Notas para el curso de Álgebra Lineal II

Centro de Matemática
Facultad de Ciencias
Universidad de la República

Andrés Abella

3 de diciembre de 2020

Introducción

Estas son notas para el curso de Álgebra Lineal II de la Licenciatura en Matemática de la Facultad de Ciencias. Tratan sobre operadores, productos internos y formas bilineales simétricas en espacios vectoriales, con énfasis en los casos de dimensión finita. Como prerrequisito, se asume que el lector tiene conocimientos básicos de espacios vectoriales, transformaciones lineales y matrices asociadas a transformaciones lineales. Si bien la mayoría de los resultados están escritos para $\Bbbk$-espacios vectoriales en los cuales $\Bbbk$ es un cuerpo arbitrario, para que estas notas sean más amigables todos los ejemplos son para $\Bbbk = \mathbb{R}$ o $\mathbb{C}$. Lo que sigue es una breve descripción de cada capítulo.

En el Capítulo 1 se estudian los operadores diagonalizables y sus propiedades básicas.

En el Capítulo 2 se estudian los espacios vectoriales con producto interno. En particular se prueba la existencia de bases ortonormales, de complementos ortogonales de subespacios, y el teorema de Riesz.

En el Capítulo 3 se estudian los operadores en espacios con producto interno. Las demostraciones están escritas para espacios complejos; las correspondientes a los espacios reales se obtienen como un caso particular de las anteriores. Si el lector está interesado solo en los espacios reales, entonces puede olvidarse de las referencias a los complejos, y las definiciones y demostraciones funcionan esencialmente igual. El único punto donde tendría problemas es en la prueba de que un operador real autoadjunto es diagonalizable sobre una base ortonormal, dado que la prueba en el caso complejo utiliza el "teorema fundamental del álgebra" que no es válido en los reales. Para solucionar esto se dan dos pruebas de ese teorema, una de las cuales no requiere números complejos.

En el Capítulo 4 se estudian las formas bilineales simétricas y las formas cuadráticas, con énfasis en el caso real.

En el Capítulo 5 se estudian las relaciones polinomiales que puede verificar un operador. En particular se introduce el polinomio minimal y se ve el teorema de Cayley-Hamilton.

En el Capítulo 6 se estudia la forma de Jordan. En la primera parte se estudian las propiedades de las bases de Jordan. En la segunda parte se ven técnicas de cálculo para obtener la forma de Jordan en dimensiones bajas. En la tercera parte se prueba el teorema de Jordan y se ve la forma de obtener bases de Jordan en el caso general.

El Capítulo 7 es un apéndice en el cual se ven algunos temas necesarios para seguir estas notas y otros que las profundizan o extienden. Lo que sigue es una breve descripción de cada una de sus secciones.

- Sección 7.1. Algunas propiedades de polinomios que no se suelen estudiar en los cursos de enseñanza secundaria, en particular se ven los polinomios de Lagrange y la identidad de Bézout
- Sección 7.2. Las matrices elementales y su relación con la invertibilidad de matrices.
- Sección 7.3. Las matrices en bloques y sus propiedades básicas.
- Sección 7.4. Las operaciones básicas con subespacios: intersección, suma y suma directa.
- Sección 7.5. Las proyecciones y su relación con las sumas directas de subespacios.
- Sección 7.6. La descomposición espectral de un operador. Primero en forma general para un operador diagonalizable y luego –en el caso de espacios con producto interno– para un operador digonalizable en una base ortonormal.

- Sección 7.7. Se aplican las técnicas de producto interno para el estudio de los movimientos del plano y del espacio.
- Sección 7.8. Se aplica la teoría de formas bilineales para el estudio de las cónicas en el plano y de las cuádricas en el espacio.
- Sección 7.9. Un algoritmo para hallar la forma de Jordan y la prueba de la unicidad de la forma de Jordan.
- Sección 7.10. A partir de la forma de Jordan compleja se obtiene la forma de Jordan real.
- Sección 7.11. Formas multilineales alternadas. Esto se necesita para entender las formas diferenciales, que se estudian en cursos posteriores.

De este apéndice, para poder leer las notas hay que conocer las secciones 7.3, 7.4 y 7.5. La sección 7.2 se usa en el estudio de las formas bilineales simétricas; si no se conoce, entonces se deben admitir ciertos resultados. De la sección 7.1, el teorema de Bézout se utiliza en el Capítulo 5, mientras que los polinomios elementales aparecen en el estudio de la descomposición espectral. La sección 7.6 da otra forma de ver a los operadores diagonalizables y es básica para la extensión de las propiedades que vimos de operadores en espacios con producto interno a espacios de dimensión infinita. Las secciones 7.7 y 7.8 son aplicaciones clásicas del álgebra lineal al estudio de problemas geométricos. Las secciones 7.9 y 7.10 profundizan el estudio de la forma de Jordan y cierran algunas preguntas que quedaban abiertas en el Capítulo 6. La ultima sección está pensada principalmente para su aplicación en cursos posteriores.

Parte de este material está inspirado en el libro *Linear Algebra*, de S. M. Friedberg, A. J. Insel y L. E. Spence, de la editorial Prentice Hall.

Contenidos

Capítulo 1

Diagonalización

En este tema $\Bbbk$ es un cuerpo y V es un $\Bbbk$-espacio vectorial no nulo de dimensión finita n. Si $\mathcal{B}$ es una base de V, supondremos siempre que $\mathcal{B}$ es una *base ordenada* es decir una base en la cual hemos elegido un orden para sus elementos.

A las transformaciones lineales de V en V les llamaremos también *operadores*. Al espacio de las transformaciones lineales de V en V lo denotaremos $\mathcal{L}(V)$ y a la transformación lineal identidad por Id o Id_V. Al espacio de las matrices cuadradas de orden n con coeficientes en $\Bbbk$ lo denotaremos $M_n(\Bbbk)$ y a la matriz identidad I o I_n.

1.1. Operadores diagonalizables

Si $T \in \mathcal{L}(V)$ y $\mathcal{B}$ y $\mathcal{C}$ son bases de V, escribiremos ${}_{\mathcal{C}}[T]_{\mathcal{B}}$ a la matriz asociada a T de la base $\mathcal{B}$ en la base $\mathcal{C}$ y $[T]_{\mathcal{B}}$ para ${}_{\mathcal{B}}[T]_{\mathcal{B}}$. Si $A \in M_n(\Bbbk)$, escribiremos L_A a la transformación lineal de $\Bbbk^n$ en $\Bbbk^n$ definida por $L_A(v) = Av$, $\forall v \in \Bbbk^n$. Observar que si $\mathcal{C} = \{e_1, \dots, e_n\}$ es la base canónica de $\Bbbk^n$ y $A \in M_n(\Bbbk)$, es

$$[L_A]_{\mathcal{C}} = A.$$

Si $v_1, \dots, v_n \in \Bbbk^n$, el símbolo $[v_1|\cdots|v_n]$ denotará a la matriz cuadrada de orden n cuyas columnas son los vectores $v_1, \dots, v_n$ escritos en forma vertical. Por ejemplo, si $v_1 = (1, 2)$ y $v_2 = (0, 3)$, entonces $[v_1|v_2] = \left(\begin{smallmatrix} 1 & 0 \\ 2 & 3 \end{smallmatrix}\right)$.

Definición 1.1.1. Diremos que dos matrices $A, B \in M_n(\Bbbk)$ son *semejantes* y escribimos $A \simeq B$, si existe una matriz invertible $Q \in M_n(\Bbbk)$ tal que $A = QBQ^{-1}$.

Queda como ejercicio el verificar que la relación de semejanza es de equivalencia en $M_n(\Bbbk)$.

Proposición 1.1.2. *Sean $T \in \mathcal{L}(V)$ y $\mathcal{B}$ una base de V.*

1. *Si $\mathcal{B}'$ es otra base de V, entonces $[T]_{\mathcal{B}'} \simeq [T]_{\mathcal{B}}$. Explícitamente, $[T]_{\mathcal{B}'} = P\,[T]_{\mathcal{B}}\,P^{-1}$, siendo $P = {}_{\mathcal{B}'}[\mathrm{Id}]_{\mathcal{B}}$.*
2. *Si $A \in M_n(\Bbbk)$ es tal que $A \simeq [T]_{\mathcal{B}}$, entonces existe una base $\mathcal{B}'$ de V tal que $A = [T]_{\mathcal{B}'}$.*

Dem.

1. Si $P = {}_{\mathcal{B}'}[\mathrm{Id}]_{\mathcal{B}}$, entonces $[T]_{\mathcal{B}'} = {}_{\mathcal{B}'}[\mathrm{Id}]_{\mathcal{B}}\,[T]_{\mathcal{B}}\,{}_{\mathcal{B}}[\mathrm{Id}]_{\mathcal{B}'} = P\,[T]_{\mathcal{B}'}\,P^{-1}$.
2. Supongamos $A = Q\,[T]_{\mathcal{B}}\,Q^{-1}$, siendo $Q^{-1} = (q_{ij})$ y $\mathcal{B} = \{v_1, \dots, v_n\}$. Definimos $w_j := \sum_{i=1}^n q_{ij}\,v_i$, para todo $j = 1 \dots, n$. Como la matriz Q^{-1} es invertible, entonces el conjunto $\mathcal{B}'' = \{w_1, \dots, w_n\}$ es una base de V y ${}_{\mathcal{B}}[\mathrm{Id}]_{\mathcal{B}''} = Q^{-1}$. Luego

$$A = Q\,[T]_{\mathcal{B}}\,Q^{-1} = {}_{\mathcal{B}''}[\mathrm{Id}]_{\mathcal{B}} \cdot [T]_{\mathcal{B}} \cdot {}_{\mathcal{B}}[\mathrm{Id}]_{\mathcal{B}''} = [T]_{\mathcal{B}''}. \qquad \square$$

Corolario 1.1.3. *Dos matrices son semejantes si y solo si representan a una misma transformación lineal.*

Dem. Sean A y B en $M_n(\Bbbk)$ tales que $A \simeq B$. Si consideramos $T = L_A \in \mathcal{L}(\Bbbk^n)$, es $A = [T]_\mathcal{C}$ siendo $\mathcal{C}$ la base canónica de $\Bbbk^n$. Luego $B \simeq [T]_\mathcal{C}$ y la parte 2 de la Proposición 1.1.2 implica que existe $\mathcal{B}$ base de $\Bbbk^n$ tal que $B = [T]_\mathcal{B}$. El recíproco es la parte 1 de la Proposición 1.1.2. □

Recordar que una matriz cuadrada $A = (a_{ij})$ se dice *diagonal* si $a_{ij} = 0$, para todo $i \neq j$. Las matrices diagonales son las más simples para operar, por ejemplo valen las fórmulas siguientes.

$$\begin{pmatrix} a_1 & \cdots & 0 \\ & \ddots & \\ 0 & \cdots & a_n \end{pmatrix} \begin{pmatrix} b_1 & \cdots & 0 \\ & \ddots & \\ 0 & \cdots & b_n \end{pmatrix} = \begin{pmatrix} a_1 b_1 & \cdots & 0 \\ & \ddots & \\ 0 & \cdots & a_n b_n \end{pmatrix}; \quad \begin{pmatrix} a_1 & \cdots & 0 \\ & \ddots & \\ 0 & \cdots & a_n \end{pmatrix}^k = \begin{pmatrix} (a_1)^k & \cdots & 0 \\ & \ddots & \\ 0 & \cdots & (a_n)^k \end{pmatrix}, \ \forall k;$$

$$\begin{vmatrix} a_1 & \cdots & 0 \\ & \ddots & \\ 0 & \cdots & a_n \end{vmatrix} = a_1 \cdots a_n; \quad \text{si } a_1 \cdots a_n \neq 0 \ \Rightarrow \ \begin{pmatrix} a_1 & \cdots & 0 \\ & \ddots & \\ 0 & \cdots & a_n \end{pmatrix}^{-1} = \begin{pmatrix} (a_1)^{-1} & \cdots & 0 \\ & \ddots & \\ 0 & \cdots & (a_n)^{-1} \end{pmatrix}.$$

Definición 1.1.4. Un operador $T \in \mathcal{L}(V)$ es *diagonalizable* si existe una base $\mathcal{B}$ de V en la cual $[T]_\mathcal{B}$ es diagonal.

Ejemplo 1.1.5. Sea r una recta del plano $\mathbb{R}^2$ que pasa por el origen y T la simetría axial respecto a r. Sean $0 \neq u \in r$ y $0 \neq v \in \mathbb{R}^2$ tales que u y v son ortogonales. Luego $\mathcal{B} = \{u, v\}$ es una base de $\mathbb{R}^2$ y la matriz asociada a T en la base $\mathcal{B}$ es $\left(\begin{smallmatrix} 1 & 0 \\ 0 & -1 \end{smallmatrix}\right)$. Luego la simetría axial de eje r es diagonalizable.

Definición 1.1.6. Una matriz $A \in M_n(\Bbbk)$ es *diagonalizable* si es semejante a una matriz diagonal.

Proposición 1.1.7. *Sea $T \in \mathcal{L}(V)$.*

1. *Si T es diagonalizable y $\mathcal{B}$ es una base cualquiera de V, entonces $[T]_\mathcal{B}$ es diagonalizable.*
2. *Si existe una base $\mathcal{B}$ de V tal que $[T]_\mathcal{B}$ es diagonalizable, entonces T es diagonalizable.*

Dem. Si T es diagonalizable, entonces existe una base $\mathcal{C}$ de V en la cual $[T]_\mathcal{C}$ es diagonal. Si $\mathcal{B}$ es una base cualquiera de V, la parte 1 de la Proposición 1.1.2 implica que $[T]_\mathcal{B}$ es semejante a la matriz diagonal $[T]_\mathcal{C}$.

Si existe una base $\mathcal{B}$ de V tal que $[T]_\mathcal{B}$ es diagonalizable, entonces $[T]_\mathcal{B}$ es semejante a una matriz diagonal D. Luego la parte 2 de la Proposición 1.1.2 implica que existe una base $\tilde{\mathcal{B}}$ de V tal que $[T]_{\tilde{\mathcal{B}}} = D$. □

Corolario 1.1.8. *Una matriz A es diagonalizable si y solo si la transformación lineal L_A es diagonalizable.*

Dem. La matriz asociada a L_A en la base canónica es la matriz A. □

1.2. Valores y vectores propios

Definición 1.2.1. Dado $T \in \mathcal{L}(V)$, un escalar $\lambda \in \Bbbk$ se dice un *valor propio* de T si existe $0 \neq v \in V$ tal que $T(v) = \lambda v$, el vector v se dice que es un *vector propio* de T asociado a λ.

Ejemplo 1.2.2. En el caso en que $T : \mathbb{R}^2 \to \mathbb{R}^2$ es la simetría axial del ejemplo 1.1.5, nosotros vimos que existían vectores no nulos u y v tales que $T(u) = u$ y $T(v) = -v$; luego u y v son vectores propios de T con valores propios 1 y -1, respectivamente.

Notar que para la definición de valor y vector propio de no se necesita dimensión finita.

Ejemplo 1.2.3. Sea $C^\infty(\mathbb{R}) = \{f : \mathbb{R} \to \mathbb{R} \mid f \text{ tiene derivada de todos los órdenes}\}$. Definimos $T \in \mathcal{L}(C^\infty(\mathbb{R}))$ por $T(f) = f'$. Entonces $\lambda \in \mathbb{R}$ es un valor propio de T si y solo si existe $f \neq 0$ tal que $f' = \lambda f$, es decir si y solo si f es una solución no nula de la ecuación diferencial $x' = \lambda x$; luego todo $\lambda \in \mathbb{R}$ es un valor propio de T y los vectores propios correspondientes a λ son las funciones de la forma $f(t) = c\, e^{\lambda t}$, $\forall c \neq 0$.

Proposición 1.2.4. *Un operador $T \in \mathcal{L}(V)$ es diagonalizable si y solo si existe una base $\mathcal{B}$ de V formada por vectores propios de T. En ese caso las entradas diagonales de $[T]_\mathcal{B}$ son valores propios de T.*

Dem. El operador $T \in \mathcal{L}(V)$ es diagonalizable si y solo si existe una base $\mathcal{B}$ tal que $[T]_\mathcal{B}$ es de la forma

$$[T]_\mathcal{B} = \begin{pmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n \end{pmatrix}. \tag{1.1}$$

Si $\mathcal{B} = \{v_1, \dots, v_n\}$, entonces vale (1.1) si y solo si $T(v_i) = \lambda_i\, v_i$, para todo $i = 1, \dots, n$, es decir, si y solo si los elementos de $\mathcal{B}$ son vectores propios de T. $\square$

Ejemplo 1.2.5. Considerar el caso de la simetría axial vista en los ejemplos 1.1.5 y 1.2.2.

Definición 1.2.6. Dada $A \in M_n(\Bbbk)$, un escalar $\lambda \in \Bbbk$ se dice un *valor propio* de A si existe $0 \neq v \in \Bbbk^n$ tal que $Av = \lambda v$, el vector v se dice que es un *vector propio* de A asociado a λ. Es claro que los valores y vectores propios de la matriz A coinciden con los del operador L_A.

Corolario 1.2.7. *Una matriz $A \in M_n(\Bbbk)$ es diagonalizable si y solo si existe una base $\mathcal{B}$ de $\Bbbk^n$ formada por vectores propios de A. En este caso, si $\mathcal{B} = \{v_1, \dots, v_n\}$ y $\lambda_1, \dots, \lambda_n$ son los valores propios correspondientes, es*

$$A = QDQ^{-1}, \quad \textit{siendo} \quad D = \begin{pmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n \end{pmatrix} \quad \textit{y} \quad Q = [v_1 | \cdots | v_n]\,.$$

Dem. La primera afirmación se deduce de la proposición anterior aplicada al operador $L_A : \Bbbk^n \to \Bbbk^n$. Para la segunda, si $\mathcal{C}$ es la base canónica de $\Bbbk^n$, es $A = [L_A]_\mathcal{C} = {}_\mathcal{C}[\mathrm{Id}]_\mathcal{B} \cdot [L_A]_\mathcal{B} \cdot {}_\mathcal{B}[\mathrm{Id}]_\mathcal{C} = QDQ^{-1}$. $\square$

Ejemplo 1.2.8. Sea $A = \left(\begin{smallmatrix} 1 & 3 \\ 4 & 2 \end{smallmatrix}\right) \in M_2(\mathbb{R})$. Consideremos $v_1 = (1, -1)$ y $v_2 = (3, 4)$. Observar que $\mathcal{B} = \{v_1, v_2\}$ es una base de $\mathbb{R}^2$ y que $Av_1 = -2\, v_1$ y $Av_2 = 5\, v_2$, luego A es diagonalizable y vale

$$\begin{pmatrix} 1 & 3 \\ 4 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 3 \\ -1 & 4 \end{pmatrix} \begin{pmatrix} -2 & 0 \\ 0 & 5 \end{pmatrix} \begin{pmatrix} 1 & 3 \\ -1 & 4 \end{pmatrix}^{-1}.$$

1.3. Polinomio característico

Proposición 1.3.1. *Si $A, B \in M_n(\Bbbk)$, son matrices semejantes, entonces* $\det(A) = \det(B)$.

Dem. Sea $Q \in M_n(\Bbbk)$ invertible tal que $A = QBQ^{-1}$. Entonces

$$\det(A) = \det\left(QBQ^{-1}\right) = \det(Q)\det(B)\det\left(Q^{-1}\right) = \det(Q)\det(B)\det(Q)^{-1} = \det(B). \quad \square$$

Luego usando la proposición 1.1.2 se deduce el siguiente resultado.

Corolario 1.3.2. *Si $T \in \mathcal{L}(V)$ y $\mathcal{B}$, $\mathcal{C}$ son dos bases de V, entonces* $\det([T]_{\mathcal{B}}) = \det([T]_{\mathcal{C}})$. □

Visto el corolario anterior, tiene sentido la definición siguiente.

Definición 1.3.3. Si $T \in \mathcal{L}(V)$, definimos su *determinante* por $\det T := \det([T]_{\mathcal{B}})$ siendo $\mathcal{B}$ una base cualquiera de V.

Ejemplo 1.3.4. Sea $V = \mathbb{R}_2[x]$ y $T \in \mathcal{L}(V)$ definida por $T\big(p(x)\big) = p'(x)$. Si $\mathcal{B} = \big\{1, x, x^2\big\}$, es

$$[T]_{\mathcal{B}} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{pmatrix} \quad \Rightarrow \quad \det T = 0.$$

Proposición 1.3.5. *Sean $T, S \in \mathcal{L}(V)$.*

1. *Vale* $\det(T \circ S) = \det T \, \det S$.
2. *El operador T es invertible si y solo si* $\det T \neq 0$ *y en ese caso* $\det\big(T^{-1}\big) = (\det T)^{-1}$.
3. *Para todo λ en $\Bbbk$ y toda base $\mathcal{B}$ de V vale* $\det(T - \lambda \operatorname{Id}) = \det(A - \lambda I)$, *siendo* $A = [T]_{\mathcal{B}}$.

Dem. Sea $\mathcal{B}$ una base de V.

(1). $\det(T \circ S) = \det([T \circ S]_{\mathcal{B}}) = \det([T]_{\mathcal{B}}[S]_{\mathcal{B}}) = \det([T]_{\mathcal{B}}) \det([S]_{\mathcal{B}}) = \det T \, \det S$.

(2). La primera afirmación se deduce de lo siguiente

$$\det T \neq 0 \quad \Leftrightarrow \quad \det([T]_{\mathcal{B}}) \neq 0 \quad \Leftrightarrow \quad [T]_{\mathcal{B}} \text{ es invertible} \quad \Leftrightarrow \quad T \text{ es invertible.}$$

Si T es invertible, entonces $\det\big(T^{-1}\big) = \det\big(\big[T^{-1}\big]_{\mathcal{B}}\big) = \det\Big([T]_{\mathcal{B}}^{-1}\Big) = \det([T]_{\mathcal{B}})^{-1} = (\det T)^{-1}$.

(3). Si $\lambda \in \Bbbk$, es $\det(T - \lambda \operatorname{Id}) = \det([T - \lambda \operatorname{Id}]_{\mathcal{B}}) = \det([T]_{\mathcal{B}} - \lambda\, [\operatorname{Id}]_{\mathcal{B}}) = \det(A - \lambda I)$. □

Definición 1.3.6. Si $A \in M_n(\Bbbk)$, definimos su *polinomio característico* por

$$\chi_A(t) := \det(A - t\, I) \in \Bbbk[t].$$

Si $T \in \mathcal{L}(V)$, definimos su *polinomio característico* por

$$\chi_T(t) := \det(T - t \operatorname{Id}) \in \Bbbk[t].$$

Observación 1.3.7. De la Proposición 1.3.5 se deduce que si $A \in M_n(\Bbbk)$, $T \in \mathcal{L}(V)$ y $\mathcal{B}$ es una base de V, entonces

$$\chi_T(t) = \chi_{[T]_{\mathcal{B}}}(t), \quad \chi_A(t) = \chi_{L_A}(t).$$

Notar que si

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix}, \text{ es } \chi_A(t) = \begin{vmatrix} a_{11} - t & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} - t \end{vmatrix}.$$

Es un ejercicio el probar que vale la fórmula siguiente

$$\chi_A(t) = (-1)^n t^n + (-1)^{n-1} \operatorname{tr}(A)\, t^{n-1} + \cdots + \det(A).$$

Luego $\operatorname{gr} \chi_A(t) = n$ si $A \in M_n(\Bbbk)$ y $\operatorname{gr} \chi_T(t) = n$ si $T \in \mathcal{L}(V)$ y $\dim V = n$.

Ejemplo 1.3.8. Si $A = \left(\begin{smallmatrix} 1 & 1 \\ 4 & 1 \end{smallmatrix}\right)$, es $\chi_A(t) = \left|\begin{smallmatrix} 1-t & 1 \\ 4 & 1-t \end{smallmatrix}\right| = t^2 - 2t - 3$.

Ejemplo 1.3.9. Si $T \in \mathcal{L}(\mathbb{R}_2[x])$ está definida por $T(p(x)) = p'(x)$, (ejemplo 1.3.4) entonces

$$\chi_T(t) = \begin{vmatrix} -t & 1 & 0 \\ 0 & -t & 2 \\ 0 & 0 & -t \end{vmatrix} = -t^3.$$

Proposición 1.3.10. *Sea $T \in \mathcal{L}(V)$ y $\lambda \in \Bbbk$. Son equivalentes:*

1. *El escalar λ es un valor propio de T.*
2. $\operatorname{Ker}(T - \lambda \operatorname{Id}) \neq \{0\}$.
3. *El operador $T - \lambda \operatorname{Id}$ no es invertible.*
4. *El escalar λ es raíz de $\chi_T(t)$.*

Dem.

$$\begin{aligned} \lambda \text{ es valor propio de } T &\Leftrightarrow \exists\, v \neq 0 :\ T(v) = \lambda v \quad \Leftrightarrow \quad \exists\, v \neq 0 :\ v \in \operatorname{Ker}(T - \lambda \mathrm{Id}) \\ &\Leftrightarrow \operatorname{Ker}(T - \lambda \mathrm{Id}) \neq \{0\} \quad \Leftrightarrow \quad T - \lambda \mathrm{Id} \text{ no es inyectiva} \\ &\Leftrightarrow T - \lambda \mathrm{Id} \text{ no es biyectiva} \quad \Leftrightarrow \quad T - \lambda \mathrm{Id} \text{ no es invertible} \\ &\Leftrightarrow \det(T - \lambda \mathrm{Id}) = 0 \quad \Leftrightarrow \quad \chi_T(\lambda) = 0. \quad \square \end{aligned}$$

Observación 1.3.11. De la prueba anterior se deduce que si λ es un valor propio de $T \in \mathcal{L}(V)$, entonces $v \in V$ es un vector propio de T correspondiente a λ si y solo si $v \neq 0$ y $v \in \operatorname{Ker}(T - \lambda \operatorname{Id})$.

Corolario 1.3.12. *Si $T \in \mathcal{L}(V)$ y $\dim V = n$, entonces T tiene a lo más n valores propios.*

Dem. Los valores propios de T son las raíces de $\chi_T(t)$, y como este polinomio tiene grado n, entonces tiene a lo más n raíces. $\square$

Definición 1.3.13. Si $A \in M_n(\Bbbk)$, escribiremos $\operatorname{Ker}(A) := \{v \in K^n :\ Av = 0\}$. Notar $\operatorname{Ker}(A) = \operatorname{Ker}(L_A)$, siendo $L_A \in \mathcal{L}(K^n)$ la transformación lineal asociada a la matriz A.

Con la definición anterior tenemos que si $A \in M_n(\Bbbk)$ y λ es un valor propio de A, entonces $v \in \Bbbk^n$ es un vector propio de A correspondiente a λ si y solo si $v \neq 0$ y $v \in \operatorname{Ker}(A - \lambda I)$.

Ejemplo 1.3.14. Consideremos la matriz $A = \left(\begin{smallmatrix} 1 & 1 \\ 4 & 1 \end{smallmatrix}\right)$ del ejemplo 1.3.8. Es $\chi_A(t) = t^2 - 2t - 3 = (t-3)(t+1)$, luego A tiene valores propios $\lambda_1 = 3$ y $\lambda_2 = -1$. Operando obtenemos $\operatorname{Ker}(A - 3I) = [(1,2)]$ y $\operatorname{Ker}(A + I) = [(1,-2)]$. Notar que $\mathcal{B} = \{(1,2), (1,-2)\}$ es LI, luego $\mathcal{B}$ es una base de $\mathbb{R}^2$ formada por vectores propios de A. Esto implica que A es diagonalizable y vale

$$A = QDQ^{-1}, \text{ siendo } D = \begin{pmatrix} 3 & 0 \\ 0 & -1 \end{pmatrix} \text{ y } Q = \begin{pmatrix} 1 & 1 \\ 2 & -2 \end{pmatrix}.$$

Ejemplo 1.3.15. Sea $T \in \mathcal{L}(\mathbb{R}_2[x])$ definida por $T(p(x)) = p'(x)$ (Ejemplo 1.3.9). Es $\chi_T(t) = -t^3$, luego 0 es el único valor propio de T. Los vectores propios de T son los polinomios constantes no nulos, luego no existe una base de $\mathbb{R}_2[x]$ formada por vectores propios de T y T no es diagonalizable.

Recordar que si $\mathcal{B} = \{v_1, \ldots, v_n\}$ es una base de V, entonces el mapa *coordenadas* $\operatorname{coord}_{\mathcal{B}} : V \to \Bbbk^n$ definido por

$$\operatorname{coord}_{\mathcal{B}}(v) = (x_1, \ldots, x_n) \quad \text{si} \quad v = x_1 v_1 + \cdots + x_n v_n,$$

es un isomorfismo lineal.

Proposición 1.3.16. *Sea $T \in \mathcal{L}(V)$, $\mathcal{B}$ una base de V, $A = [T]_{\mathcal{B}}$ y λ un valor propio de T. Entonces $v \in V$ es un vector propio de T correspondiente a λ si y solo si* $\operatorname{coord}_{\mathcal{B}}(v) \in \Bbbk^n$ *es un vector propio de A correspondiente a λ.*

Dem. Utilizando que $\operatorname{coord}_{\mathcal{B}} : V \to \Bbbk^n$ es un isomorfismo, tenemos que dado $v \in V$, es $v \neq 0$ si y solo si $\operatorname{coord}_{\mathcal{B}}(v) \neq 0$ y

$$\begin{aligned} T(v) = \lambda\, v &\Leftrightarrow \operatorname{coord}_{\mathcal{B}}(T(v)) = \operatorname{coord}_{\mathcal{B}}(\lambda\, v) \Leftrightarrow [T]_{\mathcal{B}} \operatorname{coord}_{\mathcal{B}}(v) = \lambda \operatorname{coord}_{\mathcal{B}}(v) \\ &\Leftrightarrow A \operatorname{coord}_{\mathcal{B}}(v) = \lambda \operatorname{coord}_{\mathcal{B}}(v). \quad \square \end{aligned}$$

Ejemplo 1.3.17. Sea $T \in \mathcal{L}(\mathbb{R}_2[x])$ definida por $T(p(x)) = p(x) + x\, p'(x) + p'(x)$. Considerando la base $\mathcal{C} = \{1, x, x^2\}$ de $\mathbb{R}_2[x]$, es

$$A = [T]_{\mathcal{C}} = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 2 & 2 \\ 0 & 0 & 3 \end{pmatrix} \quad \Rightarrow \quad \chi_T(t) = \chi_A(t) = -(t-1)(t-2)(t-3),$$

luego los valores propios de T son 1, 2 y 3. Operando con la matriz A obtenemos:

$$\operatorname{Ker}(L_A - \mathrm{Id}) = [(1,0,0)], \quad \operatorname{Ker}(L_A - 2\mathrm{Id}) = [(1,1,0)], \quad \operatorname{Ker}(L_A - 3\mathrm{Id}) = [(1,2,1)].$$

Luego aplicando la proposición anterior deducimos

$$\operatorname{Ker}(T - \mathrm{Id}) = [1], \quad \operatorname{Ker}(T - 2\mathrm{Id}) = [1 + x], \quad \operatorname{Ker}(T - 3\mathrm{Id}) = \left[1 + 2x + x^2\right].$$

El conjunto $\mathcal{B} = \{1,\ 1+x,\ 1+2x+x^2\}$ es LI y por lo tanto $\mathcal{B}$ es una base de $\mathbb{R}_2[x]$ formada por vectores propios de T; luego T es diagonalizable y la matriz asociada a T en la base $\mathcal{B}$ es

$$[T]_{\mathcal{B}} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix}.$$

Ejemplo 1.3.18. *Rotación.* La rotación de ángulo θ ($0 \leq \theta < 2\pi$) es la transformación lineal $R_\theta : \mathbb{R}^2 \to \mathbb{R}^2$ definida por $R_\theta(x, y) = (x \cos\theta - y \operatorname{sen}\theta,\ x \operatorname{sen}\theta + y \cos\theta)$. Es

$$R_\theta \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \cos\theta & -\operatorname{sen}\theta \\ \operatorname{sen}\theta & \cos\theta \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}, \quad \chi_{R_\theta}(t) = t^2 - (2\cos\theta)t + 1.$$

El discriminante de la ecuación $t^2 - (2\cos\theta)t + 1 = 0$ es

$$\Delta = 4\cos^2\theta - 4 = 4\big(\cos^2\theta - 1\big) \leq 0.$$

Observar que $\Delta = 0$ si y solo si $\theta = 0, \pi$. Si $\theta = 0$, es $R_\theta = \mathrm{Id}$ y todos los vectores no nulos de $\mathbb{R}^2$ son vectores propios de R_θ con valor propio 1. Si $\theta = \pi$, es $R_\theta = -\mathrm{Id}$ (simetría central de centro en el origen) y todos los vectores no nulos de $\mathbb{R}^2$ son vectores propios de R_θ con valor propio -1. Si $\theta \neq 0, \pi$, es $\Delta < 0$ y R_θ no tiene valores propios por lo cual no es diagonalizable.

Teorema 1.3.19. *Sea $T \in \mathcal{L}(V)$, $\lambda_1, \ldots, \lambda_k$ valores propios distintos de T y $v_1, \ldots, v_k$ vectores propios correspondientes. Entonces $\{v_1, \ldots, v_k\}$ es LI.*

Dem. Lo probaremos por inducción en k.

Si $k=1$, entonces $\{v_1\}$ es LI porque como v_1 es un vector propio, es $v_1 \neq 0$.

Supongamos que el resultado es cierto para $k-1 \geq 1$ y sean $\lambda_1, \ldots, \lambda_k$ valores propios distintos de T y $v_1, \ldots, v_k$ vectores propios correspondientes. Sean $a_1, \ldots, a_k \in \Bbbk$ tales que

$$a_1 v_1 + \cdots + a_{k-1} v_{k-1} + a_k v_k = 0. \tag{1.2}$$

Aplicando T en ambos lados de la ecuación (1.2) y teniendo en cuenta que $v_1, \ldots, v_k$ son vectores propios obtenemos

$$a_1 \lambda_1 v_1 + \cdots + a_{k-1} \lambda_{k-1} v_{k-1} + a_k \lambda_k v_k = 0. \tag{1.3}$$

Multiplicando la ecuación (1.2) por $-\lambda_k$ obtenemos

$$-a_1 \lambda_k v_1 - \cdots - a_{k-1} \lambda_k v_{k-1} - a_k \lambda_k v_k = 0. \tag{1.4}$$

Sumando la ecuación (1.3) y la (1.4) obtenemos

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

Por la hipótesis de inducción el conjunto $\{v_1, \ldots, v_{k-1}\}$ es LI, luego

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

Como $\lambda_i \neq \lambda_k$ para todo $i = 1, \ldots, k-1$, es

$$a_1 = \cdots = a_{k-1} = 0.$$

Substituyendo $a_1, \ldots, a_{k-1}$ por 0 en la ecuación (1.2) obtenemos $a_k v_k = 0$ y como $v_k \neq 0$, es $a_k = 0$; esto completa la prueba de que $\{v_1, \ldots, v_k\}$ es LI. $\square$

Corolario 1.3.20. *Si la dimensión de V es n y $T \in \mathcal{L}(V)$ tiene n valores propios distintos, entonces T es diagonalizable.*

Dem. Si $\lambda_1, \ldots, \lambda_n$ son los valores propios de T y $v_1, \ldots, v_n$ son vectores propios correspondientes, entonces la proposición anterior implica que $\mathcal{B} = \{v_1, \ldots, v_n\}$ es LI y por lo tanto $\mathcal{B}$ es una base de V formada por vectores propios de T. Luego la proposición 1.2.4 implica que T es diagonalizable. $\square$

Ejemplo 1.3.21. Si $A = \left(\begin{smallmatrix} 1 & 1 \\ 1 & 1 \end{smallmatrix}\right)$, es $\chi_A(t) = t(t-2)$. Luego A es diagonalizable y vale $A \simeq \left(\begin{smallmatrix} 0 & 0 \\ 0 & 2 \end{smallmatrix}\right)$.

Observación 1.3.22. La matriz identidad y la matriz nula son ejemplos de matrices diagonales -luego diagonalizables- que tienen un solo valor propio, 1 y 0 respectivamente. Estos ejemplos nos muestran que el corolario anterior nos da una condición suficiente pero no necesaria para que un operador sea diagonalizable.

Definición 1.3.23. Decimos que un polinomio no constante $f(x) \in \Bbbk[x]$ *se escinde* en $\Bbbk[X]$ si existen escalares $a, a_1, \ldots, a_n$ tales que $f(x) = a(x - a_1) \cdots (x - a_n)$ (pueden haber elementos repetidos en $a_1, \ldots, a_n$).

Observación 1.3.24. Para abreviar diremos "$f(x)$ se escinde en $\Bbbk$" en vez de "$f(x)$ se escinde en $\Bbbk[x]$". También a veces diremos simplemente "$f(x)$ se escinde" cuando no sea necesario especificar el cuerpo $\Bbbk$.

Observar que si $f(x)$ se escinde en $\Bbbk$, entonces agrupando los términos repetidos obtenemos $f(x) = a(x - b_1)^{n_1} \cdots (x - b_r)^{n_r}$, con $n_i \geq 1$, para todo $i = 1, \ldots, r$, siendo $b_1, \ldots, b_r$ las distintas raíces de f.

Ejemplo 1.3.25.
1. $x^2 - 1 = (x+1)(x-1)$ se escinde en $\mathbb{Q}$.
2. $x^2 - 2$ no se escinde en $\mathbb{Q}$.

3. $x^2-2=\left(x+\sqrt{2}\right)\left(x-\sqrt{2}\right)$ se escinde en $\mathbb{R}$.

4. x^2+1 y $\left(x^2+1\right)(x-2)$ no se escinden en $\mathbb{R}$.

5. $x^2+1=(x+i)(x-i)$ se escinde en $\mathbb{C}$.

Observación 1.3.26. En $\mathbb{C}[x]$ todo polinomio no constante se escinde. Este es el llamado *teorema fundamental del álgebra*. La prueba de este resultado se encuentra en cualquier texto de análisis complejo.

Proposición 1.3.27. *Sea $T \in \mathcal{L}(V)$. Si T es diagonalizable, entonces χ_T se escinde.*

Dem. Sea $\mathcal{B}$ una base de V en la cual $[T]_{\mathcal{B}}$ es diagonal.

$$[T]_{\mathcal{B}} = \begin{pmatrix} a_1 & & \\ & \ddots & \\ & & a_n \end{pmatrix} \quad \Rightarrow \quad \chi_T(t) = (a_1-t)\cdots(a_n-t) = (-1)^n(t-a_1)\cdots(t-a_n). \quad \square$$

Corolario 1.3.28. *Sea $T \in \mathcal{L}(V)$. Si χ_T no se escinde en $\Bbbk$, entonces T no es diagonalizable.* $\square$

Ejemplo 1.3.29. Si $A = \left(\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right) \in M_2(\mathbb{R})$, es $\chi_A(t) = t^2+1$ que no se escinde en $\mathbb{R}$, luego A no es diagonalizable en $M_2(\mathbb{R})$. Observar que si consideramos $A \in M_2(\mathbb{C})$, es $\chi_A(t) = t^2+1 = (t-i)(t+i)$ que se escinde en $\mathbb{C}$, luego A es diagonalizable en $M_2(\mathbb{C})$ y es semejante a $\left(\begin{smallmatrix} i & 0 \\ 0 & -i \end{smallmatrix}\right)$.

Observación 1.3.30. Que el polinomio característico se escinda en $\Bbbk$ es una condición necesaria pero no suficiente para que un operador sea diagonalizable (ver más adelante el ejemplo 1.4.3).

1.4. Multiplicidad geométrica y algebraica

Definición 1.4.1. Sea $T \in \mathcal{L}(V)$ y λ un valor propio de T.

1. El *subespacio propio* asociado al valor propio λ es $E_\lambda := \operatorname{Ker}(T-\lambda \mathrm{Id})$.

2. La *multiplicidad geométrica* de λ es $MG(\lambda) := \dim E_\lambda$.

3. La *multiplicidad algebraica* de λ es $MA(\lambda) := \max\left\{h : \ (t-\lambda)^h \text{ divide a } \chi_T(t)\right\}$.

Si $A \in M_n(\Bbbk)$ y λ es un valor propio de A, entonces se define la multiplicidad algebraica y geométrica de λ como las correspondientes al operador $L_A \in \mathcal{L}(\Bbbk^n)$.

Observación 1.4.2. Sea $T \in \mathcal{L}(V)$.

1. El escalar 0 es un valor propio de T si y solo si $\operatorname{Ker}(T) \neq \{0\}$; en ese caso es $E_0 = \operatorname{Ker}(T)$.

2. El subespacio propio asociado a un valor propio λ consiste en el vector nulo y los vectores propios correspondientes a λ.

3. Si λ es un valor propio de T es

$$MG(\lambda) = \dim \operatorname{Ker}(T-\lambda \mathrm{Id}) = \dim V - \dim \operatorname{Im}(T-\lambda \mathrm{Id}) = \dim V - \operatorname{rango}(T-\lambda \mathrm{Id}).$$

   En particular si $A \in M_n(\Bbbk)$ y λ es un valor propio de A, entonces $MG(\lambda) = n - \operatorname{rango}(A-\lambda I)$.

4. Si λ es un valor propio de T, es $MA(\lambda) = m$ si y solo si $\chi_T(t) = (t-\lambda)^m p(t)$ con $p(\lambda) \neq 0$.

5. Si $\chi_T(t)$ se escinde y se escribe de la forma $\chi_T(t) = (-1)^n (t-\lambda_1)^{n_1} \cdots (t-\lambda_h)^{n_h}$ con $\lambda_i \neq \lambda_j$ si $i \neq j$, entonces $MA(\lambda_i) = n_i$, para todo $i = 1, \ldots, h$.

El siguiente es un ejemplo de un operador que no es diagonalizable pero su polinomio característico se escinde.

Ejemplo 1.4.3. Sea $T \in \mathcal{L}(\mathbb{R}_2[x])$ definida por $T(p(x)) = p'(x)$. En el ejemplo 1.3.15 probamos que $\chi_T(t) = -t^3$, luego 0 es el único valor propio de T y $E_0 = \operatorname{Ker}(T) = \mathbb{R} \subset \mathbb{R}_2[x]$. Entonces $MA(0) = 3$ y $MG(0) = 1$. Observar que si fuese T diagonalizable, al ser 0 su único valor propio, la matriz en la cual se diagonalizaría sería la matriz nula, por lo cual T sería el operador nulo. Como $T \neq 0$, deducimos que T no es diagonalizable.

Ejemplo 1.4.4. Sea

$$A = \begin{pmatrix} 3 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \end{pmatrix} \in M_4(\mathbb{R}).$$

Es $\chi_A(t) = (t-3)^2 \big(t^2+1\big)$, luego 3 es el único valor propio de A. Notar $MG(3) = 4 - \operatorname{rango}(A - 3I) = 4 - 2 = 2$. Luego $MA(3) = MG(3) = 2$.

Teorema 1.4.5. *Sea $T \in \mathcal{L}(V)$ y λ un valor propio de T, entonces*

$$1 \leq MG(\lambda) \leq MA(\lambda).$$

Dem. Como λ es un valor propio de T, es $E_\lambda \neq \{0\}$ y luego $MG(\lambda) \geq 1$.

Sean $\mathcal{B}_1$ una base de E_λ y $\mathcal{B}_2$ un conjunto LI de V disjunto con $\mathcal{B}_1$ tales que $\mathcal{B} = \mathcal{B}_1 \cup \mathcal{B}_2$ es una base de V. Observar que $T(v) = \lambda v$ para todo $v \in \mathcal{B}_1$, luego la matriz asociada a T en la base $\mathcal{B}$ es una matriz en bloques de la forma

$$[T]_\mathcal{B} = \begin{pmatrix} \lambda I_m & B \\ 0 & C \end{pmatrix}$$

siendo $m = \#\mathcal{B}_1 = \dim E_\lambda = MG(\lambda)$. Notar que $[T]_\mathcal{B}$ es una matriz en bloques, luego

$$\chi_T(t) = \begin{vmatrix} (\lambda - t) I_m & B \\ 0 & C - tI \end{vmatrix} = |(\lambda - t) I_m| \cdot |C - tI| = (\lambda - t)^m \chi_C(t).$$

Como puede ser $\chi_C(\lambda) = 0$, deducimos $MA(\lambda) \geq m = MG(\lambda)$. □

Corolario 1.4.6. *Sea $T \in \mathcal{L}(V)$ y λ un valor propio de T, entonces $MA(\lambda) = 1$ implica $MG(\lambda) = 1$.* □

En lo que sigue usaremos la suma directa de subespacios (§7.4).

Proposición 1.4.7. *Sea $T \in \mathcal{L}(V)$ y $\lambda_1, \ldots, \lambda_k$ valores propios distintos de T. Entonces $E_{\lambda_1}, \ldots, E_{\lambda_k}$ son subespacios independientes.*

Dem. Sean $v_1 \in E_{\lambda_1}, \ldots, v_k \in E_{\lambda_k}$ tales que $v_1 + \cdots + v_k = 0$. Si existiese algún $i \in \{1, \ldots, k\}$ tal que $v_i \neq 0$, eventualmente reordenando los subíndices existiría algún l, $1 \leq l \leq k$, tal que $v_i \neq 0$, para todo $i = 1, \ldots, l$ y $v_{l+1} = \cdots = v_k = 0$. Luego es

$$v_1 + \cdots + v_l = 0, \tag{1.5}$$

con $v_i \neq 0$, para todo $i = 1, \ldots, l$. Para cada $i \in \{1, \ldots, l\}$ es $0 \neq v_i \in E_{\lambda_i}$, luego v_i es un vector propio de T correspondiente al valor propio λ_i y por lo tanto el teorema 1.3.19 implica que $\{v_1, \ldots, v_l\}$ es LI. Esto contradice (1.5). Luego es $v_1 = \cdots = v_k = 0$ y por lo tanto $E_{\lambda_1}, \ldots, E_{\lambda_k}$ son subespacios independientes. □

El siguiente teorema da condiciones necesarias y suficientes para que un operador sea diagonalizable.

Teorema 1.4.8. *Sean $T \in \mathcal{L}(V)$ y $\lambda_1, \ldots, \lambda_h$ sus valores propios. Las siguientes afirmaciones son equivalentes.*

1. *El operador T es diagonalizable.*
2. *El polinomio χ_T se escinde y $MG(\lambda_i) = MA(\lambda_i)$ para todo $i = 1, \ldots, h$.*
3. *Vale $V = \sum_{i=1}^h E_{\lambda_i}$.*
4. *Vale $V = \bigoplus_{i=1}^h E_{\lambda_i}$.*

Dem. Observemos primero que la proposición 1.4.7 nos dice que $E_{\lambda_1}, \ldots, E_{\lambda_k}$ son subespacios independientes y por lo tanto su suma es directa. Luego las afirmaciones 3 y 4 son equivalentes.

$(1 \Rightarrow 2)$. Como T es diagonalizable, sabemos que χ_T se escinde y que V admite una base $\mathcal{B}$ formada por vectores propios. Agrupando los vectores propios correspondientes a cada valor propio, podemos suponer que esta base es de la forma $\mathcal{B} = \left\{v_1^1, \ldots, v_{n_1}^1, \ldots, v_1^h, \ldots, v_{n_h}^h\right\}$, siendo $\lambda_1, \ldots, \lambda_h$ los valores propios de T, con $\lambda_i \neq \lambda_j$ si $i \neq j$. Luego

$$[T]_{\mathcal{B}} = \begin{pmatrix} \lambda_1 & & & & & & \\ & \ddots & & & & & \\ & & \lambda_1 & & & & \\ & & & \ddots & & & \\ & & & & \lambda_h & & \\ & & & & & \ddots & \\ & & & & & & \lambda_h \end{pmatrix}.$$

Notar que cada valor propio λ_i aparece exactamente n_i veces en $[T]_{\mathcal{B}}$. Luego

$$\chi_T(t) = (\lambda_1 - t)^{n_1} \cdots (\lambda_h - t)^{n_h} = (-1)^n (t - \lambda_1)^{n_1} \cdots (t - \lambda_h)^{n_h}.$$

Observar que por la forma de la matriz $[T]_{\mathcal{B}}$ es $\left\{v_1^i, \ldots, v_{n_i}^i\right\} \subset E_{\lambda_i}$, luego $MG(\lambda_i) \geq n_i = MA(\lambda_i)$. Pero siempre vale $MG(\lambda_i) \leq MA(\lambda_i)$, luego $MG(\lambda_i) = MA(\lambda_i)$, para todo $i = 1, \ldots, h$.

$(2 \Rightarrow 4)$. Sea $\chi_T(t) = (-1)^n (t - \lambda_1)^{n_1} \cdots (t - \lambda_h)^{n_h}$ con $\lambda_i \neq \lambda_j$ si $i \neq j$. Entonces

$$\dim\left(\bigoplus_{i=1}^h E_{\lambda_i}\right) = \sum_{i=1}^h \dim E_{\lambda_i} = \sum_{i=1}^h MG(\lambda_i) = \sum_{i=1}^h MA(\lambda_i) = \sum_{i=1}^h n_i = \operatorname{gr} \chi_T(t) = \dim V \;\Rightarrow\; \bigoplus_{i=1}^h E_{\lambda_i} = V.$$

$(4 \Rightarrow 1)$. Si $\mathcal{B}_i$ es una base de E_{λ_i} para todo $i = 1, \ldots, h$, entonces $\mathcal{B} = \mathcal{B}_1 \cup \cdots \cup \mathcal{B}_h$ es una base de V formada por vectores propios de T. Luego la proposición 1.2.4 implica que T es diagonalizable. □

La siguiente es una aplicación interesante del teorema anterior para probar un resultado que no es obvio a partir de las definiciones.

Proposición 1.4.9. *Sean $A, D \in M_n(\mathbb{R})$ tales que D es una matriz diagonal, Si $A \simeq D$ en $M_n(\mathbb{C})$, entonces $A \simeq D$ en $M_n(\mathbb{R})$.*

Dem. Si es $A \simeq D$, entonces es $\chi_A(t) = (-1)^n (t - \lambda_1)^{n_1} \cdots (t - \lambda_h)^{n_h}$, siendo $\lambda_1, \ldots, \lambda_h \in \mathbb{R}$ las distintas entradas diagonales de D, y $n_i = n - \operatorname{rango}(A - \lambda_i I)$, para todo $i = 1, \ldots, h$. Como es $A - \lambda_i I \in M_n(\mathbb{R})$ y el rango de una matriz real es el mismo pensada en $M_n(\mathbb{R})$ o $M_n(\mathbb{C})$, entonces deducimos que $A \simeq D$ en $M_n(\mathbb{C})$ equivale con $A \simeq D$ en $M_n(\mathbb{R})$. □

Capítulo 2

Espacios con producto interno

En este tema el cuerpo de base $\Bbbk$ será $\mathbb{R}$ o $\mathbb{C}$. Diremos que un $\Bbbk$-espacio vectorial es *real* si $\Bbbk = \mathbb{R}$ y que es *complejo* si $\Bbbk = \mathbb{C}$. Como es usual pensamos $\mathbb{R} \subset \mathbb{C}$. Recordar que si $z = a + ib \in \mathbb{C}$, entonces su conjugado es $\overline{z} = a - ib \in \mathbb{C}$. Notar que $z \in \mathbb{R} \subset \mathbb{C}$ si y solo si $\overline{z} = z$.

2.1. Definiciones y propiedades básicas

El producto interno es la generalización del concepto de producto escalar a un espacio vectorial arbitrario.

Definición 2.1.1. Sea V un espacio vectorial. Un *producto interno* en V es una función $\langle\ ,\ \rangle : V \times V \to \Bbbk$ que verifica:

1. $\langle u + v, w\rangle = \langle u, w\rangle + \langle v, w\rangle,\ \forall u, v, w \in V$.
2. $\langle a\, u, v\rangle = a\langle u, v\rangle,\ \forall a \in \Bbbk,\ u, v \in V$.
3. $\langle u, v\rangle = \overline{\langle v, u\rangle},\ \forall u, v \in V$.
4. $\langle v, v\rangle > 0,\ \forall v \neq 0$.

Un *espacio con producto interno* es un par $(V, \langle\ ,\ \rangle)$ en el cual V es un espacio vectorial y $\langle\ ,\ \rangle : V \times V \to \Bbbk$ es un producto interno en V.

Observación 2.1.2.

1. Si $\Bbbk = \mathbb{R}$, la condición (3) queda en $\langle u, v\rangle = \langle v, u\rangle,\ \forall u, v \in V$.
2. Las dos primeras condiciones nos dicen que todo producto interno es lineal en la primera componente. Notar que si el espacio es real entonces también es lineal en la segunda componente, pero si el espacio es complejo esto deja de ser cierto (ver la proposición 2.1.6).

Ejemplo 2.1.3. En $\Bbbk^n$ el producto interno *usual* es

$$\langle (x_1, \dots, x_n), (y_1, \dots, y_n)\rangle = x_1\, \overline{y_1} + x_2\, \overline{y_2} + \cdots + x_n\, \overline{y_n} = \sum_{i=1}^{n} x_i\, \overline{y_i}.$$

Cuando consideremos a $\Bbbk^n$ como espacio vectorial con producto interno nos estaremos refiriendo siempre al producto interno usual, a menos que explicitemos lo contrario. En el caso $\Bbbk = \mathbb{R}$, al producto interno usual también se le suele llamar *producto escalar* y la fórmula anterior queda

$$\langle (x_1, \dots, x_n), (y_1, \dots, y_n)\rangle = x_1\, y_1 + x_2\, y_2 + \cdots + x_n\, y_n = \sum_{i=1}^{n} x_i\, y_i.$$

Notar que este producto es la generalización del producto escalar de $\mathbb{R}^2$ y $\mathbb{R}^3$.

El siguiente ejemplo es una generalización del anterior.

Ejemplo 2.1.4. Si $A = (a_{ij}), B = (b_{ij}) \in M_{m\times n}(\Bbbk)$, entonces definimos $\langle A, B\rangle = \sum_{i,j=1}^{m,n} a_{ij}\overline{b_{ij}}$. Es fácil de probar que la fórmula anterior define un producto interno en $M_{m\times n}(\Bbbk)$. Observar que si identificamos $M_{m\times n}(\Bbbk)$ con $\Bbbk^{mn}$ mediante el isomorfismo $T : M_{m\times n}(\Bbbk) \to \Bbbk^{mn}$ definido por

$$T\big((a_{ij})\big) = (a_{11}, \dots, a_{1n}, \ \dots \ , a_{m1}, \dots, a_{mn}),$$

entonces este producto en $M_{m\times n}(\Bbbk)$ coincide con el producto interno usual de $\Bbbk^{mn}$.

Ejemplo 2.1.5. Sea $C\big([0,1]\big) := \{f : [0,1] \to \mathbb{R} : \ f \text{ es continua}\}$, entonces

$$\langle f, g\rangle = \int_0^1 f(t)\, g(t)\, dt$$

define un producto interno en $C\big([0,1]\big)$. Probaremos solo la propiedad 4, que es la única que requiere cierta elaboración. Sea $f \in C\big([0,1]\big)$ tal que $f \neq 0$. Esto último implica que existe $x_0 \in [0,1]$ tal que $f(x_0) \neq 0$. Supongamos $x_0 \in (0,1)$. Como f^2 es continua, entonces existen $\epsilon > 0$ y $\delta > 0$ tales que $(x_0-\delta, x_0+\delta) \subset [0,1]$ y $f^2(x) > \epsilon$, para todo $x \in (x_0-\delta, x_0+\delta)$. Luego

$$\langle f, f\rangle = \int_0^1 f^2(t)\, dt = \int_0^{x_0-\delta} f^2(t)\, dt + \int_{x_0-\delta}^{x_0+\delta} f^2(t)\, dt + \int_{x_0+\delta}^1 f^2(t)\, dt \geq \int_{x_0-\delta}^{x_0+\delta} f^2(t)\, dt \geq 2\delta\epsilon > 0.$$

Los casos $x_0 = 0$ y $x_0 = 1$ se prueban en forma análoga.

Notar que si en vez de funciones continuas consideramos funciones integrables, entonces la propiedad 4 deja de valer y el anterior no es un producto interno.

De ahora en adelante $V = (V, \langle\ ,\ \rangle)$ es un espacio con producto interno.

Proposición 2.1.6. *Valen las siguientes propiedades.*

1. $\langle u, v+w\rangle = \langle u, v\rangle + \langle u, w\rangle$, *para todo* $u, v, w \in V$.
2. $\langle u, a\,v\rangle = \overline{a}\langle u, v\rangle$, *para todo* $u, v \in V$ *y* $a \in \Bbbk$.
3. $\langle v, v\rangle = 0$ *si y solo si* $v = 0$.
4. *Si* $\langle v, u\rangle = 0$ *para todo* $v \in V$, *entonces* $u = 0$.
5. *Si* $\langle v, u\rangle = \langle v, w\rangle$ *para todo* $v \in V$, *entonces* $u = w$.

Dem.

1. $\langle u, v+w\rangle = \overline{\langle v+w, u\rangle} = \overline{\langle v, u\rangle + \langle w, u\rangle} = \overline{\langle v, u\rangle} + \overline{\langle w, u\rangle} = \langle u, v\rangle + \langle u, w\rangle$.
2. $\langle u, a\,v\rangle = \overline{\langle a\,v, u\rangle} = \overline{a\,\langle v, u\rangle} = \overline{a}\,\overline{\langle v, u\rangle} = \overline{a}\,\langle u, v\rangle$.
3. Para que sea menos confuso, en esta parte escribiremos o al vector nulo. Por un lado, es $\langle o, o\rangle = \langle 0o, o\rangle = 0\langle o, o\rangle = 0$. Por el otro lado, si $v \neq o$, entonces $\langle v, v\rangle > 0$ y por lo tanto $\langle v, v\rangle \neq 0$. Luego el único caso en que vale $\langle v, v\rangle = 0$ es cuando $v = o$.
4. Si $\langle v, u\rangle = 0$ para todo $v \in V$, entonces tomando $v = u$ es $\langle u, u\rangle = 0$, luego $u = 0$.
5. Si $\langle v, u\rangle = \langle v, w\rangle$ para todo $v \in V$, entonces es $0 = \langle v, u\rangle - \langle v, w\rangle = \langle v, u-w\rangle$ para todo $v \in V$; luego es $u - w = 0$ y resulta $u = w$. □

Observación 2.1.7. De la afirmación 3 de la proposición anterior se deduce que la condición 4 de la definición de producto interno equivale a las dos siguientes.

$$\langle v,v\rangle \geq 0,\ \forall v\in V \quad \text{y} \quad \langle v,v\rangle = 0 \ \Rightarrow\ v=0.$$

Definición 2.1.8. Si $v \in V$, definimos su *norma* por $||v|| := \sqrt{\langle v,v\rangle}$. La norma define una función $||\ || : V \to \mathbb{R}$.

Ejemplo 2.1.9. Si $V = \Bbbk^n$ es $||(x_1,\dots,x_n)|| = \sqrt{|x_1|^2+|x_2|^2+\cdots+|x_n|^2} = \sqrt{\sum_{i=1}^n |x_i|^2}$.

Observación 2.1.10. Si $z = a+ib \in \mathbb{C}$ con $a,b\in\mathbb{R}$, entonces $\operatorname{Re}(z) = a$ e $\operatorname{Im}(z) = b$ son su *parte real* e *imaginaria*, respectivamente. Observar que para todo $z \in \mathbb{C}$ vale

$$|\operatorname{Re}(z)| \leq |z|, \quad |\operatorname{Im}(z)| \leq |z|, \quad z+\overline{z} = 2\operatorname{Re}(z), \quad z-\overline{z} = 2i\operatorname{Im}(z) \quad \text{y} \quad z\overline{z} = |z|^2.$$

Proposición 2.1.11. *Vale la siguiente igualdad*

$$||u+v||^2 = ||u||^2 + 2\operatorname{Re}(\langle u,v\rangle) + ||v||^2, \quad \forall u,v\in V. \tag{2.1}$$

Dem.

$$\begin{aligned} ||u+v||^2 &= \langle u+v,u+v\rangle = \langle u,u\rangle + \langle u,v\rangle + \langle v,u\rangle + \langle v,v\rangle = ||u||^2 + \langle u,v\rangle + \overline{\langle u,v\rangle} + ||v||^2 \\ &= ||u||^2 + 2\operatorname{Re}(\langle u,v\rangle) + ||v||^2. \quad \square \end{aligned}$$

Observar que si $\Bbbk = \mathbb{R}$, la relación (2.1) queda en

$$||u+v||^2 = ||u||^2 + 2\langle u,v\rangle + ||v||^2, \quad \forall u,v\in V. \tag{2.2}$$

Proposición 2.1.12. *Valen las siguientes propiedades.*

1. $||a\,v|| = |a|\,||v||$, *para todo* $v\in V$ *y* $a\in\Bbbk$.
2. $||v|| \geq 0$ *para todo* $v \in V$ *y* $||v|| = 0$ *si y solo si* $v = 0$.
3. Desigualdad de Cauchy-Schwarz.

   $$|\langle u,v\rangle| \leq ||u||\,||v||, \quad \forall u,v\in V.$$

   Además vale la igualdad si y solo si $\{u,v\}$ *es LD.*
4. Desigualdad triangular.

   $$||u+v|| \leq ||u|| + ||v||, \quad \forall u,v\in V.$$

Dem. Las dos primeras afirmaciones se deducen inmediatamente de la definición de norma. Para probar la desigualdad de Cauchy-Schwarz, notar primero que si $v=0$, entonces $\{u,v\} = \{u,0\}$ es LD y

$$|\langle u,v\rangle| = |\langle u,0\rangle| = 0 = ||u||\,||0|| = ||u||\,||v||.$$

Supongamos ahora $v \neq 0$. Aplicando la fórmula (2.1) obtenemos

$$\begin{aligned} \left\|u - \frac{\langle u,v\rangle}{||v||^2}v\right\|^2 &= ||u||^2 - 2\operatorname{Re}\left(\left\langle u, \frac{\langle u,v\rangle}{||v||^2}v\right\rangle\right) + \left\|\frac{\langle u,v\rangle}{||v||^2}v\right\|^2 = ||u||^2 - 2\operatorname{Re}\left(\frac{\overline{\langle u,v\rangle}}{||v||^2}\langle u,v\rangle\right) + \frac{|\langle u,v\rangle|^2}{||v||^4}||v||^2 \\ &= ||u||^2 - 2\frac{|\langle u,v\rangle|^2}{||v||^2} + \frac{|\langle u,v\rangle|^2}{||v||^2} = ||u||^2 - \frac{|\langle u,v\rangle|^2}{||v||^2}. \end{aligned}$$

Luego

$$0 \leq \left\|u - \frac{\langle u,v\rangle}{||v||^2}v\right\|^2 = ||u||^2 - \frac{|\langle u,v\rangle|^2}{||v||^2}. \tag{2.3}$$

De (2.3) se obtiene inmediatamente la desigualdad de Cauchy-Schwarz.

Si $\{u,v\}$ es LD, como $v\neq 0$ entonces existe $a\in\Bbbk$ tal que $u=a\,v$, luego

$$|\langle u,v\rangle| = |\langle a\,v,v\rangle| = |a\,\langle v,v\rangle| = |a|\,||v||^2 = |a|\,||v||\,||v|| = ||a\,v||\,||v|| = ||u||\,||v||.$$

Recíprocamente si $|\langle u,v\rangle| = ||u||\,||v||$ y $v\neq 0$ entonces (2.3) implica $\left\|u-\frac{\langle u,v\rangle}{||v||^2}v\right\|^2=0$, luego $u=\langle u,v\rangle/||v||^2 v$.

La desigualdad triangular se deduce de la desigualdad Cauchy-Schwarz:

$$\begin{aligned}||u+v||^2 &= ||u||^2+2\operatorname{Re}(\langle u,v\rangle)+||v||^2 \leq ||u||^2+2|\langle u,v\rangle|+||v||^2 \leq ||u||^2+2||u||\,||v||+||v||^2\\ &= (||u||+||v||)^2 \quad\Rightarrow\quad ||u+v||\leq ||u||+||v||. \quad \square\end{aligned}$$

Aplicación 2.1.1. *En el caso $V=\Bbbk^n$, la desigualdad de Cauchy-Schwarz y la desigualdad triangular nos dan las siguientes relaciones.*

$$\left|\sum_{i=1}^n x_i\,\overline{y_i}\right| \leq \sqrt{\sum_{i=1}^n |x_i|^2}\sqrt{\sum_{i=1}^n |y_i|^2},\qquad \sqrt{\sum_{i=1}^n |x_i+y_i|^2}\leq \sqrt{\sum_{i=1}^n |x_i|^2}+\sqrt{\sum_{i=1}^n |y_i|^2}.$$

Si $\Bbbk=\mathbb{R}$ entonces las fórmulas quedan más simples:

$$\left|\sum_{i=1}^n x_iy_i\right| \leq \sqrt{\sum_{i=1}^n x_i^2}\sqrt{\sum_{i=1}^n y_i^2},\qquad \sqrt{\sum_{i=1}^n (x_i+y_i)^2}\leq \sqrt{\sum_{i=1}^n x_i^2}+\sqrt{\sum_{i=1}^n y_i^2}.$$

2.2. Ortogonalidad

En esta sección V es un espacio con producto interno.

Definición 2.2.1. Un *versor* es un vector de norma 1. Si dos vectores $u,v\in V$ verifican $\langle u,v\rangle=0$, entonces decimos que son *ortogonales* y escribimos $u\perp v$.

Observaciones 2.2.2. Si v es un vector no nulo, entonces $\frac{v}{\|v\|}:=\frac{1}{\|v\|}v$ es un versor colineal con v. Notar que el vector nulo es el único vector de V que es ortogonal a todos los vectores de V (proposición 2.1.6).

Definición 2.2.3. Un subconjunto S de V se dice *ortogonal* si para todo u,v en S, con $u\neq v$, es $u\perp v$; si además los elementos de S son versores, entonces decimos que S es un conjunto *ortonormal*. Observar que si $S=\{v_1,\dots,v_n\}$, entonces S es ortonormal si y solo si $\langle v_i,v_j\rangle=\delta_{ij}$ para todo i,j. Una *base ortonormal* de V es un conjunto ortonormal que es base del espacio V.

Ejemplos 2.2.4.

1. La base canónica de $\Bbbk^n$ es una base ortonormal.
2. El conjunto $S=\{(1,1,0),\,(1,-1,1),\,(-1,1,2)\}\subset\mathbb{R}^3$ es ortogonal pero no es ortonormal. Calculando las normas de los elementos de S obtenemos $\|(1,1,0)\|=\sqrt{2}$, $\|(1,-1,1)\|=\sqrt{3}$, $\|(-1,1,2)\|=\sqrt{6}$. Luego $S'=\left\{\frac{1}{\sqrt{3}}(1,1,0),\,\frac{1}{\sqrt{2}}(1,-1,1),\,\frac{1}{\sqrt{6}}(-1,1,2)\right\}\subset\mathbb{R}^3$ es un conjunto ortonormal.

A continuación veremos algunas propiedades de los conjuntos ortonormales.

Proposición 2.2.5 (Teorema de Pitágoras)**.** *Si u y v son ortogonales, entonces* $||u+v||^2=||u||^2+||v||^2$.

Dem. Aplicar la fórmula (2.1). $\square$

Corolario 2.2.6. *Si* $\{v_1,\dots,v_n\}$ *es un conjunto ortogonal, es* $||\sum_{i=1}^n v_i||^2 = \sum_{i=1}^n ||v_i||^2$.

Dem. Lo demostramos por inducción en n. Si $n=2$ es el teorema de Pitágoras. Supongamos que vale para n y consideremos un conjunto ortogonal $\{v_1,\dots,v_n,v_{n+1}\}$. Como v_{n+1} es ortogonal a $v_1,\dots,v_n$ entonces resulta ortogonal a $\sum_{i=1}^n v_i$. Luego aplicando Pitágoras y la hipótesis de inducción obtenemos

$$\left|\left|\sum_{i=1}^{n+1} v_i\right|\right|^2 = \left|\left|\left(\sum_{i=1}^{n} v_i\right) + v_{n+1}\right|\right|^2 = \left|\left|\sum_{i=1}^{n} v_i\right|\right|^2 + ||v_{n+1}||^2 = \sum_{i=1}^{n} ||v_i||^2 + ||v_{n+1}||^2 = \sum_{i=1}^{n+1} ||v_i||^2. \quad \square$$

Proposición 2.2.7. *Sea* $S=\{v_1,\dots,v_n\}$ *un conjunto ortogonal de vectores no nulos y* $u \in [v_1,\dots,v_n]$.

$$\text{Si } u = \sum_{i=1}^n a_i\, v_i, \ \text{ entonces } a_i = \frac{\langle u, v_i\rangle}{||v_i||^2}, \ \forall i = 1,\dots,n.$$

Dem.

$$u = \sum_{j=1}^n a_j\, v_j \quad \Rightarrow \quad \langle u, v_i\rangle = \left\langle \sum_{j=1}^n a_j\, v_j, v_i \right\rangle = \sum_{j=1}^n a_j \langle v_j, v_i\rangle = a_i ||v_i||^2 \quad \Rightarrow \quad a_i = \frac{\langle u, v_i\rangle}{||v_i||^2}. \quad \square$$

Corolario 2.2.8. *Si S es un subconjunto ortogonal (finito o infinito) de V formado por vectores no nulos, entonces S es LI.*

Dem. Si $v_1,\dots,v_n \in S$ y $a_1,\dots,a_n \in \Bbbk$ son tales que $\sum_{i=1}^n a_i\, v_i = 0$ entonces $a_i = \frac{\langle 0, v_i\rangle}{||v_i||^2} = 0$ para todo $i = 1,\dots,n$. $\square$

Observación 2.2.9. La proposición anterior implica que si V tiene dimensión finita y $\mathcal{B} = \{v_1,\dots,v_n\}$ es una base ortonormal de V, entonces vale

$$u = \sum_{i=1}^n \langle u, v_i\rangle\, v_i, \quad \forall u \in V. \tag{2.4}$$

Definición 2.2.10. Sea $\mathcal{B}$ un conjunto (posiblemente infinito) ortonormal en V. Si $w \in V$, entonces los escalares $\langle w, v\rangle$, con $v \in \mathcal{B}$, se llaman los *coeficientes de Fourier* de w respecto a $\mathcal{B}$.

Teorema 2.2.11 (Método de Gram-Schmidt)**.** *Sea* $\{v_1,\dots,v_n\}$ *un conjunto LI. Definimos* $\{w_1,\dots,w_n\}$ *mediante*

$$\begin{aligned}
w_1 &= v_1 \\
w_2 &= v_2 - \frac{\langle v_2, w_1\rangle}{||w_1||^2} w_1 \\
w_3 &= v_3 - \frac{\langle v_3, w_1\rangle}{||w_1||^2} w_1 - \frac{\langle v_3, w_2\rangle}{||w_2||^2} w_2 \\
&\vdots \\
w_n &= v_n - \frac{\langle v_n, w_1\rangle}{||w_1||^2} w_1 - \cdots - \frac{\langle v_n, w_{n-1}\rangle}{||w_{n-1}||^2} w_{n-1}.
\end{aligned}$$

Entonces $\{w_1,\dots,w_n\}$ *es un conjunto ortogonal LI que verifica* $[w_1,\dots,w_n] = [v_1,\dots,v_n]$.

Dem. Sean $S_n = \{v_1,\dots,v_n\}$ y $S'_n = \{w_1,\dots,w_n\}$, siendo $w_1,\dots,w_n$ definidos como arriba. Demostraremos por inducción en n, que si S_n es un conjunto LI, entonces S'_n es un conjunto ortogonal de vectores no nulos (y por lo tanto LI) que verifica $[S'_n] = [S_n]$.

Para $n = 1$ es $S_1 = S_1' = \{v_1\}$ y se cumple la tesis. Supongamos que vale para $n-1$. Si S_n es un conjunto LI, entonces S_{n-1} es un conjunto LI y se aplica la hipótesis de inducción, por lo cual S_{n-1}' es un conjunto ortogonal de vectores no nulos que verifica $\left[S_{n-1}'\right] = [S_{n-1}]$.
Consideremos w_n. Si fuese $w_n = 0$ entonces sería $v_n \in [S_{n-1}'] = [S_{n-1}]$ y $S_n = S_{n-1} \cup \{v_n\}$ resultaría LD, contradiciendo que S es LI. Luego es $w_n \neq 0$.

El conjunto $S_{n-1}' = \{w_1, \ldots, w_{n-1}\}$ es ortogonal. Probaremos que w_n es ortogonal a $w_1, \ldots, w_{n-1}$. Sea $i \in \{1, \ldots, n-1\}$, entonces

$$\begin{aligned}\langle w_n, w_i\rangle &= \langle v_n, w_i\rangle - \frac{\langle v_n, w_1\rangle}{||w_1||^2}\langle w_1, w_i\rangle - \cdots - \frac{\langle v_n, w_i\rangle}{||w_i||^2}\langle w_i, w_i\rangle - \cdots - \frac{\langle v_n, w_{n-1}\rangle}{||w_{n-1}||^2}\langle w_{n-1}, w_i\rangle \\ &= \langle v_n, w_i\rangle - \langle v_n, w_i\rangle = 0.\end{aligned}$$

Luego $S_n' = \{w_1, \ldots, w_{n-1}, w_n\}$ es ortogonal.

De $\left[S_{n-1}'\right] = [S_{n-1}]$ se deduce que $w_1, \ldots, w_{n-1} \in [S_{n-1}] \subset [S_n]$. Por otro lado tenemos que es $w_n \in [w_1, \ldots, w_{n-1}, v_n] \subset S_n$, luego $[S_n'] \subset [S_n]$. Como S_n' es un conjunto ortogonal de vectores no nulos, es LI y S_n también es LI. Luego $\dim[S_n'] = \dim[S_n] = n$ y por lo tanto $[S_n'] = [S_n]$. $\square$

Observación 2.2.12. Si $S = \{v_1, \ldots, v_l, v_{l+1}, \ldots, v_n\}$ es un conjunto LI tal que $\{v_1, \ldots, v_l\}$ es un conjunto ortonormal, entonces aplicándole a S el método de Gram-Schmidt obtenemos $w_i = v_i$, para todo $i = 1, \ldots, l$.

Corolario 2.2.13. *Si la dimensión de V es finita, entonces V admite una base ortonormal.*

Dem. Sea $\mathcal{C} = \{v_1, \ldots, v_n\}$ una base cualquiera de V. Aplicando el método de Gram-Schmidt obtenemos una base ortogonal $\mathcal{C}' = \{u_1, \ldots, u_n\}$. Si definimos $w_i = u_i/||u_i||$, $i = 1\ldots, n$, entonces $\mathcal{B} = \{w_1, \ldots, w_n\}$ es una base ortonormal de V. $\square$

Observación 2.2.14. Si W es un subespacio de V, entonces W es un espacio vectorial y la restrición del producto interno de V a W le da a W una estructura de espacio vectorial con producto interno.

Ejemplo 2.2.15. Consideremos el plano $W = \{(x, y, z) \in \mathbb{R}^3 :\ 2x+3y-z = 0\}$, veremos de hallar una base ortonormal de W. Empezamos por encontrar una base cualquiera de W, por ejemplo $\mathcal{B}_1 = \{(1, 0, 2),\ (0, 1, 3)\}$. A esta base le aplicamos el proceso de Gram-Schmidt.

$$\begin{aligned}w_1 &= (1, 0, 2),\\ w_2 &= (0, 1, 3) - \frac{\langle (0, 1, 3), (1, 0, 2)\rangle}{\|(1, 0, 2)\|^2}(1, 0, 2) = (0, 1, 3) - \frac{6}{5}(1, 0, 2) = \left(-\frac{6}{5}, 1, \frac{3}{5}\right).\end{aligned}$$

Luego $\mathcal{B}_2 = \left\{(1, 0, 2), \left(-\frac{6}{5}, 1, \frac{3}{5}\right)\right\}$ es una base ortogonal de W. Notar que es $\left(-\frac{6}{5}, 1, \frac{3}{5}\right) = \frac{1}{5}(-6, 5, 3)$. Luego $\mathcal{B}_3 = \{(1, 0, 2),\ (-6, 5, 3)\}$ también es una base ortogonal de W. Finalmente normalizando los vectores de $\mathcal{B}_3$ obtenemos que $\mathcal{B}_4 = \left\{\frac{1}{\sqrt{5}}(1, 0, 2),\ \frac{1}{\sqrt{70}}(-6, 5, 3)\right\}$ es una base ortonormal de W.

Observación 2.2.16. Supongamos que la dimensión de V es finita y que $\mathcal{B} = \{v_1, \ldots, v_n\}$ es una base ortonormal de V. Sean $v = \sum_{i=1}^n x_i\, v_i$ y $w = \sum_{i=1}^n y_i\, v_i$, entonces

$$\langle v, w\rangle = \left\langle \sum_{i=1}^n x_i\, v_i, \sum_{j=1}^n y_j\, v_j \right\rangle = \sum_{i,j=1}^n x_i\, \overline{y_j}\ \langle v_i, v_j\rangle = \sum_{i,j=1}^n x_i\, \overline{y_j}\, \delta_{ij} = \sum_{i=1}^n x_i\, \overline{y_i}.$$

Luego

$$\langle v, w\rangle = \sum_{i=1}^n x_i\, \overline{y_i}, \quad \|v\|^2 = \sum_{i=1}^n |x_i|^2.$$

Notar la analogía con el caso de $\mathbb{k}^n$ con el producto interno usual.

De la observación anterior y la fórmula (2.4) se obtiene lo siguiente.

Proposición 2.2.17 (Identidad de Parseval)**.** *Sea V un espacio de dimensión finita y $\mathcal{B} = \{v_1, \ldots, v_n\}$ una base ortonormal de V. Entonces*

$$\langle v, w\rangle = \sum_{i=1}^{n} \langle v, v_i\rangle \overline{\langle w, v_i\rangle}, \quad \forall v, w \in V.$$

En particular, $\|v\|^2 = \sum_{i=1}^{n} |\langle v, v_i\rangle|^2$. □

El teorema de Riesz. Recordar que el espacio dual de V es $V^* = \mathcal{L}(V, \mathbb{k})$. Si $w \in V$, entonces el mapa $\alpha : V \to \mathbb{k}$ definido por $\alpha(v) = \langle v, w\rangle$, para todo $v \in V$, es una transformación lineal, luego $\alpha \in V^*$.

Ejemplo 2.2.18. Si $\alpha \in (\mathbb{R}^n)^*$, entonces existen $a_1, \ldots, a_n \in \mathbb{R}$ tales que $\alpha(x_1, \ldots, x_n) = a_1x_1 + \cdots + a_nx_n$, para todo $(x_1, \ldots, x_n) \in \mathbb{R}^n$. Luego considerando el producto interno usual (el escalar) en $\mathbb{R}^n$, vemos que podemos escribir α de la forma

$$\alpha(x_1, \ldots, x_n) = \langle (x_1, \ldots, x_n), (a_1, \ldots, a_n)\rangle, \quad \forall (x_1, \ldots, x_n) \in \mathbb{R}^n.$$

Si ahora en vez de $\mathbb{R}^n$ consideramos $\mathbb{C}^n$, entonces manteniendo las notaciones anteriores obtenemos

$$\alpha(x_1, \ldots, x_n) = a_1x_1 + \cdots + a_nx_n = \langle (x_1, \ldots, x_n), (\overline{a_1}, \ldots, \overline{a_n})\rangle, \quad \forall (x_1, \ldots, x_n) \in \mathbb{C}^n.$$

Luego todo $\alpha \in (\mathbb{k}^n)^*$ (siendo $\mathbb{k} = \mathbb{R}$ o $\mathbb{C}$) es de la forma $\alpha(v) = \langle v, w\rangle$, para cierto $w \in \mathbb{k}^n$.

El siguiente resultado muestra que si estamos en un espacio con producto interno de dimensión finita, entonces eso siempre sucede.

Teorema 2.2.19 (Riesz)**.** *Si V es de dimensión finita y $\alpha \in V^*$, entonces existe un único $w \in V$ tal que*

$$\alpha(v) = \langle v, w\rangle, \quad \forall v \in V. \tag{2.5}$$

Dem. Sea $\mathcal{B} = \{v_1, \ldots, v_n\}$ una base ortonormal de V. Si existe $w \in V$ que verifica (2.5), entonces aplicando la fórmula (2.4) obtenemos

$$w = \sum_{i=1}^{n} \langle w, v_i\rangle v_i = \sum_{i=1}^{n} \overline{\langle v_i, w\rangle}\, v_i = \sum_{i=1}^{n} \overline{\alpha(v_i)}\, v_i.$$

Luego $w = \sum_{i=1}^{n} \overline{\alpha(v_i)}\, v_i$. Esto prueba la unicidad de w. Probaremos ahora que el vector w así definido verifica (2.5). Observemos primero que como es $w = \sum_{i=1}^{n} \langle w, v_i\rangle v_i$, entonces vale $\overline{\alpha(v_i)} = \langle w, v_i\rangle$ y por lo tanto $\alpha(v_i) = \overline{\langle w, v_i\rangle}$, para todo i.

Sea $v \in V$ arbitrario. Es $v = \sum_{i=1}^{n} \langle v, v_i\rangle v_i$, luego aplicando la identidad de Parseval y la linealidad de α obtenemos

$$\langle v, w\rangle = \sum_{i=1}^{n} \langle v, v_i\rangle \overline{\langle w, v_i\rangle} = \sum_{i=1}^{n} \langle v, v_i\rangle\, \alpha(v_i) = \alpha\left(\sum_{j=1}^{n} \langle v, v_j\rangle v_j\right) = \alpha(v). \quad \square$$

Observación 2.2.20. Si $\mathcal{B} = \{v_1, \ldots, v_n\}$ es una base ortonormal de V, entonces vale $\langle v_i, v_j\rangle = \delta_{ij}$, para todo i, j. Luego si definimos $\alpha_i \in V^*$ mediante $\alpha_i(v) = \langle v, v_i\rangle$, para todo $v \in V$, entonces es $\alpha_i(v_j) = \delta_{ij}$, para todo i, j. Luego la base dual de $\mathcal{B}$ es $\{\alpha_1, \ldots, \alpha_n\}$.

El complemento ortogonal.

Definición 2.2.21. Si S es un subconjunto cualquiera de V, el conjunto

$$S^\perp := \{v \in V : \ v \perp w, \ \forall w \in S\}$$

se llama el *complemento ortogonal* de S.

Ejemplo 2.2.22. Valen $\{0\}^\perp = V$ y $V^\perp = \{0\}$.

Observación 2.2.23. Es fácil de probar que $S^\perp$ siempre es un subespacio de V (aunque S no lo sea).

Proposición 2.2.24. *Si S es un subconjunto cualquiera de V, entonces $S^\perp = [S]^\perp$.*

Dem. Por comodidad de notación lo probaremos para el caso en que $S = \{w_1, \ldots, w_n\}$ es un conjunto finito, pero el resultado vale en general.

Es fácil de probar que $S \subset [S]$ implica $[S]^\perp \subset S^\perp$, así que solo probaremos la otra inclusión. Sea $v \in S^\perp$. Si $w = \sum_{i=1}^n a_i w_i \in [S]$, entonces $\langle w, v\rangle = \langle \sum_{i=1}^n a_i w_i, v\rangle = \sum_{i=1}^n a_i \langle w_i, v\rangle = \sum_{i=1}^n a_i 0 = 0$. Como esto vale para todo $w \in [S]$, concluimos $v \in S^\perp$. Luego $S^\perp \subset [S]^\perp$. $\square$

Observación 2.2.25. Este último resultado implica que si W es un subespacio de V y $\{w_1, \ldots, w_n\}$ es un generador de W, entonces $v \in W^\perp$ si y solo si $v \perp w_i$ para todo $i = 1, \ldots, n$.

Teorema 2.2.26. *Si W es un subespacio de dimensión finita de V, entonces $V = W \oplus W^\perp$.*

Dem. Sea $\mathcal{B} = \{w_1, \ldots, w_n\}$ una base ortonormal de W, si $v \in V$ es

$$v = \sum_{i=1}^n \langle v, w_i\rangle \ w_i + v - \sum_{i=1}^n \langle v, w_i\rangle \ w_i.$$

Observar que $\langle v - \sum_{i=1}^n \langle v, w_i\rangle \ w_i, w_j\rangle = \langle v, w_j\rangle - \sum_{i=1}^n \langle v, w_i\rangle \ \langle w_i, w_j\rangle = 0$ para todo $j = 1, \ldots, n$, luego por la observación anterior es $v - \sum_{i=1}^n \langle v, w_i\rangle \ w_i \in W^\perp$ y claramente $\sum_{i=1}^n \langle v, w_i\rangle \ w_i \in W$, de donde $V = W + W^\perp$. Por otro lado si $w \in W \cap W^\perp$ es $w \perp w$ y luego $w = 0$; esto prueba $W \cap W^\perp = \{0\}$. $\square$

Corolario 2.2.27. *Si la dimensión de V es finita y $W \subset V$ es un subespacio, entonces*

$$\dim W + \dim W^\perp = \dim V. \quad \square$$

Definición 2.2.28. Sea W un subespacio de dimensión finita de V. Sean $P_W \in \mathcal{L}(V)$ y $P_{W^\perp} \in \mathcal{L}(V)$ las proyecciones asociadas a la descomposición $V = W \oplus W^\perp$, siendo $\operatorname{Im} P_W = W$ y $\operatorname{Im} P_{W^\perp} = W^\perp$. La proyección P_W se llama la *proyección ortogonal* sobre el espacio W. Observar que si $\mathcal{B} = \{w_1, \ldots, w_n\}$ es una base ortonormal de W, entonces

$$P_W(v) = \sum_{i=1}^n \langle v, w_i\rangle \ w_i, \quad P_{W^\perp}(v) = v - \sum_{i=1}^n \langle v, w_i\rangle \ w_i, \quad \forall v \in V.$$

Ejemplo 2.2.29. Sea $W = \{(x, y, z, t) \in \mathbb{R}^4 : \ x - y + z = 0 \text{ y } 2y - z + t = 0\}$ y consideramos $\mathbb{R}^4$ con el producto escalar. Veamos de hallar $W^\perp$. Lo primero es encontrar una base cualquiera de W, por ejemplo $\mathcal{B} = \{(-1, 0, 1, 1), (1, 1, 0, -2)\}$. Sabemos que es $W^\perp = \mathcal{B}^\perp$. Luego

$(x, y, z, t) \in W^\perp \Leftrightarrow \langle (x, y, z, t), (1, 1, 0, -2)\rangle = \langle (x, y, z, t), (-1, 0, 1, 1)\rangle = 0 \Leftrightarrow x+y-2t = 0$ y $-x+z+t = 0$, luego $W^\perp = \{(x,y,z,t) \in \mathbb{R}^4 : \ x+y-2t=0 \text{ y } -x+z+t=0\}$. Vamos a hallar la proyección P_W. Aplicando Gram-Schmidt a la base $\mathcal{B}$ y normalizando obtenemos $\mathcal{B}_1 = \left\{\frac{1}{\sqrt{3}}(-1,0,1,1), \frac{1}{\sqrt{3}}(0,1,1,-1)\right\}$ base ortonormal de W. Luego aplicando la fórmula de arriba obtenemos

$$\begin{aligned} P_W(x,y,z,t) &= \frac{1}{3}(-x+z+t)(-1,0,1,1) + \frac{1}{3}(y+z-t)(0,1,1,-1) \\ &= \frac{1}{3}(x-z-t,\, y+z-t,\, -x+y+2z,\, -x-y+2t), \quad \forall (x,y,z,t) \in \mathbb{R}^4. \end{aligned}$$

Corolario 2.2.30. *Sea W un subespacio de dimensión finita de V, $P_W \in \mathcal{L}(V)$ la proyección ortogonal sobre W y $v \in V$. Entonces:*

1. *$\forall w \in W$ es*
$$||v-w|| \geq ||v-P_W(v)||. \tag{2.6}$$

2. *Si $w_0 \in W$ verifica*
$$||v-w|| \geq ||v-w_0||, \tag{2.7}$$
$\forall w \in W$, entonces $w_0 = P_W(v)$.

Dem. Observar que si $w \in W$, es $v - P_W(v) = P_{W^\perp}(v) \in W^\perp$ y $P_W(v) - w \in W$, luego aplicando el teorema de Pitágoras resulta

$$||v-w||^2 = ||v-P_W(v)+P_W(v)-w||^2 = ||v-P_W(v)||^2 + ||P_W(v)-w||^2 \geq ||v-P_W(v)||^2. \tag{2.8}$$

Esto implica la primera afirmación. Para la segunda afirmación, tomando $w = P_W(v)$ en (2.7) resulta $||v-P_W(v)|| \geq ||v-w_0||$. Por otro lado tomando $w = w_0$ en (2.6) resulta $||v-w_0|| \geq ||v-P_W(v)||$. Luego $||v-w_0|| = ||v-P_W(v)||$. Teniendo en cuenta esta última relación, tomando $w = w_0$ en (2.8) deducimos $||P_W(v)-w_0|| = 0$ y luego $P_W(v) = w_0$. $\square$

Definición 2.2.31. 1. Sean v y w en V, definimos la *distancia* entre v y w por $d(v,w) = \|v-w\|$.

2. Sea S un subconjunto no vacío de V y $v \in V$. Se define la *distancia* de v a S por
$$d(v,S) := \text{ínf}\{d(v,w) : \ w \in S\}.$$

Observación 2.2.32. Si escribimos la relación (2.6) en términos de distancias, es

$$d(v,w) \geq d(v,P_W(v)), \quad \forall w \in W.$$

Esto implica que $d(v,P_W(v))$ es una cota inferior para $\{d(v,w) : \ w \in W\}$, pero como $P_W(v) \in W$, resulta

$$d(v,P_W(v)) = \text{mín}\{d(v,w) : \ w \in W\} = d(v,W).$$

La relación (2.7) nos dice que $P_W(v)$ es el único elemento de W que realiza la distancia de v a W.

Corolario 2.2.33 (Desigualdad de Bessel)**.** *Sea $\{v_1,\ldots,v_n\}$ un subconjunto ortonormal de V, entonces*

$$||v||^2 \geq \sum_{i=1}^n |\langle v, v_i\rangle|^2, \quad \forall v \in V.$$

Dem. Sea $W = [v_1,\ldots,v_n]$. El teorema 2.2.26 implica $V = W \oplus W^\perp$, luego aplicando Pitágoras es

$$||v||^2 = ||P_W(v)+v-P_W(v)||^2 = ||P_W(v)||^2 + ||v-P_W(v)||^2 \geq ||P_W(v)||^2 = \sum_{i=1}^n |\langle v, v_i\rangle|^2. \quad \square$$

Capítulo 3

Operadores en espacios con producto interno

En este capítulo V será siempre un espacio vectorial (real o complejo) con producto interno y de dimensión finita. Nuestro objetivo es estudiar las propiedades de los operadores en V. Como sabemos, los operadores diagonalizables tienen buenas propiedades y también las tienen las bases ortonormales. Por eso nos concentraremos primero en estudiar los operadores que son diagonalizables en bases ortonormales. Luego aplicaremos lo anterior a las proyecciones ortogonales y finalmente estudiaremos los operadores que preservan el producto interno. El siguiente resultado será usado con frecuencia.

Proposición 3.0.1. *Sean $T \in \mathcal{L}(V)$, $\mathcal{B} = \{v_1, \ldots, v_n\}$ una base ortonormal de V y $[T]_{\mathcal{B}} = (a_{ij})$. Entonces $a_{ij} = \langle T(v_j), v_i \rangle$, para todo $i, j = 1, \ldots, n$.*

Dem. Si $[T]_{\mathcal{B}} = (a_{ij})$ es $T(v_i) = \sum_{j=1}^n a_{ji} v_j$, para todo $i = 1, \ldots, n$. Como $\mathcal{B}$ es ortonormal, entonces aplicando la fórmula (2.4) sabemos que también vale $T(v_i) = \sum_{j=1}^n \langle T(v_i), v_j \rangle v_j$. Luego $a_{ji} = \langle T(v_i), v_j \rangle$. □

3.1. Operadores autoadjuntos

Empezamos definiendo un concepto que veremos que está directamente relacionado con el problema que nos interesa estudiar.

Definición 3.1.1. Un operador $T \in \mathcal{L}(V)$ es *autoadjunto* si verifica

$$\langle T(u), v \rangle = \langle u, T(v) \rangle, \quad \forall u, v \in V.$$

Proposición 3.1.2. *Sea $T \in \mathcal{L}(V)$ y $\mathcal{B} = \{v_1, \ldots, v_n\}$ una base de V. Entonces T es autoadjunto si y solo si verifica*

$$\langle T(v_i), v_j \rangle = \langle v_i, T(v_j) \rangle, \quad \forall i, j = 1, \ldots, n. \tag{3.1}$$

Dem. Es claro que vale el directo, luego lo único que hay que probar es el recíproco. Si $u, v \in V$ son vectores arbitrarios, y escribimos $u = \sum_{i=1}^n x_i v_i$ y $v = \sum_{j=1}^n y_j v_j$, entonces

$$\langle T(u), v \rangle = \left\langle \sum_{i=1}^n x_i T(v_i), \sum_{j=1}^n y_j v_j \right\rangle = \sum_{i=j=1}^n x_i \overline{y_j} \langle T(v_i), v_j \rangle,$$

$$\langle u, T(v) \rangle = \left\langle \sum_{i=1}^n x_i v_i, \sum_{j=1}^n y_j T(v_j) \right\rangle = \sum_{i=j=1}^n x_i \overline{y_j} \langle v_i, T(v_j) \rangle.$$

Luego (3.1) implica que T es autoadjunto. □

Definición 3.1.3. Una matriz $A = (a_{ij}) \in M_n(\mathbb{C})$ se dice que es *hermitiana* si verifica $a_{ij} = \overline{a_{ji}}$, para todo i, j. Notar que las entradas diagonales de una matriz hermitiana siempre son números reales.

Observación 3.1.4. Si $A = (a_{ij}) \in M_n(\mathbb{R}) \subset M_n(\mathbb{C})$, entonces la condición de ser hermitiana queda en $a_{ij} = a_{ji}$, para todo i, j. Luego una matriz real es hermitiana si y solo si es simétrica.

Notar que las matrices hermitianas coinciden con las simétricas solo en el caso real. Por ejemplo, si consideramos $A = \left(\begin{smallmatrix} 3 & 1+2i \\ 1-2i & 4 \end{smallmatrix}\right)$ y $B = \left(\begin{smallmatrix} 3 & 1+2i \\ 1+2i & 4 \end{smallmatrix}\right)$, entonces A es hermitiana pero no es simétrica, mientras que B es simétrica pero no es hermitiana.

Proposición 3.1.5. *Sea $T \in \mathcal{L}(V)$, $\mathcal{B}$ base ortonormal de V y $A = [T]_{\mathcal{B}}$. Entonces T es autoadjunto si y solo si A es simétrica en el caso real o hermitiana en el caso complejo.*

Dem. Sea $\mathcal{B} = \{v_1, \dots, v_n\}$. Si $A = (a_{ij})$, entonces la proposición 3.0.1 implica que es $a_{ji} = \langle T(v_i), v_j \rangle$. Luego vale $a_{ij} = \overline{a_{ji}}$, para todo i, j, si y solo si vale $\langle T(v_j), v_i \rangle = \overline{\langle T(v_i), v_j \rangle}$, para todo i, j, lo cual equivale a la fórmula (3.1). □

La proposición anterior sirve para determinar los operadores autoadjuntos de $\mathbb{k}^n$.

Corolario 3.1.6. *Sean $A \in M_n(\mathbb{k})$ y el operador asociado $L_A \in \mathcal{L}(\mathbb{k}^n)$. Si en $\mathbb{k}^n$ consideramos el producto interno usual, entonces $L_A : \mathbb{k}^n \to \mathbb{k}^n$ es autoadjunto, es decir vale*

$$\langle Ax, y \rangle = \langle x, Ay \rangle, \quad \forall x, y \in \mathbb{k}^n,$$

si y solo si A es simétrica en el caso real o hermitiana en el caso complejo.

Dem. La base canónica $\mathcal{B}$ de $\mathbb{k}^n$ es ortonormal y $A = [L_A]_{\mathcal{B}}$. □

El siguiente resultado explica nuestro interés en los operadores autoadjuntos.

Corolario 3.1.7. *Sea $T \in \mathcal{L}(V)$. Si $\mathbb{k} = \mathbb{R}$ y T es diagonalizable en una base ortonormal o $\mathbb{k} = \mathbb{C}$ y T es diagonalizable en una base ortonormal con valores propios reales, entonces T es autoadjunto.*

Dem. En el caso real, si existe una base ortonormal $\mathcal{B}$ de V tal que $[T]_{\mathcal{B}}$ es una matriz diagonal, entonces $[T]_{\mathcal{B}}$ es simétrica (real) y por lo tanto T es autoadjunto.

En el caso complejo pasa lo mismo, porque como $[T]_{\mathcal{B}}$ es una matriz diagonal y sus entradas diagonales son los valores propios de T (que son reales), entonces $[T]_{\mathcal{B}}$ es una matriz simétrica real y por lo tanto es una matriz hermitiana compleja, lo cual implica que T es autoadjunto. □

Observación 3.1.8. En el caso real, el ser autoadjunto es una condición necesaria para que el operador sea diagonalizable en una base ortonormal. En el caso complejo no lo es, como lo muestra el siguiente ejemplo.

Ejemplo 3.1.9. Sea $A = \left(\begin{smallmatrix} i & 0 \\ 0 & 1 \end{smallmatrix}\right) \in M_2(\mathbb{C})$. Considerando $\mathbb{C}^2$ con el producto interno usual, el operador $L_A : \mathbb{C}^2 \to \mathbb{C}^2$ es diagonalizable en una base ortonormal (la canónica), pero como A no es hermitiana, deducimos que L_A no es autoadjunto.

Observación 3.1.10. El siguiente resultado muestra que los operadores autoadjuntos solo pueden tener valores propios reales.

Proposición 3.1.11. *Si $T \in \mathcal{L}(V)$ es autoadjunto y $\lambda \in \mathbb{C}$ es un valor propio de T, entonces $\lambda \in \mathbb{R}$.*

Dem. Sea $v \in V$ un vector propio correspondiente a λ. Entonces

$$\lambda \|v\|^2 = \langle \lambda v, v \rangle = \langle T(v), v \rangle = \langle v, T(v) \rangle = \langle v, \lambda v \rangle = \overline{\lambda} \|v\|^2.$$

Como es $\|v\| \neq 0$, deducimos $\overline{\lambda} = \lambda$; luego $\lambda \in \mathbb{R}$. □

A continuación probaremos el recíproco del corolario 3.1.7. Para eso necesitamos el siguiente resultado.

Teorema 3.1.12. *Si $A \in M_n(\mathbb{R})$ es simétrica, entonces A tiene algún valor propio real.*

Dem. Veremos dos pruebas.

Prueba 1 (sin usar números complejos). Sea $\langle\ ,\ \rangle$ el producto interno usual (escalar) de $\mathbb{R}^n$. Definimos $f : \mathbb{R}^n \to \mathbb{R}$ mediante $f(x) = \langle x, Ax\rangle$, para todo $x \in \mathbb{R}^n$. Observar que si $A = (a_{ij})$, entonces

$$f(x_1, \ldots, x_n) = \sum_{i=j=1}^{n} a_{ij} x_i x_j, \quad \forall (x_1, \ldots, x_n) \in \mathbb{R}^n,$$

luego f es una función polinomial y por lo tanto es continua.

Sea $S = \{x \in \mathbb{R}^n : \ \|x\| = 1\}$. Como S es compacto y f es continua, entonces el teorema de Weierstrass nos dice que f tiene un mínimo en S, es decir, existe v_0 en S tal que $f(v_0) \leq f(x)$, para todo x en S. Notar que al ser $\|v_0\| = 1$ se tiene $v_0 \neq 0$.

Sea w un vector arbitrario de $\mathbb{R}^n$ tal que $w \perp v_0$ y $\|w\| = 1$. Definimos $\alpha : \mathbb{R} \to \mathbb{R}^n$ mediante

$$\alpha(t) = \cos(t)\, v_0 + \operatorname{sen}(t)\, w, \qquad \forall t \in \mathbb{R}.$$

Observar que al ser $\|v_0\| = \|w\| = 1$ y $v_0 \perp w$, el teorema de Pitágoras implica $\|\alpha(t)\|^2 = \cos^2 t + \operatorname{sen}^2 t = 1$, luego $\alpha(t) \in S$ y por lo tanto $f(v_0) \leq f\big(\alpha(t)\big)$, para todo t en $\mathbb{R}$.

Sea $\beta = f \circ \alpha : \mathbb{R} \to \mathbb{R}$. Por lo anterior es $\beta(0) \leq \beta(t)$, para todo t en $\mathbb{R}$; esto nos dice que β tiene un mínimo absoluto en 0, y como β es derivable, deducimos $\beta'(0) = 0$. Para calcular $\beta'(0)$, empezamos observando que vale $\alpha'(t) = -\operatorname{sen}(t)\, v_0 + \cos(t)\, w$ y por lo tanto $\alpha(0) = v_0$ y $\alpha'(0) = w$. Por otro lado es $\beta(t) = \langle \alpha(t), A\alpha(t)\rangle$, luego $\beta'(t) = \langle \alpha'(t), A\alpha(t)\rangle + \langle \alpha(t), A\alpha'(t)\rangle$. Sustituyendo t por 0 y usando que A es simétrica obtenemos

$$0 = \beta'(0) = \langle w, Av_0\rangle + \langle v_0, Aw\rangle = \langle w, Av_0\rangle + \langle Av_0, w\rangle = 2\langle w, Av_0\rangle.$$

Luego Av_0 es ortogonal con w, para todo w en $\mathbb{R}^n$ tal que $w \perp v_0$ y $\|w\| = 1$. Pero es fácil de probar que esto implica que Av_0 es ortogonal con w, para todo w en $\mathbb{R}^n$ tal que $w \perp v_0$. Esto implica $Av_0 \in \{v_0\}^{\perp\perp} = [v_0]$ y por lo tanto existe λ en $\mathbb{R}$ tal que $Av_0 = \lambda v_0$.

Prueba 2 (usando números complejos). Consideremos $L_A : \mathbb{C}^n \to \mathbb{C}^n$. Como A es simétrica real, entonces A pensada en $M_n(\mathbb{C})$ es hermitiana y por lo tanto $L_A : \mathbb{C}^n \to \mathbb{C}^n$ es autoadjunto. Como estamos en $\mathbb{C}$ sabemos que L_A tiene algún valor propio λ y al ser L_A autoadjunto deducimos que es $\lambda \in \mathbb{R}$. Como los valores propios de L_A coinciden con los de A, concluimos que A tiene algún valor propio en $\mathbb{R}$. $\square$

Corolario 3.1.13. *Si $T \in \mathcal{L}(V)$ es autoadjunto, entonces T tiene algún valor propio.*

Dem. Si $\Bbbk = \mathbb{C}$ el resultado es inmediato, dado que el polinomio característico de T siempre tiene alguna raíz, la cual es un valor propio de T (esto no requiere que T sea autoadjunto). Supongamos ahora que es $\Bbbk = \mathbb{R}$. Sea $\mathcal{B}$ una base ortonormal de V. Como $T \in \mathcal{L}(V)$ es autoadjunto, entonces $[T]_{\mathcal{B}}$ es una matriz simétrica real; luego el teorema anterior implica que $[T]_{\mathcal{B}}$ tiene algún valor propio (real) y por lo tanto T tiene algún valor propio. $\square$

Observación 3.1.14. Si $T \in \mathcal{L}(V)$, un subespacio $W \subset V$ se dice T-*invariante* si verifica $T(w) \in W$, para todo $w \in W$. En ese caso $T|_W : W \to W$ definida por $(T|_W)(w) = T(w)$, para todo $w \in W$ es un operador en W. Notar que si $T \in \mathcal{L}(V)$ es autoadjunto y $W \subset V$ es un subespacio T-invariante, entonces $T|_W \in \mathcal{L}(W)$ también es autoadjunto.

Proposición 3.1.15. *Si $T \in \mathcal{L}(V)$ es autoadjunto, entonces T es diagonalizable en una base ortonormal.*

Dem. La prueba es por inducción en $n = \dim V$. Para $n = 1$ la prueba es trivial, porque en ese caso todo vector no nulo es un vector propio[1] de T. Luego tomando $v \in V$ con $\|v\| = 1$, obtenemos que $\{v\}$ es

[1] Si $\dim V = 1$, entonces toda transformación lineal de V en V es de la forma $T(v) = \lambda\, v$ para algún $\lambda \in \Bbbk$.

una base ortonormal de V formada por vectores propios de T.

Sea ahora $n > 1$ y supongamos razonando inductivamente que si tenemos un operador autoadjunto en un espacio de dimensión $n-1$, entonces existe una base ortonormal del espacio formada por vectores propios del operador.

Como T es autoadjunto, entonces el corolario anterior implica que existe λ valor propio de T. Sea v un vector propio de T correspondiente a λ y consideremos $W := [v]^\perp = \{w \in V \mid \langle w, v\rangle = 0\}$. Sabemos que W es un subespacio de V de dimensión $n-1$. Sea $w \in W$,

$$\langle T(w), v\rangle = \langle w, T(v)\rangle = \langle w, \lambda\, v\rangle = \lambda \langle w, v\rangle = \lambda\, 0 = 0 \quad \Rightarrow \quad T(w) \in W.$$

Luego W es T-invariante y como T es autoadjunto se deduce que $T|_W \in \mathcal{L}(W)$ es autoadjunto. Aplicando la hipótesis inductiva a $T|_W$, deducimos que existe $\{v_1, \dots, v_{n-1}\}$ base ortonormal de W formada por vectores propios de T. Sea $v_n = \frac{v}{||v||}$, como $v_1, \dots, v_{n-1} \in W = [v]^\perp$ y $v_n \in [v]$ es $v_i \perp v_n$, $\forall i = 1, \dots, n-1$. Luego $\mathcal{B} = \{v_1, \dots, v_{n-1}, v_n\}$ es una base ortonormal de V formada por vectores propios de T. $\square$

Corolario 3.1.16. *Si $A \in M_n(\mathbb{R})$ es simétrica, entonces A es diagonalizable en $M_n(\mathbb{R})$.*

Dem. Como $A \in M_n(\mathbb{R})$ es simétrica, entonces $L_A : \mathbb{R}^n \to \mathbb{R}^n$ es autoadjunto (con el producto escalar), luego el operador L_A es diagonalizable y sabemos que eso equivale a que la matriz A sea diagonalizable. $\square$

Juntando el corolario 3.1.7 y las proposiciones 3.1.11 y 3.1.15, obtenemos el siguiente resultado.

Teorema 3.1.17. *Sea $T \in \mathcal{L}(V)$.*

1. *Si $\Bbbk = \mathbb{R}$, entonces T es autoadjunto si y solo si T es diagonalizable en una base ortonormal.*
2. *Si $\Bbbk = \mathbb{C}$, entonces T es autoadjunto si y solo si T es diagonalizable en una base ortonormal y sus valores propios son reales.* $\square$

Observación 3.1.18. En el caso real, el teorema anterior nos dice que ser autoadjunto es una condición necesaria y suficiente para que un operador sea diagonalizable en una base ortonormal. En el caso complejo es una condición suficiente, pero vimos en el ejemplo 3.1.9 que no es necesaria. Para obtener una condición necesaria y suficiente en el caso complejo vamos a tener que introducir la normalidad, que estudiaremos más adelante.

3.2. Proyecciones ortogonales

En lo que sigue veremos de caracterizar las proyecciones ortogonales como operadores.

Proposición 3.2.1. *Sea $P \in \mathcal{L}(V)$. Entonces P es la proyección ortogonal sobre algún subespacio si y solo si P es una proyección y es un operador autoadjunto.*

Dem. Sea $W \subset V$ un subespacio y P_W la proyección ortogonal sobre W. Probaremos que P_W es autoadjunto. Sean $v_1, v_2 \in V$ arbitrarios. Escribimos $v_1 = w_1 + w_1'$ y $v_2 = w_2 + w_2'$, donde $w_1, w_2 \in W$, $w_1', w_2' \in W^\perp$, $i = 1, 2$. Entonces

$$\begin{aligned} \langle P_W(v_1), v_2\rangle &= \langle w_1, w_2 + w_2'\rangle = \langle w_1, w_2\rangle + \langle w_1, w_2'\rangle = \langle w_1, w_2\rangle, \\ \langle v_1, P_W(v_2)\rangle &= \langle w_1 + w_1', w_2\rangle = \langle w_1, w_2\rangle + \langle w_1', w_2\rangle = \langle w_1, w_2\rangle. \end{aligned}$$

Luego $\langle P_W(v_1), v_2\rangle = \langle v_1, P_W(v_2)\rangle$, $\forall v_1, v_2 \in V$.

Sea ahora P una proyección que además es un operador autoadjunto. La relación $P^2 = P$ implica $V = \text{Im } P \oplus \text{Ker } P$ y P es la proyección sobre $\text{Im}(P)$ en la dirección de $\text{Ker } P$. Para ver que P es la proyección ortogonal sobre $\text{Im}(P)$, lo único que nos falta probar es $\text{Ker } P = (\text{Im } P)^\perp$. Sean $u \in \text{Im } P$ (luego $P(u) = u$) y $v \in \text{Ker } P$, es

$$\langle u, v\rangle = \langle P(u), v\rangle = \langle u, P(v)\rangle = \langle u, 0\rangle = 0 \quad \Rightarrow \quad \text{Ker } P \subset (\text{Im } P)^\perp.$$

De $V = \text{Im } P \oplus \text{Ker } P$ y $V = \text{Im } P \oplus (\text{Im } P)^\perp$ se deduce que $\dim \text{Ker } P = \dim(\text{Im } P)^\perp$, luego la relación $\text{Ker } P \subset (\text{Im } P)^\perp$ implica $\text{Ker } P = (\text{Im } P)^\perp$. Luego $P = P_W$, siendo $W = \text{Im } P$. □

Definición 3.2.2. Un operador $P \in \mathcal{L}(V)$ es una *proyección ortogonal* si es una proyección y además es un operador autoadjunto. La proposición anterior prueba que toda proyección ortogonal es la proyección ortogonal sobre algún subespacio (este subespacio es la imagen de la proyección).

Definición 3.2.3. Sean U, W subespacios de V, decimos U y W son *ortogonales* y escribimos $U \perp W$ si vale $u \perp w$, para todo $u \in U$ y $w \in W$.

Proposición 3.2.4. *Consideremos las siguientes familias de datos.*

(A) *$W_1, \dots, W_k$ son subespacios de V tales que $V = \bigoplus_{i=1}^k W_i$ y $W_i \perp W_j$, si $i \neq j$.*

(B) *$P_1, \dots, P_k \in \mathcal{L}(V)$ son operadores que verifican las siguientes condiciones.*

1. *$P_1, \dots, P_k$ son proyecciones ortogonales.*
2. *$P_i \circ P_j = 0$, para todo $i \neq j$, con $i, j = 1, \dots, k$.*
3. *$\text{Id} = P_1 + \cdots + P_k$.*

Entonces **(A)** *y* **(B)** *son equivalentes.*

Dem. Sean $W_1, \dots, W_k$ como en **(A)**. Definimos $P_i(v) = w_i$, siendo $v = \sum_{i=1}^k w_i$, $w_i \in W_i$, $i = 1, \dots, k$, entonces la proposición 7.5.2 implica que $P_1, \dots, P_k$ son proyecciones que verifican las condiciones 2) y 3) de **(B)**. Para ver que las P_i son proyecciones ortogonales, solo resta probar que los P_i son autoadjuntos.

Sean $u = \sum_{j=1}^k u_j$ y $v = \sum_{j=1}^k v_j$ dos vectores arbitrarios, con $u_j, v_j \in W_j$, para todo j. Entonces como los subespacios W_j son ortogonales dos a dos, obtenemos

$$\langle P_i(u), v\rangle = \left\langle u_i, \sum_{j=1}^k v_j \right\rangle = \sum_{j=1}^k \langle u_i, v_j\rangle = \langle u_i, v_i\rangle; \quad \langle u, P_i(v)\rangle = \left\langle \sum_{j=1}^k u_j, v_i \right\rangle = \sum_{j=1}^k \langle u_j, v_i\rangle = \langle u_i, v_i\rangle.$$

Luego cada P_i es autoadjunto.

Supongamos ahora que tenemos operadores $P_1, \dots, P_k$ como en **(B)**. La proposición 7.5.2 implica que es $V = \bigoplus_{i=1}^k W_i$, siendo $W_i = \text{Im}(P_i)$, para todo i. Luego solo nos resta probar que es $W_i \perp W_j$, si $i \neq j$.

Sean $w_i \in W_i$ y $w_j \in W_j$, con $i \neq j$. Teniendo en cuenta que las P_l son autoadjuntas y verifican $P_l|_{W_l} = \text{Id}$, obtenemos

$$\langle w_i, w_j\rangle = \langle P_i(w_i), P_j(w_j)\rangle = \big\langle w_i, P_i\big(P_j(w_j)\big)\big\rangle = \langle w_i, (P_i \circ P_j)(w_j)\rangle = \langle w_i, 0\rangle = 0.$$

Luego W_i y W_j son ortogonales. □

Observación 3.2.5. Trabajando un poco se ve que la proposición 3.2.1 es un caso particular de la proposición 3.2.4, pero incluimos su prueba por ser más fácil de entender.

3.3. El adjunto de un operador

En la sección anterior definimos que un operador T es autoadjunto si verifica $\langle T(v), w\rangle = \langle v, T(w)\rangle$, para todo $v, w \in V$. El siguiente resultado generaliza esa idea.

Teorema 3.3.1. *Si $T \in \mathcal{L}(V)$, entonces existe un único $T^* \in \mathcal{L}(V)$ tal que*

$$\langle T(v), w\rangle = \langle v, T^*(w)\rangle, \quad \forall v, w \in V. \tag{3.2}$$

Dem. Vamos a razonar en forma similar a la prueba del teorema de Riesz. Sea $\mathcal{B} = \{v_1, \dots, v_n\}$ una base ortonormal de V. Supongamos que existe $T^* \in \mathcal{L}(V)$ que verifica (3.2) y sea $w \in V$ arbitrario. Entonces

$$T^*(w) = \sum_{i=1}^{n} \langle T^*(w), v_i\rangle\, v_i = \sum_{i=1}^{n} \overline{\langle v_i, T^*(w)\rangle}\, v_i = \sum_{i=1}^{n} \overline{\langle T(v_i), w\rangle}\, v_i = \sum_{i=1}^{n} \langle w, T(v_i)\rangle\, v_i.$$

Luego si existe $T^* \in \mathcal{L}(V)$ que verifica (3.2), entonces $T^*(w)$ está determinado por la fórmula anterior, para todo $w \in V$. Esto prueba la unicidad. Para la existencia, definamos $T^* : V \to V$ mediante

$$T^*(w) = \sum_{i=1}^{n} \langle w, T(v_i)\rangle\, v_i, \quad \forall w \in V. \tag{3.3}$$

La prueba de que T^* así definida es lineal, es simple y queda como ejercicio. Sean $v, w \in V$ arbitrarios. Aplicando las fórmulas (2.4) y (3.3) obtenemos

$$\langle v, T^*(w)\rangle = \left\langle \sum_{i=1}^{n} \langle v, v_i\rangle v_i, \sum_{j=1}^{n} \langle w, T(v_j)\rangle\, v_j \right\rangle = \sum_{i=j=1}^{n} \langle v, v_i\rangle \overline{\langle w, T(v_j)\rangle} \langle v_i, v_j\rangle = \sum_{i=1}^{n} \langle v, v_i\rangle \langle T(v_i), w\rangle;$$

$$\langle T(v), w\rangle = \left\langle T\left(\sum_{i=1}^{n} \langle v, v_i\rangle v_i\right), w \right\rangle = \left\langle \sum_{i=1}^{n} \langle v, v_i\rangle T(v_i), w \right\rangle = \sum_{i=1}^{n} \langle v, v_i\rangle \langle T(v_i), w\rangle.$$

Luego $\langle v, T^*(w)\rangle = \langle T(v), w\rangle$, para todo $v, w \in V$. □

Definición 3.3.2. El operador T^* se llama el *adjunto* del operador T.

Observación 3.3.3. La propiedad que caracteriza al adjunto vale también intercambiando los lugares de T y T^*. Es decir, si $T \in \mathcal{L}(V)$ entonces vale

$$\langle v, T(w)\rangle = \langle T^*(v), w\rangle, \quad \forall v, w \in V.$$

Esto se obtiene tomando conjugados en la fórmula (3.2).

Observación 3.3.4. Un operador T es autoadjunto si y solo si $T^* = T$ (de ahí viene el término “autoadjunto”).

A continuación veremos algunas propiedades de los adjuntos.

Proposición 3.3.5. *Sean $T, S \in \mathcal{L}(V)$ y $a \in \Bbbk$. Entonces*

$$(aT + S)^* = \bar{a}T^* + S^*, \qquad (T \circ S)^* = S^* \circ T^*, \qquad (T^*)^* = T, \qquad \mathrm{Id}^* = \mathrm{Id}.$$

Dem. Consideremos la primera igualdad.

$$\begin{aligned}\langle (aT + S)(v), w\rangle &= \langle aT(v) + S(v), w\rangle = a\langle T(v), w\rangle + \langle S(v), w\rangle = a\langle v, T^*(w)\rangle + \langle v, S^*(w)\rangle \\ &= \langle v, \overline{a}T^*(w) + S^*(w)\rangle = \langle v, (\overline{a}T^* + S^*)(w)\rangle, \quad \forall v, w \in V,\end{aligned}$$

luego la unicidad del adjunto implica $(aT+S)^* = \bar{a}T^* + S^*$. La prueba de la segunda afirmación es la siguiente

$$\langle (T\circ S)(v), w\rangle = \big\langle T\big(S(v)\big), w\big\rangle = \langle S(v), T^*(w)\rangle = \big\langle v, S^*\big(T^*(w)\big)\big\rangle = \langle v, (S^*\circ T^*)(w)\rangle\,, \quad \forall v,w\in V,$$

Luego $(T\circ S)^* = S^*\circ T^*$. Para la tercera, usando la observación 3.3.3 obtenemos

$$\langle T^*(v), w\rangle = \langle v, T(w)\rangle, \quad \forall v,w\in V,$$

luego $(T^*)^* = T$. La cuarta igualdad es inmediata. $\square$

Definición 3.3.6. Si $A = (a_{ij})\in M_n(\mathbb{C})$, definimos $\overline{A} := (\overline{a_{ij}})\in M_n(\mathbb{C})$ y $A^* := \overline{A}^t \in M_n(\mathbb{C})$, es decir, $A^* = (b_{ij})$, siendo $b_{ij} = \overline{a_{ji}}$. La matriz A^* se llama la *adjunta* de A. Notar que si $A\in M_n(\mathbb{C})$, entonces vale $A^* = A^t$ si y solo si $A\in M_n(\mathbb{R})$.

Proposición 3.3.7. *Si $T\in\mathcal{L}(V)$ y $\mathcal{B}$ es una base ortonormal de V entonces $[T^*]_{\mathcal{B}} = ([T]_{\mathcal{B}})^*$.*

Dem. Sea $[T]_{\mathcal{B}} = (a_{ij})$ y $[T^*]_{\mathcal{B}} = (c_{ij})$. Entonces $a_{ij} = \langle T(v_j), v_i\rangle = \langle v_j, T^*(v_i)\rangle = \overline{\langle T^*(v_i), v_j\rangle} = \overline{c_{ji}}$, $\square$

Corolario 3.3.8. *Si $A\in M_n(\Bbbk)$, entonces $(L_A)^* = L_{A^*}$ en $\mathcal{L}(\Bbbk^n)$, es decir vale*

$$\langle Ax, y\rangle = \langle x, A^*y\rangle, \quad \forall x,y\in\Bbbk^n,$$

considerando en $\Bbbk^n$ el producto interno usual.

Dem. La base canónica $\mathcal{B}$ es ortonormal respecto al producto interno usual y $[L_A]_{\mathcal{B}} = A$, luego $[(L_A)^*]_{\mathcal{B}} = A^*$, lo cual implica $(L_A)^* = L_{A^*}$. $\square$

Ejemplo 3.3.9. Consideremos $\mathbb{C}^2$ con el producto interno usual y $T:\mathbb{C}^2\to\mathbb{C}^2$ definido por $T(x,y) = (2ix+3y, x-y)$, $(x,y)\in\mathbb{C}^2$. Si $\mathcal{B}$ es la base canónica de $\mathbb{C}^2$, entonces

$$[T]_{\mathcal{B}} = \begin{pmatrix} 2i & 3\\ 1 & -1\end{pmatrix} \quad\Rightarrow\quad [T^*]_{\mathcal{B}} = ([T]_{\mathcal{B}})^* = \overline{\begin{pmatrix} 2i & 3\\ 1 & -1\end{pmatrix}^t} = \begin{pmatrix} -2i & 1\\ 3 & -1\end{pmatrix},$$

luego $T^*(x,y) = (-2ix+y, 3x-y)$.

De las proposiciones 3.3.5 y 3.3.7 se deduce directamente el siguiente resultado.

Corolario 3.3.10. *Sean $A,B\in M_n(\Bbbk)$ y $a\in\Bbbk$. Entonces*

$$(aA+B)^* = \bar{a}A^* + B^*, \qquad (AB)^* = B^*A^*, \qquad (A^*)^* = A, \qquad I^* = I. \quad \square$$

Proposición 3.3.11. *1. Si $T\in\mathcal{L}(V)$, entonces* $\det(T^*) = \overline{\det(T)}$.

2. Si $A\in M_n(\Bbbk)$, entonces $\det(A^*) = \overline{\det(A)}$.

Dem. Por la proposición 3.3.7 alcanza con probar la segunda afirmación.

$$\det(A^*) = \det\left(\overline{A}^t\right) = \det\left(\overline{A}\right) = \overline{\det(A)}.$$

En lo anterior usamos $\det\left(\overline{A}\right) = \overline{\det(A)}$, lo cual es fácil de probar. $\square$

El siguiente resultado relaciona los valores propios de un operador con los de su adjunto.

Proposición 3.3.12. *Sea $T\in\mathcal{L}(V)$. Si λ es un valor propio de T, entonces $\overline{\lambda}$ es un valor propio de T^*.*

Dem.

$$\chi_{T^*}\left(\overline{\lambda}\right) = \det\left(T^* - \overline{\lambda}\mathrm{Id}\right) = \det[(T-\lambda\mathrm{Id})^*] = \overline{\det(T-\lambda\mathrm{Id})} = \overline{\chi_T(\lambda)}.$$

Entonces si λ es valor propio de T, es $\chi_{T^*}(\overline{\lambda}) = \overline{\chi_T(\lambda)} = \overline{0} = 0$. Luego $\overline{\lambda}$ es un valor propio de T^*. $\square$

Observación 3.3.13. La proposición anterior dice que los valores propios de T^* son los conjugados de los valores propios de T, pero no dice nada sobre los vectores propios correspondientes, los cuales no suelen estar relacionados.

3.4. Operadores normales

Definición 3.4.1. Un operador $T \in \mathcal{L}(V)$ es *normal* si verifica $T \circ T^* = T^* \circ T$. Análogamente, decimos que una matriz $A \in M_n(\Bbbk)$ es *normal* si verifica $A\,A^* = A^*\,A$

La prueba del siguiente resultado es inmediata a partir de la proposición 3.3.7.

Proposición 3.4.2. *Si $T \in \mathcal{L}(V)$ y $\mathcal{B}$ es una base ortonormal de V, entonces T es normal si y solo si $[T]_{\mathcal{B}}$ es normal.* $\square$

Observación 3.4.3. Es claro que si $T \in \mathcal{L}(V)$ es autoadjunto, entonces T es normal. El siguiente ejemplo muestra que el recíproco es falso.

Ejemplo 3.4.4. Sea $0 < \theta < \pi$ y $A = \left(\begin{smallmatrix} \cos\theta & -\operatorname{sen}\theta \\ \operatorname{sen}\theta & \cos\theta \end{smallmatrix}\right) \in M_2(\mathbb{R})$. Es $AA^t = A^tA = \mathrm{Id}$, luego A es normal. Por otro lado es claro que A no es simétrica, luego la rotación $L_A : \mathbb{R}^2 \to \mathbb{R}^2$ es normal pero no es autoadjunto.

El siguiente resultado muestra que la normalidad es una condición necesaria para que un operador sea diagonalizable en una base ortonormal.

Proposición 3.4.5. *Si $T \in \mathcal{L}(V)$ es diagonalizable en una base ortonormal de V, entonces T es normal.*

Dem. Sea $\mathcal{B}$ base ortonormal de V tal que $[T]_{\mathcal{B}} = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$, en que esta última es una forma abreviada de escribir la matriz diagonal cuya diagonal principal es $(\lambda_1, \ldots, \lambda_n)$. La proposición 3.3.7 implica $[T^*]_{\mathcal{B}} = \operatorname{diag}\big(\overline{\lambda_1}, \ldots, \overline{\lambda_n}\big)$. Multiplicando estas matrices obtenemos

$$[T]_{\mathcal{B}}[T^*]_{\mathcal{B}} = [T^*]_{\mathcal{B}}[T]_{\mathcal{B}} = \operatorname{diag}\big(|\lambda_1|^2, \ldots, |\lambda_n|^2\big)\,.$$

En particular esto implica $T \circ T^* = T^* \circ T$. $\square$

Observación 3.4.6. En el caso real la proposición anterior no nos dice nada nuevo, ya que en ese caso sabemos que el operador es autoadjunto (teorema 3.1.17), y eso es más fuerte que ser normal. Lo interesante de la misma es que nos muestra que en el caso complejo, la normalidad es una condición necesaria para que un operador sea diagonalizable en una base ortonormal. A continuación veremos que vale también el recíproco. Para eso necesitamos probar algunas propiedades de los operadores normales, que son interesantes en sí mismas y valen tanto para espacios reales como complejos.

Proposición 3.4.7. *Sea $T \in \mathcal{L}(V)$ un operador normal. Entonces*

1. $||T(v)|| = ||T^*(v)||$, $\forall v \in V$.
2. *Si v es un vector propio de T correspondiente al valor propio λ, entonces v es un vector propio de T^* correspondiente al valor propio $\overline{\lambda}$.*
3. *Si $\lambda_1 \neq \lambda_2$ son valores propios de T con vectores propios correspondientes v_1 y v_2, entonces $v_1 \perp v_2$.*

Dem.

1. $||T(v)||^2 = \langle T(v), T(v)\rangle = \langle v, T^*(T(v))\rangle = \langle v, T(T^*(v))\rangle = \langle T^*(v), T^*(v)\rangle = ||T^*(v)||^2$.
2. Veamos primero que $T - \lambda\mathrm{Id}$ es normal, para todo $\lambda \in \Bbbk$.

$$\begin{aligned}(T - \lambda\mathrm{Id}) \circ (T - \lambda\mathrm{Id})^* &= (T - \lambda\mathrm{Id}) \circ (T^* - \overline{\lambda}\mathrm{Id}) = T \circ T^* - \lambda\, T^* - \overline{\lambda}\, T + |\lambda|^2\mathrm{Id} \\ &= T^* \circ T - \lambda\, T^* - \overline{\lambda}\, T + |\lambda|^2\mathrm{Id} = (T - \lambda\mathrm{Id})^* \circ (T - \lambda\mathrm{Id}).\end{aligned}$$

Luego aplicando la parte 1) a $T - \lambda\mathrm{Id}$, obtenemos que si $T(v) = \lambda v$, es

$$0 = ||T(v) - \lambda v|| = ||(T - \lambda\mathrm{Id})(v)|| = ||(T - \lambda\mathrm{Id})^*(v)|| = \left\|T^*(v) - \overline{\lambda} v\right\|.$$

Luego $T^*(v) = \overline{\lambda}\, v$.

3. La afirmación se deduce del cálculo siguiente

$$\lambda_1\langle v_1, v_2\rangle = \langle \lambda_1 v_1, v_2\rangle = \langle T(v_1), v_2\rangle = \langle v_1, T^*(v_2)\rangle = \langle v_1, \overline{\lambda_2} v_2\rangle = \lambda_2\langle v_1, v_2\rangle.$$

Como $\lambda_1 \neq \lambda_2$, resulta $\langle v_1, v_2\rangle = 0$. □

Lema 3.4.8. *Sea $T \in \mathcal{L}(V)$ y $W \subset V$ un subespacio que es T-invariante y T^*-invariante. Si consideramos la restricción $T|_W \in \mathcal{L}(W)$, entonces $(T|_W)^* = T^*|_W$. Si además T es normal, entonces también lo es $T|_W$.*

Dem. Si $w_1, w_2 \in W$, es

$$\langle T|_W(w_1), w_2\rangle = \langle T(w_1), w_2\rangle = \langle w_1, T^*(w_2)\rangle = \langle w_1, T^*|_W(w_2)\rangle, \ \forall w_1, w_2 \in W.$$

Luego $(T|_W)^* = T^*|_W$. Por lo tanto si T es normal, es

$$T|_W \circ (T|_W)^* = T|_W \circ T^*|_W = (T \circ T^*)|_W = (T^* \circ T)|_W = T^*|_W \circ T|_W = (T|_W)^* \circ T|_W. \quad \square$$

El siguiente resultado muestra que en el caso complejo vale el recíproco de la proposición 3.4.5. La prueba es muy similar a la de la proposición 3.1.15.

Proposición 3.4.9. *Si $\Bbbk = \mathbb{C}$ y $T \in \mathcal{L}(V)$ es normal, entonces T es diagonalizable en una base ortonormal*

Dem. La prueba es por inducción en $n = \dim V$.

Si $n = 1$, entonces tomando $v \in V$ con $\|v\| = 1$, obtenemos que $\{v\}$ es una base ortonormal de V formada por vectores propios de T.

Sea ahora $n > 1$ y supongamos razonando inductivamente que si tenemos un operador normal en un espacio de dimensión $n-1$, entonces existe una base ortonormal del espacio formada por vectores propios del operador.

Como $\Bbbk = \mathbb{C}$ existe λ valor propio de T. Sea v un vector propio de T correspondiente a λ y consideremos $W := [v]^\perp = \{w \in V \mid \langle w, v\rangle = 0\}$. Sabemos que W es un subespacio de V y que $\dim W = n - 1$. Sea $w \in W$,

$$\begin{aligned} \langle T(w), v\rangle &= \langle w, T^*(v)\rangle = \left\langle w, \overline{\lambda}\, v\right\rangle = \lambda\,\langle w, v\rangle = \lambda\, 0 = 0 \quad \Rightarrow \quad T(w) \in W, \\ \langle T^*(w), v\rangle &= \langle w, T(v)\rangle = \langle w, \lambda\, v\rangle = \overline{\lambda}\,\langle w, v\rangle = \overline{\lambda}\, 0 = 0 \quad \Rightarrow \quad T^*(w) \in W. \end{aligned}$$

Luego W es T y T^*-invariante, como T es normal, el lema anterior implica que $T|_W \in \mathcal{L}(W)$ es normal. Tenemos que $T|_W \in \mathcal{L}(W)$ normal y que $\dim W = n - 1$ entonces por la hipótesis inductiva existe $\{v_1, \dots, v_{n-1}\}$ base ortonormal de W formada por vectores propios de $T|_W$ (y por lo tanto vectores propios de T). Sea $v_n = \frac{v}{||v||}$, como $v_1, \dots, v_{n-1} \in W = [v]^\perp$ y $v_n \in [v]$ es $v_i \perp v_n$, para todo $i = 1, \dots, n-1$. Luego $\mathcal{B} = \{v_1, \dots, v_{n-1}, v_n\}$ es una base ortonormal de V formada por vectores propios de T. □

En resumen, podemos juntar los teoremas 3.1.17 y las proposiciones 3.4.5 y 3.4.9 en el siguiente resultado.

Teorema 3.4.10. *Un operador $T \in \mathcal{L}(V)$ es diagonalizable en una base ortonormal si y solo si T es normal en el caso complejo o autoadjunto en el caso real.* □

A continuación veremos un par de resultados sobre la relación entre normal y autoadjunto.

Corolario 3.4.11. *Sea $\Bbbk = \mathbb{C}$ y $T \in \mathcal{L}(V)$. Entonces T es autoadjunto si y solo si es normal y todos sus valores propios son reales.*

Dem. Lo único que hay que probar es el recíproco. Como T es normal, entonces existe una base ortonormal $\mathcal{B}$ de V formada por vectores propios de T. Además por hipótesis sabemos que los valores propios de T son reales, luego el teorema 3.1.17 implica que T es autoadjunto. $\square$

Corolario 3.4.12. *Sea* $\Bbbk = \mathbb{R}$ *y* $T \in \mathcal{L}(V)$. *Entonces* T *es autoadjunto si y solo si es normal y su polinomio característico se escinde en* $\mathbb{R}$.

Dem. Si T es autoadjunto, entonces T es diagonalizable y por lo tanto su polinomio característico se escinde en $\mathbb{R}$. Supongamos ahora que $T \in \mathcal{L}(V)$ es normal y su polinomio característico se escinde en $\mathbb{R}$. Sea $\mathcal{B} = \{v_1, \ldots, v_n\}$ una base ortonormal de V y $A = [T]_{\mathcal{B}}$. Como es $\chi_A(t) = \chi_T(t)$ y $\chi_T(t)$ se escinde en $\mathbb{R}$, entonces deducimos que los valores propios de A son reales. Además, como T es normal, entonces $A \in M_n(\mathbb{R})$ es una matriz normal. Luego $L_A \in \mathcal{L}(\mathbb{C}^n)$ es un operador normal cuyos valores propios son reales y por lo tanto el corolario anterior implica que L_A es autoadjunto. Luego $[T]_{\mathcal{B}} = A \in M_n(\mathbb{R})$ es hermitiana y por lo tanto simétrica. Como $\mathcal{B}$ es una base ortonormal, esto implica que T es autoadjunto. $\square$

Observación 3.4.13. El corolario anterior nos dice que si V es un espacio real y T es un operador en V que es normal pero no es autoadjunto, entonces T no es diagonalizable, dado que su polinomio característico necesariamente tiene raíces complejas que no son reales. Esto es lo que ocurre con las rotaciones en el plano.

3.5. Isometrías

Definición 3.5.1. Una *isometría* es una transformación lineal $T : V \to V$, que verifica

$$\langle T(u), T(v) \rangle = \langle u, v \rangle, \quad \forall u, v \in V.$$

A las isometrías también se les llama operadores *ortogonales* en el caso real o *unitarios* en el caso complejo.

Observación 3.5.2. La definición de isometría se puede hacer igual para una transformación lineal $T : V \to W$ entre dos espacios vectoriales con producto interno. Pero nosotros nos vamos a concentrar solo en los operadores en un espacio con producto interno de dimensión finita V.

Proposición 3.5.3. *Si* $T \in \mathcal{L}(V)$, *entonces* T *es una isometría si y solo si* T *preserva la norma*

$$\|T(v)\| = \|v\|, \quad \forall v \in V.$$

Dem. Si T es una isometría y $v \in V$, entonces $\|T(v)\| = \sqrt{\langle T(v), T(v) \rangle} = \sqrt{\langle v, v \rangle} = \|v\|$.

Recíprocamente, si T preserva la norma entonces de la linealidad de T y de las fórmulas de polarización

$$\langle u, v \rangle = \frac{1}{4}\left(\|u+v\|^2 - \|u-v\|^2\right), \ \Bbbk = \mathbb{R}; \quad \langle u, v \rangle = \frac{1}{4}\sum_{k=0}^{k=3} i^k \left\|u + i^k v\right\|^2, \quad \Bbbk = \mathbb{C}; \quad \forall u, v \in V,$$

se deduce que T preserva el producto interno. Por ejemplo, en el caso real es

$$\begin{aligned}\langle T(u), T(v) \rangle &= \frac{1}{4}\left(\|T(u)+T(v)\|^2 - \|T(u)-T(v)\|^2\right) = \frac{1}{4}\left(\|T(u+v)\|^2 - \|T(u-v)\|^2\right) \\ &= \frac{1}{4}\left(\|u+v\|^2 - \|u-v\|^2\right) = \langle u, v \rangle.\end{aligned}$$

El caso complejo es análogo. $\square$

Ejemplo 3.5.4. Sea $\theta \in \mathbb{R}$ y consideremos el operador $T = L_A : \mathbb{R}^2 \to \mathbb{R}^2$ con $A = \left(\begin{smallmatrix} \cos\theta & -\operatorname{sen}\theta \\ \operatorname{sen}\theta & \cos\theta \end{smallmatrix}\right)$. Entonces

$$\begin{aligned}\|T(x,y)\|^2 &= \big\|\big(x\cos(\theta) - y\operatorname{sen}(\theta), x\operatorname{sen}(\theta) + y\cos(\theta)\big)\big\|^2 \\ &= \big(x\cos(\theta) - y\operatorname{sen}(\theta)\big)^2 + \big(x\operatorname{sen}(\theta) + y\cos(\theta)\big)^2 \\ &= x^2\big(\cos^2(\theta) + \operatorname{sen}^2(\theta)\big) + y^2\big(\cos^2(\theta) + \operatorname{sen}^2(\theta)\big) = x^2 + y^2 = \|(x,y)\|^2.\end{aligned}$$

Luego $\|T(x,y)\| = \|(x,y)\|$, para todo $(x,y) \in \mathbb{R}^2$, y por lo tanto T es una isometría. Notar que T es la rotación en el plano $\mathbb{R}^2$ de centro en el origen y ángulo θ.

Proposición 3.5.5. *Sea $T \in \mathcal{L}(V)$. Las siguientes afirmaciones son equivalentes:*

1. *T es una isometría.*
2. *Para toda base ortonormal $\mathcal{B}$ de V, el conjunto $T(\mathcal{B})$ es una base ortonormal de V.*
3. *Existe una base ortonormal $\mathcal{B}$ de V tal que $T(\mathcal{B})$ es una base ortonormal de V.*

Dem. $(1 \Rightarrow 2)$. Sea $\mathcal{B} = \{v_1, \dots, v_n\}$ una base ortonormal de V, es $\langle T(v_i), T(v_j)\rangle = \langle v_i, v_j\rangle = \delta_{ij}$, luego $T(\mathcal{B}) = \{T(v_1), \dots, T(v_n)\}$ es un conjunto ortonormal (y por lo tanto es LI) con la misma cantidad de elementos que $\mathcal{B}$, luego $T(\mathcal{B})$ es una base ortonormal de V.

$(2 \Rightarrow 3)$. Esto es obvio, dado que V siempre admite una base ortonormal.

$(3 \Rightarrow 1)$. Sea $\mathcal{B} = \{w_1, \dots, w_n\}$ una base ortonormal de V tal que $T(\mathcal{B}) = \{T(w_1), \dots, T(w_n)\}$ es también una base ortonormal de V.

Sea $v \in V$. Como $\mathcal{B}$ es una base de V, entonces existen únicos $a_1, \dots, a_n \in \Bbbk$ tales que $v = \sum_{i=1}^n a_i\, w_i$, luego $T(v) = \sum_{i=1}^n a_i\, T(w_i)$. Como $\mathcal{B}$ es una base ortonormal es $\|v\|^2 = \sum_{i=1}^n |a_i|^2$ y como $T(\mathcal{B})$ es una base ortonormal es $\|T(v)\|^2 = \sum_{i=1}^n |a_i|^2$. Luego $\|T(v)\| = \|v\|$, para todo $v \in V$ y se aplica la proposición anterior. □

Notar que la proposición anterior prueba que toda isometría es un isomorfismo. El siguiente resultado completa esta relación.

Proposición 3.5.6. *Sea $T \in \mathcal{L}(V)$. Las siguientes afirmaciones son equivalentes*

$$T \text{ es una isometría}; \qquad T^* \circ T = \mathrm{Id}; \qquad T \circ T^* = \mathrm{Id}; \qquad T \text{ es un isomorfismo y } T^{-1} = T^*.$$

Dem. La equivalencia entre las tres últimas condiciones es una propiedad bien conocida de los operadores en espacios de dimensión finita. Probaremos la equivalencia entre la primera y la segunda. Observar que es

$$\langle T(u), T(v)\rangle = \langle u, T^*(T(v))\rangle = \langle u, (T^* \circ T)(v)\rangle, \quad \forall u, v \in V.$$

Luego T es una isometría si y solo si

$$\langle u, (T^* \circ T)(v)\rangle = \langle u, v\rangle, \quad \forall u, v \in V \quad \Leftrightarrow \quad (T^* \circ T)(v) = v, \quad \forall v \in V \quad \Leftrightarrow \quad T^* \circ T = \mathrm{Id}. \quad \square$$

Observación 3.5.7. Es claro que la función identidad Id_V es una isometría y es fácil de probar que la composición de isometrías es una isometría, y que la inversa de una isometría es también una isometría. Luego las isometrías de V en V forman un grupo, con la operación de composición.

Observaciones 3.5.8. Sea T una isometría.

1. Vimos que vale $T \circ T^* = T^* \circ T = \mathrm{Id}$; luego T es normal.

2. Si $\lambda \in \Bbbk$ es un valor propio de T, entonces $|\lambda| = 1$ ($\lambda = \pm 1$, si $\Bbbk = \mathbb{R}$):

$$\text{sea } 0 \neq v \in V : T(v) = \lambda\, v \quad \Rightarrow \quad \|v\| = \|T(v)\| = |\lambda|\,\|v\| \quad \Rightarrow \quad |\lambda| = 1.$$

3. Vale $|\det(T)| = 1$ ($\det T = \pm 1$, si $\Bbbk = \mathbb{R}$):

$$1 = \det(\mathrm{Id}) = \det(T \circ T^*) = \det(T)\,\det(T^*) = \det(T)\,\overline{\det(T)} = |\det(T)|^2 \quad \Rightarrow \quad |\det(T)| = 1.$$

4. Si $\Bbbk = \mathbb{C}$, como T es normal, entonces T es diagonalizable en una base ortonormal. Si $\Bbbk = \mathbb{R}$ esto no es necesariamente cierto. Por ejemplo, la rotación de centro en el origen y ángulo θ (con $\theta \neq k\pi$, $k \in \mathbb{Z}$) es una isometría y no es diagonalizable.

Proposición 3.5.9. *Sea $T \in \mathcal{L}(V)$.*

1. *Si $\Bbbk = \mathbb{C}$, entonces T es unitario si y solo si T es normal y $|\lambda| = 1$ para todo valor propio λ de T.*

2. *Si $\Bbbk = \mathbb{R}$, entonces T es ortogonal y diagonalizable si y solo si T es autoadjunto y $\lambda = \pm 1$ para todo valor propio λ de T.*

Dem. El directo de la primer parte es inmediato, y el de la segunda se deduce de que al ser ortogonal es normal, luego T es normal y diagonalizable lo cual implica autoadjunto (corolario 3.4.12). Probaremos ahora los recíprocos. En ambos casos la hipótesis implica que existe una base ortonormal $\mathcal{B} = \{v_1, \ldots, v_n\}$ de V formada por vectores propios de T. Luego $T(v_i) = \lambda_i v_i$, para todo i, siendo $\lambda_1, \ldots, \lambda_n$ los valores propios correspondientes. Notar que en ambos casos vale $|\lambda_i| = 1$, para todo i (siendo $|\ |$ el módulo en el caso complejo o el valor absoluto en el caso real). Luego

$$\langle T(v_i), T(v_j)\rangle = \langle \lambda_i v_i, \lambda_j v_j\rangle = \lambda_i \overline{\lambda_j}\langle v_i, v_j\rangle = \lambda_i\overline{\lambda_j} 0 = 0, \quad \forall i \neq j,$$
$$\|T(v_i)\| = \|\lambda_i v_i\| = |\lambda_i|\,\|v_i\| = 1, \quad \forall i.$$

Entonces $T(\mathcal{B})$ es una base ortonormal y por lo tanto T es una isometría. $\square$

Matrices ortogonales y unitarias.

Definición 3.5.10. Sea $\Bbbk$ un cuerpo arbitrario, una matriz $A \in M_n(\Bbbk)$ se dice *ortogonal* si $A^t A = A\,A^t = I$. Si $\Bbbk = \mathbb{C}$, una matriz $A \in M_n(\mathbb{C})$ se dice *unitaria* si $A^* A = A\,A^* = I$.

Observaciones 3.5.11.

1. La ortogonalidad se puede definir para matrices con coeficientes en un cuerpo cualquiera, mientras que el ser unitaria se define solo para matrices complejas. Si a una matriz real A la consideramos como matriz compleja, entonces es unitaria si y solo si es ortogonal, ya que en ese caso es $A^* = A^t$. Es claro que esto es falso para matrices complejas que no son reales.

2. Como una matriz cuadrada es invertible si y solo si es invertible solo por la derecha o la izquierda, se deduce que dada $A \in M_n(\Bbbk)$, las condiciones

$$A \text{ es ortogonal}; \quad A^t A = I; \quad A\,A^t = I; \quad A \text{ es invertible y } A^{-1} = A^t$$

son equivalentes (esto vale en todo cuerpo $\Bbbk$). Análogamente, si $A \in M_n(\mathbb{C})$, entonces las condiciones

$$A \text{ es unitaria}; \quad A^* A = I; \quad A\,A^* = I; \quad A \text{ es invertible y } A^{-1} = A^*$$

son equivalentes.

Proposición 3.5.12. *Sea $T \in \mathcal{L}(V)$, $\mathcal{B}$ una base ortonormal de V y $A = [T]_\mathcal{B}$. Entonces T es una isometría si y solo si A es ortogonal en el caso real o unitaria en el caso complejo.*

Dem. Sea $A = [T]_\mathcal{B}$. Como $\mathcal{B}$ es una base ortonormal, es $[T^*]_\mathcal{B} = A^*$. Luego $T \circ T^* = \mathrm{Id} \Leftrightarrow [T \circ T^*]_\mathcal{B} = [\mathrm{Id}]_\mathcal{B} \Leftrightarrow [T]_\mathcal{B}\,[T^*]_\mathcal{B} = I \Leftrightarrow A\,A^* = I$. $\square$

Usando la proposición 3.5.6 y las observaciones anteriores se obtiene directamente el siguiente resultado.

Corolario 3.5.13. *Sea $A \in M_n(\Bbbk)$ y consideramos el operador $L_A \in \mathcal{L}(\Bbbk^n)$. Entonces L_A es una isometría si y solo si A es ortogonal en el caso real o unitaria en el caso complejo.* $\square$

Proposición 3.5.14. *Sea $A \in M_n(\Bbbk)$. Las siguientes afirmaciones son equivalentes:*

1. *La matriz A es unitaria si $\Bbbk = \mathbb{C}$ o es ortogonal $\Bbbk = \mathbb{R}$.*
2. *Las filas de A forman una base ortonormal de $\Bbbk^n$.*
3. *Las columnas de A forman una base ortonormal de $\Bbbk^n$.*

Dem. Observar primero que si $A = (a_{ij}) \in M_n(\Bbbk)$, entonces vale

$$A^*A = I \;\Leftrightarrow\; \sum_{k=1}^{n} a_{ki}\overline{a_{kj}} = \delta_{ij}; \quad AA^* = I \;\Leftrightarrow\; \sum_{k=1}^{n} a_{ik}\overline{a_{jk}} = \delta_{ij}, \quad \forall i, j.$$

Luego $AA^* = I$ si y solo si las columnas de A son una base ortonormal de $\Bbbk^n$, y $A^*A = I$ si y solo si las filas de A son una base ortonormal de $\Bbbk^n$. $\square$

El siguiente resultado relaciona bases ortonormales con matrices ortogonales o unitarias.

Proposición 3.5.15. *Sean $\mathcal{B}$ y $\mathcal{C}$ bases de $\Bbbk^n$ y $A = {}_{\mathcal{B}}[\mathrm{Id}]_{\mathcal{C}}$. Entonces:*

1. *Si $\mathcal{B}$ y $\mathcal{C}$ son bases ortonormales, entonces A es unitaria si $\Bbbk = \mathbb{C}$, u ortogonal si $\Bbbk = \mathbb{R}$.*
2. *Si A es unitaria si $\Bbbk = \mathbb{C}$ u ortogonal si $\Bbbk = \mathbb{R}$ y $\mathcal{B}$ o $\mathcal{C}$ es una base ortonormal, entonces la otra base también es ortonormal.*

Dem. Probaremos que vale lo siguiente.

1. Si $\mathcal{B}$ es una base ortonormal, entonces $\mathcal{C}$ es una base ortonormal si y solo si $A^*A = I$.
2. Si $\mathcal{C}$ es una base ortonormal, entonces $\mathcal{B}$ es una base ortonormal si y solo si $AA^* = I$.

De estas dos afirmaciones se deduce claramente la tesis.

Sean $\mathcal{B} = \{v_1, \dots, v_n\}$, $\mathcal{C} = \{w_1, \dots, w_n\}$ y $A = {}_{\mathcal{B}}[\mathrm{Id}]_{\mathcal{C}} = (a_{ij})$. Luego $w_i = \sum_{k=1}^n a_{ki}v_k$, para todo $i = 1, \dots, n$. Supongamos que $\mathcal{B}$ es una base ortonormal. Si tomamos dos elementos de $\mathcal{C}$ $w_i = \sum_{k=1}^n a_{ki}v_k$ y $w_j = \sum_{k=1}^n a_{kj}v_k$, entonces como $\mathcal{B}$ ortonormal (recordar la observación 2.2.16) vale $\langle w_i, w_j\rangle = \sum_{k=1}^n a_{ki}\overline{a_{kj}}$. Luego $\mathcal{C}$ es una base ortonormal si y solo si $\sum_{k=1}^n a_{ki}\overline{a_{kj}} = \delta_{ij}$, para todo i, j, lo cual equivale a $A^*A = I$.

Por otro lado, si $\mathcal{C}$ es una base ortonormal, como es $A^{-1} = {}_{\mathcal{C}}[\mathrm{Id}]_{\mathcal{B}}$, deducimos aplicando lo anterior que $\mathcal{B}$ es una base ortonormal si y solo si $(A^{-1})^*A^{-1} = I$. Pero tomando inversos, obtenemos que esta última igualdad equivale a $AA^* = I$. Así $\mathcal{B}$ es una base ortonormal si y solo si $AA^* = I$. $\square$

Definición 3.5.16.

1. Dos matrices $A, B \in M_n(\mathbb{C})$ son *unitariamente equivalentes* si existe una matriz unitaria $Q \in M_n(\mathbb{C})$ tal que $A = QBQ^*$.
2. Dos matrices $A, B \in M_n(\Bbbk)$ ($\Bbbk$ cuerpo arbitrario) son *ortogonalmente equivalentes* si existe una matriz ortogonal $Q \in M_n(\Bbbk)$ tal que $A = QBQ^t$.

Proposición 3.5.17. *La relación de ser unitariamente equivalentes es de equivalencia en $M_n(\mathbb{C})$ y la de ser ortogonalmente equivalentes lo es en $M_n(\Bbbk)$ ($\Bbbk$ cuerpo arbitrario).*

Dem. Lo probaremos solo para el caso ortogonal, dejando el unitario (que es análogo) como ejercicio.

Notar que el ser ortogonalmente equivalentes es un caso particular de ser semejantes, ya que en ese caso vale $A = QBQ^t = QBQ^{-1}$. Recordando cómo es que se prueba que la semejanza es una relación de equivalencia, vemos lo único que nos falta probar es que la matriz identidad I es ortogonal (lo cual es obvio), que si Q es ortogonal, entonces Q^{-1} también lo es y que si Q_1, Q_2 son ortogonales, entonces Q_1Q_2 también lo es.

Es un ejercicio el probar que si $A \in M_n(\Bbbk)$ es invertible, entonces A^t es invertible y vale $(A^t)^{-1} = (A^{-1})^t$. Luego si Q es ortogonal, entonces

$$QQ^t = I \Rightarrow\ (QQ^t)^{-1} = I\ \Rightarrow\ (Q^t)^{-1}Q^{-1} = I\ \Rightarrow\ (Q^{-1})^tQ^{-1} = I.$$

Esto muestra que Q^{-1} es una matriz ortogonal. Además, si Q_1, Q_2 son ortogonales, entonces de

$$(Q_1Q_2)^tQ_1Q_2 = Q_2^tQ_1^tQ_1Q_2 = Q_2^tIQ_2 = Q_2^tQ_2 = I,$$

deducimos que Q_1Q_2 es ortogonal. □

Teorema 3.5.18. *Consideramos matrices en $M_n(\Bbbk)$.*

1. *Si $\Bbbk = \mathbb{C}$, entonces una matriz es normal si y solo si es unitariamente equivalente a una matriz diagonal.*

2. *Si $\Bbbk = \mathbb{R}$, entonces una matriz es simétrica si y solo si es ortogonalmente equivalente a una matriz diagonal.*

Dem.

1. ($\Rightarrow$): Sea $A \in M_n(\mathbb{C})$ y consideremos $L_A : \mathbb{C}^n \to \mathbb{C}^n$. Como A es normal, entonces L_A es normal, luego existe una base ortonormal $\mathcal{B}$ de $\mathbb{C}^n$ tal que $[L_A]_\mathcal{B} = D$, siendo D una matriz diagonal. Sea $\mathcal{C}$ la base canónica de $\mathbb{C}^n$ y $Q = {}_\mathcal{C}[\mathrm{Id}]_\mathcal{B}$. Entonces

   $$A = [L_A]_\mathcal{C} = {}_\mathcal{C}[\mathrm{Id}]_\mathcal{B}\, [L_A]_\mathcal{B}\, {}_\mathcal{B}[\mathrm{Id}]_\mathcal{C} = QDQ^{-1}.$$

   Como C y $\mathcal{B}$ son bases ortonormales de $\mathbb{C}^n$ y $Q = {}_\mathcal{C}[\mathrm{Id}]_\mathcal{B}$, entonces Q es unitaria y $A = QDQ^*$.

   ($\Leftarrow$): Sean Q y D en $M_n(\mathbb{C})$, con Q unitaria y D diagonal. Consideremos $A = QDQ^*$. Observar que vale $A^* = (QDQ^*)^* = Q^{**}\, D^*\, Q^* = Q\overline{D}Q^*$, luego

   $$A\,A^* = (QDQ^*)\big(Q\overline{D}Q^*\big) = QD\overline{D}Q^* \quad \text{y} \quad A^*A = \big(Q\overline{D}Q^*\big)(QDQ^*) = Q\overline{D}DQ^*.$$

   Como D es diagonal, vale $D\overline{D} = \overline{D}D$, luego la fórmula de arriba implica $AA^* = A^*A$.

2. ($\Rightarrow$): Es la misma idea que en 1. Si $A \in M_n(\mathbb{R})$ es simétrica, entonces $L_A \in \mathcal{L}(\mathbb{R}^n)$ es autoadjunta y por lo tanto es diagonalizable en una base ortonormal. El resto sigue igual.

   ($\Leftarrow$): Sea $A = QDQ^t$, con Q ortogonal y D diagonal. Es $A^t = \big(QDQ^t\big)^t = Q^{tt}D^tQ^t = QDQ^t = A$, luego $A^t = A$. □

Observación 3.5.19. Respecto al enunciado del teorema anterior, notar que si $A \in M_n(\mathbb{R})$ es ortogonalmente equivalente a una matriz diagonal D, entonces A y D son semejantes. Luego las entradas diagonales de D son los valores propios de A. Lo mismo ocurre en el caso complejo.

Ejemplo 3.5.20. Sea $A = \begin{pmatrix} 4 & 2 & 2 \\ 2 & 4 & 2 \\ 2 & 2 & 4 \end{pmatrix}$. Como A es una matriz simétrica real, sabemos que es ortogonalmente equivalente a una matriz diagonal. Vamos a hallar una matriz ortogonal Q y una matriz diagonal D tales que $A = QDQ^t$.

Es un ejercicio el verificar que vale $\chi_A(t) = -(t-2)^2(t-8)$ y que los subespacios propios de A son

$$E_2 = [(-1,1,0),(-1,0,1)], \quad E_8 = [(1,1,1)].$$

Aplicando Gram-Schmidt a la base $\{(-1,1,0),(-1,0,1)\}$ de E_2 obtenemos

$$E_2 = [(-1/2,1,-1/2),(-1,0,1)] = [(-1,2,-1),(-1,0,1)]\,.$$

Luego $\{(-1,2,-1),(-1,0,1),(1,1,1)\}$ es una base ortogonal de $\mathbb{R}^3$. Normalizando obtenemos

$$\left\{\left(-\frac{1}{\sqrt{6}},\frac{2}{\sqrt{6}},-\frac{1}{\sqrt{6}}\right),\left(-\frac{1}{\sqrt{2}},0,\frac{1}{\sqrt{2}}\right),\left(\frac{1}{\sqrt{3}},\frac{1}{\sqrt{3}},\frac{1}{\sqrt{3}}\right)\right\}$$

que es una base ortonormal de $\mathbb{R}^3$ formada por vectores propios de A.

Entonces si $D = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 8 \end{pmatrix}$ y $Q = \begin{pmatrix} -\frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}} \\ \frac{2}{\sqrt{6}} & 0 & \frac{1}{\sqrt{3}} \\ -\frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}} \end{pmatrix}$, es $A = QDQ^t$; es decir vale

$$\begin{pmatrix} 4 & 2 & 2 \\ 2 & 4 & 2 \\ 2 & 2 & 4 \end{pmatrix} = \begin{pmatrix} -\frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}} \\ \frac{2}{\sqrt{6}} & 0 & \frac{1}{\sqrt{3}} \\ -\frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}} \end{pmatrix} \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 8 \end{pmatrix} \begin{pmatrix} -\frac{1}{\sqrt{6}} & \frac{2}{\sqrt{6}} & -\frac{1}{\sqrt{6}} \\ -\frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} \end{pmatrix}.$$

Una aplicación del teorema anterior es el siguiente resultado.

Proposición 3.5.21. *Sean $A, B \in M_n(\mathbb{R})$ matrices simétricas. Si A y B son semejantes, entonces son ortogonalmente equivalentes.*

Dem. Si A y B son semejantes, entonces tienen los mismos valores propios $\lambda_1 \ldots, \lambda_n$ (pueden haber repetidos). Como A y B son simétricas, entonces como vimos en el teorema anterior vale que A y B son ortogonalmente equivalentes a la matriz diagonal $D = \operatorname{diag}(\lambda_1 \ldots, \lambda_n)$. Luego A y B son ortogonalmente equivalentes entre sí. $\square$

Observación 3.5.22. El resultado anterior es falso si se quita la condición de que las matrices sean simétricas. Por ejemplo, los valores propios de $A = \left(\begin{smallmatrix} 2 & 1 \\ 0 & 3 \end{smallmatrix}\right)$ son 2 y 3, luego A es semejante a $B = \left(\begin{smallmatrix} 2 & 0 \\ 0 & 3 \end{smallmatrix}\right)$. Sin embargo, como B es diagonal, entonces es fácil de probar que para toda matriz $Q \in M_2(\mathbb{R})$ vale que QBQ^t es una matriz simétrica. Como A no es simétrica, entonces A no puede ser ortogonalmente equivalente con B.

Capítulo 4

Formas bilineales simétricas

En este capítulo asumiremos siempre que los espacios son de dimensión finita y que las bases están ordenadas.

4.1. Formas bilineales

Sea $\Bbbk$ un cuerpo arbitrario y V un $\Bbbk$-espacio. Una *forma bilineal* en V es una función $\varphi : V \times V \to \Bbbk$ que es lineal en cada una de sus variables, es decir, que verifica

$$\varphi(au+v,w) = a\varphi(u,w) + \varphi(v,w), \quad \varphi(w,au+v) = a\varphi(w,u) + \varphi(w,v),$$

para todo $a \in \Bbbk$, $u,v,w \in V$. Al conjunto de las formas bilineales en V lo escribiremos $\operatorname{Bil}(V)$.

Es un ejercicio el verificar que el conjunto de las formas bilineales es un subespacio del espacio de las funciones con dominio $V \times V$ y codominio $\Bbbk$, en el cual las operaciones se definen punto a punto:

$$(f+g)(u,v) = f(u,v) + g(u,v), \quad (af)(u,v) = af(u,v), \quad \forall a \in \Bbbk,\ u,v \in V.$$

Luego $\operatorname{Bil}(V)$ con las operaciones anteriores es un $\Bbbk$-espacio vectorial.

Ejemplo 4.1.1. Un producto interno en un espacio vectorial real $\langle\ ,\ \rangle : V \times V \to \mathbb{R}$ es una forma bilineal.

Ejemplo 4.1.2. La función $\varphi : \Bbbk^2 \times \Bbbk^2 \to \Bbbk$ definida por

$$\varphi\big((x,y),(x',y')\big) = 2xx' + 3xy' + 4yx' - yy', \quad \forall (x,y),(x',y') \in \Bbbk^2,$$

es una forma bilineal.

Definición 4.1.3. Sea $A \in M_n(\Bbbk)$. Definimos $\beta_A : \Bbbk^n \times \Bbbk^n \to \Bbbk$ por $\beta_A(x,y) = x^t A y$, siendo $x,y \in \Bbbk^n$ vectores columna (es decir estamos pensando $\Bbbk^n = M_{n\times 1}(\Bbbk)$).

Notar que las propiedades del producto de matrices implican directamente que β_A es una forma bilineal. En coordenadas, si escribimos

$$x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \quad y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}, \quad A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix},$$

entonces

$$\beta_A\big((x_1,\cdots,x_n),(y_1,\cdots,y_n)\big) = \sum_{i,j=1}^{n} a_{ij}x_iy_j, \quad \forall (x_1,\cdots,x_n),(y_1,\cdots,y_n) \in \Bbbk^n. \tag{4.1}$$

El siguiente resultado nos permite construir nuevas formas bilineales a partir de una forma dada. La prueba es directa y queda como ejercicio.

Proposición 4.1.4. *Sea V un espacio vectorial y $\varphi : V \times V \to \Bbbk$ una forma bilineal. Sea W otro espacio y $T, S : W \to V$ dos transformaciones lineales. Si definimos $\psi : W \times W \to \Bbbk$ mediante*

$$\psi(w_1, w_2) = \varphi\big(T(w_1), S(w_2)\big), \quad \forall w_1, w_2 \in W,$$

Entonces ψ es una forma bilineal en W. □

De ahora en adelante V es un espacio de dimensión n.

Observación 4.1.5. Sea $\mathcal{B}$ una base de V y $A \in M_n(\Bbbk)$. El mapa coordenadas $\operatorname{coord}_{\mathcal{B}} : V \to \Bbbk^n$ es un isomorfismo lineal, luego aplicando la proposición anterior a $\varphi = \beta_A \in \operatorname{Bil}(\Bbbk^n)$ obtenemos que el mapa $\psi : V \times V \to \Bbbk$ definido por

$$\psi(u, v) = \operatorname{coord}_{\mathcal{B}}(u)^t A \operatorname{coord}_{\mathcal{B}}(v), \quad \forall u, v \in V,$$

es una forma bilineal en V. A continuación probaremos que toda forma bilineal en V se puede escribir de esa manera. Para eso necesitamos la siguiente definición.

Definición 4.1.6. Sea $\varphi \in \operatorname{Bil}(V)$ y $\mathcal{B} = \{v_1, \ldots, v_n\}$ una base de V. La *matriz asociada* a la forma bilineal φ en la base $\mathcal{B}$ es la matriz $M_{\mathcal{B}}(\varphi) \in M_n(\Bbbk)$ definida por

$$M_{\mathcal{B}}(\varphi) := \begin{pmatrix} \varphi(v_1, v_1) & \cdots & \varphi(v_1, v_n) \\ \vdots & & \vdots \\ \varphi(v_n, v_1) & \cdots & \varphi(v_n, v_n) \end{pmatrix}.$$

Llamaremos *representación matricial* de una forma bilineal φ toda matriz asociada a φ en alguna base del espacio. Notar que si $\varphi = \beta_A \in \operatorname{Bil}(\Bbbk^n)$, entonces $A = M_{\mathcal{C}}(\varphi)$, siendo $\mathcal{C}$ la base canónica de $\Bbbk^n$.

Proposición 4.1.7. *Sea $\varphi \in \operatorname{Bil}(V)$ y $\mathcal{B}$ una base de V. Entonces*

$$\varphi(u, v) = \operatorname{coord}_{\mathcal{B}}(u)^t M_{\mathcal{B}}(\varphi) \operatorname{coord}_{\mathcal{B}}(v), \quad \forall u, v \in V.$$

Además, si una matriz $A \in M_n(\Bbbk)$ verifica

$$\varphi(u, v) = \operatorname{coord}_{\mathcal{B}}(u)^t A \operatorname{coord}_{\mathcal{B}}(v), \quad \forall u, v \in V, \tag{4.2}$$

entonces $A = M_{\mathcal{B}}(\varphi)$.

Dem. Sean $u = \sum_{i=1}^n x_i v_i$ y $v = \sum_{j=1}^n y_j v_j$. Entonces por la bilinealidad de φ es

$$\varphi(u, v) = \varphi\left(\sum_{i=1}^n x_i v_i, \sum_{j=1}^n y_j\, v_j\right) = \sum_{i,j=1}^n x_i y_j \varphi(v_i, v_j) = \operatorname{coord}_{\mathcal{B}}(u)^t M_{\mathcal{B}}(\varphi) \operatorname{coord}_{\mathcal{B}}(v), \quad \forall u, v \in V.$$

Esto prueba la primera afirmación. Para la segunda, si $A = (a_{ij}) \in M_n(\Bbbk)$ verifica la fórmula (4.2) y $\mathcal{B} = \{v_1, \ldots, v_n\}$, entonces

$$\varphi(v_i, v_j) = \operatorname{coord}_{\mathcal{B}}(v_i)^t A \operatorname{coord}_{\mathcal{B}}(v_j) = e_i^t A e_j = a_{ij}, \quad \forall i, j,$$

siendo $\{e_1, \ldots, e_n\}$ la base canónica de $\Bbbk^n$. Luego $A = M_{\mathcal{B}}(\varphi)$. □

Corolario 4.1.8. *Si* $\varphi \in \operatorname{Bil}(\Bbbk^n)$ *entonces existe una única matriz* $A \in M_n(\Bbbk)$ *tal que* $\varphi = \beta_A$.

Dem. Sea $\mathcal{C}$ la base canónica de $\Bbbk^n$ y $A = M_{\mathcal{C}}(\varphi)$. Entonces

$$\varphi(u,v) = \operatorname{coord}_{\mathcal{C}}(u)^t M_{\mathcal{C}}(\varphi) \operatorname{coord}_{\mathcal{C}}(v) = u^t A v = \beta_A(u,v), \quad \forall u, v \in V.$$

La unicidad de A se deduce de la unicidad en la proposición 4.1.7. $\square$

Ejemplo 4.1.9. Si consideramos la forma bilineal $\varphi : \Bbbk^2 \times \Bbbk^2 \to \Bbbk$ definida en el ejemplo 4.1.2, es

$$\varphi\big((x,y),(x',y')\big) = 2xx' + 3xy' + 4yx' - yy' = (x,y)\begin{pmatrix} 2 & 3 \\ 4 & -1 \end{pmatrix}\begin{pmatrix} x' \\ y' \end{pmatrix}.$$

Luego $\varphi = \beta_A$, siendo $A = \left(\begin{smallmatrix} 2 & 3 \\ 4 & -1 \end{smallmatrix}\right)$.

A partir de la proposición anterior y observando que la correspondencia $\varphi \mapsto M_{\mathcal{B}}(\varphi)$ ($\mathcal{B}$ base fija de V) es lineal, se deduce inmediatamente el siguiente resultado.

Corolario 4.1.10. *Si* $\mathcal{B}$ *es una base de* V*, entonces la correspondencia que a cada forma bilineal* φ *le asocia la matriz* $M_{\mathcal{B}}(\varphi)$*, es un isomorfismo del espacio de las formas bilineales* $\operatorname{Bil}(V)$ *en el de las matrices* $M_n(\Bbbk)$*. Luego* $\dim \operatorname{Bil}(V) = n^2$. $\square$

El siguiente resultado relaciona las matrices asociadas a una misma forma bilineal en bases distintas.

Proposición 4.1.11. *Sean* $\mathcal{B}$ *y* $\mathcal{C}$ *dos bases de* V *y* $\varphi \in \operatorname{Bil}(V)$*. Entonces*

$$M_{\mathcal{C}}(\varphi) = \big({}_{\mathcal{B}}[\mathrm{Id}]_{\mathcal{C}}\big)^t M_{\mathcal{B}}(\varphi)\, {}_{\mathcal{B}}[\mathrm{Id}]_{\mathcal{C}}.$$

Dem. Sean $u, v \in V$, entonces usando la fórmula de cambio de coordenadas de $\mathcal{C}$ a $\mathcal{B}$ obtenemos

$$\begin{aligned} \varphi(u,v) &= \operatorname{coord}_{\mathcal{B}}(u)^t M_{\mathcal{B}}(\varphi) \operatorname{coord}_{\mathcal{B}}(v) = \big({}_{\mathcal{B}}[\mathrm{Id}]_{\mathcal{C}} \operatorname{coord}_{\mathcal{C}}(u)\big)^t M_{\mathcal{B}}(\varphi)\big({}_{\mathcal{B}}[\mathrm{Id}]_{\mathcal{C}} \operatorname{coord}_{\mathcal{C}}(v)\big) \\ &= \operatorname{coord}_{\mathcal{C}}(u)^t \big({}_{\mathcal{B}}[\mathrm{Id}]_{\mathcal{C}}\big)^t M_{\mathcal{B}}(\varphi)\, {}_{\mathcal{B}}[\mathrm{Id}]_{\mathcal{C}} \operatorname{coord}_{\mathcal{C}}(v). \end{aligned}$$

Como esto vale para todo $u, v \in V$, entonces la tesis se deduce aplicando la proposición 4.1.7. $\square$

Definición 4.1.12. Dos matrices A y B en $M_n(\Bbbk)$ se dicen *congruentes* si existe una matriz invertible Q en $M_n(\Bbbk)$ tal que $A = Q^t B Q$.

Ejercicio 4.1.13. Probar que la congruencia es una relación de equivalencia en $M_n(\Bbbk)$.

Proposición 4.1.14. *Sea* $\varphi \in \operatorname{Bil}(V)$ *y* $\mathcal{B}$ *una base de* V*. Si* $A \in M_n(\Bbbk)$ *es congruente con* $M_{\mathcal{B}}(\varphi)$*, entonces existe* $\mathcal{C}$ *base de* V *tal que* $M_{\mathcal{C}}(\varphi) = A$.

Dem. Sea $Q = (q_{ij}) \in M_n(\Bbbk)$ invertible tal que $A = Q^t M_{\mathcal{B}}(\varphi)\, Q$. Si $\mathcal{B} = \{v_1, \dots, v_n\}$, definimos $\mathcal{C} = \{w_1, \dots, w_n\}$ mediante

$$w_j = \sum_{i=1}^n q_{ij} v_i, \quad j = 1, \dots, n.$$

Como Q es invertible, resulta que $\mathcal{C}$ es una base de V y ${}_{\mathcal{B}}[\mathrm{Id}]_{\mathcal{C}} = Q$. Luego

$$A = Q^t M_{\mathcal{B}}(\varphi)\, Q = \big({}_{\mathcal{B}}[\mathrm{Id}]_{\mathcal{C}}\big)^t M_{\mathcal{B}}(\varphi)\, {}_{\mathcal{B}}[\mathrm{Id}]_{\mathcal{C}} = M_{\mathcal{C}}(\varphi). \quad \square$$

Corolario 4.1.15. *Dos matrices son congruentes si y solo si representan a una misma forma bilineal.*

Dem. Sean A y A' en $M_n(\Bbbk)$ dos matrices congruentes. Consideramos la forma bilineal $\beta_{A'} \in \mathrm{Bil}(\Bbbk^n)$. Si $\mathcal{B}$ es la base canónica de $\Bbbk^n$, entonces es $A' = M_{\mathcal{B}}(\beta_{A'})$. Luego A es congruente con $M_{\mathcal{B}}(\beta_{A'})$ y la proposición anterior implica que es $A = M_{\mathcal{C}}(\beta_{A'})$ para alguna base $\mathcal{C}$ de $\Bbbk^n$. El recíproco es la Proposición 4.1.11. □

Observaciones 4.1.16.

1. Si $A, B \in M_n(\Bbbk)$ y $Q \in M_n(\Bbbk)$ es una matriz invertible tal que $A = Q^t BQ$, entonces $A = PBP^t$, siendo $P = Q^t$ una matriz invertible, y es claro que estas dos condiciones son equivalentes. Luego tenemos dos formas equivalentes de definir la congruencia (lo mismo sucede con la semejanza). Nosotros usaremos siempre la primera porque se adapta mejor al trabajo con formas bilineales.

2. Hay dos relaciones de equivalencia en $M_n(\Bbbk)$ que son similares y pueden dar lugar a confusión. Sean $A, B \in M_n(\Bbbk)$. Por un lado, usando la observación anterior podemos decir que A y B son congruentes si existe una matriz invertible $Q \in M_n(\Bbbk)$ tal que $A = QBQ^t$. Pero por otro lado decimos que A y B son ortogonalmente equivalentes si existe una matriz ortogonal $P \in M_n(\Bbbk)$ tal que $A = PBP^t$ (recordar que P es ortogonal si es invertible y $P^{-1} = P^t$). Luego dos matrices ortogonalmente equivalentes son también congruentes, pero el recíproco no es necesariamente cierto. De hecho si A y B son ortogonalmente equivalentes por medio de una matriz ortogonal P, entonces A y B son congruentes ($A = PBP^t$) y también son semejantes ($A = PBP^{-1}$), pero la congruencia y la semejanza son relaciones distintas. El ser ortogonalmente equivalentes es una especie de "cruce de caminos" entre la congruencia y la semejanza.

4.2. Formas bilineales simétricas

Decimos que un cuerpo $\Bbbk$ tiene *característica* 2 y escribimos $\operatorname{car}\Bbbk = 2$ si en $\Bbbk$ se verifica $1 + 1 = 0$. Un ejemplo es el conjunto $\mathbb{F}_2$ formado por dos elementos que llamamos 0 y 1, en el cual definimos una suma y un producto mediante la siguiente tabla

$$\begin{array}{cccc} 0+0=0, & 1+1=0, & 0+1=1, & 1+0=1, \\ 0\cdot 0=0, & 0\cdot 1=0, & 1\cdot 0=0, & 1\cdot 1=1. \end{array}$$

Es un ejercicio el verificar que $\mathbb{F}_2$ con estas operaciones es un cuerpo y claramente $\operatorname{car}\mathbb{F}_2 = 2$.

Ejemplos de cuerpos con característica distinta de 2 son $\mathbb{Q}$, $\mathbb{R}$ y $\mathbb{C}$. En un cuerpo $\Bbbk$ con $\operatorname{car}\Bbbk \neq 2$ es $2 := 1 + 1 \neq 0$ y luego 2 es invertible en $\Bbbk$.

En este capítulo, de ahora en adelante V es un $\Bbbk$-espacio de dimensión finita y $\operatorname{car}\Bbbk \neq 2$.

Definición 4.2.1. Una forma bilineal φ se dice *simétrica* si verifica

$$\varphi(u, v) = \varphi(v, u), \quad \forall u, v \in V.$$

Ejemplo 4.2.2. Todo producto interno real $\langle\ ,\ \rangle : V \times V \to \mathbb{R}$ es una forma bilineal simétrica.

Proposición 4.2.3. *Sea $\varphi \in \mathrm{Bil}(V)$. Las siguientes afirmaciones son equivalentes.*

1. *La forma φ es simétrica.*
2. *Para toda base $\mathcal{B}$ de V, la matriz $M_{\mathcal{B}}(\varphi)$ es simétrica.*
3. *Existe una base $\mathcal{B}$ de V tal que la matriz $M_{\mathcal{B}}(\varphi)$ es simétrica.*

Dem. $1 \Rightarrow 2$: Sea $\mathcal{B} = \{v_1, \ldots, v_n\}$ una base de V y $M_{\mathcal{B}}(\varphi) = (a_{ij})$, entonces

$$a_{ij} = \varphi(v_i, v_j) = \varphi(v_j, v_i) = a_{ji}, \quad \forall i, j = 1, \ldots, n$$

luego $M_{\mathcal{B}}(\varphi) = (a_{ij})$ es simétrica.

$2 \Rightarrow 3$: Esto es obvio (todo espacio vectorial tiene una base).

$3 \Rightarrow 1$: Sea $\mathcal{B} = \{v_1, \dots, v_n\}$ una base de V tal que $M_{\mathcal{B}}(\varphi)$ es simétrica. Notar que esto último equivale a $\varphi(v_i, v_j) = \varphi(v_j, v_i)$, $\forall i, j$. Sean $u, v \in V$ arbitrarios, entonces existen escalares $a_i, b_i \in \Bbbk$, $i = 1, \dots, n$ tales que $u = \sum_{i=1}^n a_i v_i$ y $v = \sum_{i=1}^n b_i v_i$. Luego

$$\begin{aligned} \varphi(u, v) &= \varphi\left(\sum_{i=1}^n a_i v_i, \sum_{j=1}^n b_j v_j\right) = \sum_{i=1}^n a_i \sum_{j=1}^n b_j \varphi(v_i, v_j) = \sum_{i=1}^n a_i \sum_{j=1}^n b_j \varphi(v_j, v_i) \\ &= \varphi\left(\sum_{j=1}^n b_j v_j, \sum_{i=1}^n a_i v_i\right) = \varphi(v, u). \quad \square \end{aligned}$$

Al conjunto de las formas bilineales simétricas en V lo escribiremos $\mathrm{Bil}_S(V)$.

Observación 4.2.4. Es un ejercicio el verificar que $\mathrm{Bil}_S(V)$ es un subespacio de $\mathrm{Bil}(V)$, luego $\mathrm{Bil}_S(V)$ es un espacio vectorial. De la proposición anterior se deduce que si $\dim(V) = n$, entonces $\dim \mathrm{Bil}_S(V) = \frac{n(n+1)}{2}$.

Ejemplo 4.2.5. Como la matriz identidad $I \in M_n(\Bbbk)$ es simétrica, entonces $\beta_I \in \mathrm{Bil}(\Bbbk^n)$ es una forma bilineal simétrica. Explícitamente β_I se escribe

$$\beta_I\big((x_1, \dots, x_n), (y_1, \dots, y_n)\big) = x_1 y_1 + \dots + x_n y_n.$$

Notar que si $\Bbbk = \mathbb{C}$, entonces β_I es una forma bilineal simétrica pero no es un producto interno.

Ejemplo 4.2.6. Consideramos en $\mathbb{R}^2$ las formas bilineales φ y ψ definidas por

$$\varphi\big((x, y), (x', y')\big) = 2xx' + 3xy' + 3yx' - 5yy', \quad \psi\big((x, y), (x', y')\big) = xx' - 3xy' + 2yx' - 4yy'.$$

Observar que podemos escribir

$$\varphi\big((x, y), (x', y')\big) = \begin{pmatrix} x & y \end{pmatrix} \begin{pmatrix} 2 & 3 \\ 3 & -5 \end{pmatrix} \begin{pmatrix} x' \\ y' \end{pmatrix}, \quad \psi\big((x, y), (x', y')\big) = \begin{pmatrix} x & y \end{pmatrix} \begin{pmatrix} 1 & -3 \\ 2 & -4 \end{pmatrix} \begin{pmatrix} x' \\ y' \end{pmatrix}.$$

Luego observando las matrices asociadas deducimos que φ es simétrica y que ψ no lo es.

Formas cuadráticas.

Definición 4.2.7. Si $\varphi : V \times V \to \Bbbk$ es una forma bilineal simétrica, entonces la función $\Phi : V \to \Bbbk$ definida por $\Phi(v) = \varphi(v, v)$ para todo $v \in V$, se llama la *forma cuadrática* asociada a φ.

Ejemplo 4.2.8. La forma cuadrática Φ asociada a la forma bilineal simétrica φ del ejemplo anterior, está definida por

$$\Phi(x, y) = 2x^2 + 6xy - 5y^2, \quad \forall (x, y) \in \mathbb{R}^2.$$

Ejemplo 4.2.9. Si V es un espacio real y $\varphi = \langle\ ,\ \rangle$ es un producto interno en V, entonces su forma cuadrática asociada Φ es el cuadrado de la norma: $\Phi(v) = \|v\|^2$, para todo $v \in V$.

Observación 4.2.10. Como las formas bilineales simétricas en V forman un subespacio del espacio de las funciones de $V \times V$ en $\Bbbk$, entonces se prueba fácilmente que las formas cuadráticas en V forman un subespacio del espacio de las funciones de V en $\Bbbk$.

Proposición 4.2.11. *Sea $\varphi \in \mathrm{Bil}_S(V)$ y $\Phi : V \to \Bbbk$ la forma cuadrática asociada a φ. Entonces*

1. $\Phi(av) = a^2\Phi(v)$, *para todo* $a \in \Bbbk$, $v \in V$. *Luego* $\Phi(0) = 0$.

2. $\Phi(u+v) = \Phi(u) + 2\varphi(u,v) + \Phi(v)$, *para todo* $u, v \in V$.

Dem. La prueba de la primera afirmación es simple y queda como ejercicio. La segunda se deduce del cálculo siguiente

$$\Phi(u+v) = \varphi(u+v, u+v) = \varphi(u,u) + \varphi(u,v) + \varphi(v,u) + \varphi(v,v) = \Phi(u) + 2\varphi(u,v) + \Phi(v). \quad \square$$

Observación 4.2.12. Despejando[1] $\varphi(u,v)$ en la fórmula de la segunda afirmación de la proposición anterior, obtenemos la *fórmula de polarización*

$$\varphi(u,v) = \frac{1}{2}\big(\Phi(u+v) - \Phi(u) - \Phi(v)\big), \quad \forall u, v \in V.$$

Luego toda forma bilineal simétrica queda determinada por su forma cuadrática asociada. En este sentido, dada una forma cuadrática Φ, diremos que la correspondiente forma bilineal simétrica φ es la *forma bilineal asociada* a Φ.

Usando la fórmula de polarización se deduce inmediatamente el siguiente resultado.

Proposición 4.2.13. *Una forma bilineal simétrica es nula si y solo si lo es su forma cuadrática asociada.* $\square$

Nuestro próximo objetivo es caracterizar las formas cuadráticas en $\Bbbk^n$.

Definición 4.2.14. Consideremos $\Bbbk[x_1, \ldots, x_n]$ el espacio de los polinomios en n variables. Un polinomio $p \in \Bbbk[x_1, \ldots, x_n]$ se dice que es *homogéneo de grado 2* si es de la forma $p = \sum_{1 \le i \le j \le n} a_{ij} x_i x_j$, siendo los a_{ij} escalares arbitrarios.

Es fácil de probar que los polinomios homogéneos de grado 2 forman un subespacio de $\Bbbk[x_1, \ldots, x_n]$.

Ejemplo 4.2.15. Si $n = 2$, entonces un polinomio homogéneo de grado 2 en las variables x, y es un polinomio de la forma

$$ax^2 + bxy + cy^2, \quad a, b, c \in \Bbbk.$$

Si $n = 3$, un polinomio homogéneo de grado 2 en las variables x, y, z es un polinomio de la forma

$$ax^2 + by^2 + cz^2 + dxy + exz + fyz, \quad a, b, \ldots, f \in \Bbbk.$$

Proposición 4.2.16. *Existe una correspondencia uno a uno entre*

1. las matrices simétricas $n \times n$ *con coeficientes en* $\Bbbk$,

2. las formas bilineales simétricas en $\Bbbk^n$,

3. las formas cuadráticas en $\Bbbk^n$,

4. los polinomios homogéneos de grado 2 en n *variables con coeficientes en* $\Bbbk$.

Dem. La correspondencia entre formas bilineales simétricas y matrices simétricas está dada por asociar a una matriz simétrica A la forma β_A, y asociar a una forma bilineal simétrica φ la matriz asociada $M_{\mathcal{C}}(\varphi)$, siendo $\mathcal{C}$ la base canónica de $\Bbbk^n$. Estas correspondencias son inversas una de la otra (proposición 4.1.8).

[1] Para poder hacer este despeje es necesaria la condición $\operatorname{car} \Bbbk \neq 2$.

La correspondencia entre formas bilineales simétricas y formas cuadráticas vale en general y está dada por las fórmulas siguientes

$$\Phi(v) = \varphi(v, v), \quad \varphi(u, v) = \frac{1}{2}\big(\Phi(u + v) - \Phi(u) - \Phi(v)\big), \quad \forall u, v \in V$$

En el caso particular de $\Bbbk^n$, si $\varphi = \beta_A$, siendo $A = (a_{ij}) \in M_n(\Bbbk)$ una matriz simétrica, entonces a φ le corresponde la forma cuadrática Φ definida por

$$\Phi(x_1, \ldots, x_n) = \sum_{i=j=1}^{n} a_{ij}x_ix_j = \sum_{i\le j} b_{ij}x_ix_j, \quad b_{ij} := \begin{cases} a_{ii} & \text{si } i = j, \\ 2a_{ij} & \text{si } i < j. \end{cases}$$

Es claro que este proceso podemos revertirlo, definiendo los a_{ij} a partir de los b_{ij} mediante $a_{ii} := b_{ii}$, para todo i y $a_{ij} = a_{ji} := \frac{1}{2}b_{ij}$, si $i < j$. Luego toda función $\Phi : \Bbbk^n \to \Bbbk$ definida por una fórmula del tipo $\Phi(x_1, \ldots, x_n) = \sum_{i\le j} b_{ij}x_ix_j$ es una forma cuadrática en $\Bbbk^n$. Finalmente lo único que resta es observar que $p = \sum_{i\le j} b_{ij}x_ix_j$ es un polinomio homogéneo. Luego la fórmula que define a una forma cuadrática en $\Bbbk^n$ está dada por un polinomio homogéneo. □

Observación 4.2.17. En el caso de $\Bbbk^2$, si la matriz simétrica es $A = \left(\begin{smallmatrix} a & b \\ b & c \end{smallmatrix}\right)$, entonces la forma bilineal simétrica φ, la forma cuadrática Φ y el polinomio homogéneo p que corresponden a A son

$$\varphi\big((x, y), (x', y')\big) = axx' + bxy' + bx'y + cyy', \quad \Phi(x, y) = ax^2 + 2bxy + cy^2, \quad p = ax^2 + 2bxy + cy^2.$$

Notar que la única diferencia entre Φ y p, es que Φ es una función de $\Bbbk \times \Bbbk$ en $\Bbbk$, mientras que p es un polinomio en dos variables.

Ejemplo 4.2.18. Si $\Phi : \mathbb{R}^2 \to \mathbb{R}$ es la forma cuadrática definida por $\Phi(x, y) = x^2 + 6xy + 5y^2$, para todo $(x, y) \in \mathbb{R}^2$, entonces la forma bilineal asociada φ es

$$\varphi\big((x, y), (x', y')\big) = xx' + 3xy' + 3yx' + 5yy', \quad \forall (x, y), (x', y') \in \mathbb{R}^2.$$

Radical, degeneramiento y rango. Si consideramos la forma $\varphi \in \mathrm{Bil}_S(\mathbb{R}^3)$ definida por

$$\varphi\big((x, y, z), (x', y', z')\big) = xx' + yy',$$

entonces el vector $(0, 0, 1)$ es no nulo y verifica $\varphi\big((x, y, z), (0, 0, 1)\big) = 0$, para todo $(x, y, z) \in \mathbb{R}^3$. Esto, que no sucede en los productos internos, da lugar a la siguiente definición.

Definición 4.2.19. Dos vectores $u, v \in V$ son *φ-ortogonales* si $\varphi(u, v) = 0$. Si $\varphi \in \mathrm{Bil}_S(V)$, entonces su *radical*[2] es el conjunto $\mathrm{Rad}\,(\varphi)$ formado por los vectores que son φ-ortogonales a todos los vectores de V. Luego

$$\mathrm{Rad}\,(\varphi) = \{v \in V : \ \varphi(v, u) = 0, \ \forall u \in V\} = \{v \in V : \ \varphi(u, v) = 0, \ \forall u \in V\}.$$

Es fácil de probar que $\mathrm{Rad}\,(\varphi)$ es un subespacio de V. Decimos que φ es *no degenerada* si $\mathrm{Rad}\,(\varphi) = \{0\}$, es decir si

$$\varphi(u, v) = 0, \ \forall v \in V \quad \Rightarrow \quad u = 0.$$

Si W es un subespacio de V, decimos que φ es *no degenerada en* W si la restricción de φ a W es no degenerada.

Observación 4.2.20. Si $\varphi \in \mathrm{Bil}_S(V)$ y $\mathcal{B} = \{v_1, \ldots, v_n\}$ es un conjunto generador de V, entonces la bilinealidad de φ implica que un vector está en el radical de φ si y solo si es φ-ortogonal a $v_1, \ldots, v_n$. Luego

$$\mathrm{Rad}\,(\varphi) = \{v \in V : \ \varphi(v, v_i) = 0, \ \forall i = 1, \ldots, n\}.$$

[2] Al radical también se le suele llamar el *núcleo* de la forma bilineal simétrica.

Ejemplo 4.2.21. Consideremos la forma $\varphi = \beta_I : \Bbbk^n \times \Bbbk^n \to \Bbbk$ definida por

$$\varphi((x_1, \ldots, x_n), (y_1, \ldots, y_n)) = x_1 y_1 + \cdots + x_n y_n.$$

Aplicando la observación anterior con $\mathcal{B}$ la base canónica se obtiene $\operatorname{Rad}(\varphi) = \{0\}$. Luego φ es no degenerada.

Para hallar el radical es útil el siguiente resultado.

Proposición 4.2.22. *Sea $A \in M_n(\Bbbk)$ una matriz simétrica. Consideramos la forma bilineal simétrica $\beta_A \in \operatorname{Bil}_S(\Bbbk^n)$ y el operador $L_A \in \mathcal{L}(\Bbbk^n)$. Entonces el radical de β_A coincide con el núcleo de L_A.*

Dem. Observar primero que es

$$\beta_A(u, v) = u^t A v = \beta_I(u, Av), \quad \forall u, v \in \Bbbk^n.$$

Entonces usando que β_I es no degenerada, deducimos

$$v \in \operatorname{Rad}(\varphi) \;\Leftrightarrow\; \beta_A(u, v) = 0,\ \forall u \in \Bbbk^n \;\Leftrightarrow\; \beta_I(u, Av) = 0,\ \forall u \in \Bbbk^n \;\Leftrightarrow\; Av = 0 \;\Leftrightarrow\; v \in \operatorname{Ker}(L_A). \quad \square$$

Ejemplo 4.2.23. Consideremos en $\mathbb{R}^3$ las formas cuadráticas Φ_1 y Φ_2 definidas por $\Phi_1(x, y, z) = x^2 + y^2 - z^2$ y $\Phi_2(x, y, z) = x^2 + y^2$. Las formas bilineales asociadas son β_{A_1} y β_{A_2}, siendo A_1 y A_2 las matrices diagonales $A_1 = \operatorname{diag}(1, 1, -1)$ y $A_2 = \operatorname{diag}(1, 1, 0)$. Los núcleos de las transformaciones lineales asociadas son $\operatorname{Ker}(L_{A_1}) = \{(0, 0, 0)\}$ y $\operatorname{Ker}(L_{A_2}) = [(0, 0, 1)]$. Luego $\operatorname{Rad}(\varphi_1) = \{(0, 0, 0)\}$ y $\operatorname{Rad}(\varphi_2) = [(0, 0, 1)]$. Esto implica que φ_1 es no degenerada.

Observación 4.2.24. La proposición anterior sirve también para hallar el radical en el caso de una forma en un espacio vectorial arbirtrario V. Si $\varphi \in \operatorname{Bil}_S(V)$, $\mathcal{B}$ es una base de V y $A = M_{\mathcal{B}}(\varphi)$, entonces es

$$\varphi(u, v) = \beta_A\big(\operatorname{coord}_{\mathcal{B}}(u), \operatorname{coord}_{\mathcal{B}}(v)\big), \quad \forall u, v \in V.$$

Luego $v \in \operatorname{Rad}(\varphi)$ si y solo si $\operatorname{coord}_{\mathcal{B}}(v) \in \operatorname{Rad}(\beta_A) = \operatorname{Ker}(L_A)$. Luego $\operatorname{Rad}(\varphi) = (\operatorname{coord}_{\mathcal{B}})^{-1}\big(\operatorname{Ker}(L_A)\big)$.

Definición 4.2.25. Dada $\varphi \in \operatorname{Bil}_S(V)$, su *rango* se define mediante $\operatorname{rango}(\varphi) := \operatorname{rango}(M_{\mathcal{B}}(\varphi))$, siendo $\mathcal{B}$ una base cualquiera de V. Esta definición tiene sentido porque dos matrices congruentes siempre tienen el mismo rango. Si Φ es la forma cuadrática asociada a φ, entonces también definimos $\operatorname{rango}(\Phi) := \operatorname{rango}(\varphi)$.

Proposición 4.2.26. *Si $\varphi \in \operatorname{Bil}_S(V)$, entonces*

$$\dim \operatorname{Rad}(\varphi) + \operatorname{rango}(\varphi) = \dim V.$$

Dem. Sean $\mathcal{B}$ una base de V, $n = \dim V$ y $A = M_{\mathcal{B}}(\varphi) \in M_n(\Bbbk)$. Como $\operatorname{coord}_{\mathcal{B}} : V \to \Bbbk^n$ es un isomorfismo, entonces de la observación 4.2.24 deducimos $\dim \operatorname{Rad}(\varphi) = \dim \operatorname{Ker}(L_A)$. Por otro lado es $\operatorname{rango}(\varphi) = \operatorname{rango}(A) = \dim \operatorname{Im}(L_A)$. Luego la tesis se deduce de $\dim \operatorname{Ker}(L_A) + \dim \operatorname{Im}(L_A) = \dim \Bbbk^n$. $\square$

Corolario 4.2.27. *Una forma bilineal $\varphi \in \operatorname{Bil}_S(V)$ es no degenerada si y solo si* $\operatorname{rango}(\varphi) = \dim V$. *Luego si $\mathcal{B}$ es una base de V, entonces φ es no degenerada si y solo si* $\det\big(M_{\mathcal{B}}(\varphi)\big) \neq 0$. $\square$

Ejemplo 4.2.28. Sea $\varphi \in \operatorname{Bil}_S(\mathbb{R}^3)$ tal que su forma cuadrática asociada es

$$\Phi(x, y, z) = -7x^2 - 2y^2 + 7z^2 + 8xy + 2xz - 4yz, \quad \forall (x, y, z) \in \mathbb{R}^3.$$

Si $\mathcal{C}$ es la base canónica de $\mathbb{R}^3$, es $M_{\mathcal{C}}(\varphi) = A = \begin{pmatrix} -7 & 4 & 1 \\ 4 & -2 & -2 \\ 1 & -2 & 7 \end{pmatrix}$. Como es $\det A = 0$, deducimos que φ degenera. Más precisamente vale $\operatorname{rango}(A) = 2$, luego la proposición 4.2.26 nos dice que el radical de φ tiene dimensión 1. Para hallarlo aplicamos la proposición 4.2.22, obteniendo $\operatorname{Rad}(\varphi) = \operatorname{Ker}(L_A) = [(3, 5, 1)]$.

4.3. Diagonalización

De ahora en adelante φ es una forma bilineal simétrica en V y Φ es su forma cuadrática asociada. Para evitar discusiones triviales en general asumiremos $V \neq \{0\}$.

Definición 4.3.1. Una base $\mathcal{B} = \{v_1, \dots, v_n\}$ de V se dice *φ-ortogonal* si $\varphi(v_i, v_j) = 0$ para todo $i \neq j$.

Observación 4.3.2. Si $\mathcal{B}$ es una base de V, entonces $\mathcal{B}$ es φ-ortogonal si y solo si la matriz $M_{\mathcal{B}}(\varphi)$ es diagonal. Explícitamente, si $\mathcal{B} = \{v_1, \dots, v_n\}$ es una base φ-ortogonal de V, entonces

$$M_{\mathcal{B}}(\varphi) = \begin{pmatrix} \Phi(v_1) & 0 & \cdots & 0 \\ 0 & \Phi(v_2) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \Phi(v_n) \end{pmatrix}.$$

Luego las expresiones de φ y Φ con coordenadas en esta base son

$$\varphi(u,v) = \sum_{i=1}^{n} \Phi(v_i) x_i y_i, \quad \Phi(u) = \sum_{i=1}^{n} \Phi(v_i) x_i^2, \tag{4.3}$$

siendo $u = \sum_{i=1}^{n} x_i v_i$ y $v = \sum_{i=1}^{n} y_i v_i$.

Ejemplo 4.3.3. Consideremos la forma cuadrática $\Phi : \mathbb{R}^3 \to \mathbb{R}$ definida por $\Phi(x,y,z) = x^2 - y^2$ y sea $\varphi \in \mathrm{Bil}_S(\mathbb{R}^3)$ la forma bilineal correspondiente. Si $\mathcal{B}$ es la base canónica de $\mathbb{R}^3$, es

$$M_{\mathcal{B}}(\varphi) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Luego $\mathcal{B}$ es una base φ-ortogonal.

Definición 4.3.4. Si W es un subespacio de V, la *restricción de φ a W* es la función $\varphi|_{W\times W} : W \times W \to \Bbbk$, definida por $\varphi|_{W\times W}(w_1, w_2) := \varphi(w_1, w_2)$, para todo $w_1, w_2 \in W$. Claramente $\varphi|_{W\times W} \in \mathrm{Bil}_S(W)$.

Teorema 4.3.5. *Si $\varphi \in \mathrm{Bil}_S(V)$, entonces existe una base $\mathcal{B}$ de V tal que $M_{\mathcal{B}}(\varphi)$ es una matriz diagonal.*

Dem. Tenemos que demostrar es que existe una base φ-ortogonal. Lo probaremos por inducción en $n = \dim V$.

Si $n = 1$, entonces toda base $\mathcal{B} = \{v\}$ de V es φ-ortogonal. Supongamos ahora que vale la tesis si la dimensión del espacio es $n-1$ y sea $\varphi \in \mathrm{Bil}_S(V)$ con $\dim V = n$. Si $\varphi = 0$, entonces toda base de V es φ-ortogonal. Supongamos ahora que $\varphi \neq 0$. Por la proposición 4.2.13 es $\Phi \neq 0$, luego existe $u \in V$ tal que $\Phi(u) \neq 0$.

Definimos $\alpha \in V^*$ mediante $\alpha(v) = \varphi(v,u)$, para todo $v \in V$. Observar que $\alpha(u) = \varphi(u,u) = \Phi(u) \neq 0$, luego $\alpha \neq 0$ y por lo tanto $\dim \mathrm{Ker}\,(\alpha) = n-1$. Aplicando la hipótesis de inducción a φ restringida a $\mathrm{Ker}\,(\alpha)$ tenemos que existe $\{v_1, \dots, v_{n-1}\}$ base de $\mathrm{Ker}\,(\alpha)$ tal que $\varphi(v_i, v_j) = 0$ si $i \neq j$.

Como $u \notin \mathrm{Ker}\,(\alpha) = [v_1, \dots, v_{n-1}]$, entonces el conjunto $\mathcal{B} = \{v_1, \dots, v_{n-1}, u\}$ es LI, y al tener n elementos es base de V. Como $\{v_1, \dots, v_{n-1}\} \subset \mathrm{Ker}\,(\alpha)$, resulta que $v_1, \dots, v_{n-1}$ son φ-ortogonales con u. Luego $\mathcal{B}$ es una base φ-ortogonal de V. $\square$

Corolario 4.3.6. *Una matriz es simétrica si y solo si es congruente con una matriz diagonal.*

Dem. Si una matriz A es congruente con una matriz diagonal D, entonces existe una matriz invertible Q tal que $A = Q^t DQ$. Luego

$$A^t = (Q^t DQ)^t = Q^t D^t Q^{tt} = Q^t DQ = A$$

entonces A es simétrica. Recíprocamente, si A es simétrica entonces la forma bilineal asociada $\beta_A : \Bbbk^n \times \Bbbk^n \to \Bbbk$ es simétrica. Luego el teorema anterior implica que existe $\mathcal{B}$ base de $\Bbbk^n$ tal que $D = M_{\mathcal{B}}(\beta_A)$ es una matriz diagonal. Entonces si $\mathcal{C}$ es la base canónica de $\Bbbk^n$ y $Q = {}_{\mathcal{B}}[\mathrm{Id}]_{\mathcal{C}}$, obtenemos

$$A = M_{\mathcal{C}}(\beta_A) = \left({}_{\mathcal{B}}[\mathrm{Id}]_{\mathcal{C}}\right)^t M_{\mathcal{B}}(\beta_A)\, {}_{\mathcal{B}}[\mathrm{Id}]_{\mathcal{C}} = Q^t DQ. \quad \square$$

Algoritmo de diagonalización Dada una forma bilineal simétrica φ en V, el teorema 4.3.5 nos asegura que siempre existe una base φ-ortogonal de V, es decir una base en la cual la matriz asociada a φ es diagonal. En lo que sigue veremos cómo hallarla.

Sea $\varphi \in \mathrm{Bil}_S(V)$. Empezamos tomando una base cualquiera $\mathcal{C}$ de V y considerando $A = M_{\mathcal{C}}(\varphi) \in M_n(\Bbbk)$. Si $Q \in M_n(\Bbbk)$ es una matriz invertible cualquiera y definimos $D = Q^t AQ$, entonces la proposición 4.1.14 nos dice que existe una base $\mathcal{B}$ de V tal que $D = M_{\mathcal{B}}(\varphi)$. Luego para obtener una base φ-ortogonal de V, alcanza con encontrar una matriz invertible Q de forma tal que $D = Q^t AQ$ sea una matriz diagonal. La idea para esto es ir transformando A con matrices elementales $Q_1, \dots, Q_l$ hasta llegar a una matriz diagonal D,

$$A \sim Q_1^t AQ_1 \sim Q_2^t Q_1^t AQ_1Q_2 = (Q_1Q_2)^t A(Q_1Q_2) \sim \cdots \sim (Q_1 \cdots Q_l)^t A(Q_1 \cdots Q_l) = D.$$

Observar que si Q es una matriz elemental de tipo I, II o III, entonces $Q^t AQ$ es la matriz que se obtiene realizando la operación elemental correspondiente en las filas y columnas de A.

Mostraremos el método mediante un ejemplo. Consideremos en $\mathbb{R}^3$ la forma cuadrática Φ definida por

$$\Phi(x, y, z) = -7x^2 - 2y^2 + 7z^2 + 8xy + 2xz - 4yz, \quad \forall (x, y, z) \in \mathbb{R}^3.$$

Sea $\varphi \in \mathrm{Bil}_S(\mathbb{R}^3)$ la forma bilineal simétrica asociada a Φ. Si $\mathcal{C}$ es la base canónica de $\mathbb{R}^3$, es

$$M_{\mathcal{C}}(\varphi) = A = \begin{pmatrix} -7 & 4 & 1 \\ 4 & -2 & -2 \\ 1 & -2 & 7 \end{pmatrix}.$$

Para hallar su forma diagonal, primero sumamos a la tercer columna de A la segunda multiplicada por -1 y luego sumamos a la tercer fila la segunda multiplicada por -1.

$$\begin{pmatrix} -7 & 4 & 1 \\ 4 & -2 & -2 \\ 1 & -2 & 7 \end{pmatrix} \Rightarrow \begin{pmatrix} -7 & 4 & -3 \\ 4 & -2 & 0 \\ 1 & -2 & 9 \end{pmatrix} \Rightarrow \begin{pmatrix} -7 & 4 & -3 \\ 4 & -2 & 0 \\ -3 & 0 & 9 \end{pmatrix}. \tag{4.4}$$

Ahora le sumamos a la primer columna la segunda multiplicada por 2 y luego sumamos a la primer fila la segunda multiplicada por 2.

$$\begin{pmatrix} -7 & 4 & -3 \\ 4 & -2 & 0 \\ -3 & 0 & 9 \end{pmatrix} \Rightarrow \begin{pmatrix} 1 & 4 & -3 \\ 0 & -2 & 0 \\ -3 & 0 & 9 \end{pmatrix} \Rightarrow \begin{pmatrix} 1 & 0 & -3 \\ 0 & -2 & 0 \\ -3 & 0 & 9 \end{pmatrix}. \tag{4.5}$$

Finalmente le sumamos a la tercer columna la primera multiplicada por 3 y luego sumamos a la tercer fila la primera multiplicada por 3.

$$\begin{pmatrix} 1 & 0 & -3 \\ 0 & -2 & 0 \\ -3 & 0 & 9 \end{pmatrix} \Rightarrow \begin{pmatrix} 1 & 0 & 0 \\ 0 & -2 & 0 \\ -3 & 0 & 0 \end{pmatrix} \Rightarrow \begin{pmatrix} 1 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & 0 \end{pmatrix} = D. \tag{4.6}$$

Observar que la matriz obtenida en (4.4) corresponde a calcular $E_1^t A E_1$ siendo $E_1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{pmatrix}$. La matriz obtenida en (4.5) corresponde a calcular $E_2^t E_1^t A E_1 E_2$, siendo $E_2 = \begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$. Finalmente la matriz diagonal D obtenida en (4.6) corresponde a calcular $E_3^t E_2^t E_1^t A E_1 E_2 E_3$, siendo $E_3 = \begin{pmatrix} 1 & 0 & 3 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$. Luego es $D = Q^t A Q$, siendo $Q = E_1 E_2 E_3 = \begin{pmatrix} 1 & 0 & 3 \\ 2 & 1 & 5 \\ 0 & 0 & 1 \end{pmatrix}$. Observar que si consideramos $\mathcal{B} = \{(1,2,0),\ (0,1,0),\ (3,5,1)\}$, entonces $\mathcal{B}$ es una base de $\mathbb{R}^3$ tal que ${}_{\mathcal{C}}[\mathrm{Id}]_{\mathcal{B}} = Q$. Luego $\mathcal{B}$ es una base φ-ortogonal de $\mathbb{R}^3$ y la matriz diagonal correspondiente $M_{\mathcal{B}}(\varphi) = D$ está dada por (4.6). Notar que la matriz Q que verifica $D = Q^t A Q$, es $Q = E_1 E_2 E_3$, y se obtiene realizando las operaciones elementales correspondientes a E_1, E_2, E_3, en las columnas de la matriz identidad I. Esta última observación es la base del siguiente algoritmo.

Un método para obtener la matriz Q es el siguiente, escribimos a la izquierda la matriz A y a la derecha la matriz identidad I, luego realizamos operaciones elementales en A y cada vez que hacemos una operación elemental en las colummnas de A, también la hacemos en las columnas de I, pero cuando hacemos operaciones elementales en las filas de A no hacemos nada en la matriz I. Al finalizar el algoritmo, cuando en la izquierda obtenemos la matriz diagonal D, en la derecha está la matriz de congruencia Q. Lo mostraremos en el ejemplo anterior. En lo que sigue $C_3 - C_2$, es que a la columna 3 le vamos a restar la columna 2, etc.

$$
\begin{array}{rrr|rrrc}
-7 & 4 & 1 & 1 & 0 & 0 & \\
4 & -2 & -2 & 0 & 1 & 0 & C_3 - C_2 \\
1 & -2 & 7 & 0 & 0 & 1 & \\
\hline
-7 & 4 & -3 & 1 & 0 & 0 & \\
4 & -2 & 0 & 0 & 1 & -1 & F_3 - F_2 \\
1 & -2 & 9 & 0 & 0 & 1 & \\
\hline
-7 & 4 & -3 & & & & \\
4 & -2 & 0 & & & & C_1 + 2C_2 \\
-3 & 0 & 9 & & & & \\
\hline
1 & 4 & -3 & 1 & 0 & 0 & \\
0 & -2 & 0 & 2 & 1 & -1 & F_1 + F_2 \\
-3 & 0 & 9 & 0 & 0 & 1 & \\
\hline
1 & 0 & -3 & & & & \\
0 & -2 & 0 & & & & C_3 + 3C_1 \\
-3 & 0 & 9 & & & & \\
\hline
1 & 0 & 0 & 1 & 0 & 3 & \\
0 & -2 & 0 & 2 & 1 & 5 & F_3 + F_1 \\
-3 & 0 & 0 & 0 & 0 & 1 & \\
\hline
1 & 0 & 0 & & & & \\
0 & -2 & 0 & & & & \\
0 & 0 & 0 & & & &
\end{array}
$$

Luego $D = Q^t A Q$, siendo $D = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & 0 \end{pmatrix}$ y $Q = \begin{pmatrix} 1 & 0 & 3 \\ 2 & 1 & 5 \\ 0 & 0 & 1 \end{pmatrix}$.

El siguiente reaultado permite obtener $\operatorname{Rad}(\varphi)$ cuando se conoce una base φ-ortogonal.

Proposición 4.3.7. *Sea $\varphi \in \operatorname{Bil}_S(V)$ y $\mathcal{B} = \{v_1, \dots, v_n\}$ una base φ-ortogonal de V. Supongamos que ordenamos $\mathcal{B}$ de forma tal que*

$$\Phi(v_i) \neq 0, \quad \forall i = 1, \dots, r; \quad \Phi(v_i) = 0, \quad \forall i = r+1, \dots, n.$$

Entonces $\operatorname{Rad}(\varphi) = [v_{r+1}, \dots, v_n]$.

Dem. Sea $A = M_{\mathcal{B}}(\varphi) \in M_n(\mathbb{k})$. La matriz A es diagonal y sus entradas diagonales son $\Phi(v_1), \dots, \Phi(v_n)$. Luego $\operatorname{rango}(\varphi) = \operatorname{rango} A = r$ y por lo tanto $\dim \operatorname{Rad}(\varphi) = n - r$. Además como la base $\mathcal{B}$ es φ-ortogonal y $\Phi(v_{r+1}) = \cdots = \Phi(v_n) = 0$, entonces vale $\varphi(v_i, v_j) = 0$, para todo $i = r+1, \dots, n$ y para todo $j = 1, \dots, n$. Esto implica $[v_{r+1}, \dots, v_n] \subset \operatorname{Rad}(\varphi)$, y como ambos subespacios tienen la misma dimensión, coinciden. $\square$

Ejemplo 4.3.8. En el ejemplo anterior obtuvimos $D = Q^t A Q$ siendo $D = \operatorname{diag}(1, -2, 0)$. Luego la proposición anterior nos dice que la última columna de Q es una base del radical de φ, es decir $\operatorname{Rad}(\varphi) = [(3, 5, 1)]$. Notar que este es el mismo resultado que obtuvimos en el ejemplo 4.2.28, pero aplicando un método completamente distinto.

4.4. Formas bilineales simétricas reales

En esta sección el cuerpo de base $\mathbb{k}$ es $\mathbb{R}$ y V es un $\mathbb{R}$-espacio vectorial de dimensión finita.

Definición 4.4.1. Una base $\mathcal{B} = \{v_1, \dots, v_n\}$ de V se dice *φ-ortonormal* si es una base φ-ortogonal y además $\Phi(v_i) \in \{-1, 0, 1\}$, para todo $i = 1, \dots, n$.

Observación 4.4.2. Si $\mathcal{B} = \{v_1, \dots, v_n\}$ es una base φ-ortonormal de V, entonces la matriz $M_{\mathcal{B}}(\varphi)$ es diagonal

$$M_{\mathcal{B}}(\varphi) = \begin{pmatrix} \Phi(v_1) & 0 & \cdots & 0 \\ 0 & \Phi(v_2) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \Phi(v_n) \end{pmatrix}$$

y sus entradas diagonales $\Phi(v_1), \dots, \Phi(v_n)$ son 0, 1 o -1. Notar que reordenando $v_1, \dots, v_n$ podemos asumir genéricamente que existen $1 \leq p \leq r \leq n$ tales que

$$\Phi(v_i) = 1,\ \forall i = 1, \dots, p; \quad \Phi(v_i) = -1,\ \forall i = p+1, \dots, r; \quad \Phi(v_i) = 0,\ \forall i = r+1, \dots, n.$$

Luego las expresiones de φ y Φ con coordenadas en esta base son

$$\varphi(u, v) = x_1 y_1 + \cdots + x_p y_p - x_{p+1} y_{p+1} - \cdots - x_r y_r; \quad \Phi(u) = x_1^2 + \cdots + x_p^2 - x_{p+1}^2 - \cdots - x_r^2.$$

siendo $u = \sum_{i=1}^n x_i v_i$ y $v = \sum_{i=1}^n y_i v_i$.

Ejemplo 4.4.3. Si consideramos la forma bilineal $\varphi \in \operatorname{Bil}_S(\mathbb{R}^3)$ asociada a la forma cuadrática $\Phi : \mathbb{R}^3 \to \mathbb{R}$ definida por $\Phi(x, y, z) = x^2 - y^2$, entonces la base canónica es una base φ-ortonormal de $\mathbb{R}^3$.

Proposición 4.4.4. *Si $\mathbb{k} = \mathbb{R}$ y $\varphi \in \operatorname{Bil}_S(V)$, entonces existe una base φ-ortonormal de V.*

Dem. Sea $\mathcal{B} = \{v_1, \dots, v_n\}$ una base φ-ortogonal de V que suponemos ordenada de forma tal que existen $1 \leq p \leq r \leq n$ tales que

$$\Phi(v_i) > 0,\ \forall i = 1, \dots, p; \quad \Phi(v_i) < 0,\ \forall i = p+1, \dots, r; \quad \Phi(v_i) = 0,\ \forall i = r+1, \dots, n.$$

Luego si definimos $w_1, \dots, w_n$ por

$$w_i = \frac{1}{\sqrt{\Phi(v_i)}} v_i,\ \forall i = 1, \dots, p; \quad w_i = \frac{1}{\sqrt{-\Phi(v_i)}} v_i,\ \forall i = p+1, \dots, r; \quad w_i = v_i,\ \forall i = r+1, \dots, n,$$

entonces es claro que w_i es φ-ortogonal con w_j si $i \neq j$, y vale

$$\Phi(w_i) = 1,\ \forall i = 1, \dots, p; \quad \Phi(w_i) = -1,\ \forall i = p+1, \dots, r; \quad \Phi(w_i) = 0,\ \forall i = r+1, \dots, n.$$

Luego $\mathcal{C} = \{w_1, \dots, w_n\}$ es una base φ-ortonormal de V. $\square$

Ejemplo 4.4.5. En la sección anterior vimos que si consideramos $\Phi : \mathbb{R}^3 \to \mathbb{R}$ definida por

$$\Phi(x,y,z) = -7x^2 - 2y^2 + 7z^2 + 8xy + 2xz - 4yz, \quad \forall (x,y,z) \in \mathbb{R}^3,$$

entonces $\mathcal{B} = \{(1,2,0),\ (0,1,0),\ (3,5,1)\}$ es una base φ-ortogonal y $\Phi(1,2,0) = 1$, $\Phi(0,1,0) = -2$ y $\Phi(3,5,1) = 0$. “Normalizando” el segundo vector obtenemos $\Phi\left(0, \frac{1}{\sqrt{2}}, 0\right) = -1$. Luego

$$\mathcal{B}' = \left\{(1,2,0),\ \left(0, \frac{1}{\sqrt{2}}, 0\right),\ (3,5,1)\right\}$$

es una base φ-ortonormal de $\mathbb{R}^3$.

Observaciones 4.4.6.

1. La existencia de bases φ-ortonormales es un dato intresante pero no es particulamente útil, dado que a diferencia de lo que ocurre en espacios con producto interno, para formas bilineales en general las bases φ-ortonormales no tienen grandes ventajas sobre las bases φ-ortogonales. Por ejemplo no hay un análogo de la fórmula $u = \sum_{i=1}^{n} \langle u, v_i \rangle v_i$, para todo $u \in V$ que se cumple si $\{v_1, \dots, v_n\}$ es una base ortonormal de V. Por esto es que en general se trabaja con bases φ-ortogonales.

2. La existencia de bases φ-ortonormales depende del cuerpo $\Bbbk$. Por ejemplo, consideremos $\mathbb{Q}^2$ como $\mathbb{Q}$-espacio y φ la forma bilineal asociada a la forma cuadrática $\Phi(x,y) = 2x^2$. Como no existe ningún $x \in \mathbb{Q}$ tal que $2x^2 = \pm 1$, entonces no existen bases φ-ortonormales de $\mathbb{Q}^2$. Notar que la base canónica de $\mathbb{Q}^2$ es φ-ortogonal, pero no hay forma de normalizarla.

A continuación introducimos varias definiciones que se suelen usar al trabajar con formas bilineales simétricas y formas cuadráticas reales.

Definición 4.4.7. Sea Φ una forma cuadrática en V y φ la forma bilineal simétrica asociada.

1. Φ es *definida positiva* si $\Phi(v) > 0$, para todo $v \neq 0$.

2. Φ es *definida negativa* si $\Phi(v) < 0$, para todo $v \neq 0$.

3. Φ es *definida*, si es definida positiva o es definida negativa.

4. Φ es *semidefinida positiva* si $\Phi(v) \geq 0$, para todo $v \in V$ y existe $0 \neq v_0 \in V$ tal que $\Phi(v_0) = 0$.

5. Φ es *semidefinida negativa* si $\Phi(v) \leq 0$, para todo $v \in V$ y existe $0 \neq v_0 \in V$ tal que $\Phi(v_0) = 0$.

6. Φ es *semidefinida*, si es semidefinida positiva o es semidefinida negativa.

7. Φ es *no definida* si existen v_1 y v_2 en V tales que $\Phi(v_1) > 0$ y $\Phi(v_2) < 0$.

8. Φ es *no degenerada* si lo es φ, es decir si $\varphi(u,v) = 0$ para todo $v \in V$, implica $u = 0$.

Decimos que una forma bilineal simétrica verifica alguna de las propiedades anteriores si lo hace su forma cuadrática asociada.

Observación 4.4.8. Una forma bilineal simétrica definida positiva es lo mismo que un producto interno real.

Ejemplo 4.4.9. Consideremos las siguientes formas cuadráticas en $\mathbb{R}^3$.

1. $\Phi(x,y,z) = x^2 + y^2 + z^2$, es definida positiva (la forma bilineal asociada es el producto escalar).
2. $\Phi(x,y,z) = -x^2 - y^2 - z^2$, es definida negativa.
3. $\Phi(x,y,z) = x^2 + y^2$, es semidefinida positiva.
4. $\Phi(x,y,z) = -x^2 - y^2$, es semidefinida negativa.
5. $\Phi(x,y,z) = x^2 - y^2$, es no definida y degenerada.
6. $\Phi(x,y,z) = x^2 - y^2 - z^2$, es no definida y no degenerada.

Usando las fórmulas (4.3) es fácil de probar el siguiente resultado.

Proposición 4.4.10. *Sea Φ una forma cuadrática en V y φ la forma bilineal simétrica asociada. Si $\mathcal{B} = \{v_1, \ldots, v_n\}$ es una base φ-ortogonal de V, entonces vale lo siguiente.*

1. *Φ es* definida positiva *si y solo si $\Phi(v_i) > 0$, para todo $i = 1, \ldots, n$.*
2. *Φ es* definida negativa *si y solo si $\Phi(v_i) < 0$, para todo $i = 1, \ldots, n$.*
3. *Φ es* semidefinida positiva *si y solo si $\Phi(v_i) \geq 0$, para todo $i = 1, \ldots, n$ y existe i_0 tal que $\Phi(v_{i_0}) = 0$.*
4. *Φ es* semidefinida negativa *si y solo si $\Phi(v_i) \leq 0$, para todo $i = 1, \ldots, n$ y existe i_0 tal que $\Phi(v_{i_0}) = 0$.*
5. *Φ es* no definida *si y solo si existen i y j tales que $\Phi(v_i) > 0$ y $\Phi(v_j) < 0$.*
6. *Φ es* no degenerada *si y solo si existe algún i tal que $\Phi(v_i) = 0$.* □

Combinando la proposición anterior y la proposición 4.3.7 obtenemos el siguiente resultado.

Corolario 4.4.11. *Sea Φ una forma cuadrática en V. Si Φ es definida, entonces es no degenerada, y si Φ es semidefinida, entonces degenera.* □

Teorema 4.4.12 (Ley de inercia de Sylvester)**.** *La cantidad de entradas positivas, negativas y nulas de una matriz diagonal asociada a una forma cuadrática no depende de la representación diagonal.*

Dem. Sea Φ una forma cuadrática en V y φ su forma bilineal asociada. Sea $\mathcal{B} = \{v_1, \ldots, v_n\}$ una base φ-ortogonal de V, que podemos considerar ordenada de forma tal que

$$\Phi(v_i) > 0,\ \forall i = 1, \ldots, p; \quad \Phi(v_i) < 0,\ \forall i = p+1, \ldots, r; \quad \Phi(v_i) = 0,\ \forall i = r+1, \ldots, n,$$

para ciertos $1 \leq p \leq r \leq n$. La proposición 4.3.7 nos dice que es $\operatorname{Rad}(\varphi) = [v_{r+1}, \ldots, v_n]$ y es claro que el rango de φ es r, dado que r es la cantidad de entradas no nulas de la diagonal principal de $M_{\mathcal{B}}(\varphi)$.

Sea $\mathcal{C} = \{w_1, \ldots, w_n\}$ otra base φ-ortogonal de V ordenada como $\mathcal{B}$. Entonces existe $1 \leq q \leq r$ tal que

$$\Phi(w_i) > 0,\ \forall i = 1, \ldots, q; \quad \Phi(w_i) < 0,\ \forall i = q+1, \ldots, r; \quad \Phi(w_i) = 0,\ \forall i = r+1, \ldots, n.$$

Para probar la tesis lo único que hay que demostrar es que vale $p = q$. Consideremos los siguientes subespacios

$$V_+ = [v_1, \ldots, v_p], \quad V_- = [v_{p+1}, \ldots, v_r], \quad W_+ = [w_1, \ldots, w_q], \quad W_- = [w_{q+1}, \ldots, w_r].$$

Lo que tenemos que probar es que vale $\dim V_+ = \dim W_+$. Notar que como $\mathcal{B}$ y $\mathcal{C}$ son bases de V, deducimos $V = V_+ \oplus V_- \oplus \operatorname{Rad}(\varphi)$ y $V = W_+ \oplus W_- \oplus \operatorname{Rad}(\varphi)$, siendo $\operatorname{Rad}(\varphi) = [v_{r+1}, \ldots, v_n] = [w_{r+1}, \ldots, w_n]$.

Observar que $\{v_1, \ldots, v_p\}$ es una base φ-ortogonal de V_+ y vale $\Phi(v_i) > 0$, para todo $i = 1, \ldots, p$. Luego Φ es definida positiva en V_+. El mismo razonamiento prueba que Φ es definida negativa en V_-. Considerando $W_\pm$ en lugar de $V_\pm$, deducimos que Φ es definida positiva en W_+ y definida negativa en W_-.

Probaremos que los subespacios W_+ y $V_- \oplus \operatorname{Rad}(\varphi)$ son independientes. Sea $v \in W_+ \cap \big(V_- \oplus \operatorname{Rad}(\varphi)\big)$. Como $v \in V_- \oplus \operatorname{Rad}(\varphi)$, entonces existen únicos $v_0 \in \operatorname{Rad}(\varphi)$ y $v_- \in V_-$ tales que $v = v_- + v_0$. Luego

$$\Phi(v) = \Phi(v_- + v_0) = \Phi(v_-) + 2\varphi(v_-, v_0) + \Phi(v_0) = \Phi(v_-) \leq 0.$$

Por otro lado, como $v \in W_+$ y Φ es definida positiva en W_+, si fuese $v \neq 0$ sería $\Phi(v) > 0$ lo cual nos lleva a una contradicción. Entonces la única posibilidad es $v = 0$ y por lo tanto $W_+ \cap \big(V_- \oplus \operatorname{Rad}(\varphi)\big) = \{0\}$.

La independencia de W_+ y $V_- \oplus \operatorname{Rad}(\varphi)$ implica que la suma $W_+ + \big(V_- \oplus \operatorname{Rad}(\varphi)\big)$ es directa, luego

$$\dim V \geq \dim(W_+ \oplus V_- \oplus \operatorname{Rad}(\varphi)) = \dim W_+ + \dim V_- + \dim \operatorname{Rad}(\varphi).$$

Por otro lado, de $V = V_+ \oplus V_- \oplus \operatorname{Rad}(\varphi)$, deducimos $\dim V = \dim V_+ + \dim V_- + \dim \operatorname{Rad}(\varphi)$. Comparando esta fórmula con la desigualdad anterior, deducimos $\dim V_+ \geq \dim W_+$. Intercambiando los roles de $V_\pm$ y $W_\pm$ obtenemos $\dim W_+ \geq \dim V_+$, luego $\dim V_+ = \dim W_+$. $\square$

Definición 4.4.13. Dada $\varphi \in \operatorname{Bil}_S(V)$, decimos que dos subespacios U y W de V son *φ-ortogonales* si $\varphi(u, w) = 0$ para todo $u \in U$ y $w \in W$. Una *φ-descomposición* de V es una descomposición del tipo $V = V_+ \oplus V_- \oplus \operatorname{Rad}(\varphi)$, en la cual V_+ y V_- son φ-ortogonales, Φ es definida positiva en V_+ y es definida negativa en V_-.

Observación 4.4.14. La prueba del teorema anterior nos muestra que toda base φ-ortogonal $\mathcal{B}$ induce una φ-descomposición $V = V_+ \oplus V_- \oplus \operatorname{Rad}(\varphi)$, de forma tal que las cantidades de entradas diagonales positivas, negativas y nulas de $M_\mathcal{B}(\varphi)$ coinciden respectivamente con $\dim V_+$, $\dim V_-$ y $\dim \operatorname{Rad}(\varphi)$, Recíprocamente, si tenemos una φ-descomposición $V = V_+ \oplus V_- \oplus \operatorname{Rad}(\varphi)$ y tomamos bases φ-ortogonales de V_+, V_-, $\operatorname{Rad}(\varphi)$ y las unimos, entonces obtenemos una base φ-ortogonal de V que mediante el proceso anterior da lugar a esa φ-descomposición.

En general existen distintas φ-descomposiciones $V = V_+ \oplus V_- \oplus \operatorname{Rad}(\varphi)$, pero la ley de inercia de Sylvester muestra que en esas φ-descomposiciones hay unicidad en las dimensiones de V_+ y V_-.

Ejemplo 4.4.15. Consideremos la forma cuadrática $\Phi : \mathbb{R}^3 \to \mathbb{R}$ definida por $\Phi(x, y, z) = x^2 - y^2$, para todo $(x, y, z) \in \mathbb{R}^3$. Es fácil de probar que vale $\operatorname{Rad}(\varphi) = [(0, 0, 1)]$ y que si definimos $V_+ = [(1, 0, 0)]$, $V_- = [(0, 1, 0)]$, $W_+ = [(2, 1, 0)]$ y $W_- = [(1, 2, 0)]$, entonces $\mathbb{R}^3 = V_+ \oplus V_- \oplus \operatorname{Rad}(\varphi)$ y $\mathbb{R}^3 = W_+ \oplus W_- \oplus \operatorname{Rad}(\varphi)$ son dos φ-descomposiciones distintas de $\mathbb{R}^3$.

Definición 4.4.16. Sea Φ una forma cuadrática en V y φ su forma bilineal simétrica asociada. Sea $\mathcal{B}$ una base φ-ortogonal de V. El *índice positivo (negativo) de inercia* de Φ es la cantidad de entradas diagonales positivas (negativas) de $M_\mathcal{B}(\varphi)$. Para abreviar llamaremos *índice* al índice positivo de inercia. La *signatura* de Φ es diferencia entre la cantidad de entradas diagonales positivas y negativas de $M_\mathcal{B}(\varphi)$. La signatura, el índice y el rango son los *invariantes* de Φ.

Observación 4.4.17. Con las notaciones anteriores, si el número de entradas positivas de $M_\mathcal{B}(\varphi)$ es p y el de entradas negativas es q, entonces el índice de Φ es p, el rango es $r = p + q$ y la signatura es $s = p - q$. Luego $s + r = 2p$. Por lo tanto para obtener los tres invariantes alcanza con conocer dos de ellos.

Terminamos este capítulo con algunos resultados que se obtienen usando que toda matriz simétrica real es ortogonalmente equivalente a una matriz diagonal (teorema 3.5.18).

Teorema 4.4.18. *Sea $\varphi \in \operatorname{Bil}_S(V)$. Entonces las cantidades de entradas diagonales positivas, negativas y nulas de una representación matricial diagonal arbitraria de φ coinciden respectivamente con las cantidades de valores propios positivos, negativos y nulos de una representación matricial arbitraria de φ.*

Dem. Sea $\mathcal{B}$ una base arbitraria de V y $A = M_{\mathcal{B}}(\varphi)$. La matriz A es simétrica real, luego existen matrices $D, Q \in M_n(\mathbb{R})$ tales que D es diagonal, Q es ortogonal y $D = Q^{-1}AQ = Q^tAQ$.

Como D es ortogonalmente equivalente con $A = M_{\mathcal{B}}(\varphi)$, entonces existe $\mathcal{C}$ base de V tal que $D = M_{\mathcal{C}}(\varphi)$. Por otro lado como A y D son semejantes, entonces tienen los mismos valores propios, los cuales coinciden con las entradas diagonales de D. Luego encontramos una base $\mathcal{C}$ tal que $D = M_{\mathcal{C}}(\varphi)$ es diagonal y las entradas diagonales de D son los valores propios de A. Luego la tesis se obtiene aplicando la ley de inercia de Sylvester. $\square$

Corolario 4.4.19. *Sea $A \in M_n(\mathbb{R})$ una matriz simétrica. Entonces la cantidad de entradas diagonales positivas, negativas y nulas de cualquier matriz diagonal congruente con A coincide respectivamente con la cantidad de valores propios positivos, negativos y nulos de A.*

Dem. La matriz A es una representación matricial de $\beta_A \in \operatorname{Bil}_S(\mathbb{R}^n)$ (en la base canónica), luego cualquier matriz diagonal que sea congruente con A es también una representación matricial de β_A y por lo tanto se aplica el teorema anterior. $\square$

Observación 4.4.20. Combinando este corolario con el algoritmo de la sección 4.3, se obtiene una forma simple de conocer la cantidad de valores propios positivos, negativos y nulos de una matriz simétrica real.

Teorema 4.4.21. *Sea $\varphi \in \operatorname{Bil}_S(V)$, siendo V un espacio vectorial real con producto interno. Entonces existe una base ortonormal $\mathcal{B}$ de V tal que $M_{\mathcal{B}}(\varphi)$ es diagonal.*

Dem. Sea $\mathcal{C}$ una base ortonormal cualquiera de V y $A = M_{\mathcal{C}}(\varphi)$. La matriz A es simétrica real, luego existen matrices $D, Q \in M_n(\mathbb{R})$ tales que D es diagonal, Q es ortogonal y $D = Q^{-1}AQ = Q^tAQ$.

Como Q es invertible, entonces existe $\mathcal{B}$ base de V tal que $Q = {}_{\mathcal{C}}[\operatorname{Id}]_{\mathcal{B}}$. Luego

$$D = Q^tAQ = \left({}_{\mathcal{C}}[\operatorname{Id}]_{\mathcal{B}}\right)^t M_{\mathcal{C}}(\varphi)\,{}_{\mathcal{C}}[\operatorname{Id}]_{\mathcal{B}} = M_{\mathcal{B}}(\varphi).$$

Además como ${}_{\mathcal{C}}[\operatorname{Id}]_{\mathcal{B}}$ es ortogonal y la base $\mathcal{C}$ es ortonormal, entonces $\mathcal{B}$ es una base ortonormal. En resumen, encontramos una base ortonormal $\mathcal{B}$ tal que $M_{\mathcal{B}}(\varphi) = D$ es una matriz diagonal. $\square$

Observación 4.4.22. El teorema anterior es un resultado de diagonalización simultánea. En esencia lo que nos dice es que si en un espacio real de dimensión finita tenemos dos formas bilineales simétricas y una de ellas es definida positiva, entonces existe una base en la cual ambas se diagonalizan.

Capítulo 5

Transformaciones lineales y polinomios

En el capítulo 1 estudiamos la diagonalización de operadores en espacios de dimensión finita. Este capítulo y el siguiente profundizan en el tema, y exploran el caso no diagonalizable.

En lo que sigue T es un operador arbitrario fijo en un $\Bbbk$-espacio vectorial no nulo de dimensión finita V.

5.1. Subespacios invariantes

Recordemos que un subespacio W de V se dice T*-invariante* si verifica $T(W) \subset W$. En este caso escribimos $T|_W : W \to W$ a la restricción de T a W, es decir a la función definida por $T|_W(w) = T(w)$ para todo $w \in W$. Notar que si W es T-invariante, entonces $T|_W \in \mathcal{L}(W)$.

Ejemplo 5.1.1. Si $T \in \mathcal{L}(\mathbb{R}^3)$ está definida por $T(x,y,z) = (x+y,\, y+z,\, 0)$, entonces los subespacios $W_1 = \{(x,y,0): \ x,y \in \mathbb{R}\}$ y $W_2 = \{(x,0,0): \ x \in \mathbb{R}\}$ son T-invariantes.

Proposición 5.1.2. *Los subespacios* $\{0\}$*,* V*,* $\operatorname{Ker}(T)$ *e* $\operatorname{Im}(T)$ *son* T*-invariantes. Si* $\lambda \in \Bbbk$ *es un valor propio de* T*, entonces el subespacio propio* $E_\lambda = \operatorname{Ker}(T - \lambda I)$ *es* T*-invariante.*

Dem. Ejercicio. □

Observación 5.1.3. Si T es diagonalizable, entonces es $V = \bigoplus_{i=1}^h E_{\lambda_i}$, siendo $\lambda_1, \ldots, \lambda_h$ los distintos valores propios de T. Luego si $\mathcal{B}$ es una base de V formada uniendo bases de $E_{\lambda_1}, \ldots, E_{\lambda_h}$, entonces

$$[T]_{\mathcal{B}} = \begin{pmatrix} A_1 & & \\ & \ddots & \\ & & A_h \end{pmatrix}, \quad A_i = \lambda_i I_{n_i},\ \forall i = 1, \ldots, h,$$

siendo n_i la multiplicidad de λ_i, para todo $i = 1, \ldots, h$.

El siguiente resultado generaliza la observación anterior. Su prueba es directa.

Proposición 5.1.4. *Supongamos que tenemos una descomposición* $V = W_1 \bigoplus \cdots \bigoplus W_h$ *en que cada* W_i *es* T*-invariante. Si* $\mathcal{B}_i$ *es una base de* W_i*,* $i = 1, \ldots, h$*, y definimos* $\mathcal{B} = \mathcal{B}_1 \cup \cdots \cup \mathcal{B}_h$*, entonces* $[T]_{\mathcal{B}}$ *es una matriz en bloques de la forma*

$$[T]_{\mathcal{B}} = \begin{pmatrix} A_1 & & \\ & \ddots & \\ & & A_h \end{pmatrix}, \quad A_i = [T|_{W_i}]_{\mathcal{B}_i},\ \forall i = 1, \ldots, h. \quad \square$$

Ejemplo 5.1.5. Consideremos $T \in \mathcal{L}(\mathbb{R}^3)$ definida por $T(x,y,z) = (x-z, 2y, x+z)$, para todo $(x,y,z) \in \mathbb{R}^3$. Notar que valen $T(x,0,z) = (x-z,0,x+z)$ y $T(0,y,0) = (0,2y,0)$, luego $W_1 = \{(x,y,z) \in \mathbb{R}^3 : \ y = 0\}$ y $W_2 = \{(x,y,z) \in \mathbb{R}^3 : \ x = z = 0\}$ son subespacios T-invariantes. Si $\{e_1, e_2, e_3\}$ es la base canónica de $\mathbb{R}^3$, entonces es claro que $\mathcal{B}_1 = \{e_1, e_3\}$ es base de W_1 y $\mathcal{B}_2 = \{e_2\}$ es base de W_2. Si consideramos $\mathcal{B} = \mathcal{B}_1 \cup \mathcal{B}_2 = \{e_1, e_3, e_2\}$, entonces obtenemos

$$[T]_{\mathcal{B}} = \begin{pmatrix} 1 & -1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}, \quad [T|_{W_1}]_{\mathcal{B}_1} = \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}, \quad [T|_{W_2}]_{\mathcal{B}_2} = \begin{pmatrix} 2 \end{pmatrix}$$

Notar que es $\chi_T(t) = -(t-2)\big(t^2 - 2t + 2\big)$ y este polinomio no se escinde, luego T no es diagonalizable.

5.2. Polinomios evaluados en operadores

Definimos $T^n \in \mathcal{L}(V)$, $n = 0, 1, \ldots$, mediante $T^0 = \mathrm{Id}$, $T^n = \underbrace{T \circ \cdots \circ T}_{n}$, $\forall n = 1, 2, \ldots$

Si $p(t) = \sum_{i=0}^n a_i \, t^i \in \Bbbk[t]$, definimos $p(T) \in \mathcal{L}(V)$ mediante

$$p(T) = \sum_{i=0}^n a_i T^i = a_n \, T^n + a_{n-1} \, T^{n-1} + \cdots + a_1 \, T + a_0 \mathrm{Id}.$$

Ejemplo 5.2.1. Si consideramos los polinomios $p(t) = 1$ y $q(t) = t$, entonces $p(T) = \mathrm{Id}$ y $q(T) = T$.

Proposición 5.2.2. *Sean $\lambda \in \Bbbk$ y $p(t), q(t) \in \Bbbk[t]$. Si $r(t) = \lambda \, p(t) + q(t)$ y $s(t) = p(t) q(t)$, entonces*

$$r(T) = \lambda \, p(T) + q(T), \qquad s(T) = p(T) \circ q(T).$$

Dem. Podemos suponer $p(t) = \sum_{i=0}^n a_i \, t^i$ y $q(t) = \sum_{i=0}^n b_i \, t^i$ (puede ser $a_n = 0$ o $b_n = 0$).

Es $r(t) = \sum_{i=0}^n (\lambda \, a_i + b_i) t^i$, luego

$$r(T) = \sum_{i=0}^n (\lambda \, a_i + b_i) \, T^i = \sum_{i=0}^n \lambda \, a_i \, T^i + \sum_{i=0}^n b_i \, T^i = \lambda \, p(T) + q(T).$$

Es $s(t) = a_n \, b_n \, t^{2n} + (a_{n-1} \, b_1 + a_1 \, b_{n-1}) \, t^{2n-1} + \cdots + (a_1 \, b_0 + a_0 \, b_1) \, t + a_0 \, b_0$, luego

$$\begin{aligned} s(T) &= a_n \, b_n \, T^{2n} + (a_{n-1} \, b_1 + a_1 \, b_{n-1}) \, T^{2n-1} + \cdots + (a_1 \, b_0 + a_0 \, b_1) \, T + a_0 \, b_0 \, \mathrm{Id} \\ &= \left(\sum_{i=0}^n a_i \, T^i \right) \left(\sum_{i=0}^n b_i \, T^i \right) = p(T) \, q(T). \quad \square \end{aligned}$$

Como el producto de polinomios es conmutativo, la proposición anterior implica el siguiente resultado.

Corolario 5.2.3. *Si $p, q \in \Bbbk[t]$, entonces $p(T) \circ q(T) = q(T) \circ p(T)$.* $\square$

Proposición 5.2.4. *Si $p(t) \in \Bbbk[t]$, entonces* $\mathrm{Ker} \ p(T)$ *e* $\mathrm{Im} \ p(T)$ *son subespacios T-invariantes.*

Dem. Sea $v \in \mathrm{Ker} \ p(T)$, entonces

$$p(T)(T(v)) = (p(T) \circ T)(v) = (T \circ p(T))(v) = T(p(T)(v)) = T(0) = 0 \ \Rightarrow \ T(v) \in \mathrm{Ker} \ p(T).$$

Sea $u = p(T)(w) \in \mathrm{Im} \ p(T)$, entonces

$$T(u) = T(p(T)(w)) = (T \circ p(T)) \, (w) = (p(T) \circ T) \, (w) = p(T)(T(w)) \ \Rightarrow \ T(u) \in \mathrm{Im} \ p(T). \quad \square$$

Proposición 5.2.5. *Existe un polinomio no nulo $p(t)$ tal que $p(T) = 0$.*

Dem. Como $\mathcal{L}(V)$ tiene dimensión finita, entonces el conjunto $\{\mathrm{Id}, T, T^2, \dots\}$ es LD y por lo tanto existe una cantidad finita de escalares no todos nulos $a_0, a_1, \dots, a_n$ tales que

$$a_0 \mathrm{Id} + a_1 T + \cdots + a_n T^n = 0;$$

luego vale $p(T) = 0$, siendo $p(t) = a_0 + a_1 t + \cdots + a_n t^n$. □

Ejemplo 5.2.6. Si T es una proyección ($T^2 = T$), entonces vale $p(T) = 0$, siendo $p(t) = t^2 - t$.

Observación 5.2.7. Observar que si vale $p(T) = 0$ y multiplicamos a $p(t)$ por otro polinomio cualquiera $q(t)$, entonces el polinomio resultante $r(t) = p(t)q(t)$ verifica $r(T) = p(T) \circ q(T) = 0 \circ q(T) = 0$. Luego el polinomio que se anula en T no es único.

Los resultados que siguen permiten descomponer V en suma directa de subespacios T-invariantes.

Proposición 5.2.8. *Sea $p(t) \in \Bbbk[t]$ tal que $p(T) = 0$. Si escribimos $p(t) = p_1(t)\, p_2(t)$ con $p_1(t)$ y $p_2(t)$ tales que* $\operatorname{mcd}(p_1(t), p_2(t)) = 1$*, entonces*

$$V = \mathrm{Ker}\ p_1(T) \oplus \mathrm{Ker}\ p_2(T).$$

Dem. Como $\operatorname{mcd}(p_1(t), p_2(t)) = 1$, entonces (proposición 7.1.4) existen $m_1(t)$ y $m_2(t)$ en $\Bbbk[t]$ tales que $m_1(t)\, p_1(t) + m_2(t)\, p_2(t) = 1$, luego

$$\mathrm{Id} = m_1(T) \circ p_1(T) + m_2(T) \circ p_2(T) = p_1(T) \circ m_1(T) + p_2(T) \circ m_2(T).$$

Sea $v \in V$, entonces $v = \mathrm{Id}(v) = p_1(T)\big(m_1(T)(v)\big) + p_2(T)\big(m_2(T)(v)\big)$ luego

$$V = \mathrm{Im}\, p_1(T) + \mathrm{Im}\, p_2(T).$$

Observar que es $0 = p(T) = p_1(T) \circ p_2(T)$. Entonces para todo v en V es $0 = p_1(T)(p_2(T)(v))$ y resulta $\mathrm{Im}\, p_2(T) \subset \mathrm{Ker}\ p_1(T)$. Análogamente de $0 = p(T) = p_2(T) \circ p_1(T)$ se deduce $\mathrm{Im}\, p_1(T) \subset \mathrm{Ker}\ p_2(T)$. Luego

$$V = \mathrm{Im}\, p_1(T) + \mathrm{Im}\, p_2(T) \subset \mathrm{Ker}\ p_2(T) + \mathrm{Ker}\ p_1(T) \subset V,$$

de donde deducimos

$$V = \mathrm{Ker}\ p_1(T) + \mathrm{Ker}\ p_2(T).$$

Si $v \in \mathrm{Ker}\ p_1(T) \cap \mathrm{Ker}\ p_2(T)$ es $p_1(T)(v) = p_2(T)(v) = 0$, entonces

$$v = \mathrm{Id}(v) = m_1(T)(p_1(T)(v)) + m_2(T)(p_2(T)(v)) = m_1(T)(0) + m_2(T)(0) = 0.$$

Luego $\mathrm{Ker}\ p_1(T) \cap \mathrm{Ker}\ p_2(T) = \{0\}$. □

Ejemplo 5.2.9. Volviendo al caso en que T es una proyección, sabemos que vale $p(T) = 0$, siendo $p(t) = t^2 - t$. Como $p(t)$ se factoriza $p(t) = t(t-1)$ y claramente t y $t-1$ son primos entre sí, deducimos

$$V = \mathrm{Ker}\,(T) \oplus \mathrm{Ker}\,(T - \mathrm{Id}),$$

siendo $\mathrm{Ker}\,(T - \mathrm{Id}) = \{v \in V : \ T(v) = v\}$. Notar que alguno de estos subespacios puede ser nulo, lo cual corresponde a los casos $T = 0$ y $T = \mathrm{Id}$. Observar también que esto implica que T es diagonalizable.

Teorema 5.2.10. *Sea $p(t) \in \Bbbk[t]$ tal que $p(T) = 0$. Si escribimos $p(t) = p_1(t) \cdots p_h(t)$ con $p_1(t), \dots, p_h(t)$ tales que* $\operatorname{mcd}(p_i(t), p_j(t)) = 1$*, $\forall i \neq j$, entonces*

$$V = \mathrm{Ker}\ p_1(T) \oplus \cdots \oplus \mathrm{Ker}\ p_h(T).$$

Dem. Lo demostraremos por inducción en h. Para $h = 2$ es la proposición anterior.

Supongamos que se cumple para $h-1$ y sea $p(t) = p_1(t)\cdots p_{h-1}(t)\,p_h(t)$ con $\operatorname{mcd}(p_i(t), p_j(t)) = 1$, $\forall i \neq j$.

Si escribimos $q(t) = p_1(t)\cdots p_{h-1}(t) \in \Bbbk[t]$, entonces es $p(t) = q(t)\,p_h(t)$ y $\operatorname{mcd}(q(t), p_h(t)) = 1$. Luego el lema anterior aplicado a $q(t)$ y $p_h(t)$ implica

$$V = \operatorname{Ker}\, q(T) \oplus \operatorname{Ker}\, p_h(T).$$

Sea $W = \operatorname{Ker}\, q(T)$. El subespacio W es T-invariante. Consideremos $T|_W \in \mathcal{L}(W)$. Es

$$q(T|_W) = q(T)|_W = q(T)|_{\operatorname{Ker}\, q(T)} = 0.$$

Luego $q(T|_W) = 0$ y por lo tanto podemos aplicar la hipótesis inductiva a $T|_W \in \mathcal{L}(W)$ y $q(t) \in \Bbbk[t]$, para concluir que vale $W = \bigoplus_{i=1}^{h-1} \operatorname{Ker}\, p_i(T|_W)$. Entonces es $V = \left(\bigoplus_{i=1}^{h-1} \operatorname{Ker}\, p_i(T|_W)\right) \oplus \operatorname{Ker}\, p_h(T)$.
Para $i = 1, \ldots, h-1$ es

$$\operatorname{Ker}\, p_i(T|_W) = \operatorname{Ker}\, p_i(T)|_W = W \cap \operatorname{Ker}\, p_i(T) = \operatorname{Ker}\, p_i(T),$$

La última igualdad se deduce de que si $v \in \operatorname{Ker}\, p_i(T)$, entonces

$$q(T)(v) = \big(p_1(T) \circ \cdots \circ p_{h-1}(T)\big)(v) = \big(p_1(T) \circ \cdots \circ p_{i-1}(T) \circ p_{i+1}(T) \circ \cdots \circ p_{h-1}(T)\big)\big(p_i(T)(v)\big) = 0.$$

Luego

$$\operatorname{Ker}\, p_i(T) \subset \operatorname{Ker}\, q(T) = W \quad \Rightarrow \quad \operatorname{Ker}\, p_i(T) \cap W = \operatorname{Ker}\, p_i(T).$$

Entonces $V = \left(\bigoplus_{i=1}^{h-1} \operatorname{Ker}\, p_i(T)\right) \oplus \operatorname{Ker}\, p_h(T) = \bigoplus_{i=1}^{h} \operatorname{Ker}\, p_i(T)$. $\square$

Polinomios evaluados en matrices. Como es de esperar, todo lo que vimos anteriormente para operadores se aplica también a matrices. A continuación desarrollamos brevemente la versión matricial.

Si $A \in M_n(\Bbbk)$ y $p = \sum_{i=0}^{n} a_i\, t^i \in \Bbbk[t]$, entonces definimos $p(A) \in M_n(\Bbbk)$ mediante

$$p(A) = \sum_{i=0}^{n} a_i A^i = a_n\, A^n + a_{n-1}\, A^{n-1} + \cdots + a_1\, A + a_0 I.$$

Proposición 5.2.11. *Sea $T \in \mathcal{L}(V)$ y $\mathcal{B}$ una base de V. Entonces $[p(T)]_{\mathcal{B}} = p([T]_{\mathcal{B}})$, $\forall p \in \Bbbk[t]$.*

Dem. Si $p = \sum_{i=0}^{n} a_i\, t^i$, es $[p(T)]_{\mathcal{B}} = \left[\sum_{i=0}^{n} a_i\, T^i\right]_{\mathcal{B}} = \sum_{i=0}^{n} a_i\, ([T]_{\mathcal{B}})^i = p([T]_{\mathcal{B}})$. $\square$

Corolario 5.2.12. *Si $A \in M_n(\Bbbk)$, es $p(L_A) = L_{p(A)} \in \mathcal{L}(\Bbbk^n)$ para todo $p \in \Bbbk[t]$.*

Dem. Si $\mathcal{B}$ es la base canónica de $\Bbbk^n$, es $A = [L_A]_{\mathcal{B}}$. Aplicando la proposición anterior obtenemos $[p(L_A)]_{\mathcal{B}} = p(A)$, luego $p(L_A) = L_{p(A)}$. $\square$

Ejemplo 5.2.13. Sea $T \in \mathcal{L}(\mathbb{R}^2)$ definida por $T(x, y) = (x, 2x + y)$ y consideremos $p = t^3 + t^2 - t + 2 \in \mathbb{R}[t]$. Es $T = L_A$, siendo $A = \left(\begin{smallmatrix} 1 & 0 \\ 2 & 1 \end{smallmatrix}\right)$. Luego

$$p(A) = A^3 + A^2 - A + 2I = \begin{pmatrix} 3 & 0 \\ 8 & 3 \end{pmatrix} \quad \Rightarrow \quad p(T)(x, y) = (3x,\, 8x + 3y),\ \forall (x, y) \in \mathbb{R}^2.$$

Observación 5.2.14. La proposición 5.2.11 y el corolario 5.2.12 nos permiten deducir propiedades de polinomios aplicados a operadores, a partir de propiedades de polinomios aplicados a matrices, y viceversa. Los dos resultados siguientes se obtienen aplicando esa observación.

Proposición 5.2.15. *Sean $A \in M_n(\Bbbk)$, $\lambda \in \Bbbk$ y $p, q \in \Bbbk[t]$. Si $r(t) = \lambda\, p(t) + q(t)$ y $s(t) = p(t)q(t)$, entonces*

$$r(A) = \lambda\, p(A) + q(A), \qquad s(A) = p(A)\, q(A). \quad \square$$

Corolario 5.2.16. *Si $p, q \in \Bbbk[t]$ y $A \in M_n(\Bbbk)$, entonces $p(A)\, q(A) = q(A)\, p(A)$.* $\square$

5.3. El teorema de Cayley-Hamilton

En la sección anterior probamos que existe un polinomio no nulo $p(t)$ tal que $p(T)=0$. En lo que sigue veremos que el polinomio característico de T verifica esa propiedad.

Proposición 5.3.1. *Sea W un subespacio T-invariante de V. Si $\mathcal{B}_W$ es una base de W y la completamos a una base $\mathcal{B}$ de V, entonces $[T]_{\mathcal{B}} = \left(\begin{smallmatrix} A & B \\ 0 & D \end{smallmatrix}\right)$, siendo $A = [T|_W]_{\mathcal{B}_W}$.*

Dem. Ejercicio. $\square$

Proposición 5.3.2. *Si $W \subset V$ es un subespacio T-invariante, entonces $\chi_{T|_W}(t)$ divide a $\chi_T(t)$.*

Dem. Sea $\mathcal{B}$ una base de V como en la proposición anterior. Entonces es $[T]_{\mathcal{B}} = \left(\begin{smallmatrix} A & B \\ 0 & D \end{smallmatrix}\right)$, siendo $A = [T|_W]_{\mathcal{B}_W}$. Luego aplicando la proposición 7.3.1 obtenemos

$$\chi_T(t) = \begin{vmatrix} A - t\,I & B \\ 0 & D - t\,I \end{vmatrix} = |A - t\,I|\,|D - t\,I| = \chi_{T|_W}(t)\,p(t).$$

Luego $\chi_T(t) = \chi_{T|_W}(t)\,p(t)$, siendo $p(t) = |D - t\,I|$. $\square$

Ejemplo 5.3.3. Sea $T \in \mathcal{L}\big(\mathbb{R}^4\big)$ definida por

$$T(x,y,z,t) = (x + y + 2z - t,\ y + t,\ 2z - t,\ z + t).$$

Consideremos el subespacio $W = \{(x,y,0,0) :\ x,y \in \mathbb{R}\}$. Observar que $T(x,y,0,0) = (x+y,\ y,\ 0,\ 0) \in W$, luego W es T-invariante. Si $\mathcal{B} = \{e_1,e_2,e_3,e_4\}$ es la base canónica de $\mathbb{R}^4$, entonces $\mathcal{B}_W = \{e_1,e_2\}$ es base de W y obtenemos

$$[T|_W]_{\mathcal{B}_W} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}, \quad [T]_{\mathcal{B}} = \left(\begin{array}{cc|cc} 1 & 1 & 2 & -1 \\ 0 & 1 & 0 & 1 \\ \hline 0 & 0 & 2 & -1 \\ 0 & 0 & 1 & 1 \end{array}\right).$$

Luego $\chi_{T|_W}(t) = (1-t)^2$ y $\chi_T(t) = (1-t)^2\big(t^2 - 3t + 3\big)$.

Definición 5.3.4. Sea $v \in V$. Llamamos *subespacio T-cíclico* generado por v a

$$S_{v,T} := \big[v,\ T(v),\ T^2(v),\ \ldots\big] = \left\{\sum_{i=0}^{n} a_i\,T^i(v) :\ a_i \in \Bbbk,\ i = 0,\ldots,n,\ n \in \mathbb{N}\right\}.$$

Proposición 5.3.5. *El subespacio $S_{v,T}$ es el menor subespacio T-invariante de V que contiene a v.*

Dem. Es claro que vale $v \in S_{v,T}$. Por otro lado $T\big(\sum_{i=0}^n a_i\,T^i(v)\big) = \sum_{i=0}^n a_i\,T^{i+1}(v)$, luego $S_{v,T}$ es T-invariante.

Sea W un subespacio T-invariante que contiene a v. Como W es T-invariante y $v \in W$, entonces $T(v) \in W$. Luego $T^2(v) = (T \circ T)(v) = T(T(v)) \in W$ y por inducción se prueba que $T^n(v) \in W$, para todo $n \in \mathbb{N}$. Como W es un subespacio, esto implica $S_{v,T} = \big[v,\ T(v),\ T^2(v),\ \ldots\big] \subset W$. $\square$

Proposición 5.3.6. *Sea $0 \neq v \in V$, $W = S_{v,T}$ y $h = \dim W$. Entonces:*

1. *El conjunto $\big\{v, T(v), \ldots, T^{h-1}(v)\big\}$ es base de W.*
2. *Si escribimos $T^h(v) = a_0\,v + a_1\,T(v) + \cdots + a_{h-1}T^{h-1}(v)$, entonces*

$$\chi_{T|_W}(t) = (-1)^h\Big(t^h - a_{h-1}\,t^{h-1} - \cdots - a_1\,t - a_0\Big).$$

Dem. Sea j el menor entero positivo tal que $\big\{v, T(v), \ldots, T^j(v)\big\}$ es LD; como $v \neq 0$, es $j \geq 1$. Observar que $\big\{v, T(v), \ldots, T^{j-1}(v)\big\}$ es LI y $\big\{v, T(v), \ldots, T^{j-1}(v), T^j(v)\big\}$ es LD, luego $T^j(v) \in \big[v, T(v), \ldots, T^{j-1}(v)\big]$.

Afirmación: Para todo $p \in \mathbb{N}$ se cumple que $T^{j+p}(v) \in \left[v, T(v), \dots, T^{j-1}(v)\right]$.

Lo probaremos por inducción en p. Si $p = 0$ sabemos que es cierto. Supongamos que $T^{j+p}(v) \in \left[v, T(v), \dots, T^{j-1}(v)\right]$. Entonces

$$T^{j+p+1}(v) = \left(T \circ T^{j+p}\right)(v) = T\left(T^{j+p}(v)\right) \in \left[T(v), T^2(v), \dots, T^j(v)\right].$$

Como es $\left\{T(v), T^2(v), \dots, T^{j-1}(v), T^j(v)\right\} \subset \left[v, T(v), T^2(v), \dots, T^{j-1}(v)\right]$, deducimos que $T^{j+p+1}(v) \in \left[v, T(v), T^2(v), \dots, T^{j-1}(v)\right]$.

Luego $W = \left[v, T(v), \dots, T^{j-1}(v)\right]$, $j = h = \dim W$ y el conjunto $\mathcal{B} = \left\{v, T(v), T^2(v), \dots, T^{h-1}(v)\right\}$ es una base de W. Esto prueba la primer afirmación. Para la segunda, observar que si escribimos $T^h(v) = a_0\, v + a_1\, T(v) + \cdots + a_{h-1} T^{h-1}(v)$, entonces

$$[T|_W]_{\mathcal{B}} = \begin{pmatrix} 0 & 0 & \cdots & 0 & a_0 \\ 1 & 0 & \cdots & 0 & a_1 \\ 0 & 1 & \cdots & 0 & a_2 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & a_{h-2} \\ 0 & 0 & \cdots & 1 & a_{h-1} \end{pmatrix} \quad \Rightarrow \quad \chi_{T|_W}(t) = \begin{vmatrix} -\lambda & 0 & \cdots & 0 & a_0 \\ 1 & -\lambda & \cdots & 0 & a_1 \\ 0 & 1 & \cdots & 0 & a_2 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \cdots & -\lambda & a_{h-2} \\ 0 & 0 & \cdots & 1 & a_{h-1} - \lambda \end{vmatrix}.$$

Usando esta última fórmula se obtiene la fórmula para $\chi_{T|_W}(t)$ de la tesis (la misma se puede probar por inducción en h, desarrollando el determinante por la primera fila o columna). $\square$

Teorema 5.3.7 (Cayley-Hamilton)**.** *El polinomio característico de T se anula en T,* i. e. $\chi_T(T) = 0$.

Dem. Lo que tenemos que probar es que $\chi_T(T) : V \to V$ verifica $\chi_T(T)\,(v) = 0$ para todo v en V.

Sea $v \in V$ arbitrario fijo. Si $v = 0$ el resultado es obvio. Supongamos ahora $v \neq 0$ y sea $W = S_{v,T}$. Como W es T-invariante, entonces $\chi_{T|_W}$ divide a χ_T. Sea $p(t) \in \Bbbk[t]$ tal que $\chi_T(t) = p(t)\,\chi_{T|_W}(t)$. Sean $h = \dim W$ y $T^h(v) = a_0\, v + a_1\, T(v) + \cdots + a_{h-1}\, T^{h-1}(v)$. Sabemos por la proposición anterior que vale $\chi_{T|_W}(t) = (-1)^h \left(t^h - a_{h-1}\, t^{h-1} - a_1\, t - \cdots - a_0\right)$. Luego

$$\chi_{T|_W}(T)\,(v) = (-1)^h \Big(T^h(v) - a_{h-1}\, T^{h-1}(v) - \cdots a_1\, T(v) - a_0\, v\Big) = 0.$$

Entonces

$$\chi_T(T)\,(v) = \left(p(T) \circ \chi_{T|_W}(T)\right)\,(v) = p(T)\,\left(\chi_{T|_W}(T)\,(v)\right) = p(T)\,(0) = 0. \quad \square$$

Corolario 5.3.8. *Si $A \in M_n(\Bbbk)$, entonces $\chi_A(A) = 0$.* $\square$

Aplicando el teorema 5.2.10 obtenemos el siguiente resultado.

Corolario 5.3.9. *Si el polinomio característico de T se escinde (lo cual siempre sucede si $\Bbbk = \mathbb{C}$) y escribimos $\chi_T(t) = (-1)^n (t - \lambda_1)^{n_1} \cdots (t - \lambda_h)^{n_h}$, para ciertos $\lambda_i \neq \lambda_j$ si $i \neq j$. Entonces*

$$V = \bigoplus_{i=1}^{h} \operatorname{Ker}\,(T - \lambda_i \mathrm{Id})^{n_i}. \quad \square$$

Notar que en la fórmula de arriba los subespacios son T-invariantes, luego se puede aplicar la proposición 5.1.4 para obtener una base $\mathcal{B}$ de V tal que $[T]_{\mathcal{B}}$ sea una matriz diagonal en bloques.

5.4. Polinomio minimal

En las secciones anteriores probamos que siempre existe un polinomio no nulo $p(t)$ tal que $p(T) = 0$; además vimos en el teorema de Cayley-Hamilton que el polinomio caraterístico $\chi_T(t)$ es uno de ellos. Nuestro interés en este tipo de polinomios es porque aplicando el teorema 5.2.10 obtenemos una descomposición del espacio en subespacios T-invariantes y luego aplicando la proposición 5.1.4 obtenemos una base del espacio en la cual la matriz asociada a T es una matriz diagonal en bloques. Como el polinomio $p(t)$ no es único, distintos polinomios pueden dar lugar a distintas descomposiciones. En el corolario 5.3.9 vimos la descomposición asociada al polinomio característico. En esta sección estudiaremos otro polinomio de este tipo, el minimal, que es útil para el estudio de si un operador es diagonalizable, y también ayuda para calcular la forma de Jordan, que veremos en la sección siguiente.

Recordar que un polinomio $p(t)$ es *mónico* si es no nulo y el coeficiente de su término de mayor grado es 1, es decir si es de la forma $p(t) = t^m + a_{m-1}\, t^{m-1} + \cdots + a_1\, t + a_0,$ con $m \geq 1$.

Teorema 5.4.1. *Existe un único polinomio $m_T(t) \in \mathbb{k}[t]$ que verifica:*

1. $m_T(T) = 0$.
2. *$m_T(t)$ es de grado mínimo entre los polinomios no nulos que se anulan en T.*
3. *$m_T(t)$ es mónico.*

Además $m_T(t)$ divide a todo polinomio que se anule en T.

Dem. La proposición 5.2.5 muestra que existe algún polinomio no nulo que se anula en T. Luego tomando un polinomio de grado mínimo que verifique lo anterior y dividiéndolo por el coeficiente de su teŕmino de mayor grado, obtenemos un polinomio $m_T(t)$ que verifica las tres condiciones.

Probaremos ahora que $m_T(t)$ divide a todo polinomio que se anule en T. Sea $q(t) \in \mathbb{k}[t]$ que verifique $q(T) = 0$. Dividimos $q(t)$ por $m_T(t)$ y es $q(t) = m_T(t)\, d(t) + r(t)$ con $r(t) = 0$ o $r(t) \neq 0$ y gr $r(t) <$ gr $m(t)$. Teniendo en cuenta la primer condición obtenemos:

$$r(T) = q(T) - m_T(T) \circ d(T) = 0 - 0 \circ d(T) = 0.$$

Si fuese $r(t) \neq 0$ tendríamos una contradicción con la minimalidad del grado de $m_T(t)$, luego necesariamente es $r(t) = 0$ y por lo tanto $m_T(t)$ divide a $q(t)$.

Ahora probaremos la unicidad. Supongamos ahora que $p(t)$ es otro polinomio que verifica las tres condiciones. Como $p(t)$ verifica la primer condición, entonces $m_T(t)$ divide a $p(t)$. Cambiando los roles de $p(t)$ y $m_T(t)$ obtenemos también que $p(t)$ divide a $m_T(t)$. Luego ambos se dividen mutuamente y por lo tanto existe $a \in \mathbb{k}$ tal que $p(t) = a\, m_T(t)$. Como $p(t)$ y $m(t)$ son mónicos, deducimos $a = 1$; luego $p(t) = m_T(t)$. $\square$

Definición 5.4.2. El polinomio $m_T(t)$ del teorema anterior se llama el *polinomio minimal* de T.

Ejemplos 5.4.3. Los siguientes son casos en que es fácil de hallar el polinomio minimal de $T \in \mathcal{L}(V)$.

1. Es $m_T(t) = t$ si y solo si $T = 0$.
2. Vale $m_T(t) = t - \lambda$, para cierto $\lambda \in \mathbb{k}$, si y solo si $T = \lambda \mathrm{Id}$. En particular $m_{\mathrm{Id}}(t) = t - 1$.
3. El operador T es una proyección tal que $T \neq 0$ y $T \neq \mathrm{Id}$ si y solo si $m_T(t) = t^2 - t$.

Observación 5.4.4. El teorema de Cayley-Hamilton nos dice que el polinomio característico de T se anula en T, luego el polinomio minimal de T divide al característico.

En forma análoga a lo anterior, se prueba que dada una matriz $A \in M_n(\mathbb{k})$ existe un único polinomio $m_A(t) \in \mathbb{k}[t]$ llamado el *polinomio minimal* de A que verifica

1. $m_A(A) = 0$,
2. $m_A(t)$ es de grado mínimo entre los polinomios no nulos que se anulan en A,
3. $m_A(t)$ es mónico.

Además, si $q(t) \in \mathbb{k}[t]$ es tal que $q(A) = 0$, entonces $m_A(t)$ divide a $q(t)$; en particular $m_A(t)$ divide a $\chi_A(t)$.

Proposición 5.4.5. *Sea $\mathcal{B}$ una base de V y $A = [T]_{\mathcal{B}}$. Entonces $m_T(t) = m_A(t)$.*

Dem. Es fácil de probar que vale $p([T]_{\mathcal{B}}) = [p(T)]_{\mathcal{B}}$, para todo $p(t) \in \mathbb{k}[t]$. Usando eso obtenemos

$$\begin{aligned} 0 = m_A(A) = m_A([T]_{\mathcal{B}}) = [m_A(T)]_{\mathcal{B}} \quad &\Rightarrow \quad m_A(T) = 0 \quad \Rightarrow \quad m_T(t)|m_A(t), \\ m_T(A) = m_T([T]_{\mathcal{B}}) = [m_T(T)]_{\mathcal{B}} = [0]_{\mathcal{B}} = 0 \quad &\Rightarrow \quad m_T(A) = 0 \quad \Rightarrow \quad m_A(t)|m_T(t). \end{aligned}$$

Luego existe $a \in \mathbb{k}$ tal que $m_T(t) = a\, m_A(t)$. Como ambos polinomios son mónicos la única posibilidad es $a = 1$ y por lo tanto $m_T(t) = m_A(t)$. $\square$

Corolario 5.4.6. *Si $A \in M_n(\mathbb{k})$, entonces $m_{L_A}(t) = m_A(t)$.*

Dem. Es $[L_A]_{\mathcal{B}} = A$, siendo $\mathcal{B}$ la base canónica de $\mathbb{k}^n$. $\square$

Ejemplo 5.4.7. Si consideremos una matriz escalar $A = \begin{pmatrix} \lambda & & \\ & \ddots & \\ & & \lambda \end{pmatrix} \in M_n(\mathbb{k})$, entonces es $L_A = \lambda \mathrm{Id}$. Luego $m_A(t) = t - \lambda$. Notar que también vale el recíproco: si $m_A(t) = t - \lambda$, entonces $A = \lambda I$.

La proposición y el corolario anteriores implican que los polinomios minimales de las matrices y de los operadores verifican el mismo tipo de propiedades.

Observación 5.4.8. Si $A = \begin{pmatrix} A_1 & & \\ & \ddots & \\ & & A_h \end{pmatrix}$ es una matriz diagonal en bloques, entonces $A^l = \begin{pmatrix} A_1^l & & \\ & \ddots & \\ & & A_h^l \end{pmatrix}$, para todo $l \in \mathbb{N}$. Esto implica $p(A) = \begin{pmatrix} p(A_1) & & \\ & \ddots & \\ & & p(A_h) \end{pmatrix}$, para todo $p(t) \in \mathbb{k}[t]$. En particular esto último vale también para matrices diagonales (el caso en que $A_1, \ldots, A_h$ son matrices 1×1).

Teorema 5.4.9. *El operador T es diagonalizable si y solo si $m_T(t)$ es de la forma*

$$m_T(t) = (t - \lambda_1) \cdots (t - \lambda_h)\,,$$

con $\lambda_i \neq \lambda_j$ si $i \neq j$.

Dem. ($\Rightarrow$) Supongamos que existe $\mathcal{B}$ base de V tal que

$$[T]_{\mathcal{B}} = \begin{pmatrix} \lambda_1 & & & & & & \\ & \ddots & & & & & \\ & & \lambda_1 & & & & \\ & & & \ddots & & & \\ & & & & \lambda_h & & \\ & & & & & \ddots & \\ & & & & & & \lambda_h \end{pmatrix}, \text{con } \lambda_i \neq \lambda_j \text{ si } i \neq j.$$

Dado un polinomio arbitrario $p(t)$, vale

$$[p(T)]_{\mathcal{B}} = p([T]_{\mathcal{B}}) = \begin{pmatrix} p(\lambda_1) & & & & & & \\ & \ddots & & & & & \\ & & p(\lambda_1) & & & & \\ & & & \ddots & & & \\ & & & & p(\lambda_h) & & \\ & & & & & \ddots & \\ & & & & & & p(\lambda_h) \end{pmatrix}. \tag{5.1}$$

Luego

$$p(T) = 0 \quad \Leftrightarrow \quad p(\lambda_1) = \cdots = p(\lambda_h) = 0 \quad \Leftrightarrow \quad \lambda_1, \ldots, \lambda_h \text{ son raíces de } p(t) \quad \Leftrightarrow \quad (t-\lambda_1)\cdots(t-\lambda_h)|p(t).$$

Luego el polinomio mónico de menor grado que se anula en T es $(t-\lambda_1)\cdots(t-\lambda_h)$.

($\Leftarrow$) Supongamos que vale $m_T(t) = (t-\lambda_1)\cdots(t-\lambda_h)$, con $\lambda_i \neq \lambda_j$ si $i \neq j$. Sabemos que $m_T(t)$ verifica $m_T(T) = 0$, y como $\lambda_i \neq \lambda_j$ si $i \neq j$, es $\operatorname{mcd}(t-\lambda_i, t-\lambda_j) = 1$ si $i \neq j$. Entonces aplicando el teorema 5.2.10 a $m_T(t)$ obtenemos

$$V = \operatorname{Ker}(T - \lambda_1 \mathrm{Id}) \oplus \cdots \oplus \operatorname{Ker}(T - \lambda_h \mathrm{Id}) = E_{\lambda_1} \oplus \cdots \oplus E_{\lambda_h}.$$

Esto implica que T es diagonalizable. $\square$

Corolario 5.4.10. *Si existe un polinomio de la forma $p(t) = a(t-\lambda_1)\cdots(t-\lambda_k)$, con $\lambda_i \neq \lambda_j$ si $i \neq j$ y $0 \neq a \in \Bbbk$ tal que $p(T) = 0$, entonces T es diagonalizable.*

Dem. Como $p(T) = 0$, entonces $m_T(t)$ divide a $p(t)$ y por lo tanto se aplica el teorema anterior. $\square$

Ejemplo 5.4.11. Sea $T \in \mathcal{L}(\mathbb{R}_2[x])$ definida por $T(p(x)) = p'(x)$. Observar que vale

$$T\big(ax^2+bx+c\big) = 2ax+b, \quad T^2\big(ax^2+bx+c\big) = 2a, \quad T^3\big(ax^2+bx+c\big) = 0.$$

Luego $T^3 = 0$ y $T^2 \neq 0$. Esto implica $m_T(t) = t^3$ y por lo tanto T no es diagonalizable.

Sabemos que el polinomio minimal $m_T(t)$ divide al característico $\chi_T(t)$. Esto implica que las raíces de $m_T(t)$ son también raíces de $\chi_T(t)$. A continuación veremos que vale también el recíproco.

Lema 5.4.12. *Sea $p(t) \in \Bbbk[t]$ y v un vector propio de T correspondiente a un valor propio λ, entonces*

$$p(T)(v) = p(\lambda)\, v.$$

Dem. Observar que $T(v) = \lambda\, v$ implica $T^2(v) = T(T(v)) = T(\lambda\, v) = \lambda\, T(v) = \lambda^2\, v$. Razonando por inducción se prueba que vale $T^n(v) = \lambda^n\, v$, para todo $n \in \mathbb{N}$. Luego si $p(t) = \sum_{i=0}^{n} a_i\, t^i$, entonces

$$p(T)\,(v) = \left(\sum_{i=0}^{n} a_i\, T^i\right)(v) = \sum_{i=0}^{n} a_i\, T^i(v) = \sum_{i=0}^{n} a_i\, \lambda^i\, v = \left(\sum_{i=0}^{n} a_i\, \lambda^i\right) v = p(\lambda)\, v. \quad \square$$

Teorema 5.4.13. *El polinomio característico de T y el polinomio minimal de T tienen las mismas raíces.*

Dem. Ya vimos que las raíces del polinomio minimal de son también raíces del característico. Veremos ahora el recíproco. Si $\lambda \in \Bbbk$ es tal que $\chi_T(\lambda) = 0$, entonces λ es un valor propio de T y por lo tanto existe $v \neq 0$ tal que $T(v) = \lambda\, v$. Por el lema anterior es $m_T(\lambda)\, v = m_T(T)\,(v) = 0(v) = 0$ y como es $v \neq 0$, deducimos $m_T(\lambda) = 0$. $\square$

Observación 5.4.14. Este teorema implica que si $T \in \mathcal{L}(V)$ es tal que $\chi_T(t)$ escinde y se escribe de la forma

$$\chi_T(t) = (-1)^n (t-\lambda_1)^{n_1} \cdots (t-\lambda_h)^{n_h}\,,\ \text{con } \lambda_i \neq \lambda_j \text{ si } i \neq j,$$

entonces su polinomio minimal es de la forma

$$m_T(t) = (t-\lambda_1)^{m_1} \cdots (t-\lambda_h)^{m_h}\,,\ \text{con } 1 \le m_i \le n_i,\ i = 1, \ldots, h.$$

Ejemplo 5.4.15. Sea $T \in \mathcal{L}(\mathbb{R}^3)$ definida por $T(x, y, z) = (3x - y, 2y, x - y + 2z)$. Notar que es $T = L_A$, siendo $A = \begin{pmatrix} 3 & -1 & 0 \\ 0 & 2 & 0 \\ 1 & -1 & 2 \end{pmatrix}$. Es $\chi_T(t) = -(t-2)^2(t-3)$, luego hay solo dos posibilidades para el minimal

$$m_T(t) = \begin{cases} (t-2)(t-3) \\ (t-2)^2(t-3) \end{cases}.$$

Calculando el producto obtenemos $(A - 2I)(A - 3I) = 0$, luego $m_T(t) = (t-2)(t-3)$. Esto implica que T es diagonalizable y al ser $\chi_T(t) = -(t-2)^2(t-3)$, deducimos que existe una base $\mathcal{B}$ de $\mathbb{R}^3$ tal que $[T]_{\mathcal{B}} = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix}$. Para determinar $\mathcal{B}$ hay que hallar los vectores propios, como vimos en diagonalización.

Lo que sigue es la versión matricial de lo visto anteriormente, que resumimos en la siguiente proposición.

Proposición 5.4.16. *Sea $A \in M_n(\Bbbk)$.*

1. *La matriz A es diagonalizable si y solo si su polinomio minimal es de la forma $(t-\lambda_1)\cdots(t-\lambda_h)$, con $\lambda_i \neq \lambda_j$ si $i \neq j$.*

2. *Si existe un polinomio de la forma $p(t) = a(t-\lambda_1)\cdots(t-\lambda_k) \in \Bbbk[t]$, con $\lambda_i \neq \lambda_j$ si $i \neq j$ y $0 \neq a \in \Bbbk$, tal que $p(A) = 0$, entonces A es diagonalizable.*

3. *Los polinomios $m_A(t)$ y $\chi_A(t)$ tienen las mismas raíces.* $\square$

Ejemplo 5.4.17. Sea $A \in M_n(\mathbb{R})$ tal que $A^3 = A$. Entonces vale $p(A) = 0$, siendo $p(t) = t^3 - t$. Observar que es $p(t) = t^3 - t = t(t-1)(t+1)$, luego A es diagonalizable.

Ejemplo 5.4.18. Veamos cómo determinar todas las matrices reales 2×2 que verifican $A^2 - 3A + 2I = 0$.

Sea $p(t) = t^2 - 3t + 2$. Es $p(A) = 0$, luego $m_A(t)$ divide a $p(t)$. Como $p(t) = (t-1)(t-2)$, entonces $m_A(t)$ puede ser $t-1$, $t-2$ o $(t-1)(t-2)$. Si $m_A(t) = t-1$, entonces $A = I$. Si $m_A(t) = t-2$, entonces $A = 2I$. Si $m_A(t) = (t-1)(t-2)$, entonces A es diagonalizable con valores propios 1 y 2, luego es semejante a $\left(\begin{smallmatrix} 1 & 0 \\ 0 & 2 \end{smallmatrix}\right)$.

Capítulo 6

Forma de Jordan

En el capítulo 1 estudiamos los operadores diagonalizables; en el mismo vimos que un operador es diagonalizable si y solo si su polinomio característico se escinde y para cada uno de sus valores propios la multiplicidad geométrica coincide con la algebraica. En este capítulo estudiaremos los operadores cuyo polinomio característico se escinde, pero que no necesariamente verifican la segunda condición. Esto cubre todos los operadores en espacios complejos (de dimensión finita). Usando esto se cubre también el caso real, el cual está tratado en la sección 7.10 del apéndice. Trabajaremos siempre en un espacio de dimensión finita V sobre un cuerpo arbitrario $\Bbbk$. Las bases serán siempre bases ordenadas. Si $A \in M_n(\Bbbk)$, entonces diremos que n es el *orden* de A.

6.1. Forma de Jordan

Recordemos que un operador $T \in \mathcal{L}(V)$ es diagonalizable si y sólo si existe una base $\mathcal{B}$ de V formada por vectores propios de T. En ese caso si $\mathcal{B} = \{v_1, \dots, v_n\}$, es

$$[T]_{\mathcal{B}} = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}, \quad T(v_i) = \lambda_i \, v_i, \ \forall i = 1, \dots, n.$$

También vimos que no todo operador es diagonalizable, aún si el polinomio característico se escinde. En este capítulo probaremos que vale el siguiente resultado.

Teorema 6.1.1 (Jordan)**.** *Si $T \in \mathcal{L}(V)$ es tal que $\chi_T(t)$ se escinde en $\Bbbk$, entonces existe una base $\mathcal{B}$ de V tal que*

$$[T]_{\mathcal{B}} = \begin{pmatrix} J_1 & & & \\ & J_2 & & \\ & & \ddots & \\ & & & J_h \end{pmatrix}, \ J_i = \begin{pmatrix} \lambda_i & 1 & & & \\ & \lambda_i & 1 & & \\ & & \ddots & \ddots & \\ & & & \lambda_i & 1 \\ & & & & \lambda_i \end{pmatrix} \in M_{p_i}(\Bbbk), \ i = 1, \dots, h. \quad \square \tag{6.1}$$

La prueba de este teorema la veremos más adelante (teorema 6.3.13). Respecto a las fórmulas en (6.1), notar primero que como $[T]_{\mathcal{B}}$ es triangular superior, entonces los escalares $\lambda_1, \dots, \lambda_h$ son los valores propios de T. Cada matriz J_i es un *bloque de Jordan* correspondiente al valor propio λ_i (pueden haber varios bloques correspondientes al mismo valor propio). Cuando ordenamos los bloques que forman $[T]_{\mathcal{B}}$ de forma tal que ponemos juntos los bloques correspondientes a los mismos valores propios y estos los ordenamos en órdenes decrecientes, entonces a $[T]_{\mathcal{B}}$ le llamamos la *forma de Jordan*[1] de T y decimos que $\mathcal{B}$ es una *base de Jordan* de V correspondiente a T.

Observación 6.1.2. Como en $\mathbb{C}$ todo polinomio se escinde, del teorema anterior se deduce que en un espacio vectorial complejo todo operador admite una base de Jordan.

Proposición 6.1.3. *Si existe una base de Jordan correspondiente a T, entonces $\chi_T(t)$ se escinde en $\Bbbk$.*

Dem. Supongamos que existe una base $\mathcal{B}$ de V tal que $[T]_{\mathcal{B}}$ es como en (6.1). Entonces

$$\chi_T(t)=\prod_{i=1}^{h}\det(J_i-t\mathrm{Id})=\prod_{i=1}^{h}(\lambda_i-t)^{p_i}=(-1)^n\prod_{i=1}^{h}(t-\lambda_i)^{p_i}\,,\ \text{siendo } n=\dim V.\quad\square$$

La proposición anterior implica que para que exista una base de Jordan correspondiente a T es necesario que el polinomio característico de T se escinda.

Ejemplo 6.1.4. Una forma de Jordan de un operador es una matriz del tipo

$$A=\begin{pmatrix}2&1&0&&&&&\\0&2&1&&&&&\\0&0&2&&&&&\\&&&2&&&&\\&&&&3&1&&\\&&&&0&3&&\\&&&&&&0&1\\&&&&&&0&0\end{pmatrix}.$$

en el cual estamos omitiendo los ceros fuera de los bloques de Jordan de A. En este caso hay dos bloques de Jordan correspondientes al valor propio 2, uno correspondiente a 3 y otro correspondiente a 0.

Observación 6.1.5. Sea $T\in\mathcal{L}(V)$ y supongamos que tenemos una base de Jordan $\mathcal{B}$ como en el teorema anterior. Mantenemos las notaciones del teorema. Separando la base $\mathcal{B}$ en partes correspondientes a cada bloque J_i obtenemos una descomposición $\mathcal{B}=\mathcal{C}_1\sqcup\cdots\sqcup\mathcal{C}_h$ en que cada subconjunto $\mathcal{C}_i$ tiene p_i elementos. Para simplificar las notaciones, omitiendo el subíndice i podemos escribir cada subconjunto $\mathcal{C}_i$ de la forma $\mathcal{C}=\{v_1,v_2,\dots,v_p\}$, y si escribimos el bloque de Jordan correspondiente de la forma

$$J=\begin{pmatrix}\lambda&1&&&\\&\lambda&1&&\\&&\ddots&\ddots&\\&&&\lambda&1\\&&&&\lambda\end{pmatrix},$$

entonces vale

$$T(v_1)=\lambda v_1,\quad T(v_2)=v_1+\lambda v_2,\quad\cdots\quad,T(v_{p-1})=v_{p-2}+\lambda v_{p-1},\quad T(v_p)=v_{p-1}+\lambda v_p.\tag{6.2}$$

De estas fórmulas se deduce

$$(T-\lambda\mathrm{Id})(v_1)=0,\quad v_1=(T-\lambda\mathrm{Id})(v_2),\quad\cdots,\quad v_{p-2}=(T-\lambda\mathrm{Id})(v_{p-1}),\quad v_{p-1}=(T-\lambda\mathrm{Id})(v_p).$$

[1]Más adelante probaremos que la forma de Jordan es única a menos de reordenar los valores propios, lo cual justifica el referirnos a $[T]_{\mathcal{B}}$ como **la** forma de Jordan de T (ver la observación 7.9.3).

La última fórmula permite obtener v_{p-1} en función de v_p, sustituyendo ese valor en la penúltima fórmula obtenemos v_{p-2} en función de v_p, y siguiendo de esa forma obtenemos todos los vectores $v_1, \ldots, v_{p-1}$ en función de v_p,

$$(T - \lambda\mathrm{Id})^p(v_p) = 0, \quad v_1 = (T - \lambda\mathrm{Id})^{p-1}(v_p), \quad \cdots, \quad v_{p-2} = (T - \lambda\mathrm{Id})^2(v_p), \quad v_{p-1} = (T - \lambda\mathrm{Id})(v_p). \tag{6.3}$$

Luego si escribimos $v = v_p$, entonces obtenemos que $\mathcal{C}$ tiene la forma

$$\mathcal{C} = \left\{(T - \lambda\mathrm{Id})^{p-1}(v),\, (T - \lambda\mathrm{Id})^{p-2}(v),\, \ldots, (T - \lambda\mathrm{Id})^2(v),\, (T - \lambda\mathrm{Id})(v),\, v\right\},\ \text{con}\ (T - \lambda\mathrm{Id})^p(v) = 0.$$

Notar que si $\mathcal{C}$ es como arriba, entonces escribiendo $v_p = v$ y definiendo $v_1, v_2, \ldots, v_{p-1}$ mediante (6.3) obtenemos vectores $v_1, v_2, \ldots, v_p$ que verifican (6.2).

Definición 6.1.6. Sea $T \in \mathcal{L}(V)$ y $\lambda \in \Bbbk$. Un *ciclo* de T correspondiente a λ es un conjunto ordenado de la forma

$$\mathcal{C} = \left\{(T - \lambda\mathrm{Id})^{p-1}(v),\, (T - \lambda\mathrm{Id})^{p-2}(v),\, \ldots, (T - \lambda\mathrm{Id})^2(v),\, (T - \lambda\mathrm{Id})(v),\, v\right\} \tag{6.4}$$

siendo $v \in \operatorname{Ker}(T - \lambda\mathrm{Id})^p \setminus \operatorname{Ker}(T - \lambda\mathrm{Id})^{p-1}$ para cierto $p \geq 1$. La *longitud* de $\mathcal{C}$ es p y su *vector inicial* es $(T - \lambda\mathrm{Id})^{p-1}(v)$.

Observación 6.1.7. Una base de Jordan es una base obtenida como unión disjunta de ciclos de T. Luego para poder encontrarla tenemos que saber hallar los ciclos correspondientes. Los siguientes resultados van en esa dirección.

En lo que sigue $T \in \mathcal{L}(V)$ es un operador arbitrario fijo.

Observación 6.1.8. Si $v \in V$ es tal que $(T - \lambda\mathrm{Id})^k(v) = 0$ para algún $\lambda \in \Bbbk$ y $k \in \mathbb{N}$, entonces

$$(T - \lambda\mathrm{Id})^{k+1}(v) = (T - \lambda\mathrm{Id})\big((T - \lambda\mathrm{Id})^k(v)\big) = (T - \lambda\mathrm{Id})(0) = 0.$$

Luego $\operatorname{Ker}(T - \lambda\mathrm{Id})^k \subset \operatorname{Ker}(T - \lambda\mathrm{Id})^{k+1}$. Esto implica

$$\operatorname{Ker}(T - \lambda\mathrm{Id}) \subset \operatorname{Ker}(T - \lambda\mathrm{Id})^2 \subset \operatorname{Ker}(T - \lambda\mathrm{Id})^3 \subset \cdots$$

Notar que como estamos en dimensión finita, entonces esas inclusiones no pueden ser siempre estrictas. Más adelante veremos que esto está vinculado con la existencia de los ciclos de las bases de Jordan.

Proposición 6.1.9. *Sea $\lambda \in \Bbbk$ y $\mathcal{C}$ un ciclo correspondiente a λ. Entonces λ es un valor propio de T y el vector inicial de $\mathcal{C}$ es un vector propio correspondiente a λ.*

Dem. Escribimos $\mathcal{C}$ como en (6.4). Si $v_1 = (T - \lambda\mathrm{Id})^{p-1}(v)$, entonces de $v \in \operatorname{Ker}(T - \lambda\mathrm{Id})^p \setminus \operatorname{Ker}(T - \lambda\mathrm{Id})^{p-1}$ deducimos $v_1 \neq 0$ y $(T - \lambda\mathrm{Id})(v_1) = (T - \lambda\mathrm{Id})^p(v) = 0$. Luego v_1 es un vector propio de T con valor propio λ. $\square$

Proposición 6.1.10. *Sea λ un valor propio de T y $\mathcal{C}$ un ciclo correspondiente a λ. Entonces $\mathcal{C}$ es un conjunto LI.*

Dem. Supongamos que $\mathcal{C}$ es como en (6.4). Sean $a_0, \ldots, a_{p-1} \in \Bbbk$ tales que

$$a_{p-1}(T - \lambda\mathrm{Id})^{p-1}(v) + \cdots + a_2(T - \lambda\mathrm{Id})^2(v) + a_1(T - \lambda\mathrm{Id})(v) + a_0 v = 0. \tag{6.5}$$

Como $v \in \operatorname{Ker}(T - \lambda\mathrm{Id})^p$, entonces aplicando $(T - \lambda\mathrm{Id})^{p-1}$ a (6.5) obtenemos $a_0(T - \lambda\mathrm{Id})^{p-1}(v) = 0$. Como es $v \notin \operatorname{Ker}(T - \lambda\mathrm{Id})^{p-1}$, concluimos $a_0 = 0$. Luego (6.5) queda en

$$a_{p-1}(T - \lambda\mathrm{Id})^{p-1}(v) + \cdots + a_2(T - \lambda\mathrm{Id})^2(v) + a_1(T - \lambda\mathrm{Id})(v) = 0.$$

Ahora aplicando $(T - \lambda\mathrm{Id})^{p-2}$ a esta expresión y razonando como antes deducimos $a_1 = 0$. Es claro que ese razonamiento podemos seguirlo aplicando hasta obtener $a_0 = \cdots = a_{p-1} = 0$. Luego $\mathcal{C}$ es LI. $\square$

Proposición 6.1.11. *Sean $\mathcal{C}_1, \cdots, \mathcal{C}_h$ ciclos de T correspondientes a un mismo valor propio λ. Si sus vectores iniciales forman un conjunto LI, entonces $\mathcal{C}_1 \cup \cdots \cup \mathcal{C}_h$ es un conjunto LI.*

Dem. Para simplificar la notación lo probaremos solo cuando hay dos ciclos $\mathcal{C}_1$ y $\mathcal{C}_2$ de longitudes pequeñas, pero va a ser claro que la prueba se extiende sin problemas al caso general. Empezamos considerando primero el caso en que $\mathcal{C}_1$ y $\mathcal{C}_2$ tienen la misma longitud y que esa longitud es 3. Luego

$$\mathcal{C}_1 = \left\{(T-\lambda\mathrm{Id})^2(v), (T-\lambda\mathrm{Id})(v),\, v\right\}, \quad \mathcal{C}_2 = \left\{(T-\lambda\mathrm{Id})^2(w), (T-\lambda\mathrm{Id})(w),\, w\right\},$$

siendo $v, w \in \mathrm{Ker}\,(T-\lambda\mathrm{Id})^3 \setminus \mathrm{Ker}\,(T-\lambda\mathrm{Id})^2$. Sean $a_0, a_1, a_2, b_0, b_1, b_2 \in \Bbbk$ tales que

$$a_2(T-\lambda\mathrm{Id})^2(v) + a_1(T-\lambda\mathrm{Id})(v) + a_0 v + b_2(T-\lambda\mathrm{Id})^2(w) + b_1(T-\lambda\mathrm{Id})(w) + b_0 w = 0. \tag{6.6}$$

Aplicando $(T-\lambda\mathrm{Id})^2$ a (6.6) obtenemos $a_0(T-\lambda\mathrm{Id})^2(v) + b_0(T-\lambda\mathrm{Id})^2(w) = 0$. Como por hipótesis estos vectores son LI, deducimos $a_0 = b_0 = 0$. Luego sustituimos estos valores en (6.6) y seguimos razonando como en la proposición anterior hasta obtener que todos los escalares en (6.6) son nulos.

Ahora consideraremos el caso en que $\mathcal{C}_1$ y $\mathcal{C}_2$ tienen distintas longitudes, por ejemplo suponiendo que la longitud de $\mathcal{C}_1$ es 5 y la de $\mathcal{C}_2$ es 3. Luego

$$\begin{aligned}
\mathcal{C}_1 &= \left\{(T-\lambda\mathrm{Id})^4(v), (T-\lambda\mathrm{Id})^3(v), (T-\lambda\mathrm{Id})^2(v), (T-\lambda\mathrm{Id})(v),\, v\right\},\\
\mathcal{C}_2 &= \left\{(T-\lambda\mathrm{Id})^2(w), (T-\lambda\mathrm{Id})(w),\, w\right\},
\end{aligned}$$

siendo $v \in \mathrm{Ker}\,(T-\lambda\mathrm{Id})^5 \setminus \mathrm{Ker}\,(T-\lambda\mathrm{Id})^4$ y $w \in \mathrm{Ker}\,(T-\lambda\mathrm{Id})^3 \setminus \mathrm{Ker}\,(T-\lambda\mathrm{Id})^2$. Sean $a_0, \ldots, a_4, b_0, \ldots, b_2 \in \Bbbk$ tales que

$$\begin{aligned}
a_4(T-\lambda\mathrm{Id})^4(v) + a_3(T-\lambda\mathrm{Id})^3(v) + a_2(T-\lambda\mathrm{Id})^2(v) + a_1(T-\lambda\mathrm{Id})(v) + a_0 v +\\
b_2(T-\lambda\mathrm{Id})^2(w) + b_1(T-\lambda\mathrm{Id})(w) + b_0 w = 0.
\end{aligned} \tag{6.7}$$

En este caso aplicando $(T-\lambda\mathrm{Id})^3$ a la expresión anterior obtenemos

$$a_1(T-\lambda\mathrm{Id})^4(v) + a_0(T-\lambda\mathrm{Id})^3(v) = 0.$$

Como $\mathcal{C}_1$ es LI, deducimos $a_0 = a_1 = 0$. Luego de (6.7) deducimos

$$a_4(T-\lambda\mathrm{Id})^4(v) + a_3(T-\lambda\mathrm{Id})^3(v) + a_2(T-\lambda\mathrm{Id})^2(v) + b_2(T-\lambda\mathrm{Id})^2(w) + b_1(T-\lambda\mathrm{Id})(w) + b_0 w = 0. \tag{6.8}$$

Sea $u = (T-\lambda\mathrm{Id})^2(v)$, entonces la expresión anterior queda

$$a_4(T-\lambda\mathrm{Id})^2(u) + a_3(T-\lambda\mathrm{Id})(u) + a_2 u + b_2(T-\lambda\mathrm{Id})^2(w) + b_1(T-\lambda\mathrm{Id})(w) + b_0 w = 0.$$

Notar $u \in \mathrm{Ker}\,(T-\lambda\mathrm{Id})^3 \setminus \mathrm{Ker}\,(T-\lambda\mathrm{Id})^2$. Luego volvimos al caso $p = q = 3$ que estudiamos anteriormente. Esto implica que todos los escalares en (6.8) son nulos, lo cual termina de probar que $\mathcal{C}_1 \cup \mathcal{C}_2$ es LI. $\square$

El siguiente teorema resume los resultados anteriores y agrega más información

Teorema 6.1.12. *Sean $\mathcal{C}_1, \cdots, \mathcal{C}_h$ ciclos de T correspondientes a valores propios que pueden ser distintos. Si los ciclos correspondientes a los mismos valores propios son tales que sus vectores iniciales forman un conjunto LI, entonces $\mathcal{C}_1 \cup \cdots \cup \mathcal{C}_h$ es un conjunto LI.*

Dem. De nuevo para simplificar la notación vamos a hacer la prueba en un caso particular. Supongamos que tenemos dos ciclos $\mathcal{C}_1$ y $\mathcal{C}_2$ correspondientes a un valor propio λ y tenemos tres ciclos $\mathcal{D}_1$, $\mathcal{D}_2$ y $\mathcal{D}_3$ correspondientes a otro valor propio μ. Como los vectores iniciales de $\mathcal{C}_1$ y $\mathcal{C}_2$ forman un conjunto LI, entonces $\mathcal{C}_1 \cup \mathcal{C}_2$ es un conjunto LI y por la misma razón $\mathcal{D}_1 \cup \mathcal{D}_2 \cup \mathcal{D}_3$ es LI.

Supongamos es

$$\mathcal{C}_1 = \left\{(T-\lambda\mathrm{Id})^{p-1}(v),\, (T-\lambda\mathrm{Id})^{p-2}(v),\, \dots, (T-\lambda\mathrm{Id})^{2}(v),\, (T-\lambda\mathrm{Id})(v),\, v\right\} \tag{6.9}$$

siendo $v \in \operatorname{Ker}(T-\lambda\mathrm{Id})^p \setminus \operatorname{Ker}(T-\lambda\mathrm{Id})^{p-1}$ para cierto $p \geq 1$. Sea W_1 el subespacio generado por $\mathcal{C}_1$. Como $\mathcal{C}_1$ es LI, entonces es base de W_1. Entonces razonando como en la observación 6.1.5 obtenemos que W_1 es T-invariante, y

$$[T|_{W_1}]_{\mathcal{C}_1} = \begin{pmatrix} \lambda & 1 & & & \\ & \lambda & 1 & & \\ & & \ddots & \ddots & \\ & & & \lambda & 1 \\ & & & & \lambda \end{pmatrix} \in M_p(\Bbbk),$$

Esto implica $\chi_{T|W_1}(t) = (-1)^p(t-\lambda)^p$. Como W_1 es T-invariante, entonces $\chi_{T|W_1}(t)$ divide a $\chi_T(t)$, luego es $\chi_T(t) = (t-\lambda)^m q(t)$, para cierto $m \geq p$ y cierto polinomio $q(t)$. Notar que es $\mathcal{C}_1 \subset \operatorname{Ker}(T-\lambda\mathrm{Id})^p \subset \operatorname{Ker}(T-\lambda\mathrm{Id})^m$, luego $\mathcal{C}_1 \subset \operatorname{Ker}(T-\lambda\mathrm{Id})^m$. Aplicando el mismo razonamiento al otro ciclo $\mathcal{C}_2$ concluimos que es $\chi_T(t) = (t-\lambda)^m q(t)$, para cierto m mayor o igual que el máximo de las longitudes de $\mathcal{C}_1$ y $\mathcal{C}_2$, y vale $\mathcal{C}_1 \cup \mathcal{C}_2 \subset \operatorname{Ker}(T-\lambda\mathrm{Id})^m$.

Razonando análogamente con los ciclos $\mathcal{D}_1$, $\mathcal{D}_2$ y $\mathcal{D}_3$ correspondientes a μ, concluimos que el polinomio característico de T se puede factorizar de la forma $\chi_T(t) = (t-\lambda)^m(t-\mu)^r h(t)$, para cierto polinomio $h(t)$ que es primo con $(t-\lambda)^m$ y $(t-\mu)^r$, y que valen $\mathcal{C}_1 \cup \mathcal{C}_2 \subset \operatorname{Ker}(T-\lambda\mathrm{Id})^m$ y $\mathcal{D}_1 \cup \mathcal{D}_2 \cup \mathcal{D}_3 \subset \operatorname{Ker}(T-\mu\mathrm{Id})^r$. Como $(t-\lambda)^m$, $(t-\mu)^r$ y $h(t)$ son polinomios primos entre sí, entonces el teorema 5.2.10 implica que vale $V = \operatorname{Ker}(T-\lambda\mathrm{Id})^m \oplus \operatorname{Ker}(T-\lambda\mathrm{Id})^r \oplus \operatorname{Ker} h(T)$. Luego los subespacios $\operatorname{Ker}(T-\lambda\mathrm{Id})^m$ y $\operatorname{Ker}(T-\lambda\mathrm{Id})^r$ son independientes. Como $\mathcal{C}_1 \cup \mathcal{C}_2 \subset \operatorname{Ker}(T-\lambda\mathrm{Id})^m$ y $\mathcal{D}_1 \cup \mathcal{D}_2 \cup \mathcal{D}_3 \subset \operatorname{Ker}(T-\lambda\mathrm{Id})^r$ son conjuntos LI, concluimos que su unión $\mathcal{C}_1 \cup \mathcal{C}_2 \cup \mathcal{D}_1 \cup \mathcal{D}_2 \cup \mathcal{D}_3$ es también LI. □

Semejanza de matrices y forma de Jordan. Sin referirnos a transformaciones lineales, un *bloque de Jordan* es una matriz de la forma

$$J = \begin{pmatrix} \lambda & 1 & & & \\ & \lambda & 1 & & \\ & & \ddots & \ddots & \\ & & & \lambda & 1 \\ & & & & \lambda \end{pmatrix},$$

siendo $\lambda \in \Bbbk$ un escalar arbitrario, y una *matriz de Jordan* es una matriz diagonal en bloques de la forma $A = \begin{pmatrix} J_1 & & \\ & \ddots & \\ & & J_k \end{pmatrix}$, siendo $J_1, \dots, J_k$ bloques de Jordan arbitrarios.

Definición 6.1.13. Si $A \in M_n(\Bbbk)$ es tal que su polinomio característico $\chi_A(t) \in \Bbbk[t]$ se escinde, entonces definimos la *forma de Jordan* de A como la de $L_A \in \mathcal{L}(\Bbbk^n)$.

Observar que si J es la forma de Jordan de A y $\mathcal{B} = \{v_1, \dots, v_n\}$ es la base de Jordan para L_A correspondiente, entonces es

$$A = QJQ^{-1}, \quad Q = [v_1|\cdots|v_n].$$

Del teorema de Jordan se deduce el siguiente.

Corolario 6.1.14. *Si $A \in M_n(\Bbbk)$ es tal que $\chi_A(t) \in \Bbbk[t]$ se escinde, entonces A es semejante a una matriz de Jordan.* $\square$

Observación 6.1.15. Sean $A, B \in M_n(\mathbb{C})$. El corolario anterior implica que si A y B tienen la misma forma de Jordan, entonces son semejantes. Más adelante en la sección 7.9 se prueba que vale también el recíproco. Luego A y B son semejantes si y solo si tienen la misma forma de Jordan (a menos de reordenar los bloques). Vale un resultado análogo para matrices reales, ver la sección 7.10.

6.2. Cálculo de la forma de Jordan

En esta sección veremos algunos resultados que en dimensiones bajas nos permiten determinar la forma de Jordan. En particular esto simplifica el hallar la base de Jordan.

Observación 6.2.1. Notar que si $A = \begin{pmatrix} A_1 & & \\ & \ddots & \\ & & A_k \end{pmatrix}$ es una matriz en bloques, entonces el rango de A es la suma de los rangos de $A_1, \ldots, A_k$.

Proposición 6.2.2. *Sea $T \in \mathcal{L}(V)$ un operador cuyo polinomio característico se escinde y sean $\lambda_1, \ldots, \lambda_h$ los valores propios distintos de T. Sea $\mathcal{B}$ una base de Jordan correspondiente a T. Ordenamos $\mathcal{B}$ agrupando los ciclos que corresponden a los mismos valores propios. Luego obtenemos*

$$[T]_{\mathcal{B}} = \begin{pmatrix} A_1 & & \\ & \ddots & \\ & & A_h \end{pmatrix},$$

siendo $A_1, \ldots, A_h$ de la forma $A_i = \begin{pmatrix} J_1 & & \\ & \ddots & \\ & & J_k \end{pmatrix}$, en que $J_1, \ldots, J_k$ son bloques de Jordan correspondientes al mismo valor propio λ_i. Entonces para cada $i = 1, \ldots, h$, vale

- *El orden del bloque A_i es la multiplicidad algebraica de λ_i.*
- *La cantidad k de bloques de Jordan contenidos en A_i es la multiplicidad geométrica de λ_i.*

Dem. Sea $n = \dim V$ y n_i el orden de A_i, para cada $i = 1, \ldots, h$. Observar que cada matriz A_i es triangular superior con λ_i en la diagonal principal, luego

$$\chi_T(t) = \chi_{[T]_{\mathcal{B}}}(t) = \prod_{i=1}^{h} \det(A_i - tI) = \prod_{i=1}^{h} (\lambda_i - t)^{n_i} = (-1)^n \prod_{i=1}^{h} (t - \lambda_i)^{n_i}.$$

Como $\lambda_i \neq \lambda_j$ si $i \neq j$, deducimos $n_i = MA(\lambda_i)$, para todo $i = 1, \ldots, h$. Esto prueba la primer afirmación.

Sea $i \in \{1, \ldots, h\}$ arbitrario fijo. Como en A_i están todos los bloques de Jordan correspondientes al valor propio λ_i, obtenemos

$$A_i - \lambda_i I = \begin{pmatrix} \tilde{J}_1 & & \\ & \ddots & \\ & & \tilde{J}_k \end{pmatrix}, \quad \tilde{J}_l = \begin{pmatrix} 0 & 1 & & & \\ & 0 & 1 & & \\ & & \ddots & \ddots & \\ & & & 0 & 1 \\ & & & & 0 \end{pmatrix}, \; l = 1, \ldots, k.$$

Si el bloque $\tilde{J}_l$ tiene orden m_l, entonces $\operatorname{rango}\left(\tilde{J}_l\right) = m_l - 1$. Luego

$$\operatorname{rango}(A_i - \lambda_i I) = \sum_{l=1}^{k} \operatorname{rango}\left(\tilde{J}_l\right) = \sum_{l=1}^{k}(m_l - 1) = \sum_{l=1}^{k} m_l - k = n_i - k,$$

luego la cantidad de bloques de Jordan contenidos en A_i es $k = n_i - \operatorname{rango}(A_i - \lambda_i I)$.

Por otro lado, si $j \neq i$, entonces la matriz $A_j - \lambda_i I$ es triangular superior con $\lambda_j - \lambda_i \neq 0$ en la diagonal principal, luego $A_j - \lambda_i I \in M_{n_j}(\Bbbk)$ es invertible y por lo tanto $\operatorname{rango}(A_j - \lambda_i I) = n_j$, para todo $j \neq i$. Esto implica

$$\begin{aligned}
MG(\lambda_i) &= \dim V - \operatorname{rango}(T - \lambda_i \mathrm{Id}) = n - \operatorname{rango}([T]_{\mathcal{B}} - \lambda_i I) = \sum_{j=1}^{h} n_j - \sum_{j=1}^{h} \operatorname{rango}(A_j - \lambda_i I) \\
&= \sum_{j=1}^{h} n_j - \left(\sum_{j\neq i} \operatorname{rango}(A_j - \lambda_i I) + \operatorname{rango}(A_i - \lambda_i I)\right) = \sum_{j=1}^{h} n_j - \left(\sum_{j\neq i} n_j + \operatorname{rango}(A_i - \lambda_i I)\right) \\
&= n_i - \operatorname{rango}(A_i - \lambda_i I) = k.
\end{aligned}$$

Esto prueba la segunda afirmación. □

Observación 6.2.3. Con las notaciones de la proposición anterior, la suma de los órdenes de los bloques $J_1, \ldots, J_k$ coincide con el orden de A_i (que a su vez coincide con la multiplicidad algebraica de λ_i).

Ejemplo 6.2.4. Sea $T \in \mathcal{L}(\mathbb{R}^3)$ definida por $T(x, y, z) = (3x + y - 2z, -x + 5z, -x - y + 4z)$. Notar que es $T = L_A$, siendo $A = \begin{pmatrix} 3 & 1 & -2 \\ -1 & 0 & 5 \\ -1 & -1 & 4 \end{pmatrix}$. El polinomio característico de T es $\chi_T(t) = -(t-3)(t-2)^2$, luego los valores propios de T son 2 y 3 y la forma de Jordan de T tiene un bloque A_1 de orden 1 correspondiente al valor propio 3 y un bloque A_2 de orden 2 correspondiente al valor propio 2. Es

$$A - 2\,I = \begin{pmatrix} 1 & 1 & -2 \\ -1 & -2 & 5 \\ -1 & -1 & 2 \end{pmatrix} \text{ y } \operatorname{rango}(A - 2\,I) = 2.$$

Luego $\mathrm{MG}(2) = 1$. Esto implica A_2 tiene un solo bloque de Jordan, luego la forma de Jordan de T es

$$J = \begin{pmatrix} 3 & 0 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix}.$$

Ahora obtendremos una base de Jordan $\mathcal{B}$. Como hay dos bloques de Jordan, entonces $\mathcal{B}$ consta de dos ciclos. Al valor propio 3 le corresponde un ciclo de longitud 1 que es lo mismo que un vector propio, y al valor propio 2 le corresponde un ciclo de longitud 2. Luego la base $\mathcal{B}$ tiene la forma

$$\mathcal{B} = \{u,\, (T - 2\mathrm{Id})(v),\, v\}, \quad u \in \operatorname{Ker}(T - 3\mathrm{Id}),\ v \in \operatorname{Ker}(T - 2\mathrm{Id})^2 \setminus \operatorname{Ker}(T - 2\mathrm{Id}).$$

Operando obtenemos

$$\operatorname{Ker}(T - 3\mathrm{Id}) = [(-1, 2, 1)], \quad \operatorname{Ker}(T - 2\mathrm{Id}) = [(-1, 3, 1)], \quad \operatorname{Ker}(T - 2\mathrm{Id})^2 = [(1, 0, 2), (0, 1, 1)].$$

Elegimos $(0, 1, 1) \in \operatorname{Ker}(T - 2\mathrm{Id})^2 \setminus \operatorname{Ker}(T - 2\mathrm{Id})$. Luego

$$\mathcal{B} = \{u,\, (T - 2\mathrm{Id})(v),\, v\} = \{(-1, 2, 1), (-1, 3, 1), (0, 1, 1)\}$$

Notar que $\mathcal{B}$ es unión de ciclos correspondientes a valores propios distintos, luego es LI y por lo tanto $\mathcal{B}$ es una base de Jordan tal que $[T]_{\mathcal{B}} = J$.

Ejemplo 6.2.5. Sea $T \in \mathcal{L}(\mathbb{R}_2[x])$ definida por $T(p(x)) = -p(x) - p'(x)$.

Si $\mathcal{C} = \{x^2, x, 1\}$ es
$$A = [T]_{\mathcal{C}} = \begin{pmatrix} -1 & 0 & 0 \\ -2 & -1 & 0 \\ 0 & -1 & -1 \end{pmatrix}.$$
Luego $\chi_T(t) = -(t+1)^3$ y el único valor propio de T es -1 con $\mathrm{MA}(-1) = 3$. Además,
$$MG(-1) = 3 - \mathrm{rango}(A+I) = 3 - \mathrm{rango}\begin{pmatrix} 0 & 0 & 0 \\ -2 & 0 & 0 \\ 0 & -1 & 0 \end{pmatrix} = 1,$$
luego hay un solo bloque de Jordan correspondiente al valor propio -1 y la forma de Jordan de T es
$$J = \begin{pmatrix} -1 & 1 & 0 \\ 0 & -1 & 1 \\ 0 & 0 & -1 \end{pmatrix}.$$
Esto nos dice que la base de Jordan es un ciclo de longitud 3 de la forma $\mathcal{B} = \big\{(T+\mathrm{Id})^2(v), (T+\mathrm{Id})(v), v\big\}$, siendo $v \in \mathrm{Ker}\,(T+\mathrm{Id})^3 \setminus \mathrm{Ker}\,(T+\mathrm{Id})^2$. Observar que vale
$$(T+\mathrm{Id})\big(p(x)\big) = -p'(x), \quad (T+\mathrm{Id})^2\big(p(x)\big) = p''(x), \quad (T+\mathrm{Id})^3\big(p(x)\big) = 0, \quad \forall p(x) \in \mathbb{R}_2[x].$$
Luego $\mathrm{Ker}\,(T+\mathrm{Id})^3 = \mathbb{R}_2[x]$ y por lo tanto $v \in \mathbb{R}_2[x] \setminus \mathrm{Ker}\,(T+\mathrm{Id})^2$. Como $(T+\mathrm{Id})^2\big(p(x)\big) = p''(x)$, deducimos que $\mathrm{Ker}\,(T+\mathrm{Id})^2 = \{ax+b : \ a, b \in \mathbb{R}\}$ y por lo tanto $x^2 \in \mathbb{R}_2[x] \setminus \mathrm{Ker}\,(T+\mathrm{Id})^2$. Entonces $\mathcal{B} = \{2, -2x, x^2\}$ es una base de Jordan para T y $J = [T]_{\mathcal{B}}$.

Ejemplo 6.2.6. Sea $A = \begin{pmatrix} 2 & 0 & 0 \\ 0 & -1 & 3 \\ 0 & -3 & 5 \end{pmatrix} \in M_3(\mathbb{R})$.

Veremos de hallar su forma de Jordan y la matriz de semejanza correspondiente. Es $\chi_A(t) = -(t-2)^3$, luego 2 es el único valor propio de A. Operando obtenemos $MG(2) = 3 - \mathrm{rango}(A - 2I) = 2$, por lo cual A tiene dos bloques de Jordan. Como la suma de los órdenes de los dos bloques tiene que dar 3, deducimos que la forma de Jordan de A es
$$J = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}.$$
Luego la base de Jordan correspondiente está formada por dos ciclos de longitudes 2 y 1, es decir tiene la forma $\mathcal{B} = \{(A-2\mathrm{Id})(u), u, v\}$ siendo $u \in \mathrm{Ker}\,(A-2\mathrm{Id})^2 \setminus \mathrm{Ker}\,(A-2\mathrm{Id})$ y $v \in \mathrm{Ker}\,(A-2\mathrm{Id})$. Observar que es
$$A - 2I = \begin{pmatrix} 0 & 0 & 0 \\ 0 & -3 & 3 \\ 0 & -3 & 3 \end{pmatrix}, \quad (A-2I)^2 = 0,$$
luego
$$\mathrm{Ker}\,(A-2\mathrm{Id}) = [(1,0,0), (0,1,1)], \quad \mathrm{Ker}\,(A-2\mathrm{Id})^2 = \mathbb{R}^3$$
Así que es $u \in \mathbb{R}^3 \setminus \mathrm{Ker}\,(A-2\mathrm{Id})$. Podemos tomar como u cualquier vector que no esté en $\mathrm{Ker}\,(A-2\mathrm{Id})$, por ejemplo $u = (0,0,1)$. Es $(A-2\mathrm{Id})(u) = (0,3,3) \in \mathrm{Ker}\,(A-2\mathrm{Id})$. Ahora tenemos que encontrar un vector en $v \in \mathrm{Ker}\,(A-2\mathrm{Id})$ que no sea colineal con $(0,3,3)$, por ejemplo $v = (1,0,0)$. Entonces el teorema 6.1.12 implica que $\mathcal{B} = \{(0,3,3), (0,0,1), (1,0,0)\}$ es LI, luego $\mathcal{B}$ es una base de Jordan tal que $[L_A]_{\mathcal{B}} = J$. Finalmente usando las fórmulas de cambio de base deducimos $A = QJQ^{-1}$, siendo
$$Q = {}_{\mathcal{C}}[\mathrm{Id}]_{\mathcal{B}} = \begin{pmatrix} 0 & 0 & 1 \\ 3 & 0 & 0 \\ 3 & 1 & 0 \end{pmatrix}.$$

En lo que sigue veremos que el polinomio minimal también brinda información sobre la forma de Jordan.

Proposición 6.2.7. *1. Si*

$$J = \begin{pmatrix} \lambda & 1 & & & \\ & \lambda & 1 & & \\ & & \ddots & \ddots & \\ & & & \lambda & 1 \\ & & & & \lambda \end{pmatrix} \in M_n(\Bbbk).$$

Entonces $m_J(t) = (t-\lambda)^n$.

2. Si

$$A = \begin{pmatrix} J_1 & & \\ & \ddots & \\ & & J_k \end{pmatrix} \in M_n(\Bbbk), \quad J_i = \begin{pmatrix} \lambda & 1 & & & \\ & \lambda & 1 & & \\ & & \ddots & \ddots & \\ & & & \lambda & 1 \\ & & & & \lambda \end{pmatrix} \in M_{p_i}(\Bbbk),$$

con $p_1 \geq p_2 \geq \cdots \geq p_k$. Entonces $m_A(t) = (t-\lambda)^{p_1}$.

3. Supongamos

$$A = \begin{pmatrix} A_1 & & \\ & \ddots & \\ & & A_h \end{pmatrix} \in M_n(\Bbbk),$$

en que cada bloque $A_i \in M_{n_i}(\Bbbk)$ es como en la parte anterior correspondiente a cierto escalar λ_i. Supongamos que los λ_i son distintos entre sí y que q_i es el tamaño del primer bloque de Jordan contenido en A_i, para todo i . Entonces

$$m_A(t) = (t-\lambda_1)^{q_1} \cdots (t-\lambda_h)^{q_h}.$$

Dem. (1). Sea

$$\tilde{J} = J - \lambda I = \begin{pmatrix} 0 & 1 & & & \\ & 0 & 1 & & \\ & & \ddots & \ddots & \\ & & & 0 & 1 \\ & & & & 0 \end{pmatrix} \in M_n(\Bbbk).$$

Si consideramos $T = L_{\tilde{J}} \in \mathcal{L}(\Bbbk^n)$, entonces es $T(x_1, x_2, \dots, x_n) = (x_2, \dots, x_n, 0)$, para todo $(x_1, \dots, x_n) \in \Bbbk^n$. Luego

$$\begin{aligned} T(x_1, x_2, \dots, x_n) &= (x_2, \dots, x_n, 0), \\ T^2(x_1, x_2, \dots, x_n) &= (x_3, \dots, x_n, 0, 0), \\ &\vdots \\ T^{n-1}(x_1, x_2, \dots, x_n) &= (x_n, 0, \dots, 0), \\ T^n(x_1, x_2, \dots, x_n) &= (0, \dots, 0). \end{aligned}$$

Lo anterior implica[2] $T^n = 0$ y $T^{n-1} \neq 0$, lo cual equivale a $(J-\lambda I)^n = 0$ y $(J-\lambda I)^{n-1} \neq 0$. De $(J-\lambda I)^n = 0$ deducimos $m_J(t) = (t-\lambda)^r$, para algún $1 \leq r \leq n$. Pero como es $(J-\lambda I)^{n-1} \neq 0$, concluimos $m_J(t) = (t-\lambda)^n$.

[2] Esto también se puede probar calculando las potencias de $\tilde{J}$, observando que en las potencias sucesivas la diagonal de unos va subiendo hasta desaparecer al llegar a $\tilde{J}^n$.

(2). Por la parte anterior sabemos que es $m_{J_i}(t) = (t-\lambda)^{p_i}$, para todo $i = 1, \dots, k$. Como estamos asumiendo $p_1 \geq p_2 \geq \cdots \geq p_k$, es

$$m_{J_1}(J_i) = (J_i - \lambda I)^{p_1} = (J_i - \lambda I)^{p_1 - p_i} \cdot (J_i - \lambda I)^{p_i} = (J_i - \lambda I)^{p_1 - p_i} \cdot 0 = 0, \quad \forall i = 1, \dots, k.$$

Luego

$$m_{J_1}(A) = \begin{pmatrix} m_{J_1}(J_1) & & \\ & \ddots & \\ & & m_{J_1}(J_k) \end{pmatrix} = 0. \tag{6.10}$$

Entonces $m_A(t)$ divide a $m_{J_1}(t) = (t-\lambda)^{p_1}$ y por lo tanto es de la forma $m_A(t) = (t-\lambda)^l$, para cierto $l \leq p_1$. Pero si consideramos $p(t) = (t-\lambda)^l$ con $l < p_1$, entonces $p(J_1) = (J_1 - \lambda I)^l \neq 0$. Luego razonando como en (6.10) obtenemos $p(A) \neq 0$. Esto implica $m_A(t) = (t-\lambda)^{p_1}$.

(3). Como la matriz A es triangular superior, entonces $\chi_A(t) = (-1)^n (t-\lambda_1)^{n_1} \cdots (t-\lambda_h)^{n_h}$, con $\lambda_i \neq \lambda_j$ si $i \neq j$. Luego $m_A(t) = (t-\lambda_1)^{m_1} \cdots (t-\lambda_h)^{m_h}$, siendo $1 \leq m_i \leq n_i$, $\forall i = 1, \dots, h$.

Sea $p(t) = (t-\lambda_1)^{l_1} \cdots (t-\lambda_h)^{l_h}$ un polinomio arbitrario con raíces $\lambda_1, \dots, \lambda_h$. Es

$$p(A) = \begin{pmatrix} p(A_1) & & \\ & \ddots & \\ & & p(A_h) \end{pmatrix}.$$

Luego

$$\begin{aligned} p(A) = 0 &\Leftrightarrow p(A_i) = 0,\ \forall i = 1, \dots, h \\ &\Leftrightarrow m_{A_i}(t) | \, p(t),\ \forall i = 1, \dots, h \\ &\Leftrightarrow (t-\lambda_i)^{q_i} | (t-\lambda_1)^{l_1} \cdots (t-\lambda_h)^{l_h},\ \forall i = 1, \dots, h \\ &\Leftrightarrow q_i \leq l_i,\ \forall i = 1, \dots, h. \end{aligned}$$

Dado que $m_A(t)$ es el polinomio de grado mínimo que verifica esta condición, deducimos que es $m_A(t) = (t-\lambda_1)^{q_1} \cdots (t-\lambda_h)^{q_h}$. □

Corolario 6.2.8. *Si el polinomio característico de un operador se escinde, entonces para cada valor propio λ, el exponente de $t-\lambda$ en la descomposición factorial del polinomio minimal coincide con el tamaño del mayor bloque de Jordan correspondiente a λ.* □

Ejemplo 6.2.9. Sea $A \in M_3(\mathbb{R})$ tal que $(A-5I)^2 = 0$, $A \neq 5I$. Esto implica $m_A(t) = (t-5)^2$ y por lo tanto $\chi_A(t) = -(t-5)^3$. Sabemos que la forma de Jordan de A tiene orden 3 y en la misma solo hay bloques correspondientes al valor propio 5, y de estos el más grande tiene orden 2. Luego la forma de Jordan es

$$J = \begin{pmatrix} 5 & 1 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & 5 \end{pmatrix}.$$

Ejemplo 6.2.10. Sea $A \in M_4(\mathbb{R})$ tal que $(A-5I)^2 = 0$, $A \neq 5I$. Como antes deducimos $m_A(t) = (t-5)^2$, luego $\chi_A(t) = (t-5)^4$. Por la forma del polinomio minimal de A, sabemos que el tamaño del mayor bloque de la forma de Jordan de A es 2, luego las posibles formas de Jordan de A son

$$\begin{pmatrix} 5 & 1 & 0 & 0 \\ 0 & 5 & 0 & 0 \\ 0 & 0 & 5 & 1 \\ 0 & 0 & 0 & 5 \end{pmatrix}, \quad \begin{pmatrix} 5 & 1 & 0 & 0 \\ 0 & 5 & 0 & 0 \\ 0 & 0 & 5 & 0 \\ 0 & 0 & 0 & 5 \end{pmatrix}.$$

Para determinar la forma de Jordan necesitamos algún dato más. Por ejemplo, si sabemos el rango de $A-5I$, entonces la forma de Jordan de A es la primera si $\operatorname{rango}(A-5I)=2$ y es la segunda si $\operatorname{rango}(A-5I)=1$.

Ejemplo 6.2.11. Sea $T=L_A\in\mathcal{L}(\mathbb{R}^6)$, siendo

$$A=\begin{pmatrix}3&0&1&0&0&0\\0&3&0&1&0&0\\0&0&3&0&0&0\\0&0&0&3&0&0\\0&0&0&0&8&-9\\0&0&0&0&4&-4\end{pmatrix}.$$

Notar que A es una matriz diagonal en bloques

$$A=\begin{pmatrix}B_1&0\\0&B_2\end{pmatrix},\quad B_1=\begin{pmatrix}3&0&1&0\\0&3&0&1\\0&0&3&0\\0&0&0&3\end{pmatrix},\quad B_2=\begin{pmatrix}8&-9\\4&-4\end{pmatrix}.$$

Luego $\chi_T(t)=\chi_A(t)=\chi_{B_1}(t)\chi_{B_2}(t)=(t-3)^4(t-2)^2$. Esto implica que si $\mathcal{B}$ es una base de Jordan correspondiente, entonces $[T]_\mathcal{B}$ es de la forma

$$[T]_\mathcal{B}=\begin{pmatrix}A_1&0\\0&A_2\end{pmatrix},\quad A_1\in M_4(\mathbb{R}),\ A_2\in M_2(\mathbb{R}),$$

siendo A_1 el bloque correspondiente al valor propio 3 y A_2 el correspondiente al valor propio 2. Calculando obtenemos $\operatorname{rango}(A-2I)=5$ y $\operatorname{rango}(A-3I)=4$. Luego las multiplicidades geométricas de los valores propios son $\operatorname{MG}(2)=1$ y $\operatorname{MG}(3)=2$. De $\operatorname{MG}(2)=1$ deducimos que hay un solo bloque de Jordan correspondiente al valor propio 2, luego $A_2=\left(\begin{smallmatrix}2&1\\0&2\end{smallmatrix}\right)$. Al ser $\operatorname{MG}(3)=2$, sabemos que hay dos bloques de Jordan correspondientes al valor propio 3, luego hay dos posibilidades para A_1 que son las siguientes

$$\begin{pmatrix}3&1&0&0\\0&3&0&0\\0&0&3&1\\0&0&0&3\end{pmatrix},\quad\begin{pmatrix}3&1&0&0\\0&3&1&0\\0&0&3&0\\0&0&0&3\end{pmatrix}.$$

De acuerdo al corolario anterior, el polinomio minimal de T es, en el primer caso $(t-3)^2(t-2)^2$, y en el segundo $(t-3)^3(t-2)^2$. Se verifica fácilmente que vale $(B_1-3I)^2=0$, luego el minimal es el primero y por lo tanto la forma de Jordan es

$$[T]_\mathcal{B}=\begin{pmatrix}3&1&0&0&0&0\\0&3&0&0&0&0\\0&0&3&1&0&0\\0&0&0&3&0&0\\0&0&0&0&2&1\\0&0&0&0&0&2\end{pmatrix}.$$

Esto implica que la base de Jordan $\mathcal{B}$ está formada por tres ciclos de longitud 2, dos de ellos corresponden al valor propio 3 y uno al valor propio 2. Es decir es del tipo

$$\mathcal{B}=\{(T-3\mathrm{Id})(u),\,u,\,(T-3\mathrm{Id})(v),\,v,\,(T-2\mathrm{Id})(w),\,w\}$$

siendo $u, v \in \operatorname{Ker}(T-3\mathrm{Id})^2 \setminus \operatorname{Ker}(T-3\mathrm{Id})$ y $w \in \operatorname{Ker}(T-2\mathrm{Id})^2 \setminus \operatorname{Ker}(T-2\mathrm{Id})$. Realizando los cálculos obtenemos

$$\operatorname{Ker}(T-3\mathrm{Id}) = [e_1, e_2],\ \operatorname{Ker}(T-3\mathrm{Id})^2 = [e_1, e_2, e_3, e_4],\ \operatorname{Ker}(T-2\mathrm{Id}) = [3e_5+2e_6],\ \operatorname{Ker}(T-2\mathrm{Id})^2 = [e_5, e_6],$$

siendo $\{e_1, \dots, e_6\}$ la base canónica de $\mathbb{R}^6$. Si elegimos $u = e_3$, $v = e_4$ y $w = e_5$, entonces es $(T-3\mathrm{Id})(e_3) = e_1$, $(T-3\mathrm{Id})(e_4) = e_2$ y $(T-2\mathrm{Id})(e_5) = 6e_5 + 4e_6$, luego una base de Jordan es

$$\mathcal{B} = \{e_1,\ e_3,\ e_2,\ e_4,\ 6e_5 + 4e_6,\ e_5\}.$$

Observación 6.2.12. Es un ejercicio el verificar que si todos los valores propios del operador tienen multiplicidad algebraica menor que 7, entonces para obtener su forma de Jordan alcanza con aplicar lo que vimos en esta sección. Si alguno tiene multiplicidad algebraica 7, entonces hay solo un caso que no sabemos determinar, que es cuando los órdenes de los bloques de Jordan correspondientes son $(3,2,2)$ y $(3,3,1)$, todos los otros casos se obtienen sin problema. En casos como estos de multiplicidad mayor que 6, lo que se puede hacer es hallar la base de Jordan y a partir de ahí deducir la forma de Jordan. La existencia de la base de Jordan y la forma general para hallarla se ven en la sección siguiente. Más adelante en la sección 7.9 se muestra un algoritmo general para obtener la forma de Jordan sin pasar por hallar la base de Jordan. Este algoritmo también prueba la unicidad de la forma de Jordan.

6.3. Existencia de la base de Jordan

En esta sección probaremos el teorema de Jordan y mostraremos cómo se obtiene una base de Jordan en el caso general. Empezamos considerando el caso en que el operador es nilpotente.

Definición 6.3.1. Un operador $T \in \mathcal{L}(V)$ se dice *nilpotente* si existe un entero positivo k tal que $T^k = 0$. El *orden* de nilpotencia de T es el menor entero k que verifica $T^k = 0$. Luego T es nilpotente de orden p si y solo si $T^p = 0$ y $T^{p-1} \neq 0$.

Análogamente, decimos que una matriz $A \in M_n(\Bbbk)$ es *nilpotente* si existe $k \in \mathbb{Z}^+$ tal que $A^k = 0$; su *orden* de nilpotencia es el menor k tal que $A^k = 0$.

Observación 6.3.2. Si $T \in \mathcal{L}(V)$ y $\mathcal{B}$ es una base de V, entonces es claro que T es nilpotente si y solo si $[T]_{\mathcal{B}}$ es nilpotente; en ese caso T y $[T]_{\mathcal{B}}$ tienen el mismo orden de nilpotencia.

Observación 6.3.3. Es claro que si $T \in \mathcal{L}(V)$, entonces T es nilpotente de orden p si y solo si $m_T(t) = t^p$. Como el polinomio minimal y característico tienen las mismas raíces, concluimos que 0 es el único valor propio de un operador nilpotente.

Ejemplo 6.3.4. Sea $T \in \mathcal{L}(\mathbb{R}_2[x])$ definida por $T(p(x)) = p'(x)$. En el ejemplo 5.4.11 vimos que T es nipotente de orden 3 y por lo tanto $m_T(t) = t^3$; luego $\chi_T(t) = -t^3$.

Lo siguiente muestra que un operador nilpotente es en cierto sentido lo opuesto a ser diagonalizable

Proposición 6.3.5. *Si $T \in \mathcal{L}(V)$ es nilpotente y diagonalizable, entonces $T = 0$.*

Dem. Si T es nilpotente, entonces $m_T(t) = t^p$, para cierto $p \geq 1$. Por otro lado si T es diagonalizable entonces $m_T(t)$ no tiene raíces múltiples. Luego la única posibilidad es $p = 1$, lo cual equivale a $T = 0$. □

Para evitar discutir casos triviales, al trabajar con un operador nilpotente asumiremos generalmente que no es el operador nulo, luego su orden de nilpotencia es mayor o igual a 2. Lo que vamos a probar es que si $T \in \mathcal{L}(V)$ es nilpotente, entonces existe una base $\mathcal{B}$ de V tal que la matriz asociada tiene la forma

$$[T]_{\mathcal{B}} = \begin{pmatrix} J_1 & & & \\ & J_2 & & \\ & & \ddots & \\ & & & J_h \end{pmatrix}, \quad J_i = \begin{pmatrix} 0 & 1 & & & \\ & 0 & 1 & & \\ & & \ddots & \ddots & \\ & & & 0 & 1 \\ & & & & 0 \end{pmatrix}, \quad i = 1, \dots, h.$$

Empezamos con una proposición que prueba el recíproco de la afirmación anterior y además brinda información sobre la cantidad de bloques que aparecen en la descomposición.

Proposición 6.3.6. *Si*

$$A = \begin{pmatrix} J_1 & & \\ & \ddots & \\ & & J_k \end{pmatrix} \in M_n(\Bbbk), \quad J_i = \begin{pmatrix} 0 & 1 & & & \\ & 0 & 1 & & \\ & & \ddots & \ddots & \\ & & & 0 & 1 \\ & & & & 0 \end{pmatrix} \in M_{p_i}(\Bbbk), \ \forall i = 1, \dots, k. \tag{6.11}$$

siendo $p_1 \geq p_2 \geq \cdots \geq p_k$, entonces A es nilpotente de orden p_1 y la cantidad de bloques contenidos en A es $k = n - \operatorname{rango}(A)$.

Dem. La primer afirmación se deduce de la proposición 6.2.7, que en este caso implica $m_A(t) = t^{p_1}$. Para la segunda, aplicando la proposición 6.2.2 sabemos que la cantidad de bloques contenidos en A es la multiplicidad geométrica de 0 como valor propio de A, lo cual coincide con $n - \operatorname{rango}(A)$. $\square$

A continuación veremos que, dado un operador nilpotente, siempre se puede obtener una base del espacio de forma tal que su matriz asociada sea como en (6.11). Los siguientes resultados van en esa dirección.

Proposición 6.3.7. *Sea $T \in \mathcal{L}(V)$ un operador nilpotente de orden p. Entonces*

$$\{0\} \neq \operatorname{Ker}(T) \subsetneq \cdots \subsetneq \operatorname{Ker}\left(T^{p-1}\right) \subsetneq \operatorname{Ker}(T^p) = V. \tag{6.12}$$

Dem. Observemos primero que $\operatorname{Ker}(T^m) \subset \operatorname{Ker}\left(T^{m+1}\right)$, $\forall m \in \mathbb{N}$. En efecto, si $w \in \operatorname{Ker}(T^m)$ es

$$T^m(w) = 0 \quad \Rightarrow \quad T^{m+1}(w) = T(T^m(w)) = T(0) = 0 \quad \Rightarrow \quad w \in \operatorname{Ker}\left(T^{m+1}\right).$$

Afirmación: Si para algún l es $\operatorname{Ker}\left(T^{l+1}\right) = \operatorname{Ker}\left(T^l\right)$, entonces $\operatorname{Ker}\left(T^{l+m}\right) = \operatorname{Ker}\left(T^l\right), \quad \forall m \geq 1$.

Lo probaremos por indución en m. Si $m = 1$ es la hipótesis. Supongamos ahora que para algún $m \geq 1$ es $\operatorname{Ker}\left(T^l\right) = \operatorname{Ker}\left(T^{l+m}\right)$. Queremos probar $\operatorname{Ker}\left(T^l\right) = \operatorname{Ker}\left(T^{l+m+1}\right)$. Por la observación anterior $\operatorname{Ker}\left(T^l\right) = \operatorname{Ker}\left(T^{l+m}\right) \subset \operatorname{Ker}\left(T^{l+m+1}\right)$, así que solo falta probar la otra inclusión. Sea $w \in \operatorname{Ker}\left(T^{l+m+1}\right)$,

$$\begin{aligned} &0 = T^{l+m+1}(w) = T^{l+1}(T^m(w)) \quad \Rightarrow \quad T^m(w) \in \operatorname{Ker}\left(T^{l+1}\right) = \operatorname{Ker}\left(T^l\right) \quad \Rightarrow \\ &0 = T^l\big(T^m(w)\big) = T^{l+m}(w) \quad \Rightarrow \quad w \in \operatorname{Ker}\left(T^{l+m}\right) = \operatorname{Ker}\left(T^l\right) \quad \Rightarrow \quad w \in \operatorname{Ker}\left(T^l\right). \end{aligned}$$

Esto concluye la prueba de la afirmación.

Observemos que como T es nilpotente de orden p, entonces $T^p = 0$ y $T^{p-1} \neq 0$. Luego $\operatorname{Ker}\left(T^{p-1}\right) \subsetneq \operatorname{Ker}(T^p) = V$ y la afirmación anterior implica que necesariamente $\operatorname{Ker}\left(T^{l-1}\right) \subsetneq \operatorname{Ker}\left(T^l\right)$, para todo $l = 1, 2, \ldots, p-1$. De acá se deduce inmediatamente (6.12). $\square$

Lema 6.3.8. *Sea $T \in \mathcal{L}(V)$ un operador nilpotente de orden[3] p.*
Si $2 \leq q \leq p$ y tenemos un conjunto $\{v_1, \ldots, v_k\} \subset \operatorname{Ker}(T^q)$ que es LI y $[v_1, \ldots, v_k] \cap \operatorname{Ker}\left(T^{q-1}\right) = \{0\}$, entonces $\{T(v_1), \ldots, T(v_k)\} \subset \operatorname{Ker}\left(T^{q-1}\right)$, es LI y $[T(v_1), \ldots, T(v_k)] \cap \operatorname{Ker}\left(T^{q-2}\right) = \{0\}$.

Dem. Como $v_i \in \operatorname{Ker}(T^q)$, es

$$0 = T^q(v_i) = T^{q-1}(T(v_i)) \quad \Rightarrow \quad T(v_i) \in \operatorname{Ker}\left(T^{q-1}\right), \ \forall i = 1, \ldots, k,$$

luego $\{T(v_1), \ldots, T(v_k)\} \subset \operatorname{Ker}\left(T^{q-1}\right)$.

Afirmación 1: vale $[T(v_1), \ldots, T(v_k)] \cap \operatorname{Ker}(T^{q-2}) = \{0\}$.

Sea $w \in [T(v_1), \ldots, T(v_k)] \cap \operatorname{Ker}(T^{q-2})$, luego existen $a_i \in \mathbb{k}$, $i = 1, \ldots, k$ tales que $w = \sum_{i=1}^k a_i\, T(v_i)$ y $w \in \operatorname{Ker}\left(T^{q-2}\right)$. Entonces

$$0 = T^{q-2}(w) = T^{q-2}\left(\sum_{i=1}^k a_i\, T(v_i)\right) = T^{q-1}\left(\sum_{i=1}^k a_i\, v_i\right) \Rightarrow$$
$$\sum_{i=1}^k a_i\, v_i \in \operatorname{Ker}\left(T^{q-1}\right) \cap [v_1, \ldots, v_k] = \{0\}.$$

luego $\sum_{i=1}^k a_i\, v_i = 0$ y por lo tanto $w = \sum_{i=1}^k a_i\, T(v_i) = T\left(\sum_{i=1}^k a_i\, v_i\right) = T(0) = 0$.

Afirmación 2: el conjunto $\{T(v_1), \ldots, T(v_k)\}$ es LI.

Sean $b_i \in \mathbb{k}$, $i = 1, \ldots, k$ tales que $\sum_{i=1}^k b_i\, T(v_i) = 0$. Entonces

$$0 = \sum_{i=1}^k b_i\, T(v_i) = T\left(\sum_{i=1}^k b_i\, v_i\right) \quad \Rightarrow \quad \sum_{i=1}^k b_i\, v_i \in \operatorname{Ker} T \subset \operatorname{Ker}\left(T^{q-1}\right),$$

luego

$$\sum_{i=1}^k b_i\, v_i \in \operatorname{Ker}\left(T^{q-1}\right) \cap [v_1, \ldots, v_k] = \{0\} \quad \Rightarrow \quad \sum_{i=1}^k b_i\, v_i = 0.$$

Como $\{v_1, \ldots, v_k\}$ es LI, resulta $b_i = 0$, $\forall i = 1, \ldots, k$. Esto prueba que $\{T(v_1), \ldots, T(v_k)\}$ es LI. $\square$

Lema 6.3.9. *Supongamos que W es un subespacio de V y $\{v_1, \ldots, v_k\} \subset V$ es un conjunto LI tales que $W \cap [v_1, \ldots, v_k] = \{0\}$. Entonces existen[4] vectores $u_1, \ldots, u_h \in V$ tales que $\{v_1, \ldots, v_k, u_1, \ldots, u_h\}$ es LI y $V = W \oplus [v_1, \ldots, v_k, u_1, \ldots, u_h]$.*

Dem. Supongamos $W \oplus [v_1, \ldots, v_k] \subsetneq V$. Si $\{w_1, \ldots, w_l\}$ es una base de W, entonces $W \cap [v_1, \ldots, v_k] = \{0\}$ implica que el conjunto $\{w_1, \ldots, w_l, v_1, \ldots, v_k\}$ es LI y $l + k < \dim V$. Entonces existen $u_1, \ldots, u_h$ tales que $\{w_1, \ldots, w_l, v_1, \ldots, v_k, u_1, \ldots, u_h\}$ es base de V. Luego

$$V = [w_1, \ldots, w_l] \oplus [v_1, \ldots, v_k, u_1, \ldots, u_h] = W \oplus [v_1, \ldots, v_k, u_1, \ldots, u_h]. \quad \square$$

El siguiente es la versión del teorema de Jordan para operadores nilpotentes.

[3] Recordar que estamos asumiendo $T \neq 0$, luego $p \geq 2$.
[4] Estamos asumiendo que puede ser $h = 0$, es decir que puede valer $V = W \oplus [v_1, \ldots, v_k]$.

Teorema 6.3.10. *Si $T \in \mathcal{L}(V)$, entonces T es nilpotente si y solo si existe una base $\mathcal{B}$ de V tal que*

$$[T]_{\mathcal{B}} = \begin{pmatrix} J_1 & & \\ & \ddots & \\ & & J_k \end{pmatrix}, \quad J_i = \begin{pmatrix} 0 & 1 & & & \\ & 0 & 1 & & \\ & & \ddots & \ddots & \\ & & & 0 & 1 \\ & & & & 0 \end{pmatrix} \in M_{p_i}(\mathbb{k}),\ \forall i = 1, \ldots, k. \tag{6.13}$$

Dem. El recíproco ya lo vimos es la proposición 6.3.6, así que lo único que hay que probar es el directo. Sea T un operador nilpotente de orden p. Aplicando la proposición 6.3.7 sabemos que vale

$$\{0\} \neq \operatorname{Ker}(T) \subsetneq \cdots \subsetneq \operatorname{Ker}(T^p) = V.$$

Para simplificar la notación supondremos $p = 3$, pero la demostración es completamente general. Tenemos

$$\{0\} \neq \operatorname{Ker}(T) \subsetneq \operatorname{Ker}\left(T^2\right) \subsetneq \operatorname{Ker}\left(T^3\right) = V.$$

Consideremos $\operatorname{Ker}\left(T^2\right) \subsetneq \operatorname{Ker}\left(T^3\right) = V$. Sea $\{v_1, \ldots, v_m\}$ un conjunto LI tal que

$$\operatorname{Ker}\left(T^3\right) = \operatorname{Ker}\left(T^2\right) \oplus [v_1, \ldots, v_m].$$

Entonces $\{v_1, \ldots, v_m\}$ es LI y $[v_1, \ldots, v_m] \cap \operatorname{Ker}\left(T^2\right) = \{0\}$, luego el lema 6.3.8 implica

$$\{T(v_1), \ldots, T(v_m)\} \subset \operatorname{Ker}\left(T^2\right), \quad \{T(v_1), \ldots, T(v_m)\} \text{ es LI} \quad \text{y} \quad [T(v_1), \ldots, T(v_m)] \cap \operatorname{Ker}(T) = \{0\}.$$

Consideremos $\operatorname{Ker}(T) \subsetneq \operatorname{Ker}\left(T^2\right)$. Por el lema 6.3.9 sabemos que existen $u_1, \ldots, u_q$ en $\operatorname{Ker}\left(T^2\right)$ tales que:

$$\{T(v_1), \ldots, T(v_m), u_1, \ldots, u_q\} \text{ es LI y } \operatorname{Ker}\left(T^2\right) = \operatorname{Ker}(T) \oplus [T(v_1), \ldots, T(v_m), u_1, \ldots, u_q].$$

Luego $[T(v_1), \ldots, T(v_m), u_1, \ldots, u_q] \cap \operatorname{Ker}(T) = \{0\}$.

Ahora repetimos el procedimiento anterior con $\{T(v_1), \ldots, T(v_m), u_1, \ldots, u_q\}$ en lugar de $\{v_1, \ldots, v_m\}$:

Primero aplicamos el lema 6.3.8 para deducir que el conjunto $\{T^2(v_1), \ldots, T^2(v_m), T(u_1), \ldots, T(u_q)\}$ está contenido en $\operatorname{Ker}(T)$ y es LI. Luego existen $w_1, \ldots, w_r$ en $\operatorname{Ker}(T)$ tales que

$$\left\{T^2(v_1), \ldots, T^2(v_m), T(u_1), \ldots, T(u_q), w_1, \ldots, w_r\right\}$$

es base de $\operatorname{Ker}(T)$. Sean

$$\begin{aligned} \mathcal{B}_1 &= \{v_1, \ldots, v_m\}, \\ \mathcal{B}_2 &= \{T(v_1), \ldots, T(v_m), u_1, \ldots, u_q\}, \\ \mathcal{B}_3 &= \left\{T^2(v_1), \ldots, T^2(v_m), T(u_1), \ldots, T(u_q), w_1, \ldots, w_r\right\} \end{aligned}$$

Tenemos que:

$$V = \operatorname{Ker}\left(T^3\right) = \operatorname{Ker}\left(T^2\right) \oplus [\mathcal{B}_1], \qquad \operatorname{Ker}\left(T^2\right) = \operatorname{Ker}(T) \oplus [\mathcal{B}_2], \qquad \operatorname{Ker}(T) = [\mathcal{B}_3].$$

Luego, $V = [\mathcal{B}_1] \oplus [\mathcal{B}_2] \oplus [\mathcal{B}_3]$ y $\mathcal{B} = \mathcal{B}_1 \cup \mathcal{B}_2 \cup \mathcal{B}_3$ es base de V. Reordenando la base $\mathcal{B}$ se tiene:

$$\mathcal{B} = \left\{T^2(v_1), T(v_1), v_1, \ldots, T^2(v_m), T(v_m), v_m, T(u_1), u_1, \ldots, T(u_q), u_q, w_1, \ldots, w_r\right\}$$

Observar que $\mathcal{B}_3 \subset \operatorname{Ker}(T)$, esto implica:

$$T^3(v_1) = \cdots = T^3(v_m) = 0, \quad T^2(u_1) = \cdots = T^2(u_q) = 0, \quad T(w_1) = \cdots = T(w_r) = 0.$$

Luego[5]

$$[T]_{\mathcal{B}} = \begin{pmatrix} 0 & 1 & 0 &&&&&&&&&&&& \\ 0 & 0 & 1 &&&&&&&&&&&& \\ 0 & 0 & 0 &&&&&&&&&&&& \\ &&& \ddots &&&&&&&&&&& \\ &&&& 0 & 1 & 0 &&&&&&&& \\ &&&& 0 & 0 & 1 &&&&&&&& \\ &&&& 0 & 0 & 0 &&&&&&&& \\ &&&&&&& 0 & 1 &&&&&& \\ &&&&&&& 0 & 0 &&&&&& \\ &&&&&&&&& \ddots &&&&& \\ &&&&&&&&&& 0 & 1 &&& \\ &&&&&&&&&& 0 & 0 &&& \\ &&&&&&&&&&&& 0 && \\ &&&&&&&&&&&&& \ddots & \\ &&&&&&&&&&&&&& 0 \end{pmatrix}. \quad \square$$

Observación 6.3.11. De acuerdo a la proposición 6.3.6, la cantidad de bloques de $[T]_{\mathcal{B}}$ en (6.13), coincide con $\dim V - \operatorname{rango}(T) = \dim \operatorname{Ker}(T)$.

Corolario 6.3.12. *Si $T \in \mathcal{L}(V)$ y $\dim V = n$, entonces T es nilpotente si y solo si $\chi_T(t) = (-1)^n t^n$.*

Dem. El recíproco es inmediato a partir del teorema de Cayley-Hamilton. El directo se deduce inmediatamente de que en (6.13) la matriz $[T]_{\mathcal{B}}$ es triangular superior con ceros en la diagonal principal. $\square$

A continuación probaremos la versión general del teorema de Jordan.

Teorema 6.3.13. *Si $T \in \mathcal{L}(V)$, entonces existe una base de Jordan correspondiente a T si y solo si $\chi_T(t)$ se escinde en $\Bbbk$.*

Dem. El directo ya lo probamos en la observación 6.1.3. Consideremos ahora el recíproco. Supongamos que vale $\chi_T(t) = (-1)^n (t-\lambda_1)^{n_1} \cdots (t-\lambda_h)^{n_h}$, con $\lambda_i \neq \lambda_j$ si $i \neq j$. Como vimos en el corolario 5.3.9, esta factorización de $\chi_T(t)$ da lugar a una descomposición de V en subespacios T-invariantes

$$V = \bigoplus_{i=1}^{h} W_i, \quad \text{siendo } W_i = \operatorname{Ker}(T - \lambda_i \mathrm{Id})^{n_i},\ i = 1, \ldots, h.$$

Luego la proposición 5.1.4 implica que si $\mathcal{B}_i$ es una base de W_i, entonces $\mathcal{B} = \mathcal{B}_1 \cup \cdots \cup \mathcal{B}_h$ es una base de V y la matriz asociada a T en $\mathcal{B}$ tiene la forma

$$[T]_{\mathcal{B}} = \begin{pmatrix} A_1 && \\ & \ddots & \\ && A_h \end{pmatrix}, \quad \text{siendo } A_i = [T|_{W_i}]_{\mathcal{B}_i},\ \forall i = 1, \ldots, h. \tag{6.14}$$

A continuación veremos cómo obtener $\mathcal{B}_i$ en forma adecuada.

[5]Notar que en $[T]_{\mathcal{B}}$ siempre van a haber bloques de orden 3, dado que es $\operatorname{Ker}(T^2) \subsetneq \operatorname{Ker}(T^3) = V$; pero podría ser $q = 0$ o $r = 0$, en cuyo caso no habrían bloques de orden 2 o 1, respectivamente.

Empezamos observando que $(T|_{W_i} - \lambda_i \operatorname{Id}_{W_i})^{n_i} = \left(T - \lambda_i \operatorname{Id}|_{W_i}\right)^{n_i} = (T - \lambda_i \operatorname{Id})^{n_i}|_{W_i} = 0$. Luego $T|_{W_i} - \lambda_i \operatorname{Id}_{W_i} \in \mathcal{L}(W_i)$ es nilpotente. Entonces aplicando el teorema 6.3.10 a $T|_{W_i} - \lambda_i \operatorname{Id}_{W_i}$ deducimos que existe $\mathcal{B}_i$ base de W_i tal que

$$[T|_{W_i} - \lambda_i \operatorname{Id}_{W_i}]_{\mathcal{B}_i} = \begin{pmatrix} \tilde{J}_1^i & & \\ & \ddots & \\ & & \tilde{J}_{k_i}^i \end{pmatrix}, \quad \text{donde } \tilde{J}_j^i = \begin{pmatrix} 0 & 1 & & & \\ & 0 & 1 & & \\ & & \ddots & \ddots & \\ & & & 0 & 1 \\ & & & & 0 \end{pmatrix}, \ \forall j = 1, \dots, k_i.$$

Observar que $T|_{W_i} = T|_{W_i} - \lambda_i \operatorname{Id}_{W_i} + \lambda_i \operatorname{Id}_{W_i}$, luego $[T|_{W_i}]_{\mathcal{B}_i} = [T|_{W_i} - \lambda_i \operatorname{Id}_{W_i}]_{\mathcal{B}_i} + \lambda_i I$ y por lo tanto

$$A_i = [T|_{W_i}]_{\mathcal{B}_i} = \begin{pmatrix} J_1^i & & \\ & \ddots & \\ & & J_{k_i}^i \end{pmatrix}, \quad \text{donde } J_j^i = \begin{pmatrix} \lambda_i & 1 & & & \\ & \lambda_i & 1 & & \\ & & \ddots & \ddots & \\ & & & \lambda_i & 1 \\ & & & & \lambda_i \end{pmatrix}, \ \forall j = 1, \dots, k_i.$$

Finalmente, si $\mathcal{B} = \mathcal{B}_1 \cup \cdots \cup \mathcal{B}_h$, es

$$[T]_{\mathcal{B}} = \begin{pmatrix} A_1 & & \\ & \ddots & \\ & & A_h \end{pmatrix}, \quad A_i = \begin{pmatrix} J_1^i & & \\ & \ddots & \\ & & J_{k_i}^i \end{pmatrix}, \quad J_j^i = \begin{pmatrix} \lambda_i & 1 & & & \\ & \lambda_i & 1 & & \\ & & \ddots & \ddots & \\ & & & \lambda_i & 1 \\ & & & & \lambda_i \end{pmatrix}, \tag{6.15}$$

para todo $i = 1, \dots, h$ y $j = 1, \dots, k_i$. □

Ejemplo 6.3.14. Sea $T = L_A \in \mathcal{L}(\mathbb{R}^7)$, siendo A la matriz diagonal bloques siguiente

$$A = \begin{pmatrix} 2 & 1 & 0 & & & & \\ 0 & 2 & 0 & & & & \\ 1 & 0 & 2 & & & & \\ & & & 2 & 0 & 1 & 0 \\ & & & 0 & 2 & 0 & 0 \\ & & & 0 & 0 & 2 & 0 \\ & & & 1 & 0 & 0 & 2 \end{pmatrix}.$$

El polinomio característico es $\chi_T(t) = -(t-2)^7$. Operando obtenemos

$$A - 2I = \begin{pmatrix} 0 & 1 & 0 & & & & \\ 0 & 0 & 0 & & & & \\ 1 & 0 & 0 & & & & \\ & & & 0 & 0 & 1 & 0 \\ & & & 0 & 0 & 0 & 0 \\ & & & 0 & 0 & 0 & 0 \\ & & & 1 & 0 & 0 & 0 \end{pmatrix}, \quad (A-2I)^2 = \begin{pmatrix} 0 & 0 & 0 & & & & \\ 0 & 0 & 0 & & & & \\ 0 & 1 & 0 & & & & \\ & & & 0 & 0 & 0 & 0 \\ & & & 0 & 0 & 0 & 0 \\ & & & 0 & 0 & 0 & 0 \\ & & & 0 & 0 & 1 & 0 \end{pmatrix}, \quad (A-2I)^3 = 0.$$

Luego es $m_A(t) = (t-2)^3$ y $7 - \operatorname{rango}(A - 2I) = 7 - 4 = 3$. Esto nos dice que hay tres bloques de Jordan correspondientes al valor propio 2 y el bloque mayor tiene tamaño 3. Entonces los órdenes de los bloques de Jordan pueden ser $(3, 2, 2)$ o $(3, 3, 1)$. Para determinarlo vamos a hallar una base de Jordan. Sabemos que es

$$\{0\} \subsetneq \operatorname{Ker}(T - 2\operatorname{Id}) \subsetneq \operatorname{Ker}(T - 2\operatorname{Id})^2 \subsetneq \operatorname{Ker}(T - 2\operatorname{Id})^3 = \mathbb{R}^7.$$

Operando obtenemos que una base de $\operatorname{Ker}(T-2\mathrm{Id})$ es $\{e_3, e_5, e_7\}$ y una base de $\operatorname{Ker}(T-2\mathrm{Id})^2$ es $\{e_1, e_3, e_4, e_5, e_7\}$. Para obtener una base de Jordan tenemos que completar $\{e_1, e_3, e_4, e_5, e_7\}$ a una base de $\mathbb{R}^7$. La opción más fácil es agregarle $\{e_2, e_6\}$. Sabemos que cada uno de esos vectores genera un ciclo de longitud 3, luego los órdenes de los bloques de Jordan son $(3,3,1)$ y por lo tanto la forma de Jordan de T es

$$J = \begin{pmatrix} 2 & 1 & 0 & & & & \\ 0 & 2 & 1 & & & & \\ 0 & 0 & 2 & & & & \\ & & & 2 & 1 & 0 & \\ & & & 0 & 2 & 1 & \\ & & & 0 & 0 & 2 & \\ & & & & & & 2 \end{pmatrix}.$$

Para hallar la base de Jordan calculamos los ciclos anteriores

$$\{(A-2I)^2(e_2),\, (A-2I)(e_2),\, e_2\} = \{e_3,\, e_1,\, e_2\}; \quad \{(A-2I)^2(e_6),\, (A-2I)(e_6),\, e_6\} = \{e_7,\, e_4,\, e_6\}.$$

Notar que e_3 y e_7 –que son los vectores iniciales de estos ciclos– están en $\operatorname{Ker}(T-2\mathrm{Id})$. Luego para obtener la base de Jordan alcanza con agregar un vector de $\operatorname{Ker}(T-2\mathrm{Id})$ que sea linealmente independiente con estos dos. La opción más fácil es agregar e_5, luego

$$\mathcal{B} = \{e_3,\, e_1,\, e_2,\, e_7,\, e_4,\, e_6,\, e_5\}$$

es una base de $\mathbb{R}^7$ tal que $[T]_{\mathcal{B}} = J$.

Observación 6.3.15. Supongamos que $T \in \mathcal{L}(V)$ es tal que $\chi_T(t)$ se escinde y se escribe de la forma $\chi_T(t) = (-1)^n \prod_{i=1}^{h}(t-\lambda_i)^{n_i}$, con $\lambda_i \neq \lambda_j$ si $i \neq j$. Luego su minimal es de la forma $m_T(t) = \prod_{i=1}^{h}(t-\lambda_1)^{m_i}$, con $1 \leq m_i \leq n_i$, para todo $i = 1, \ldots, h$. Como vale $\chi_T(T) = m_T(T) = 0$, sabemos que esto implica

$$V = \bigoplus_{i=1}^{h} \operatorname{Ker}(T-\lambda_i\mathrm{Id})^{n_i} \quad \text{y} \quad V = \bigoplus_{i=1}^{h} \operatorname{Ker}(T-\lambda_i\mathrm{Id})^{m_i}.$$

Veremos que esas dos descomposiciones coinciden. Al ser $m_i \leq n_i$, deducimos $\operatorname{Ker}(T-\lambda_i\mathrm{Id})^{m_i} \subset \operatorname{Ker}(T-\lambda_i\mathrm{Id})^{n_i}$. Pero si recordamos la prueba del teorema 6.3.10, vemos que si $\mathcal{B}$ es una base de Jordan, entonces el máximo de las longitudes de los ciclos en $\mathcal{B}$ correspondientes a λ_i coincide con el índice de nilpotencia de $T-\lambda_i\mathrm{Id}$ en $\operatorname{Ker}(T-\lambda_i\mathrm{Id})^{n_i}$, que a su vez concide con el exponente de $t-\lambda_i$ en $m_T(t)$. Luego es $\operatorname{Ker}(T-\lambda_i\mathrm{Id})^{m_i} = \operatorname{Ker}(T-\lambda_i\mathrm{Id})^{n_i}$, para todo $i = 1, \ldots, h$.

Capítulo 7

Apéndice

7.1. Polinomios

En esta sección veremos algunos resultados sobre polinomios que necesitamos para el estudio de las transformaciones lineales. El cuerpo $\Bbbk$ es arbitrario.

Proposición 7.1.1. *Si $c_1, \ldots, c_k$ elementos distintos de $\Bbbk$, entonces existen $p_1(x), \ldots, p_k(x) \in \Bbbk[x]$ tales que $p_i(c_j) = \delta_{ij}$, para todo $i, j = 1, \ldots, k$.*

Dem. Definimos $p_1(x), \ldots, p_k(x)$ mediante

$$p_i(x) = \prod_{j \neq i} \frac{x - c_j}{c_i - c_j} = \frac{(x - c_1)}{(c_i - c_1)} \cdots \frac{(x - c_{i-1})}{(c_i - c_{i-1})} \frac{(x - c_{i+1})}{(c_i - c_{i+1})} \cdots \frac{(x - c_k)}{(c_i - c_k)}, \quad \forall i = 1, \ldots, k.$$

$$\text{Luego} \quad p_i(c_l) = \prod_{j \neq i} \frac{c_l - c_j}{c_i - c_j} = \left\{ \begin{array}{ll} 0, & l \neq i \\ \prod_{j \neq i} \frac{c_i - c_j}{c_i - c_j} = 1, & l = i \end{array} \right\} = \delta_{il}, \ \forall i, l = 1, \ldots, k. \quad \square$$

Los polinomios $p_1(x), \ldots, p_k(x)$ obtenidos en la proposición anterior se llaman los *polinomios de Lagrange* asociados a los números $c_1, \ldots, c_k$.

Proposición 7.1.2. *Sean $c_1, \ldots, c_k$ elementos distintos de $\Bbbk$ y $b_1, \ldots, b_k$ elementos de $\Bbbk$. Entonces existe un polinomio $p(x) \in \Bbbk[x]$ tal que $p(c_i) = b_i$, para todo $i = 1, \ldots, k$.*

Dem. Sea $p(x) = \sum_{i=1}^k b_i \, p_i(x)$, siendo $p_1(x), \ldots, p_k(x)$ los polinomios de Lagrange asociados a $c_1, \ldots, c_k$. Luego $p(c_l) = \sum_{i=1}^k b_i \, p_i(c_l) = \sum_{i=1}^k b_i \, \delta_{il} = b_l$, para todo $l = 1, \ldots, k$. $\square$

Recordamos algunas definiciones. Un polinomio $p(x)$ es *mónico* si es no nulo y el coeficiente de su término de mayor grado es 1, es decir si es de la forma $p(x) = x^m + a_{m-1} \, x^{m-1} + \cdots + a_1 \, x + a_0$, con $m \geq 1$.

Si dos polinomios $p(x)$ y $q(x)$ verifican que existe un tercer polinomio $r(x)$ tal que $p(x) = q(x) r(x)$, entonces decimos que $q(x)$ *divide* a $p(x)$ y escribimos $q(x) | p(x)$.

Si $p(x)$ y $q(x)$ en $\Bbbk[x]$ son no simultáneamente nulos, entonces el *máximo común divisor* de $p(x)$ y $q(x)$ es el único polinomio mónico $d(x)$ que verifica $d(x)|p(x)$, $d(x)|q(x)$ y si $m(x) \in \Bbbk[x]$ es tal que $m(x)|p(x)$, $m(x)|q(x)$, entonces $m(x)|d(x)$. Escribimos $d(x) = \operatorname{mcd}(p(x), q(x))$.

Observar que si $p(x) \neq 0$ entonces $\operatorname{mcd}(p(x), 0) = \frac{1}{a} \, p(x)$, siendo a el coeficiente del término de mayor grado de $p(x)$.

En los cursos básicos se prueba que si tenemos $p(x)$ y $q(x)$ en $\mathbb{k}[x]$ con $q(x)$ no nulo y gr $p(x) \geq$ gr $q(x)$, entonces $\operatorname{mcd}(p(x), q(x)) = \operatorname{mcd}(q(x), r(x))$, siendo $r(x)$ el resto de dividir $p(x)$ por $q(x)$. Iterando este proceso obtenemos el *algoritmo de Euclides*, que permite obtener el máximo común divisor de dos polinomios:

$$\left.\begin{array}{rcl} p(x) & = & q(x)\, d_1(x) + r_1(x), \quad \text{gr } r_1(x) < \text{gr } q(x) \\ q(x) & = & r_1(x)\, d_2(x) + r_2(x), \quad \text{gr } r_2(x) < \text{gr } r_1(x) \\ r_1(x) & = & r_2(x)\, d_3(x) + r_3(x), \quad \text{gr } r_3(x) < \text{gr } r_2(x) \\ r_2(x) & = & r_3(x)\, d_4(x) + r_4(x), \quad \text{gr } r_4(x) < \text{gr } r_3(x) \\ & \vdots & \\ r_{n-2}(x) & = & r_{n-1}(x)\, d_n(x) + r_n(x), \quad \text{gr } r_n(x) < \text{gr } r_{n-1}(x) \\ r_{n-1}(x) & = & r_n(x)\, d_{n+1}(x), \quad r_{n+1}(x) = 0 \end{array}\right\} \Rightarrow \operatorname{mcd}(p(x), q(x)) = \frac{1}{a} r_n(x). \tag{7.1}$$

siendo a el coeficiente del término de mayor grado de $r_n(x)$.

Ejemplo 7.1.3. Sean $p(x) = x^6 - x^4 + 2\,x^3 + 2\,x^2 + x + 1$ y $q(x) = x^3 + x^2 + x + 1$ en $\mathbb{R}[x]$. Aplicando el algoritmo de Euclides obtenemos

$$\begin{array}{rcl} x^6 - x^4 + 2\,x^3 + 2\,x^2 + x + 1 & = & (x^3 + x^2 + x + 1)(x^3 - x^2 - x + 3) + x^2 - x - 2 \\ x^3 + x^2 + x + 1 & = & (x^2 - x - 2)(x + 2) + 5x + 5 \\ x^2 - x - 2 & = & (5\,x + 5)(\frac{1}{5}x - \frac{2}{5}) \end{array}$$

luego $\operatorname{mcd}(p(x), q(x)) = x + 1$.

Proposición 7.1.4 (Identidad de Bézout)**.** *Sean $p(x)$ y $q(x)$ en $\mathbb{k}[x]$ no simultáneamente nulos. Entonces existen $a(x)$ y $b(x)$ en $\mathbb{k}[x]$ tales que*

$$\operatorname{mcd}(p(x), q(x)) = a(x)\, p(x) + b(x)\, q(x). \tag{7.2}$$

Dem. Si $q(x) = 0$ entonces $\operatorname{mcd}(p(x), 0) = \frac{1}{a}\, p(x)$, siendo a el coeficiente del término de mayor grado de $p(x)$. Luego es $a(x) = \frac{1}{a}$ y $b(x) = 0$.

Supongamos que tenemos $p(x)$ y $q(x)$ en $\mathbb{k}[x]$ no nulos con gr $p(x) \geq$ gr $q(x)$. Consideremos el algoritmo de Euclides (7.1). Probaremos por inducción que para todo $l \in \{1, 2, \ldots, n\}$ existen polinomios $a_l(x)$ y $b_l(x)$ tales que

$$r_l(x) = a_l(x)\, p(x) + b_l(x)\, q(x).$$

Luego la relación (7.2) se obtiene considerando el caso $l = n$ y dividiendo por el coeficiente del término de mayor grado de $r_n(x)$.

Observar que de la primera ecuación de (7.1) deducimos $r_1(x) = d(x)\, p(x) - q(x)$, si llamamos $a_1(x) = d(x)$ y $b_1(x) = -1$, es

$$r_1(x) = a_1(x)\, p(x) + b_1(x)\, q(x).$$

De la segunda ecuación de (7.1) obtenemos

$$\begin{aligned} r_2(x) &= -d_2(x)\, r_1(x) + q(x) = -d_2(x)\, (a_1(x)\, p(x) + b_1(x)\, q(x)) + q(x) \\ &= (-d_2(x)\, a_1(x))\, p(x) + (1 - d_2(x)\, b_1(x))\, q(x). \end{aligned}$$

Si llamamos $a_2(x) = -d_2(x)\, a_1(x)$ y $b_2(x) = 1 - d_2(x)\, b_1(x)$, es

$$r_2(x) = a_2(x)\, p(x) + b_2(x)\, q(x).$$

Razonando inductivamente, sea $h < n$ y supongamos que para todo $l \leq h$ es $r_l(x) = a_l(x)\, p(x) + b_l(x)\, q(x)$. Despejando $r_{h+1}(x)$ en (7.1) obtenemos

$$\begin{aligned}
r_{h+1}(x) &= -r_h(x)\, d_{h+1}(x) + r_{h-1}(x) \\
&= -(a_h(x)\, p(x) + b_h(x)\, q(x))\, d_{h+1}(x) + a_{h-1}(x)\, p(x) + b_{h-1}(x)\, q(x) \\
&= (-a_h(x)\, d_{h+1}(x) + a_{h-1}(x))\, p(x) + (-b_h(x)\, d_{h+1}(x) + b_{h-1}(x))\, q(x) \\
&= a_{h+1}(x)\, p(x) + b_{h+1}(x)\, q(x),
\end{aligned}$$

siendo $a_{h+1} = -a_h(x)\, d_{h+1}(x) + a_{h-1}(x)$ y $b_{h+1}(x) = -b_h(x)\, d_{h+1}(x) + b_{h-1}(x)$. □

Ejemplo 7.1.5. Consideremos de nuevo $p(x) = x^6 - x^4 + 2\,x^3 + 2\,x^2 + x + 1$ y $q(x) = x^3 + x^2 + x + 1$ como en el ejemplo 7.1.3. Ahí vimos que es $\operatorname{mcd}(p(x), q(x)) = x + 1$ y que el algoritmo de Euclides respectivo es

$$\begin{array}{rcl}
x^6 - x^4 + 2\,x^3 + 2\,x^2 + x + 1 &=& (x^3 + x^2 + x + 1)(x^3 - x^2 - x + 3) + x^2 - x - 2 \\
x^3 + x^2 + x + 1 &=& (x^2 - x - 2)(x + 2) + 5x + 5 \\
x^2 - x - 2 &=& (5\,x + 5)(\frac{1}{5}x - \frac{2}{5})
\end{array}$$

Observar que esto lo podemos escribir

$$\begin{array}{rcl}
p(x) &=& q(x)\, d_1(x) + r_1(x), \quad \operatorname{gr} r_1(x) < \operatorname{gr} q(x) \\
q(x) &=& r_1(x)\, d_2(x) + r_2(x), \quad \operatorname{gr} r_2(x) < \operatorname{gr} r_1(x) \\
r_1(x) &=& r_2(x)\, d_3(x), \quad r_3(x) = 0.
\end{array}$$

Despejando $r_2(x)$ de las dos primeras ecuaciones obtenemos

$$r_2(x) = -d_2(x)\, p(x) + (1 + d_1(x)\, d_2(x))\, q(x), \tag{7.3}$$

siendo $r_2(x) = 5x + 5$, $d_1(x) = x^3 - x^2 - x + 3$ y $d_2(x) = x + 2$. Si substituimos $r_2(x)$, $d_1(x)$ y $d_2(x)$ en (7.3) llegamos a

$$5x + 5 = -(x + 2)\, p(x) + (x^4 + x^3 - 3x^2 + x + 7)\, q(x),$$

luego $x + 1 = -\frac{1}{5}(x + 2)\, p(x) + \frac{1}{5}(x^4 + x^3 - 3x^2 + x + 7)\, q(x)$ (verficarlo!).

Proposición 7.1.6. *Sean $p(x)$ y $q(x)$ polinomios no nulos. Entonces vale* $\operatorname{mcd}(p(x), q(x)) = 1$ *si y solo si existen $a(x)$ y $b(x)$ en $\Bbbk[x]$ tales que*

$$a(x)\, p(x) + b(x)\, q(x) = 1.$$

Dem. Si $\operatorname{mcd}(p(x), q(x)) = 1$, entonces la proposición anterior nos prueba que existen $a(x)$ y $b(x)$ en $\Bbbk[x]$ tales que $a(x)\, p(x) + b(x)\, q(x) = 1$.

Supongamos que existen polinomios $a(x)$ y $b(x)$ que verifican $a(x)\, p(x) + b(x)\, q(x) = 1$. Si $m(x)$ divide a $p(x)$ y a $q(x)$, entonces $m(x)$ divide a $a(x)\, p(x) + b(x)\, q(x) = 1$, luego $m(x) \in \Bbbk$. Esto prueba que es $\operatorname{mcd}(p(x), q(x)) = 1$. □

7.2. Matrices elementales

En esta sección $\Bbbk$ es un cuerpo arbitrario.

Definición 7.2.1. Sea $A \in M_{m \times n}$. Se llaman *operaciones elementales* en A a las siguientes:

1. *Operación de tipo I.* Intercambiar dos filas o columnas de A.

2. *Operación de tipo II.* Multiplicar una fila o columna de A por una constante no nula.

3. *Operación de tipo III.* Sumarle a una fila o columna un múltiplo de otra fila o columna, respectivamente.

Definición 7.2.2. Se llaman *matrices elementales* de tipo I, II o III a las matrices que se obtienen aplicando las operaciones elementales correspondientes a la matriz identidad.

Es decir que las matrices elementales son las siguientes:

$$\text{Tipo I:}\quad \begin{pmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & 0 & \cdots & 1 & & \\ & & \vdots & \ddots & \vdots & & \\ & & 1 & \cdots & 0 & & \\ & & & & & \ddots & \\ & & & & & & 1 \end{pmatrix};\quad \text{Tipo II:}\quad \begin{pmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & 1 & & & & \\ & & & p & & & \\ & & & & 1 & & \\ & & & & & \ddots & \\ & & & & & & 1 \end{pmatrix},\quad p \neq 0;$$

$$\text{Tipo III:}\quad \begin{pmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & 1 & \cdots & q & & \\ & & & \ddots & \vdots & & \\ & & & & 1 & & \\ & & & & & \ddots & \\ & & & & & & 1 \end{pmatrix} \quad \text{o} \quad \begin{pmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & 1 & & & & \\ & & \vdots & \ddots & & & \\ & & q & \cdots & 1 & & \\ & & & & & \ddots & \\ & & & & & & 1 \end{pmatrix},\quad q \in \Bbbk.$$

Proposición 7.2.3. *Realizar una operación elemental de tipo I, II o III en las filas (columnas) de una matriz A, equivale a multiplicar a A por la izquierda (derecha) con una matriz elemental del mismo tipo.* □

Ejemplo 7.2.4. Sea $A = \begin{pmatrix} 1 & 2 & 3 \\ 5 & 6 & 7 \\ 9 & 1 & 2 \\ 4 & 5 & 6 \end{pmatrix}$. Consideremos $E_1 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$ matriz elemental de tipo I. Multiplicando obtenemos $E_1 \cdot A = \begin{pmatrix} 1 & 2 & 3 \\ 9 & 1 & 2 \\ 5 & 6 & 7 \\ 4 & 5 & 6 \end{pmatrix}$. Luego multiplicar la matriz A por la izquierda con la matriz E_1 equivale a intercambiar las filas 2 y 3 de A. Por otro lado considerando $E_2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{pmatrix}$ matriz elemental de tipo III, obtenemos $A \cdot E_2 = \begin{pmatrix} 1 & 2 & 1 \\ 5 & 6 & 1 \\ 9 & 1 & 1 \\ 4 & 5 & 1 \end{pmatrix}$. Luego multiplicar la matriz A por la derecha con la matriz E_2 equivale a restarle la columna 2 a la columna 3.

Corolario 7.2.5. *Las matrices elementales de tipo I, II y III son invertibles y sus inversas son también matrices elementales del mismo tipo. Explícitamente*

Matrices de tipo I:

$$\begin{pmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & 0 & \cdots & 1 & & \\ & & \vdots & \ddots & \vdots & & \\ & & 1 & \cdots & 0 & & \\ & & & & & \ddots & \\ & & & & & & 1 \end{pmatrix}^{-1} = \begin{pmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & 0 & \cdots & 1 & & \\ & & \vdots & \ddots & \vdots & & \\ & & 1 & \cdots & 0 & & \\ & & & & & \ddots & \\ & & & & & & 1 \end{pmatrix}.$$

Matrices de tipo II:

$$\begin{pmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & 1 & & & & \\ & & & p & & & \\ & & & & 1 & & \\ & & & & & \ddots & \\ & & & & & & 1 \end{pmatrix}^{-1} = \begin{pmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & 1 & & & & \\ & & & p^{-1} & & & \\ & & & & 1 & & \\ & & & & & \ddots & \\ & & & & & & 1 \end{pmatrix}, \quad p \neq 0.$$

Matrices de tipo III:

$$\begin{pmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & 1 & \cdots & q & & \\ & & & \ddots & \vdots & & \\ & & & & 1 & & \\ & & & & & \ddots & \\ & & & & & & 1 \end{pmatrix}^{-1} = \begin{pmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & 1 & \cdots & -q & & \\ & & & \ddots & \vdots & & \\ & & & & 1 & & \\ & & & & & \ddots & \\ & & & & & & 1 \end{pmatrix}, \quad q \in \Bbbk.$$

La fórmula para la inversa de la otra matriz de tipo III es análoga. □

Observación 7.2.6. 1. La traspuesta de una matriz elemental, es una matriz elemental del mismo tipo.

2. Si realizar una operación elemental en las columnas de A corresponde a multiplicar por la derecha a A con una cierta matriz elemental E, entonces realizar la misma operación elemental en las filas de A corresponde a multiplicar por la izquierda a A con E^t.

Definición 7.2.7. Sean $A, B \in M_{m\times n}$. Decimos que A es *equivalente* con B y escribimos $A \sim B$ si A se puede obtener realizando operaciones elementales en las filas y columnas de B. De la proposición 7.2.3 se deduce que $A \sim B$ si y solo si existen matrices elementales $E_1, \dots, E_k \in M_m$ y $F_1, \dots, F_h \in M_n$ tales que

$$A = E_k \cdots E_1 B F_1 \cdots F_h.$$

Proposición 7.2.8. *La relación $\sim$ es de equivalencia en $M_{m\times n}$, es decir verifica*

$$A \sim A; \quad A \sim B \;\Rightarrow\; B \sim A; \quad A \sim B \text{ y } B \sim C \;\Rightarrow\; A \sim C. \quad \square$$

El siguiente resultado relaciona la equivalencia de matrices y la invertibilidad.

Proposición 7.2.9. *Sean $A, B \in M_n$ tales que $A \sim B$. Entonces A es invertible si y solo si B es invertible.*

Dem. Las matrices elementales son invertibles, y el producto de matrices invertibles es también una matriz invertible. □

Una pregunta natural es la siguiente. Dada una matriz arbitraria A ¿cuál es la forma más simple que podemos obtener, realizando operaciones elementales en filas y columnas de A? La respuesta la da el siguiente teorema.

Teorema 7.2.10. *Si $A \in M_{m\times n}$, entonces existe r, $0 \leq r \leq min\{m,n\}$, tal que $A \sim \left(\begin{smallmatrix} I_r & 0 \\ 0 & 0 \end{smallmatrix}\right)$.*

En lo anterior, se entiende que si $r = 0$ es $A \sim 0$, lo cual equivale a $A = 0$; también puede ser $A \sim \left(\begin{smallmatrix} I_r & 0 \end{smallmatrix}\right)$, $A \sim \left(\begin{smallmatrix} I_r \\ 0 \end{smallmatrix}\right)$ o $A \sim I_r$.

Dem. Sea $A \in M_{m\times n}$. Si $A = 0$, entonces es $r = 0$ y ya está probado. En lo que sigue asumiremos siempre $A \neq 0$. Supongamos primero que es $A \in M_2$. Como es $A \neq 0$, realizando operaciones elementales podemos obtener

$$A \sim \begin{pmatrix} 1 & b \\ c & d \end{pmatrix} \sim \begin{pmatrix} 1 & b \\ 0 & d - cb \end{pmatrix}.$$

Luego,

$$\text{si } d = cb, \text{ es } A \sim \begin{pmatrix} 1 & b \\ 0 & 0 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}; \quad \text{si } d \neq cb, \text{ es } A \sim \begin{pmatrix} 1 & b \\ 0 & d - cb \end{pmatrix} \sim \begin{pmatrix} 1 & b \\ 0 & 1 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Luego el teorema es válido cuando $m = n = 2$. También el teorema es muy fácil de probar para matrices fila $1 \times m$ o matrices columna $n \times 1$. Supongamos que tenemos ahora una matriz $A \in M_{m\times n}$, que no es de las formas ya estudiadas, y razonamos inductivamente suponiendo que el resultado es válido para matrices de tamaño $(m-1) \times n$ o $m \times (n-1)$. Como es $A \neq 0$, realizando operaciones elementales podemos obtener

$$A \sim \begin{pmatrix} 1 & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{2m} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \sim \begin{pmatrix} 1 & a_{12} & \cdots & a_{1n} \\ 0 & a'_{22} & \cdots & a'_{2n} \\ \vdots & \vdots & & \vdots \\ 0 & a'_{m2} & \cdots & a'_{mn} \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & a''_{22} & \cdots & a''_{2n} \\ \vdots & \vdots & & \vdots \\ 0 & a''_{m2} & \cdots & a''_{mn} \end{pmatrix}.$$

Sea $B = \begin{pmatrix} a''_{22} & \cdots & a''_{2n} \\ \vdots & & \vdots \\ a''_{m2} & \cdots & a''_{mn} \end{pmatrix}$. Si $B = 0$, entonces $A \sim \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 0 \end{pmatrix}$ y ya está probado. Si $B \neq 0$, entonces por la hipótesis inductiva sabemos que es $B \sim \left(\begin{smallmatrix} I_r & 0 \\ 0 & 0 \end{smallmatrix}\right)$, para cierto r. Luego $A = \left(\begin{smallmatrix} 1 & 0 \\ 0 & B \end{smallmatrix}\right) \sim \left(\begin{smallmatrix} I_{r+1} & 0 \\ 0 & 0 \end{smallmatrix}\right)$. $\square$

Nota: La relación de equivalencia $\sim$ parte al conjunto M_n en clases de equivalencia. El teorema 7.2.10 nos dice que a lo más hay $n + 1$ clases que corresponden a las matrices $\left(\begin{smallmatrix} I_r & 0 \\ 0 & 0 \end{smallmatrix}\right)$, para $r = 0, 1, \ldots, n$. Usando el concepto del rango de un matriz se prueba que estas matrices no son equivalentes entre sí, por lo que hay exactamente $n + 1$ clases de equivalencia.

En lo que sigue aplicaremos esta relación de equivalencia al estudio de la invertibilidad de matrices.

Proposición 7.2.11. *Una matriz $A \in M_n$ es invertible si y solo si $A \sim I_n$.*

Dem. Sabemos que vale $A \sim \left(\begin{smallmatrix} I_r & 0 \\ 0 & 0 \end{smallmatrix}\right)$, para cierto $0 \leq r \leq n$. Es claro que $\left(\begin{smallmatrix} I_r & 0 \\ 0 & 0 \end{smallmatrix}\right) \in M_n$ es invertible si y solo si $r = n$. Luego la proposición 7.2.9 implica que A es invertible si y solo si $r = n$. Observar que la misma proposición implica que no puede ser $\left(\begin{smallmatrix} I_r & 0 \\ 0 & 0 \end{smallmatrix}\right) \sim I_n$ con $r < n$. Esto implica la tesis. $\square$

Teorema 7.2.12. *Una matriz cuadrada es invertible si y solo si es producto de matrices elementales.*

Dem. Por la proposición anterior, sabemos que $A \in M_n$ es invertible si y solo si $A \sim I_n$. Esta última afirmación quiere decir que existen matrices elementales $E_1, \ldots, E_k$, $F_1, \ldots, F_h \in M_n$ tales que

$$A = E_k \cdots E_1 I_n F_1 \cdots F_h = E_k \cdots E_1 F_1 \cdots F_h. \quad \square$$

Corolario 7.2.13. *Dos matrices $A, B \in M_{m\times n}$ son equivalentes si y solo si existen matrices invertibles $U, V \in M_n$ tales que $A = UBV$.* $\square$

Corolario 7.2.14. *Sea $A \in M_n$.*

1. *Si A es equivalente a una matriz triangular superior de la forma $\begin{pmatrix} * & & * \\ & \ddots & \\ 0 & & * \end{pmatrix}$, cuyas entradas en la diagonal principal son todas no nulas, entonces A es invertible.*
2. *Si A es equivalente a una matriz triangular superior de la forma*
$$\begin{pmatrix} * & \cdots & * & * & \cdots & * \\ & \ddots & \vdots & \vdots & & \vdots \\ & & * & * & \cdots & * \\ & & & 0 & \cdots & 0 \\ & & & & \ddots & \vdots \\ & & & & & 0 \end{pmatrix},$$
enonces A no es invertible.

Vale lo mismo con matrices triangulares inferiores en vez de triangulares superiores.

Dem. La segunda afirmación es inmediata (si T es una amtriz triangular de ese tiepo, entonces TB no puede dar la matriz identidad, para ninguna matriz B). Si una matriz A está en las hipótesis de la primer afirmación, entonces es fácil probar que realizando operaciones elementales obtenemos $A \sim I_n$; luego A es invertible. La última afirmación de la tesis se deduce de las anteriores aplicando la traspuesta. $\square$

Aplicación 7.2.1. *Un método rápido para saber si una matriz es invertible es realizar operaciones elementales en sus filas y columnas hasta obtener una matriz que tenga forma triangular (en forma parecida a la escalerización de un sistema de ecuaciones), y luego aplicar el corolario anterior.*

Ejemplo 7.2.15. Sea $A = \begin{pmatrix} -7 & 4 & 1 \\ 4 & -2 & -2 \\ 1 & -2 & 7 \end{pmatrix}$. Si primero le sumamos a la primer columna la segunda multiplicada por 2 y luego a la tercer fila le sumamos la primera multiplicada por 3, y finalmente a la tercer fila le sumamos la segunda multiplicada por 5, obtenemos:

$$\begin{pmatrix} -7 & 4 & 1 \\ 4 & -2 & -2 \\ 1 & -2 & 7 \end{pmatrix} \sim \begin{pmatrix} 1 & 4 & 1 \\ 0 & -2 & -2 \\ -3 & -2 & 7 \end{pmatrix} \sim \begin{pmatrix} 1 & 4 & 1 \\ 0 & -2 & -2 \\ 0 & 10 & 10 \end{pmatrix} \sim \begin{pmatrix} 1 & 4 & 1 \\ 0 & -2 & -2 \\ 0 & 0 & 0 \end{pmatrix}.$$

Como la última matriz tiene una fila de ceros, entonces no es invertible y por lo tanto A tampoco lo es. Notar que este método para determinar la invertibilidad es muy rápido, porque podemos trabajar con las columnas o filas alternadamente. Eso no ocurre con el método que veremos a continuación

Observación 7.2.16. Notar que este método para determinar la invertibilidad es muy rápido, dado que podemos trabajar con las columnas o filas alternadamente. Eso no ocurre con el método que veremos a continuación para calcular la inversa, en el cual hay que elegir trabajar con columnas o filas, pero no se puede mezclar.

Aplicación 7.2.2 (Método para obtener la matriz inversa.)**.** *La proposición 7.2.11 junto con el teorema 7.2.12 nos dan un algoritmo para hallar la inversa de una matriz: si $A^{-1}A = I$ y $E_1, \ldots, E_r$ son matrices elementales, entonces $A^{-1}(AE_1 \cdots E_r) = E_1 \cdots E_r$. Luego si elegimos $E_1, \ldots, E_r$ de forma tal que $AE_1 \cdots E_r = I$, entonces $A^{-1} = E_1 \cdots E_r$. Observar que multiplicar por la derecha por matrices elementales equivale a hacer las operaciones elementales respectivas en las columnas, por lo tanto si realizando ciertas operaciones elementales en las columnas de una matriz A llegamos a la matriz identidad ($AE_1 \cdots E_r = I$), entonces realizando las mismas operaciones elementales en las columnas de la matriz identidad obtenemos $A^{-1} = I \cdot E_1 \cdots E_r$. De la misma forma, operar con las filas equivale a multiplicar por la izquierda, y por lo tanto partiendo de la otra igualdad $AA^{-1} = I$ se obtiene un algoritmo análogo operando con filas en vez de columnas.*

Notar que para trabajar con columnas partimos de $A^{-1}A = I$ y para las filas de $AA^{-1} = I$; es por eso que en el algoritmo no podemos realizar operaciones elementales con filas y columnas al mismo tiempo.

A continuación veremos algunos ejemplos donde estudiamos la invertibilidad de matrices y calculamos las inversas. En lo que sigue, si por ejemplo escribimos $F_2 - F_1$ quiere decir que a la fila 2 le estamos restando la fila 1, análogamente $C_3 + 2C_4$ quiere decir que a la columna 3 le estamos sumando la columna 4 multiplicada por 2. Notar que en las sumas o restas anteriores siempre escribimos primero la fila o columna que es modificada. Si escribimos $F_2 \leftrightarrow F_3$, quiere decir que intercambiamos la fila 2 con la fila 3, y análogamente para columnas.

Ejemplos 7.2.17. Sea $A = \begin{pmatrix} 2 & 3 & 1 \\ -1 & 0 & -1 \\ 1 & 1 & 1 \end{pmatrix}$. Primero estudiamos su invertibilidad.

$$\begin{pmatrix} 2 & 3 & 1 \\ -1 & 0 & -1 \\ 1 & 1 & 1 \end{pmatrix} \underset{C_1 - C_3}{\sim} \begin{pmatrix} 1 & 3 & 1 \\ 0 & 0 & -1 \\ 0 & 1 & 1 \end{pmatrix} \underset{F_2 \leftrightarrow F_3}{\sim} \begin{pmatrix} 1 & 3 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & -1 \end{pmatrix}.$$

Como la última matriz es triangular superior y los elementos de su diagonal principal son no nulos, deducimos que A es invertible.

Para calcular su inversa tenemos que elegir primero si trabajar con columnas o con filas. En este caso trabajaremos con columnas.

2	3	1	1	0	0	
−1	0	−1	0	1	0	$C_1 - C_3$
1	1	1	0	0	1	
1	3	1	1	0	0	
0	0	−1	0	1	0	$C_2 - C_3$
0	1	1	−1	0	1	
1	2	1	1	0	0	
0	1	−1	0	1	0	$C_3 - C_1$
0	0	1	−1	−1	1	
1	2	0	1	0	−1	
0	1	−1	0	1	0	$C_2 - 2C_1$
0	0	1	−1	−1	2	
1	0	0	1	−2	−1	
0	1	−1	0	1	0	$C_3 + C_2$
0	0	1	−1	1	2	
1	0	0	1	−2	−3	
0	1	0	0	1	1	
0	0	1	−1	1	3	

Luego la inversa es $A^{-1} = \begin{pmatrix} 1 & -2 & -3 \\ 0 & 1 & 1 \\ -1 & 1 & 3 \end{pmatrix}$.

Sea $A = \begin{pmatrix} 2 & 2 & 1 \\ -2 & -1 & -2 \\ 1 & 0 & 1 \end{pmatrix}$.

Dejamos como ejercicio el probar que A es invertible y mostraremos cómo hallar su inversa operando con las filas. También iremos un poco más rápido, haciendo a veces más de una operación por paso.

$$
\begin{array}{|ccc|ccc|c}
2 & 2 & 1 & 1 & 0 & 0 & \\
-2 & -1 & -2 & 0 & 1 & 0 & F_1 \leftrightarrow F_3 \\
1 & 0 & 1 & 0 & 0 & 1 & \\
\hline
1 & 0 & 1 & 0 & 0 & 1 & \\
-2 & -1 & -2 & 0 & 1 & 0 & F_2 + 2F_1 \text{ y } F_3 - 2F_1 \\
2 & 2 & 1 & 1 & 0 & 0 & \\
\hline
1 & 0 & 1 & 0 & 0 & 1 & \\
0 & -1 & 0 & 0 & 1 & 2 & F_3 + 2F_2 \\
0 & 2 & -1 & 1 & 0 & -2 & \\
\hline
1 & 0 & 1 & 0 & 0 & 1 & \\
0 & -1 & 0 & 0 & 1 & 2 & (-1)F_2 \text{ y } (-1)F_3 \\
0 & 0 & -1 & 1 & 2 & 2 & \\
\hline
1 & 0 & 1 & 0 & 0 & 1 & \\
0 & 1 & 0 & 0 & -1 & -2 & F_1 - F_3 \\
0 & 0 & 1 & -1 & -2 & -2 & \\
\hline
1 & 0 & 0 & 1 & 2 & 3 & \\
0 & 1 & 0 & 0 & -1 & -2 & F_1 - F_3 \\
0 & 0 & 1 & -1 & -2 & -2 &
\end{array}
$$

Luego la inversa es $A^{-1} = \begin{pmatrix} 1 & 2 & 3 \\ 0 & -1 & -2 \\ -1 & -2 & -2 \end{pmatrix}$.

Sea $A = \begin{pmatrix} 3 & 1 & 1 \\ 2 & 3 & 2 \\ 3 & 3 & 3 \end{pmatrix}$. Dejamos como ejercicio el probar que vale $\det(A) = 6$, luego A es invertible.

Trabajaremos con columnas.

$$\begin{array}{ccc|ccc|c}
3 & 1 & 1 & 1 & 0 & 0 & \\
2 & 3 & 2 & 0 & 1 & 0 & C_1 - C_3 \text{ y } C_2 - C_3 \\
3 & 3 & 3 & 0 & 0 & 1 & \\
\hline
2 & 0 & 1 & 1 & 0 & 0 & \\
0 & 1 & 2 & 0 & 1 & 0 & \frac{1}{2}C_1 \\
0 & 0 & 3 & -1 & -1 & 1 & \\
\hline
1 & 0 & 1 & 1/2 & 0 & 0 & \\
0 & 1 & 2 & 0 & 1 & 0 & C_3 - C_1 \\
0 & 0 & 3 & -1/2 & -1 & 1 & \\
\hline
1 & 0 & 0 & 1/2 & 0 & -1/2 & \\
0 & 1 & 2 & 0 & 1 & 0 & C_3 - 2C_2 \\
0 & 0 & 3 & -1/2 & -1 & 3/2 & \\
\hline
1 & 0 & 0 & 1/2 & 0 & -1/6 & \\
0 & 1 & 0 & 0 & 1 & -2/3 & \frac{1}{3}C_3 \\
0 & 0 & 1 & -1/2 & -1 & 7/6 &
\end{array}$$

Luego la inversa es $A^{-1} = \begin{pmatrix} 1/2 & 0 & -1/6 \\ 0 & 1 & -2/3 \\ -1/2 & -1 & 7/6 \end{pmatrix}$.

Observación 7.2.18. Notar que en los ejemplos anteriores lo que hicimos primero fue transformar la matriz A en una matriz triangular (triangular superior en todos los casos, pero podría ser también triangular inferior) y luego llevarla a una forma diagonal. Es importante verificar que la matriz hallada es realmente la matriz inversa (haciendo el producto), dado que es muy fácil cometer algún error de cuentas.

Observación 7.2.19. Conviene remarcar que tenemos dos situaciones distintas:

1. Si queremos investigar si una matriz es invertible, entonces podemos usar operaciones elementales en sus filas y/o columnas llevándola a una matriz triangular, y luego aplicamos el corolario 7.2.14.

2. Si ya sabemos que una matriz es invertible y queremos hallar su inversa, entonces aplicamos el algoritmo de arriba, pero tenemos que elegir trabajar con filas o columnas, y no podemos mezclar!.

7.3. Matrices en bloques

En esta sección estudiaremos las propiedades de matrices cuyos elementos están agrupados en submatrices (bloques) de la matriz original. Esto tiene varias aplicaciones, en particular se usa para el estudio del polinomio minimal y de la forma de Jordan, que veremos en los capítulos 5 y 6.

Empezamos considerando un caso particular. Dadas cuatro matrices

$$A = (a_{ij}) \in M_{p\times p}(\Bbbk), \quad B = (b_{ij}) \in M_{p\times q}(\Bbbk), \quad C = (c_{ij}) \in M_{q\times p}(\Bbbk), \quad D = (d_{ij}) \in M_{q\times q}(\Bbbk),$$

las podemos combinar para obtener una matriz $\left(\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right) \in M_{n\times n}(\Bbbk)$, con $n = p + q$, definida por

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix} := \begin{pmatrix} a_{11} & \cdots & a_{1p} & b_{11} & \cdots & b_{1q} \\ \vdots & & \vdots & \vdots & & \vdots \\ a_{p1} & \cdots & a_{pp} & b_{p1} & \cdots & b_{pq} \\ c_{11} & \cdots & c_{1p} & d_{11} & \cdots & d_{1q} \\ \vdots & & \vdots & \vdots & & \vdots \\ c_{q1} & \cdots & c_{qp} & d_{q1} & \cdots & d_{qq} \end{pmatrix}.$$

Notar que lo que hicimos para cuatro matrices, se puede aplicar a una cantidad arbitraria de matrices que tengan tamaños adecuados para poder combinarse correctamente. Una matriz expresada de esa forma diremos que es una *matriz en bloques.* Es claro que toda matriz se puede expresar como matriz en bloques (en general de muchas formas). Lo interesante de esto es cuando esos bloques tienen alguna forma especial.

Para evitar confusiones, cuando escribamos una matriz en bloques, a las submatrices que la componen las escribiremos siempre con mayúsculas. Así $\left(\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right)$ representa una matriz en bloques y no a una matriz 2×2 (a menos que todos los bloques sean de tamaño 1×1).

Las matrices que más nos interesan son las matrices diagonales en bloques, es decir, las de la forma

$$\begin{pmatrix} A_1 & 0 & \cdots & 0 \\ 0 & A_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & A_h \end{pmatrix},$$

en que cada A_i es una matriz de tamaño $n_i \times n_i$, $i = 1, \ldots, h$ y las otras entradas son nulas. Es un ejercicio el probar que valen las fórmulas siguientes

$$\begin{pmatrix} A_1 & 0 & \cdots & 0 \\ 0 & A_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & A_h \end{pmatrix} + c \begin{pmatrix} B_1 & 0 & \cdots & 0 \\ 0 & B_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & B_h \end{pmatrix} = \begin{pmatrix} A_1 + cB_1 & 0 & \cdots & 0 \\ 0 & A_2 + cB_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & A_h + cB_h \end{pmatrix},$$

$$\begin{pmatrix} A_1 & 0 & \cdots & 0 \\ 0 & A_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & A_h \end{pmatrix} \begin{pmatrix} B_1 & 0 & \cdots & 0 \\ 0 & B_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & B_h \end{pmatrix} = \begin{pmatrix} A_1B_1 & 0 & \cdots & 0 \\ 0 & A_2B_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & A_hB_h \end{pmatrix}$$

En particular de la fórmula para el producto se obtiene la fórmula para las potencias

$$\begin{pmatrix} A_1 & 0 & \cdots & 0 \\ 0 & A_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & A_h \end{pmatrix}^k = \begin{pmatrix} A_1^k & 0 & \cdots & 0 \\ 0 & A_2^k & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & A_h^k \end{pmatrix}, \quad \forall k \in \mathbb{N}.$$

Notar la analogía de estas fórmulas con las correspondientes en matrices diagonales. A continuación veremos que el cálculo de determinantes se simplifica para ciertos tipos de matrices en bloques.

Proposición 7.3.1. *Si* $X = \left(\begin{smallmatrix} A & B \\ 0 & D \end{smallmatrix}\right)$*, siendo* A *y* D *matrices cuadradas, entonces* $\det X = \det A \det D$.

Dem. Supongamos $A = (a_{ij}) \in M_{p\times p}(\Bbbk)$. Si $p = 1$, entonces $A = (a_{11})$. Luego desarrollando $\det X$ por la primer comlumna obtenemos

$$\det X = a_{11} \det D = \det A \det D.$$

El caso general se deduce por inducción en p, desarrollando $\det X$ por su primer columna. Como este cálculo es muy engorroso, lo mostraremos solo para $p = 3$, que es ilustrativo del paso inductivo, dejando la prueba general como ejercicio. Supongamos entonces que tenemos probado el resultado para $p = 2$ y consideramos el caso $p = 3$. Desarrollando $\det X$ por la primer columna obtenemos

$$
\begin{vmatrix}
a_{11} & a_{12} & a_{13} & b_{11} & \cdots & b_{1q} \\
a_{21} & a_{22} & a_{23} & b_{21} & \cdots & b_{2q} \\
a_{31} & a_{32} & a_{33} & b_{31} & \cdots & b_{3q} \\
0 & 0 & 0 & d_{11} & \cdots & d_{1q} \\
\vdots & \vdots & \vdots & \vdots & & \vdots \\
0 & 0 & 0 & d_{q1} & \cdots & d_{qq}
\end{vmatrix} =
$$

$$
= a_{11} \begin{vmatrix}
a_{22} & a_{23} & b_{21} & \cdots & b_{2q} \\
a_{32} & a_{33} & b_{31} & \cdots & b_{3q} \\
0 & 0 & d_{11} & \cdots & d_{1q} \\
\vdots & \vdots & \vdots & & \vdots \\
0 & 0 & d_{q1} & \cdots & d_{qq}
\end{vmatrix} - a_{21} \begin{vmatrix}
a_{12} & a_{13} & b_{11} & \cdots & b_{1q} \\
a_{32} & a_{33} & b_{31} & \cdots & b_{3q} \\
0 & 0 & d_{11} & \cdots & d_{1q} \\
\vdots & \vdots & \vdots & & \vdots \\
0 & 0 & d_{q1} & \cdots & d_{qq}
\end{vmatrix} + a_{31} \begin{vmatrix}
a_{12} & a_{13} & b_{11} & \cdots & b_{1q} \\
a_{22} & a_{23} & b_{21} & \cdots & b_{2q} \\
0 & 0 & d_{11} & \cdots & d_{1q} \\
\vdots & \vdots & \vdots & & \vdots \\
0 & 0 & d_{q1} & \cdots & d_{qq}
\end{vmatrix}
$$

$$
= a_{11} \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} |D| - a_{21} \begin{vmatrix} a_{12} & a_{13} \\ a_{32} & a_{33} \end{vmatrix} |D| + a_{31} \begin{vmatrix} a_{12} & a_{13} \\ a_{22} & a_{23} \end{vmatrix} |D| \quad \text{(usando la hipótesis inductiva)}
$$

$$
= \left(a_{11} \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{21} \begin{vmatrix} a_{12} & a_{13} \\ a_{32} & a_{33} \end{vmatrix} + a_{31} \begin{vmatrix} a_{12} & a_{13} \\ a_{22} & a_{23} \end{vmatrix} \right) |D| = |A||D|. \quad \square
$$

Proposición 7.3.2. *Si X es una matriz triangular superior en bloques del tipo*

$$
X = \begin{pmatrix}
A_{11} & A_{12} & \cdots & A_{1h} \\
0 & A_{22} & \cdots & A_{2h} \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & A_{hh}
\end{pmatrix},
$$

en que $A_{11}, \dots, A_{hh}$ son matrices cuadradas, entonces $\det X = \prod_{i=1}^{h} \det A_{ii}$.

Dem. El caso $h = 2$ es la proposición anterior. El caso general se prueba por inducción en h, observando que podemos escribir $X = \left(\begin{smallmatrix} A11 & B \\ 0 & T \end{smallmatrix}\right)$, siendo $T = \begin{pmatrix} A_{22} & \cdots & A_{2h} \\ \vdots & \ddots & \vdots \\ 0 & \cdots & A_{hh} \end{pmatrix}$ y $B = \begin{pmatrix} A_{12} & \cdots & A_{1h} \end{pmatrix}$. $\square$

Corolario 7.3.3. *Si X es una matriz diagonal en bloques del tipo*

$$
X = \begin{pmatrix}
A_1 & 0 & \cdots & 0 \\
0 & A_2 & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & A_h
\end{pmatrix},
$$

en que $A_1, \dots, A_h$ son matrices cuadradas, entonces $\det X = \prod_{i=1}^{h} \det A_i$. $\square$

7.4. Suma directa de subespacios

Sea V un espacio vectorial y W_1, W_2 dos subespacios de V. Definimos su *intersección* $W_1 \cap W_2$ y su *suma* $W_1 + W_2$, mediante

$$
W_1 \cap W_2 := \{v \in V : \ v \in W_1 \text{ y } v \in W_2\}, \qquad W_1 + W_2 := \{w_1 + w_2 : \ w_1 \in W_1, \ w_2 \in W_2\}.
$$

Proposición 7.4.1. *Si W_1, W_2 son subespacios de V, entonces $W_1 \cap W_2$ y $W_1 + W_2$ son también subespacios.*

Dem. Ejercicio. $\square$

Observación 7.4.2. Dados dos subespacios W_1, W_2, su intersección $W_1 \cap W_2$ es el mayor subespacio contenido en W_1 y W_2, mientras que $W_1 + W_2$ es el menor subespacio que contiene a W_1 y W_2.

Ejemplos 7.4.3. Consideremos el espacio $\mathbb{R}^3$.

1. Si $W_1 = \{(x, y, z) : \ z = 0\}$ (plano Oxy) y $W_2 = \{(x, y, z) : \ y = 0\}$ (plano Oxz), entonces

$$W_1 \cap W_2 = \{(x, y, z) : \ y = z = 0\} \text{ (eje Ox)}, \qquad W_1 + W_2 = \mathbb{R}^3.$$

2. Si $W_1 = \{(x, y, z) : \ y = z = 0\}$ (eje Ox) y $W_2 = \{(x, y, z) : \ x = z = 0\}$ (eje Oy), entonces

$$W_1 \cap W_2 = \{(0, 0, 0)\}, \qquad W_1 + W_2 = \{(x, y, z) : \ z = 0\} \text{ (plano Oxy)}.$$

La siguiente proposición relaciona las dimensiones de $W_1 + W_2$ y $W_1 \cap W_2$, con las de W_1 y W_2.

Proposición 7.4.4. *Si W_1, W_2 son subespacios de dimensión finita de un espacio V, entonces vale*

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

Dem. Sea $\mathcal{B} = \{u_1, \dots, u_n\}$ una base de $W_1 \cap W_2$. Como $\mathcal{B}$ es un subconjunto LI de W_1 y de W_2, entonces existen $v_1, \dots, v_p \in W_1$ y $w_1, \dots, w_q \in W_2$, tales que si

$$\mathcal{B}_1 = \{u_1, \dots, u_n, v_1, \dots, v_p\} \quad \text{y} \quad \mathcal{B}_2 = \{u_1, \dots, u_n, w_1, \dots, w_q\},$$

entonces $\mathcal{B}_1$ es base de W_1 y $\mathcal{B}_2$ es base de W_2. Probaremos que $\mathcal{C} = \{u_1, \dots, u_n, v_1, \dots, v_p, w_1, \dots, w_q\}$ es base de $W_1 + W_2$. La prueba de que $\mathcal{C}$ es un conjunto generador de $W_1 + W_2$ es fácil y queda como ejercicio. Veamos que $\mathcal{C}$ es LI. Sean escalares a_i, b_j, c_k tales que

$$a_1 u_1 + \cdots + a_n u_n + b_1 v_1 + \cdots + b_p v_p + c_1 w_1 \cdots + c_q w_q = 0. \tag{7.4}$$

Luego

$$c_1 w_1 \cdots + c_q w_q = -(a_1 u_1 + \cdots + a_n u_n + b_1 v_1 + \cdots + b_p v_p). \tag{7.5}$$

Como el lado izquierdo de (7.5) está en W_2 y el derecho en W_1, deducimos $c_1 w_1 \cdots + c_q w_q \in W_1 \cap W_2$. Luego existen escalares d_i tales que

$$c_1 w_1 \cdots + c_q w_q = d_1 u_1 + \cdots + d_n u_n \quad \Rightarrow \quad c_1 w_1 \cdots + c_q w_q + (-d_1) u_1 + \cdots + (-d_n) u_n = 0.$$

Dado que $\mathcal{B}_2$ es LI, deducimos $c_1 = \cdots = c_q = d_1 = \cdots = d_n = 0$. Sustituyendo estos valores en (7.4) obtenemos

$$a_1 u_1 + \cdots + a_n u_n + b_1 v_1 + \cdots + b_p v_p = 0.$$

Ahora usando que $\mathcal{B}_1$ es LI deducimos $a_1 = \cdots = a_n = b_1 = \cdots = b_p = 0$. Esto prueba que $\mathcal{C}$ es LI y por lo tanto es base de $W_1 + W_2$. Luego

$$\dim(W_1 + W_2) = n + p + q = (n + p) + (n + q) - n = \dim W_1 + \dim W_2 - \dim(W_1 \cap W_2). \quad \square$$

La proposición anterior se suele usar para calcular la dimensión de la suma de dos subespacios, ya que en general es más fácil determinar la intersección que la suma.

Ejemplo 7.4.5. Sea $V = \mathbb{R}^4$ y consideramos los subespacios

$$W_1 = \left\{(x, y, z, t) \in \mathbb{R}^4 : \ x = 0, \ y + z + t = 0\right\} \quad \text{y} \quad W_2 = \left\{(x, y, z, t) \in \mathbb{R}^4 : \ t = 0\right\}.$$

Es fácil de probar que es $\dim W_1 = 2$ y $\dim W_2 = 3$. Además

$$W_1 \cap W_2 = \left\{(x, y, z, t) \in \mathbb{R}^4 : \ x = 0, \ y + z + t = 0, \ t = 0\right\} = \left\{(x, y, z, t) \in \mathbb{R}^4 : \ x = t = 0, \ y + z = 0\right\},$$

luego $\dim(W_1 \cap W_2) = 1$ y por lo tanto

$$\dim(W_1 + W_2) = \dim W_1 + \dim W_2 - \dim(W_1 \cap W_2) = 2 + 3 - 1 = 4 = \dim \mathbb{R}^4.$$

Luego $W_1 + W_2 = \mathbb{R}^4$.

Definición 7.4.6. Sean W_1, W_2 dos subespacios de un espacio V. Decimos que un subespacio U de V es *suma directa* de W_1 y W_2 y escribimos $U = W_1 \oplus W_2$ si se cumple $U = W_1 + W_2$ y $W_1 \cap W_2 = \{0\}$.

Ejemplo 7.4.7. Si consideramos los ejemplos 7.4.3, deducimos del primero que el espacio $\mathbb{R}^3$ es suma del plano Oxy con el plano Oxz, pero esta suma no es directa, y del segundo que el plano Oxy es suma directa de la recta Ox y la recta Oy. En el ejercicio 7.4.5 vemos también que $\mathbb{R}^4$ es suma de W_1 y W_2, pero esta suma no es directa.

Proposición 7.4.8. *Si W_1, W_2 son dos subespacios de V, entonces $V = W_1 \oplus W_2$ si y solo si todo vector de V se escribe en forma única como suma de un vector de W_1 con uno de W_2.*

Dem. Ver la prueba de la proposición 7.4.14. ☐

De la proposición 7.4.4 y de su demostración, se deduce inmediatamente el siguiente resultado.

Proposición 7.4.9. *Sea V un espacio de dimensión finita y W_1, W_2 dos subespacios de V, tales que $V = W_1 \oplus W_2$. Valen las siguientes afirmaciones.*

1. *Si $\mathcal{B}_1$ y $\mathcal{B}_2$ son bases respectivas de W_1 y W_2, entonces su unión $\mathcal{B}_1 \cup \mathcal{B}_2$ es una base de V.*
2. *$dim V = \dim W_1 + \dim W_2$.* ☐

Corolario 7.4.10. *Sean W_1, W_2 dos subespacios de un espacio de dimensión finita V tales que $W_1 \cap W_2 = \{0\}$. Si $\dim V = \dim W_1 + \dim W_2$, entonces $V = W_1 \oplus W_2$.*

Dem. La proposición 7.4.9 implica $\dim(W_1 \oplus W_2) = \dim W_1 + \dim W_2 = \dim V$, luego $W_1 \oplus W_2 = V$. ☐

Este último corolario nos da un método fácil para probar que un espacio es suma directa de dos subespacios dados.

Ejemplo 7.4.11. Consideremos el espacio $\mathbb{R}^3$ y los subespacios

$$W_1 = \{(x, y, z) : \ 2x + y + z = 0\} \qquad \text{y} \qquad W_2 = \{(x, y, z) : \ y = z = 0\}.$$

Si $(x, y, z) \in W_1 \cap W_2$, entonces vale

$$\begin{cases} 2x + y + z = 0 \\ \qquad y = z = 0 \end{cases} \quad \Rightarrow \quad x = y = z = 0 \quad \Rightarrow \quad (x, y, z) = (0, 0, 0).$$

Luego $W_1 \cap W_2 = \{(0, 0, 0)\}$. Por otro lado W_1 es un plano y W_2 una recta, así que es $\dim W_1 = 2$ y $\dim W_2 = 1$ y por lo tanto $\dim W_1 + \dim W_2 = 2 + 1 = 3 = \dim \mathbb{R}^3$; luego $\mathbb{R}^3 = W_1 \oplus W_2$.

Observación 7.4.12. Dados dos subespacios W_1, W_2 de V, no alcanza con que valga $\dim V = \dim W_1 + \dim W_2$ para que la suma sea directa, hay que probar que también vale $W_1 \cap W_2 = \{0\}$. Por ejemplo, si el espacio es $\mathbb{R}^3$ y consideramos los subespacios

$$W_1 = \{(x, y, z) : \ z = 0\} \qquad \text{y} \qquad W_2 = \{(x, y, z) : \ y = z = 0\},$$

entonces W_1 es un plano y W_2 una recta, y por lo tanto $\dim W_1 + \dim W_2 = 2 + 1 = 3 = \dim \mathbb{R}^3$. Pero W_2 está contenido en W_1, luego $W_1 + W_2 = W_1 \subsetneq \mathbb{R}^3$ y por lo tanto nunca puede ser $W_1 \oplus W_2 = \mathbb{R}^3$.

Generalización. A continuación veremos que todos los resultados anteriores para la suma e intersección de dos subespacios se pueden generalizar a una cantidad finita arbitraria de subespacios.

Sea V un espacio vectorial y $W_1, \dots, W_m$ una cantidad finita de subespacios de V. Definimos su *intersección* $\bigcap_{i=1}^m W_i$ y su *suma* $\sum_{i=1}^m W_i$, mediante

$$\bigcap_{i=1}^{m} W_i := \{v \in V : \ v \in W_i, \ \forall i = 1, \dots, n\}, \qquad \sum_{i=1}^{m} W_i := \left\{ \sum_{i=1}^{m} w_i : \ w_i \in W_i, \ \forall i = 1, \dots, n \right\}.$$

La prueba de la siguiente proposición es simple y queda como ejercicio.

Proposición 7.4.13. *Si $W_1, \dots, W_m$ son subespacios de V, entonces $\bigcap_{i=1}^m W_i$ y $\sum_{i=1}^m W_i$ son también subespacios.* □

La definición general de suma directa es un poco más delicada que para el caso de dos subespacios. Para definirla introducimos previamente el siguiente concepto.

Decimos que una familia[1] $\{W_1, \dots, W_m\}$ de subespacios de V es *independiente* si verifica la siguiente condición.

$$\text{Si } w_i \in W_i, \ \forall i = 1, \dots, m \text{ y } \sum_{i=1}^{m} w_i = 0, \text{ entonces } w_1 = \cdots = w_m = 0.$$

En lo que sigue escribimos

$$\sum_{j \neq i} W_j := W_1 + \cdots + W_{i-1} + W_{i+1} + \cdots + W_m, \quad \forall i = 1, \dots, m.$$

Proposición 7.4.14. *Sea $\{W_1, \dots, W_m\}$ una familia de subespacios de V. Entonces las siguientes condiciones son equivalentes:*

1. *La familia $\{W_1, \dots, W_m\}$ es independiente.*
2. *Para cada $i = 1, \dots, m$, vale $W_i \cap \sum_{j \neq i} W_j = \{0\}$.*
3. *Si $W = \sum_{i=1}^m W_i$, entonces para todo $w \in W$, existen únicos $w_i \in W_i$, $i = 1, \dots, m$, tales que $w = \sum_{i=1}^m w_i$.*

Dem. ($1 \Rightarrow 2$). Supongamos que la familia es independiente. Sea $i \in \{1, \dots, m\}$. Si $w \in W_i \cap \sum_{j \neq i} W_j$, entonces $w \in W_i$ y existen $w_j \in W_j$, para todo $j \neq i$ tales que

$$w = w_1 + \cdots + w_{i-1} + w_{i+1} + \cdots + w_m \quad \Rightarrow \quad w_1 + \cdots + w_{i-1} + (-w) + w_{i+1} + \cdots + w_m = 0.$$

Como la familia es independiente, entonces todos los sumandos de la suma anterior son nulos, luego $w = 0$.

($2 \Rightarrow 1$). Supongamos ahora que vale la segunda condición. Sean $w_i \in W_i$, $i = 1, \dots, m$, tales que $\sum_{i=1}^m w_i = 0$. Para cada $i \in \{1, \dots, m\}$, es

$$-w_i = w_1 + \cdots + w_{i-1} + w_{i+1} + \cdots + w_m.$$

Como $-w_i \in W_i$ y $w_1 + \cdots + w_{i-1} + w_{i+1} + \cdots + w_m \in \sum_{j \neq i} W_j$, deducimos que es $-w_i \in W_i \cap \sum_{j \neq i} W_j = \{0\}$. Luego $w_i = 0$, para todo $i = 1, \dots, m$.

[1] En este texto llamaremos *familia* a un conjunto de conjuntos.

(3 $\Rightarrow$ 1). Supongamos vale la tercer afirmación. Sean $w_i \in W_i$, $i = 1, \ldots, m$, tales que $\sum_{i=1}^m w_i = 0$. Como $0 \in W_i$ para todo $i = 1, \ldots, m$ y $\sum_{i=1}^m 0 = 0$, entonces la unicidad implica que es $w_i = 0$, para todo $i = 1, \ldots, m$. Luego la familia $\{W_1, \ldots, W_m\}$ es independiente.

(1 $\Rightarrow$ 3). Supongamos ahora que la familia $\{W_1, \ldots, W_m\}$ es independiente. Si un elemento $w \in W$ se escribe de dos formas $w = \sum_{i=1}^m w_i$ y $w = \sum_{i=1}^m w_i'$, con $w_i, w_i' \in W_i$, para todo $i = 1, \ldots, m$, entonces

$$\sum_{i=1}^m (w_i - w_i') = 0, \text{ siendo } w_i - w_i' \in W_i, \ \forall i = 1, \ldots, m.$$

Luego $w_i - w_i' = 0$ y por lo tanto $w_i = w_i'$, para todo $i = 1, \ldots, m$. □

Observación 7.4.15. Una familia con solo dos subespacios W_1 y W_2 es independiente si y solo si $W_1 \cap W_2 = \{0\}$, que es la condición que pedimos para que la suma $W_1 + W_2$ sea directa. Una familia de tres subespacios $\{W_1, W_2, W_3\}$ es independiente si y solo si W_1, W_2, W_3 verifican

$$W_1 \cap (W_2 + W_3) = \{0\}, \quad W_2 \cap (W_1 + W_3) = \{0\}, \quad W_3 \cap (W_1 + W_2) = \{0\}.$$

Notar que las condiciones $W_1 \cap W_2 = W_1 \cap W_3 = W_2 \cap W_3 = \{0\}$ no implican las condiciones anteriores. Por ejemplo, si consideramos $V = \mathbb{k}^3$ y

$$W_1 = \{(x, 0, 0) : \ x \in \mathbb{k}\}, \quad W_2 = \{(0, y, 0) : \ y \in \mathbb{k}\}, \quad W_3 = \{(x, x, 0) : \ x \in \mathbb{k}\},$$

entonces es fácil de probar que las intersecciones dos a dos son triviales, pero $W_1 + W_2 = \{(x, y, 0); \ x, y \in \mathbb{k}\}$, luego $W_3 \subset W_1 + W_2$ y por lo tanto $W_3 \cap (W_1 + W_2) = W_3 \neq \{0\}$.

Definición 7.4.16. Decimos que un subespacio W de V es *suma directa* de una familia de subespacios $\{W_1, \ldots, W_m\}$ y escribimos $W = \bigoplus_{i=1}^m W_i$, si $W = \sum_{i=1}^m W_i$ y la familia $\{W_1, \ldots, W_m\}$ es independiente.

Observación 7.4.17. Lo que distingue la suma directa de la suma, y la hace más interesante, es la unicidad de la tercer afirmación de la proposición 7.4.14. Es similar a la diferencia entre una base y un conjunto generador.

Corolario 7.4.18. *Sea $\{W_1, \ldots, W_m\}$ una familia de subespacios de un espacio V. Entonces $V = \bigoplus_{i=1}^m W_i$ si y solo si para todo $v \in V$, existen únicos $w_i \in W_i$, $i = 1, \ldots, m$, tales que $v = \sum_{i=1}^m w_i$.* □

Ejemplo 7.4.19. Es fácil probar usando el corolario 7.4.18 que vale $M_n = W_1 \oplus W_2 \oplus W_3$, siendo W_1 es el espacio de las matrices estrictamente triangulares inferiores, W_2 el de las matrices estrictamente triangulares superiores y W_3 el de las matrices diagonales, es decir

$$W_1 = \{(a_{ij}) : \ a_{ij} = 0 \text{ si } i \leq j\}, \quad W_2 = \{(a_{ij}) : \ a_{ij} = 0 \text{ si } i \geq j\}, \quad W_1 = \{(a_{ij}) : \ a_{ij} = 0 \text{ si } i \neq j\}.$$

Diagramáticamente, la descomposición $M_n = W_1 \oplus W_2 \oplus W_3$ consiste en escribir $\left(\begin{smallmatrix} * & * \\ * & * \end{smallmatrix}\right) = \left(\begin{smallmatrix} 0 & 0 \\ * & 0 \end{smallmatrix}\right) + \left(\begin{smallmatrix} 0 & * \\ 0 & 0 \end{smallmatrix}\right) + \left(\begin{smallmatrix} * & 0 \\ 0 & * \end{smallmatrix}\right)$.

El siguiente resultado es fácil de probar.

Proposición 7.4.20. *Sean $W, W_1, \ldots, W_m$ subespacios de V tales que $W = \sum_{i=1}^m W_i$. Si $\mathcal{G}_i$ es un conjunto generador de W_i, para todo $i = 1, \ldots, m$, entonces su unión $\bigcup_{i=1}^m \mathcal{G}_i$ es un conjunto generador de W.* □

A continuación veremos cómo se relaciona la suma directa con la dimensión.

Proposición 7.4.21. *Sean $W, W_1, \dots, W_m$ subespacios de dimensión finita de un espacio V, tales que $W = \bigoplus_{i=1}^m W_i$. Valen las siguientes afirmaciones.*

1. *Si $\mathcal{B}_i$ es una base de W_i, para todo $i = 1, \dots, m$, entonces $\mathcal{B} = \bigcup_{i=1}^m \mathcal{B}_i$ es una base de W.*
2. $\dim W = \sum_{i=1}^m \dim W_i$.

Dem. Por la proposición anterior sabemos que $\mathcal{B}$ es un conjunto generador de W, lo que resta probar es que es LI. Para cada $i = 1, \dots, m$, sea $\mathcal{B}_i = \{w_1^i, \dots, w_{l_i}^i\}$. Supongamos que tenemos escalares a_j^i tales que

$$a_1^1 w_1^1 + \cdots + a_{l_1}^1 w_{l_1}^1 + \cdots + a_1^m w_1^m + \cdots + a_{l_m}^m w_{l_m}^m = 0.$$

Como $a_1^i w_1^i + \cdots + a_{l_i}^i w_{l_i}^i \in W_i$ (para todo i) y la familia $\{W_1, \dots, W_m\}$ es independiente, deducimos que es

$$a_1^i w_1^i + \cdots + a_{l_i}^i w_{l_i}^i = 0, \quad \forall i = 1, \dots, m.$$

Luego como cada $\mathcal{B}_i$ es LI, deducimos que es $a_j^i = 0$, para todo i, j y por lo tanto $\mathcal{B}$ es LI. La segunda afirmación se deduce inmediatamente de la primera. $\square$

Corolario 7.4.22. *Sean $W_1, \dots, W_m$ subespacios de un espacio de dimensión finita V. Si $\{W_1, \dots, W_m\}$ es una familia independiente y vale $\dim V = \sum_{i=1}^m \dim W_i$, entonces $V = \bigoplus_{i=1}^m W_i$.* $\square$

7.5. Proyecciones

En esta sección V es un espacio vectorial de dimensión posiblemente infinita. A continuación definiremos las proyecciones y estudiaremos su relación con las sumas directas. Las proyecciones son los elementos básicos para definir la descomposición espectral, que veremos más adelante.

Sean U y W subespacios de V tales que $V = U \oplus W$. Definimos una función $T : V \to V$ mediante $T(v) = u$, si $v = u + w$, con $u \in U$, $w \in W$, para todo $v \in V$. Es un ejercicio el probar que T es una transformación lineal. La función T se llama la *proyección sobre U en la dirección de W*.

Observar que si $v = u + w$, con $u \in U$, $w \in W$, entonces $T(v) = u = u + 0$, con $u \in U$, $0 \in W$; luego $T(T(v)) = u = T(v)$, para todo $v \in V$, es decir $T^2 = T$, siendo $T^2 := T \circ T$.

Una *proyección* es un operador $T \in \mathcal{L}(V)$ que verifica $T^2 = T$. El comentario anterior muestra que para toda descomposición $V = U \oplus W$, la proyección sobre U en la dirección de W es una proyección. La proposición siguiente muestra que también vale el recíproco.

Proposición 7.5.1. *Sea $T \in \mathcal{L}(V)$ tal que $T^2 = T$, entonces $V = \operatorname{Im}(T) \oplus \operatorname{Ker}(T)$ y T es la proyección sobre $\operatorname{Im}(T)$ en la dirección de $\operatorname{Ker}(T)$.*

Dem. Sea $v \in V$. Observar que la condición $T^2 = T$ implica $v - T(v) \in \operatorname{Ker}(T)$. Luego escribiendo $v = T(v) + v - T(v)$, deducimos $V = \operatorname{Im}(T) + \operatorname{Ker}(T)$. Por otro lado, si $v \in \operatorname{Ker}(T) \cap \operatorname{Im}(T)$, entonces $T(v) = 0$ y existe $u \in V$ tal que $v = T(u)$. Luego, $0 = T(v) = T\big(T(u)\big) = T^2(u) = T(u) = v$. Esto prueba que vale $\operatorname{Ker}(T) \cap \operatorname{Im}(T) = \{0\}$ y por lo tanto $V = \operatorname{Im}(T) \oplus \operatorname{Ker}(T)$.

Sea $v \in V = \operatorname{Im}(T) \oplus \operatorname{Ker}(T)$. Por lo que vimos anteriormente es $v = T(v) + v - T(v)$, con $T(v) \in \operatorname{Im}(T)$ y $v - T(v) \in \operatorname{Ker}(T)$. Luego la proyección sobre $\operatorname{Im}(T)$ en la dirección de $\operatorname{Ker}(T)$ de v es $T(v)$. $\square$

El siguiente resultado generaliza lo anterior a una cantidad finita arbitraria de subespacios.

Proposición 7.5.2. *Consideremos las siguientes familias de datos.*

(A) $W_1, \ldots, W_k$ *son subespacios de* V *tales que* $V = \bigoplus_{i=1}^k W_i$.

(B) $P_1, \ldots, P_k \in \mathcal{L}(V)$ *son operadores que verifican las siguientes condiciones.*

1. $P_i^2 = P_i$, *para todo* $i = 1, \ldots, k$.
2. $P_i \circ P_j = 0$, *para todo* $i \neq j$, *con* $i, j = 1, \ldots, k$.
3. $\mathrm{Id} = P_1 + \cdots + P_k$.

Entonces **(A)** *y* **(B)** *son equivalentes.*

Dem. Sean $W_1, \ldots, W_k$ como en **(A)** y definimos $P_1, \ldots, P_k : V \to V$ por $P_i(v) = w_i$, si $v = \sum_{j=1}^k w_j$, con $w_j \in W_j$, para todo $j = 1, \ldots, k$ y todo $v \in V$. Notar que para cada $i \in \{1, \ldots, n\}$ es

$$V = W_i \oplus \left(\bigoplus_{j \neq i} W_j \right) \quad \text{y} \quad v = \sum_{j=1}^k w_j = w_i + \sum_{j \neq i} w_j, \quad \text{con } w_l \in W_l, \ \forall l = 1, \ldots, k.$$

Luego $P_i : V \to V$ es la proyección sobre W_i en la dirección de $\bigoplus_{j \neq i} W_j$; esto prueba que P_i es una proyección ($P_i \in \mathcal{L}(V)$ y $P_i^2 = P_i$), para todo $i = 1, \ldots, k$. Observar que por como definimos los operadores P_i, es

$$v = \sum_{i=1}^k P_i(v), \ \forall v \in V \quad \Rightarrow \quad \mathrm{Id} = P_1 + \cdots + P_k.$$

Finalmente, si $v \in V$, entonces $P_i(v) \in W_i$, por lo tanto su descomposición correspondiente a $V = \bigoplus_{i=1}^k W_i$ es $P_i(v) = 0 + \cdots + P_i(v) + \cdots + 0$, es decir, la componente de $P_i(v)$ en el subespacio W_j es 0 si $j \neq i$. Esto implica que vale $P_j\big(P_i(v)\big) = 0$, para todo $j \neq i$ y $v \in V$; luego $P_j \circ P_i = 0$, para todo $j \neq i$.

Supongamos ahora que $P_1, \ldots, P_k \in \mathcal{L}(V)$ verifican las condiciones de **(B)** y definimos $W_i = \mathrm{Im}\,(P_i)$, para todo i. De $\mathrm{Id} = \sum_{i=1}^k P_i$ deducimos $v = \sum_{i=1}^k P_i(v) \in \sum_{i=1}^k W_i$, para todo $v \in V$. Luego $V = \sum_{i=1}^k W_i$. Solo resta probar que $W_1, \ldots, W_k$ son independientes. Sean $w_i \in W_i$, $i = 1, \ldots, k$, tales que $w_1 + \cdots + w_k = 0$. Al ser $w_i \in W_i = \mathrm{Im}\,(P_i)$, entonces existe $v_i \in V$ tal que $w_i = P_i(v_i)$, para todo i. Luego para cada i, es

$$0 = P_i\left(\sum_{j=1}^k w_j \right) = P_i\left(\sum_{j=1}^k P_j(v_j) \right) = \sum_{j=1}^k P_i(P_j(v_j)) = \sum_{j=1}^k (P_i \circ P_j)(v_j) = P_i(v_i) = w_i.$$

Por lo tanto es $w_1 = \cdots = w_k = 0$. Esto muestra que $W_1, \ldots, W_k$ son independientes y concluye la prueba de que es $V = \bigoplus_{i=1}^k W_i$. □

Observación 7.5.3. En las hipótesis de la observación anterior. Notar que si partimos de $W_1, \ldots, W_k$ y les asociamos las correspondientes proyecciones $P_1, \ldots, P_k$, entonces $\mathrm{Im}\,(P_i) = W_i$, para todo $i = 1, \ldots, k$. Recíprocamente, si partimos de proyecciones $P_1, \ldots, P_k$ y definimos $W_i = \mathrm{Im}\,(P_i)$, para todo $i = 1, \ldots, k$, entonces las proyecciones asociadas a la descomposición $V = \bigoplus_{i=1}^k W_i$, son exactamente $P_1, \ldots, P_k$. Es decir, las correspondencias que vimos entre proyecciones y sumas directas, son inversas una de la otra.

7.6. La descomposición espectral de un operador

En el capítulo 1 estudiamos las propiedades de los operadores que son diagonalizables y en el 3 lo hicimos con los operadores en espacios con producto interno que son diagonalizables en bases ortonormales. En esta sección veremos que lo anterior está vinculado con poder descomponer el operador como combinación lineal de proyecciones. Primero lo estudiaremos para espacios vectoriales en general y luego para espacios con producto interno. En esta sección los espacios son siempre de dimensión finita.

Operadores diagonalizables En el capítulo 1 vimos que un operador T en un espacio V es diagonalizable si y solo si V se descompone en suma directa de subespacios propios de T. A su vez sabemos que descomponer el espacio en suma directa de subespacios equivale a tener una cierta descomposición del operador identidad en suma de proyecciones. El siguiente teorema combina ambos resultados.

Teorema 7.6.1. *Sea V un espacio vectorial y T un operador en V. Si T es diagonalizable, entonces existen únicos escalares distintos $\lambda_1,\dots,\lambda_h$ y operadores no nulos $P_1,\dots,P_h \in \mathcal{L}(V)$ tales que*

1. $P_i^2 = P_i$, $\forall i = 1,\dots,h$ *(los operadores P_i son proyecciones).*
2. $P_i \circ P_j = 0$ *si* $i \neq j$.
3. $\mathrm{Id} = P_1 + \cdots + P_h$.
4. $T = \lambda_1 P_1 + \cdots + \lambda_h P_h$.

Dem. Existencia. Como T es diagonalizable, entonces es $V = \bigoplus_{i=1}^h E_{\lambda_i}$, siendo $\lambda_1,\dots,\lambda_h$ los distintos valores propios de T. Para cada $i \in \{1,\dots,h\}$, definimos $P_i : V \to V$ por $P_i(v) = v_i$ si $v = v_1 + \cdots + v_h$, con $v_i \in E_{\lambda_i}$, $i = 1,\dots,h$. Por las relaciones que conocemos entre proyecciones y sumas directas, sabemos que $P_1,\dots,P_h$ verifican las condiciones 1, 2 y 3. Además como $E_{\lambda_i} \neq \{0\}$, es $P_i \neq 0$, $\forall i = 1,\dots,n$.

Si $v \in V$, es $v = P_1(v) + \cdots + P_h(v)$ con $P_i(v) \in E_{\lambda_i}$, $\forall i = 1,\dots,h$. Luego

$$\begin{aligned} T(v) &= T(P_1(v) + \cdots + P_h(v)) = T(P_1(v)) + \cdots + T(P_h(v)) = \lambda_1 P_1(v) + \cdots + \lambda_h P_h(v) \\ &= (\lambda_1 P_1 + \cdots + \lambda_h P_h)(v), \ \forall v \in V \quad \Rightarrow \quad T = \lambda_1 P_1 + \cdots + \lambda_h P_h. \end{aligned}$$

Unicidad. Sea $T \in \mathcal{L}(V)$ tal que existen escalares distintos $\lambda_1,\dots,\lambda_h$ y operadores no nulos $P_1,\dots,P_h$ que verifican las condiciones 1, 2, 3 y 4. Probaremos que T es diagonalizable, que $\lambda_1,\dots,\lambda_h$ son los distintos valores propios de T y que $P_1,\dots,P_h$ son las proyecciones asociadas a la descomposición $V = \bigoplus_{i=1}^h E_{\lambda_i}$.

Sea $W_i = \mathrm{Im}(P_i)$, $i = 1,\dots,h$. Las condiciones 1, 2 y 3 implican que $V = \bigoplus_{i=1}^h W_i$ y que si $v = w_1 + \cdots + w_h$, con $w_i \in W_i$, entonces $P_i(v) = w_i$, $\forall i = 1,\dots h$. Además, como $P_i \neq 0$, es $W_i \neq \{0\}$ $\forall i = 1,\dots,h$.

Si $v \in W_i$, es $P_i(v) = v$, luego

$$T(v) = \left(\sum_{j=1}^h \lambda_j P_j\right)(v) = \sum_{j=1}^h \lambda_j P_j(v) = \sum_{j=1}^h \lambda_j P_j(P_i(v)) = \sum_{j=1}^h \lambda_j (P_j \circ P_i)(v) = \lambda_i P_i^2(v) = \lambda_i P_i(v) = \lambda_i v.$$

Como $W_i \neq \{0\}$, la relación anterior implica que λ_i es un valor propio de T y que $W_i \subset E_{\lambda_i}$, para todo $i = 1,\dots,h$. Además, como $\lambda_1,\dots,\lambda_h$ son distintos entre sí, entonces $E_{\lambda_1},\dots,E_{\lambda_h}$ son independientes, luego

$$V = \bigoplus_{i=1}^h W_i = \sum_{i=1}^h W_i \subset \sum_{i=1}^h E_{\lambda_i} = \bigoplus_{i=1}^h E_{\lambda_i} \subset V. \tag{7.6}$$

Luego $V = \bigoplus_{i=1}^h W_i = \bigoplus_{i=1}^h E_{\lambda_i}$. Esto implica que T es diagonalizable y que $\lambda_1, \ldots, \lambda_h$ son los distintos valores propios de T. Observar que la igualdad anterior implica $\dim V = \sum_{i=1}^h \dim W_i = \sum_{i=1}^h \dim E_{\lambda_i}$. Por otro lado de $W_i \subset E_{\lambda_i}$ se deduce que $\dim W_i \leq \dim E_{\lambda_i}$ para todo $i = 1, \ldots, h$. De estas dos últimas relaciones se tiene que $\dim W_i = \dim E_{\lambda_i}$ y como $W_i \subset E_{\lambda_i}$, es $W_i = E_{\lambda_i}$ para todo $i = 1, \ldots, h$. Esto implica que $P_1, \ldots, P_h$ son las proyecciones asociadas a la descomposición $V = \bigoplus_{i=1}^h E_{\lambda_i}$. $\square$

Definición 7.6.2. Dado un operador T, el *espectro*[2] de T es el conjunto de sus valores propios. Si T es diagonalizable, a la descomposición $T = \lambda_1 P_1 + \cdots + \lambda_h P_h$ dada por el teorema anterior, se le llama la *descomposición espectral* de T.

El teorema anterior prueba que todo operador diagonalizable T tiene una descomposición espectral. A continuación veremos que las proyecciones que aparecen en esa descomposición espectral se pueden obtener como polinomios evaluados en T.

Proposición 7.6.3. *Si $T = \sum_{i=1}^h \lambda_i P_i$ es la descomposición espectral de un operador diagonalizable y $q(x) \in \Bbbk[x]$, entonces $q(T) = \sum_{j=1}^h q(\lambda_j) P_j$.*

Dem. Empezamos calculando T^2.

$$T^2 = \left(\sum_{i=1}^h \lambda_i P_i\right)^2 = \left(\sum_{i=1}^h \lambda_i P_i\right) \circ \left(\sum_{j=1}^h \lambda_j P_j\right) = \sum_{i=1}^h \lambda_i \sum_{j=1}^h \lambda_j P_i \circ P_j = \sum_{i=1}^h \lambda_i \lambda_i P_i \circ P_i = \sum_{i=1}^h \lambda_i^2 P_i.$$

Luego $T^2 = \sum_{i=1}^h \lambda_i^2 P_i$. En forma análoga se prueba por inducción que vale $T^n = \sum_{i=1}^h \lambda_i^n P_i$, para todo $n \in \mathbb{N}$. Luego si $q(x) = \sum_{i=1}^m a_i x^i \in \Bbbk[x]$, entonces

$$q(T) = \sum_{j=1}^m a_j T^j = \sum_{j=1}^m a_j \left(\sum_{i=1}^h \lambda_i^j P^i\right) = \sum_{j=1}^m \sum_{i=1}^h a_j \lambda_i^j P^i = \sum_{i=1}^h \left(\sum_{j=1}^m a_j \lambda_i^j\right) P^i = \sum_{i=1}^h q(\lambda_i) P_i. \quad \square$$

Proposición 7.6.4. *Sea $T \in \mathcal{L}(V)$ un operador diagonalizable y consideremos su descomposición espectral $T = \sum_{i=1}^k \lambda_i P_i$. Entonces cada P_i es un polinomio en T.*

Dem. Dado que $\lambda_1, \ldots, \lambda_k$ son escalares distintos, entonces existen polinomios $q_1(x), \ldots, q_k(x)$ llamados *polinomios de Lagrange* que verifican $q_i(\lambda_j) = \delta_{ij}$, para todo $i, j = 1, \ldots, k$ (ver la sección 7.1); luego

$$q_i(T) = \sum_{j=1}^k q_i(\lambda_j) P_j = \sum_{j=1}^k \delta_{ij} P_j = P_i \quad \Rightarrow \quad P_i = q_i(T),\ \forall i = 1, \ldots k, \quad \square$$

Ejemplo 7.6.5. Sea $T \in \mathcal{L}(\mathbb{R}^3)$ definida por $T(x, y, z) = (4x + z, 2x + 3y + 2z, x + 4z)$. Observar que es $T = L_A$, siendo

$$A = \begin{pmatrix} 4 & 0 & 1 \\ 2 & 3 & 2 \\ 1 & 0 & 4 \end{pmatrix} \quad \Rightarrow \quad \chi_T(t) = \begin{vmatrix} 4-t & 0 & 1 \\ 2 & 3-t & 2 \\ 1 & 0 & 4-t \end{vmatrix} = -(t-3)^2(t-5). \tag{7.7}$$

La multiplicidad algebraica de cada valor propio la obtenemos aplicando la parte 3 de la observación 1.4.2,

$$MG(3) = 3 - \operatorname{rango}(A - 3I) = 2, \quad MG(5) = 3 - \operatorname{rango}(A - 5I) = 1.$$

[2] El espectro de un operador es un concepto que se usa principalmente en dimensión infinita. En dimensión finita coincide con el que definimos acá.

Luego χ_T escinde y vale $MG(3) = MA(3) = 2$ y $MG(5) = MA(5) = 1$; esto implica que T es diagonalizable y que su descomposición espectral es $T = 3P_3+5P_5$, siendo P_3 y P_5 las proyecciones asociadas a $\mathbb{R}^3 = E_3 \oplus E_5$. Notar que estamos escribiendo P_3 y P_5, en vez de P_1 y P_2, dado que para cálculos concretos esta notación es más sugerente (es $\operatorname{Im} P_3 = E_3$ y $\operatorname{Im} P_5 = E_5$). Los polinomios de Lagrange correspondientes a (3,5) son

$$q_3(x) = \frac{x-5}{3-5} = -\frac{1}{2}(x-5), \qquad q_5(x) = \frac{x-3}{5-3} = \frac{1}{2}(x-3).$$

Luego la proposición anterior nos dice que las proyecciones P_3 y P_5 se obtienen mediante

$$P_3 = q_3(T) = -\frac{1}{2}(T - 5\mathrm{Id}), \qquad P_5 = q_5(T) = \frac{1}{2}(T - 3\mathrm{Id}).$$

Para obtener P_3 y P_5 explícitamente, observar que si $\mathcal{B}$ es la base canónica de $\mathbb{R}^3$ y A es la matriz definida en (7.7), entonces

$$[P_3]_{\mathcal{B}} = -\frac{1}{2}(A-5I) = -\frac{1}{2}\begin{pmatrix} -1 & 0 & 1 \\ 2 & -2 & 2 \\ 1 & 0 & -1 \end{pmatrix}, \qquad [P_5]_{\mathcal{B}} = \frac{1}{2}(A-3I) = \frac{1}{2}\begin{pmatrix} 1 & 0 & 1 \\ 2 & 0 & 2 \\ 1 & 0 & 1 \end{pmatrix}. \tag{7.8}$$

Luego

$$P_3(x,y,z) = \left(\frac{x-z}{2}, y-x-z, \frac{z-x}{2}\right), \quad P_5(x,y,z) = \left(\frac{x+z}{2}, x+z, \frac{z+x}{2}\right)$$

Notar que obtuvimos las proyecciones sin tener que hallar los subespacios propios de T. De hecho conociendo las proyecciones podemos obtener los subespacios propios. Sabemos que es $\operatorname{Im} P_3 = E_3$ y $\operatorname{Im} P_5 = E_5$, luego de las fórmulas (7.8) deducimos

$$E_3 = \operatorname{Im} P_3 = [(-1,2,1), (0,2,0)], \quad E_5 = \operatorname{Im} P_5 = [(1,2,1)].$$

Notar que los generadores anteriores son bases de los subespacios propios respectivos.

Operadores diagonalizables en bases ortonormales Lo que sigue es la versión del teorema 7.6.1 para operadores en espacios con producto interno.

Teorema 7.6.6 (Teorema espectral)**.** *Sea V un espacio vectorial con producto interno. Supongamos que $T \in \mathcal{L}(V)$ es un operador normal si $\Bbbk = \mathbb{C}$ o autoadjunto si $\Bbbk = \mathbb{R}$. Entonces existen únicos $\lambda_1, \dots, \lambda_k \in \Bbbk$ distintos y $P_1, \dots, P_k \in \mathcal{L}(V)$ no nulos, tales que:*

1. $P_i^2 = P_i = P_i^*$, $\forall i = 1, \dots, k$.
2. $P_i \circ P_j = 0$ *si* $i \neq j$.
3. $\mathrm{Id} = \sum_{i=1}^k P_i$.
4. $T = \sum_{i=1}^k \lambda_i P_i$.

Dem. Como T es normal si $\Bbbk = \mathbb{C}$ o autoadjunto si $\Bbbk = \mathbb{R}$, entonces el teorema 3.4.10 implica que T es diagonalizable en una base ortonormal.

Al ser T diagonalizable, entonces el teorema 7.6.1 implica que existen únicos $\lambda_1, \dots, \lambda_k \in \Bbbk$ distintos y $P_1, \dots, P_k \in \mathcal{L}(V)$ no nulos, tales que se verifican las propiedades 1,2,3,4, salvo eventualmente la condición $P_i = P_i^*$, para todo i. Recordar que $P_1, \dots, P_k$ son las proyecciones asociadas a la descomposición de V en suma directa de subespacios propios $V = \bigoplus_{i=1}^k E_{\lambda_i}$. Como T es normal ($\Bbbk = \mathbb{C}$ o $\mathbb{R}$), es $E_{\lambda_i} \perp E_{\lambda_j}$, para todo $i \neq j$. Luego la proposición 3.2.4 implica que las proyecciones P_i son proyecciones ortogonales. Esto concluye la prueba de la existencia. La unicidad se deduce inmediatamente de la unicidad en el caso clásico (el caso de V sin producto interno y T un operador diagonalizable). $\square$

Observación 7.6.7. Sea $T \in \mathcal{L}(V)$ normal si $\Bbbk = \mathbb{C}$ o autoadjunto si $\Bbbk = \mathbb{R}$. Sean $\lambda_1, \dots, \lambda_k$ los valores propios distintos de T y $P_1, \dots, P_k$ las proyecciones ortogonales sobre los subespacios propios correspondientes. Recordar que a la descomposición

$$T = \lambda_1 P_1 + \cdots + \lambda_k P_k$$

se le llama la descomposición espectral de T y a $\{\lambda_1, \dots, \lambda_k\}$ el espectro de T. Recordar también que en la Proposición 1.22 probamos que cada P_i es un polinomio en T.

Corolario 7.6.8. *Sea $\Bbbk = \mathbb{C}$ y $T \in \mathcal{L}(V)$. Entonces T es normal si y solo si $T^* = p(T)$ para algún $p \in \mathbb{C}[x]$.*

Dem. Si $T^* = p(T)$ con $p = \sum_{i=0}^{n} a_i\, x^i$, entonces

$$T^* \circ T = p(T) \circ T = \left(\sum_{i=0}^{n} a_i\, T^i\right) \circ T = \sum_{i=0}^{n} a_i\, T^{i+1} = T \circ \sum_{i=0}^{n} a_i\, T^i = T \circ p(T) = T \circ T^*,$$

luego T es normal.

Si T es normal y $T = \sum_{i=1}^{k} \lambda_i\, P_i$ es la descomposición espectral de T, entonces $T^* = \sum_{i=1}^{k} \overline{\lambda_i}\, P_i$. Como $\lambda_1, \dots, \lambda_k$ son distintos entre sí, entonces la proposición 7.1.2 implica que existe $p \in \Bbbk[x]$ tal que $p(\lambda_i) = \overline{\lambda_i}$, $\forall i = 1, \dots, k$. Entonces aplicando la proposición 7.6.3 obtenemos

$$p(T) = \sum_{i=1}^{k} p(\lambda_i)\, P_i = \sum_{i=1}^{k} \overline{\lambda_i}\, P_i = T^*. \quad \square$$

Observación 7.6.9. El corolario anterior no requiere el teorema espectral, y se puede probar usando la matriz asociada al operador en una base ortonormal de vectores propios, en vez de la descomposición espectral. Sin embargo lo ubicamos en este lugar por dos razones, una es para mostrar cómo se aplica el teorema espectral y la otra es que requiere la evaluación de polinomios en transformaciones lineales, que recién lo vimos en el capítulo 5, que es posterior al de espacios con producto interno.

7.7. Movimientos rígidos

En esta sección V es un espacio vectorial real no nulo con producto interno. En forma análoga a lo que sucede en $\mathbb{R}^2$ y $\mathbb{R}^3$, el producto interno $\langle\ ,\ \rangle$ induce una norma $\|\ \|$, y esta induce una distancia d entre los puntos de V, definiendo $d(p, q) = \|p - q\|$, para todo $p, q \in V$. Estudiaremos las funciones de V en V que preservan esta distancia, poniendo énfasis en los casos en que V es el plano $\mathbb{R}^2$ o el espacio $\mathbb{R}^3$. Como siempre el producto interno en estos últimos es el producto escalar.

Definición 7.7.1. Un *movimiento rígido* (abreviadamente *movimiento*) en V es una función $M : V \to V$ que preserva la distancia, es decir, que verifica $\|M(u) - M(v)\| = \|u - v\|$, para todo $u, v \in V$.

Ejemplos 7.7.2.

1. La función identidad claramente es un movimiento.
2. Una *traslación* es una función $\mathcal{T}_v : V \to V$ de la forma $\mathcal{T}_v(p) = p + v$, para todo $p \in V$, donde $v \in V$ es un vector fijo. Notar que $\mathcal{T}_v$ es una función invertible con inversa $\mathcal{T}_v^{-1} = \mathcal{T}_{-v}$. Es fácil de probar que toda traslación es un movimiento.
3. Si $T : V \to V$ es una isometría (es decir, T es una transformación lineal que verifica $\langle T(u), T(v)\rangle = \langle u, v\rangle$, para todo $u, v \in V$), entonces T preserva la norma (proposición 3.5.3) y por lo tanto

   $$\|T(u) - T(v)\| = \|T(u - v)\| = \|u - v\|, \quad \forall u, v \in V.$$

   Luego toda isometría es un movimiento.

Además de las traslaciones, en el plano y el espacio existen otros movimientos que describiremos a continuación. En las rotaciones se asume que el ángulo de giro es tomado en sentido positivo.

Ejemplos 7.7.3. 1. *Movimientos en el plano.*

   a) *Rotación* de centro en un punto $p \in \mathbb{R}^2$ y ángulo $\theta \in \mathbb{R}$. A la rotación de centro p y ángulo π se le llama *simetría central* de centro p. Notaciones: $R_{p,\theta}$ y $C_p = R_{p,\pi}$.

   b) *Simetría axial* de eje la recta $l \subset \mathbb{R}^2$. Notación: S_l.

   c) *Simetría deslizante* o *antitraslación* $\mathcal{T}_v \circ S_l$, siendo v paralelo a l.

2. *Movimientos en el espacio.*

   a) *Rotación* de eje la recta $l \subset \mathbb{R}^3$ y ángulo $\theta \in \mathbb{R}$. Notación: $R_{l,\theta}$.

   b) *Movimiento helicoidal* $\mathcal{T}_v \circ R_{l,\theta}$, siendo v paralelo a l.

   c) *Simetría especular* respecto al plano $\mathcal{P} \subset \mathbb{R}^3$. Notación: $S_{\mathcal{P}}$.

   d) *Simetría deslizante* $\mathcal{T}_v \circ S_{\mathcal{P}}$, siendo v paralelo al plano $\mathcal{P}$.

   e) *Simetría rotacional* $R_{l,\theta} \circ S_{\mathcal{P}}$, siendo l ortogonal al plano $\mathcal{P}$.

Observaciones 7.7.4. Respecto de los movimientos en el plano y en el espacio.

1. Una traslación de vector nulo es la función identidad, y lo mismo sucede con una rotación de ángulo nulo, tanto en el plano como en el espacio. Luego en el plano una simetría axial es un caso particular de simetría deslizante.

2. Consideremos movimientos del espacio. Una simetría especular es un caso particular de simetría deslizante o rotacional. Una simetría axial es una rotación de ángulo π. Una simetría central es una simetría rotacional de ángulo π.

Proposición 7.7.5. *La composición de movimientos es un movimiento.*

Dem. Sean M_1, M_2 movimientos en V y $u, v \in V$. Entonces

$$\|(M_2 \circ M_1)(u) - (M_2 \circ M_1)(v)\| = \big\|M_2\big(M_1(u)\big) - M_2\big(M_1(v)\big)\big\| = \|M_1(u) - M_1(v)\| = \|u - v\|. \quad \square$$

Proposición 7.7.6. *Todo movimiento M en V se escribe de forma única como $M(v) = T(v) + v_0$ para todo $v \in V$, donde v_0 es un vector fijo y $T \in \mathcal{L}(V)$ es una isometría. Luego los movimientos de V se obtienen componiendo isometrías con traslaciones.*

Dem. Para probar la unicidad, notar que si $v_0 \in V$ y $T \in \mathcal{L}(V)$ son tales que $M(v) = T(v) + v_0$, para todo $v \in V$, entonces tomando $v = 0$ deducimos que es $v_0 = M(0)$ y por lo tanto despejando T obtenemos $T(v) = M(v) - v_0 = M(v) - M(0)$, para todo $v \in V$. Luego v_0 y T quedan determinados por M.

Prueba de la existencia de la descomposición. Sea $v_0 = M(0)$. Definimos una función $T : V \to V$ mediante $T(v) = M(v) - v_0$, para todo $v \in V$. Luego $M(v) = T(v) + v_0$ para todo $v \in V$. A continuación probaremos las siguientes afirmaciones.

1. $\|T(u) - T(v)\| = \|u - v\|$, para todo $u, v \in V$;

2. $\|T(v)\| = \|v\|$, para todo $v \in V$;

3. $\langle T(u), T(v)\rangle = \langle u, v\rangle$, para todo $u, v \in V$;

4. $T(au+v) = aT(u) + T(v)$, para todo $a \in \mathbb{R}$, $u, v \in V$. Es decir, T es lineal.

Notar que de las dos últimas afirmaciones se obtiene que T es una isometría, lo cual completa la prueba de la proposición. La prueba de las dos primeras afirmaciones es la siguiente.

$$\|T(u) - T(v)\| = \left\|\big(M(u) - v_0\big) - \big(M(v) - v_0\big)\right\| = \|M(u) - M(v)\| = \|u - v\|,$$
$$\|T(v)\| = \|M(v) - v_0\| = \|M(v) - M(0)\| = \|v - 0\| = \|v\|.$$

Para probar la tercera, observar que de la fórmula (2.2) se deduce que para todo $u, v \in V$ vale

$$\|u - v\|^2 = \|u\|^2 + \|v\|^2 - 2\langle u, v\rangle \quad \Rightarrow \quad \langle u, v\rangle = \frac{1}{2}\left\{\|u\|^2 + \|v\|^2 - \|u - v\|^2\right\}. \tag{7.9}$$

Luego de (7.9) y las dos primeras afirmaciones se deduce la tercera. Para probar la cuarta usamos (7.9) y las tres afirmaciones anteriores. Sean $u, v \in V$ y $a \in \mathbb{R}$, entonces

$$\begin{aligned}
&\left\|T(au+v) - \big(aT(u) + T(v)\big)\right\|^2 = \\
&= \left\|\big(T(au+v) - T(v)\big) - aT(u)\right\|^2 \\
&= \|T(au+v) - T(v)\|^2 + \|aT(u)\|^2 - 2\langle T(au+v) - T(v), aT(u)\rangle \\
&= \|(au+v) - v\|^2 + a^2\|T(u)\|^2 - 2a\big(\langle T(au+v), T(u)\rangle - \langle T(v), T(u)\rangle\big) \\
&= \|au\|^2 + a^2\|u\|^2 - 2a\big(\langle au+v, u\rangle - \langle v, u\rangle\big) = 0.
\end{aligned}$$

Luego $\left\|T(au+v) - \big(aT(u) + T(v)\big)\right\|^2 = 0$ y esto equivale a la afirmación 4. □

Aplicación 7.7.1. *En $\mathbb{R}^2$ las isometrías son los operadores de la forma L_A, en que $A \in M_2(\mathbb{R})$ es ortogonal. Luego los movimientos en $\mathbb{R}^2$ son las funciones $M : \mathbb{R}^2 \to \mathbb{R}^2$ de la forma*

$$M(x, y) = (ax + by + m,\; cx + dy + n), \quad \forall (x, y) \in \mathbb{R}^2,$$

en que $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ es una matriz ortogonal y $m, n \in \mathbb{R}$ son arbitrarios. En forma análoga se describen los movimientos en $\mathbb{R}^n$, para todo n.

Para operar con movimientos escritos de la forma $M = \mathcal{T}_v \circ T$ es útil la siguiente proposición. En la misma no es necesario que los operadores sean isometrías.

Proposición 7.7.7. *Sean $v, v_1, v_2 \in V$ y $T, T_1, T_2 \in \mathcal{L}(V)$. Valen las siguientes relaciones.*

1. *Si $v \in V$ y $T \in \mathcal{L}(V)$, entonces $T \circ \mathcal{T}_v = \mathcal{T}_{T(v)} \circ T$.*
2. *$M_1 = \mathcal{T}_{v_1} \circ T_1$ y $M_2 = \mathcal{T}_{v_2} \circ T_2$, entonces $M_1 \circ M_2 = \mathcal{T}_u \circ (T_1 \circ T_2)$, siendo $u = v_1 + T_1(v_2)$.*
3. *Si $M = \mathcal{T}_v \circ T$ y T es invertible, entonces M es invertible y $M^{-1} = \mathcal{T}_w \circ T^{-1}$, siendo $w = -T^{-1}(v)$.*

Dem. La prueba de la primera igualdad es la siguiente.

$$(T \circ \mathcal{T}_v)(x) = T\big(\mathcal{T}_v(x)\big) = T(x+v) = T(x) + T(v) = \mathcal{T}_{T(v)}\big(T(x)\big) = (\mathcal{T}_{T(v)} \circ T)(x), \quad \forall x \in V.$$

Usando esa fórmula se obtienen las otras dos.

$$M_1 \circ M_2 = \mathcal{T}_{v_1} \circ T_1 \circ \mathcal{T}_{v_2} \circ T_2 = \mathcal{T}_{v_1} \circ \mathcal{T}_{T_1(v_2)} \circ T_1 \circ T_2 = \mathcal{T}_{v_1 + T_1(v_2)} \circ (T_1 \circ T_2),$$
$$M^{-1} = (\mathcal{T}_v \circ T)^{-1} = T^{-1} \circ (\mathcal{T}_v)^{-1} = T^{-1} \circ \mathcal{T}_{-v} = \mathcal{T}_{-T^{-1}(v)} \circ T^{-1}. \quad \square$$

Observación 7.7.8. Si $M = \mathcal{T}_v \circ T$ es un movimiento y a T le llamamos la *parte lineal* de M, entonces la proposición anterior muestra que la parte lineal de la composición de dos movimientos es la composición de sus partes lineales, y la parte lineal de la inversa de un movimiento es la inversa de su parte lineal.

En lo que sigue asumiremos que V es de dimensión finita.

Proposición 7.7.9. *Los movimientos en V forman un grupo respecto a la composición.*

Dem. Ya sabemos que la composición de movimientos es un movimiento y que la identidad también lo es, luego lo único que nos resta probar es que si M es un movimiento, entonces M es una función invertible y su inversa M^{-1} también es un movimiento. Sabemos que es $M = \mathcal{T}_v \circ T$, para cierto $v \in V$ y cierta isometría $T \in \mathcal{L}(V)$. Como V es de dimensión finita, entonces T es invertible y T^{-1} también es una isometría (observación 3.5.7). Luego aplicando la proposición anterior obtenemos que $M = \mathcal{T}_v \circ T$ es invertible y $M^{-1} = \mathcal{T}_w \circ T^{-1}$, siendo $w = -T^{-1}(v)$, lo cual implica que M^{-1} es también un movimiento. $\square$

De ahora en adelante las bases son bases ordenadas. Recordar que si $\mathcal{B}$ y $\mathcal{C}$ son bases de V, entonces ${}_{\mathcal{B}}[\mathrm{Id}]_{\mathcal{C}}$ es una matriz invertible y por lo tanto $\det {}_{\mathcal{B}}[\mathrm{Id}]_{\mathcal{C}} \neq 0$,

Orientación. En las rotaciones, tanto del plano como del espacio, es necesario fijar un sentido de giro como positivo, para poder indicar hacia dónde estamos girando. Pensando en esto introducimos la siguiente definición.

Definición 7.7.10. Sean $\mathcal{B}$ y $\mathcal{C}$ bases de V. Si $\det {}_{\mathcal{B}}[\mathrm{Id}]_{\mathcal{C}} > 0$, entonces $\mathcal{B}$ y $\mathcal{C}$ tienen la *misma orientación* y si $\det {}_{\mathcal{B}}[\mathrm{Id}]_{\mathcal{C}} < 0$, entonces tienen *distinta orientación*.

Observaciones 7.7.11.

1. Tener *la misma orientación* es una relación de equivalencia en el conjunto de las bases de V. Esto se prueba fácilmente usando las fórmulas siguientes

$$\det {}_{\mathcal{B}}[\mathrm{Id}]_{\mathcal{B}} = \det I = 1; \quad \left(\det {}_{\mathcal{B}}[\mathrm{Id}]_{\mathcal{C}}\right)^{-1} = \det {}_{\mathcal{C}}[\mathrm{Id}]_{\mathcal{B}}; \quad \det {}_{\mathcal{B}}[\mathrm{Id}]_{\mathcal{C}} \cdot \det {}_{\mathcal{C}}[\mathrm{Id}]_{\mathcal{D}} = \det {}_{\mathcal{B}}[\mathrm{Id}]_{\mathcal{D}}.$$

2. Si $\mathcal{B} = \{v_1, v_2, \ldots, v_n\}$ es una base de V, entonces $\mathcal{B}^{\mathrm{op}} = \{-v_1, v_2, \ldots, v_n\}$ es otra base de V y vale $\det {}_{\mathcal{B}}[\mathrm{Id}]_{\mathcal{B}^{\mathrm{op}}} = -1$. Luego $\mathcal{B}$ y $\mathcal{B}^{\mathrm{op}}$ tienen distinta orientación.

3. Sea $\mathcal{B}$ una base de V fija y consideremos $\mathcal{B}^{\mathrm{op}}$ como en el ítem anterior. Si $\mathcal{C}$ es otra base, entonces

$$\det {}_{\mathcal{B}^{\mathrm{op}}}[\mathrm{Id}]_{\mathcal{C}} = \det {}_{\mathcal{B}^{\mathrm{op}}}[\mathrm{Id}]_{\mathcal{B}} \cdot \det {}_{\mathcal{B}}[\mathrm{Id}]_{\mathcal{C}} = -\det {}_{\mathcal{B}}[\mathrm{Id}]_{\mathcal{C}}.$$

   De acá deducimos que necesariamente es $\det {}_{\mathcal{B}}[\mathrm{Id}]_{\mathcal{C}} > 0$ o $\det {}_{\mathcal{B}^{\mathrm{op}}}[\mathrm{Id}]_{\mathcal{C}} > 0$ y por lo tanto $\mathcal{C}$ tiene la misma orientación que $\mathcal{B}$ o que $\mathcal{B}^{\mathrm{op}}$. Luego hay exactamente dos orientaciones en V.

Un *espacio orientado* es un espacio V en el cual hemos hecho una elección de una de sus dos orientaciones. Las bases que tienen esa orientación decimos que tienen sentido *positivo* y las otras que tienen sentido *negativo*. En $\mathbb{R}^n$ conveniremos que el sentido positivo es el de la base canónica; luego en $\mathbb{R}^2$ el sentido positivo es el antihorario (de e_1 a e_2), mientras que en $\mathbb{R}^3$ viene dado por la *regla de la mano derecha*: cerrando esa mano de e_1 a e_2, el pulgar apunta en la dirección de e_3.

Observación 7.7.12. En lo que sigue veremos que esta definición refleja la idea intuitiva que tenemos de orientación. Empezamos con el caso de $\mathbb{R}^2$. Es fácil de probar que una base y la correspondiente base “normalizada” (obtenida dividiendo cada vector por su norma) tienen el mismo sentido, así que podemos restringirnos a bases normalizadas. Sea $\mathcal{B} = \{f_1, f_2\}$ una base tal que $\|f_1\| = \|f_2\| = 1$ y sea $\mathcal{C} = \{e_1, e_2\}$ la base canónica. Como f_1, f_2 son versores, entonces están en el círculo de radio 1 centrado en el origen y por lo tanto existen $\alpha, \beta \in [0, 2\pi]$ tales que $f_1 = (\cos\alpha, \operatorname{sen}\alpha)$ y $f_2 = (\cos\beta, \operatorname{sen}\beta)$. Luego

$$\det {}_{\mathcal{C}}[\mathrm{Id}]_{\mathcal{B}} = \begin{vmatrix} \cos\alpha & \cos\beta \\ \operatorname{sen}\alpha & \operatorname{sen}\beta \end{vmatrix} = \cos\alpha \operatorname{sen}\beta - \cos\beta \operatorname{sen}\alpha = \operatorname{sen}\theta,$$

siendo $\theta = \beta - \alpha$ el ángulo entre f_1 y f_2. Supongamos que es $\beta > \alpha$ y por lo tanto $0 < \theta < 2\pi$. Entonces $\mathcal{B}$ es una base positiva si $0 < \theta < \pi$ y es negativa si $\pi < \theta < 2\pi$. Es decir que $\mathcal{B} = \{f_1, f_2\}$ es positiva si el ángulo de f_1 a f_2 (medido en sentido antihorario) es menor que π y es negativa en caso contrario. El caso $\beta < \alpha$ se deduce del anterior, dado que las bases $\mathcal{B} = \{f_1, f_2\}$ y $\mathcal{B}' = \{f_2, f_1\}$ tienen distinto sentido.

A continuación esbozamos el caso de $\mathbb{R}^3$, dejando los detalles para el lector. Sean $\mathcal{B} = \{f_1, f_2, f_3\}$ una base formada por versores y $\mathcal{C} = \{e_1, e_2, e_3\}$ la base canónica. Sea $\mathcal{B}_1 = \{g_1, g_2, g_3\}$ la base obtenida rotando la base $\mathcal{B}$ de forma tal que el subespacio generado por g_1 y g_2 coincida con el plano xy. Las rotaciones preservan el sentido del despacio (ver más adelante la proposición 7.7.20), luego $\mathcal{B}$ y $\mathcal{B}_1$ tienen el mismo sentido. Consideremos $\mathcal{B}_2 = \{g_1, g_2, e_3\}$. Es claro que $\mathcal{B}_2$ es una base y si $g_3 = (a, b, c)$, entonces $\det({}_{\mathcal{B}_2}[\mathrm{Id}]_{\mathcal{B}_1}) = c$. Como $\mathcal{C}$ es la base canónica y g_3 es un versor, entonces $c = \cos\theta$, siendo $0 < \theta < \pi$ el ańgulo entre e_3 y g_3. Luego $\mathcal{B}_1$ y $\mathcal{B}_2$ tienen el mismo sentido si y solo si g_3 está "del lado de arriba" del plano xy. Supongamos que estamos en ese caso y por lo tanto $\mathcal{B}$ y $\mathcal{B}_2$ tienen el mismo sentido. Si $g_1 = (x_1, y_1, 0)$ y $g_2 = (x_2, y_2, 0)$, entonces $\det({}_{\mathcal{C}}[\mathrm{Id}]_{\mathcal{B}_2}) = \left|\begin{smallmatrix} x_1 & x_2 \\ y_1 & y_2 \end{smallmatrix}\right|$. Luego $\det({}_{\mathcal{C}}[\mathrm{Id}]_{\mathcal{B}_2})$ coincide con el determinante de la matriz de cambio de base de $\{g_1, g_2\}$ a la base canónica de $\mathbb{R}^2$, pensando $\{g_1, g_2\} \subset \mathbb{R}^2$. Luego, asumiendo que g_3 está del lado de arriba del plano xy, obtenemos que $\mathcal{B}_2$ (y por lo tanto $\mathcal{B}$) es una base positiva si y solo si $\{g_1, g_2\}$ es una base positiva de $\mathbb{R}^2$. Esto coincide con la "regla de la mano derecha" descrita anteriormente.

Isometrías directas e indirectas. De ahora en adelante fijamos una orientación en V.

Observación 7.7.13. Sea $T \in \mathcal{L}(V)$ una isometría. Recordar que el determinante de T se define mediante $\det T := \det([T]_{\mathcal{B}}$, siendo $\mathcal{B}$ una base cualquiera de V (esto no depende de la elección de la base). Notar que si $\mathcal{B}$ es una base de V, entonces ${}_{\mathcal{B}}[\mathrm{Id}]_{T(\mathcal{B})} = [T]_{\mathcal{B}}$. Luego $\det {}_{\mathcal{B}}[\mathrm{Id}]_{T(\mathcal{B})} = \det T$, para toda base $\mathcal{B}$ de V.

De la observación anterior se deduce el siguiente resultado.

Proposición 7.7.14. *Si $T \in \mathcal{L}(V)$ es una isometría, entonces las siguientes afirmaciones son equivalentes.*

1. $\det T > 0$ *(*$\det T < 0$*).*
2. *Existe una base $\mathcal{B}$ de V tal que $T(\mathcal{B})$ tiene la misma (distinta) orientación que $\mathcal{B}$.*
3. *Para toda base $\mathcal{B}$ de V se cumple que $T(\mathcal{B})$ tiene la misma (distinta) orientación que $\mathcal{B}$.* □

En base a la proposición anterior damos la siguiente definición.

Definición 7.7.15. Sea $T \in \mathcal{L}(V)$ una isometría. Si $\det T > 0$ entonces diremos que T es una isometría *directa* y si $\det T < 0$ entonces diremos que T es una isometría *indirecta*.

Por la proposición anterior sabemos que T es directa si y solo si preserva la orientación de V, es decir, lleva bases positivas en positivas y negativas en negativas. Análogamente, T es indirecta si y solo si invierte la orientación de V, es decir, intercambia las bases positivas con las negativas.

Si M es un movimiento de V y lo escribimos de la forma $M(v) = T(v) + v_0$, donde v_0 es un vector fijo y $T \in \mathcal{L}(V)$ es una isometría, entonces diremos que M es un movimiento *directo* o *indirecto* de acuerdo a que T sea una isometría directa o indirecta, respectivamente.

Proposición 7.7.16. *Los movimientos directos forman un subgrupo del grupo de los movimientos. Es decir, la identidad es un movimientos directo, el inverso de un movimiento directo es también un movimiento directo y la composición de movimientos directos es un movimiento directo.*

Dem. Es claro que la identidad (que es una isometría) es un movimiento directo. Las propiedades del determinante implican que la composición de isometrías directas es una isometría directa y que lo mismo ocurre con la inversa de una isometría directa. Luego usando las fórmulas de la proposición 7.7.7 deducimos que lo mismo ocurre con los movimientos directos. □

Observación 7.7.17. Las traslaciones siempre son movimientos directos (independientemente de la orientación elegida).

Movimientos en el plano y el espacio. En lo que sigue clasificaremos los movimientos en el plano y en el espacio. Para ello, por la proposición 7.7.6 alcanza con determinar las isometrías de $\mathbb{R}^2$ y $\mathbb{R}^3$, y los resultados de componer esas isometrías con las traslaciones. La identificación de los movimientos directos se obtiene aplicando la proposición 7.7.14. Recordar que si T es una isometría en $\mathbb{R}^n$, entonces $\det T = \pm 1$.

Movimientos en el plano.

Proposición 7.7.18. *Si T es una isometría de $\mathbb{R}^2$, entonces T es una rotación si* $\det T = 1$ *o una simetría axial si* $\det T = -1$.

Dem. Sea $T \in \mathcal{L}(\mathbb{R}^2)$ una isometría. Sabemos que es $T = L_A$, siendo $A = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ una matriz ortogonal. Además $\det A = \det T = \pm 1$.

Supongamos $\det A = 1$. Sabemos que las columnas de A forman una base ortonormal de $\mathbb{R}^2$, luego (b,d) es ortogonal con (a,c). Como $(-c,a)$ es no nulo y también es ortogonal con (a,c), deducimos que existe $k \in \mathbb{R}$ tal que $(b,d) = k(-c,a)$. La condición $\det A = 1$ junto con $1 = \|(a,c)\|^2 = a^2 + c^2$ implican $k = 1$, luego $(b,d) = (-c,a)$. Como $a^2 + c^2 = 1$, entonces existe un único $\alpha \in [0, 2\pi)$ tal que $a = \cos\alpha$ y $c = \operatorname{sen}\alpha$. Luego

$$A = \begin{pmatrix} a & -c \\ c & a \end{pmatrix} = \begin{pmatrix} \cos\alpha & -\operatorname{sen}\alpha \\ \operatorname{sen}\alpha & \cos\alpha \end{pmatrix}$$

y por lo tanto T es la rotación de centro en el origen y ángulo α.

Si $\det A = -1$, es $\chi_T(t) = t^2 - (a+d)t - 1$, luego su discriminante es positivo y por lo tanto tiene dos raíces reales cuyo producto es -1. Como los valores propios de A solo pueden ser ± 1, deducimos que existe $\mathcal{B} = \{w_1, w_2\}$ base ortonormal de $\mathbb{R}^2$ que verifica

$$T(w_1) = w_1, \quad T(w_2) = -w_2.$$

Luego T es la simetría axial de eje el subespacio $[w_1]$. □

Corolario 7.7.19. *Todo movimiento del plano es una traslación, rotación, simetría axial o simetría deslizante. Los directos son los dos primeros.*

Dem. Por la proposición 7.7.18 sabemos que las isometrías del plano son rotaciones con centros en el origen o simetrías axiales con ejes que pasan por el origen. Luego los movimientos del plano se obtienen componiendo traslaciones con esas rotaciones o simetrías axiales. La descripción final se obtiene observando que la composición de una rotación con una traslación da otra rotación con centro trasladado (a menos que la rotación sea la identidad, en cuyo caso da una traslación) y la composición de una traslación con una simetría axial da una simetría axial si el vector de traslación es ortogonal al eje de la simetría axial o una antitraslación en caso contrario. □

Movimientos en el espacio.

Proposición 7.7.20. *Si T es una isometría de $\mathbb{R}^3$, entonces T es una rotación si* $\det T = 1$, *o una simetría especular o rotacional si* $\det T = -1$.

Dem. Como el polinomio característico de T es de grado 3, entonces el teorema de Bolzano implica que tiene alguna raíz real λ. Sea v_0 un vector propio de T correspondiente a λ y $W = [v_0]^\perp$. Es fácil de probar que W es T invariante, luego $T|_W \in \mathcal{L}(W)$ y $\dim W = 2$.

Si $\mathcal{C} = \{w_1, w_2\}$ es una base de W, entonces $\mathcal{B} = \{v_0, w_1, w_2\}$ es una base de $\mathbb{R}^3$ y

$$[T]_\mathcal{B} = \begin{pmatrix} \lambda & 0 & 0 \\ 0 & a & b \\ 0 & c & d \end{pmatrix}, \quad [T|_W]_\mathcal{C} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}.$$

esto implica $\det T = \lambda \det(T|_W)$. Como T es una isometría sabemos que valen $\lambda = \pm 1$ y $\det T = \pm 1$. Discutiremos según estas posibilidades.

1. Si $\det T = 1$ y $\lambda = 1$, es $\det(T|_W) = 1$; luego $T|_W$ es una rotación en el plano W y por lo tanto T es una rotación en el espacio de eje $[v_0]$.

2. Si $\det T = 1$ y $\lambda = -1$, es $\det(T|_W) = -1$; luego $T|_W$ es una simetría axial de eje $[u_0]$ en el plano W y por lo tanto T es una simetría axial en el espacio con eje $[u_0]$, que es lo mismo que una rotación de eje $[u_0]$ y ángulo π.

3. Si $\det T = -1$ y $\lambda = 1$, es $\det(T|_W) = -1$; luego $T|_W$ es una simetría axial de eje $[u_0]$ en el plano W y por lo tanto T es una simetría especular respecto al subespacio $[v_0, u_0]$.

4. Si $\det T = -1$ y $\lambda = -1$, es $\det(T|_W) = 1$; luego $T|_W$ es una rotación en el plano W y por lo tanto T es una rotación en el espacio de eje $[v_0]$ compuesta con una simetría especular respecto a W. $\square$

Observación 7.7.21. Considerando el espacio, notar que toda traslación $\mathcal{T}_v$ se puede factorizar de la forma $\mathcal{T}_v = S_{\mathcal{P}_1} \circ S_{\mathcal{P}_2}$, siendo $\mathcal{P}_1$ y $\mathcal{P}_2$ dos planos paralelos que distan entre sí $\|v/2\|$. Análogamente, toda rotación $R_{l,\theta}$ se puede factorizar de la forma $R_{l,\theta} = S_{\mathcal{P}_1} \circ S_{\mathcal{P}_2}$, siendo $\mathcal{P}_1$ y $\mathcal{P}_2$ dos planos que se cortan en l y forman un ángulo $\theta/2$. Estas descomposiciones se usan en la prueba siguiente.

Corolario 7.7.22. *Todo movimiento del espacio es una traslación, rotación, movimiento helicoidal, simetría especular, simetría deslizante o simetría rotacional. Los directos son los tres primeros.*

Dem. Sabemos que si M es un movimiento del espacio, entonces es $M = \mathcal{T}_v \circ U$, siendo U una isometría de $\mathbb{R}^3$. La proposición 7.7.20 nos determina todas las posibilidades para U. Consideremos primero el caso $U = R_{l,\theta}$. Si v es paralelo a l, entonces M es un movimiento helicoidal. Si v es ortogonal a l, entonces es fácil de probar que M es una rotación $M = R_{l',\theta}$, siendo l' una recta paralela a l. En el caso general con v no ortogonal a l, escribimos $v = v_1 + v_2$, con $0 \neq v_1$ paralelo a l y v_2 ortogonal a l. Luego

$$M = \mathcal{T}_v \circ R_{l,\theta} = \mathcal{T}_{v_1+v_2} \circ R_{l,\theta} = \mathcal{T}_{v_1} \circ \mathcal{T}_{v_2} \circ R_{l,\theta} = \mathcal{T}_{v_1} \circ R_{l',\theta}.$$

con l' paralela a l y por lo tanto a v_1. Luego M es un movimiento helicoidal.

Consideremos ahora el caso en que $U = S_{\mathcal{P}}$, siendo $\mathcal{P}$ un plano que pasa por el origen. Si v es ortogonal a $\mathcal{P}$, entonces es fácil de probar que es $M = S_{\mathcal{P}'}$, siendo $\mathcal{P}'$ un plano paralelo a $\mathcal{P}$. Si v es paralelo a $\mathcal{P}$, entonces $M = \mathcal{T}_v \circ S_{\mathcal{P}}$ es una simetría con deslizamiento. En el caso general con v no ortogonal a l, escribiendo $v = v_1 + v_2$ como antes obtenemos que es $M = \mathcal{T}_{v_1} \circ S'_{\mathcal{P}}$, con v_1 paralelo a $\mathcal{P}'$, luego M es una simetría con deslizamiento.

El caso que nos resta es cuando $U = R_{l,\theta} \circ S_{\mathcal{P}} = S_{\mathcal{P}} \circ R_{l,\theta}$ es una simetría rotacional. Si v es ortogonal a $\mathcal{P}$, entonces es fácil de probar que es $M = S_{\mathcal{P}'} \circ R_{l,\theta}$, siendo $\mathcal{P}'$ un plano paralelo a $\mathcal{P}$; luego M es una simetría rotacional. Si v es paralelo a l, entonces es $M = \mathcal{T}_v \circ R_{l,\theta} \circ S_{\mathcal{P}} = R_{l',\theta} \circ S_{\mathcal{P}}$, luego M es también una simetría rotacional. Por el mismo argumento de antes, si v es genérico, entonces escribiendo $v = v_1 + v_2$, con v_1 paralelo a l y v_2 ortogonal a l, deducimos que M es una simetría rotacional. $\square$

7.8. Cónicas y cuádricas

En esta sección veremos la clasificación de las cónicas en el plano y de las cuádricas en el espacio. El producto interno es el producto escalar.

Empezamos con una definición general.

Definición 7.8.1. Una *cuádrica* en $\mathbb{R}^n$ ($n \geq 2$) es un conjunto de la forma

$$\mathcal{S} = \{(x_1, \ldots, x_n) \in \mathbb{R}^n : \ p(x_1, \ldots, x_n) = 0\}$$

donde p es un polinomio en n-variables de grado 2. Cuando $n = 2$ se dice que S es una *cónica*.

A nosotros nos interesa describir las cónicas en el plano $\mathbb{R}^2$ y las cuádricas en el espacio $\mathbb{R}^3$. En lo que sigue asumimos cierta familiaridad con las cónicas, por lo que no entraremos en mucho detalle de las mismas.

Ejemplo 7.8.2. Ejemplos de cónicas.

1. *Circunferencia:* $x^2 + y^2 = r^2$, $r > 0$ es el radio.
2. *Elipse:* $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$, $a, b > 0$. Notar que la circunferencia es un caso particular de elipse.
3. *Hipérbola:* $\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1$, $a, b > 0$.
4. *Parábola:* $y = ax^2 + bx + c$, $a, b, c \in \mathbb{R}$, $a \neq 0$.
5. *Dos rectas paralelas:* $x^2 + 2xy + y^2 - 1 = 0 \quad \Leftrightarrow \quad (x + y + 1)(x + y - 1) = 0$.
6. *Una recta doble:* $4x^2 + 4xy + y^2 - 4x - 2y + 1 = 0 \quad \Leftrightarrow \quad (2x + y - 1)^2 = 0$.

Ejemplo 7.8.3. Ejemplos de cuádricas.

1. *Esfera:* $x^2 + y^2 + z^2 = 1$.
2. *Cono circular:* $x^2 + y^2 - z^2 = 0$.
3. *Paraboloide:* $z = x^2 + y^2$.
4. *Dos planos secantes:* $x^2 + 2xy + y^2 - z^2 = 0 \quad \Leftrightarrow \quad (x + y + z)(x + y - z) = 0$.

Observación 7.8.4. Las cónicas en el plano son curvas y las cuádricas en el espacio son superficies. A estas últimas también se les llama *superficies cuádricas*.

A continuación veremos algunas propiedades generales de las cuádricas en $\mathbb{R}^n$ y luego nos restringiremos a las cónicas y las cuádricas.

Si en un polinomio de grado 2 separamos los términos de grado 2, 1 y 0, entonces siempre lo podemos escribir de la forma siguiente

$$p(x_1, \ldots, x_n) = \sum_{i=1}^{n} a_{ii} x_i^2 + \sum_{1 \leq i < j \leq n} 2a_{ij} x_i x_j + \sum_{i=1}^{n} b_i x_i + c.$$

La función $\Phi : \mathbb{R}^n \to \mathbb{R}$ definida por $\Phi(x_1, \ldots, x_n) = \sum_{i=1}^{n} a_{ii} x_i^2 + \sum_{1 \leq i < j \leq n} 2a_{ij} x_i x_j$ es una forma cuadrática que puede describirse mediante $\Phi(X) = X^t A X$, para todo $X \in \mathbb{R}^n$, siendo $A = (a_{ij}) \in M_n(\mathbb{R})$ una matriz simétrica y $X = (x_1, \ldots, x_n)$ escrito en forma de columna. Usando el producto escalar $\langle \ , \ \rangle$, podemos escribir

$$\Phi(X) = \langle X, AX \rangle = \langle AX, X \rangle, \quad \forall X \in \mathbb{R}^n.$$

Notar que al ser A una matriz simétrica, entonces vale $\langle X, AY \rangle = \langle AX, Y \rangle$, para todo $X, Y \in \mathbb{R}^n$. De la misma forma es $\sum_{i=1}^{n} b_i x_i = \langle B, X \rangle$, siendo $B = (b_1, \ldots, b_n)$. Luego una cuádrica en $\mathbb{R}^n$ es un conjunto de la forma

$$\mathcal{S} = \{X \in \mathbb{R}^n : \ \langle AX, X \rangle + \langle B, X \rangle + c = 0\},$$

siendo $A = (a_{ij}) \in M_n(\mathbb{R})$ una matriz simétrica no nula, $B = (b_1, \ldots, b_n)$ un vector de $\mathbb{R}^n$ y c un escalar. En lo que sigue asumiremos siempre que $\mathcal{S}$ está descrita de la forma anterior.

Definición 7.8.5. Una cuádrica $\mathcal{S}$ se dice *no degenerada* si lo es la forma cuadrática asociada Φ, es decir si $\det A \neq 0$. Decimos que $X_0 \in \mathbb{R}^3$ es un *centro* de $\mathcal{S}$ si verifica $2AX_0 + B = 0$.

Observaciones 7.8.6.
1. La cuádrica $\mathcal{S}$ es no degenerada si solo si lo es su forma cuadrática Φ.
2. Si $\mathcal{S}$ es no degenerada, entonces tiene un único centro $X_0 = -\frac{1}{2}A^{-1}B$. En caso contrario $\mathcal{S}$ puede tener infinitos centros o no tener ninguno (ver el ejemplo 7.8.12).

El siguiente resultado muestra que los centros definidos arriba son centros de simetría de la cuádrica.

Proposición 7.8.7. *Sea $\mathcal{S}$ una cuádrica que admite un centro X_0. Si un punto está en $\mathcal{S}$, entonces su simétrico respecto a X_0 también está en $\mathcal{S}$.*

Dem. Supongamos $X_1 \in \mathcal{S}$, luego X_1 verifica

$$\langle AX_1, X_1\rangle + \langle B, X_1\rangle + c = 0.$$

Observar que si X_2 es el simétrico de X_1 respecto a X_0, entonces $X_0 = \frac{1}{2}(X_1 + X_2)$, luego $X_2 = 2X_0 - X_1$. Calculando obtenemos

$$\begin{aligned}
\langle AX_2, X_2\rangle + \langle B, X_2\rangle + c &= \langle A(2X_0 - X_1), 2X_0 - X_1\rangle + \langle B, 2X_0 - X_1\rangle + c \\
&= \langle 2AX_0 - AX_1, 2X_0 - X_1\rangle + \langle B, 2X_0 - X_1\rangle + c \\
&= \langle 2AX_0 - AX_1 + B, 2X_0 - X_1\rangle + c \\
&= \langle -AX_1, 2X_0 - X_1\rangle + c = -\langle AX_1, 2X_0\rangle + \langle AX_1, X_1\rangle + c \\
&= -\langle X_1, 2AX_0\rangle + \langle AX_1, X_1\rangle + c = \langle X_1, B\rangle + \langle AX_1, X_1\rangle + c = 0.
\end{aligned}$$

Luego $\langle AX_2, X_2\rangle + \langle B, X_2\rangle + c = 0$ y por lo tanto X_2 está en $\mathcal{S}$. $\square$

A continuación probaremos que si una cuádrica $\mathcal{S}$ tiene un centro X_0, entonces trasladando el origen de coordenadas a X_0 obtenemos que la ecuación que define a $\mathcal{S}$ expresada en las nuevas coordendas carece de su parte lineal.

Proposición 7.8.8. *Sea $\mathcal{S}$ una cuádrica que admite un centro X_0. Consideremos el cambio de variables $X = \tilde{X} + X_0$ (traslación de ejes). Entonces la ecuación que define a $\mathcal{S}$ en las nuevas coordenadas es*

$$\langle A\tilde{X}, \tilde{X}\rangle + d = 0, \tag{7.10}$$

siendo $d = \langle AX_0, X_0\rangle + \langle B, X_0\rangle + c$.

Dem. Operamos,

$$\begin{aligned}
\langle AX, X\rangle + \langle B, X\rangle + c &= \langle A(\tilde{X} + X_0), \tilde{X} + X_0\rangle + \langle B, \tilde{X} + X_0\rangle + c \\
&= \langle A\tilde{X} + AX_0, \tilde{X} + X_0\rangle + \langle B, \tilde{X} + X_0\rangle + c \\
&= \langle A\tilde{X}, \tilde{X}\rangle + \langle A\tilde{X}, X_0\rangle + \langle AX_0, \tilde{X}\rangle + \langle AX_0, X_0\rangle + \langle B, \tilde{X}\rangle + \langle B, X_0\rangle + c \\
&= \langle A\tilde{X}, \tilde{X}\rangle + \langle \tilde{X}, AX_0\rangle + \langle \tilde{X}, B\rangle + \langle \tilde{X}, AX_0\rangle + \langle AX_0, X_0\rangle + \langle B, X_0\rangle + c \\
&= \langle A\tilde{X}, \tilde{X}\rangle + \langle \tilde{X}, 2AX_0 + B\rangle + \langle AX_0, X_0\rangle + \langle B, X_0\rangle + c \\
&= \langle A\tilde{X}, \tilde{X}\rangle + d, \ \text{siendo } d = \langle AX_0, X_0\rangle + \langle B, X_0\rangle + c.
\end{aligned}$$

Luego a $\langle AX, X\rangle + \langle B, X\rangle + c = 0$ le corresponde $\langle A\tilde{X}, \tilde{X}\rangle + d = 0$. $\square$

Observación 7.8.9. En la fórmula (7.10) vale $d = 0$ si y solo si el centro X_0 está en $\mathcal{S}$.

En lo que sigue aplicaremos lo visto anteriormente para determinar cómo son las cónicas y las cuádricas.

Cónicas Una *cónica* es un subconjunto de $\mathbb{R}^2$ de la forma

$$\mathcal{C} = \left\{(x,y) \in \mathbb{R}^2 : \; ax^2 + by^2 + 2cxy + dx + ey + f = 0\right\},$$

siendo $a, \dots, f$ números reales tales que a, b, c no son todos nulos. En la notación anterior es

$$\mathcal{C} = \{X \in \mathbb{R}^2 : \; \langle AX, X\rangle + \langle B, X\rangle + f = 0\},$$

siendo $X = \binom{x}{y}$, $A = \left(\begin{smallmatrix} a & b \\ b & c \end{smallmatrix}\right)$ y $B = \binom{d}{e}$.

Teorema 7.8.10. *Si una cónica $\mathcal{C}$ no es el conjunto vacío ni está formada por un solo punto, entonces es una elipse, hipérbola, parábola, dos rectas o una recta doble.*

Dem. Mantenemos para $\mathcal{C}$ las notaciones de arriba. La matriz A es simétrica real, luego existe una matriz ortogonal Q tal que $Q^tAQ = D$, siendo $D = \left(\begin{smallmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{smallmatrix}\right)$ una matriz diagonal cuyas entradas diagonales son los valores propios de A.

Supongamos $\det A \neq 0$. Sea $(x_0, y_0) \in \mathbb{R}^2$ el (único) centro de $\mathcal{C}$. Si realizamos el cambio de variable $(x, y) = (\tilde{x}, \tilde{y}) + (x_0, y_0)$, entonces la proposición 7.8.8 nos dice que la ecuación de $\mathcal{C}$ queda en

$$\mathcal{C} : \; \langle A\tilde{X}, \tilde{X}\rangle + g = 0.$$

siendo $g = ax_0^2 + by_0^2 + 2cx_0y_0 + dx_0 + ey_0 + f$. Explícitamente esa ecuación se escribe $a\tilde{x}^2 + b\tilde{y}^2 + 2c\tilde{x}\tilde{y} + g = 0$. Ahora realizamos el cambio de variable $\tilde{X} = Q\hat{X}$. Como Q es una matriz ortogonal, entonces este cambio de variables corresponde a una rotación de ejes si $\det Q = 1$ o a realizar una simetría axial a los ejes si $\det Q = -1$. Operando obtenemos $\langle A\tilde{X}, \tilde{X}\rangle = \langle AQ\hat{X}, Q\hat{X}\rangle = \langle Q^tAQ\hat{X}, \hat{X}\rangle = \langle D\hat{X}, \hat{X}\rangle$. Luego la ecuación de $\mathcal{C}$ queda

$$\mathcal{C} : \; \langle D\hat{X}, \hat{X}\rangle + g = 0.$$

Siendo $D = \left(\begin{smallmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{smallmatrix}\right)$, si escribimos $\tilde{X} = (\tilde{x}, \tilde{y})$, entonces la ecuación de $\mathcal{C}$ queda

$$\mathcal{C} : \; \lambda_1\hat{x}^2 + \lambda_2\hat{y}^2 + g = 0.$$

Consideremos primero el caso $g = 0$ ($\mathcal{C}$ contiene su centro). La ecuación de $\mathcal{C}$ es $\lambda_1\hat{x}^2 + \lambda_2\hat{y}^2 = 0$. Si $\operatorname{sg} \lambda_1 = \operatorname{sg} \lambda_2$, entonces $\mathcal{C}$ está formada solo por un punto (que es su centro). Si $\operatorname{sg} \lambda_1 \neq \operatorname{sg} \lambda_2$, entonces podemos escribir la ecuación de $\mathcal{C}$ de la forma

$$h^2\hat{x} - k^2\hat{y}^2 = 0 \quad \Leftrightarrow \quad (h\hat{x} - k\hat{y})(h\hat{x} + k\hat{y}) = 0$$

Luego $\mathcal{C}$ es la unión de dos rectas que se cortan en el centro de $\mathcal{C}$.

Consideremos ahora el caso $g \neq 0$. La ecuación de $\mathcal{C}$ se puede escribir de la forma

$$\left(\frac{-\lambda_1}{g}\right)\hat{x}^2 + \left(\frac{-\lambda_2}{g}\right)\hat{y}^2 = 1.$$

Si $\operatorname{sg} \lambda_1 = \operatorname{sg} \lambda_2 = \operatorname{sg} g$, entonces $\mathcal{C}$ es el conjunto vacío. Si $\operatorname{sg} \lambda_1 = \operatorname{sg} \lambda_2 \neq \operatorname{sg} g$, entonces $\mathcal{C}$ es una elipse. Si $\operatorname{sg} \lambda_1 \neq \operatorname{sg} \lambda_2$, entonces $\mathcal{C}$ es una hipérbola.

Supongamos ahora $\det A = 0$. Luego $\lambda_1\lambda_2 = 0$ y al ser $A \neq 0$, deducimos que uno y solo uno de λ_1, λ_2 es nulo. Supongamos $\lambda_1 \neq 0$ y $\lambda_2 = 0$ (el otro caso es análogo intercambiando las variables).

Razonando como antes, si escribimos

$$\mathcal{C}:\ \langle AX, X\rangle + \langle B, X\rangle + f = 0,$$

y hacemos el cambio de variables $X = Q\hat{X}$, entonces

$$\begin{aligned}\langle AX, X\rangle + \langle B, X\rangle + f &= \langle AQ\hat{X}, Q\hat{X}\rangle + \langle B, Q\hat{X}\rangle + f = \langle Q^tAQ\hat{X}, \hat{X}\rangle + \langle Q^tB, \hat{X}\rangle + f \\ &= \langle D\hat{X}, \hat{X}\rangle + \langle \beta, \hat{X}\rangle + f, \text{ siendo } \beta = Q^tB.\end{aligned}$$

Como es $D = \left(\begin{smallmatrix}\lambda_1 & 0\\ 0 & 0\end{smallmatrix}\right)$, entonces probamos que realizando el cambio de variables $\binom{x}{y} = Q\binom{\hat{x}}{\hat{y}}$ obtenemos

$$\mathcal{C}:\ \lambda_1\hat{x}^2 + \beta_1\hat{x} + \beta_2\hat{y} + f = 0,$$

siendo $\binom{\beta_1}{\beta_2} = Q^t\binom{b_1}{b_2}$. Si $\beta_2 \neq 0$, entonces $\mathcal{C}$ es una parábola. Si $\beta_2 = 0$, entonces la ecuación de $\mathcal{C}$ queda

$$\mathcal{C}:\ \lambda_1\hat{x}^2 + \beta_1\hat{x} + f = 0,$$

Consideremos el discriminante $\Delta = \beta_1^2 - 4\lambda_1 f$. Si $\Delta > 0$, entonces $\mathcal{C}$ son dos rectas paralelas de ecuaciones $\hat{x} = \frac{-\beta_1 \pm \sqrt{\Delta}}{2\lambda_1}$. Si $\Delta = 0$, entonces $\mathcal{C}$ es una recta (doble) de ecuación $\hat{x} = -\frac{\beta_1}{2}$ y si $\Delta < 0$, entonces $\mathcal{C}$ es el conjunto vacío. Esto concluye la clasificación de las cónicas en el plano. □

Observación 7.8.11. En el caso no degenerado en el cual hay un solo centro, la cónica es una elipse, hipérbola o dos rectas que se cortan. Mientras que los casos en que hay degeneramiento son los de la parábola, que no tiene centros, y los de dos rectas paralelas o una recta doble, que tienen infinitos centros.

Cuádricas Una *cuádrica* o *superficie cuádrica* en $\mathbb{R}^3$ es un conjunto de la forma

$$\mathcal{S} = \left\{(x,y,z) \in \mathbb{R}^3:\ a_{11}\,x^2 + a_{22}\,y^2 + a_{33}\,z^2 + 2\,a_{12}\,xy + 2\,a_{13}\,xz + 2\,a_{23}\,yz + b_1\,x + b_2\,y + b_3\,z + c = 0\right\},$$

siendo $A = \begin{pmatrix} a_{11} & a_{12} & a_{13}\\ a_{12} & a_{22} & a_{23}\\ a_{13} & a_{23} & a_{33}\end{pmatrix} \neq 0$ y $b_1, b_2, b_3, c \in \mathbb{R}$. Como en el caso general, escribimos

$$\mathcal{S} = \left\{X \in \mathbb{R}^3:\ \langle AX, X\rangle + \langle B, X\rangle + c = 0\right\},$$

siendo $B = (b_1, b_2, b_3)$ y $X = (x, y, z)$.

Ejemplo 7.8.12. Los siguientes son ejemplos de superficies cuádricas.

Cuádricas con un solo centro.

1. *Hiperboloide de una hoja*: $\frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2} = 1$.
2. *Hiperboloide de dos hojas*: $\frac{x^2}{a^2} - \frac{y^2}{b^2} - \frac{z^2}{c^2} = 1$.
3. *Elipsoide:* $\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1$.
4. *Cono:* $\frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2} = 0$.

Cuádricas sin centro.

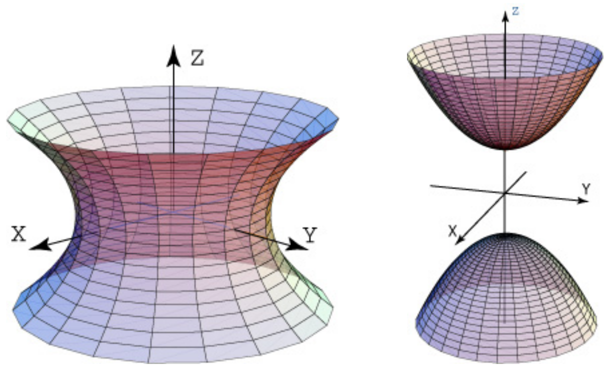

Figura 7.1: Hiperboloide de una hoja y de dos hojas.

1. *Paraboloide elíptico:* $z = \frac{x^2}{a^2} + \frac{y^2}{b^2}$.

2. *Paraboloide hiperbólico:* $z = \frac{x^2}{a^2} - \frac{y^2}{b^2}$.

3. *Cilindro parabólico:* $z = \frac{x^2}{a^2}$.

Cuádricas con un eje de centros.

1. *Cilindro elíptico:* $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$.

2. *Cilindro hiperbólico:* $\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1$.

3. *Dos planos secantes:* $\frac{x^2}{a^2} - \frac{y^2}{b^2} = 0 \quad (\frac{x}{a} - \frac{y}{b} = 0$ y $\frac{x}{a} + \frac{y}{b} = 0)$.

4. *Una recta (doble):* $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 0 \quad (x = 0$ e $y = 0)$.

Cuádricas con un plano de centros.

1. *Dos planos paralelos:* $\frac{x^2}{a^2} = 1 \quad (x = a$ y $x = -a)$.

2. *Un plano (doble):* $\frac{x^2}{a^2} = 0 \quad (x = 0)$.

Las ecuaciones anteriores se llaman las ecuaciones *reducidas* de las cuádricas.

Las figuras siguientes fueron tomadas de la revista digital costarricense *Matemática, Educación e Internet*, dependiente del Tecnológico de Costa Rica (TEC).

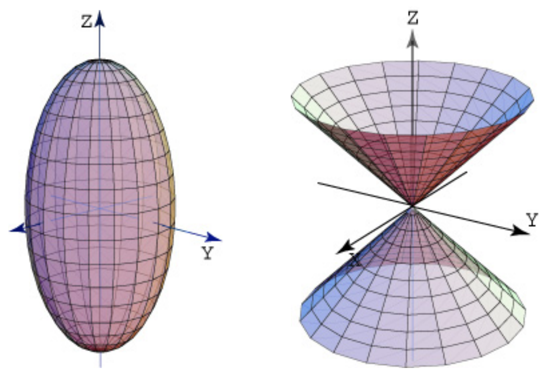

Figura 7.2: Elipsoide y cono.

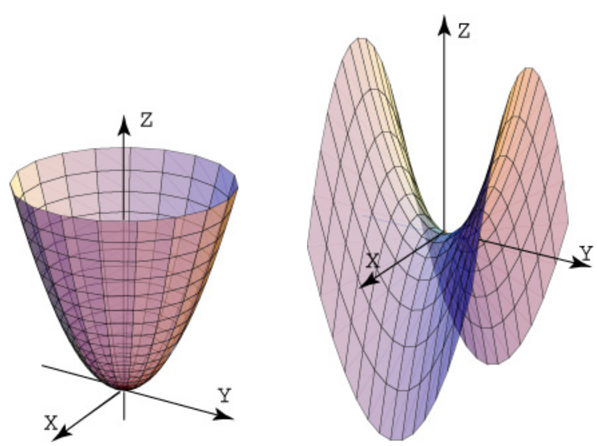

Figura 7.3: Paraboloide elíptico e hiperbólico.

Teorema 7.8.13. *Si una cuádrica $\mathcal{S}$ no es el conjunto vacío ni está formada por un solo punto, entonces coincide con una de las descritas en el ejemplo 7.8.12, a menos de trasladar, rotar o intercambiar los ejes de coordenadas o simetrizar respecto a un plano coordenado.*

Dem. Supongamos que la ecuación de $\mathcal{S}$ es

$$\mathcal{S}: \ \langle AX, X\rangle + \langle B, X\rangle + c = 0,$$

siendo $0 \neq A \in M_3(\mathbb{R})$ simétrica, $B \in \mathbb{R}^3$ y $c \in \mathbb{R}$. Como A es simétrica, entonces existe $Q \in M_3(\mathbb{R})$ matriz ortogonal tal que $Q^t A Q = D$, siendo

$$D = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix} \in M_3(\mathbb{R}).$$

Observar que λ_1, λ_2 y λ_3 son los valores propios de A y $\det(A) = \lambda_1\,\lambda_2\,\lambda_3$.

Empezamos el estudio discutiendo según $\det(A)$ es o no cero.

Caso 1. $\det(A) \neq 0$ ($\mathcal{S}$ no degenerada). La matriz A es invertible, luego $\mathcal{S}$ tiene un único centro $X_0 \in \mathbb{R}^3$. Realizamos el cambio de variables $X = \tilde{X} + X_0$ (traslación de ejes), luego la proposición 7.8.8 implica que la ecuación de $\mathcal{S}$ respecto a las coordenadas $\tilde{X}$ es

$$\mathcal{S}: \ \langle A\tilde{X}, \tilde{X}\rangle + d = 0.$$

siendo $d = \langle AX_0, X_0\rangle + \langle B, X_0\rangle + c$. Realizamos el cambio de variables $\tilde{X} = Q\,\hat{X}$. Como Q es una matriz ortogonal, esto corresponde a realizar un giro de ejes si $\det(Q) = 1$ o un giro de ejes seguido de una simetría especular si $\det(Q) = -1$ (ver la proposición 7.7.20). Si $\hat{X} = (\hat{x}, \hat{y}, \hat{z})$, entonces

$$\langle A\tilde{X}, \tilde{X}\rangle = \langle AQ\hat{X}, Q\hat{X}\rangle = \langle Q^t AQ\hat{X}, \hat{X}\rangle = \langle D\hat{X}, \hat{X}\rangle = \lambda_1\hat{x}^2 + \lambda_2\hat{y}^2 + \lambda_3\hat{z}^2.$$

Entonces la ecuación de $\mathcal{S}$ respecto a las coordenadas $\hat{x}$, $\hat{y}$ y $\hat{z}$ es

$$\mathcal{S}: \ \lambda_1\hat{x}^2 + \lambda_2\hat{y}^2 + \lambda_3\hat{z}^2 + d = 0. \tag{7.11}$$

Consideremos primero el caso en el cual $d = 0$ ($\mathcal{S}$ contiene su centro).

1. Si λ_1, λ_2 y λ_3 tienen el mismo signo, entonces $\mathcal{S}$ es un punto ($\mathcal{S} = \{X_0\}$).
2. Si λ_1, λ_2 y λ_3 no tienen el mismo signo, entonces $\mathcal{S}$ es un cono.

Consideremos ahora el caso en que $d \neq 0$. Observar que la fórmula (7.11) se puede escribir

$$\mathcal{S}: \ \frac{-\lambda_1}{d}\hat{x}^2 + \frac{-\lambda_2}{d}\hat{y}^2 + \frac{-\lambda_3}{d}\hat{z}^2 = 1.$$

Luego a menos de intercambiar $\hat{x}$, $\hat{y}$ y $\hat{z}$ tenemos los casos siguientes:

1. Si λ_1, λ_2, λ_3 y d tienen el mismo signo, entonces $\mathcal{S}$ es el conjunto vacío.
2. Si λ_1, λ_2, λ_3 tienen el mismo signo y este signo es distinto del signo de d, entonces $\mathcal{S}$ es un elipsoide.
3. Si $\operatorname{sg}\lambda_1 = \operatorname{sg}\lambda_2 \neq \operatorname{sg}\lambda_3 = \operatorname{sg} d$, ($\operatorname{sg}$ es el signo) entonces $\mathcal{S}$ es un hiperboloide de una hoja.
4. Si $\operatorname{sg}\lambda_1 = \operatorname{sg}\lambda_2 \neq \operatorname{sg}\lambda_3 \neq \operatorname{sg} d$, entonces $\mathcal{S}$ es un hiperboloide de dos hojas.

Caso 2. $\det(A) = 0$**.** Realizamos primero el cambio de variables $X = Q\tilde{X}$. Entonces

$$\begin{aligned}\langle AX, X\rangle + \langle B, X\rangle + c &= \langle AQ\tilde{X}, Q\tilde{X}\rangle + \langle B, Q\tilde{X}\rangle + c = \langle Q^t AQ\tilde{X}, \tilde{X}\rangle + \langle Q^t B, \tilde{X}\rangle + c \\ &= \langle D\tilde{X}, \tilde{X}\rangle + \langle Q^t B, \tilde{X}\rangle + c.\end{aligned}$$

Sean $\tilde{X} = \begin{pmatrix} \tilde{x} \\ \tilde{y} \\ \tilde{z} \end{pmatrix}$ y $\begin{pmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \end{pmatrix} = Q^t \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}$. Entonces la ecuación de $\mathcal{S}$ respecto a las coordenadas $\tilde{x}$, $\tilde{y}$ y $\tilde{z}$ es

$$\mathcal{S}:\ \lambda_1\tilde{x}^2 + \lambda_2\tilde{y}^2 + \lambda_3\tilde{z}^2 + \beta_1\tilde{x} + \beta_2\tilde{y} + \beta_3\tilde{z} + c = 0.$$

Observar que $A \neq 0$ implica que no puede ser $\lambda_1 = \lambda_2 = \lambda_3 = 0$; luego a menos de intercambiar los ejes tenemos dos posibilidades: $\lambda_1 \neq 0$, $\lambda_2 = \lambda_3 = 0$ y $\lambda_1 \neq 0$, $\lambda_2 \neq 0$, $\lambda_3 = 0$.

Caso 2.1. $\lambda_1 \neq 0$**,** $\lambda_2 = \lambda_3 = 0$**.** Es

$$\mathcal{S}:\ \lambda_1\tilde{x}^2 + \beta_1\tilde{x} + \beta_2\tilde{y} + \beta_3\tilde{z} + c = 0.$$

Caso 2.1.1. $\beta_2 = \beta_3 = 0$**.** Es:

$$\mathcal{S}:\ \lambda_1\tilde{x}^2 + \beta_1\tilde{x} + c = 0.$$

Sea $\Delta = \beta_1^2 - 4\lambda_1 c$. Tenemos las siguientes posibilidades:

1. Si $\Delta < 0$, entonces $\mathcal{S}$ es el conjunto vacío.
2. Si $\Delta > 0$, entonces $\mathcal{S}$ son dos planos paralelos: $\tilde{x} = \frac{-\beta_1+\sqrt{\Delta}}{2\lambda_1}$ y $\tilde{x} = \frac{-\beta_1-\sqrt{\Delta}}{2\lambda_1}$.
3. Si $\Delta = 0$, entonces $\mathcal{S}$ es un plano (doble): $\tilde{x} = \frac{-\beta_1}{2\lambda_1}$.

Caso 2.1.2. $\beta_2 \neq 0$ **o** $\beta_3 \neq 0$**.** Si realizamos el cambio de coordenadas

$$\begin{cases} \tilde{x} = \hat{x} \\ \tilde{y} = (\cos\omega)\hat{y} - (\operatorname{sen}\omega)\hat{z} \\ \tilde{z} = (\operatorname{sen}\omega)\hat{y} + (\cos\omega)\hat{z} \end{cases},$$

que es un giro de ángulo ω alrededor del eje $O\tilde{x}$, obtenemos

$$\mathcal{S}:\ \lambda_1\hat{x}^2 + \beta_1\hat{x} + (\beta_2\cos\omega + \beta_3\operatorname{sen}\omega)\hat{y} + (\beta_3\cos\omega - \beta_2\operatorname{sen}\omega)\hat{z} + c = 0. \tag{7.12}$$

Observar que la ser $(\beta_2, \beta_3) \neq (0, 0)$, siempre existe $\omega \in [0, 2\pi)$ tal que $\beta_2\cos\omega + \beta_3\operatorname{sen}\omega = 0$. En efecto esta última ecuación equivale a que el vector $(\cos\omega, \operatorname{sen}\omega)$ sea ortogonal al vector (β_2, β_3), y para obtener ω alcanza con considerar el vector $(\beta_3, -\beta_2)$ (que claramente es ortogonal a (β_2, β_3)) y normalizarlo; notar que esto último equivale a intersectar la recta que pasa por el origen y es paralela a $(\beta_3, -\beta_2)$ con la circunferencia de centro en el origen y radio 1. Así es $(\beta_3, -\beta_2) = \mu(\cos\omega, \operatorname{sen}\omega)$, siendo $\mu = \sqrt{\beta_2^2 + \beta_3^2}$.

Eligiendo ω de esa manera, esta última relación junto con $(\beta_2, \beta_3) \neq (0, 0)$ implican que $(\cos\omega, \operatorname{sen}\omega)$ no puede ser ortogonal al vector $(\beta_3, -\beta_2)$, es decir que el coeficiente de $\hat{y}$ en (7.12) es no nulo. Luego la ecuación de $\mathcal{S}$ respecto a las coordenadas $\hat{x}$, $\hat{y}$ y $\hat{z}$ queda

$$\mathcal{S}:\ \lambda_1\hat{x}^2 + \beta_1\hat{x} + \gamma\hat{y} + c = 0, \quad \text{con } \gamma = \beta_2 a + \beta_3 b \neq 0.$$

Despejando $\hat{y}$ en la ecuación anterior obtenemos

$$\mathcal{S}:\ \hat{y} = \frac{-\lambda_1}{\gamma}\hat{x}^2 + \frac{-\beta_1}{\gamma}\hat{x} + \frac{-c}{\gamma},$$

luego $\mathcal{S}$ es un cilindro parabólico.

Caso 2.2. $\lambda_1 \neq 0$**,** $\lambda_2 \neq 0$ **y** $\lambda_3 = 0$**.** Es $\mathcal{S}: \ \lambda_1 \tilde{x}^2 + \lambda_2 \tilde{y}^2 + \beta_1 \tilde{x} + \beta_2 \tilde{y} + \beta_3 \tilde{z} + c = 0$.

Si hacemos la traslación de ejes $\begin{cases} \tilde{x} = \hat{x} - \frac{\beta_1}{2\lambda_1} \\ \tilde{y} = \hat{y} - \frac{\beta_2}{2\lambda_2} \\ \tilde{z} = \hat{z} \end{cases}$, entonces la ecuación de $\mathcal{S}$ respecto a las coordenadas $\hat{x}$, $\hat{y}$ y $\hat{z}$ queda

$$\mathcal{S}: \ \lambda_1 \hat{x}^2 + \lambda_2 \hat{y}^2 + \beta_3 \hat{z} + \eta = 0, \tag{7.13}$$

siendo $\eta = c - \frac{\beta_1^2}{4\lambda_1} - \frac{\beta_2^2}{4\lambda_2}$.

Caso 2.2.1. $\beta_3 = 0$**.** Es $\mathcal{S}: \ \lambda_1 \hat{x}^2 + \lambda_2 \hat{y}^2 + \eta = 0$

1. Si $\eta = 0$, entonces tenemos dos posibilidades:
   a) Si sg $\lambda_1 =$ sg λ_2, entonces $\mathcal{S}$ es la recta (doble) $\hat{x} = 0$ y $\hat{y} = 0$.
   b) Si sg $\lambda_1 \neq$ sg λ_2, entonces $\mathcal{S}$ son dos planos secantes $\hat{y} = \sqrt{\frac{-\lambda_1}{\lambda_2}}\hat{x}$ e $\hat{y} = -\sqrt{\frac{-\lambda_1}{\lambda_2}}\hat{x}$.
2. Si $\eta \neq 0$ entonces tenemos tres posibilidades:
   a) Si sg $\lambda_1 =$ sg $\lambda_2 =$ sg η, entonces $\mathcal{S}$ es el conjunto vacío.
   b) Si sg $\lambda_1 =$ sg $\lambda_2 \neq$ sg η, entonces $\mathcal{S}$ es un cilindro elíptico.
   c) Si sg $\lambda_1 \neq$ sg λ_2, entonces $\mathcal{S}$ es un cilindro hiperbólico.

Caso 2.2.2. $\beta_3 \neq 0$**.** Es $\mathcal{S}: \ \lambda_1 \hat{x}^2 + \lambda_2 \hat{y}^2 + \beta_3 \hat{z} + \eta = 0$.

Haciendo la traslación de ejes $\begin{cases} \hat{x} = \overline{x} \\ \hat{y} = \overline{y} \\ \hat{z} = \overline{z} - \eta/\beta_3 \end{cases}$, obtenemos $\mathcal{S}: \ \lambda_1 \overline{x}^2 + \lambda_2 \overline{y}^2 + \beta_3 \overline{z} = 0$. Luego

1. Si sg $\lambda_1 =$ sg λ_2, entonces $\mathcal{S}$ es un paraboloide elíptico.
2. Si sg $\lambda_1 \neq$ sg λ_2, entonces $\mathcal{S}$ es un paraboloide hiperbólico.

Esto concluye la clasificación de las superficies cuádricas. $\square$

Resumen.

Sea $\mathcal{S}:\ \Phi(X)+\alpha(X)+c$, siendo $\Phi(X)=X^tAX,\ \alpha(X)=B^tX$. Sean λ_1, λ_2 y λ_3 los valores propios de A. Sean D matriz diagonal y Q matriz ortogonal tales que $D=Q^tAQ$.

- $\det(A)\neq 0$. Sean X_0 la única solución de $2AX+B=0$ y $d=\Phi(X_0)+\alpha(X_0)+c$.

  Cambio de variables $X=Q\hat{X}+X_0\ \Rightarrow\ \mathcal{S}:\ \lambda_1\hat{x}^2+\lambda_2\hat{y}^2+\lambda_3\hat{z}^2+d=0$.

  - $d=0$ ($\mathcal{S}$ contiene a su centro X_0).
    - λ_1, λ_2 y λ_3 tienen el mismo signo, $\mathcal{S}$ es un punto.
    - λ_1, λ_2 y λ_3 tienen distinto signo, $\mathcal{S}$ es un cono.
  - $d\neq 0$.
    - λ_1, λ_2 y λ_3 tienen el mismo signo y coincide con el de d, $\mathcal{S}$ es el conjunto vacío.
    - λ_1, λ_2 y λ_3 tienen el mismo signo y es distinto del signo de d, $\mathcal{S}$ es un elipsoide.
    - $\operatorname{sg}\lambda_1=\operatorname{sg}\lambda_2\neq\operatorname{sg}\lambda_3=\operatorname{sg}d$, $\mathcal{S}$ es un hiperboloide de una hoja.
    - $\operatorname{sg}\lambda_1=\operatorname{sg}\lambda_2\neq\operatorname{sg}\lambda_3\neq\operatorname{sg}d$, $\mathcal{S}$ es un hiperboloide de dos hojas.

- $\det(A)=0$, luego $0\in\{\lambda_1,\lambda_2,\lambda_3\}$. Sea $\beta=(\beta_1,\beta_2,\beta_3):=Q^tB$.

  Cambio de variables $X=Q\tilde{X}\ \Rightarrow\ \mathcal{S}:\ \lambda_1\tilde{x}^2+\lambda_2\tilde{y}^2+\lambda_3\tilde{z}^2+\beta_1\tilde{x}+\beta_2\tilde{y}+\beta_3\tilde{z}+c=0$.

  - $\lambda_1\neq 0$, $\lambda_2=\lambda_3=0$. Luego $\mathcal{S}:\ \lambda_1\tilde{x}^2+\beta_1\tilde{x}+\beta_2\tilde{y}+\beta_3\tilde{z}+c=0$.
    - $\beta_2\neq 0$ o $\beta_3\neq 0$, $\mathcal{S}$ es un cilindro parabólico.
    - $\beta_2=\beta_3=0$. Sea $\Delta=\beta_1^2-4\lambda_1c$.
      - $\Delta<0$, $\mathcal{S}$ es el conjunto vacío.
      - $\Delta>0$, $\mathcal{S}$ son dos planos paralelos.
      - $\Delta=0$, $\mathcal{S}$ es un plano (doble).
  - $\lambda_1\neq 0$, $\lambda_2\neq 0$, $\lambda_3=0$. Luego $\mathcal{S}:\ \lambda_1\tilde{x}^2+\lambda_2\tilde{y}^2+\beta_1\tilde{x}+\beta_2\tilde{y}+\beta_3\tilde{z}+c=0$.
    - $\beta_3=0$. Sea $\eta=c-\frac{\beta_1^2}{4\lambda_1}-\frac{\beta_2^2}{4\lambda_2}$.
      - $\eta=0$.
        - $\operatorname{sg}\lambda_1=\operatorname{sg}\lambda_2$, $\mathcal{S}$ es una recta (doble).
        - $\operatorname{sg}\lambda_1\neq\operatorname{sg}\lambda_2$, $\mathcal{S}$ son dos planos secantes.
      - $\eta\neq 0$.
        - $\operatorname{sg}\lambda_1=\operatorname{sg}\lambda_2=\operatorname{sg}\eta$, $\mathcal{S}$ es el conjunto vacío.
        - $\operatorname{sg}\lambda_1=\operatorname{sg}\lambda_2\neq\operatorname{sg}\eta$, $\mathcal{S}$ es un cilindro elíptico.
        - $\operatorname{sg}\lambda_1\neq\operatorname{sg}\lambda_2$, $\mathcal{S}$ es un cilindro hiperbólico.
    - $\beta_3\neq 0$.
      - $\operatorname{sg}\lambda_1=\operatorname{sg}\lambda_2$, $\mathcal{S}$ es un paraboloide elíptico.
      - $\operatorname{sg}\lambda_1=\operatorname{sg}\lambda_2$, $\mathcal{S}$ es un paraboloide hiperbólico.

Observación 7.8.14. Si en la cuádrica $\mathcal{S}:\ X^tAX+B^tX+c=0$ se verifica $\det A=0$, entonces 0 es valor propio de A y por lo tanto se pueden hallar explícitamente todos los valores y vectores propios de A, y por lo tanto se pueden hallar explícitamente las matrices D y Q tales que $A=Q^tDQ$.

Ejemplo 7.8.15. Sea $\mathcal{S}:\ 2x^2+y^2-z^2+2xz+4x-4y+2z+3=0$. Es $A=\begin{pmatrix}2&0&1\\0&1&0\\1&0&-1\end{pmatrix}$ y $\det(A)=-3$. Calculando obtenemos $X_0=(-1,2,0)$ y $d=-3$. Como A es simétrica real, sabemos por el corolario 4.4.18 que los signos de sus valores propios coinciden con los signos de las entradas diagonales de cualquier matriz diagonal congruente con A. Luego para saber el signo de sus valores propios podemos aplicar el algoritmo de la sección 7.2. Observar que no necesitamos obtener la matriz de congruencia, así que el algoritmo es más simple y solo debemos realizar las operaciones elementales en las filas y columnas de A. En nuestro caso obtenemos

$$A=\begin{pmatrix}2&0&1\\0&1&0\\1&0&-1\end{pmatrix}\sim\begin{pmatrix}1&0&0\\0&2&1\\0&1&-1\end{pmatrix}\sim\begin{pmatrix}1&0&0\\0&-1&1\\0&1&2\end{pmatrix}\sim\begin{pmatrix}1&0&0\\0&-1&0\\0&0&3\end{pmatrix}.$$

Luego hay dos valores propios positivos y uno negativo, como el signo de d es negativo, deducimos que $\mathcal{S}$ es un hiperboloide de una hoja.

Observar que en este caso podemos hallar explícitamente los valores propios, estos son 1, $\frac{1+\sqrt{13}}{2}$ y $\frac{1-\sqrt{13}}{2}$, luego aplicando (7.11) tenemos que la ecuación reducida de $\mathcal{S}$ es

$$\frac{1}{3}\hat{x}^2+\left(\frac{1+\sqrt{13}}{6}\right)\hat{y}^2-\left(\frac{\sqrt{13}-1}{6}\right)\hat{z}^2=1.$$

Ejemplo 7.8.16. Sea $\mathcal{S}:\ x^2+4y^2+z^2-2xz-2x+2z=0$. Es $A=\begin{pmatrix}1&0&-1\\0&4&0\\-1&0&1\end{pmatrix}$ y $\det(A)=0$. Luego uno de los valores propios es cero. Calculando el polinomio característico de A obtenemos los otros valores propios. En este caso es $\chi_A(t)=-t(2-t)(4-t)$ y los valores propios son 0, 2 y 4. Operando obtenemos que $\{(1,0,1),\ (1,0.-1),\ (0,1,0)\}$ es una base de vectores propios correspondientes. Normalizando esta base obtenemos la matriz Q. Luego

$$D=\begin{pmatrix}4&0&0\\0&2&0\\0&0&0\end{pmatrix},\quad D=\begin{pmatrix}0&\frac{1}{\sqrt{2}}&\frac{1}{\sqrt{2}}\\1&0&0\\0&-\frac{1}{\sqrt{2}}&\frac{1}{\sqrt{2}}\end{pmatrix},\quad A=Q^tDQ.$$

Calculando $\beta=Q^tB$ obtenemos $\beta=\left(0,-2\sqrt{2},0\right)$, luego $\beta_3=0$. Calculando $\eta=c-\frac{\beta_1^2}{4\lambda_1}-\frac{\beta_2^2}{4\lambda_2}$ obtenemos $\eta=-1$. Luego de (7.13) deducimos que la ecuación reducida de $\mathcal{S}$ es

$$4\hat{x}^2+2\hat{y}^2=1.$$

Esto nos dice que $\mathcal{S}$ es un cilindro elíptico.

7.9. Un algoritmo para determinar la forma de Jordan

La proposición 6.2.2 nos permite calcular la forma de Jordan en casos de dimensión baja. En lo que sigue veremos cómo obtener la forma de Jordan en el caso general y también probaremos su unicidad.

Lema 7.9.1. *Si*

$$A=\begin{pmatrix}J_1&&\\&\ddots&\\&&J_k\end{pmatrix}\in M_n(\Bbbk),\quad J_i=\begin{pmatrix}0&1&&&\\&0&1&&\\&&\ddots&\ddots&\\&&&0&1\\&&&&0\end{pmatrix}\in M_{p_i}(\Bbbk),\ \forall i=1,\dots,k. \tag{7.14}$$

siendo $p = p_1 \geq p_2 \geq \cdots \geq p_k$, entonces la cantidad de bloques de Jordan de tamaño r contenidos en A es $l_r - l_{r+1}$, siendo $l_r = \operatorname{rango}\big(A^{r-1}\big) - \operatorname{rango}(A^r)$, $r = 1, 2 \ldots$.

Dem. Recordar que en la primera parte de la proposición 6.3.6 probamos que si J es un bloque de Jordan de tamaño j con ceros en su diagonal principal, entonces

$$\operatorname{rango}(J) = j - 1,\ \operatorname{rango}\big(J^2\big) = j - 2,\ \ldots,\ \operatorname{rango}\big(J^{j-1}\big) = 1,\ \operatorname{rango}\big(J^j\big) = 0.$$

En particular $\operatorname{rango}(J_i^m) = 0$, para todo $m \geq j$. Para cada $r = 1, 2, \ldots$, sea m_r la cantidad de bloques de Jordan de tamaño r contenidos en A. Observar que $m_r = 0$ si $j \notin \{p_1, \ldots, p_k\}$. Como para cada $j = 1, \ldots, p$ hay m_j bloques de tamaño j y cada bloque de tamaño j tiene rango $j - 1$, es

$$\operatorname{rango}(A) = \sum_{i=1}^{k} \operatorname{rango}(J_i) = m_2 + 2m_3 + 3m_4 + \cdots + (p-2)m_{p-1} + (p-1)m_p.$$

Notar que la suma empieza en m_2 porque los bloques de tamaño 1 son nulos. Observar que vale:

$$A^r = \begin{pmatrix} J_1^r & & \\ & \ddots & \\ & & J_k^r \end{pmatrix} \in M_n(\Bbbk), \quad \forall r = 0, 1, \ldots. \tag{7.15}$$

Considerando (7.15) con $r = 2$, cada bloque J^2 de tamaño j tiene rango $j - 2$, luego

$$\operatorname{rango}\big(A^2\big) = \sum_{i=1}^{k} \operatorname{rango}\big(J_i^2\big) = m_3 + 2m_4 + 3m_5 + \cdots + (p-3)m_{p-1} + (p-2)m_p.$$

La suma empieza en m_3 porque los bloques de tamaño 2 al elevar al cuadrado son nulos. Así seguimos obteniendo

$$\begin{aligned}
\operatorname{rango}(A) &= m_2 + 2m_3 + 3m_4 + \cdots + (p-2)m_{p-1} + (p-1)m_p,\\
\operatorname{rango}\big(A^2\big) &= m_3 + 2m_4 + 3m_5 + \cdots + (p-3)m_{p-1} + (p-2)m_p,\\
\operatorname{rango}\big(A^3\big) &= m_4 + 2m_5 + 3m_6 + \cdots + (p-4)m_{p-1} + (p-3)m_p,\\
&\vdots\\
\operatorname{rango}\big(A^{p-3}\big) &= m_{p-2} + 2m_{p-1} + 3m_p,\\
\operatorname{rango}\big(A^{p-2}\big) &= m_{p-1} + 2m_p,\\
\operatorname{rango}\big(A^{p-1}\big) &= m_p,\\
\operatorname{rango}(A^p) &= 0.
\end{aligned}$$

Notar que vale

$$m_1 + 2m_2 + 3m_3 + 4m_4 + \cdots + (p-1)m_{p-1} + pm_p = n = \operatorname{rango}(I) = \operatorname{rango}(A^0).$$

Luego $l_r = 0$ si $r > p$ y

$$\begin{aligned}
l_p &= \operatorname{rango}\big(A^{p-1}\big) - \operatorname{rango}(A^p) = m_p,\\
l_{p-1} &= \operatorname{rango}\big(A^{p-2}\big) - \operatorname{rango}\big(A^{p-1}\big) = m_p + m_{p-1},\\
l_{p-2} &= \operatorname{rango}\big(A^{p-3}\big) - \operatorname{rango}\big(A^{p-2}\big) = m_p + m_{p-1} + m_{p-2},\\
&\vdots\\
l_3 &= \operatorname{rango}\big(A^{2}\big) - \operatorname{rango}\big(A^{3}\big) = m_p + m_{p-1} + m_{p-2} + \cdots + m_3,\\
l_2 &= \operatorname{rango}\big(A^{1}\big) - \operatorname{rango}\big(A^{2}\big) = m_p + m_{p-1} + m_{p-2} + \cdots + m_3 + m_2,\\
l_1 &= \operatorname{rango}\big(A^{0}\big) - \operatorname{rango}\big(A^{1}\big) = m_p + m_{p-1} + m_{p-2} + \cdots + m_3 + m_2 + m_1.
\end{aligned}$$

Finalmente, restando miembro a miembro las ecuaciones anteriores obtenemos $m_r = l_r - l_{r+1}$, para todo $r = 1, 2, \ldots$. $\square$

El siguiente teorema da un algoritmo para hallar la forma de Jordan.

Teorema 7.9.2. *Sea $T \in \mathcal{L}(V)$ tal que su polinomio característico escinde y $\mathcal{B}$ una base de Jordan para T. Supongamos que*

$$[T]_{\mathcal{B}} = \begin{pmatrix} A_1 & & \\ & \ddots & \\ & & A_h \end{pmatrix}, \tag{7.16}$$

y cada bloque A_i es de la forma

$$A_i = \begin{pmatrix} J_1 & & \\ & \ddots & \\ & & J_k \end{pmatrix}, \quad J_l = \begin{pmatrix} \lambda_i & 1 & & & \\ & \lambda_i & 1 & & \\ & & \ddots & \ddots & \\ & & & \lambda_i & 1 \\ & & & & \lambda_i \end{pmatrix}, \ l = 1, \ldots, k, \tag{7.17}$$

siendo $\lambda_1, \ldots, \lambda_h$ los valores propios distintos de T. Entonces para cada $i = 1, \ldots, h$ vale

- *El tamaño del bloque A_i es la multiplicidad algebraica de λ_i.*
- *La cantidad de bloques de Jordan contenidos en A_i es la multiplicidad geométrica de λ_i.*
- *La cantidad de bloques de Jordan de tamaño r contenidos en A_i es $l_r - l_{r+1}$, siendo $l_r = \operatorname{rango}(T - \lambda_i \mathrm{Id})^{r-1} - \operatorname{rango}(T - \lambda_i \mathrm{Id})^r$, $r = 1, 2 \ldots$.*

Dem. Sea $n = \dim V$. Observar que las dos primeras afirmaciones ya las probamos en la proposición 6.2.2, luego solo resta probar la tercera. Recordar que la cantidad de bloques de Jordan de tamaño r contenidos en A_i coincide con la de $A_i - \lambda_i I$. Por el lema anterior, si llamamos m_r a la cantidad de bloques de Jordan de tamaño r contenidos en $A_i - \lambda_i I$, es $m_r = l_r - l_{r+1}$, siendo
$l_r = \operatorname{rango}(A_i - \lambda_i \mathrm{Id})^{r-1} - \operatorname{rango}(A_i - \lambda_i \mathrm{Id})^r$.

De (7.16) deducimos

$$[(T - \lambda_i I)^r]_{\mathcal{B}} = \begin{pmatrix} (A_1 - \lambda_i I)^r & & \\ & \ddots & \\ & & (A_h - \lambda_i I)^r \end{pmatrix}, \quad \forall r = 1, 2, \ldots.$$

Luego $\operatorname{rango}((T-\lambda_i I)^r) = \sum_{j=1}^h \operatorname{rango}((A_j-\lambda_i I)^r)$, para todo $r=1,2,\ldots$. Observar que si $j\neq i$, entonces $A_j-\lambda_i I \in M_{n_j}(\Bbbk)$ es invertible y por lo tanto $(A_j-\lambda_i I)^r \in M_{n_j}(\Bbbk)$ es invertible para todo r, luego $\operatorname{rango}((A_j-\lambda_i I)^r)=n_j$, para todo r y todo $j\neq i$. Esto implica

$$\operatorname{rango}((T-\lambda_i I)^r) = \sum_{j=1}^h \operatorname{rango}((A_j-\lambda_i I)^r) = \sum_{j\neq i} n_j + \operatorname{rango}((A_i-\lambda_i I)^r)\,, \quad \forall r=1,2,\ldots.$$

Operando obtenemos

$$\begin{aligned}
&\operatorname{rango}\big((T-\lambda_i I)^{r-1}\big) - \operatorname{rango}((T-\lambda_i I)^r) = \\
&= \left(\sum_{j\neq i} n_j + \operatorname{rango}\big((A_i-\lambda_i I)^{r-1}\big)\right) - \left(\sum_{j\neq i} n_j + \operatorname{rango}((A_i-\lambda_i I)^r)\right) \\
&= \operatorname{rango}\big((A_i-\lambda_i I)^{r-1}\big) - \operatorname{rango}((A_i-\lambda_i I)^r) = l_r.
\end{aligned}$$

Esto prueba la tercer afirmación del teorema. $\square$

Observación 7.9.3. El teorema anterior prueba que la forma de Jordan de T no depende de la elección de la base de Jordan. Es decir, si ordenamos los bloques de Jordan correspondientes al mismo valor propio en tamaños decrecientes, entonces la forma de Jordan de T es única a menos de reordenar $A_1,\ldots,A_h$.

Ejemplo 7.9.4. Sea V un espacio vectorial de dimensión 8 y $T\in\mathcal{L}(V)$ del cual se sabe

$$\chi_T(t)=(t-2)^8, \quad \operatorname{rango}(T-2\mathrm{Id})=4, \quad \operatorname{rango}(T-2\mathrm{Id})^2=2, \quad \operatorname{rango}(T-2\mathrm{Id})^3=1.$$

Veamos que con esos datos podemos determinar su forma de Jordan J. Por la forma del polinomio característico sabemos que J está formada solo por bloques de Jordan correspondientes al valor propio 2. La cantidad de bloques de J es $\mathrm{MG}(2)=8-4=4$. Para saber sus tamaños aplicamos el teorema 7.9.2:

$$l_1=\mathrm{MG}(2)=4, \quad l_2=\operatorname{rango}(T-2\mathrm{Id})-\operatorname{rango}(T-2\mathrm{Id})^2=2, \quad l_3=\operatorname{rango}(T-2\mathrm{Id})^2-\operatorname{rango}(T-2\mathrm{Id})^3=1.$$

Luego la cantidad de bloques de tamaño 1 es $l_1-l_2=2$ y la cantidad de bloques de tamaño 2 es $l_2-l_3=1$. Sabemos que J tiene 4 bloques y que de los cuales hay dos bloques de tamaño 1 y un bloque de tamaño 2, luego el que resta solo puede ser un bloque de Jordan de tamaño 4 y por lo tanto

$$J=\begin{pmatrix}
2&1&0&0&0&0&0&0\\
0&2&1&0&0&0&0&0\\
0&0&2&1&0&0&0&0\\
0&0&0&2&0&0&0&0\\
0&0&0&0&2&1&0&0\\
0&0&0&0&0&2&0&0\\
0&0&0&0&0&0&2&0\\
0&0&0&0&0&0&0&2
\end{pmatrix}.$$

Definición 7.9.5. Si $A\in M_n(\Bbbk)$ y $\chi_A(t)=\chi_{L_A}(t)\in\Bbbk[t]$ escinde, definimos la *forma de Jordan* de A como la de $L_A\in\mathcal{L}(\Bbbk^n)$.

Observar que si J es la forma de Jordan de A y $\mathcal{B}=\{v_1,\ldots,v_n\}$ es la base de Jordan para L_A correspondiente, entonces es

$$A=QJQ^{-1}, \quad Q=[v_1|\cdots|v_n].$$

Ejemplo 7.9.6. Sea
$$A = \begin{pmatrix} 2 & -1 & 0 & 1 \\ 0 & 3 & -1 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & -1 & 0 & 3 \end{pmatrix} \in M_4(\mathbb{R}).$$
Es $\chi_A(t) = (t-2)^3(t-3)$, luego
$$J = \begin{pmatrix} A_1 & \\ & 3 \end{pmatrix}, \quad A_1 \in M_3(\mathbb{R}).$$
Para el valor propio 2, es $\mathrm{MG}(2) = 4 - \mathrm{rango}(A - 2I) = 2$. Luego A_1 tiene dos bloques y como $A_1 \in M_3(\mathbb{R})$ esto determina A_1. Entonces la forma de Jordan de A es
$$J = \begin{pmatrix} 2 & 1 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \end{pmatrix}.$$
Para hallar la matriz de semejanza Q tenemos que hallar la base de Jordan correspondiente. Observar que esta base es de la forma
$$\mathcal{B} = \{(A-2I)\,v, v, u, w\},$$
siendo $\{(A-2I)\,v, u\}$ una base de $\operatorname{Ker}(L_A - 2\,\mathrm{id}\,)$, $\operatorname{Ker}(L_A - 2\,\mathrm{id}\,)^2 = \operatorname{Ker}(L_A - 2\,\mathrm{id}\,) \oplus [v]$ y $\{w\}$ una base de $\operatorname{Ker}(L_A - 3\,\mathrm{id}\,)$.

Operando obtenemos
$$A - 2I = \begin{pmatrix} 0 & -1 & 0 & 1 \\ 0 & 1 & -1 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & -1 & 0 & 1 \end{pmatrix}, \quad (A-2I)^2 = \begin{pmatrix} 0 & -2 & 1 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & -2 & 1 & 1 \end{pmatrix}, \quad A - 3I = \begin{pmatrix} -1 & -1 & 0 & 1 \\ 0 & 0 & -1 & 0 \\ 0 & 1 & -2 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix}.$$
Luego
$$\begin{aligned} \operatorname{Ker}(L_A - 2\,\mathrm{id}\,) &= \{(x,y,z,t) : \ y = z = t\} = [(1,0,0,0),(0,1,1,1)], \\ \operatorname{Ker}(L_A - 2\,\mathrm{id}\,)^2 &= \{(x,y,z,t) : \ -2y + z + t = 0\} = [(1,0,0,0),(0,1,0,2),(0,0,1,-1)], \\ \operatorname{Ker}(L_A - 3\,\mathrm{id}\,) &= \{(x,y,z,t) : \ x = t \text{ y } y = z = 0\} = [(1,0,0,1)]. \end{aligned}$$
Observar que $v = (0,1,0,2) \in \operatorname{Ker}(L_A - 2\,\mathrm{id}\,)^2 \setminus \operatorname{Ker}(L_A - 2\,\mathrm{id}\,)$, luego
$$\operatorname{Ker}(L_A - 2\,\mathrm{id}\,)^2 = \operatorname{Ker}(L_A - 2\,\mathrm{id}\,)^2 \oplus [v].$$
Entonces sabemos que $(L_A - 2\,\mathrm{id}\,)(v) = (1,1,1,1) \in \operatorname{Ker}(L_A - 2\,\mathrm{id}\,)$. Observar que si $u = (1,0,0,0)$, el conjunto
$$\{(L_A - 2\,\mathrm{id}\,)(v), u\} = \{(1,1,1,1),(1,0,0,0)\} \subset \operatorname{Ker}(L_A - 2\,\mathrm{id}\,)$$
y es LI, luego es base de $\operatorname{Ker}(L_A - 2\,\mathrm{id}\,)$. Si tomamos $w = (1,0,0,1) \in \operatorname{Ker}(L_A - 3\,\mathrm{id}\,)$, obtenemos que el conjunto
$$\mathcal{B} = \{(1,1,1,1),(0,1,0,2),(1,0,0,0),(1,0,0,1)\}$$
es una base de Jordan y si $Q = \begin{pmatrix} 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 1 & 2 & 0 & 1 \end{pmatrix}$, es $A = QJQ^{-1}$.

Del teorema 7.9.2 se deduce inmediatamente la siguiente.

Proposición 7.9.7. *Sea $T \in \mathcal{L}(V)$, $\mathcal{B}$ una base de V y $A = [T]_{\mathcal{B}}$. Supongamos que $\chi_A(t) = \chi_T(t)$ escinde. Entonces la forma de Jordan de T coincide con la forma de Jordan de A.* $\square$

En el teorema y corolario siguientes asumiremos que en las formas de Jordan, los bloques de Jordan correspondientes a los mismos valores propios están ordenados en tamaño decreciente.

Teorema 7.9.8. *Sean A y B en $M_n(\mathbb{k})$ tales que sus polinomios característicos escinden. Entonces A y B son semejantes si y solo si tienen la misma forma de Jordan*[3].

Dem. Si A y B tienen la misma forma de Jordan J, entonces A y B son semejantes a J y por lo tanto son semejantes entre sí.

Si A y B son semejantes, entonces existen $T \in \mathcal{L}(\mathbb{k}^n)$ y $\mathcal{B}$ y $\mathcal{C}$ bases de $\mathbb{k}^n$ tales que $A = [T]_{\mathcal{B}}$ y $B = [T]_{\mathcal{C}}$. Luego la proposición 7.9.7 implica que A y B tienen la misma forma de Jordan. $\square$

Corolario 7.9.9. *Si A y B están en $M_n(\mathbb{C})$, entonces A y B son semejantes si y solo si tienen la misma forma de Jordan.* $\square$

7.10. Forma de Jordan real

A continuación desarrollaremos la forma de Jordan *real* para una matriz con coeficientes en $\mathbb{R}$. Si X es un vector de $\mathbb{C}^n$ o una matriz con coeficiente complejos, escribiremos $\overline{X}$ al vector o matriz que se obtiene conjugando todas las entradas de X.

Sea $A \in M_n(\mathbb{R})$ y supongamos que la descomposición factorial de $\chi_A(t)$ en $\mathbb{R}[t]$ es

$$\chi_A(t) = (-1)^n \prod_{i=1}^{h} (t-\lambda_i)^{n_i} \prod_{j=1}^{k} \left(t^2 + c_j t + d_j\right)^{m_j}, \quad \lambda_i, c_j, d_j \in \mathbb{R},\ c_j^2 - 4\,d_j < 0,\ \forall i, j. \tag{7.18}$$

Sea $T = L_A \in \mathcal{L}(\mathbb{R}^n)$, es decir $T(v) = Av$, $\forall v \in \mathbb{R}^n$. Consideremos $W_i = \operatorname{Ker} q_i(T)$ y $V_j = \operatorname{Ker} p_j(T)$, siendo $q_i(t) = (t-\lambda_i)^{n_i}$ y $p_j(t) = \left(t^2 + c_j t + d_j\right)^{m_j}$, $i = 1, \ldots, h$, $j = 1, \ldots, k$. Como los factores de $\chi_A(t)$ en la descomposición (7.18) son primos entre sí, entonces el Teorema 5.2.10 implica

$$\mathbb{R}^n = \bigoplus_{i=1}^{h} W_i \oplus \bigoplus_{j=1}^{k} V_j.$$

Sean $\mathcal{B}_i$ una base cualquiera de W_i y $\mathcal{C}_j$ una base cualquiera de V_j. Al ser W_i y V_j subespacios T-invariantes, el conjunto $\mathcal{B} := \mathcal{B}_1 \cup \cdots \mathcal{B}_h \cup \mathcal{C}_1 \cup \cdots \cup \mathcal{C}_k$ es una base de $\mathbb{R}^n$. Luego aplicando la Proposición 5.1.4 obtenemos

$$[T]_{\mathcal{B}} = \begin{pmatrix} A_1 & & & & & \\ & \ddots & & & & \\ & & A_h & & & \\ & & & B_1 & & \\ & & & & \ddots & \\ & & & & & B_k \end{pmatrix},$$

[3] Se entiende que la misma a menos de permutar los bloques correspondientes a distintos valores propios.

siendo $A_i = [T|_{W_i}]_{\mathcal{B}_i} \in M_{n_i}(\mathbb{R})$, $B_j = \big[T|_{V_j}\big]_{\mathcal{C}_j} \in M_{m_j}(\mathbb{R})$. Notar que para cada i, j vale

$$\chi_{A_i} = (-1)^{n_i}(t-\lambda_i)^{n_i} \quad \text{y} \quad \chi_{B_j}(t) = \big(t^2 + c_j t + d_j\big)^{m_j}.$$

Sabemos que cada matriz A_i se descompone en bloques de Jordan correspondientes al valor propio λ_i. El siguiente teorema nos da una forma de descomponer las matrices B_j.

Teorema 7.10.1. *Sea $A \in M_n(\mathbb{R})$ tal que su polinomio característico es*[4] $\chi_A(t) = \big(t^2 + ct + d\big)^m$, *con $c^2 - 4d < 0$. Escribimos*

$$t^2 + c\,t + d = (t-a)^2 + b^2, \quad \textit{siendo} \quad a = -\frac{c}{2},\ b = \frac{\sqrt{4d - c^2}}{2}.$$

Notar que es $b \geq 0$ y que $a \pm ib$ son las raíces complejas (conjugadas) de $\chi_A(t)$.

Luego A es semejante en $M_n(\mathbb{R})$ a una matriz en bloques de la forma

$$\begin{pmatrix} F & I & & & \\ & F & I & & \\ & & \ddots & \ddots & \\ & & & F & I \\ & & & & F \end{pmatrix}, \quad F = \begin{pmatrix} a & b \\ -b & a \end{pmatrix}, \quad I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}. \tag{7.19}$$

Dem. Sean $T \in \mathcal{L}(\mathbb{R}^n)$ definida por $T(v) = Av$, $v \in \mathbb{R}^n$ y $T_\mathbb{C} \in \mathcal{L}(\mathbb{C}^n)$ definida por $T_\mathbb{C}(v) = Av$, $v \in \mathbb{C}^n$.

Consideremos $T_\mathbb{C} \in \mathcal{L}(\mathbb{C}^n)$. Es

$$\chi_{T_\mathbb{C}}(t) = \chi_A(t) = \big(t^2 + ct + d\big)^m = \big((t-a)^2 + b^2\big)^m = (t-\alpha)^m (t-\overline{\alpha})^m,$$

siendo $a = -\frac{c}{2}$, $b = \sqrt{d - \frac{c^2}{4}}$, $\alpha = a + i\,b$ y $\overline{\alpha} = a - i\,b$.

Como $\chi_{T_\mathbb{C}}(t) = (t-\alpha)^m(t-\overline{\alpha})^m$, entonces el Teorema 5.2.10 implica $\mathbb{C}^n = U_\alpha \oplus U_{\overline{\alpha}}$, siendo $U_\alpha = \operatorname{Ker}(T_\mathbb{C} - \alpha\,\mathrm{id})^m$ y $U_{\overline{\alpha}} = \operatorname{Ker}(T_\mathbb{C} - \overline{\alpha}\,\mathrm{id})^m$. A continuación veremos que los elementos de $U_{\overline{\alpha}}$ son los conjugados de los de U_α.

Afirmación: $U_{\overline{\alpha}} = \{\overline{v} :\ v \in U_\alpha\}$.

$$\begin{aligned} v \in U_{\overline{\alpha}} &\Leftrightarrow (T_\mathbb{C} - \overline{\alpha}\,\mathrm{id})^m(v) = 0 \Leftrightarrow (A - \overline{\alpha} I)^m \cdot v = 0 \Leftrightarrow \overline{(A - \overline{\alpha} I)^m \cdot v} = \overline{0} \\ &\Leftrightarrow \overline{(A - \overline{\alpha} I)^m} \cdot \overline{v} = 0 \Leftrightarrow \big(\,\overline{A - \overline{\alpha} I}\,\big)^m \cdot \overline{v} = 0 \Leftrightarrow (A - \alpha I)^m \cdot \overline{v} = 0 \\ &\Leftrightarrow (T_\mathbb{C} - \alpha\,\mathrm{id})^m(\overline{v}) = 0 \Leftrightarrow \overline{v} \in U_\alpha. \end{aligned}$$

La forma de Jordan para el caso de operadores complejos nos dice que existe una base de U_α formada por unión de ciclos de $T_\mathbb{C}$ correspondientes al valor propio α y lo mismo para $U_{\overline{\alpha}}$ respecto al valor propio $\overline{\alpha}$. Lo que sigue muestra cómo obtener los ciclos correspondientes a $\overline{\alpha}$ a partir de los de α.

Afirmación: Si $\{w_1, \ldots, w_r\}$ es un ciclo para $T_\mathbb{C}$ correspondiente al valor propio α, entonces $\{\overline{w_1}, \ldots, \overline{w_r}\}$ es un ciclo para $T_\mathbb{C}$ correspondiente al valor propio $\overline{\alpha}$.

$$\left\{ \begin{array}{l} T_\mathbb{C}(w_1) = \alpha\, w_1 \\ T_\mathbb{C}(w_l) = w_{l-1} + \alpha\, w_l \end{array} \right. \Rightarrow \left\{ \begin{array}{l} A \cdot w_1 = \alpha\, w_1 \\ A \cdot w_l = w_{l-1} + \alpha\, w_l \end{array} \right. \Rightarrow \left\{ \begin{array}{l} \overline{A \cdot w_1} = \overline{\alpha\, w_1} \\ \overline{A \cdot w_l} = \overline{w_{l-1} + \alpha\, w_l} \end{array} \right.$$

$$\Rightarrow \left\{ \begin{array}{l} A \cdot \overline{w_1} = \overline{\alpha}\, \overline{w_1} \\ A \cdot \overline{w_l} = \overline{w_{l-1}} + \overline{\alpha}\, \overline{w_l} \end{array} \right. \Rightarrow \left\{ \begin{array}{l} T_\mathbb{C}(\overline{w_1}) = \overline{\alpha}\, \overline{w_1} \\ T_\mathbb{C}(\overline{w_l}) = \overline{w_{l-1}} + \overline{\alpha}\, \overline{w_l},\ \forall l = 2, \ldots, r \end{array} \right.$$

y es claro que $w_1 \neq 0$ implica $\overline{w_1} \neq 0$, luego $\{\overline{w_1}, \ldots, \overline{w_r}\}$ es un ciclo.

[4] Observar que falta un factor $(-1)^n$ en $\chi_A(t)$. Pero como el grado de $\chi_A(t)$ es $2m$, entonces $n = 2m$ y por lo tanto $(-1)^n = 1$.

Sea $\mathcal{C} = \mathcal{C}_1 \cup \cdots \cup \mathcal{C}_p$ una base de U_α en la cual los $\mathcal{C}_1, \ldots, \mathcal{C}_p$ son ciclos disjuntos para $T_{\mathbb{C}}|_{U_\alpha}$ correspondientes al valor propio α (es decir, $\mathcal{C}$ es una base de Jordan de U_α para $T_{\mathbb{C}}|_{U_\alpha}$). Para cada $l = 1, \ldots, p$, si $\mathcal{C}_l = \{w_1, \ldots, w_r\}$, escribimos $\overline{\mathcal{C}_l} = \{\overline{w_1}, \ldots, \overline{w_r}\}$. De las afirmaciones anteriores se deduce que $\overline{\mathcal{C}} := \overline{\mathcal{C}_1} \cup \cdots \cup \overline{\mathcal{C}_p}$ es una base de $U_{\overline{\alpha}}$ en la cual los $\overline{\mathcal{C}_1}, \ldots, \overline{\mathcal{C}_p}$ son ciclos disjuntos para $T_{\mathbb{C}}|_{U_{\overline{\alpha}}}$ correspondientes al valor propio $\overline{\alpha}$. Luego la matriz asociada a $T_{\mathbb{C}}$ en la base $\mathcal{C} \cup \overline{\mathcal{C}}$ es

$$[T_{\mathbb{C}}]_{\mathcal{C}\cup\overline{\mathcal{C}}} = \begin{pmatrix} J_1 & & & & & \\ & \ddots & & & & \\ & & J_p & & & \\ & & & \overline{J_1} & & \\ & & & & \ddots & \\ & & & & & \overline{J_p} \end{pmatrix}, \quad J_l = \begin{pmatrix} \alpha & 1 & & & \\ & \alpha & 1 & & \\ & & \ddots & \ddots & \\ & & & \alpha & 1 \\ & & & & \alpha \end{pmatrix}, \ \forall l.$$

Si ahora reordenamos los elementos de $\mathcal{C} \cup \overline{\mathcal{C}}$ en la forma $\mathcal{D} = \left(\mathcal{C}_1 \cup \overline{\mathcal{C}_1}\right) \cup \cdots \cup \left(\mathcal{C}_p \cup \overline{\mathcal{C}_p}\right)$, entonces la matriz asociada a $T_{\mathbb{C}}$ en la base $\mathcal{D}$ es

$$[T_{\mathbb{C}}]_{\mathcal{D}} = \begin{pmatrix} J_1 & & & & \\ & \overline{J_1} & & & \\ & & \ddots & & \\ & & & J_p & \\ & & & & \overline{J_p} \end{pmatrix}.$$

Cada submatriz de $[T_{\mathbb{C}}]_{\mathcal{D}}$ de la forma $\left(\begin{smallmatrix} J_j & \\ & \overline{J_j} \end{smallmatrix}\right)$ corresponde a una parte de la base $\mathcal{D}$ de la forma

$$\mathcal{C}_j \cup \overline{\mathcal{C}}_j = \{w_1, \ldots, w_r, \overline{w_1}, \ldots, \overline{w_r}\}.$$

Como $\mathcal{C}_j$ y $\overline{\mathcal{C}}_j$ son ciclos, es

$$\left.\begin{array}{ll} T_{\mathbb{C}}(w_1) &= \alpha\, w_1, \quad T_{\mathbb{C}}(w_l) = w_{l-1} + \alpha\, w_l, \\ T_{\mathbb{C}}(\overline{w_1}) &= \overline{\alpha}\, \overline{w_1}, \quad T_{\mathbb{C}}(\overline{w_l}) = \overline{w_{l-1}} + \overline{\alpha}\, \overline{w_l} \end{array}\right\},\ l = 2, \ldots, r. \tag{7.20}$$

Sea $\mathcal{E}_j = \{w_1, \overline{w_1}, w_2, \overline{w_2}, \ldots, w_r, \overline{w_r}\}$. Observar que $\mathcal{E}_j$ y $\mathcal{C}_j \cup \overline{\mathcal{C}}_j$ coinciden a menos del orden, luego generan el mismo subespacio de $\mathbb{C}^n$. Es un ejercicio el verificar que si definimos $\mathcal{E} := \mathcal{E}_1 \cup \cdots \cup \mathcal{E}_p$, entonces es

$$[T_{\mathbb{C}}]_{\mathcal{E}} = \begin{pmatrix} E & I & & & \\ & E & I & & \\ & & \ddots & \ddots & \\ & & & E & I \\ & & & & E \end{pmatrix}, \quad E = \begin{pmatrix} \alpha & 0 \\ 0 & \overline{\alpha} \end{pmatrix}, \quad I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Consideremos una parte $\mathcal{E}_j = \{w_1, \overline{w_1}, w_2, \overline{w_2}, \ldots, w_r, \overline{w_r}\}$ de la base $\mathcal{E}$. Cada vector w_l se puede descomponer de la forma $w_l = p_l + i\, q_l$, $l = 1, \ldots, r$, siendo p_l, q_l vectores de $\mathbb{R}^n$. Sea $\mathcal{F}_j = \{p_1, q_1, p_2, q_2, \ldots, p_r, q_r\}$, $j = 1, \ldots, p$ y $\mathcal{F} = \mathcal{F}_1 \cup \cdots \cup \mathcal{F}_p$. Observar que valen las siguientes relaciones

$$\begin{array}{ll} w_l = p_l + i\, q_l, & p_l = \frac{1}{2}(w_l + \overline{w_l}) \\ \overline{w_l} = p_l - i\, q_l, & q_l = -\frac{i}{2}(w_l - \overline{w_l}) \end{array}, \quad l = 1, \ldots, r. \tag{7.21}$$

Esto implica que $\mathcal{E}_j$ y $\mathcal{F}_j$ generan el mismo subespacio de $\mathbb{C}^n$ (que es $T_{\mathbb{C}}$-invariante), $\forall j = 1, \ldots, p$. Luego el subespacio generado por $\mathcal{F}$ coincide con el subespacio generado por $\mathcal{E}$ que es $\mathbb{C}^n$ y por lo tanto $\mathcal{F}$ es una base de $\mathbb{C}^n$. Como el subespacio generado por cada $\mathcal{F}_j$ es $T_{\mathbb{C}}$-invariante, es

$$[T_{\mathbb{C}}]_{\mathcal{F}} = \begin{pmatrix} [T_{\mathbb{C}}]_{\mathcal{F}_1} & & & \\ & [T_{\mathbb{C}}]_{\mathcal{F}_2} & & \\ & & \ddots & \\ & & & [T_{\mathbb{C}}]_{\mathcal{F}_p} \end{pmatrix}.$$

Es un ejercicio el verificar que las relaciones (7.20) y (7.21) implican

$$[T_{\mathbb{C}}]_{\mathcal{F}_j} = \begin{pmatrix} F & I & & & \\ & F & I & & \\ & & \ddots & \ddots & \\ & & & F & I \\ & & & & F \end{pmatrix}, \quad F = \begin{pmatrix} a & b \\ -b & a \end{pmatrix}, \quad I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Luego obtuvimos una base $\mathcal{F}$ de $\mathbb{C}^n$ formada por vectores de $\mathbb{R}^n$ tal que $[T_{\mathbb{C}}]_{\mathcal{F}}$ es una matriz con coeficientes reales. Notar que $\mathcal{F}$ está contenido en $\mathbb{R}^n$, es[5] LI y tiene n elementos, luego $\mathcal{F}$ es una base de $\mathbb{R}^n$ como $\mathbb{R}$-espacio vectorial y por lo tanto $[T]_{\mathcal{F}} = [T_{\mathbb{C}}]_{\mathcal{F}}$. Finalmente la tesis se deduce de que siendo A la matriz asociada a T en la base canónica de $\mathbb{R}^n$, resulta que A y $[T]_{\mathcal{F}}$ son semejantes en $M_n(\mathbb{R})$. □

Definición 7.10.2. Sea $A \in M_n(\mathbb{R})$ y supongamos que la descomposición factorial de $\chi_A(t)$ en $\mathbb{R}[t]$ es

$$\chi_A(t) = (-1)^n \prod_{i=1}^{h} (t-\lambda_i)^{n_i} \prod_{j=1}^{k} \left(t^2 + c_j t + d_j\right)^{m_j}, \quad \lambda_i, c_j, d_j \in \mathbb{R},\ c_j^2 - 4\,d_j < 0,\ \forall i,j.$$

Esta factorización implica que A es semejante a una matriz en bloques de la forma

$$\begin{pmatrix} A_1 & & & & & \\ & \ddots & & & & \\ & & A_h & & & \\ & & & B_1 & & \\ & & & & \ddots & \\ & & & & & B_k \end{pmatrix}. \tag{7.22}$$

De acuerdo a la teoría general de la forma de Jordan para operadores cuyo polinomio característico escinde, sabemos que cada bloque A_i podemos tomarlo como una matriz en bloques de Jordan correspondientes al valor propio λ_i, ordenados en tamaño decreciente. Por otro lado, el teorema 7.10.1 nos dice que cada bloque B_j se puede tomar descompuesto en bloques de la forma (7.19), ordenados también en tamaño decreciente. Una matriz del tipo (7.22) con los A_i y B_j elegidos de la forma descrita anteriormente se dice que es una *forma de Jordan real* de la matriz A.

Observación 7.10.3. Notar que en la forma de Jordan real estamos eligiendo los bloques de la forma $\left(\begin{smallmatrix} a & b \\ -b & a \end{smallmatrix}\right)$ con $b \geq 0$. Esto lo hacemos para obtener unicidad en la forma de Jordan, ya que las matrices reales del tipo

$$A = \begin{pmatrix} a & b \\ -b & a \end{pmatrix}, \quad B = \begin{pmatrix} a & -b \\ b & a \end{pmatrix},$$

[5] Es claro que $\mathcal{F}$ es LI en $\mathbb{C}^n$, pero esto equivale a que el determinante de la matriz cuyas columnas son los vectores de $\mathcal{F}$ sea distinto de cero (lo cual no depende de estar en $\mathbb{R}$ o en $\mathbb{C}$), luego $\mathcal{F}$ es LI en $\mathbb{R}^n$.

son semejantes. Para verlo, observar que si definimos $T : \mathbb{R}^2 \to \mathbb{R}^2$ de forma tal que $A = [T]_{\{e_1,e_2\}}$, entonces $B = [T]_{\{e_2,e_1\}}$. Esta semejanza refleja el hecho de que ambas matrices tienen el mismo polinomio característico

$$\chi_A(t) = \chi_B(t) = (t-a)^2 + b^2 = (t-\alpha)(t-\overline{\alpha}), \quad \alpha = a + ib$$

y por lo tanto tienen la misma forma de Jordan compleja.

El siguiente es la versión en real del corolario 7.9.9.

Proposición 7.10.4. *Si A y B están en $M_n(\mathbb{R})$, entonces A y B son semejantes en $M_n(\mathbb{R})$ si y solo si tienen la misma forma de Jordan real.*

Dem. Es claro que si A y B tienen la misma forma de Jordan real, entonces son semejantes en $M_n(\mathbb{R})$. Para el recíproco, observar que si A y B son semejantes en $M_n(\mathbb{R})$, entonces lo son en $M_n(\mathbb{C})$; luego el corolario 7.9.9 implica que A y B tienen la misma forma de Jordan compleja y por lo tanto tienen la misma forma de Jordan real. $\square$

A continuación veremos el cálculo de la forma de Jordan real en algunos casos concretos.

Ejemplo 7.10.5. Sea $A = \left(\begin{smallmatrix} -1 & 17 \\ -2 & 5 \end{smallmatrix}\right)$. Su polinomio característico es

$$\chi_A(t) = t^2 - 4t + 29 = (t-2)^2 + 25 = (t-2-5i)(t-2+5i).$$

Luego las formas de Jordan compleja $J_A^{\mathbb{C}}$ y real $J_A^{\mathbb{R}}$ de A son

$$J_A^{\mathbb{C}} = \begin{pmatrix} 2+5i & 0 \\ 0 & 2-5i \end{pmatrix}, \quad J_A^{\mathbb{R}} = \begin{pmatrix} 2 & 5 \\ -5 & 2 \end{pmatrix}.$$

Ejemplo 7.10.6. Sea $A \in M_4(\mathbb{R})$ y supongamos que $\chi_A(t) = (t^2 - 4t + 29)(t^2 + 9)$. Es

$$\chi_A(t) = \left((t-2)^2 + 25\right)(t^2 + 9) = (t-2-5i)(t-2+5i)(t-3i)(t+3i).$$

Luego

$$J_A^{\mathbb{C}} = \begin{pmatrix} 2+5i & & & \\ & 2-5i & & \\ & & 3i & \\ & & & -3i \end{pmatrix}, \quad J_A^{\mathbb{R}} = \begin{pmatrix} 2 & 5 & & \\ -5 & 2 & & \\ & & 0 & 3 \\ & & -3 & 0 \end{pmatrix}.$$

Ejemplo 7.10.7. Sea $A \in M_4(\mathbb{R})$ y supongamos que $\chi_A(t) = t^4 - 1$. Es

$$t^4 - 1 = (t-1)(t+1)(t^2+1) = (t-1)(t+1)(t-i)(t+i).$$

Luego

$$J_A^{\mathbb{C}} = \begin{pmatrix} 1 & & & \\ & -1 & & \\ & & i & \\ & & & -i \end{pmatrix}, \quad J_A^{\mathbb{R}} = \begin{pmatrix} 1 & & & \\ & -1 & & \\ & & 0 & 1 \\ & & -1 & 0 \end{pmatrix}.$$

Notar que en los tres ejemplos anteriores la matriz A es diagonalizable en $\mathbb{C}$ pero no lo es en $\mathbb{R}$.

Ejemplo 7.10.8. Sea $A \in M_4(\mathbb{R})$ y supongamos que $\chi_A(t) = (t-2)^2(t^2+9)$. Observar que tenemos dos posibilidades

$$J_A^{\mathbb{C}} = \begin{pmatrix} 2 & & & \\ & 2 & & \\ & & 3i & \\ & & & -3i \end{pmatrix}, \; J_A^{\mathbb{R}} = \begin{pmatrix} 2 & & & \\ & 2 & & \\ & & 0 & 3 \\ & & -3 & 0 \end{pmatrix} \quad \text{o} \quad J_A^{\mathbb{C}} = \begin{pmatrix} 2 & 1 & & \\ & 2 & & \\ & & 3i & \\ & & & -3i \end{pmatrix}, \; J_A^{\mathbb{R}} = \begin{pmatrix} 2 & 1 & & \\ & 2 & & \\ & & 0 & 3 \\ & & -3 & 0 \end{pmatrix}.$$

En el primer caso el polinomio minimal de A es $m_A(t) = (t-2)(t^2+9)$ y en el segundo es $m_A(t) = (t-2)^2(t^2+9)$. Luego para esta matriz, conociendo el polinomio minimal podemos deducir la forma de Jordan real (sin necesitar calcular la forma de Jordan compleja).

Ejemplo 7.10.9. Sea $A \in M_4(\mathbb{R})$ y supongamos que $\chi_A(t) = (t^2-4t+29)^2$. Es

$$\chi_A(t) = \left((t-2)^2+25\right)^2 = (t-2-5i)^2(t-2+5i)^2,$$

En este caso el teorema anterior implica que las multiplicidades geométricas de los valores propios $2+5i$ y $2-5i$ son las mismas. Luego tenemos dos posibilidades para la forma de Jordan[6]:

$$J_A^{\mathbb{C}} = \begin{pmatrix} 2+5i & 0 & 0 & 0 \\ 0 & 2+5i & 0 & 0 \\ 0 & 0 & 2-5i & 0 \\ 0 & 0 & 0 & 2-5i \end{pmatrix} \text{ y } J_A^{\mathbb{R}} = \begin{pmatrix} 2 & 5 & 0 & 0 \\ -5 & 2 & 0 & 0 \\ 0 & 0 & 2 & 5 \\ 0 & 0 & -5 & 2 \end{pmatrix} \text{ o}$$
$$J_A^{\mathbb{C}} = \begin{pmatrix} 2+5i & 1 & 0 & 0 \\ 0 & 2+5i & 0 & 0 \\ 0 & 0 & 2-5i & 1 \\ 0 & 0 & 0 & 2-5i \end{pmatrix} \text{ y } J_A^{\mathbb{R}} = \begin{pmatrix} 2 & 5 & 1 & 0 \\ -5 & 2 & 0 & 1 \\ 0 & 0 & 2 & 5 \\ 0 & 0 & -5 & 2 \end{pmatrix}.$$

Queda como ejercicio el verificar que la primera y segunda forma corresponden respectivamente a

$$m_A(t) = t^2-4t+29 \quad \text{y} \quad m_A(t) = (t^2-4t+29)^2.$$

Una forma de obtener la base de Jordan real $\mathcal{B}_{\mathbb{R}}$, es obtener previamente la base $\mathcal{B}_{\mathbb{C}}$ correspondiente a la forma de Jordan compleja en la forma habitual y luego aplicarle el procedimiento descrito en la prueba del teorema anterior. De hecho basta con obtener la mitad de la base, es decir los vectores que corresponden al valor propio $2+5i$, dado que sabemos que los que corresponden a $2-5i$ son los conjugados de los anteriores. Por ejemplo, si

$$\{(1+3i,2,i,3-5i),\ (2-3i,3+i,1-i,2i)\} \tag{7.23}$$

es una base de $\operatorname{Ker}\left(L_A-(2+5i)\operatorname{id}\right)^2$, entonces la base correspondiente a la forma de Jordan compleja es

$$\mathcal{B}_{\mathbb{C}} = \{(1+3i,2,i,3-5i),\ (2-3i,3+i,1-i,2i),\ (1-3i,2,-i,3+5i),\ (2+3i,3-i,1+i,-2i)\}.$$

La base de Jordan real se obtiene de (7.23) descomponiendo cada vector en su parte real e imaginaria,

$$\begin{aligned}(1+3i,2,i,3-5i) &= (1,2,0,3)+i(3,0,1,-5),\\ (2-3i,3+i,1-i,2i) &= (2,3,1,0)+i(-3,1,-1,2).\end{aligned}$$

Luego

$$\mathcal{B}_{\mathbb{R}} = \{(1,2,0,3),\ (3,0,1,-5),\ (2,3,1,0),\ (-3,1,-1,2)\}.$$

Finalmente a partir de $\mathcal{B}_{\mathbb{R}}$ obtenemos la matriz $Q = \begin{pmatrix} 1 & 3 & 2 & -3 \\ 2 & 0 & 3 & 1 \\ 0 & 1 & 1 & -1 \\ 3 & -5 & 0 & 2 \end{pmatrix}$ que verifica $J_A^{\mathbb{R}} = Q^{-1}A\,Q$.

[6]En este caso estamos escribiendo todos los ceros en las matrices, para dar un ejemplo de cómo se escriben realmente.

Ejemplo 7.10.10. Sea $A \in M_6(\mathbb{R})$ y supongamos que $\chi_A(t) = (t^2+25)^3$. Para el polinomio minimal de A tenemos tres posibilidades:

$$m_A(t) = t^2+25, \quad m_A(t) = (t^2+25)^2, \quad m_A(t) = (t^2+25)^3.$$

Las correspondientes formas de Jordan reales son

$$\begin{pmatrix} 0 & 5 & & & & \\ -5 & 0 & & & & \\ & & 0 & 5 & & \\ & & -5 & 0 & & \\ & & & & 0 & 5 \\ & & & & -5 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 5 & 1 & 0 & & \\ -5 & 0 & 0 & 1 & & \\ & & 0 & 5 & & \\ & & -5 & 0 & & \\ & & & & 0 & 5 \\ & & & & -5 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 5 & 1 & 0 & & \\ -5 & 0 & 0 & 1 & & \\ & & 0 & 5 & 1 & 0 \\ & & -5 & 0 & 0 & 1 \\ & & & & 0 & 5 \\ & & & & -5 & 0 \end{pmatrix}.$$

7.11. Formas multilineales alternadas

El objetivo de esta sección es dar una introducción lo más elemental posible al tema de las formas multilineales alternadas. Para un tratamiento más avanzado se sugiere leer algún texto sobre formas diferenciales.

Sea $\Bbbk$ un cuerpo arbitrario, V un $\Bbbk$-espacio vectorial y n un entero positivo. Una *n-forma multilineal* en V es una función $\alpha : \underbrace{V \times \cdots \times V}_{n} \to \Bbbk$ que verifica:

$$\alpha(v_1, \ldots, a\, v_i + w_i, \ldots, v_n) = a\, \alpha(v_1, \ldots, v_i, \ldots, v_n) + \alpha(v_1, \ldots, w_i, \ldots, v_n)$$

para todo $a \in \Bbbk$, $v_1, \ldots, v_n, w_i \in V$ y todo $i = 1, \ldots, n$.

Observar que una forma multilineal es una función que es lineal en cada variable, en particular una 1-forma multilineal es simplemente un elemento del espacio dual V^*.

Es un ejercicio el verificar que el conjunto de las n-formas multilineales es un subespacio del espacio de las funciones con dominio $V \times \cdots \times V$ y codominio $\Bbbk$, en el cual las operaciones se definen punto a punto:

$$(f+g)(v_1, \ldots, v_n) = f(v_1, \ldots, v_n) + g(v_1, \ldots, v_n), \quad (a\, f)(v_1, \ldots, v_n) = a\, f(v_1, \ldots, v_n).$$

Luego el conjunto de las n-formas multilineales es un espacio vectorial.

Una k-forma ω en V se dice *alternada* si verifica:

$$\omega(v_1, \ldots, v_i, \ldots, v_j, \ldots, v_k) = 0 \tag{7.24}$$

cada vez que existan $i \neq j$ tales que $v_i = v_j$.

Observar que si ω es una k-forma alternada, entonces verifica:

$$\omega(v_1, \ldots, v_i, \ldots, v_j, \ldots, v_k) = -\omega(v_1, \ldots, v_j, \ldots, v_i, \ldots, v_k)\,, \quad \forall i \neq j,\ v_1, \ldots, v_k \in V. \tag{7.25}$$

En efecto, es

$$\begin{aligned} 0 &= \omega(v_1, \ldots, v_i + v_j, \ldots, v_i + v_j, \ldots, v_k) \\ &= \omega(v_1, \ldots, v_i, \ldots, v_i, \ldots, v_k) + \omega(v_1, \ldots, v_i, \ldots, v_j, \ldots, v_k) + \omega(v_1, \ldots, v_j, \ldots, v_i, \ldots, v_k) \\ &\quad + \omega(v_1, \ldots, v_j, \ldots, v_j, \ldots, v_k) \\ &= 0 + \omega(v_1, \ldots, v_i, \ldots, v_j, \ldots, v_k) + \omega(v_1, \ldots, v_j, \ldots, v_i, \ldots, v_k) + 0 \\ &= \omega(v_1, \ldots, v_i, \ldots, v_j, \ldots, v_k) + \omega(v_1, \ldots, v_j, \ldots, v_i, \ldots, v_k)\,, \end{aligned}$$

luego $\omega(v_1, \ldots, v_i, \ldots, v_j, \ldots, v_k) = -\omega(v_1, \ldots, v_j, \ldots, v_i, \ldots, v_k)$.

Recíprocamente, si $\operatorname{car}\Bbbk \neq 2$ y ω es una k-forma multilineal que verifica (7.25) entonces verifica (7.24):

Sean $v_1,\ldots,v_i,\ldots,v_j,\ldots,v_k \in V$ con $v_i = v_j$ e $i \neq j$. Utilizando (7.25) y que es $v_i = v_j$ obtenemos:

$$\omega(v_1,\ldots,v_i,\ldots,v_j,\ldots,v_k) = -\omega(v_1,\ldots,v_j,\ldots,v_i,\ldots,v_k) = -\omega(v_1,\ldots,v_i,\ldots,v_j,\ldots,v_k)\,.$$

Luego $2\,\omega(v_1,\ldots,v_i,\ldots,v_j,\ldots,v_k) = 0$ y como $\operatorname{car}\Bbbk \neq 2$ es $\omega(v_1,\ldots,v_i,\ldots,v_j,\ldots,v_k) = 0$.

Observar que lo anterior prueba que (7.24) y (7.25) son equivalentes si $\operatorname{car}\Bbbk \neq 2$.

Escribiremos

$$\operatorname{Alt}_k(V) = \{\omega : \underbrace{V\times\cdots\times V}_{k} \to \Bbbk :\ \omega \text{ es una } k\text{-forma multilineal alternada}\}, \quad k = 1,2,\ldots$$

Es un ejercicio el verificar que $\operatorname{Alt}^k(V)$ es un subespacio del espacio de las k-formas multilineales en V, luego $\operatorname{Alt}^k(V)$ es un espacio vectorial. Observar que $\operatorname{Alt}^1(V) = V^*$. Definimos $\operatorname{Alt}^0(V) := \Bbbk$.

Ejemplo 7.11.1. Sea $\omega : \Bbbk^3 \times \Bbbk^3 \to \Bbbk$ definida por $\omega((x,y,z),(x',y',z')) = x\,y' - y\,x' + 2\,y\,z' - 2\,y'\,z$. Observar que ω se puede escribir como el siguiente producto matricial:

$$\begin{aligned}\omega\big((x,y,z),(x',y',z')\big) &= x\,y' + y\big(-x' + 2\,z'\big) - 2\,y'\,z \\ &= (x,y,z)\begin{pmatrix} y' \\ -x' + 2\,z' \\ -2\,y' \end{pmatrix} = (x,y,z)\begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 2 \\ 0 & -2 & 0 \end{pmatrix}\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix}.\end{aligned} \tag{7.26}$$

De esta fórmula se deduce que ω es una forma bilineal y es fácil probar que ω es alternada. Observar que la matriz en la igualdad (7.26) es antisimétrica, es decir verifica $A^t = -A$. Veremos que esto vale en general.

Recordar que si V tiene dimensión finita n, $\mathcal{B} = \{v_1,\ldots,v_n\}$ es una base de V y $\omega \in \operatorname{Bil}(V)$, entonces la matriz asociada a la forma bilineal ω en la base $\mathcal{B}$ es

$$M_{\mathcal{B}}(\omega) = \begin{pmatrix} \omega(v_1,v_1) & \cdots & \omega(v_1,v_n) \\ \vdots & & \vdots \\ \omega(v_n,v_1) & \cdots & \omega(v_n,v_n) \end{pmatrix}.$$

En la proposición 4.1.7 probamos que vale

$$\omega(u,v) = \operatorname{coord}_{\mathcal{B}}(u)^t M_{\mathcal{B}}(\omega)\operatorname{coord}_{\mathcal{B}}(v), \quad \forall u,v \in V. \tag{7.27}$$

Luego es fácil probar que ω es una 2-forma alternada si y solo si la matriz $M_{\mathcal{B}}(\omega)$ es antisimétrica, es decir

$$\omega(v_i,v_j) = -\omega(v_j,v_i)\,,\ \forall i \neq j; \qquad \omega(v_i,v_i) = 0,\ \forall i.$$

Ejemplo 7.11.2. Observando $\underbrace{\Bbbk^n \times\cdots\times \Bbbk^n}_{n} \simeq M_n(\Bbbk)$, podemos definir $\det : \Bbbk^n \times\cdots\times \Bbbk^n \to \Bbbk$ mediante

$$\det\!\big((x_{11},\ldots,x_{n1}),\ldots,(x_{1n},\ldots,x_{nn})\big) = \begin{vmatrix} x_{11} & \cdots & x_{1n} \\ \vdots & & \vdots \\ x_{n1} & \cdots & x_{nn} \end{vmatrix}.$$

De las propiedades del determinante se deduce inmediatamente que det es una n-forma alternada en $\Bbbk^n$.

De ahora en adelante supondremos que V es un espacio vectorial de dimensión finita n.

Proposición 7.11.3. *Para todo $k > n$ es $\operatorname{Alt}^k(V) = \{0\}$.*

Dem. Consideremos $\{e_1, \ldots, e_n\}$ una base de V y sea $\omega \in \operatorname{Alt}^k(V)$, con $k > n$. Sean $v_1, \ldots, v_k \in V$. Escribiendo $v_1 = \sum_{j_1=1}^n a_{j_1} e_{j_1}, \ldots, v_k = \sum_{j_k=1}^n a_{j_k} e_{j_k}$, es

$$\omega(v_1, \ldots, v_k) = \omega\left(\sum_{j_1=1}^n a_{j_1} e_{j_1}, \ldots, \sum_{j_k=1}^n a_{j_k} e_{j_k}\right) = \sum_{j_1=\cdots=j_k=1}^n a_{j_1} \cdots a_{j_k} \omega(e_{j_1}, \ldots, e_{j_k}).$$

Como $k > n$, entonces en $\{e_{j_1}, \ldots, e_{j_k}\}$ necesariamente hay elementos repetidos; luego $\omega(e_{j_1}, \ldots, e_{j_k}) = 0$, $\forall j_1, \ldots, j_k$ y por lo tanto $\omega(v_1, \ldots, v_k) = 0$. Como $v_1, \ldots, v_k$ son arbitrarios, se concluye que es $\omega = 0$ $\square$

Proposición 7.11.4. *Sea $\omega \in \operatorname{Alt}^k(V)$, $k \leq n$. Si existe $\{e_1, \ldots, e_n\}$ base de V tal que*

$$\omega(e_{i_1}, \ldots, e_{i_k}) = 0,\ \forall i_1, \ldots, i_k \in \{1, \ldots, n\} \text{ tales que } 1 \leq i_1 < i_2 < \cdots < i_k \leq n, \tag{7.28}$$

entonces $\omega = 0$.

Dem. Supongamos ω verifica (7.28). Sean $v_1, \ldots, v_k \in V$. Escribiendo $v_1 = \sum_{j_1=1}^n a_{j_1} e_{j_1}, \ldots, v_k = \sum_{j_k=1}^n a_{j_k} e_{j_k}$, es

$$\omega(v_1, \ldots, v_k) = \omega\left(\sum_{j_1=1}^n a_{j_1} e_{j_1}, \ldots, \sum_{j_k=1}^n a_{j_k} e_{j_k}\right) = \sum_{j_1=\cdots=j_k=1}^n a_{j_1} \cdots a_{j_k} \omega(e_{j_1}, \ldots, e_{j_k}). \tag{7.29}$$

Si existen $p \neq q$ tales que $e_{j_p} = e_{j_q}$, entonces $\omega(e_{j_1}, \ldots, e_{j_k}) = 0$. Por otro lado, si $e_{j_1}, \ldots, e_{j_k}$ son distintos entre sí, podemos reordenarlos para obtener un conjunto $e_{i_1}, \ldots, e_{i_k}$ con $1 \leq i_1 < i_2 < \cdots < i_k \leq n$. Notar que siempre podemos pasar de la k-upla $(e_{j_1}, \ldots, e_{j_k})$ a la k-upla $(e_{i_1}, \ldots, e_{i_k})$ realizando una cantidad finita de intercambios entre los elementos de $(e_{j_1}, \ldots, e_{j_k})$; cada intercambio corresponde a un cambio de signo en ω, de donde aplicando (7.28) obtenemos

$$\omega(e_{j_1}, \ldots, e_{j_k}) = \pm\, \omega(e_{i_1}, \ldots, e_{i_k}) = 0.$$

Luego todos los sumandos de (7.29) son nulos y por lo tanto $\omega(v_1, \ldots, v_k) = 0$. $\square$

Corolario 7.11.5. *Sean $\omega, \eta \in \operatorname{Alt}^k(V)$, $k \leq n$. Si existe $\{e_1, \ldots, e_n\}$ base de V tal que*

$$\omega(e_{i_1}, \ldots, e_{i_k}) = \eta(e_{i_1}, \ldots, e_{i_k}),\ \forall i_1, \ldots, i_k \in \{1, \ldots, n\} \text{ tales que } 1 \leq i_1 < i_2 < \cdots < i_k \leq n,$$

Entonces $\omega = \eta$.

Dem. Aplicar la proposición anterior a $\omega - \eta \in \operatorname{Alt}^k(V)$. $\square$

La siguiente construcción nos va a permitir hallar una base de $\operatorname{Alt}^k(V)$, para todo $k \leq n$.

Definición 7.11.6. Si $\alpha_1, \ldots, \alpha_k \in V^*$, definimos $\alpha_1 \wedge \cdots \wedge \alpha_k : \underbrace{V \times \cdots \times V}_{k} \to \Bbbk$ mediante

$$(\alpha_1 \wedge \cdots \wedge \alpha_k)(v_1, \ldots, v_k) = \begin{vmatrix} \alpha_1(v_1) & \cdots & \alpha_1(v_k) \\ \vdots & & \vdots \\ \alpha_k(v_1) & \cdots & \alpha_k(v_k) \end{vmatrix}, \quad \forall v_1, \ldots, v_v \in V. \tag{7.30}$$

Claramente $\alpha_1 \wedge \cdots \wedge \alpha_k \in \operatorname{Alt}^k(V)$, para todo $\alpha_1, \ldots, \alpha_k \in V^*$ y para todo k.

Ejemplo 7.11.7. Si $\alpha, \beta \in V^*$ es $\alpha \wedge \beta \in \operatorname{Alt}^2(V)$ y

$$(\alpha \wedge \beta)(u, v) = \begin{vmatrix} \alpha(u) & \alpha(v) \\ \beta(u) & \beta(v) \end{vmatrix} = \alpha(u)\,\beta(v) - \alpha(v)\,\beta(u), \quad \forall u, v \in V.$$

Consideremos $V = \mathbb{k}^3$, $\{e_1, e_2, e_3\}$ la base canónica y $\{e_1^*, e_2^*, e_3^*\}$ la base dual en V^*, es decir

$$e_1^*(x, y, z) = x, \quad e_2^*(x, y, z) = y, \quad e_3^*(x, y, z) = z.$$

Entonces

$$(e_1^* \wedge e_2^*)\left((x, y, z), (x', y', z')\right) = \begin{vmatrix} x & x' \\ y & y' \end{vmatrix} = x\,y' - y\,x',$$
$$(e_1^* \wedge e_3^*)\left((x, y, z), (x', y', z')\right) = \begin{vmatrix} x & x' \\ z & z' \end{vmatrix} = x\,z' - z\,x',$$
$$(e_2^* \wedge e_3^*)\left((x, y, z), (x', y', z')\right) = \begin{vmatrix} y & y' \\ z & z' \end{vmatrix} = y\,z' - z\,y',$$

Observar que si ω es la 2-forma del ejemplo 7.11.1, entonces $\omega = e_1^* \wedge e_2^* + 2\, e_2^* \wedge e_3^*$.

Ejemplo 7.11.8. Si $\alpha, \beta, \gamma \in V^*$ es $\alpha \wedge \beta \wedge \gamma \in \operatorname{Alt}^3(V)$ y

$$(\alpha \wedge \beta \wedge \gamma)(u, v, w) = \begin{vmatrix} \alpha(u) & \alpha(v) & \alpha(w) \\ \beta(u) & \beta(v) & \beta(w) \\ \gamma(u) & \gamma(v) & \gamma(w) \end{vmatrix}$$

Si consideramos de nuevo $V = \mathbb{k}^3$, $\{e_1, e_2, e_3\}$ la base canónica y $\{e_1^*, e_2^*, e_3^*\}$ la base dual en V^*, entonces

$$(e_1^* \wedge e_2^* \wedge e_3^*)\left((x, y, z), (x', y', z'), (x'', y'', z'')\right) = \begin{vmatrix} x & x' & x'' \\ y & y' & y'' \\ z & z' & z'' \end{vmatrix}.$$

La prueba de la siguiente proposición es inmediata a partir de las propiedades de los determinantes.

Proposición 7.11.9. *Para todo $\alpha_1, \ldots, \alpha_k \in V^*$, vale:*

1. *Si existen $i \neq j$ tales que $\alpha_i = \alpha_j$, entonces $\alpha_1 \wedge \cdots \wedge \alpha_i \wedge \cdots \wedge \alpha_j \wedge \cdots \wedge \alpha_k = 0$.*
2. *El intercambio de dos elementos en $\alpha_1 \wedge \cdots \wedge \alpha_k$ genera un cambio de signo, es decir*

$$\alpha_1 \wedge \cdots \wedge \alpha_i \wedge \cdots \wedge \alpha_j \wedge \cdots \wedge \alpha_k = -\alpha_1 \wedge \cdots \wedge \alpha_j \wedge \cdots \wedge \alpha_i \wedge \cdots \wedge \alpha_k. \quad \square$$

Teorema 7.11.10. *Sea $\{e_1, \ldots, e_n\}$ una base de V y $\{e_1^*, \ldots, e_n^*\}$ la base dual en V^*, entonces*

$$\mathcal{B} = \left\{e_{i_1}^* \wedge \cdots \wedge e_{i_k}^* : \ 1 \leq i_1 < \cdots < i_k \leq n\right\}$$

es una base de $\operatorname{Alt}^k(V)$, para todo $k = 1, \ldots, n$. Además vale

$$\omega = \sum_{1 \leq i_1 < \cdots < i_k \leq n} \omega(e_{i_1}, \ldots, e_{i_k})\, e_{i_1}^* \wedge \cdots \wedge e_{i_k}^*, \quad \forall \omega \in \operatorname{Alt}^k(V). \tag{7.31}$$

Dem. Sea $k \in \{1, \ldots, n\}$ y $\omega \in \operatorname{Alt}^k(V)$. Probaremos que existen únicos $a_{i_1,\ldots,i_k} \in \mathbb{k}$ tales que

$$\omega = \sum_{1 \leq i_1 < \cdots < i_k \leq n} a_{i_1,\ldots,i_k}\, e_{i_1}^* \wedge \cdots \wedge e_{i_k}^* \tag{7.32}$$

y que $a_{i_1,\ldots,i_k} = \omega(e_{i_1}, \ldots, e_{i_k})$. Esto implica que $\mathcal{B}$ es una base de $\operatorname{Alt}^k(V)$ y que vale la igualdad en (7.31).

Aplicando el Corolario 7.11.5, vemos que se verifica la igualdad en (7.32) si y solo si

$$\omega(e_{j_1},\dots,e_{j_k}) = \sum_{1\le i_1<\cdots<i_k\le n} a_{i_1,\dots,i_k}\left(e_{i_1}^*\wedge\cdots\wedge e_{i_k}^*\right)(e_{j_1},\dots,e_{j_k}), \tag{7.33}$$

para todo $j_1,\dots,j_k\in\{1,\dots,n\}$ con $1\le j_1<\cdots<j_k\le n$. Considerando el lado derecho de (7.33), vemos que para cada sumando hay dos posibilidades

$$\left(e_{i_1}^*\wedge\cdots\wedge e_{i_k}^*\right)(e_{j_1},\dots,e_{j_k}) = \begin{vmatrix} e_{i_1}^*(e_{j_1}) & \cdots & e_{i_1}^*(e_{j_k}) \\ \vdots & & \vdots \\ e_{i_k}^*(e_{j_1}) & \cdots & e_{i_k}^*(e_{j_k}) \end{vmatrix} = \begin{cases} (A) & \text{si } \exists l:\ j_l\notin\{i_1,\dots,i_k\}, \\ (B) & \text{si } \{j_1,\dots,j_k\}=\{i_1,\dots,i_k\}. \end{cases} \tag{7.34}$$

En el caso (A), es $e_{i_1}^*(e_{j_l})=\cdots=e_{i_k}^*(e_{j_l})=0$, luego en el determinante de (7.34) la columna l-ésima es nula y por lo tanto $\left(e_{i_1}^*\wedge\cdots\wedge e_{i_k}^*\right)(e_{j_1},\dots,e_{j_k})=0$.

En el caso (B) es $\{j_1,\dots,j_k\}=\{i_1,\dots,i_k\}$ siendo $1\le j_1<\cdots<j_k\le n$ y $1\le i_1<\cdots<i_k\le n$, luego necesariamente es $i_1=j_1,\ \dots,\ i_k=j_k$ y

$$\left(e_{i_1}^*\wedge\cdots\wedge e_{i_k}^*\right)(e_{j_1},\dots,e_{j_k}) = \left(e_{i_1}^*\wedge\cdots\wedge e_{i_k}^*\right)(e_{i_1},\dots,e_{i_k}) = \begin{vmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{vmatrix} = 1.$$

Entonces en la suma de la ecuación (7.33) el único término no nulo corresponde al caso $i_1=j_1,\ \dots,\ i_k=j_k$ y en ese caso es $\left(e_{i_1}^*\wedge\cdots\wedge e_{i_k}^*\right)(e_{i_1},\dots,e_{i_k})=1$, luego la ecuación (7.33) equivale a

$$\omega(e_{j_1},\dots,e_{j_k}) = a_{j_1,\dots,j_k}, \quad \forall j_1,\dots,j_k\in\{1,\dots,n\},\ 1\le j_1<\cdots<j_k\le n$$

Esto concluye la prueba del teorema. $\square$

Recordemos el símbolo combinatorio $\binom{n}{k}=\frac{n!}{(n-k)!k!}$, que está definido para todo $n,k\in\mathbb{N}$, $n\ge k$.

Corolario 7.11.11. *Si* $\dim V=n$ *y* $k\le n$, *entonces* $\dim \operatorname{Alt}^k(V)=\binom{n}{k}$. $\square$

Ejemplo 7.11.12. Si $\dim V=3$ y $\{e_1,e_2,e_3\}$ es una base de V, entonces $\operatorname{Alt}^k(V)=\{0\},\ \forall k\ge 4$ y vale

- $\dim\operatorname{Alt}^0(V)=\binom{3}{0}=1,\ \{1\}$ es una base de $\operatorname{Alt}^0(V)=\Bbbk$;
- $\dim\operatorname{Alt}^1(V)=\binom{3}{1}=3,\ \{e_1^*,\ e_2^*,\ e_3^*\}$ es una base de $\operatorname{Alt}^1(V)=V^*$;
- $\dim\operatorname{Alt}^2(V)=\binom{3}{2}=3,\ \{e_1^*\wedge e_2^*,\ e_1^*\wedge e_3^*,\ e_2^*\wedge e_3^*\}$ es una base de $\operatorname{Alt}^2(V)$;
- $\dim\operatorname{Alt}^3(V)=\binom{3}{3}=1,\ \{e_1^*\wedge e_2^*\wedge e_3^*\}$ es una base de $\operatorname{Alt}^3(V)$.

En el caso de $V=\Bbbk^3$ estas bases fueron halladas explícitamente en los ejemplos 7.11.7 y 7.11.8.

Ejemplo 7.11.13. Si $V=\Bbbk^n$ y $\{e_1,\dots,e_n\}$ es la base canónica de $\Bbbk^n$, entonces $\dim\operatorname{Alt}^n(V)=\binom{n}{n}=1$ y $\{e_1^*\wedge\cdots\wedge e_n^*\}$ es una base de $\operatorname{Alt}^n(V)$. Operando obtenemos

$$\begin{aligned}(e_1^*\wedge\cdots\wedge e_n^*)((x_{11},\dots,x_{n1}),\dots,(x_{1n},\dots,x_{nn})) &= \begin{vmatrix} e_1^*(x_{11},\dots,x_{n1}) & \cdots & e_1^*(x_{1n},\dots,x_{nn}) \\ \vdots & \ddots & \vdots \\ e_n^*(x_{11},\dots,x_{n1}) & \cdots & e_n^*(x_{1n},\dots,x_{nn}) \end{vmatrix} \\ &= \begin{vmatrix} x_{11} & \cdots & x_{1n} \\ \vdots & \ddots & \vdots \\ x_{n1} & \cdots & x_{nn} \end{vmatrix}.\end{aligned}$$

Luego $e_1^*\wedge\cdots\wedge e_n^*=\det$ (el determinante) y $\{\det\}$ es una base de $\operatorname{Alt}^n(V)$.

Definición 7.11.14. Si $T : V \to W$ es una transformación lineal y $k \geq 1$, entonces el *pullback* asociado a T es la función $T^* : \mathrm{Alt}^k(W) \to \mathrm{Alt}^k(V)$ definida por

$$T^*(\omega)(v_1, \ldots, v_k) := \omega\big(T(v_1), \ldots, T(v_k)\big), \quad \forall\, \omega \in \mathrm{Alt}^k(W),\ v_1, \ldots, v_k \in V.$$

Es fácil probar que $T^* : \mathrm{Alt}^k(W) \to \mathrm{Alt}^k(V)$ es una transformación lineal.

Proposición 7.11.15. *Propiedades del pullback.*

1. $\mathrm{Id}^* = \mathrm{Id} : \mathrm{Alt}^k(V) \to \mathrm{Alt}^k(V)$.
2. *Si $S : U \to V$ y $T : V \to W$ son transformaciones lineales, entonces $(T \circ S)^* = S^* \circ T^*$.*
3. *Si $T : V \to W$ es un isomorfismo, entonces $T^* : \mathrm{Alt}^k(W) \to \mathrm{Alt}^k(V)$ es un isomorfismo y verifica $\big(T^{-1}\big)^* = (T^*)^{-1} : \mathrm{Alt}^k(V) \to \mathrm{Alt}^k(W)$.*

Dem. La prueba de las dos primeras afirmaciones es inmediata. Para la tercera, si $\omega \in \mathrm{Alt}^k(W)$ y $u_1 \ldots, u_k \in U$, es

$$\begin{aligned}\big((T \circ S)^*(\omega)\big)(u_1, \ldots, u_k) &= \omega\big((T \circ S)(u_1), \ldots, (T \circ S)(u_k)\big) = \omega\Big(T\big(S(u_1)\big), \ldots, T\big(S(u_k)\big)\Big)\\ &= \big(T^*(\omega)\big)\big(S(u_1)\big), \ldots, \big(S(u_k)\big) = \Big(S^*\big(T^*(\omega)\big)\Big)(u_1, \ldots, u_k)\\ &= \big(S^* \circ T^*\big)(\omega)(u_1, \ldots, u_k).\end{aligned}$$

La cuarta afirmación se deduce de la segunda y la tercera: al ser $T^{-1} \circ T = \mathrm{id}$, se deduce

$$\big(T^{-1} \circ T\big)^* = \mathrm{id}^* \quad \Rightarrow \quad T^* \circ \big(T^{-1}\big)^* = \mathrm{id}\,.$$

Análogamente de $T \circ T^{-1} = \mathrm{id}$ se obtiene $\big(T^{-1}\big)^* \circ T^* = \mathrm{id}$ y por lo tanto $\big(T^{-1}\big)^*$ es la inversa de T^*. $\square$

Proposición 7.11.16. *Si $T : V \to W$ es una transformación lineal y $\alpha_1, \ldots, \alpha_k \in W^*$, entonces*

$$T^*(\alpha_1 \wedge \cdots \wedge \alpha_k) = T^*(\alpha_1) \wedge \cdots \wedge T^*(\alpha_k) \in \mathrm{Alt}^k(V).$$

Dem.

$$T^*(\alpha_1 \wedge \cdots \wedge \alpha_k)(v_1, \ldots, v_k) = (\alpha_1 \wedge \cdots \wedge \alpha_k)(T(v_1), \ldots, T(v_k)) = \begin{vmatrix} \alpha_1(T(v_1)) & \cdots & \alpha_1(T(v_k)) \\ \vdots & & \vdots \\ \alpha_k(T(v_1)) & \cdots & \alpha_k(T(v_k)) \end{vmatrix}$$

$$= \begin{vmatrix} T^*(\alpha_1)(v_1) & \cdots & T^*(\alpha_1)(v_k) \\ \vdots & & \vdots \\ T^*(\alpha_k)(v_1) & \cdots & T^*(\alpha_k)(v_k) \end{vmatrix} = \left(T^*(\alpha_1) \wedge \cdots \wedge T^*(\alpha_k)\right)(v_1, \ldots, v_k). \quad \square$$

Proposición 7.11.17. *Si $\dim V = n$, $T : V \to V$ es una transformación lineal y $\mathcal{B} = \{e_1, \ldots, e_n\}$ es una base de V, entonces*

$$T^*(e_1^* \wedge \cdots \wedge e_n^*) = \det(T)\, e_1^* \wedge \cdots \wedge e_n^*.$$

Dem. Si ${}_{\mathcal{B}}[T]_{\mathcal{B}} = (a_{ij})$, es $e_k^*(T(e_i)) = e_k^*\Big(\sum_{j=1}^n a_{ji}\, e_j\Big) = a_{ki}$, $\forall\ i, k = 1, \ldots, n$. Luego aplicando la fórmula (7.31) a $T^*(e_1^* \wedge \cdots \wedge e_n^*)$, obtenemos

$$\begin{aligned} T^*(e_1^* \wedge \cdots \wedge e_n^*) &= T^*(e_1^* \wedge \cdots \wedge e_n^*)\,(e_1, \ldots, e_n)\, e_1^* \wedge \cdots \wedge e_n^* \\ &= (e_1^* \wedge \cdots \wedge e_n^*)\,\big(T(e_1), \ldots, T(e_n)\big)\, e_1^* \wedge \cdots \wedge e_n^* \\ &= \begin{vmatrix} e_1^*(T(e_1)) & \cdots & e_1^*(T(e_n)) \\ \vdots & & \vdots \\ e_n^*(T(e_1)) & \cdots & e_n^*(T(e_n)) \end{vmatrix} e_1^* \wedge \cdots \wedge e_n^* = \begin{vmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{vmatrix} e_1^* \wedge \cdots \wedge e_n^* \\ &= \det(T)\, e_1^* \wedge \cdots \wedge e_n^*. \quad \square \end{aligned}$$

Se puede definir un producto de formas, de manera tal que $\alpha_1 \wedge \cdots \wedge \alpha_k$ es el producto de $\alpha_1, \ldots, \alpha_k \in V^*$. En lo que sigue veremos una idea de cómo es esa construcción.

Proposición 7.11.18. *Para cada para $k, l = 1, 2, \ldots$, existe una única función*

$$\mathrm{Alt}^k(V) \times \mathrm{Alt}^l(V) \xrightarrow{\wedge} \mathrm{Alt}^{k+l}(V); \quad (\omega, \eta) \mapsto \omega \wedge \eta,$$

llamada el producto cuña *que verifica las siguientes propiedades:*

1. $\omega \wedge (\eta + \nu) = \omega \wedge \eta + \omega \wedge \nu$, $\forall\, \omega \in \mathrm{Alt}^k(V)$ *y* $\eta, \nu \in \mathrm{Alt}^l(V)$.
2. $(\eta + \nu) \wedge \omega = \eta \wedge \omega + \nu \wedge \omega$, $\forall\, \omega \in \mathrm{Alt}^k(V)$ *y* $\eta, \nu \in \mathrm{Alt}^l(V)$.
3. $\omega \wedge (a\, \eta) = (a\, \omega) \wedge \eta = a\, (\omega \wedge \eta)$, $\forall\, \omega \in \mathrm{Alt}^k(V)$, $\eta \in \mathrm{Alt}^l(V)$ *y* $a \in \Bbbk$.
4. $\omega \wedge (\eta \wedge \nu) = (\omega \wedge \eta) \wedge \nu$, $\forall\, \omega \in \mathrm{Alt}^k(V)$, $\eta \in \mathrm{Alt}^l(V)$ *y* $\nu \in \mathrm{Alt}^h(V)$.
5. *Si* $k \geq 1$ *y* $\alpha_1, \ldots, \alpha_k \in \mathrm{Alt}^1(V) = V^*$, *entonces*[7] $\alpha_1 \wedge \cdots \wedge \alpha_k \in \mathrm{Alt}^k(V)$ *está definido por* (7.30). $\square$

Explícitamente, si $\omega \in \mathrm{Alt}^k(V)$, $\eta \in \mathrm{Alt}^l(V)$ y $\mathcal{B} = \{e_1, \ldots, e_n\}$ es una base de V, y escribimos

$$\omega = \sum_{1 \leq i_1 < \cdots < i_k \leq n} a_{i_1, \ldots, i_k}\, e_{i_1}^* \wedge \cdots \wedge e_{i_k}^*, \quad \eta = \sum_{1 \leq j_1 < \cdots < j_l \leq n} b_{j_1, \ldots, j_l}\, e_{j_1}^* \wedge \cdots \wedge e_{j_l}^*,$$

entonces

$$\omega \wedge \eta = \sum_{\substack{1 \leq i_1 < \cdots < i_k \leq n \\ 1 \leq j_1 < \cdots < j_l \leq n}} a_{i_1, \ldots, i_k}\, b_{j_1, \ldots, j_l}\, e_{i_1}^* \wedge \cdots \wedge e_{i_k}^* \wedge e_{j_1}^* \wedge \cdots \wedge e_{j_l}^*. \tag{7.35}$$

Notar que los factores en los sumandos de (7.35) no están ordenados y pueden haber repetidos. La prueba de la proposición anterior no es simple y no la vamos a incluir aquí (hay una en el libro "Cálculo en variedades", de Michael Spivak).

Lo último que veremos es que el pullback conmuta con el producto cuña.

Proposición 7.11.19. *Si $T : V \to W$ es una transformación lineal y $\omega \in \mathrm{Alt}^k(V)$, $\eta \in \mathrm{Alt}^l(W)$, entonces*

$$T^*(\omega \wedge \eta) = T^*(\omega) \wedge T^*(\eta). \quad \square$$

Dem. Aplicar la fórmula (7.35) y la proposición 7.11.16. $\square$

[7] Notar que podemos escribir $\alpha_1 \wedge \cdots \wedge \alpha_k$ porque el producto cuña es asociativo.