = (B_2)_{(i-n),(j-n)}$ para $n+1 \leq i, j \leq n+m$ y $A_{ij} = 0$ si $j > n$ e $i < n$. Esto se resume en la condición que $A = B_1 \oplus B_2$.

Ahora supongamos que el teorema vale para $k-1$. Sea $W = W_1 \oplus W_2 \oplus \cdots \oplus W_{k-1}$ y $\beta' = \beta_1 \cup \beta_2 \cup \cdots \cup \beta_{k-1}$. Veamos que W es T-invariante. Sea $w \in W$, de modo que existen $w_1, w_2, \ldots, w_{k-1}$ donde cada $w_i \in W_i$ tales que

$$w = w_1 + w_2 + \cdots + w_{k-1}.$$

Entonces,

$$T(w) = T(w_1) + T(w_2) + \cdots + T(w_{k-1}),$$

pero cada w_i pertenece a un subespacio T-invariante, $T(w_i) \in W_i$ y por ende $T(w) \in W$. Llegados a este punto, por la hipótesis de inducción tenemos que

$$[T_W]_{\beta'} = B_1 \oplus B_2 \oplus \cdots \oplus B_{k-1}.$$

Reduciendo esto al caso con dos subespacios, $V = W \oplus W_k$ y por lo tanto,

$$[T]_\beta = [T_W]_{\beta'} \oplus B_k = (B_1 \oplus B_2 \oplus \cdots \oplus B_{k-1}) \oplus B_k = B_1 \oplus B_2 \oplus \cdots \oplus B_k.$$

□

Proyecciones

Def. Sea P un operador sobre un espacio vectorial V. Decimos que P es una proyección si $P \circ P = P$.

Teorema 5.18. *Sean $P_1, P_2, \ldots, P_k$ proyecciones sobre un espacio vectorial V y supongamos que $W_i = \operatorname{im}(P_i)$ para $1 \leq i \leq k$. Si estas proyecciones son tales que*

1. $\sum_{i=1}^{k} P_i = \mathrm{id}_V$,

2. $P_i \circ P_j = T_0$ *si* $i \neq j$,

entonces $V = \bigoplus_{i=1}^{k} W_k$.

Demostración. Sea $v \in V$ y supongamos que $v_i = P_i(v) \in W_i$ para cada i $(1 \leq i \leq k)$. Luego

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

Esto prueba que $V = \sum_{i=1}^{k} W_i$.

Ahora, fijemos un j $(1 \leq j \leq k)$ y tomemos $v \in W_j \cap \sum_{i \neq j} W_i$. Como $v \in \sum_{i \neq j} W_i$, existen vectores $v_1, \ldots, v_{j-1}, v_{j+1}, \ldots, v_k$ donde cada $v_i \in W_i$ tales que

$$v = \sum_{i \neq j} v_i.$$

Similarmente, como $v \in W_j$, existe $v_j \in V$ tal que $v = P_j(v_j)$. Entonces,

$$v = P_j(v_j) = P_j \circ P_j(v_j) = P_j(v) = P_j \left(\sum_{i \neq j} v_i \right) = P_j \left(\sum_{i \neq j} P_i(v_i) \right) = \sum_{i \neq j} P_j \circ P_i(v_i) = 0.$$

Esto prueba que $W_j \cap \sum_{i \neq j} W_i = \{0\}$ y por ende la suma es directa. □

Lema 5.19. *Sean* $f(t)$ *y* $g(t)$ *polinomios con coeficientes en un campo* F. *Sea además* T *un operador sobre un espacio vectorial* V *sobre* F. *Entonces,*

$$f(T) \circ g(T) = g(T) \circ f(T).$$

Demostración. Basta con notar que para cualesquiera dos escalares $a, b \in F$ y cualesquiera dos enteros no negativos n, m,

$$(aT^n) \circ (bT^m) = (bT^m) \circ (aT^n).$$

□

Lema 5.20. *Sea* $f(t)$ *un polinomio con coeficientes en un campo* F *y sea* T *un operador sobre un espacio vectorial* V *sobre* F. *Entonces,* $\ker(f(T))$ *es* T*-invariante.*

Demostración. Sea $v \in \ker(f(T))$, luego por el lema anterior

$$f(T)(T(v)) = (f(T) \circ T)(v) = (T \circ f(T))(v) = T(f(T)(v)) = T(0) = 0.$$

Esto prueba que $T(v) \in \ker(f(T))$. □

Teorema 5.21. *Sea* $f(t)$ *un polinomio con coeficientes en un campo* F *y sea* T *un operador sobre un espacio vectorial* V *sobre* F *tal que* $f(T) = 0$. *Adicionalmente, supongamos que* $f_1(t)$, $f_2(t)$, $\ldots$, $f_k(t)$ *son polinomios tales que* $f(t) = f_1(t) f_2(t) \cdots f_k(t)$ *y* $\gcd(f_i(t), f_j(t)) = 1$ *para* $i \neq j$. *Entonces,*

$$V = W_1 \oplus W_2 \oplus \cdots \oplus W_k$$

donde $W_i = \ker(f_i(T))$ $(1 \leq i \leq k)$. *Más aún cada* W_i *es* T*-invariante y existen polinomios* $\Pi_1(t), \Pi_2(t), \ldots, \Pi_k(t) \in F[t]$ *tales que cada* $\Pi_i(T)$ *es la proyección sobre* W_i *a lo largo de* $\bigoplus_{j \neq i} W_j$.

Demostración. La existencia de los $\Pi_i(t)$ descritos en el teorema garantiza la condición de la suma directa a través del teorema 5.18. Definimos

$$g_i(t) = \prod_{j \neq i} f_j(t).$$

Entonces, $\gcd(g_1, g_2, \ldots, g_k) = 1$. La identidad de Bézout garantiza la existencia de $h_1(t), h_2(t), \ldots, h_k(t) \in F[t]$ tales que

$$h_1(t)g_1(t) + h_2(t)g_2(t) + \cdots + h_k(t)g_k(t) = 1.$$

Definimos $\Pi_i(t) = h_i(t)g_i(t)$ para cada i $(1 \leq i \leq k)$. Así,

$$\Pi_1(T) + \Pi_2(T) + \cdots + \Pi_k(T) = \mathrm{id}_V.$$

Resta verificar que los $\Pi_i(t)$ satisfacen las condiciones del teorema.

Veamos que cada $\Pi_i(T) \circ \Pi_j(T) = T_0$ para $i \neq j$. Tenemos que

$$\Pi_i(T) \circ \Pi_j(T) = h_i(T) \circ g_i(T) \circ h_j(T) \circ g_j(T) = h_i(T) \circ h_j(T) \circ g_i(T) \circ g_j(T)$$

Como $i \neq j$, $g_i(T)$ contiene el factor $f_j(T)$, obtenemos que

$$\begin{aligned}
\Pi_i(T) \circ \Pi_j(T) &= h_i(T) \circ h_j(T) \circ \prod_{l \neq i,j} f_l(T) \circ f_j(T) \circ g_j(T) \\
&= h_i(T) \circ h_j(T) \circ \prod_{l \neq i,j} f_l(T) \circ f(T)
\end{aligned}$$

(tratamos el producto de las funciones como su composición). Como $f(T) = T_0$, concluimos que

$$\Pi_i(T) \circ \Pi_j(T) = T_0.$$

Ahora, verifiquemos que cada $\Pi_i(T)$ es proyección. Tenemos que

$$\begin{aligned}
\Pi_i(T) &= \Pi_i(T) \circ \mathrm{id}_V = \Pi_i(T) \circ (\Pi_1(T) + \Pi_2(T) + \cdots + \Pi_k(T)) \\
&= \Pi_i(T) \circ \Pi_1(T) + \Pi_i(T) \circ \Pi_2(T) + \cdots + \Pi_i(T) \circ \Pi_i(T) + \cdots + \Pi_i(T) \circ \Pi_k(T).
\end{aligned}$$

Como $\Pi_i(T) \circ \Pi_j(T) = T_0$ siempre que $i \neq j$, la expresión anterior se reduce a

$$\Pi_i(T) = \Pi_i(T) \circ \Pi_i(T).$$

Esto verifica que $\Pi_i(T)$ es proyección.

Resta verificar que $\mathrm{im}(\Pi_i(T)) = \ker(f_i(T))$. Sea $v_i \in \ker(f_i(T))$. Entonces,

$$\begin{aligned}
v_i &= \mathrm{id}_V(v_i) = (\Pi_1(T) + \Pi_2(T) + \cdots + \Pi_k(T))(v_i) \\
&= (h_1(T) \circ g_1(T))(v_i) + (h_2(T) \circ g_2(T))(v_i) + \cdots + (h_k(T) \circ g_k(T))(v_i) \\
&= \left(h_1 \prod_{j_1 \neq 1} f_{j_1}(T) \right)(v_i) + \left(h_2 \prod_{j_1 \neq 2} f_{j_2}(T) \right)(v_i) + \cdots + \left(h_i \prod_{j_i \neq i} f_{j_i}(T) \right)(v_i) + \cdots + \left(h_k \prod_{j_k \neq k} f_{j_k}(T) \right)(v_i)
\end{aligned}$$

Todos los términos salvo el i-ésimo tienen a $f_i(T)$ como factor y por lo tanto se anulan. Así,

$$v_i = \left(h_i \prod_{j \neq i} f_j(T) \right)(v_i) = \Pi_i(T)(v_i).$$

De aquí concluimos que $v_i \in \operatorname{im}(\Pi_i(T))$ y en consecuencia $\ker(f_i(T)) \subseteq \operatorname{im}(\Pi_i(T))$. Por otro lado, tomemos $w_i \in \operatorname{im}(\Pi_i(T))$, de modo que existe $v_i \in V$ tal que $w_i = \Pi_i(T)(v_i)$, luego

$$\begin{aligned} f_i(T)(w_i) &= f_i(T)(\Pi_i(T)(v_i)) = (f_i(T) \circ \Pi_i(T))(v_i) = (f_i(T) \circ h_i(T) \circ g_i(T))(v_i) \\ &= (h_i(T) \circ f_i(T) \circ g_i(T))(v_i) = (h_i(T) \circ f(T))(v_i) \\ &= h_i(T)(0) = 0, \end{aligned}$$

lo cual nos permite concluir que $w_i \in \ker(f_i(T))$ y por ende $\operatorname{im}(\Pi_i(T)) \subseteq \ker(f_i(T))$. Esto junto con la continencia anterior implica que $\operatorname{im}(\Pi_i(T)) = \ker(f_i(T))$. □

5.4. Forma normal de Jordan

La forma normal de Jordan es especialmente relevante para describir operadores con la propiedad de nilpotencia.

Def. Sea T un operador sobre un espacio vectorial V. Decimos que T es **nilpotente** si existe un entero positivo r tal que $T^r = T_0$. Al mínimo de estos r lo llamamos **grado de nilpotencia** de T.

Teorema 5.22. *Sea T un operador nilpotente de grado r y sea $K_i = \ker(T^i)$ $(n = 0, 1, 2, \ldots, r)$. Entonces tenemos una cadena de inclusiones estrictas*

$$\{0\} = K_0 \subset K_1 \subset K_2 \subset \cdots \subset K_r = V.$$

Demostración. El caso en que $r = 1$ corresponde a T_0. Tendríamos que

$$\{0\} = K_0 < K_1 = V.$$

Supongamos ahora que $r > 1$, entonces $T^{r-1} \neq T_0$ y existe $v \in V$ tal que $T^{r-1}(v) \neq 0$. Luego, $v \in K_r \setminus K_{r-1}$. En general, para $i = 0, 1, \ldots, r-1$, $0 \neq T^{r-1}(v) = T^i(T^{r-1-i}(v))$ pero $0 = T^r(v) = T^{i+1}(T^{r-1-i}(v))$, luego $T^{r-1-i}(v) \in K_{i+1} \setminus K_i$. □

Corolario 5.22.1. *Sea T un operador nilpotente sobre un espacio vectorial V de dimensión finita. Entonces, el grado de nilpotencia de T es menor o igual a la dimensión de V.*

Observación. El núcleo de todo operador es un subespacio vectorial, la cadena de inclusiones implica que cada K_i es subespacio de K_{i+1}.

Teorema 5.23. *Sea T un operador sobre un espacio vectorial V y supongamos que $v \in V$ es un vector tal que $T^r(v) = 0$ pero $T^{r-1}(v) = 0$. Sea W el subespacio T-cíclico generado por v. Entonces,*

$$\beta = \{T^{r-1}(v), \ldots, T(v), v\}$$

es una base de de W y

$$[T_W]_\beta = \begin{bmatrix} 0 & 1 & 0 & \ldots & 0 \\ 0 & 0 & 1 & \ldots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \ldots & 1 \\ 0 & 0 & 0 & \ldots & 0 \end{bmatrix}.$$

Demostración. Por definición, $W = \operatorname{span}(\{v, T(v), T^2(v), \ldots\})$. Como $T^k(v) = 0$ para $k \geq r$, esto se reduce a $W = \operatorname{span}(\{v, T(v), T^2(v), \ldots, T^{r-1}(v)\})$. Por lo tanto $\dim(W)$ es a lo sumo r. Sean $a_0, a_1, \ldots, a_{r-1}$ escalares tales que

$$a_0 v + a_1 T(v) + \cdots + a_{r-1} T^{r-1}(v) = 0.$$

Aplicamos T^{r-1} a ambos lados de la igualdad y obtenemos que $a_0 T^{r-1}(v) = \mathit{0}$, pero $T^{r-1}(v) \neq \mathit{0}$ así que $a_0 = 0$. Inductivamente, si $a_0 = a_1 = \cdots = a_{i-1} = 0$, para $i < r-1$, entonces aplicamos T^{r-1-i} a ambos lados de la misma igualdad y concluimos que $a_i = 0$. Esto prueba que $\{v, T(v), T^2(v), \ldots, T^{r-1}(v)\}$ es linealmente independiente.

Por la definición de la base ordenada β, el i-ésimo vector de la base es $T^{r-i}(v)$ y por lo tanto su imagen es $T^{r-i+1}(v)$, esto es el $i-1$-ésimo vector de la base ($i > 1$). Así, la i-ésima columna de la representación matricial es e_{i-1} (para $i > 1$). Para $i = 1$, el vector correspondiente es $T^{r-1}(v)$, cuya imagen mediante T es $T^r(v) = \mathit{0}$ y por ende la primera columna es el vector cero. $\square$

Def. Sea J una matriz de tamaño $n \times n$ de la forma

$$J = \begin{bmatrix} 0 & 1 & 0 & \ldots & 0 \\ 0 & 0 & 1 & \ldots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \ldots & 1 \\ 0 & 0 & 0 & \ldots & 0 \end{bmatrix}.$$

Decimos que J es un **bloque de Jordan** correspondiente a 0.

Teorema 5.24. *Sea T un operador nilpotente sobre un espacio vectorial V de dimensión finita. Entonces, existe una base β para la cual $[T]_\beta$ es una matriz diagonal por bloques de Jordan.*

Demostración. Sea r el grado de nilpotencia de T y para $i = 0, 1, \ldots, r$ denotemos $K_i = \ker(T^i)$. Por el teorema 5.22, tenemos que

$$\{\mathit{0}\} = K_0 \subset K_1 \subset \cdots \subset K_r = V.$$

Sean $K_1', K_2', \ldots, K_r' \subseteq V$ tales que

$$\begin{aligned} K_1 &= K_0 \oplus K_1' \\ K_2 &= K_1 \oplus K_2' \\ &\vdots \\ K_r &= K_{r-1} \oplus K_r' \end{aligned}$$

Note que si $v_i \in K_i' \setminus \{\mathit{0}\}$, entonces $T^i(v_i) = \mathit{0}$ pero $T^{i-1}(v_i) \neq \mathit{0}$. Denotemos $n_i = \dim(K_i)$ y $m_i = \dim(K_i')$. Así, $n_i = n_{i-1} + m_i$ para $i = 1, 2, \ldots, r$ y $n = n_1 + m_2 + \cdots + m_{r-1} + m_r$. Sea $\{v_{r,1}, v_{r,2}, \ldots, v_{r,m_r}\}$ una base de K_r'. Veamos que

$$\{v_{r,1}, v_{r,2}, \ldots, v_{r,m_r}, T(v_{r,1}), T(v_{r,2}), \ldots, T(v_{r,m_r})\}$$

es linealmente independiente. Supongamos que

$$\sum_{i=1}^{m_r} a_i v_{r,i} + \sum_{j=1}^{m_r} b_j T(v_{r,j}) = \mathit{0}.$$

Es evidente que la primera sumatoria pertenece a K_i' y la segunda a K_{r-1}. Definamos

$$v = \sum_{i=1}^{m_r} a_i v_{r,i} = -\sum_{j=1}^{m_r} b_j T(v_{r,j}).$$

Es claro que $v \in K_{r-1} \cap K'_r$, pero como la suma de estos dos subespacios es directa, tenemos que $K_{r-1} \cap K'_r = \{0\}$. Por lo tanto, $v = 0$. Esto implica que $a_1 = a_2 = \cdots = a_{m_r} = 0$ y que

$$T\left(\sum_{j=1}^{m_r} b_j v_{r,j}\right) = 0.$$

Luego, $\sum_{j=1}^{m_r} b_j v_{r,j} \in K_1 \subset K_{r-1}$ pero $b_j v_{r,j} \in K_r$, así que

$$\sum_{j=1}^{m_r} b_j v_{r,j} = 0,$$

y al ser $\{v_{r,1}, v_{r,2}, \ldots, v_{r,m_r}\}$ base de K'_r, $b_1 = b_2 = \cdots = b_{m_r} = 0$.

Tenemos entonces que $\{T(v_{r,1}), T(v_{r,2}), \ldots, T(v_{r,m_r})\}$ es linealmente independiente. Completamos este conjunto a una base de K'_{r-1}:

$$\{T(v_{r,1}), T(v_{r,2}), \ldots, T(v_{r,m_r}), v_{r-1,m_r+1}, v_{r-1,m_r+2}, \ldots, v_{r-1,m_{r-1}}\}.$$

Un argumento similar al anterior muestra que

$$\{v_{r,i}\}_{i=1}^{m_r} \cup \{T(v_{r,i})\}_{i=1}^{m_r} \cup \{v_{r-1,i}\}_{i=m_r+1}^{m_{r-1}} \cup \{T^2(v_{r,i})\}_{i=1}^{m_r} \cup \{T(v_{r-1,i})\}_{i=m_r+1}^{m_{r-1}}$$

es linealmente independiente.

Así, inductivamente construimos una base con $n_1 + m_2 + \cdots + m_{r-1} + m_r$ vectores que resulta ser una base de V. Denotemos $v_{j-1,i} = T(v_{j,i})$ con $j = 1, 2, \ldots, r$ e $i = 1, 2, \ldots, m_j$. Tomemos W_i como el subespacio T-cíclico generado por $v_{k,i}$ donde k es el máximo entero posible para un i dado. Por el teorema anterior, $\beta_i = \{T^{k-1}(v_{k,i}), \ldots, T(v_{k,i}), v_{k,i}\}$ es una base de W_i y $[T_{W_i}]_{\beta_i}$ tiene la forma requerida. Por construcción tenemos que $\beta_1 \cup \beta_2 \cup \cdots \cup \beta_{n_1}$ es una base de V. Por el teorema 5.8, esto es equivalente a que $V = W_1 \oplus W_2 \oplus \cdots \oplus W_{n_1}$. Llegados a este punto, el teorema 5.17 garantiza el resultado. □

Observación. En la notación de la prueba anterior n_1 es la cantidad de bloques de tamaño mayor o igual a 1 y m_i es la cantidad de bloques de tamaño mayor o igual a i.

Lema 5.25. *Sea T un operador sobre un espacio vectorial de dimensión finita V. Supongamos además que $U = T - \lambda \mathrm{id}_V$ es nilpotente y β es una base como la de la conclusión del teorema 5.24. Entonces,*

$$[T]_\beta = [U]_\beta + \lambda I_{\dim(V)}.$$

Esta representación matricial corresponde a una matriz diagonal por bloques de Jordan con λ en la diagonal.

Demostración. Tenemos que $U = T - \lambda \mathrm{id}_V$, de modo que

$$\begin{aligned}[U]_\beta &= [T - \lambda \mathrm{id}_V]_\beta \\ &= [T]_\beta - \lambda [\mathrm{id}_V]_\beta\end{aligned}$$

Sumando $\lambda[\mathrm{id}_V]_\beta$ a ambos lados de esta igualdad y reconociendo que $[\mathrm{id}_V]_\beta = I_{\dim(V)}$, obtenemos que

$$[T]_\beta = [U]_\beta + \lambda I_{\dim(V)}.$$

La forma de la representación matricial se obtiene con un cálculo directo. □

Def. Sea J una matriz de tamaño $n \times n$ de la forma

$$J = \begin{bmatrix} \lambda & 1 & 0 & \dots & 0 & 0 \\ 0 & \lambda & 1 & \dots & 0 & 0 \\ 0 & 0 & \lambda & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & \lambda & 1 \\ 0 & 0 & 0 & \dots & 0 & \lambda \end{bmatrix}.$$

J es un **bloque de Jordan** de tamaño n correspondiente λ. El bloque de Jordan de tamaño $n \times n$ correspondiente a λ se denota $J_{n,\lambda}$.

Teorema 5.26 (Teorema de Jordan). *Sea T un operador sobre un espacio vectorial de dimensión finita V y supongamos que*

$$f(t) = (t - \lambda_1)^{m_1}(t - \lambda_2)^{m_2} \cdots (t - \lambda_k)^{m_k},$$

donde $\lambda_1, \lambda_2, \dots, \lambda_k$ son escalares, es el polinomio característico de T. Entonces existe una base β de V tal que la representación matricial de T en β es una matriz diagonal por bloques de Jordan.

Demostración. Supongamos que $\lambda_i \neq \lambda_j$ cuando $i \neq j$. Esto garantiza que m_i es la multiplicidad algebraica de λ_i y además que $\gcd((t - \lambda_i)^{m_i}, (t - \lambda_j)^{m_j}) = 1$. El teorema 5.21implica que

$$V = W_1 \oplus W_2 \oplus \cdots \oplus W_k \tag{1}$$

donde $W_i = \ker((T - \lambda_i \mathrm{id}_V)^{m_i})$ $(1 \leq i \leq k)$. Adicionalmente, ese teorema garantiza cada W_i es T-invariante. Definimos T_i como la restricción de T a W_i. El polinomio característico de T_i es $f_i(t) = (t - \lambda_i)^{m_i}$ y por el teorema de Cayley-Hamilton, tenemos que $f_i(T_i) = (T_i - \lambda_i \mathrm{id}_{W_i})^{m_i} = T_0$ y por lo tanto $T_i - \lambda_i \mathrm{id}_{W_i}$ es nilpontente de grado menor o igual a m_i. Por el lema 5.25, existe una base β_i de W_i tal que $[T_i]_{\beta_i}$ es diagonal en bloques de Jordan correspondientes a λ_i. El teorema se sigue de aplicar el teorema 5.17 a la ecuación (1). ☐

Teorema 5.27 (Descomposición de Jordan-Chevalley). *Sea T un operador sobre un espacio vectorial de dimensión finita V. Supongamos que*

$$f(t) = (t - \lambda_1)^{m_1}(t - \lambda_2)^{m_2} \cdots (t - \lambda_k)^{m_k},$$

donde $\lambda_1, \lambda_2, \dots, \lambda_k$ son escalares distintos, es el polinomio característico de T. Entonces, existen dos polinomios $f_d(t)$ y $f_n(t)$ tales que si $D = f_d(T)$ y $N = f_n(T)$,

(*a*) *$T = N + D$;*

(*b*) *D es diagonalizable y N es nilpotente;*

(*c*) *$DN = ND$.*

Demostración. Sean $f_i(t) = (t - \lambda_i)^{m_i}$ $(1 \leq i \leq k)$ y $W_i = \ker(f_i(T))$. Por el teorema 5.21, existen polinomios $\Pi_1(t), \Pi_2(t), \dots, \Pi_k(t) \in F[t]$ tales que cada $\Pi_i(T)$ es la proyección sobre W_i a lo largo de $\bigoplus_{j \neq i} W_i$. Definimos

$$f_d(t) = \sum_{i=1}^{k} \lambda_i \Pi_i(t),$$

$$f_n(t) = t - f_d(t).$$

Veamos que estos polinomios satisfacen las condiciones (a), (b) y (c).

(a) $D + N = f_d(T) + f_n(T) = f_d(T) + T - f_d(T) = T$.

(b) Por definición, cada W_i es un espacio propio de D correspondiente al valor propio λ_i. Entonces, por el teorema 5.10 D es diagonalizable.

Por otra parte, note que si $w \in W_i$, para algún i $(1 \leq i \leq k)$

$$N(w) = (T - \lambda_i \mathrm{id}_V)(w)$$

y por el teorema de Cayley-Hamilton, $N^{m_i}(w) = 0$. Como $V = \bigoplus_{i=1}^k W_i$, para todo $v \in V$ existen $w_1, w_2, \ldots, w_k$ donde cada $w_i \in W_i$ $(1 \leq i \leq k)$ tales que

$$v = w_1 + w_2 + \cdots + w_k$$

y en consecuencia $N^l(v) = 0$ donde $l = \max\{m_1, m_2, \ldots, m_k\}$. Esto prueba que N es nilpotente.

(c) Por el lema 5.19, $DN = f_d(T) \circ f_n(T) = f_n(T) \circ f_d(T) = ND$.

$\square$

6. Espacios euclídeos y producto interno

Hasta ahora, solo hemos estudiado espacios vectoriales sin una estructural adicional. En esta sección y la siguiente estudiamos espacios vectoriales dotados de una estructura adicional de significancia geométrica.

Def. Sea V un espacio vectorial sobre $\mathbb{R}$. Un **producto interno** es una función

$$\begin{aligned} \langle\ ,\ \rangle : V \times V &\to \mathbb{R} \\ (u, v) &\mapsto \langle u, v\rangle \end{aligned}$$

tal que se satisfacen las siguientes condiciones:

1. Para todo $u, v, w \in V$, $\langle u + w, v\rangle = \langle u, v\rangle + \langle w, v\rangle$.
2. Para todo $u, v \in V$ y $c \in \mathbb{R}$, $\langle cu, v\rangle = c\,\langle u, v\rangle$.
3. Para todo $u, v \in V$, $\langle u, v\rangle = \langle v, u\rangle$.
4. Para todo $v \in V$, $\langle v, v\rangle > 0$ si $v \neq 0$.

Decimos un espacio vectorial sobre $\mathbb{R}$ provisto de un producto interno es un **espacio euclídeo**.

Observación. Las propiedades 1 y 2 en la definición anterior implican que la función es lineal en la primera componente. Esto se extiende a la segunda componente también usando la propiedad de simetría (3).

Def. Sea V un espacio euclídeo y $v \in V$. Definimos la norma de v como $\|v\| = \sqrt{\langle v, v\rangle}$.

Teorema 6.1. *Sea V un espacio euclídeo. Para todo $u, v \in V$ y $c \in \mathbb{R}$, las siguientes afirmaciones son verdaderas:*

(a) $\|cv\| = |c|\,\|v\|$,

(b) $\|v\| = 0$ *si y solo si* $v = 0$,

(c) *(Desigualdad de Cauchy-Schwarz)* $|\langle u,v\rangle| \leq \|u\|\,\|v\|$,

(d) *(Desigualdad triangular)* $\|u+v\| \leq \|u\| + \|v\|$.

Demostración. (a) $\|cv\| = \sqrt{\langle cv, cv\rangle} = \sqrt{c^2\langle v,v\rangle} = |c|\sqrt{\langle v,v\rangle} = |c|\,\|v\|$.

(b) Tomemos un vector no nulo $v \in V$. Entonces,

$$\|v\|^2 = \langle v,v\rangle = \langle v+\mathit{0}, v+\mathit{0}\rangle = \langle v,v\rangle + \langle \mathit{0},v\rangle + \langle v,\mathit{0}\rangle + \langle \mathit{0},\mathit{0}\rangle .$$

Pero $\langle \mathit{0},v\rangle = \langle 0\cdot \mathit{0}, v\rangle = 0\,\langle \mathit{0},v\rangle = 0$ y lo mismo para $\langle v,\mathit{0}\rangle$, de modo que

$$\langle v,v\rangle = \langle v,v\rangle + \langle \mathit{0},\mathit{0}\rangle$$

y en consecuencia $\langle \mathit{0},\mathit{0}\rangle = 0$.

Recíprocamente, si $\|v\| = 0$ la propiedad 4 implica que $v = \mathit{0}$.

(c) Para $u = \mathit{0}$ o $v = \mathit{0}$ es evidente. Asumamos que $u, v \neq \mathit{0}$. Para todo $x \in \mathbb{R}$ tenemos que

$$\begin{aligned}
0 &\leq \|xu - v\|^2 = \langle xu - v, xu - v\rangle = \langle xu, xu - v\rangle - \langle v, xu - v\rangle \\
&\leq x(\langle u, xu\rangle - \langle u,v\rangle) - (\langle v, xu\rangle - \langle v,v\rangle) = \|u\|^2 x^2 - x\langle u,v\rangle - x\langle v,u\rangle + \|v\|^2 \\
&\leq \|u\|^2 x^2 - 2x\langle u,v\rangle + \|v\|^2 .
\end{aligned}$$

Esto es un polinomio cuadrático en x que no cambia de signo, de modo que su determinante es negativo o nulo, esto es

$$(-2\langle u,v\rangle)^2 - 4\|u\|^2\|v\|^2 \leq 0 \Rightarrow \langle u,v\rangle^2 \leq \|u\|^2\|v\|^2$$

(d) Tenemos que

$$\begin{aligned}
\|u+v\|^2 &= \langle u+v, u+v\rangle \\
&= \|u\|^2 + 2\langle u,v\rangle + \|v\|^2 .
\end{aligned}$$

Por la desigualdad de Cauchy-Schwartz esto último es menor o igual a $\|u\|^2 + 2\|u\|\,\|v\| + \|v\|^2$, i.e.

$$\begin{aligned}
\|u+v\|^2 &= \|u\|^2 + 2\langle u,v\rangle + \|v\|^2 \\
&\leq \|u\|^2 + 2\|u\|\,\|v\| + \|v\|^2 \\
&\leq (\|u\| + \|v\|)^2.
\end{aligned}$$

□

6.1. Ortogonalidad

De la desigualdad de Cauchy-Schwarz tenemos que

$$-1 \leq \frac{\langle u,v\rangle}{\|u\|\,\|v\|} \leq 1.$$

Entonces podemos definir el ángulo entre u y v como

$$\cos(\theta) = \frac{\langle u,v\rangle}{\|u\|\,\|v\|}.$$

Def. Sea V un espacio euclídeo y $u, v \in V$. Decimos que u y v son **ortogonales** si $\langle u, v\rangle = 0$. Además, decimos que u es **unitario** si $\|u\| = 1$.

Def. Sea V un espacio euclídeo y $S \subseteq V$ un subconjunto no vacío. Decimos que S es **ortogonal** para todo $u, v \in S$ con $u \neq v$ $\langle u, v\rangle = 0$. Decimos que S es **ortonormal** si es ortogonal y todo $u \in S$ es unitario.

Teorema 6.2. *Sea V un espacio euclídeo y $S = \{v_1, v_2, \ldots, v_k\}$ un subconjunto ortonormal de V. S es linealmente independiente.*

Demostración. Tomemos $0 = a_1 v_1 + a_2 v_2 + \cdots + a_k v_k$. Tenemos que

$$0 = \langle 0, v_j\rangle = \sum_{i=1}^{k} a_i \langle v_i, v_j\rangle = a_j$$

para todo $j = 1, 2, \ldots, k$, puesto que $\langle v_j, v_j\rangle = 1$. □

Teorema 6.3. *Sea V un espacio euclídeo de dimensión finita y $\beta = \{u_1, u_2, \ldots, u_n\}$ una base ortonormal de V. Sean además*

$$v_1 = \sum_{i=1}^{n} a_i u_i; \qquad v_2 = \sum_{i=1}^{n} b_i u_i.$$

Entonces,

$$\langle v_1, v_2\rangle = \sum_{i=1}^{n} a_i b_i.$$

Demostración. Tenemos que

$$\langle v_1, v_2\rangle = \left\langle \sum_{i=1}^{n} a_i u_i, \sum_{j=1}^{n} b_j u_j \right\rangle = \sum_{i=1}^{n}\sum_{j=1}^{n} a_i b_j \langle u_i, u_j\rangle = \sum_{i=1}^{n}\sum_{j=1}^{n} a_i b_j \delta_{ij} = \sum_{i=1}^{n} a_i b_i.$$

□

Observación. El teorema anterior muestra que todo espacio euclídeo de dimensión finita es idéntico a $\mathbb{R}^n$. Por el teorema 2.21, V es isomorfo a $\mathbb{R}^n$, pero la estructura que brinda el producto interno implica un grado superior de similaridad.

Teorema 6.4. *Sea V un espacio euclídeo de dimensión finita y $\beta = \{u_1, u_2, \ldots, u_n\}$ es una base ortonormal de V. Entonces, para todo $v \in V$*

$$v = \sum_{i=1}^{n} \langle v, u_i\rangle u_i.$$

Demostración. Como β es una base de V, existen $a_1, a_2, \ldots, a_n \in \mathbb{R}$ tales que $v = \sum_{i=1}^{n} a_i u_i$. En consecuencia, para $j = 1, 2, \ldots, n$.

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

□

Las bases ortonormales son muy útiles y son una representación muy natural en varias circunstancias. Es por esto que necesitamos una forma de construir bases ortonormales. La construcción de una base ortnormal a partir de cualquier base es el objetivo del proceso de Gram-Schmidt.

Lema 6.5. *Sea V un espacio euclídeo y $S = \{v_1, v_2, \ldots, v_k\}$ un subconjunto ortogonal de V conformado por vectores no nulos. Entonces, S es linealmente independiente.*

Demostración. Similar a la prueba del teorema 6.2. □

Teorema 6.6 (Ortogonalización de Gram-Schmidt)**.** *Sea V un espacio euclídeo de dimensión finita y $\beta = \{w_1, w_2, \ldots, w_k\}$ una base de V. Defina $\beta' = \{v_1, v_2, \ldots, v_k\}$, donde $v_1 = w_1$ y*

$$v_i = w_i - \sum_{j=1}^{i-1} \frac{\langle w_i, v_j \rangle}{\|v_j\|^2} v_j \qquad \text{para } 2 \leq i \leq n.$$

Entonces, β' es una base ortogonal de V. Más aún, $\gamma = \{u_1, u_2, \ldots, u_n\}$, donde $u_i = v_i / \|v_i\|$, es una base ortonormal de V.

Demostración. Primero, utilizamos un argumento inductivo para mostrar que β' es ortogonal. Queremos ver que para un k fijo ($1 \leq k < n$), $\langle v_{k+1}, v_j \rangle = 0$ ($1 \leq j \leq k$). Tomemos $k = 1$.

$$\begin{aligned}
\langle v_2, v_1 \rangle &= \left\langle w_2 - \frac{\langle w_2, v_1 \rangle}{\|v_1\|^2} v_1, v_1 \right\rangle \\
&= \langle w_2, v_1 \rangle - \frac{\langle w_2, v_1 \rangle}{\|v_1\|^2} \langle v_1, v_1 \rangle \\
&= \langle w_2, v_1 \rangle - \langle w_2, v_1 \rangle \\
&= 0
\end{aligned}$$

Ahora, supongamos que $\langle v_i, v_j \rangle = 0$ para $1 \leq i < j \leq k$. Tenemos entonces que para $1 \leq j \leq k$

$$\begin{aligned}
\langle v_{k+1}, v_j \rangle &= \left\langle w_{k+1} - \sum_{i=1}^{k} \frac{\langle w_{k+1}, v_i \rangle}{\|v_i\|^2} v_i, v_j \right\rangle \\
&= \langle w_{k+1}, v_j \rangle - \sum_{i=1}^{k} \frac{\langle w_{k+1}, v_i \rangle}{\|v_i\|^2} \langle v_i, v_j \rangle \\
&= \langle w_{k+1}, v_j \rangle - \langle w_{k+1}, v_j \rangle \\
&= 0.
\end{aligned}$$

Ahora, veamos que β' genera a V. Definimos $S_k = \{w_1, w_2, \ldots, w_k\}$ y $S'_k = \{v_1, v_2, \ldots, v_k\}$ para $1 \leq k \leq n$. Por la fórmula recursiva

$$v_i = w_i - \sum_{j=1}^{i-1} \frac{\langle w_i, v_j \rangle}{\|v_j\|^2} v_j,$$

tenemos que $\operatorname{span}(S'_k) \subseteq \operatorname{span}(S_k)$, pero por el lema anterior, S'_k es linealmente independiente y por lo tanto $\dim(\operatorname{span}(S'_k)) = \dim(\operatorname{span}(S_k)) = k$. En consecuencia, $\operatorname{span}(S'_k) = \operatorname{span}(S_k)$. Note que $S_n = \beta$ y $S'_n = \beta'$, lo cual nos permite concluir que $\operatorname{span}(\beta') = \operatorname{span}(\beta) = V$.

Finalmente, note que

$$\langle u_i, u_i \rangle = \left\langle \frac{v_i}{\|v_i\|}, \frac{v_i}{\|v_i\|} \right\rangle = \frac{\langle v_i, v_i \rangle}{\|v_i\|^2} = 1$$

para $1 \leq i \leq n$. Similarmente, si $i \neq j$,

$$\langle u_i, u_j \rangle = \left\langle \frac{v_i}{\|v_i\|}, \frac{v_j}{\|v_j\|} \right\rangle = \frac{\langle v_i, v_j \rangle}{\|v_i\| \|v_j\|} = 0.$$

Esto prueba que γ es ortonormal. Por el teorema 6.2, esto implica que γ es linealmente independiente y al tener n elementos, es una base de V. $\square$

Def. Sea V un espacio euclídeo y $S \subset V$. Definimos el conjunto ortogonal a S por

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

Teorema 6.7. *Sea V un espacio euclídeo y $S \subset V$. Entonces, $S^\perp$ es un subespacio de V.*

Demostración. Como $\langle 0, v \rangle = 0$ para todo $v \in V$, $0 \in S^\perp$. Ahora, tomemos $v, w \in S^\perp$ y $a \in \mathbb{R}$. Para todo $u \in S$,

$$\begin{aligned} \langle v + w, u \rangle &= \langle v, u \rangle + \langle w, u \rangle = 0 + 0 = 0 \\ \langle av, u \rangle &= a \langle v, u \rangle = a \cdot 0 = 0 \end{aligned}$$

luego $v + w \in S^\perp$ y $av \in S^\perp$; así, $S^\perp$ es un subespacio de V. $\square$

Teorema 6.8. *Sea V un espacio euclídeo de dimensión finita y sea W un subespacio de V. Entonces, $V = W \oplus W^\perp$.*

Demostración. Sea $n = \dim(V)$, $m = \dim(W)$ y $\{v_1, v_2, \ldots, v_n\}$ una base de V tal que los primeros m vectores de esta base formen una base de W. Sea $\{u_1, u_2, \ldots, u_n\}$ la base ortonormal obtenida a partir de β mediante el proceso de Gram-Schmidt. Por construcción, tenemos que $\operatorname{span}(\{u_1, u_2, \ldots, u_m\}) = W$ y $\operatorname{span}(\{u_{m+1}, u_{m+2}, \ldots, u_n\}) = W^\perp$. Como $\{u_{m+1}, u_{m+2}, \ldots, u_n\}$ es ortonormal, también es linealmente independiente y por lo tanto es una base de $W^\perp$. Por el teorema 5.8, esto implica que $V = W \oplus W^\perp$. $\square$

Def. Sea V un espacio euclídeo de dimensión finita y W un subespacio de V. Sea además $\beta_W = \{u_1, u_2, \ldots, u_m\}$ una base ortonormal de W. Definimos la **proyección ortogonal** sobre W, denotado como el operador lineal sobre V dado por

$$\operatorname{proj}_W(v) = \sum_{i=1}^{m} \langle v, u_i \rangle u_i \qquad \text{para todo } v \in V.$$

6.2. Operador adjunto

Def. Sea V un espacio euclídeo y T un operador lineal sobre V. Decimos que $U : V \to V$ es **operador adjunto** de T si para todo $v, w \in V$,

$$\langle U(v), w \rangle = \langle v, T(w) \rangle.$$

Decimos que T es auto-adjunto si T es un operador adjunto de T.

Observación. Si U es adjunto de T, entonces T es adjunto de U.

$$\langle T(v), w \rangle = \langle w, T(v) \rangle = \langle U(w), v \rangle = \langle v, U(w) \rangle.$$

Teorema 6.9. *Sea V un espacio euclídeo y P una proyección ortogonal sobre un subespacio W de V. P es auto-adjunto.*

Demostración. Sean $v, w \in V$. Entonces

$$\langle P(v), w\rangle = \langle P(v), (w - P(w)) + P(w)\rangle = \langle P(v), w - P(w)\rangle + \langle P(v), P(w)\rangle .$$

Como $P(v) \in W$ y $w - P(w) \in W^{\perp}$, el primer término es cero y tenemos que $\langle P(v), w\rangle = \langle P(v), P(w)\rangle$. Ahora,

$$\begin{aligned}\langle v, P(w)\rangle &= \langle (v - P(v)) + P(v), P(w)\rangle \\ &= \langle v - P(v), P(w)\rangle + \langle P(v), P(w)\rangle \\ &= \langle P(v), P(w)\rangle = \langle P(v), w\rangle .\end{aligned}$$

□

Teorema 6.10. *Sea V un espacio euclídeo de dimensión finita y sea T un operador lineal sobre V. Entonces existe una función única $T^* : V \to V$ tal que para todo $v, w \in V$,*

$$\langle T^*(v), w\rangle = \langle v, T(w)\rangle .$$

Esto es, el operador adjunto existe y es único.

Demostración. Sea $\beta = \{u_1, u_2, \ldots, u_n\}$ un base ortonormal de V. Definimos el operador $T^* : V \to V$ por

$$T^*(v) = \sum_{i=1}^{n} \langle v, T(u_i)\rangle\, u_i.$$

De esta definición tenemos que para todo $u_i, u_j \in \beta$,

$$\langle T^*(u_j), u_i\rangle = \langle u_j, T(u_i)\rangle$$

y por las propiedades del producto interno, T^* es adjunto de T.

Por otro lado, si $U : V \to V$ es adjunto de T, por el teorema 6.4,

$$\begin{aligned}U(v) &= \sum_{i=1}^{n} \langle U(v), u_i\rangle\, u_i \\ &= \sum_{i=1}^{n} \langle v, T(u_i)\rangle\, u_i \\ &= T^*.\end{aligned}$$

Esto prueba la unicidad del operador adjunto. □

Teorema 6.11. *Sea V un espacio euclídeo de dimensión finita, T un operador sobre V y β una base ortonormal. Entonces,*

$$[T^*]_\beta = ([T]_\beta)^t .$$

Demostración. Sea $\beta = \{u_1, u_2, \ldots, u_n\}$. Si $A = [T]_\beta$ y $B = [T^*]_\beta$, tenemos que

$$A_{ij} = [T(u_j)]_{\beta,i} = \langle T(u_j), u_i\rangle \qquad \text{y} \qquad B_{ji} = [T^*(u_i)]_{\beta,j} = \langle T^*(u_i), u_j\rangle .$$

Por lo tanto,

$$B_{ji} = \langle T^*(u_i), u_j\rangle = \langle u_i, T(u_j)\rangle = \langle T(u_j), u_i\rangle = A_{ij}.$$

Esto prueba que $A^t = B$. □

Corolario 6.11.1. *Sea V un espacio euclídeo de dimensión finita y T un operador sobre V. Las siguientes dos condiciones son equivalentes.*

1. *T es auto-adjunto.*
2. *La representación matricial de T respecto a cualquier base ortonormal es simétrica.*

6.3. Teorema espectral

Lema 6.12. *Sea V un espacio euclídeo de dimensión finita y T un operador lineal sobre V. Supongamos que T es auto-adjunto. Entonces,*

(*a*) *para todo $f(t) \in \mathbb{R}[t]$, $f(T)$ es auto-adjunto;*

(*b*) *si T es nilpotente, $T = T_0$;*

(*c*) *si $v, w \in V$ son vectores propios asociados a valores propios distintos, entonces $\langle v, w \rangle = 0$.*

Demostración. (*a*) Primero, mostramos por inducción que T^k es auto-adjunto para todo entero no negativo k. El caso base se establece al decir que T es auto-adjunto. Ahora, supongamos que T^{k-1} es auto-adjunto. Entonces,

$$\left\langle T^k(v), w \right\rangle = \left\langle T^{k-1}(T(v)), w \right\rangle = \left\langle T(v), T^{k-1}(w) \right\rangle = \left\langle v, T^k(w) \right\rangle.$$

Ahora, si $f(t) = \sum_{i=0}^{m} a_i t^i$, entonces para todo $v, w \in V$,

$$\begin{aligned} \langle f(T)(v), w \rangle &= \left\langle \sum_{i=0}^{m} a_i T^i(v), w \right\rangle \\ &= \sum_{i=0}^{m} a_i \left\langle T^i(v), w \right\rangle = \sum_{i=0}^{m} a_i \left\langle v, T^i(w) \right\rangle \\ &= \left\langle v, \sum_{i=0}^{m} a_i T^i(w) \right\rangle \\ &= \langle v, f(T)(w) \rangle. \end{aligned}$$

(*b*) Sea r el grado de nilpotencia de T. Supongamos, buscando una contradicción, que $r > 1$ ($r \geq 2$) y existe $v \in V$ tal que $T^{r-1}(v) \neq 0$. En tal caso,

$$\left\| T^{r-1}(v) \right\|^2 = \left\langle T^{r-1}(v), T^{r-1}(v) \right\rangle = \left\langle T^r(v), T^{r-2}(v) \right\rangle = \left\langle 0, T^{r-2}(v) \right\rangle = 0,$$

lo cual implica por el teorema 6.1 que $T^{r-1}(v) = 0$, contradicción. Por lo tanto, $r = 1$ y $T = T_0$.

(*c*) Sean $\lambda_1, \lambda_2 \in \mathbb{R}$ valores propios distintos tales que $T(v_1) = \lambda_1 v_1$ y $T(v_2) = \lambda_2 v_2$. Así,

$$\lambda_1 \langle v_1, v_2 \rangle = \langle \lambda_1 v_1, v_2 \rangle = \langle T(v_1), v_2 \rangle = \langle v_1, T(v_2) \rangle = \langle v_1, \lambda_2 v_2 \rangle = \lambda_2 \langle v_1, v_2 \rangle.$$

Luego,

$$(\lambda_1 - \lambda_2) \langle v_1, v_2 \rangle = 0,$$

pero como λ_1 y λ_2 son distintos, $\langle v_1, v_2 \rangle = 0$.

□

Teorema 6.13 (Teorema Espectral)**.** *Sea V un espacio euclídeo de dimensión finita y T un operador auto-adjunto sobre V. Suponga que*

$$f(t) = (t - \lambda_1)^{m_1} (t - \lambda_2)^{m_2} \cdots (t - \lambda_k)^{m_k},$$

donde $\lambda_1, \lambda_2, \ldots, \lambda_k \in \mathbb{R}$, es el polinomio característico de T. Entonces, existe una base ortonormal β de V tal que $[T]_\beta$ es diagonal.

Demostración. Por el teorema 5.27 existen polinomios $f_D(t)$ y $f_N(t)$ tales que $D = f_D(T)$ es diagonalizables, $N = f_N(T)$ es nilpotente y $T = D + N$. Por el lema anterior, N es auto-adjunta y $N = T_0$. En consecuencia, $T = D$ es diagonalizable.

Para cada $i = 1, 2, \dots, k$ sea β_i una base ortonormal de E_{λ_i}. Si $v_i \in E_{\lambda_i}$ y $v_j \in E_{\lambda_j}$ (con $i \neq j$), por la parte (c) del lema, v_i y v_j son ortogonales. Por ende,

$$\beta = \beta_1 \cup \beta_2 \cup \cdots \cup \beta_k$$

es una base ortonormal de V formada por vectores propios de T y $[T]_\beta$ es diagonal. □

Observación. Si T es auto-adjunta y

$$f(t) = (t - \lambda_1)^{m_1}(t - \lambda_2)^{m_2} \cdots (t - \lambda_k)^{m_k},$$

donde $\lambda_1, \lambda_2, \dots, \lambda_k \in \mathbb{R}$, es el polinomio característico de T, entonces

$$T = \lambda_1 \mathrm{proj}_{E_{\lambda_1}} + \lambda_2 \mathrm{proj}_{E_{\lambda_2}} + \cdots + \lambda_k \mathrm{proj}_{E_{\lambda_k}}.$$

En el próximo capítulo veremos que la hipótesis de que el polinomio característico se parta sobre $\mathbb{R}$ en el teorema espectral es superflua. Con el producto hermítico veremos que todo operador auto-adjunto tiene valores propios reales.

6.4. Operadores Ortogonales

Def. Sea V un espacio euclídeo y T un operador sobre V. Decimos que T es un **operador ortogonal** si para todo $v, w \in V$, $\langle T(v), T(w)\rangle = \langle v, w\rangle$.

Observación. Tenemos que

$$\begin{aligned} \|v + w\|^2 &= \langle v + w, v + w\rangle = \langle v, v\rangle + \langle w, v\rangle + \langle v, w\rangle + \langle w, w\rangle \\ &= \|v\|^2 + 2\langle v, w\rangle + \|w\|^2. \end{aligned}$$

De aquí obtenemos la siguiente expresión para el producto interno:

$$\langle v, w\rangle = \frac{\|v + w\|^2 - (\|v\|^2 + \|w\|^2)}{2}.$$

De esto podemos concluir que T es ortogonal si y solo si T preserva la norma, esto es $\|T(v)\| = \|v\|$ para todo $v \in V$.

Teorema 6.14. *Sea V es un espacio euclídeo de dimensión finita y T un operador sobre V. Entonces, las siguientes propiedades son equivalentes.*

(a) *T es ortogonal.*

(b) *T preserva la norma.*

(c) $T^*T = \mathrm{id}_V$.

(d) *La imagen mediante T de una base ortonormal es una base ortonormal.*

Demostración. En la observación anterior se demostró la equivalencia de (a) y (b). Ahora verifiquemos la equivalencia entre (a) y (c).

Sea $\{u_1, u_2, \dots, u_n\}$ una base ortonormal de V y supongamos que T es ortogonal, entonces

$$T^* \circ T(v) = \sum_{i=1}^{n} \langle (T^* \circ T)(v), u_i\rangle\, u_i = \sum_{i=1}^{n} \langle T(v), T(u_i)\rangle\, u_i = \sum_{i=1}^{n} \langle v, u_i\rangle\, u_i = v.$$

Esto prueba la condición suficiente.

Ahora, supongamos que $T^*T = \mathrm{id}_V$. Entonces,

$$\langle v, w\rangle = \langle v, T^*T(w)\rangle = \langle T(v), T(w)\rangle .$$

Esto termina de establecer la equivalencia.

Resta verificar una última equivalencia, a saber $(a) \Leftrightarrow (d)$. Sea $\{u_1, u_2, \ldots, u_n\}$ una base ortonormal de V y supongamos que T es ortogonal, entonces $\delta_{ij} = \langle u_i, u_j\rangle = \langle T(u_i), T(u_j)\rangle$. Esto prueba la ortogonalidad. Para ver que los vectores en la imagen de la base son normales, note que T preserva normas.

Recíprocamente, suponga que $\{u_1, u_2, \ldots, u_n\}$ y $\{T(u_1), T(u_2), \ldots, T(u_n)\}$ son bases ortonormales de V. Sean

$$v = \sum_{i=1}^{n} x_i u_i \qquad \text{y} \qquad w = \sum_{i=1}^{n} y_i u_i.$$

Entonces,

$$T(v) = \sum_{i=1}^{n} x_i T(u_i) \qquad \text{y} \qquad T(w) = \sum_{i=1}^{n} y_i T(u_i)$$

y $\langle T(v), T(w)\rangle = \sum\limits_{i=1}^{n} x_i y_i = \langle v, w\rangle$. □

Def. Sea $A \in M_{n\times n}(\mathbb{R})$. Decimos que A es una **matriz ortogonal** si $A^tA = I_n$.

Teorema 6.15. *Sea V un espacio euclídeo de dimensión finita y β una base ortonormal. Entonces, un operador T es ortogonal si y solo si $[T]_\beta$ es una matriz ortogonal.*

Demostración. Del teorema 6.14, tenemos que T es ortogonal si y solo $T^*T = \mathrm{id}_V$. En términos de la representación matricial, esto es

$$[T^*T]_\beta = [\mathrm{id}_V]_\beta.$$

Por el teorema 2.11, esto es

$$[T^*]_\beta [T]_\beta = I_n,$$

pero como β es ortonormal por el teorema 6.11, $[T^*]_\beta = ([T]_\beta)^t$ y así

$$([T]_\beta)^t\, [T]_\beta = I_n,$$

de modo que $[T]_\beta$ es ortogonal. □

6.5. Digresión: espacio dual

Def. Sea V un espacio vectorial sobre F. El **espacio dual** de V, denotado V^*, es el espacio vectorial de transformaciones lineales de V en F, $V^* = \mathcal{L}(V, F)$. A los elementos de V^* los llamamos **funcionales lineales**.

Teorema 6.16. *Sea V un espacio vectorial de dimensión finita. La dimensión de V es igual a la dimensión de V^*.*

Demostración. Es consecuencia del corolario 2.22.1. □

Def. Sea V un espacio vectorial de dimensión finita y $\beta = \{v_1, v_2, \ldots, v_n\}$ una base de V. La base dual de β, denotada β^*, es el conjunto $\{\lambda_1, \lambda_2, \ldots, \lambda_n\}$ donde

$$\lambda_i(v_j) = \delta_{ij} \qquad \text{para } 1 \le i, j \le n.$$

Teorema 6.17. *Sea V un espacio vectorial de dimensión finita* ($\dim(V) = n$) *y* $\beta = \{v_1, v_2, \dots, v_n\}$ *una base de V. Entonces, la base dual de β es una base de V^*.*

Demostración. Basta con verificar que β^* genera a V^*. Sea $\lambda \in V^*$ y definimos $a_i = \lambda(v_i)$. Sea

$$\mu = a_1\lambda_1 + a_2\lambda_2 + \cdots + a_n\lambda_n,$$

de modo que para cada $j = 1, 2, \dots, n$,

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

luego $\mu = \lambda$. Esto prueba que $V^* = \operatorname{span}(\{\lambda_1, \lambda_2, \dots, \lambda_n\})$. ☐

Def. Sea V un espacio vectorial sobre F. Para cada $v \in V$ definimos $\widehat{v} : V^* \to F$ por $\widehat{v}(\lambda) = \lambda(v)$ para cada $\lambda \in V^*$.

Lema 6.18. *Sea V un espacio vectorial y sea $v \in V$. Si $\widehat{v}(\lambda) = 0$ para todo $\lambda \in V^*$, entonces $v = 0$.*

Demostración. Sea $\{v_i\}_{i\in I}$ una base de V y $\{\lambda_i\}_{i\in I} \subseteq V^*$ la colección tal que $\lambda_i(v_j) = \delta_{ij}$ para todo $i, j \in I$. Como para todo $v \in V$, $v = \sum_{i\in I} \lambda_i(v)v_i$, si $\widehat{v}(\lambda) = 0$ para todo $\lambda \in V^*$, entonces en particular, $\lambda_i(v) = 0$ para todo $i \in I$ y así $v = 0$. ☐

Teorema 6.19. *Sea V un espacio vectorial. Existe una transformación lineal inyectiva*

$$\begin{aligned} \psi : V &\longrightarrow V^{**} \\ v &\longmapsto \widehat{v}. \end{aligned}$$

Si V es de dimensión finita, ψ es un isomorfismo.

Demostración. Por definición, ψ es lineal, pues para todo $v, w \in V$,

$$\psi cv + w(\lambda) = \lambda(cv + w) = c\lambda(v) + \lambda(w) = \widehat{cv}(\lambda) + \widehat{w}(\lambda) = (\psi cv + \psi w)(\lambda).$$

Por el lema anterior y el teorema 2.5, ψ es inyectiva. En el caso particular en que V es de dimensión finita, el teorema 6.16 implica que $\dim(V) = \dim(V^{**})$, pero por el lema 2.19, esto implica que ψ es un isomorfismo. ☐

Transformación dual

Def. Sean V y W espacios vectoriales sobre un campo F y sea $T \in \mathcal{L}(V, W)$. Definimos la transformación dual de T, denotada T^t, como la transformación lineal de W^* en V^* por

$$T^t(\lambda) = \lambda T$$

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

Observación. Sean V y W espacios vectoriales sobre un campo F y sea $T \in \mathcal{L}(V, W)$. La transformación dual T^t es lineal.

Demostración. $T^t(\lambda + \mu) = (\lambda + \mu)T = \lambda T + \mu T = T^t(\lambda) + T^t(\mu)$. ☐

Teorema 6.20. *Sean U, V y W espacios vectoriales sobre un campo F. Sean $T \in \mathcal{L}(V, W)$ y $S \in \mathcal{L}(W, U)$. Entonces,*

$$(ST)^t = T^t S^t.$$

Demostración. Para todo $\lambda \in U^*$ tenemos que

$$(ST)^t(\lambda) = \lambda(ST) = (\lambda S)T = T^t(\lambda S) = T^t S^t(\lambda).$$

$\square$

Teorema 6.21. *Sean V y W espacios vectoriales sobre F con bases ordenadas β y γ, respectivamente. Para toda transformación lineal $T : V \to W$, $[T^t]_{\gamma^*}^{\beta^*} = \left([T]_\beta^\gamma\right)^t$.*

Demostración. Sean $\beta = \{v_1, v_2, \dots, v_n\}$ y $\gamma = \{w_1, w_2, \dots, w_m\}$ con bases duales $\beta^* = \{\lambda_1, \lambda_2, \dots, \lambda_n\}$ y $\gamma^* = \{\mu_1, \mu_2, \dots, \mu_m\}$, respectivamente. Sea $A = [T]_\beta^\gamma$. Procedemos a expresar $T^t(\mu_j)$ como una combinación lineal de los elementos de β^*. Por el teorema 6.17,

$$T^t(\mu_j) = \mu_j T = \sum_{k=1}^n (\mu_j T)(v_k)\lambda_k.$$

Así, la entrada en la fila i, columna j de $[T^t]_{\gamma^*}^{\beta^*}$ es

$$(\mu_j T)(v_i) = \mu_j(T(v_i)) = \mu_j\left(\sum_{k=1}^n A_{ki} w_k\right) = \sum_{k=1}^n A_{ki}\mu_j(w_k) = \sum_{k=1}^n A_{ki}\delta_{jk} = A_{ji}.$$

Esto prueba que $[T^t]_{\gamma^*}^{\beta^*} = A^t$. $\square$

Hablamos del espacio dual en este momento por dos motivos. El primero es que en los espacios euclídeos, la base dual de una base ortonormal se puede expresar en términos del producto interno. Sea V un espacio euclídeo. Definimos $\imath_v : V \to F$ por

$$\imath_v(w) = \langle w, v\rangle .$$

Ahora, si V es de dimensión finita y $\beta = \{u_1, u_2, \dots, u_n\}$ es una base ortonormal de V, tenemos que

$$\beta^* = \{\imath_{u_1}, \imath_{u_2}, \dots, \imath_{u_n}\}.$$

El segundo motivo para traer a colación el espacio dual es para comparar el operador adjunto y la transformación dual. En espacios euclídeos, ambas cosas tienen la misma representación matricial en bases ortonormales. Sin embargo, cabe resaltar que son funciones muy distintas, empezando por su dominio y codominio.

El espacio dual es de particular importancia en el álgebra multilineal.

7. Espacios unitarios y producto hermítico

En esta sección exploramos espacios vectoriales con estructura adicional a la de un espacio euclídeo. Para ello, definimos una especie de producto interno, pero sobre espacios complejos.

Como vimos en la sección anterior, el producto interno configura un instrumento para hacer geometría en el espacio vectorial. Queremos extender esta posibilidad a espacios complejos. En particular, deseamos extender la noción de distancia establecida por la norma. Pero si generalizamos el producto interno al pie de la letra, nos encontramos con una inconsistencia.

Consideremos el vector $v \in V$, donde V es un espacio sobre $\mathbb{C}$. El cuadrado de la norma de iv está dado por $\langle iv, iv\rangle$, pero usando la bilinealidad del producto interno se obtiene que

$$\langle iv, iv\rangle = i^2 \langle v, v\rangle = -\langle v, v\rangle .$$

Esto quiere decir que alguno de los dos lados de la igualdad es negativo, lo cual contradice la propiedad 4 del producto interno. Es por esto que debemos modificar el producto interno para poder aplicarlo a espacio complejos.

Def. Sea V un espacio vectorial sobre $\mathbb{C}$. Un **producto hermítico** es una función

$$\begin{aligned} \langle\ ,\ \rangle : V \times V &\to \mathbb{C} \\ (u,v) &\mapsto \langle u,v\rangle \end{aligned}$$

tal que se satisfacen las siguientes condiciones:

1. Para todo $u,v,w \in V$, $\langle u+w,v\rangle = \langle u,v\rangle + \langle w,v\rangle$.
2. Para todo $u,v \in V$ y $c \in \mathbb{C}$, $\langle cu,v\rangle = c\,\langle u,v\rangle$.
3. Para todo $u,v \in V$, $\langle u,v\rangle = \overline{\langle v,u\rangle}$.
4. Para todo $v \in V$, $\langle v,v\rangle > 0$ si $v \neq \mathit{0}$.

Decimos que un espacio vectorial sobre $\mathbb{C}$ provisto de un producto hermítico es un **espacio unitario**.

Observación. La propiedad 3 es lo que le da su nombre al producto hermítico. La propiedad 3 implica que la segunda componente no es enteramente lineal, sino que es *sesquilineal*, esto es

$$\langle u, cv+w\rangle = \overline{c}\,\langle u,v\rangle + \langle u,w\rangle$$

para todo $u,v,w \in V$ y $c \in \mathbb{C}$.

La definición de la norma permanece sin alteraciones.

El ejemplo más común de un producto hermítico es el producto interno estándar de $\mathbb{C}^n$. Sean $u = (a_1,a_2,\ldots,a_n)$ y $v = (b_1,b_2,\ldots,b_n)$ vectores en $\mathbb{C}^n$. Entonces, el producto definido por

$$\langle u,v\rangle = \sum_{i=1}^{n} a_i\overline{b}_i$$

es un producto hermítico en $\mathbb{C}^n$.

Teorema 7.1. *Sea V un espacio unitario. Para todo $u,v \in V$ y $c \in \mathbb{C}$ las siguientes afirmaciones son verdaderas.*

(*a*) $\|cv\| = \|c\|\,\|v\|$,

(*b*) $\|v\| = 0$ *si y solo si* $v = \mathit{0}$,

(*c*) *(Desigualdad de Cauchy-Schwarz)* $|\langle u,v\rangle| \leq \|u\|\,\|v\|$,

(*d*) *(Desigualdad triangular)* $\|u+v\| \leq \|u\| + \|v\|$.

Demostración. (a) $\|cv\| = \sqrt{\langle cv,cv\rangle} = \sqrt{c\,\|c\|\,\langle v,v\rangle} = \|c\|\,\sqrt{\langle v,v\rangle} = \|c\|\,\|v\|$.

(b) La prueba es idéntica a su análogo para espacios euclídeos.

(c) Para $u = \mathit{0}$ o $v = \mathit{0}$ es evidente. Asumamos que $u,v \neq \mathit{0}$. Para todo $a,b \in \mathbb{C}$ tenemos que

$$\begin{aligned} 0 \leq \|au-bv\|^2 = \langle au-bv, au-bv\rangle &= a\overline{a}\,\langle u,u\rangle - a\overline{b}\,\langle u,v\rangle - \overline{a}b\,\langle v,u\rangle + b\overline{b}\,\langle v,v\rangle \\ &\leq \|a\|^2\,\|u\|^2 - (a\overline{b}\,\langle u,v\rangle + \overline{a}b\overline{\langle u,v\rangle}) + \|b\|^2\,\|v\|^2 . \end{aligned}$$

Si $a = \|u\|^2$ y $b = \langle u,v\rangle$, obtenemos que

$$0 \leq \|v\|^4\,\|u\|^2 - \|v\|^2\,\|\langle u,v\rangle\|^2 ,$$

lo cual establece el resultado.

(d) Tenemos que

$$\begin{aligned}\|u+v\|^2 &= \|u\|^2 + \langle u,v\rangle + \overline{\langle u,v\rangle} + \|v\|^2 \\ &= \|u\|^2 + 2\mathrm{Re}(\langle u,v\rangle) + \|v\|^2 \\ &\leq \|u\|^2 + 2\,\|\langle u,v\rangle\| + \|v\|^2 .\end{aligned}$$

Por la desigualdad de Cauchy-Schwartz esto último es menor o igual a $\|u\|^2 + 2\,\|u\|\,\|v\| + \|v\|^2$, i.e.

$$\begin{aligned}\|u+v\|^2 &\leq \|u\|^2 + 2\,\|u\|\,\|v\| + \|v\|^2 \\ &\leq (\|u\| + \|v\|)^2.\end{aligned}$$

□

A excepción del ángulo entre dos vectores, todas las definiciones dadas en la sección 6.1 valen también para espacios unitarios. Adicionalmente, todos los teoremas de dicha sección también se extienden a los espacios unitarios bajo las respectivas modificaciones a sus pruebas. Los detalles se dejan al lector.

7.1. Operador adjunto

Nuevamente, la definición anterior de operador adjunto se traslada directamente a los espacios unitarios. Similarmente, los teoremas 6.9 y 6.10 también valen para espacios unitarios y las pruebas son idénticas. No obstante, el teorema 6.11 sí cambia para espacio unitarios.

Teorema 7.2. *Sea V un espacio unitario de dimensión finita, T un operador sobre V y β una base ortonormal. Entonces,*

$$[T^*]_\beta = ([T]_\beta)^*,$$

donde * denota la matriz transpuesta conjugada.

Demostración. Sea $\beta = \{u_1, u_2, \ldots, u_n\}$. Si $A = [T]_\beta$ y $B = [T^*]_\beta$, tenemos que

$$A_{ij} = [T(u_j)]_{\beta,i} = \langle T(u_j), u_i\rangle \qquad \text{y} \qquad B_{ji} = [T^*(u_i)]_{\beta,j} = \langle T^*(u_i), u_j\rangle .$$

Por lo tanto,

$$B_{ji} = \langle T^*(u_i), u_j\rangle = \langle u_i, T(u_j)\rangle = \overline{\langle T(u_j), u_i\rangle} = \overline{A_{ij}}.$$

Esto prueba que $A^* = B$. □

Corolario 7.2.1. *Sea V un espacio unitario de dimensión finita y T un operador sobre V. Las siguientes dos condiciones son equivalentes.*

1. *T es auto-adjunto.*

2. *La representación matricial de T respecto a cualquier base ortonormal es hermítica, i.e. $A = A^*$.*

7.2. Teorema Espectral (Revisado)

En la sección anterior, demostramos una versión del teorema espectral con una hipótesis que resulta superflua una vez tratamos con espacios unitarios. Para verlo, primero modificamos el lema 6.12

Lema 7.3. *Sea V un espacio unitario de dimensión finita y T un operador lineal sobre V. Supongamos que T es auto-adjunto. Entonces,*

(a) *para todo $f(t) \in \mathbb{R}[t]$, $f(T)$ es auto-adjunto;*

(b) *si T es nilpotente, $T = T_0$;*

(c) *los valores propios de T son reales;*

(d) *si $v, w \in V$ son vectores propios asociados a valores propios distintos, entonces $\langle v, w\rangle = 0$.*

Demostración. Las pruebas de las partes (a), (b) y (d) son muy similares a las de la versión del lema en la sección anterior. Los detalles se dejan al lector.

Sea $T(v) = \lambda v$, donde v es un vector no nulo. Entonces,

$$\lambda \langle v, v\rangle = \langle \lambda v, v\rangle = \langle T(v), v\rangle = \langle v, T(v)\rangle = \langle v, \lambda v\rangle = \overline{\lambda} \langle v, v\rangle ,$$

luego $(\lambda - \overline{\lambda}) \|v\|^2 = 0$, pero $v \neq 0$, de modo que $\lambda = \overline{\lambda}$ o sea $\lambda \in \mathbb{R}$. Esto prueba (c). □

Llegados a este punto, enunciamos una nueva versión del teorema espectral.

Teorema 7.4 (Teorema Espectral)**.** *Sea V un espacio unitario de dimensión finita y T un operador auto-adjunto sobre V. Entonces, existe una base ortonormal β de V tal que $[T]_\beta$ es diagonal.*

Demostración. Basta con notar dos cosas: primero, que el campo de los números complejos es algebraicamente cerrado, de modo que el polinomio característico de T se puede escribir de la forma

$$f(t) = (t - \lambda_1)^{m_1}(t - \lambda_2)^{m_2} \cdots (t - \lambda_k)^{m_k},$$

donde $\lambda_1, \lambda_2, \dots, \lambda_k \in \mathbb{C}$, y segundo que por el lema anterior todos estos λ_i son de hecho reales. De aquí, nos remitimos a la prueba original en la sección anterior. □

7.3. Operadores Unitarios

Def. Sea V un espacio hermítico y T un operador sobre V. Decimos que T es un **operador unitario** si para todo $u, v \in V$,

$$\langle T(u), T(v)\rangle = \langle u, v\rangle .$$

Observación. Tenemos que

$$\begin{aligned}\|v + w\|^2 &= \langle v + w, v + w\rangle = \langle v, v\rangle + \langle w, v\rangle + \langle v, w\rangle + \langle w, w\rangle \\ &= \langle v, v\rangle + \overline{\langle v, w\rangle} + \langle v, w\rangle + \langle w, w\rangle \\ &= \|v\|^2 + 2\mathrm{Re}(\langle v, w\rangle) + \|w\|^2 ,\end{aligned}$$

de modo que

$$\mathrm{Re}(\langle v, w\rangle) = \frac{\|v + w\|^2 - (\|v\|^2 + \|w\|^2)}{2}.$$

Similarmente,

$$\begin{aligned}\|v+iw\|^2 &= \langle v+iw, v+iw\rangle = \langle v,v\rangle + i\langle w,v\rangle - i\langle v,w\rangle + \langle w,w\rangle \\ &= \langle v,v\rangle - i(\langle v,w\rangle - \overline{\langle v,w\rangle}) + \langle w,w\rangle \\ &= \|v\|^2 + 2\mathrm{Im}(\langle v,w\rangle) + \|w\|^2,\end{aligned}$$

de modo que

$$\mathrm{Im}(\langle v,w\rangle) = \frac{\|v+iw\|^2 - (\|v\|^2 + \|iw\|^2)}{2}.$$

De aquí obtenemos la siguiente expresión para el producto hermítico de dos vectores:

$$\langle v,w\rangle = \frac{\|v+w\|^2 - (\|v\|^2 + \|w\|^2)}{2} + i\frac{\|v+iw\|^2 - (\|v\|^2 + \|iw\|^2)}{2}.$$

De esto podemos concluir que T es unitario si y solo si preserva la norma. En este sentido, los operadores unitarios son análogos a los operadores ortogonales.

Asimismo, podemos ver que el producto hermítico es más que una generalización del producto interno debido a su parte imaginaria.

Teorema 7.5. *Sea T un operador lineal sobre un espacio unitario de dimensión finita V. Entonces, las siguientes afirmaciones son equivalentes.*

(*a*) *T es unitario.*

(*b*) *T preserva la norma.*

(*c*) *$TT^* = T^*T = \mathrm{id}_V$.*

(*d*) *La imagen mediante T de una base ortonormal es una base ortonormal.*

Demostración. La prueba es análoga a la del teorema 6.14. □

Def. Sea $A \in M_{n\times n}(\mathbb{C})$. Decimos que A es una **matriz unitaria** si $AA^* = I_n$.

Teorema 7.6. *Sea V un espacio euclídeo de dimensión finita y β una base ortonormal. Entonces, un operador T es unitario si y solo si $[T]_\beta$ es una matriz unitario.*

Demostración. Del teorema 7.5 tenemos que T es unitario si y solo si $T^*T = \mathrm{id}_V$. En términos de la representación matricial, esto es

$$[T^*T]_\beta = [\mathrm{id}_V]_\beta.$$

Por el teorema 2.11, esto es

$$[T^*]_\beta [T]_\beta = I_n,$$

pero como β es ortonormal por el teorema 7.2, $[T^*]_\beta = ([T]_\beta)^*$ y así

$$([T]_\beta)^*[T]_\beta = I_n,$$

de modo que $[T]_\beta$ es unitaria. □

7.4. Estructura Compleja

Def. Sea V un espacio vectorial sobre $\mathbb{R}$. Una **estructura compleja** J es un operador sobre V que satisface $J^2 = -\mathrm{id}_V$ y $\langle J(u), J(v)\rangle = \langle u, v\rangle$ para todo $u, v \in V$.

Teorema 7.7. *Sea V un espacio vectorial sobre $\mathbb{R}$ de dimensión finita. Si V admite una estructura compleja, entonces la dimensión de V es par.*

Demostración. Es fácil ver que $J^{-1} = -J$, de modo que $\det(J) \neq 0$. Así,

$$0 < \det(J)^2 = \det(J^2) = \det(-\mathrm{id}_V) = (-1)^{\dim(V)},$$

luego $2 | \dim(V)$. $\square$

La existencia de una estructura compleja permite convertir un espacio vectorial sobre $\mathbb{R}$ en uno sobre $\mathbb{C}$.

Teorema 7.8. *Sea V un espacio vectorial sobre $\mathbb{R}$ y J una estructura compleja en V. Entonces, V es un espacio vectorial sobre $\mathbb{C}$ con el siguiente producto escalar:*

$$(a + bi)v = (a\mathrm{id}_V + bJ)v.$$

Demostración. Los axiomas referentes a la suma (1-4) se siguen cumpliendo en la nueva versión de V. Resta verificar que se cumplan los axiomas 5-8.

5. $1v = \mathrm{id}_V(v) = v$.

6.
$$\begin{aligned}((a + bi)(c + di))v &= (ac - bd + (ad + bc)i)v \\ &= ((ac - bd)\mathrm{id}_V + (ad + bc)J)v \\ &= (ac\mathrm{id}_V + adJ + bcJ + bdJ^2)v \\ &= (a\mathrm{id}_V + bJ)(c\mathrm{id}_V + dJ)v \\ &= (a + bi)((c + di)v)\end{aligned}$$

7. Tenemos que $(a + bi)(v + w) = (a\mathrm{id}_V + bJ)(v + w)$, pero como $a\mathrm{id}_V + bJ$ es lineal,
$$(a + bi)(v + w) = (a\mathrm{id}_V + bJ)v + (a\mathrm{id}_V + bJ)w = (a + bi)v + (a + bi)w.$$

8.
$$\begin{aligned}((a + bi) + (c + di))v &= (a + c + (b + d)i)v \\ &= ((a + c)\mathrm{id}_V + (b + d)J)v \\ &= (a\mathrm{id}_V + bJ + c\mathrm{id}_V + dJ)v \\ &= (a\mathrm{id}_V + bJ)v + (c\mathrm{id}_V + dJ)v \\ &= (a + bi)v + (c + di)v.\end{aligned}$$

$\square$

Teorema 7.9. *Sea V un espacio vectorial sobre $\mathbb{R}$ de dimensión finita y sea J una estructura compleja en V. Suponga que $\dim(V) = 2n$ para algún entero positivo n. Entonces, existe una base V (en $\mathbb{R}$) de la forma $\beta = \{v_1, v_2, \ldots, v_n, J(v_1), \ldots, J(v_n)\}$. Más aún,*

$$[J]_\beta = \begin{bmatrix} 0 & -I_n \\ I_n & 0 \end{bmatrix}.$$

Demostración. Sea $\{v_1, v_2, \ldots, v_n\}$ una base de V en $\mathbb{C}$. Podemos ver que $V = \operatorname{span}_{\mathbb{R}}\{v_1, v_2, \ldots, v_n, J(v_1), \ldots, J(v_n)\}$, puesto que

$$\sum_{i=1}^{n} z_i v_i = \sum_{i=1}^{n} a_i v_i + b_i J(v_i),$$

donde $z_i = a_i + b_i i$ y $a_i, b_i \in \mathbb{R}$. Resta verificar que β es linealmente independiente. Suponga que $a_i, b_i \in \mathbb{R}$ son tales que

$$\sum_{i=1}^{n} a_i v_i + b_i J(v_i) = \mathit{0}.$$

Entonces, $\sum_{i=1}^{n} z_i v_i = \mathit{0}$, con $z_i = a_i + b_i i$. Así, todo $z_i = 0$ y en consecuencia $a_i = b_i = 0$ para todo i.

La prueba de la segunda parte del teorema es trivial. □

Hasta este momento hemos visto como complejizar un espacio vectorial sobre los reales. No obstante, el álgebra no solo estudia las estructuras por sí solas sino que también se interesa en ciertos tipos de funciones entre estructuras. En particular, en el álgebra lineal nos interesan las transformaciones lineales, de modo que también nos compete complejizar transformaciones.

Para ello, es necesario notar primero que para complejizar una transformación lineal entre espacios vectoriales reales es necesario que ambos espacios vectoriales admitan estructuras complejas. Adicionalmente, si $T : V \to W$ es una transformación lineal y J_1 y J_2 son estructuras complejas sobre V y W, respectivamente, para complejizar T se debe cumplir que $J_2 T = T J_1$. El siguiente teorema considera esta condición para operadores lineales.

Teorema 7.10. *Sea V un espacio vectorial sobre $\mathbb{R}$, J una estructura compleja en V y T un operador sobre V. Considere V como un espacio vectorial sobre $\mathbb{C}$ con $(a + bi)v = av + bJ(v)$. Entonces, $T \in \mathcal{L}(V)$ si y solo si $TJ = JT$.*

Demostración. Si $T \in \mathcal{L}(V)$, entonces para todo $v \in V$,

$$T(J(v)) = T(iv) = iT(v) = J(T(v)).$$

Recíprocamente, si $TJ = JT$, entonces para todo $a, b \in \mathbb{R}$ y para todo $v \in V$,

$$\begin{aligned} T((a + bi)v) &= T(a\mathrm{id}_V + bJ)(v) \\ &= (aT + bTJ)(v) \\ &= (aT + bJT)(v) \\ &= (a\mathrm{id}_V + bJ)T(v) \\ &= (a + bi)T(v), \end{aligned}$$

de modo que $T \in \mathcal{L}(V)$. □

En lo que va de esta discusión no hemos traído a colación el producto interno ni el producto hermítico, pero sin duda alguna deberíamos poder definir un producto hermítico en la complejización de un espacio euclídeo.

Teorema 7.11. *Sea V un espacio euclídeo y J una estructura compleja en V. Considere V como espacio vectorial sobre sobre $\mathbb{C}$ con $(a + bi)v = av + J(v)$. Entonces, $\langle v, w \rangle_J = \langle v, w \rangle + \langle v, J(w) \rangle i$ define un producto hermítico si y solo J es ortogonal.*

Demostración. Suponga que J es ortogonal (i.e. $J^* = J^{-1} = -J$), entonces para todo $v, w \in V$,

$$\begin{aligned}
\langle w, v\rangle_J &= \langle w, v\rangle + \langle w, J(v)\rangle\, i \\
&= \langle w, v\rangle - \langle J(w), v\rangle\, i \\
&= \langle v, w\rangle - \langle v, J(w)\rangle\, i \\
&= \overline{\langle v, w\rangle_J}.
\end{aligned}$$

Esto prueba que $\langle \bullet, \bullet\rangle_J$ es hermítico. La linealidad en la primera componente es evidente. Ahora, para todo $v \in V$,

$$\langle v, J(v)\rangle = -\langle J(v), v\rangle = -\langle v, J(v)\rangle\,,$$

de modo que $\langle v, J(v)\rangle = 0$. Si además $v \neq 0$,

$$\langle v, v\rangle_J = \langle v, v\rangle + \langle v, J(v)\rangle\, i = \langle v, v\rangle > 0.$$

Esto verifica que $\langle \bullet, \bullet\rangle_J$ es definitivamente positivo y por ende es un producto hermítico.

Recíprocamente, suponga que $\langle \bullet, \bullet\rangle_J$ es un producto hermítico, entonces para todo $v, w \in V$

$$\begin{aligned}
\langle J(v), J(w)\rangle &= \mathrm{Re}(\langle J(v), J(w)\rangle_J) \\
&= \frac{\langle J(v), J(w)\rangle_J + \langle J(v), J(w)\rangle_J}{2} \\
&= \frac{\langle iv, iw\rangle_J + \langle iv, iw\rangle_J}{2} \\
&= \frac{\langle v, w\rangle_J + \langle v, w\rangle_J}{2} \\
&= \mathrm{Re}(\langle v, w\rangle_J) \\
&= \langle v, w\rangle\,.
\end{aligned}$$

Esto prueba que J es ortogonal. □

Finalmente, deseamos trasladar operadores ortogonales a operadores unitarios al complejizar un espacio euclídeo. Para ello elaboramos el siguiente teorema.

Teorema 7.12. *Sea V un espacio euclídeo y J una estructura compleja en V. Considere V como un espacio vectorial sobre $\mathbb{C}$ con $(a + bi)v = av + bJ(v)$ y como espacio unitario con*

$$\langle v, w\rangle_J = \langle v, w\rangle + \langle v, J(w)\rangle\, i \qquad \text{para todo } v, w \in V.$$

Sea T un operador sobre V. T es unitario si y solo si T es ortogonal y $TJ = JT$.

Demostración. Supongamos que T es unitario, entonces T es un operador sobre V (visto como espacio complejo) y así $TJ = JT$. Además para todo $v \in V$,

$$\|T(v)\|^2 = \langle T(v), T(v)\rangle = \langle T(v), T(v)\rangle_J = \langle v, v\rangle_J = \langle v, v\rangle = \|v\|^2\,,$$

luego T es ortogonal.

Por el contrario, suponga $TJ = JT$ y T es ortogonal. Entonces, para todo $v, w \in V$,

$$\begin{aligned}
\langle T(v), T(w)\rangle_J &= \langle T(v), T(w)\rangle + \langle T(v), J(T(w))\rangle\, i \\
&= angT(v), T(w) + \langle T(v), T(J(w))\rangle\, i \\
&= \langle v, w\rangle + \langle v, J(w)\rangle\, i \\
&= \langle v, w\rangle_J\,,
\end{aligned}$$

de modo que T es unitario. □

8. Álgebra Multilineal

El álgebra multilineal es básicamente álgebra lineal con varios espacios vectoriales al mismo tiempo. El objeto fundamental del álgebra multilineal es una generalización del vector llamada tensor. Existen varias formas de definir los tensores. En estas notas, definimos un tensor como una función multilineal.

8.1. Tensores

Def. Sean V y W espacios vectoriales sobre un campo F. Decimos que

$$T : \underbrace{V \times V \times \cdots \times V}_{k \text{ veces}} \to W$$

es una función multilineal si para todo $i = 1, 2, \ldots, k$,

$$T(v_1, v_2, \ldots, cv_i + w, \ldots, v_k) = cT(v_1, v_2, \ldots, v_i, \ldots, v_k) + T(v_1, v_2, \ldots, w, \ldots, v_k),$$

donde $v_1, v_2, \ldots, v_k, w \in V$ y $c \in F$; esto es, si T es lineal en cada uno de sus componentes.

De ahora en adelante, denotaremos el producto cartesiano $V \times V \times \cdots \times V$ (k veces) como V^k.

Def. Sea V un espacio vectorial sobre un campo F. Un **k-tensor** es una función multilineal

$$T : V^k \to F.$$

El conjunto de k-tensores se denota $\mathcal{T}_k(V)$.

Observación. $\mathcal{T}_0(V) \simeq F$ y $\mathcal{T}_1(V) = V^*$. Adicionalmente, si $F = \mathbb{R}$ y es posible definir un producto interno, este es un 2-tensor.

Observación. Si T y S son k-tensores y $a, b \in F$, $aT + bS$ también es un k-tensor. En otras palabras, $\mathcal{T}_k(V)$ es un espacio vectorial.

Def. Sea T un k-tensor y S un l-tensor. Definimos el **producto tensorial** de T y S por

$$T \otimes S(v_1, v_2, \ldots, v_k, v_{k+1}, \ldots, v_{k+1}) = T(v_1, v_2, \ldots, v_k)S(v_{k+1}, \ldots, v_{k+l}).$$

Observación. $T \otimes S$ es un $(k+l)$-tensor. Adicionalmente, el producto tensorial no es conmutativo, pues $S \otimes T$ no coincide en general con $T \otimes S$. Sin embargo, el producto tensorial sí es asociativo; esto es, si S, T y R son tensores (de cualquier orden) $(S \otimes T) \otimes R = S \otimes (T \otimes R)$.

Teorema 8.1. *Sean V un espacio vectorial sobre un campo F, $a \in F$, $S, T \in \mathcal{T}_k(V)$ y $R \in \mathcal{T}_l(V)$. Entonces,*

(a) $T \otimes (aS) = (aT) \otimes S = a(T \otimes S)$,

(b) $(T + S) \otimes R = T \otimes R + S \otimes R$,

(c) $R \otimes (T + S) = R \otimes T + R \otimes S$.

Demostración. La prueba de (a) es trivial. Tomemos $v_1, v_2, \ldots, v_{k+l} \in V$. Entonces,

$$\begin{aligned}
(T + S) \otimes R(v_1, v_2, \ldots, v_{k+l}) &= (T + S)(v_1, \ldots, v_k) \cdot R(v_{k+1}, \ldots, v_{k+l}) \\
&= [T(v_1, \ldots, v_k) + S(v_1, \ldots, v_k)]R(v_{k+1}, \ldots, v_{k+l}) \\
&= T(v_1, \ldots, v_k)R(v_{k+1}, \ldots, v_{k+l}) + S(v_1, \ldots, v_k)R(v_{k+1}, \ldots, v_{k+l}) \\
&= T \otimes R(v_1, v_2, \ldots, v_{k+1}) + S \otimes R(v_1, v_2, \ldots, v_{k+1}).
\end{aligned}$$

Esto verifica (b). La demostración de la parte (c) es similar. □

A partir del producto tensorial podemos construir $\mathcal{T}_k(V)$ usando una base de $V^* = \mathcal{T}_1(V)$.

Teorema 8.2. *Sea V un espacio vectorial de dimensión finita y $\beta = \{v_1, v_2, \ldots, v_n\}$ una base de V. Considere la base dual $\beta^* = \{\lambda_1, \lambda_2, \ldots, \lambda_n\}$. Entonces, la colección de k-tensores*

$$\{\lambda_{i_1} \otimes \lambda_{i_2} \otimes \cdots \otimes \lambda_{i_k}\}_{i_1, i_2, \ldots, i_k = 1}^n$$

es una base de $\mathcal{T}_k(V)$. En particular, $\dim(\mathcal{T}_k(V)) = n^k$.

Demostración. Primero, veamos que $\mathcal{T}_k(V) = \operatorname{span}(\{\lambda_{i_1} \otimes \lambda_{i_2} \otimes \cdots \otimes \lambda_{i_k}\}_{i_1, i_2, \ldots, i_k = 1}^n)$. Note que para todo $i_1, \ldots, i_k, j_1, \ldots, j_k \in \{1, 2, \ldots, n\}$,

$$\lambda_{i_1} \otimes \lambda_{i_2} \otimes \cdots \otimes \lambda_{i_k}(v_{j_1}, v_{j_2}, \ldots, v_{j_k}) = \delta_{i_1 j_1} \delta_{i_2 j_2} \cdots \delta_{i_k j_k} = \begin{cases} 1 & \text{si } i_1 = j_1, \ldots, i_k = j_k, \\ 0 & \text{de lo contrario.} \end{cases}$$

Así, si $w_1, w_2, \ldots, w_k \in V$ y $w_j = \sum_{i=1}^n a_{ij} v_i$, donde $a_{ij} \in F$ y $j = 1, 2, \ldots, k$. Entonces,

$$\begin{aligned} \lambda_{i_1} \otimes \lambda_{i_2} \otimes \cdots \otimes \lambda_{i_k}(w_1, w_2, \ldots, w_k) &= \sum_{j_1, \ldots, j_k = 1}^n a_{j_1 1} \cdots a_{j_k k} \lambda_{i_1} \otimes \cdots \otimes \lambda_{i_k}(v_{j_1}, \ldots, v_{j_k}) \\ &= \sum_{j_1, \ldots, j_k = 1}^n a_{j_1 1} \cdots a_{j_k k} \delta_{i_1 j_1} \cdots \delta_{i_k j_k} \\ &= a_{i_1 1} a_{i_2 2} \cdots a_{i_k k}. \end{aligned}$$

Ahora, dado $T \in \mathcal{T}_k(V)$,

$$\begin{aligned} T(w_1, w_2, \ldots, w_k) &= \sum_{i_1, \ldots, i_k = 1}^n a_{i_1 1} a_{i_2 2} \cdots a_{i_k k} T(v_{i_1}, v_{i_2}, \ldots, v_{i_k}) \\ &= \sum_{i_1, \ldots, i_k = 1}^n T(v_{i_1}, v_{i_2}, \ldots, v_{i_k}) \lambda_{i_1} \otimes \lambda_{i_2} \otimes \cdots \otimes \lambda_{i_k}(w_1, w_2, \ldots, w_k), \end{aligned}$$

luego

$$T = \sum_{i_1, \ldots, i_k = 1}^n T(v_{i_1}, v_{i_2}, \ldots, v_{i_k}) \lambda_{i_1} \otimes \lambda_{i_2} \otimes \cdots \otimes \lambda_{i_k}.$$

Esto verifica que $\mathcal{T}_k(V) = \operatorname{span}(\{\lambda_{i_1} \otimes \lambda_{i_2} \otimes \cdots \otimes \lambda_{i_k}\}_{i_1, i_2, \ldots, i_k = 1}^n)$.

En cuanto a la independencia lineal, tomemos

$$0 = \sum_{i_1, \ldots, i_k = 1}^n c_{i_1 i_2 \ldots i_k} \lambda_{i_1} \otimes \lambda_{i_2} \otimes \cdots \otimes \lambda_{i_k}$$

y evaluamos en $(v_{j_1}, v_{j_2}, \ldots, v_{j_k})$:

$$\begin{aligned} 0 &= \sum_{i_1, \ldots, i_k = 1}^n c_{i_1 \ldots i_k} \lambda_{i_1} \otimes \cdots \otimes \lambda_{i_k}(v_{j_1}, \ldots, v_{j_k}) \\ &= c_{j_1 j_2 \ldots j_k}. \end{aligned}$$

Esto prueba la independencia lineal. □

8.2. Tensores generales

Resulta que la definición anterior de un tensor es demasiado restrictiva para ciertos propósitos. Es por ello que consideramos una clase más general de tensores. Para este fin, empleamos el teorema 6.19 y asumimos que V es dimensión finita.

Def. Sea V un espacio vectorial sobre un campo F. Un **tensor** de tipo (l,k) o (l,k)-tensor es una función multilineal

$$T : \underbrace{V^* \times V^* \times \cdots \times V^*}_{l \text{ veces}} \times \underbrace{V \times V \times \cdots \times V}_{k \text{ veces}} \to F.$$

El conjunto de (l,k)-tensores sobre V se denota por $\mathcal{T}_k^l(V)$.

Observación. Note que $\mathcal{T}_k^0(V) = \mathcal{T}_k(V)$. Adicionalmente, $\mathcal{T}_0^1(V) = V^{**} \simeq V$. Por lo tanto, una versión del producto tensorial nos permite construir $\mathcal{T}_k^l(V)$ a partir de $\widehat{v} = \psi(v)$ (véase teorema 6.19).

Def. Sean $T \in \mathcal{T}_{k_1}^{l_1}(V)$ y $S \in \mathcal{T}_{k_2}^{l_2}(V)$ tensores. Definimos el **producto tensorial** $T \otimes S \in \mathcal{T}_{k_1+k_2}^{l_1+l_2}(V)$ por

$$\begin{aligned} T \otimes S(\lambda_1, \ldots, \lambda_{l_1+l_2}, v_1, \ldots, v_{k_1+k_2}) \\ = T(\lambda_1, \ldots, \lambda_{l_1}, v_1, \ldots, v_{k_1}) S(\lambda_{l_1+1}, \ldots, \lambda_{l_1+l_2}, v_{l_1+1}, \ldots, v_{k_1+k_2}), \end{aligned}$$

para todo $\lambda_1, \ldots, \lambda_{l_1+l_2} \in V^*$ y $v_1, \ldots, v_{k_1+k_2} \in V$.

Teorema 8.3. *Sea V un espacio vectorial de dimensión finita y $\beta = \{v_1, v_2, \ldots, v_n\}$ una base de V. Considere la base dual $\beta^* = \{\lambda_1, \lambda_2, \ldots, \lambda_n\}$. Entonces, la colección de (l,k)-tensores*

$$\{\widehat{v}_{i_1} \otimes \cdots \otimes \widehat{v}_{i_l} \otimes \lambda_{j_1} \otimes \cdots \otimes \lambda_{j_k}\}_{i_1,\ldots,i_l,j_1,\ldots,j_k=1}^n$$

es una base de $\mathcal{T}_k^l(V)$. En particular, $\dim(\mathcal{T}_k^l(V)) = n^{l+k}$.

Demostración. La prueba es similar a la del teorema 8.2. □

Aprovechando el isomorfismo del teorema 6.19, no distinguimos el vector $v \in V$ de la función $\widehat{v} \in V^{**}$. Así, la base descrita en el teorema anterior se suele denotar por

$$\{v_{i_1} \otimes \cdots \otimes v_{i_l} \otimes \lambda_{j_1} \otimes \cdots \otimes \lambda_{j_k}\}_{i_1,\ldots,i_l,j_1,\ldots,j_k=1}^n.$$

8.3. Notación de tensores

De acuerdo al teorema anterior, un tensor de tipo (l,k) sobre un espacio vectorial de dimensión finita n está completamente determinado por sus valores en la base. Esto es, n^{l+k} valores determinan completamente un tensor dada una base. Para referirnos a cada una de las componentes de un tensor necesitamos $l+k$ índices, lo cual que se puede tornar muy tedioso. Sin embargo, introducimos una notación menos tediosa y que permite distinguir entre los índices asociados a vectores o a funcionales lineales (también llamados covectores).

Para introducir esta notación, considere dos bases de un espacio vectorial $\{e_i\}$ y $\{e'_i\}$. Recordemos que los vectores de la base primada se pueden expresar en términos de los vectores de la base original a través de la siguiente expresión

$$e'_j = \sum_i A_{ij} e_i,$$

donde A_{ij} son las entradas de la matriz de cambio de base. Ahora, considere un vector v y su expresión en términos de la base original, $v = \sum_i v_i e_i$, y en términos de la nueva base, $v = \sum_i v'_i e'_i$. Entonces,

$$\sum_j v'_j e'_j = \sum_{ij} v'_j A_{ij} e_i = \sum_i \left(\sum_j A_{ij} v'_j \right) e_i = \sum_i v_i e_i,$$

de modo que

$$v_i = \sum_j A_{ij} v'_j \qquad \text{ó} \qquad v'_i = \sum_j (A^{-1})_{ij} v_j.$$

Note que los vectores de la base y las componentes transforman distinto bajo un cambio de base.

Llegados a este punto es natural preguntarse qué ocurre con los funcionales lineales de la base dual. Primero, busquemos una matriz B tal que

$$\lambda'_j = \sum_i B_{ij} \lambda_i,$$

donde $\lambda_j(e_i) = \delta_{ij}$ y $\lambda'_j(e'_i) = \delta_{ij}$ para todo i, j (i.e. $\{\lambda_i\}$ es la base dual de $\{e_i\}$ y $\{\lambda'_i\}$ es la base dual de $\{e'_i\}$). Para este fin, considere lo siguiente:

$$\begin{aligned} \delta_{ij} = \lambda'_j(e'_i) = \lambda'_j \left(\sum_k A_{ki} e_k \right) &= \sum_l B_{lj} \lambda_l \left(\sum_k A_{ki} e_k \right) = \sum_{kl} A_{ki} B_{lj} \lambda_l(e_k) \\ &= \sum_{kl} A_{ki} B_{lj} \delta_{kl} = \sum_k A_{ki} B_{kj}. \end{aligned}$$

Esto quiere decir que $A^t B = I_n$, así que $B = (A^t)^{-1} = (A^{-1})^t$. Si $f \in V^*$ es cualquier funcional lineal con $f = \sum_i f_i \lambda_i$ y $f = \sum_i f'_i \lambda'_i$, tenemos que

$$\sum_j f'_j \lambda'_j = \sum_{ij} f'_j B_{ij} \lambda_i = \sum_i \left(\sum_j B_{ij} f'_j \right) \lambda_i = \sum_i f_i \lambda_i,$$

lo cual nos permite concluir que

$$f_i = \sum_j B_{ij} f'_j \qquad \text{ó} \qquad f'_i = \sum_j (B^{-1})_{ij} f_j.$$

No obstante, es más natural escribir estas expresiones en términos de la matriz de cambio de coordenadas A:

$$\lambda'_i = \sum_j B_{ji} \lambda_j = \sum_j (B^t)_{ij} \lambda_j = \sum_j (A^{-1})_{ij} \lambda_j,$$

$$f'_j = \sum_i (B^{-1})_{ji} f_i = \sum_i (A^t)_{ji} f_i = \sum_i A_{ij} f_i.$$

Note que tantos los vectores de la base como los componentes de los funcionales lineales transforman a través de A, mientras que los funcionales de la base dual y los componentes de los vectores transforman a través de A^{-1}. Esto da lugar a dos tipos de índices: **covariante** (con los vectores de la base) y **contravariante** (en contra de los vectores de la base). Los índices covariantes se denotan como subíndices y los contravariantes como superíndices. Así, los vectores de la base se denotan como e_i y los funcionales de la base dual como λ^i. Esta nueva notación hace que sea evidente como transforma el elemento solo por inspección.

A modo de ejemplo considere la siguiente expresión de un vector $v \in V$ en términos de la base $\{e_i\}$:

$$v = \sum_i v^i e_i.$$

El superíndice en los componentes v^i indica que estos transforman de forma contravariante. Del mismo modo, en la expresión para algún funcional lineal $f \in V^*$,

$$f = \sum_i f_i \lambda^i,$$

el subíndice en los componentes f_i indica su covarianza.

Aplicando esto a tensores, un tensor de tipo (l,k) se expresa en términos de una base de $\mathcal{T}_k^l(V)$ por

$$T = \sum_{\substack{i_1 \dots i_l \\ j_1 \dots j_k}} T^{i_1 i_2 \dots i_l}{}_{j_1 j_2 \dots j_k} e_{i_1} \otimes e_{i_2} \otimes \cdots \otimes e_{i_l} \otimes \lambda^{j_1} \otimes \lambda^{j_2} \otimes \cdots \otimes \lambda^{j_k}.$$

En adición a esto, adoptamos una convención suma llamada convención de Einstein, en la que si un índice se encuentra repetido como superíndice y subíndice, se asume una sumatoria sobre este índice. Adoptando esta convención un vector se expresa como combinación lineal de los vectores en una base así:

$$v = v^i e_i.$$

Similarmente, en la anterior expresión del tensor T podemos omitir la sumatoria:

$$T = T^{i_1 i_2 \dots i_l}{}_{j_1 j_2 \dots j_k} e_{i_1} \otimes e_{i_2} \otimes \cdots \otimes e_{i_l} \otimes \lambda^{j_1} \otimes \lambda^{j_2} \otimes \cdots \otimes \lambda^{j_k}.$$

8.4. Tensores alternantes

Def. Un tensor ω de tipo $(0,k)$ se dice alternante si dada una permutación $\sigma \in S_n$,

$$\omega(v_{\sigma(1)}, v_{\sigma(2)}, \dots, v_{\sigma(k)}) = \operatorname{sign}(\sigma)\omega(v_1, v_2, \dots, v_k).$$

Al conjunto de $(0,k)$-tensores alternantes lo denotamos $\Lambda^k(V)$.

Observación. El conjunto de tensores alternantes $\Lambda^k(V)$ es un subespacio vectorial de $\mathcal{T}_k(V)$.

En vista de que el producto tensorial de dos tensores alternantes no necesariamente es alternante. Por lo tanto, buscamos obtener un producto que sí respete la alternancia. Para ello, primero debemos ser capaces de construir una función que construya tensores alternantes a partir de cualquier k-tensor.

Def. Definimos la función $\operatorname{Alt} : \mathcal{T}_k(V) \to \mathcal{T}_k(V)$ por

$$\operatorname{Alt}(T)(v_1, v_2, \dots, v_k) = \frac{1}{k!} \sum_{\sigma \in S_k} \operatorname{sign}(\sigma) T(v_{\sigma(1)}, v_{\sigma(2)}, \dots, v_{\sigma(k)}).$$

Teorema 8.4. *El operador* Alt *es una proyección sobre* $\Lambda^k(V)$*:*

(*a*) *para todo* $T \in \mathcal{T}_k(V)$, $\operatorname{Alt}(T) \in \Lambda^k(V)$,

(*b*) *para todo* $\omega \in \Lambda^k(V)$, $\operatorname{Alt}(\omega) = \omega$,

(*c*) *para todo* $T \in \mathcal{T}_k(V)$, $\operatorname{Alt}(\operatorname{Alt}(T)) = \operatorname{Alt}(T)$.

Demostración.

(a) Tome $\sigma \in S_n$. Tenemos entonces que

$$\begin{aligned}
\mathrm{Alt}(T)(v_{\sigma(1)}, v_{\sigma(2)}, \ldots, v_{\sigma(k)}) &= \frac{1}{k!} \sum_{\rho \in S_k} \mathrm{sign}(\rho) T(v_{\rho(\sigma(1))}, v_{\rho(\sigma(2))}, \ldots, v_{\rho(\sigma(k))}) \\
&= \frac{1}{k!} \mathrm{sign}(\sigma) \sum_{\rho \in S_k} \mathrm{sign}(\rho)\mathrm{sign}(\sigma) T(v_{\rho(\sigma(1))}, v_{\rho(\sigma(2))}, \ldots, v_{\rho(\sigma(k))}) \\
&= \frac{1}{k!} \mathrm{sign}(\sigma) \sum_{\rho \in S_k} \mathrm{sign}(\rho\sigma) T(v_{\rho(\sigma(1))}, v_{\rho(\sigma(2))}, \ldots, v_{\rho(\sigma(k))})
\end{aligned}$$

Introducimos un cambio de variable $\tau = \rho\sigma$ y así,

$$\begin{aligned}
\mathrm{Alt}(T)(v_{\sigma(1)}, v_{\sigma(2)}, \ldots, v_{\sigma(k)}) &= \frac{1}{k!} \mathrm{sign}(\sigma) \sum_{\tau \in S_k} \mathrm{sign}(\tau) T(v_{\tau(1)}, v_{\tau(2)}, \ldots, v_{\tau(k)}) \\
&= \mathrm{sign}(\sigma) \left(\frac{1}{k!} \sum_{\tau \in S_k} \mathrm{sign}(\tau) T(v_{\tau(1)}, v_{\tau(2)}, \ldots, v_{\tau(k)}) \right) \\
&= \mathrm{sign}(\sigma) \mathrm{Alt}(T)(v_1, v_2, \ldots, v_k).
\end{aligned}$$

Esto prueba que $\mathrm{Alt}(T) \in \Lambda^k(V)$.

(b) Si $\omega \in \Lambda^k(V)$ y $\sigma \in S_n$, tenemos que $\omega(v_{\sigma(1)}, v_{\sigma(2)}, \ldots, v_{\sigma(k)}) = \mathrm{sign}(\sigma)\omega(v_1, v_2, \ldots, v_k)$, por lo tanto,

$$\begin{aligned}
\mathrm{Alt}(\omega)(v_1, v_2, \ldots, v_k) &= \frac{1}{k!} \sum_{\sigma \in S_k} \mathrm{sign}(\sigma) \omega(v_{\sigma(1)}, v_{\sigma(2)}, \ldots, v_{\sigma(k)}) \\
&= \frac{1}{k!} \sum_{\sigma \in S_k} \mathrm{sign}(\sigma)\mathrm{sign}(\sigma) \omega(v_1, v_2, \ldots, v_k) \\
&= \omega(v_1, v_2, \ldots, v_k).
\end{aligned}$$

(c) Se sigue de (a) y (b).

$\square$

Def. Sea $\omega \in \Lambda^k(V)$ y $\eta \in \Lambda^l(V)$. Definimos el **producto exterior** $\omega \wedge \eta \in \Lambda^{k+l}(V)$ por

$$\omega \wedge \eta = \frac{(k+l)!}{k!\, l!} \mathrm{Alt}(\omega \otimes \eta).$$

Lema 8.5. $\mathrm{Alt} : \mathcal{T}^k(V) \to \mathcal{T}^k(V)$ *es una transformación lineal.*

Demostración. Sean $T, S \in \mathcal{T}^k(V)$ y $c \in F$. Entonces,

$$\begin{aligned}
\mathrm{Alt}(cT+S)(v_1, v_2, \ldots, v_k) &= \frac{1}{k!}\sum_{\sigma\in S_k} \mathrm{sign}(\sigma)(cT+S)(v_{\sigma(1)}, v_{\sigma(2)}, \ldots, v_{\sigma(k)}) \\
&= \frac{1}{k!}\sum_{\sigma\in S_k} \mathrm{sign}(\sigma)[cT(v_{\sigma(1)}, v_{\sigma(2)}, \ldots, v_{\sigma(k)}) + S(v_{\sigma(1)}, v_{\sigma(2)}, \ldots, v_{\sigma(k)})] \\
&= c\frac{1}{k!}\sum_{\sigma\in S_k} \mathrm{sign}(\sigma)T(v_{\sigma(1)}, v_{\sigma(2)}, \ldots, v_{\sigma(k)}) \\
&\qquad + \frac{1}{k!}\sum_{\sigma\in S_k} \mathrm{sign}(\sigma)S(v_{\sigma(1)}, v_{\sigma(2)}, \ldots, v_{\sigma(k)}) \\
&= c\mathrm{Alt}(T) + \mathrm{Alt}(S).
\end{aligned}$$

□

Teorema 8.6. *Sean $\omega, \eta \in \Lambda^k(V)$ y $\theta \in \Lambda^l(V)$ tensores alternantes. Entonces,*

(*a*) $\omega \wedge (a\theta) = (a\omega) \wedge \theta = a(\omega \wedge \theta)$,

(*b*) $(\omega + \eta) \wedge \theta = \omega \wedge \theta + \eta \wedge \theta$,

(*c*) $\theta \wedge (\omega + \theta) = \theta \wedge \omega + \theta \wedge \eta$,

(*d*) $\omega \wedge \theta = (-1)^{kl}\theta \wedge \omega$.

Demostración. La prueba de (*a*) es trivial. Para (*b*),

$$\begin{aligned}
(\omega+\eta)\wedge\theta &= \frac{(k+l)!}{k!\,l!}\mathrm{Alt}((\omega+\eta)\otimes\theta) \\
&= \frac{(k+l)!}{k!\,l!}\mathrm{Alt}(\omega\otimes\theta + \eta\otimes\theta).
\end{aligned}$$

Como Alt es lineal, obtenemos que

$$(\omega+\eta)\wedge\theta = \frac{(k+l)!}{k!\,l!}\mathrm{Alt}(\omega\otimes\theta) + \frac{(k+l)!}{k!\,l!}\mathrm{Alt}(\eta\otimes\theta) = \omega\wedge\theta + \eta\wedge\theta.$$

La prueba de (*c*) es similar.

Finalmente, sea $\tau \in S_{k+l}$ tal que

$$\tau(i) = \begin{cases} i+k & \text{si } 1 \le i \le l, \\ i-l & \text{si } l+1 \le i \le l+k. \end{cases}$$

Es fácil ver que $\mathrm{sign}(\tau) = (-1)^{kl}$. Entonces, para todo $v_1, v_2, \ldots, v_{k+l}$

$$\begin{aligned}
\omega\wedge\theta(v_1, v_2, \ldots, v_{k+l}) &= \frac{(k+l)!}{k!\,l!}\mathrm{Alt}(\omega\otimes\theta)(v_1, v_2, \ldots, v_{k+l}) \\
&= \frac{1}{k!\,l!}\sum_{\sigma\in S_{k+l}} \mathrm{sign}(\sigma)\omega\otimes\theta(v_{\sigma(1)}, v_{\sigma(2)}, \ldots, v_{\sigma(k+l)}) \\
&= \frac{1}{k!\,l!}\sum_{\sigma\in S_{k+l}} \mathrm{sign}(\sigma)\omega(v_{\sigma(1)}, \ldots, v_{\sigma(k)})\theta(v_{\sigma(k+1)}, \ldots, v_{\sigma(k+l)}) \\
&= \frac{1}{k!\,l!}\sum_{\sigma\in S_{k+l}} \mathrm{sign}(\sigma)\omega(v_{\sigma\tau(l+1)}, \ldots, v_{\sigma\tau(k+l)})\theta(v_{\sigma\tau(1)}, \ldots, v_{\sigma\tau(l)}) \\
&= \frac{1}{k!\,l!}\sum_{\sigma\in S_{k+l}} (-1)^{kl}\mathrm{sign}(\sigma\tau)\theta(v_{\sigma\tau(1)}, \ldots, v_{\sigma\tau(l)})\omega(v_{\sigma\tau(l+1)}, \ldots, v_{\sigma\tau(k+l)}).
\end{aligned}$$

Introducimos el cambio de variable $\rho = \sigma\tau$ para obtener

$$\begin{aligned}\omega \wedge \theta(v_1, v_2, \ldots, v_{k+l}) &= (-1)^{kl} \frac{1}{k!\, l!} \sum_{\rho \in S_{k+l}} \operatorname{sign}(\rho)\theta(v_{\rho(1)}, \ldots, v_{\rho(l)})\omega(v_{\rho(l+1)}, \ldots, v_{\rho(k+l)}) \\ &= (-1)^{kl} \frac{1}{k!\, l!} \sum_{\rho \in S_{k+l}} \operatorname{sign}(\rho)\theta \otimes \omega(v_{\rho(1)}, v_{\rho(2)}, \ldots, v_{\rho(k+l)}) \\ &= (-1)^{kl} \frac{(k+l)!}{k!\, l!} \operatorname{Alt}(\theta \otimes \omega)(v_1, v_2, \ldots, v_{k+l}) \\ &= (-1)^{kl} \theta \wedge \omega(v_1, v_2, \ldots, v_{k+l}).\end{aligned}$$

Esto prueba (d). □

Lema 8.7. *Sean $R \in \mathcal{T}_m(V)$, $S \in \mathcal{T}_k(V)$ y $T \in \mathcal{T}_l(V)$. Entonces,*

(a) *si* $\operatorname{Alt}(S) = 0$, $\operatorname{Alt}(T \otimes S) = \operatorname{Alt}(S \otimes T) = 0$,

(b) $\operatorname{Alt}(\operatorname{Alt}(R \otimes S) \otimes T) = \operatorname{Alt}(R \otimes S \otimes T) = \operatorname{Alt}(R \otimes \operatorname{Alt}(S \otimes T))$.

Demostración.

(a) **Esta demostración emplea conceptos de álgebra abstracta.** Considere S_k como el subgrupo de S_{k+l} conformado por las permutaciones que fijan $k+1, k+2, \ldots, k+l$ y sean $\sigma_1, \sigma_2, \ldots, \sigma_N \in S_{k+l}$, $N = (k+l)!/(k!)$, representantes de los coconjuntos $\sigma_i S_k$, de forma que

$$S_{k+l} = \bigcup_{i=1}^{N} \sigma_i S_k,$$

es decir $\{\sigma_i S_k\}_{i=1}^N$ es una partición de S_{k+l}. Entonces, para todo $v_1, v_2, \ldots, v_{k+l} \in V$

$$\begin{aligned}&(k+l)!\operatorname{Alt}(S \otimes T)(v_1, v_2, \ldots, v_{k+l}) \\ &= \sum_{\sigma \in S_{k+l}} \operatorname{sign}(\sigma) S \otimes T(v_1, v_2, \ldots, v_{k+l}) \\ &= \sum_{i=1}^{N} \sum_{\tau \in S_k} \operatorname{sign}(\sigma_i\tau) S \otimes T(v_{\sigma_i\tau(1)}, v_{\sigma_i\tau(2)}, \ldots, v_{\sigma_i\tau(k+l)}) \\ &= \sum_{i=1}^{N} \operatorname{sign}(\sigma_i) \sum_{\tau \in S_k} \operatorname{sign}(\tau) S(v_{\sigma_i\tau(1)}, v_{\sigma_i\tau(2)}, \ldots, v_{\sigma_i\tau(k)}) T(v_{\sigma_i\tau(k+1)}, v_{\sigma_i\tau(k+2)}, \ldots, v_{\sigma_i\tau(k+l)}).\end{aligned}$$

Denotemos $w_j^{(i)} = v_{\sigma_i(j)}$ para $1 \le i \le N$ y $1 \le j \le k+l$. Así,

$$\begin{aligned}&(k+l)!\operatorname{Alt}(S \otimes T)(v_1, v_2, \ldots, v_{k+l}) \\ &= \sum_{i=1}^{N} \operatorname{sign}(\sigma_i) \sum_{\tau \in S_k} \operatorname{sign}(\tau) S(w_{\tau(1)}^{(i)}, w_{\tau(2)}^{(i)}, \ldots, w_{\tau(k)}^{(i)}) T(w_{\tau(k+1)}^{(i)}, w_{\tau(k+2)}^{(i)}, \ldots, w_{\tau(k+l)}^{(i)}) \\ &= \sum_{i=1}^{N} \operatorname{sign}(\sigma_i) \sum_{\tau \in S_k} \operatorname{sign}(\tau) S(w_{\tau(1)}^{(i)}, w_{\tau(2)}^{(i)}, \ldots, w_{\tau(k)}^{(i)}) T(w_{k+1}^{(i)}, w_{k+2}^{(i)}, \ldots, w_{k+l}^{(i)}) \\ &= \sum_{i=1}^{N} \operatorname{sign}(\sigma_i) T(w_{k+1}^{(i)}, w_{k+2}^{(i)}, \ldots, w_{k+l}^{(i)}) \operatorname{Alt}(S)(w_1^{(i)}, w_2^{(i)}, \ldots, w_k^{(i)}) \\ &= 0.\end{aligned}$$

Similarmente, si tomamos S_l como el subgrupo de S_{k+l} formado por las permutaciones que fijan $1, 2, \ldots, l$, por un método similar obtenemos que $\mathrm{Alt}(T \otimes S) = 0$.

(b) Primero, notemos que

$$\mathrm{Alt}(\mathrm{Alt}(R \otimes S) - R \otimes S) = \mathrm{Alt}(R \otimes S) - \mathrm{Alt}(R \otimes S) = 0.$$

Por (a), tenemos que

$$\begin{aligned} 0 &= \mathrm{Alt}((\mathrm{Alt}(R \otimes S) - R \otimes S) \otimes T) \\ &= \mathrm{Alt}(\mathrm{Alt}(R \otimes S) \otimes T) - \mathrm{Alt}(R \otimes S \otimes T). \end{aligned}$$

Esto prueba que $\mathrm{Alt}(\mathrm{Alt}(R \otimes S) \otimes T) = \mathrm{Alt}(R \otimes S \otimes T)$. Un procedimiento similar prueba que $\mathrm{Alt}(R \otimes \mathrm{Alt}(S \otimes T)) = \mathrm{Alt}(R \otimes S \otimes T)$.

□

Teorema 8.8. *Sea $\omega \in \Lambda^k(V)$, $\eta \in \Lambda^l(V)$ y $\theta \in \Lambda^m(V)$. Entonces,*

$$(\omega \wedge \eta) \wedge \theta = \omega \wedge (\eta \wedge \theta) = \frac{(k+l+m)!}{k!\,l!\,m!}\mathrm{Alt}(\omega \otimes \eta \otimes \theta).$$

Demostración. La prueba es directa usando el lema 8.7.

$$\begin{aligned} (\omega \wedge \eta) \wedge \theta &= \frac{(k+l+m)!}{(k+l)!m!}\mathrm{Alt}((\omega \wedge \eta) \otimes \theta) \\ &= \frac{(k+l+m)!}{(k+l)!m!}\frac{(k+l)!}{k!\,l!}\mathrm{Alt}(\omega \otimes \eta \otimes \theta). \end{aligned}$$

La otra igualdad se prueba de forma similar. □

A partir de los teoremas 8.6 y 8.8, podemos construir una base de $\Lambda^k(V)$.

Teorema 8.9. *Sea V un espacio vectorial de dimensión finita y $\beta = \{e_1, e_2, \ldots, e_n\}$ una base de V. Tomemos $\beta^* = \{\lambda^1, \lambda^2, \ldots, \lambda^n\}$, la base dual de β. La colección*

$$\{\lambda^{i_1} \wedge \lambda^{i_2} \wedge \cdots \wedge \lambda^{i_k}\}_{1 \leq i_1 < i_2 < \ldots < i_k \leq n}$$

es una base de $\Lambda^k(V)$. En particular, $\dim(\Lambda^k(V)) = \binom{n}{k}$.

Demostración. Como $\Lambda^k(V) \subset \mathcal{T}_k(V)$, sabemos que para todo $\omega \in \Lambda^k(V)$ existen $c_{j_1 j_2 \ldots j_k} \in F$ tales que

$$\omega = \sum_{j_1,\ldots,j_k=1}^{n} c_{j_1 j_2 \ldots j_k} \lambda^{j_1} \otimes \lambda^{j_2} \otimes \cdots \otimes \lambda^{j_k},$$

pero $\omega = \mathrm{Alt}(\omega)$, así que

$$\begin{aligned} \omega &= \mathrm{Alt}\left(\sum_{j_1,\ldots,j_k=1}^{n} c_{j_1 j_2 \ldots j_k} \lambda^{j_1} \otimes \lambda^{j_2} \otimes \cdots \otimes \lambda^{j_k}\right) \\ &= \sum_{j_1,\ldots,j_k=1}^{n} c_{j_1 j_2 \ldots j_k} \mathrm{Alt}(\lambda^{j_1} \otimes \lambda^{j_2} \otimes \cdots \otimes \lambda^{j_k}) \\ &= \sum_{j_1,\ldots,j_k=1}^{n} c_{j_1 j_2 \ldots j_k} k! \lambda^{j_1} \wedge \lambda^{j_2} \wedge \cdots \wedge \lambda^{j_k} \end{aligned}$$

Por el teorema 8.6 si dos índices son iguales el producto externo se anula, de modo que podemos asumir que todos los índices son distintos. En tal caso,

$$\lambda^{j_1} \wedge \lambda^{j_2} \wedge \cdots \wedge \lambda^{j_k} = \operatorname{sign}(\sigma_{j_1 j_2 \ldots j_k}) \lambda^{i_1} \wedge \lambda^{i_2} \wedge \cdots \wedge \lambda^{i_k},$$

donde $i_1 < i_2 < \ldots < i_k$ son los mismos índices $j_1, j_2, \ldots, j_k$, pero ordenados de manera ascendente y $\sigma_{j_1 j_2 \ldots j_k} \in S_k$ es la permutación que ordena estos índices. Luego

$$\omega = \sum_{1 \leq i_1 < \ldots < i_k \leq n} a_{i_1 i_2 \ldots i_k} \lambda^{i_1} \wedge \lambda^{i_2} \wedge \cdots \wedge \lambda^{i_k},$$

donde $a_{i_1 i_2 \ldots i_k} = k! \sum_{\sigma \in S_k} c_{i_{\sigma(1)} i_{\sigma(2)} \cdots i_{\sigma(k)}}$. Esto comprueba que $\{\lambda_{i_1} \wedge \lambda_{i_2} \wedge \cdots \wedge \lambda_{i_k}\}_{1 \leq i_1 < i_2 < \ldots < i_k \leq n}$ genera a $\Lambda^k(V)$.

En cuanto a la independencia lineal, notemos que si $1 \leq i_1 < i_2 < \ldots < i_k \leq n$ y $1 \leq j_1 < j_2 < \ldots < j_k \leq n$ entonces

$$\lambda^{i_1} \wedge \lambda^{i_2} \wedge \cdots \wedge \lambda^{i_k}(v_{j_1}, v_{j_2}, \ldots, v_{j_k}) = \begin{cases} 1 & \text{si } i_r = j_r \text{ para todo } r \in \{1, 2, \ldots, k\}, \\ 0 & \text{de lo contrario.} \end{cases}$$

Entonces, si $0 = \sum_{1 \leq i_1 < \ldots < i_k \leq n} a_{i_1 i_2 \ldots i_k} \lambda^{i_1} \wedge \lambda^{i_2} \wedge \cdots \wedge \lambda^{i_k}$, entonces

$$\begin{aligned} 0 &= \sum_{1 \leq i_1 < \ldots < i_k \leq n} a_{i_1 i_2 \ldots i_k} \lambda^{i_1} \wedge \lambda^{i_2} \wedge \cdots \wedge \lambda^{i_k}(v_{j_1}, v_{j_2}, \ldots, v_{j_k}) \\ &= a_{j_1 j_2 \ldots j_k}. \end{aligned}$$

□

Esta teoría de los tensores alternantes es particularmente relevante para la geometría diferencial, en particular para integrar sobre variedades, lo cual se encuentra fuera del alcance de estas notas. No obstante, hay un par de aplicaciones que vale la pena mencionar.

Determinante

Una aplicación inmediata de lo que es los tensores alternantes es una definición del determinante de una matriz. Anteriormente vimos que el determinante puede ser visto como una función de las columnas de la matriz. En este momento, estamos en la capacidad de ver que el determinante es un tensor. En particular, es un tensor alternante. Podemos ver que si $\{e_i\}_{i=1}^n$ es la base canónica de F^n y $\{\lambda^i\}_{i=1}^n$ es su dual,

$$\det = \lambda^1 \wedge \lambda^2 \wedge \cdots \wedge \lambda^n.$$

Producto cruz

Considere el tensor $\varepsilon^i_{jk} \in \mathcal{T}^1_2(\mathbb{R}^3)$ con $i, j, k \in \{1, 2, 3\}$ definido por

$$\epsilon^i_{jk} = \begin{cases} \operatorname{sign}(\sigma) & \text{si } \sigma : 1 \mapsto i,\ 2 \mapsto j,\ 3 \mapsto k \text{ es una permutación,} \\ 0 & \text{de lo contrario.} \end{cases}$$

Note que $\varepsilon^i_{jk} = 0$ siempre que $i = j$, $j = k$ ó $k = i$. Cabe resaltar además que el símbolo ε^i_{jk} denota las componentes del tensor en la base canónica más que el tensor como tal. El tensor *per se* sería

$$\varepsilon = \varepsilon^i_{jk} e_i \otimes \lambda^j \otimes \lambda^k.$$

Aquí empleamos la convención de Einstein.

A los tensores les podemos aplicar lo que se conoce como una **contracción** o evaluación incompleta. Esto se trata de evaluar el tensor en algunas de sus componentes.

Def. Sea $\{e_i\}_{i=1}^n$ una base de V y $\{\lambda^i\}_{i=1}^n$ su dual. Sean

$$t = e_{\alpha_1} \otimes e_{\alpha_2} \otimes \cdots \otimes e_{\alpha_r} \otimes \lambda^{\beta_1} \otimes \lambda^{\beta_2} \otimes \cdots \otimes \lambda^{\beta_s} \in \mathcal{T}_s^r(V)$$

y

$$T = e_{i_1} \otimes e_{i_2} \otimes \cdots \otimes e_{i_l} \otimes \lambda^{j_1} \otimes \lambda^{j_2} \otimes \cdots \otimes \lambda^{j_k} \in \mathcal{T}_k^l(V),$$

con $l \geq r$ y $k \geq s$ Definimos $T(t) \in T_{(k-s)}^{(l-r)}$, la **contracción** de T por t como

$$T(t) = \left(\prod_{m=1}^{s} \lambda^{\beta_m}(e_{i_m}) \prod_{m=1}^{r} \lambda^{j_m}(e_{\alpha_m}) \right) e_{i_{s+1}} \otimes e_{i_{s+2}} \otimes \cdots \otimes e_{i_l} \otimes \lambda^{j_{s+1}} \otimes \lambda^{j_{s+1}} \otimes \cdots \otimes \lambda^{j_k}.$$

A modo de ejemplo, tomemos

$$v = \sum_{i=1}^{3} v^i e_i \qquad \text{y} \qquad w = \sum_{i=1}^{3} w^i e_i,$$

de modo que $v \otimes w \in \mathcal{T}_0^2(\mathbb{R}^3)$ y

$$\begin{aligned}
&\varepsilon(v \otimes w) \\
&= e_1(\lambda^2(v)\lambda^3(w) - \lambda^3(v)\lambda^2(w)) - e_2(\lambda^1(v)\lambda^3(w) - \lambda^3(v)\lambda^1(w)) + e_3(\lambda^1(v)\lambda^2(w) - \lambda^2(v)\lambda^1(w)) \\
&= (v^2w^3 - v^3w^2)e_1 - (v^1w^3 - v^3w^1)e_2 + (v^1w^2 - v^2w^1)e_3.
\end{aligned}$$

Esta contracción define el **producto cruz** entre v y w. Otra posible contracción es mediante el tensor $\theta = \sum_{i=1}^3 \theta_i \lambda^i \in \mathcal{T}_0^1(V)$:

$$\begin{aligned}
\sum_{ijkm} \varepsilon^i_{jk} \theta_m \lambda^m(e_i) \lambda^j \otimes \lambda^k &= \sum_{ijk} \varepsilon^i_{jk} \lambda^j \otimes \lambda^k \sum_m \theta_m \lambda^m(e_i) \\
&= \sum_{ijk} \varepsilon^i_{jk} \theta_i \lambda^j \otimes \lambda^k \\
&= \theta_1(\lambda^2 \otimes \lambda^3 - \lambda^3 \otimes \lambda^2) - \theta_2(\lambda^1 \otimes \lambda^3 - \lambda^3 \otimes \lambda^1) + \theta_3(\lambda^1 \otimes \lambda^2 - \lambda^2 \otimes \lambda^1) \\
&= \theta_1(\lambda^2 \wedge \lambda^3) - \theta_2(\lambda^1 \wedge \lambda^3) + \theta_3(\lambda^1 \wedge \lambda^3).
\end{aligned}$$

Esta contracción elucida una estructura de tensor alternante, lo cual está asociado a la anti-simetría del producto cruz.