(Contents of)

J. Rojo
Álgebra lineal. 2a edición corregida
Editorial AC, Madrid, 1991
563 pp.

Álgebra lineal

Jesús ROJO

Doctor en Matemáticas
Departamento de Matemática Aplicada
E.T.S. de Ingenieros Industriales de Valladolid

Jesús ROJO

Álgebra lineal

AMS Subject Classification: 15-01 Clasificación Decimal: 512.5

Jesús ROJO

Departamento de Matemática Aplicada
E.T.S. de Ingenieros Industriales
Paseo del Cauce, s/n
47011 VALLADOLID , España

Contenido

Capítulo 6

Diagonalización de endomorfismos y matrices.

Además de la diagonalización propiamente dicha, este capítulo estudia la reducción a la forma triangular y a la forma de Jordan. Este último aspecto es el que requerirá mayor esfuerzo, puesto que algunos razonamientos que se emplean son difíciles. El lector deberá poner también atención a la cuestión del cuerpo ($\mathbb{R}$ o $\mathbb{C}$) en el que se reduce una matriz, y a la diferencia entre valor propio y raíz del polinimio característico, estudiando con particular atención los apartados 6.2.11 y 6.2.12.

Nota. En este capítulo y los siguientes, $\mathbb{K}$ representa siempre $\mathbb{R}$ o $\mathbb{C}$. Por lo que a este capítulo se refiere, la limitación a estos casos tiene por objeto la utilización de los resultados que se exponen en 1.13.11 y 1.13.12, ya que, si bien no son sencillos de probar, al menos serán resultados conocidos para el lector de este libro. La extensión de algunos resultados de este capítulo a cuerpos diferentes requiere substituir $\mathbb{C}$ por un cuerpo 'algebraicamente cerrado', es decir, un cuerpo para el que todo polinomio no constante posea una raíz.

6.1 Subespacios invariantes. Vectores y valores propios.

6.1.1 Sea E un espacio vectorial sobre el cuerpo $\mathbb{K}$ y $f : E \to E$ un endomorfismo de E. Un *subespacio invariante* para f es un subespacio F de E tal que

$$f(F) \subset F\,,$$

o sea, tal que

$$x \in F \quad \Rightarrow \quad f(x) \in F\,.$$

Ejemplos sencillos de tales subespacios son $\{0\}$, E, $\operatorname{Im} f$ y $\operatorname{Ker} f$, que son invariantes independientemente de la naturaleza del endomorfismo f.

Si F es un subespacio invariante para f, podemos considerar la aplicación

$$\begin{aligned} F &\to F \\ x &\to f(x) \end{aligned}$$

que envía cada vector x de F a su imagen por f; esta aplicación es un endomorfismo de F, el 'endomorfismo inducido' por f en F. Es habitual representar también por f el endomorfismo inducido.

6.1.2 Sea E un e.v. y $f \in \mathcal{L}(E)$. Supongamos que

$$E = F_1 \oplus \cdots \oplus F_p\,,$$

donde $F_1, \ldots, F_p$ son subespacios invariantes para f, con dimensiones $n_1, \ldots, n_p$ diferentes de 0. Si $(a_1, \ldots, a_{n_1})$ es una base de F_1, $(a_{n_1+1}, \ldots, a_{n_1+n_2})$ una base de F_2, etc., y consideramos la base de E

$$(a_1, \ldots, a_{n_1}, a_{n_1+1}, \ldots, a_{n_1+n_2}, \ldots)$$

formada por la reunión de dichas bases (v. 2.5.7b), entonces la matriz de f en esta base es de la forma

$$[f, (a_i)] = \begin{bmatrix} \boxed{A_1} & & & \\ & \boxed{A_2} & & \\ & & \ddots & \\ & & & \boxed{A_p} \end{bmatrix}$$

o sea, lo que se suele llamar una matriz diagonal por bloques (véase el ejercicio 4 de la sección 3.2). La matriz A_1 es la matriz del endomorfismo inducido por f en F_1 representado en la base $(a_1, \ldots, a_{n_1})$, la matriz A_2 es la matriz del endomorfismo inducido por f en F_2 representado en la base $(a_{n_1+1}, \ldots, a_{n_1+n_2})$, etc.

6.1.3 Ejemplo. Consideremos el endomorfismo de $\mathbb{R}^3$ dado por

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

es decir, por la matriz

$$A = \begin{bmatrix} 4. & 2. & 1. \\ -2. & 0 & -1. \\ -1. & -1. & 1. \end{bmatrix}.$$

Consideremos los vectores

$$a_1 = (1, -1, 0)\,, \quad a_2 = (1, 0, -1) \quad \text{y} \quad a_3 = (-1, 1, 1)\,.$$

Es fácil comprobar que

$$\begin{aligned} f(a_1) &= (\;\;2, -2, \;\;0) = 2a_1 \\ f(a_2) &= (\;\;3, -1, -2) = a_1 + 2a_2 \\ f(a_3) &= (-1, \;\;1, \;\;1) = a_3\,. \end{aligned}$$

Los subespacios

$$F_1 = \langle a_1, a_2 \rangle \quad \text{y} \quad F_2 = \langle a_3 \rangle$$

son invariantes; (a_1, a_2) y (a_3) son bases de F_1 y F_2 y (a_1, a_2, a_3) es una base de $\mathbb{R}^3$ (luego $\mathbb{R}^3 = F_1 \oplus F_2$). La matriz de f en la base (a_1, a_2, a_3) es

$$A' = \left[\begin{array}{cc|c} 2. & 1. & 0 \\ 0 & 2. & 0 \\ \hline 0 & 0 & 1. \end{array}\right] .$$

6.1.4 El objeto de este capítulo es representar endomorfismos mediante matrices sencillas, entendiendo por tales, matrices con gran número de ceros. Las matrices diagonales por bloques son las más interesantes; lo son aún más cuando se trata de matrices pura y simplemente diagonales. Vamos a estudiar el caso en que un endomorfismo se puede representar por una matriz diagonal, dejando para más tarde las matrices diagonales por bloques, que constituyen un caso más general. El ejemplo siguiente nos muestra una situación en la que es posible representar un endomorfismo mediante una matriz diagonal.

6.1.5 Ejemplo. Consideremos el endomorfismo de $\mathbb{R}^3$ dado por

$$\begin{aligned} f(e_1) &= -\ e_1 - 6e_2 + 2e_3 \\ f(e_2) &= 3e_1 + 8e_2 - 2e_3 \\ f(e_3) &= 6e_1 + 16e_2 - 4e_3 \ , \end{aligned}$$

cuya matriz en la base canónica es

$$A = \begin{bmatrix} -1. & 3. & 6. \\ -6. & 8. & 16. \\ 2. & -2. & -4. \end{bmatrix} .$$

Para la base (a_1, a_2, a_3) de $\mathbb{R}^3$ formada por los vectores

$$a_1 = (0, -2, 1)\,, \quad a_2 = (3, -2, 2) \quad \text{y} \quad a_3 = (1, 1, 0)\,,$$

la matriz de f es

$$A' = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1. & 0 \\ 0 & 0 & 2. \end{bmatrix}$$

puesto que

$$\begin{aligned} f(a_1) &= 0 \\ f(a_2) &= a_2 \\ f(a_3) &= 2a_3\,. \end{aligned}$$

En esta base, f se representa por una matriz diagonal. La matriz A' pone en evidencia algunos aspectos de f que no eran fáciles de ver con la matriz A. Así, ahora es evidente que

$$\operatorname{rg} f = 2\,, \quad \operatorname{Im} f = \langle a_2, a_3 \rangle \quad \text{y} \quad \operatorname{Ker} f = \langle a_1 \rangle\,.$$

6.1.6 A la vista del ejemplo precedente, el problema se centra en dos aspectos:

— averiguar si existe alguna base en la que el endomorfismo se represente por una matriz diagonal, y

— encontrar, cuando exista, una de tales bases.

Nótese que los tres vectores de la base del ejemplo precedente verifican que

$$f(a_i) = \lambda_i \, a_i$$

para escalares λ_i que son, en ese caso,

$$\lambda_1 = 0\,, \quad \lambda_2 = 1 \quad \text{y} \quad \lambda_3 = 2\,.$$

Es a partir de esta propiedad como trataremos de resolver los dos problemas planteados.

6.1.7 DEFINICIÓN. Sea E un e.v. sobre $\mathbb{K}$ y $f \in \mathcal{L}(E)$; sea $x \in E$. Decimos que x es un *vector propio* de f cuando

(vcp) $(\exists \lambda \in \mathbb{K}) \quad f(x) = \lambda\, x\,.$

6.1.8 Ejemplo. En el ejemplo (6.1.5), los vectores a_1, a_2 y a_3 son vectores propios de f. Hay una infinidad de vectores propios de f; compruébese, por ejemplo, que lo son todos los vectores de la forma $\mu\, a_i$ cualesquiera que sean $\mu \in \mathbb{K}$ y $i = 1, 2, 3$.

6.1.9 El vector 0 es siempre un vector propio, puesto que

$$f(0) = 0 = \lambda 0$$

cualquiera que sea $\lambda \in \mathbb{K}$; todos los escalares hacen pues cierta la igualdad de **(vcp)** para el vector 0.

No ocurre lo mismo para vectores no nulos; si $x \neq 0$ y $f(x) = \lambda x$ y $f(x) = \mu x$, entonces $\lambda x - \mu x = 0$, luego $(\lambda - \mu)x = 0$ y, como $x \neq 0$, resulta que $\lambda = \mu$. Para cada vector propio x no nulo, existe exactamente un escalar λ tal que $f(x) = \lambda x$.

El método para encontrar los vectores propios de f es el estudio de los escalares que sirven para realizar las igualdades del tipo $f(x) = \lambda x$.

6.1.10 DEFINICIÓN. Sea E un e.v. sobre $\mathbb{K}$ y $f \in \mathcal{L}(E)$; sea $\lambda \in \mathbb{K}$. Decimos que λ es un *valor propio* de f cuando

(vlp) $(\exists x \in E,\, x \neq 0) \quad f(x) = \lambda\, x\,.$

La condición $x \neq 0$ que exigimos en **(vlp)** es importante puesto que, si no, cualquier elemento de $\mathbb{K}$ sería un valor propio.

6.1.11 Ejemplo. En el ejemplo (6.1.5) los escalares $0, 1$ y 2 son valores propios de f.

6.1.12 Es también útil otra forma de considerar los valores propios. Si $f \in \mathcal{L}(E)$ y $\lambda \in \mathbb{K}$, representamos por $V(\lambda)$ el subconjunto de E dado por

$$V(\lambda) = \{x \in E \mid f(x) = \lambda x\}\,.$$

Ya hemos visto que $0 \in V(\lambda)$. Es sencillo probar que $V(\lambda)$ es un subespacio vectorial de E; es justamente el subespacio

$$\operatorname{Ker}(f - \lambda\, id_E)\,,$$

núcleo del endomorfismo $f - \lambda\, id_E$ (v. 2.1.11). Decimos que $V(\lambda)$ es el *subespacio propio* de f asociado a λ.

Podemos expresar **(vlp)** es las siguientes formas equivalentes: λ es un valor propio de f si y sólo si

(vlp1) $V(\lambda) \neq \{0\}\,,$

y también si y sólo si

(vlp2) el endomorfismo $f - \lambda\, id$ no es inyectivo.

Si $\lambda \neq \mu$, entonces

$$V(\lambda) \cap V(\mu) = \{0\}\,,$$

como se ve utilizando los argumentos de (6.1.9).

6.1.13 Ejemplo. En el ejemplo (6.1.5) se tiene

$$\begin{aligned} V(0) &= \operatorname{Ker} f = \langle a_1 \rangle \\ V(1) &= \operatorname{Ker}(f - id) = \langle a_2 \rangle \\ V(2) &= \operatorname{Ker}(f - 2\, id) = \langle a_3 \rangle\,, \end{aligned}$$

como se comprueba con facilidad. Se puede comprobar también que, si λ es distinto de $0, 1$ y 2, entonces

$$V(\lambda) = \operatorname{Ker}(f - \lambda\, id) = \{0\}\,,$$

lo que indica que no hay más valores propios que estos tres.

6.1.14 PROPOSICIÓN

Sea $f \in \mathcal{L}(E)$ y sean $\lambda_1, \ldots, \lambda_p$ valores propios de f, todos distintos entre sí. Si $x_1 \in V(\lambda_1), \ldots, x_p \in V(\lambda_p)$, y $x_1 \neq 0, \ldots, x_p \neq 0$, entonces $(x_1, \ldots, x_p)$ es un sistema libre.

Lo probaremos por inducción sobre el número p. El resultado es evidente cuando $p = 1$. Supongamos (hipótesis de recurrencia) que el resultado es cierto para $p-1$; sean entonces $x_1, \ldots, x_p$ p vectores que cumplen las hipótesis del enunciado. Si

$$\alpha_1 x_1 + \alpha_2 x_2 + \cdots + \alpha_p x_p = 0\,,$$

tenemos, por una parte, que

$$f(\alpha_1 x_1 + \alpha_2 x_2 + \cdots + \alpha_p x_p) = 0\,,$$

o sea,

$$\alpha_1 f(x_1) + \alpha_2 f(x_2) + \cdots + \alpha_p f(x_p) = 0\,,$$

luego

$$\alpha_1 \lambda_1 x_1 + \alpha_2 \lambda_2 x_2 + \cdots + \alpha_p \lambda_p x_p = 0\,.$$

Por otra parte, $\lambda_1 0 = 0$, o sea,

$$\lambda_1(\alpha_1 x_1 + \alpha_2 x_2 + \cdots + \alpha_p x_p) = 0\,,$$

luego

$$\lambda_1 \alpha_1 x_1 + \lambda_1 \alpha_2 x_2 + \cdots + \lambda_1 \alpha_p x_p = 0\,.$$

Restando ambas expresiones obtenemos

$$(\lambda_2 - \lambda_1)\alpha_2 x_2 + \cdots + (\lambda_p - \lambda_1)\alpha_p x_p = 0\,,$$

y como, por la hipótesis de recurrencia, el sistema $(x_2, \ldots, x_p)$ de $p-1$ vectores es libre, resulta que

$$(\lambda_2 - \lambda_1)\alpha_2 = \cdots = (\lambda_p - \lambda_1)\alpha_p = 0\,,$$

luego que

$$\alpha_2 = \cdots = \alpha_p = 0$$

ya que todos los λ_i son distintos. Por fin, resulta que

$$\alpha_1 x_1 = 0$$

y que $\alpha_1 = 0$. El sistema $(x_1, x_2, \ldots, x_p)$ es libre; el resultado es también cierto para p, y esto concluye la demostración.

6.1.15 Si E es un e.v. de dimensión n y $f \in \mathcal{L}(E)$, resulta inmediatamente de la proposición anterior que f posee a lo sumo n valores propios.

6.1.16 PROPOSICIÓN

Sea $f \in \mathcal{L}(E)$ y sean $\lambda_1, \ldots, \lambda_p$ valores propios de f, todos distintos entre sí. Entonces

$$V(\lambda_1) + \cdots + V(\lambda_p) = V(\lambda_1) \oplus \cdots \oplus V(\lambda_p) .$$

Si

$$0 = x_1 + \cdots + x_p$$

con $x_1 \in V(\lambda_1), \ldots, x_p \in V(\lambda_p)$, entonces $x_1 = \cdots = x_p = 0$, pues, en caso contrario, llamando $x_{i_1}, \ldots, x_{i_r}$ a aquellos vectores que no fuesen nulos, resultaría que

$$0 = x_{i_1} + \cdots + x_{i_r} ;$$

esto es imposible, ya que el sistema $(x_{i_1}, \ldots, x_{i_r})$ es libre, como consecuencia de la proposición anterior.

Sea entonces x un vector de $V(\lambda_1) + \cdots + V(\lambda_p)$ y supongamos que

$$x = x_1 + \cdots + x_p \quad \text{y} \quad x = x_1' + \cdots + x_p' ,$$

con $x_1, x_1' \in V(\lambda_1), \ldots, x_p, x_p' \in V(\lambda_p)$; se tiene

$$0 = (x_1 - x_1') + \cdots + (x_p - x_p')$$

y, como acabamos de ver, esto significa que

$$x_1 = x_1', \ldots, x_p = x_p' .$$

El subespacio suma es, pues, suma directa (v. 2.3.22).

6.1.17 TEOREMA

Sea E un e.v. de dimensión $n \neq 0$, $(a_1, \ldots, a_n)$ una base de E, $f \in \mathcal{L}(E)$ y $A = [f, (a_i)]$. Sea $\lambda \in \mathbb{K}$; las proposiciones siguientes son equivalentes:

(i) λ es un valor propio de f ;
(ii) el endomorfismo $f - \lambda\, id$ no es inversible ;
(iii) $\det(A - \lambda I_n) = 0$.

Decir que $f - \lambda\, id$ no es inversible equivale a decir que no es inyectivo, puesto que E es de dimensión finita (v. 3.1.24); se tiene así la equivalencia de (i) y (ii) (v. 6.1.12).

La equivalencia de (ii) y (iii) proviene de que $A - \lambda I_n$ es la matriz de $f - \lambda\, id$ en la base $(a_1, \ldots, a_n)$.

6.1.18 DEFINICIÓN. Si A es una matriz cuadrada con elementos en $\mathbb{K}$, los escalares $\lambda \in \mathbb{K}$ tales que

$$\det(A - \lambda I_n) = 0$$

reciben el nombre de *valores propios* de la matriz A.

Lo que demuestra el teorema precedente es que los valores propios de un endomorfismo f de un espacio vectorial de dimensión finita son los valores propios de cualquier matriz asociada a f en una base del espacio.

Dos endomorfismos del e.v. E que se representen (en bases diferentes) por la misma matriz poseen entonces los mismos valores propios.

Otra consecuencia aún más importante es que dos matrices semejantes poseen los mismos valores propios (v. 3.5.29).

6.1.19 Ejemplo. La matriz

$$A = \begin{bmatrix} -1. & 3. & 6. \\ -6. & 8. & 16. \\ 2. & -2. & -4. \end{bmatrix}$$

asociada al endomorfismo del ejemplo (6.1.5) posee como valores propios $0, 1$ y 2. Compruébese que éstos son los únicos valores $\lambda \in \mathbb{R}$ para los que el determinante de

$$A - \lambda I_3 = \begin{bmatrix} -1. - \lambda & 3. & 6. \\ -6. & 8. - \lambda & 16. \\ 2. & -2. & -4. - \lambda \end{bmatrix}$$

es nulo.

6.1.20 Sea $A \in M_{\mathbb{K}}(n)$. Los valores propios de A son los del endomorfismo $f \in \mathcal{L}(\mathbb{K}^n)$ dado por A (v. 3.5.1), es decir, los escalares $\lambda \in \mathbb{K}$ para los que existe un vector $x \in \mathbb{K}^n$, $x \neq 0$, tal que

$$A\,x = \lambda\,x\,.$$

Se suele hablar de *vectores propios* de A para referirse a los vectores $x \in \mathbb{K}^n$ tales que

$$A\,x = \lambda\,x$$

para algún $\lambda \in \mathbb{K}$, o sea, a los vectores propios del endomorfismo dado por A.

Si $\lambda \in \mathbb{K}$, el subespacio

$$V(\lambda) = \{x \in \mathbb{K}^n \mid Ax = \lambda x\}$$

de $\mathbb{K}^n$ se suele denominar *subespacio propio* de A correspondiente a λ.

6.1.21 Ejercicios.

1 Sea E un e.v. de dimensión $n \neq 0$, $f \in \mathcal{L}(E)$ y F un subespacio, no trivial, invariante para f. Se considera una base $(a_1, \dots, a_p)$ de F y se la extiende a una base $(a_1, \dots, a_p, a_{p+1}, \dots, a_n)$ de E. Demuéstrese que la matriz $[f, (a_i)]$ es triangular por bloques (véase el ejercicio 4 de la sección 3.2).

2 Para el endomorfismo

$$D : \mathbb{R}[X] \to \mathbb{R}[X]$$

(véase el ejercicio 9 de la sección 2.1) de derivación formal de polinomios, demuéstrese que los subespacios $\mathbb{R}_n[X]$ (v. 2.2.6) son invariantes.

3 Siendo F_1 y F_2 dos subespacios invariantes para un endomorfismo $f \in \mathcal{L}(E)$, pruébese que los subespacios $F_1 + F_2$ y $F_1 \cap F_2$ son también invariantes.

4 Para el endomorfismo $f \in \mathcal{L}(\mathbb{R}^2)$ dado por

$$f(x, y) = (x, 0)\,,$$

¿cuáles son los subespacios invariantes?

5 Siendo $f, g \in \mathcal{L}(E)$ dos endomorfismos del e.v. E tales que $f \circ g = g \circ f$, pruébese que $\operatorname{Im} g$ y $\operatorname{Ker} g$ son subespacios invariantes para f.

6 Sea E un e.v. Una *proyección* de E es un endomorfismo $p \in \mathcal{L}(E)$ tal que

$$p^2 = p$$

(recuérdese que $p^2 = p \circ p$). Demuéstrese que, si p es una proyección de E, entonces:

a) $\operatorname{Im} p = \{x \in E \mid p(x) = x\}$,

b) para todo $x \in E$, $x = p(x) + z$ con $z \in \operatorname{Ker} p$,

c) $E = \operatorname{Im} p \oplus \operatorname{Ker} p$,

d) si E es de dimensión $n \neq 0$, existe una base de E en la que la la matriz de p es diagonal.

7 Sea E un e.v. y F_1 y F_2 dos subespacios de E tales que

$$E = F_1 \oplus F_2\,;$$

definimos dos aplicaciones

$$p, q : E \to E$$

de la manera siguiente: si $x \in E$, se considera la única descomposición $x = x_1 + x_2$, con $x_1 \in F_1$ y $x_2 \in F_2$, y se pone

$$p(x) = x_1 \quad \text{y} \quad q(x) = x_2\,,$$

Decimos que p es la 'proyección de E sobre F_1 paralela a F_2' y que q es la 'proyección de E sobre F_2 paralela a F_1'.

a) Demuéstrese que p y q son endomorfismos de E.

b) Demuéstrese que $p^2 = p$ y $q^2 = q$ (esto es, que p y q son proyecciones), que $p + q = id_E$ y que $p \circ q = q \circ p = 0$.

c) Demuéstrese que $\operatorname{Im} p = F_1$, $\operatorname{Ker} p = F_2$, $\operatorname{Im} q = F_2$ y $\operatorname{Ker} q = F_1$.

d) Demuéstrese que F_1 y F_2 son subespacios invariantes para p y q.

e) Siendo E de dimensión finita y $(a_1, \ldots, a_r, a_{r+1}, \ldots, a_n)$ una base formada por reunión de dos bases $(a_1, \ldots, a_r)$ y $(a_{r+1}, \ldots, a_n)$ de F_1 y F_2, calcúlense las matrices de p y q en dicha base de E.

f) Consideremos el isomorfismo

$$\bar{q} : E/F_1 \to F_2$$

correspondiente a la descomposición canónica de q (v. 3.1.17) y el isomorfismo

$$\begin{aligned} g : F_2 &\to E/F_1 \\ x &\to \dot{x} \end{aligned}$$

que vimos en (2.3.18). Demuéstrese que $\bar{q}$ y g son inversos uno del otro.

g) Para la suma directa

$$\mathbb{R}^2 = F_1 \oplus F_2 \,,$$

donde $F_1 = \{(x, y) \mid x = y\}$ y $F_2 = \{(x, y) \mid y = 0\}$, y las correspondientes proyecciones p y q, calcúlense las imágenes $p(x, y)$ y $q(x, y)$ de un vector $(x, y) \in \mathbb{R}^2$.

h) Sean E y E' dos e.v., $f \in \mathcal{L}(E, E')$ una aplicación lineal y F un subespacio de E tal que

$$E = \operatorname{Ker} f \oplus F \,.$$

Se consideran las proyecciones p y q de E, correspondientes a esta suma directa. Demuéstrese que $f \circ p = 0$ y $f \circ q = f$.

8 Sea p una proyección de E (véase el ejercicio 6 de esta sección). Demuéstrese que p es justamente la proyección sobre $\operatorname{Im} p$ paralela a $\operatorname{Ker} p$ (véase también el ejercicio precedente).

9 Sea E un e.v. de dimensión finita y F_1 y F_2 subespacios de E tales que

$$E = F_1 \oplus F_2 \,;$$

sean p y q las proyecciones correspondientes (véase el ejercicio 7 de esta sección). Se tiene que

$$E^* = F_1^\perp \oplus F_2^\perp$$

(véase el ejercicio 10 de la sección 3.6); consideremos también las proyecciones p' y q' (de E^*) correspondientes a esta suma directa. Demuéstrese que, para todo $x^* \in E^*$,

$$p^t(x^*) \in F_2^\perp \,, \quad q^t(x^*) \in F_1^\perp \quad \text{y} \quad p^t(x^*) + q^t(x^*) = x^* \,.$$

Demuéstrese que $q' = p^t$ y $p' = q^t$.

10 Sea E un e.v. y F_1 y F_2 subespacios de E tales que

$$E = F_1 \oplus F_2 \,,$$

y sean p y q las correspondientes proyecciones (véase el ejercicio 7 de esta sección). Siendo $f \in \mathcal{L}(E)$, demuéstrese la equivalencia de:

(i) F_1 y F_2 son invariantes para f, y

(ii) $f \circ p = p \circ f \quad \text{y} \quad f \circ q = q \circ f$.

11 Siendo $f \in \mathcal{L}(E)$ y $x \in E$, pruébese la equivalencia de:

(i) x es un vector propio, y

(ii) el subespacio $\langle x \rangle$ es invariante.

12 Sea $f \in \mathcal{L}(E)$, $\lambda \in \mathbb{K}$ y $V(\lambda)$ el correspondiente subespacio propio. Demuéstrese que $V(\lambda)$ es un subespacio invariante para f. Demuéstrese que el endomorfismo inducido por f en $V(\lambda)$ es una homotecia de razón λ (véase el ejercicio 5 de la sección 3.4).

13 Sea $f \in \mathcal{L}(E)$ y sean $\lambda_1, \ldots, \lambda_p$ valores propios de f, todos distintos entre sí. Demuéstrese que el subespacio $V(\lambda_1) \oplus \cdots \oplus V(\lambda_p)$ es invariante.

14 En este ejercicio, f representa un endomorfismo de un e.v. E sobre $\mathbb{K}$ y A representa una matriz $n \times n$ con elementos en $\mathbb{K}$. Recuérdense las definiciones de f^r (v. 3.4.8) y A^r (v. 3.4.29) cuando $r = 0, 1, 2, \ldots$; recuérdense también las definiciones de $p(f)$ y $p(A)$ cuando $p(X) \in \mathbb{K}[X]$ es un polinomio con coeficientes en $\mathbb{K}$. Demuéstrense los resultados que siguen:

a) Si $\lambda \in \mathbb{K}$ es un valor propio de f (de A), entonces λ^2 es un valor propio de f^2 (de A^2).

b) Si $\lambda \in \mathbb{K}$ es un valor propio de f (de A), entonces λ^r es un valor propio de f^r (de A^r). (Utilícese el principio de inducción para r.)

c) Si $\lambda \in \mathbb{K}$ es un valor propio de f (de A) y $p(X) \in \mathbb{K}[X]$, entonces $p(\lambda)$ es un valor propio de $p(f)$ (de $p(A)$). (Utilícese el principio de inducción para el grado del polinomio.)

d) Sea $p(X) \in \mathbb{K}[X]$ un polinomio que anula f (que anula A), o sea, tal que $p(f) = 0$ ($p(A) = 0$). Si $\lambda \in \mathbb{K}$ es un valor propio de f (de A), entonces λ es una raíz del polinomio $p(X)$.

e) Si $f^2 = id$ ($A^2 = I_n$) y $\lambda \in \mathbb{K}$ es un valor propio de f (de A), entonces $\lambda = 1$ o $\lambda = -1$.

f) Si $f^2 = 0$ ($A^2 = 0$) y $\lambda \in \mathbb{K}$ es un valor propio de f (de A), entonces $\lambda = 0$.

15 Sea $A \in M_{\mathbb{K}}(n)$; vamos a considererar el siguiente endomorfismo del espacio vectorial $M_{\mathbb{K}}(n)$:

$$\begin{aligned} f : M_{\mathbb{K}}(n) &\to M_{\mathbb{K}}(n) \\ B &\to AB\,. \end{aligned}$$

Pruébese que los valores propios de A y los de f coinciden.

16 En el e.v. $S_{\mathbb{K}}$ ($S_{\mathbb{R}}$ o $S_{\mathbb{C}}$), formado por las sucesiones de números reales o complejos según el caso, se consideran los endomorfismos $L, R \in \mathcal{L}(S_{\mathbb{K}})$ definidos como sigue. L es el 'desplazamiento a la izquierda' de las sucesiones, esto es,

$$L(x_1, x_2, x_3, \ldots) = (x_2, x_3, x_4, \ldots)\,.$$

R es el 'desplazamiento a la derecha' de las sucesiones, esto es,

$$R(x_1, x_2, x_3, \ldots) = (0, x_1, x_2, \ldots)\,.$$

Por otra parte se consideran los subespacios

$$c_{00}\,, \quad l^1\,, \quad c_0\,, \quad c \quad \text{y} \quad l^\infty$$

de $S_{\mathbb{K}}$, definidos en (2.2.7) y en el ejercicio 11 de la sección 2.2.

Pruébese que todos estos subespacios son invariantes para L y para R.

Esto permite considerara L y R como endomorfismos de cualquiera de estos espacios (v. 6.1.1).

Calcúlense los valores propios de L y de R cuando se consideran como endomorfismos de $c_{00}, l^1, c_0, c, l^\infty$ y $S_{\mathbb{K}}$. Para cada λ que sea valor propio en cada caso, descríbase el subespacio propio $V(\lambda)$ y calcúlese la dimensión de $V(\lambda)$.

6.2 Polinomio característico.

6.2.1 Sea $A \in M_{\mathbb{K}}(n)$ una matriz cuadrada. Es sencillo averiguar cuando un elemento $\lambda \in \mathbb{K}$ es un valor propio de A; basta comprobar que

$$\det(A - \lambda I_n) = 0\,.$$

Parece, sin embargo, más dificil encontrar todos los valores propios de A, puesto que no podemos efectuar dicha comprobación para todos los elementos de $\mathbb{K}$, si éste posee una cantidad infinita de elementos (lo que es normal). Lo que se hace entonces es calcular $\det(A - \lambda I_n)$ sin precisar el valor de λ, y tratar de ver a continuación para qué valores de λ se anula.

6.2.2 Ejemplo. Para la matriz de elementos reales

$$A = \begin{bmatrix} 2. & 2. \\ 1. & 1. \end{bmatrix}$$

tenemos

$$A - \lambda I_n = \begin{bmatrix} 2. - \lambda & 2. \\ 1. & 1. - \lambda \end{bmatrix}$$

y

$$\det(A - \lambda I_n) = (2 - \lambda)(1 - \lambda) - 2 = \lambda^2 - 3\lambda\,,$$

expresión que se anula para $\lambda = 0$ y $\lambda = 3$; los valores propios de A son entonces 0 y 3.

6.2.3 Sea $A = [\alpha_i^j] \in M_{\mathbb{K}}(n)$ una matriz cuadrada. Siendo X un símbolo, consideramos la matriz

$$A - X\,I_n = \begin{bmatrix} \alpha_1^1 - X & \alpha_2^1 & \alpha_3^1 & \cdots & \alpha_n^1 \\ \alpha_1^2 & \alpha_2^2 - X & \alpha_3^2 & \cdots & \alpha_n^2 \\ \alpha_1^3 & \alpha_2^3 & \alpha_3^3 - X & \cdots & \alpha_n^3 \\ \vdots & \vdots & \vdots & & \vdots \\ \alpha_1^n & \alpha_2^n & \alpha_3^n & \cdots & \alpha_n^n - X \end{bmatrix}$$

cuyos elementos son polinomios en la indeterminada X con coeficientes en $\mathbb{K}$ (v. 1.13.1). Aun cuando los elementos diagonales de $A - X\,I_n$ no pertenecen al cuerpo $\mathbb{K}$, calculamos el determinante de esta matriz operando con los polinomios $\alpha_i^i - X$ como si se tratase de elementos de $\mathbb{K}$. (Una explicación más formalizada sobrepasa el propósito de este libro.)

$\det A - X\,I_n$ es un polinomio en X, con coeficientes en $\mathbb{K}$, y de grado n.

6.2.4 DEFINICIÓN. Sea $A \in M_{\mathbb{K}}(n)$ una matriz cuadrada. Llamamos *polinomio característico* de A y representamos por $p_A(X)$ al polinomio

$$p_A(X) = \det A - X\, I_n\,.$$

Si $\lambda \in \mathbb{K}$, tenemos que

$$p_A(\lambda) = \det A - \lambda\, I_n\,.$$

6.2.5 Ejemplo. Si

$$A = \begin{bmatrix} 2. & 2. \\ 1. & 1. \end{bmatrix}$$

entonces

$$p_A(X) = X^2 - 3X\,,$$

como vimos en (6.2.2).

Si

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

entonces

$$\begin{aligned} p_A(X) &= \begin{vmatrix} -2.-X & 4. & 5. \\ -3. & 5.-X & 5. \\ 0 & 0 & 1.-X \end{vmatrix} \\ &= (-2-X)(5-X)(1-X) + 12(1-X) \\ &= -X^3 + 4X^2 - 5X + 2\,, \end{aligned}$$

6.2.6 Algunas consideraciones sencillas nos van a permitir calcular ciertos términos de $p_A(X)$. Los términos de grado n y $n-1$ provienen exclusivamente del producto de los elementos diagonales de $A - X\, I_n$, o sea, de

$$(\alpha_1^1 - X)(\alpha_2^2 - X)\cdots(\alpha_n^n - X)\,.$$

Esto se debe a que en el resto de los $n!$ productos de la fórmula **(3)** de (4.2.4) interviene algún α_i^j con $i \neq j$, luego no intervienen $\alpha_i^i - X$ y $\alpha_j^j - X$; el resto de los productos son entonces polinomios de grado menor o igual que $n-2$.

Como

$$\begin{aligned} &(\alpha_1^1 - X)(\alpha_2^2 - X)\cdots(\alpha_n^n - X) = \\ &= (-1)^n X^n + (-1)^{n-1}(\alpha_1^1 + \alpha_2^2 + \cdots + \alpha_n^n)X^{n-1} + \cdots\,, \end{aligned}$$

resulta que el término de grado n de $p_A(X)$ es

$$(-1)^n X^n\,,$$

y el de grado $n-1$ es

$$(-1)^{n-1}(\alpha_1^1 + \alpha_2^2 + \cdots + \alpha_n^n)X^{n-1}\,,$$

o bien

$$(-1)^{n-1}\operatorname{tr} A\, X^{n-1}\,,$$

donde el escalar

$$\operatorname{tr} A = \alpha_1^1 + \alpha_2^2 + \cdots + \alpha_n^n\,,$$

suma de los elementos diagonales de A, es lo que llamamos *traza* de la matriz A.

Por otra parte, el término independiente de $p_A(X)$ vale

$$p_A(0) = \det(A - 0\, I_n) = \det A\,;$$

$p_A(X)$ es entonces de la forma

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

6.2.7 PROPOSICIÓN

Si $A, A' \in M_{\mathbb{K}}(n)$ son dos matrices semejantes, entonces

$$p_A(X) = p_{A'}(X)$$

y, en particular,

$$\operatorname{tr} A = \operatorname{tr} A'\,.$$

Existe una matriz P, $n \times n$ e inversible, tal que

$$A' = P^{-1} A\, P\,,$$

luego

$$A' - X\, I_n = P^{-1} A\, P - X\, I_n = P^{-1} A\, P - X\, I_n (P^{-1} P)$$

y, como $X\, I_n$ conmuta con cualquier matriz de $M_{\mathbb{K}}(n)$ (v. 3.4.28),

$$A' - X\, I_n = P^{-1} A\, P - P^{-1} X\, I_n\, P = P^{-1}(A - X\, I_n) P\,.$$

Entonces

$$p_{A'}(X) = \det(A' - X\, I_n) = \det(P^{-1}) p_A(X) \det P = p_A(X)\,.$$

6.2.8 Sea E un e.v. de dimensión $n \neq 0$ y $f \in \mathcal{L}(E)$. Llamamos *polinomio característico* de f y *traza* de f al polinomio característico y la traza de una matriz cualquiera asociada a f; los representamos por $p_f(X)$ y $\operatorname{tr} f$. La proposición precedente nos garantiza que esta definición no depende de la matriz asociada a f que se elija.

Recordando la definición de $\det f$ (v. 4.2.21) se tiene que

$$p_f(X) = (-1)^n X^n + (-1)^{n-1}\operatorname{tr} f\, X^{n-1} + \cdots + \det f\,.$$

6.2.9 TEOREMA

a) Sea $A \in M_{\mathbb{K}}(n)$ y $\lambda \in \mathbb{K}$ ($\mathbb{R}$ o $\mathbb{C}$). Entonces λ es un valor propio de A si y sólo si es una raíz de $p_A(X)$.

b) Sea E un e.v. de dimensión $n \neq 0$ sobre $\mathbb{K}$ ($\mathbb{R}$ o $\mathbb{C}$), $f \in \mathcal{L}(E)$ y $\lambda \in \mathbb{K}$. Entonces λ es un valor propio de f si y sólo si es una raíz de $p_f(X)$.

a) λ es un valor propio de A si y sólo si

$$p_A(\lambda) = \det(A - \lambda\, I_n) = 0\,,$$

o sea, si y sólo si λ es raíz de $p_A(X)$.

b) es una consecuencia de a), pues los valores propios de f son los de una matriz cualquiera asociada a f.

6.2.10 Se deduce del teorema precedente que una matriz $n \times n$ posee a lo sumo n valores propios, y que un endomorfismo de un e.v. de dimensión n posee a lo sumo n valores propios.

6.2.11 Cuando $\mathbb{K} = \mathbb{R}$, es importante suponer en el teorema precedente que $\lambda \in \mathbb{R}$. En el caso real pueden existir raíces de $p_f(X)$ que no sean elementos de $\mathbb{R}$; dichas raíces no serán, sin embargo, valores propios de f.

6.2.12 Ejemplo. El endomorfismo f de $\mathbb{R}^2$ dado por la matriz

$$A = \begin{bmatrix} 0 & -1. \\ 1. & 0 \end{bmatrix}$$

tiene como polinomio característico

$$p_f(X) = X^2 + 1\,,$$

polinomio cuyas raíces son los números complejos $+i$ y $-i$, que no son valores propios de f.

Sin embargo, el endomorfismo g de $\mathbb{C}^2$ dado por la misma matriz posee $+i$ y $-i$ como valores propios.

Obsérvese que

$$\begin{bmatrix} 0 & -1. \\ 1. & 0 \end{bmatrix} \begin{bmatrix} 1 \\ -i \end{bmatrix} = i \begin{bmatrix} 1 \\ -i \end{bmatrix}, \quad \text{o sea,} \quad g(1,-i) = i\,(1,-i)$$

y que

$$\begin{bmatrix} 0 & -1. \\ 1. & 0 \end{bmatrix} \begin{bmatrix} 1 \\ i \end{bmatrix} = -i \begin{bmatrix} 1 \\ i \end{bmatrix}, \quad \text{o sea,} \quad g(1,i) = -i\,(1,i)\,,$$

luego que $(1,-i)$ y $(1,i)$ son vectores (de $\mathbb{C}^2$) propios de g. Compruébese que no existe ningún vector (de $\mathbb{R}^2$) propio de f, excepción hecha del vector $(0,0)$.

6.2.13 PROPOSICIÓN

Sea E un e.v. sobre $\mathbb{K}$ de dimensión $n \neq 0$, $f \in \mathcal{L}(E)$; sea $\lambda \in \mathbb{K}$ un valor propio de f. Hemos visto que λ es raíz de $p_f(X)$.

Si k es el orden de multiplicidad de λ como raíz (v. 1.13.9), entonces

$$1 \leq \dim V(\lambda) \leq k\,.$$

Como λ es un valor propio de f, entonces $V(\lambda) \neq \{0\}$, luego $\dim V(\lambda) \geq 1$.

Representemos por h la dimensión de $V(\lambda)$; vamos a probar que $h \leq k$. Para ello completamos una base $(a_1, \dots, a_h)$ de $V(\lambda)$ hasta formar una base $(a_1, \dots, a_h, a_{h+1}, \dots, a_n)$ de E. La matriz $A = [f, (a_i)]$ es de la forma

$$\begin{array}{cc} \begin{array}{cc} \quad h \quad & \quad n-h \end{array} & \\ A = \left[\begin{array}{ccc|c} \lambda & & & \\ & \ddots & & A' \\ & & \lambda & \\ \hline & 0 & & A'' \end{array}\right] & \begin{array}{c} h \\ \\ n-h \end{array} \end{array}$$

luego

$$A - X\,I_n = \left[\begin{array}{ccc|c} \lambda - X & & & \\ & \ddots & & A' \\ & & \lambda - X & \\ \hline & 0 & & A'' - X\,I_{n-h} \end{array}\right],$$

y entonces (v. 4.2.14)

$$\begin{aligned} p_f(X) &= p_A(X) \\ &= \det(A - X\,I_n) \\ &= \det \begin{bmatrix} \lambda - X & & \\ & \ddots & \\ & & \lambda - X \end{bmatrix} \det(A'' - X\,I_{n-h}) \\ &= (\lambda - X)^h q(X)\,, \end{aligned}$$

donde $q(X)$ es un polinomio de grado $n-h$ (el polinomio característico de A''). Resulta así que λ es raíz de $p_f(X)$ con multiplicidad $\geq h$, o sea, $k \geq h$.

6.2.14 Ejemplo. Es posible, sin embargo, que

$$\dim V(\lambda) < k\,.$$

Así, si f es el endomorfismo de $\mathbb{R}^2$ dado por la matriz

$$A = \begin{bmatrix} 1. & 1. \\ 0 & 1. \end{bmatrix},$$

1 es un valor propio de f y es una raíz de multiplicidad 2 del polinomio

$$p_f(X) = (1 - X)^2 .$$

Pero

$$V(1) = \{(x, y) \in \mathbb{R}^2 \mid y = 0\} = \langle (1, 0) \rangle ,$$

luego $\dim V(1) = 1$.

6.2.15 Ejercicios.

1 a) Sea f el endomorfismo de $\mathbb{R}^3$ dado por la matriz

$$A = \begin{bmatrix} 4. & 0 & -20. \\ 2. & 0 & -10. \\ 1. & -1. & -2. \end{bmatrix} .$$

Calcúlese el polinomio característico de f y sus valores propios. Calcúlese un vector propio no nulo (de $\mathbb{R}^3$) para cada valor propio.

b) Sea g el endomorfismo de $\mathbb{C}^3$ dado por la misma matriz del apartado anterior. Calcúlese el polinomio característico de g y sus valores propios. Calcúlese un vector propio no nulo (de $\mathbb{C}^3$) para cada valor propio. Utilícese la proposición (6.1.14) para buscar una base de $\mathbb{C}^3$ formada por vectores propios de g. Compruébese que la matriz de g en dicha base es diagonal y que sus elementos diagonales son los valores propios de g.

2 Siendo f un endomorfismo de $\mathbb{R}^2$, pruébese la equivalencia de las proposiciones siguientes:

(i) los únicos subespacios invariantes para f son los triviales;
(ii) f no posee valores propios.

Búsquese un endomorfismo que verifique tales condiciones.

3 Demuéstrese que, si A es una matriz triangular, sus valores propios son los elementos diagonales de A.

4 Sea E un e.v. de dimensión $n \neq 0$. Calcúlese el polinomio característico de la homotecia h_λ, de razón λ (véase el ejercicio 5 de la sección 3.4); calcúlense los polinomios característicos del endomorfismo 0 y de id_E .

5 Pruébese que las matrices

$$A = \begin{bmatrix} 2. & & \\ & 2. & \\ & & 2. \end{bmatrix} \quad \text{y} \quad B = \begin{bmatrix} 2. & 0 & 3. \\ & 2. & -1. \\ & & 2. \end{bmatrix}$$

de $M_{\mathbb{R}}(3)$ poseen el mismo polinomio característico y que, sin embargo, no son semejantes (v. 3.5.27).

6 Sea $A \in M_{\mathbb{K}}(n)$ una matriz cuadrada y sea $\lambda \in \mathbb{K}$ un valor propio de A. Demuéstrese que la dimensión del subespacio propio $V(\lambda)$ es $n - r$, donde r el el rango de la matriz $A - \lambda I_n$.

7 Siendo $A, B \in M_{\mathbb{K}}(n)$ y $\lambda \in \mathbb{K}$, pruébese que

$$\mathrm{tr}(A+B) = \mathrm{tr}\, A + \mathrm{tr}\, B$$

y que

$$\mathrm{tr}(\lambda A) = \lambda\ \mathrm{tr}\, A$$

pero que, en general,

$$\mathrm{tr}(AB) \neq \mathrm{tr}\, A \cdot \mathrm{tr}\, B\,.$$

Pruébese que

$$\mathrm{tr}(AB) = \mathrm{tr}(BA)\,.$$

8 Sea $A \in M(n)$ una matriz cuadrada real o compleja y $\lambda_1, \ldots, \lambda_n$ las raíces complejas (iguales o distintas) del polinomio característico $p_A(X)$. Demuéstrese que

$$\mathrm{tr}\, A = \lambda_1 + \cdots + \lambda_n \quad \text{y} \quad \det A = \lambda_1 \cdots \lambda_n\,.$$

9 Sea E un e.v. de dimensión $n \neq 0$, $f \in \mathcal{L}(E)$, F un subespacio no trivial e invariante para f, r su dimensión. Se representa por f_F el endomorfismo inducido por f en F (v. 6.1.1). Pruébese que

$$p_f(X) = p_{f_F}(X)\, q(X)\,,$$

donde $q(X)$ es un polinomio de grado $n - r$. (Utilícese el ejercicio 1 de la sección 6.1.)

6.3 Diagonalización: condiciones.

6.3.1 Estamos ahora en condiciones de responder a las dos cuestiones planteadas en (6.1.6). Comencemos por proporcionarnos una forma breve de expresar el problema.

6.3.2 DEFINICIÓN. Sea E un e.v. de dimensión $n \neq 0$ sobre $\mathbb{K}$ y $f \in \mathcal{L}(E)$. Decimos que el endomorfismo f es *diagonalizable* cuando existe una base $(a_1, \ldots, a_n)$ de E tal que la matriz $[f, (a_i)]$ que representa a f en dicha base es una matriz diagonal.

Sea $A \in M_{\mathbb{K}}(n)$ una matriz cuadrada (recordemos que $\mathbb{K} = \mathbb{R}$ o $\mathbb{C}$). Decimos que A es *diagonalizable* en $\mathbb{K}$ cuando A es semejante a una matriz diagonal de $M_{\mathbb{K}}(n)$, esto es, cuando existe $P \in M_{\mathbb{K}}(n)$ inversible y tal que $P^{-1} A\, P$ es una matriz diagonal con elementos de $\mathbb{K}$.

6.3.3 Las dos nociones precedentes están profundamente relacionadas. Si f es un endomorfismo diagonalizable, $(a_1, \ldots, a_n)$ una base de E y $A = [f, (a_i)]$, entonces la matriz A es diagonalizable. En efecto, sabemos que existe una base $(b_1, \ldots, b_n)$ de E tal que la matriz

$$B = [f, (b_i)]$$

es diagonal; como

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

donde P es la matriz de paso de la base (a_i) a la base (b_i), resulta que A es diagonalizable.

6.3.4 Recíprocamente, si $A \in M_{\mathbb{K}}(n)$ es una matriz diagonalizable, cualquier endomorfismo que se represente por A es diagonalizable (en particular lo es el endomorfismo de $\mathbb{K}^n$ dado por A). En efecto; sea E un e.v. de dimensión n sobre $\mathbb{K}$, $f \in \mathcal{L}(E)$ y $(a_1, \ldots, a_n)$ una base de E, de tal manera que

$$[f, (a_i)] = A\,.$$

Puesto que A es diagonalizable, existe $P \in M_{\mathbb{K}}(n)$ inversible y tal que $P^{-1} A\, P$ es diagonal. Llamemos $(b_1, \ldots, b_n)$ a la base de E tal que $P = [(b_i), (a_i)]$; entonces

$$[f, (b_i)] = P^{-1} A\, P\,,$$

lo que demuestra que f es diagonalizable.

6.3.5 Un endomorfismo f de E es diagonalizable si y sólo si existe una base $(a_1, \ldots, a_n)$ de E formada por vectores propios de f.

En efecto; si $(a_1, \ldots, a_n)$ es una base de E formada por vectores propios, existen escalares $\lambda_1, \ldots, \lambda_n$ tales que

$$\begin{aligned} f(a_1) &= \lambda_1 a_1 \\ \ldots&\ldots\ldots\ldots \\ f(a_n) &= \lambda_n a_n \;; \end{aligned}$$

pero entonces

$$[f, (a_i)] = \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix}\,,$$

luego f es diagonalizable.

Recíprocamente, si f es diagonalizable, existe una base $(a_1, \ldots, a_n)$ de E tal que

$$[f, (a_i)] = \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix}$$

y, como esto significa que

$$\begin{aligned} f(a_1) &= \lambda_1 a_1 \\ \ldots&\ldots\ldots\ldots \\ f(a_n) &= \lambda_n a_n \;, \end{aligned}$$

resulta que los elementos de la base son vectores propios.

6.3.6 TEOREMA

Sea E un e.v. sobre $\mathbb{K}$ ($\mathbb{R}$ o $\mathbb{C}$), de dimensión $n \neq 0$; sea $f \in \mathcal{L}(E)$. El endomorfismo f es diagonalizable si y sólo si se verifican las dos propiedades siguientes

(d1) $p_f(X)$ posee n raíces en $\mathbb{K}$, iguales o distintas (más exactamente, posee raíces en $\mathbb{K}$ cuyos órdenes de multiplicidad suman n),

propiedad que siempre se verifica cuando $\mathbb{K} = \mathbb{C}$, y

(d2) para cada raíz $\lambda \in \mathbb{K}$ de $p_f(X)$,

$$\dim V(\lambda) = k\,,$$

donde k es el orden de multiplicidad de λ.

El mismo resultado es cierto para una matriz $A \in M_{\mathbb{K}}(n)$, substituyendo $p_f(X)$ por $p_A(X)$.

Supongamos en primer lugar que f es diagonalizable; existe entonces una base $(a_1, \ldots, a_n)$ de E tal que $[f, (a_i)]$ es diagonal, o sea,

$$[f, (a_i)] = \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix}.$$

Resulta así que

$$p_f(X) = (\lambda_1 - X)(\lambda_2 - X) \cdots (\lambda_n - X)\,,$$

luego $p_f(X)$ posee n raíces, $\lambda_1, \lambda_2, \ldots, \lambda_n$, en $\mathbb{K}$ (no necesariamente diferentes), lo que prueba **(d1)**. Si λ es una de estas raíces y k su orden de multiplicidad, entonces

$$\lambda_{i_1} = \lambda_{i_2} = \cdots = \lambda_{i_k} = \lambda$$

para k índices $i_1, \ldots, i_k \in \{1, \ldots, n\}$. Los correspondientes vectores de la base verifican, pues,

$$\begin{aligned} f(a_{i_1}) &= \lambda_{i_1} a_{i_1} = \lambda a_{i_1} \\ &\cdots\cdots\cdots\cdots \\ f(a_{i_k}) &= \lambda_{i_k} a_{i_k} = \lambda a_{i_k}\quad , \end{aligned}$$

es decir,

$$a_{i_1}, \ldots, a_{i_k} \in V(\lambda)\,,$$

luego

$$\dim V(\lambda) \geq k\,.$$

Como por otra parte $\dim V(\lambda) \leq k$ (v. 6.2.13), tenemos la igualdad

$$\dim V(\lambda) = k\,,$$

lo que prueba **(d2)**.

Recíprocamente, supongamos que se verifican **(d1)** y **(d2)**. Por la primera sabemos que existen elementos $\lambda_1, \dots, \lambda_p$ que son raíces de $p_f(X)$ y que sus órdenes de multiplicidad, $k_1, \dots, k_p$, suman

$$k_1 + \cdots + k_p = n\,.$$

Los λ_i son valores propios de f y cada subespacio $V(\lambda_i)$ es de dimensión k_i , puesto que se cumple **(d2)**. La suma directa (v. 6.1.16)

$$V(\lambda_1) \oplus \cdots \oplus V(\lambda_p)$$

es un subespacio de dimensión $k_1 + \cdots + k_p = n$ (v. 2.5.7), o sea,

$$V(\lambda_1) \oplus \cdots \oplus V(\lambda_p) = E\,.$$

Reuniendo bases de $V(\lambda_1), \dots, V(\lambda_p)$, obtenemos entonces una base de E, base que estará formada por vectores propios. El endomorfismo f es, pues, diagonalizable.

El resultado para matrices es evidente, teniendo en cuenta que A es diagonalizable si y sólo si lo es el endomorfismo de $\mathbb{K}^n$ dado por A (v. 6.3.3 y 6.3.4).

6.3.7 La demostración del teorema precedente nos indica además un procedimiento para obtener una base $(a_1, \dots, a_n)$ en la que un endomorfismo diagonalizable f tenga asociada una matriz diagonal. Tal base se puede formar reuniendo bases de los subespacios $V(\lambda_i)$, para los valores propios λ_i de f.

6.3.8 COROLARIO

Sea E un e.v. sobre $\mathbb{K}$ ($\mathbb{R}$ o $\mathbb{C}$), de dimensión $n \neq 0$; sea $f \in \mathcal{L}(E)$. Si $p_f(X)$ tiene n raíces distintas en $\mathbb{K}$, entonces f es diagonalizable.

El mismo resultado es cierto para una matriz $A \in M_{\mathbb{K}}(n)$, substituyendo $p_f(X)$ por $p_A(X)$.

Si $p_f(X)$ posee n raíces distintas, $\lambda_1, \dots, \lambda_n$, en $\mathbb{K}$, se verifica la propiedad **(d1)** del teorema precedente. Además cada λ_i es de multiplicidad 1, luego (v. 6.2.13)

$$1 \leq \dim V(\lambda_i) \leq 1\,,$$

o sea, $\dim V(\lambda_i) = 1$; se verifica también **(d2)**.

6.3.9 Ejemplo. Sea $f \in \mathcal{L}(\mathbb{R}^3)$ el endomorfismo dado por la matriz

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

Tenemos que

$$\begin{aligned} p_f(X) &= \begin{vmatrix} -2.-X & 4. & 5. \\ -3. & 5.-X & 5. \\ 0 & 0 & 1.-X \end{vmatrix} \\ &= (1-X)\,(\,(5-X)(-2-X)+12) \\ &= (1-X)(X^2-3X+2) \end{aligned}$$

y las raíces de este polinomio son

$$1 \text{ con multiplicidad } 2 \quad \text{y} \quad 2 \text{ con multiplicidad } 1\,;$$

se verifica entonces **(d1)**. Vamos a estudiar los subespacios propios.

$V(1) = \mathrm{Ker}(f - id)$, luego $(x, y, z) \in V(1)$ cuando

$$(A - I_3)\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix},$$

o sea,

$$\begin{bmatrix} -3. & 4. & 5. \\ -3. & 4. & 5. \\ 0 & 0 & 0 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix};$$

la solución de este sistema es

$$(x, y, z) = (4\lambda + 5\mu, 3\lambda, 3\mu)$$

y por consiguiente

$$V(1) = \langle (4,3,0), (5,0,3) \rangle\,.$$

$V(1)$ es de dimensión 2. Se cumple también **(d2)**, pues ya sabemos que la dimensión de $V(2)$ no puede ser sino 1. El endomorfismo f es diagonalizable.

Para encontrar una base en la que corresponda a f una matriz diagonal, tenemos que calcular también $V(2)$.

$V(2) = \mathrm{Ker}(f - 2\,id)$, luego $(x, y, z) \in V(2)$ cuando

$$(A - 2\,I_3)\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix},$$

o sea,

$$\begin{bmatrix} -4. & 4. & 5. \\ -3. & 3. & 5. \\ 0 & 0 & -1. \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix};$$

la solución de este sistema es

$$(x, y, z) = (\lambda, \lambda, 0)\,,$$

luego

$$V(2) = \langle (1, 1, 0) \rangle\,.$$

En la base de $\mathbb{R}^3$

$$(\,(4, 3, 0), (5, 0, 3), (1, 1, 0)\,)$$

f se representa por la matriz diagonal

$$\begin{bmatrix} 1. & & \\ & 1. & \\ & & 2. \end{bmatrix}.$$

La matriz A es diagonalizable en $\mathbb{R}$ y

$$P^{-1}A\,P = \begin{bmatrix} 1. & & \\ & 1. & \\ & & 2. \end{bmatrix},$$

siendo P la matriz

$$P = \begin{bmatrix} 4. & 5. & 1. \\ 3. & 0 & 1. \\ 0 & 3. & 0 \end{bmatrix}.$$

Compruébese este hecho, teniendo en cuenta que

$$P^{-1} = \frac{1}{3.}\begin{bmatrix} 3. & -3. & -5. \\ 0 & 0 & 1. \\ -9. & 12. & 15. \end{bmatrix}.$$

6.3.10 Ejemplo. Sea $f \in \mathcal{L}(\mathbb{R}^3)$ el endomorfismo dado por la matriz

$$A = \begin{bmatrix} 2. & -2. & 1. \\ 1. & 3. & 1. \\ 0 & 1. & 2. \end{bmatrix}.$$

Su polinomio característico vale

$$\begin{aligned} p_f(X) &= \begin{vmatrix} 2.-X & -2. & 1. \\ 1. & 3.-X & 1. \\ 0 & 1. & 2.-X \end{vmatrix} \\ &= (2-X)^2(3-X) + 1 - (2-X) + 2(2-X) \\ &= (2-X)^2(3-X) + (3-X) \\ &= (3-X)(\,(2-X)^2 + 1) \\ &= (3-X)(X^2 - 4X + 5)\,; \end{aligned}$$

las raíces son 3, $2+i$ y $2-i$. $p_f(X)$ posee únicamente una raíz en $\mathbb{R}$, luego **(d1)** no se cumple y f no es diagonalizable.

6.3.11 Ejemplo. Sea $f \in \mathcal{L}(\mathbb{R}^3)$ el endomorfismo dado por la matriz

$$A = \begin{bmatrix} 3. & 2. & 4. \\ 0 & 1. & 0 \\ -2. & 0 & -3. \end{bmatrix}.$$

Tenemos que

$$\begin{aligned} p_f(X) &= \begin{vmatrix} 3.-X & 2. & 4. \\ 0 & 1.-X & 0 \\ -2. & 0 & -3.-X \end{vmatrix} \\ &= (1-X)(\,(3-X)(-3-X)+8) \\ &= (1-X)(X^2-1) \end{aligned}$$

con raíces

$$1 \text{ con multiplicidad } 2 \quad \text{y} \quad -1 \text{ con multiplicidad } 1\,;$$

se verifica entonces **(d1)**.

$V(1) = \operatorname{Ker}(f - id)$, luego $(x, y, z) \in V(1)$ cuando

$$\begin{bmatrix} 2. & 2. & 4. \\ 0 & 0 & 0 \\ -2. & 0 & -4. \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix};$$

la solución de este sistema es

$$(x, y, z) = (-2\lambda, 0, \lambda)$$

y por consiguiente

$$V(1) = \langle(-2, 0, 1)\rangle\,;$$

$V(1)$ es de dimensión 1 y no se verifica **(d2)**. f no es diagonalizable.

6.3.12 Ejemplo. Sea $f \in \mathcal{L}(\mathbb{R}^3)$ el endomorfismo del ejemplo (6.1.5), dado por la matriz

$$A = \begin{bmatrix} -1. & 3. & 6. \\ -6. & 8. & 16. \\ 2. & -2. & -4. \end{bmatrix}.$$

Vimos en dicho ejemplo que f es diagonalizable, puesto que encontramos una base en la que la matriz de f es diagonal.

El polinomio característico de f es

$$\begin{aligned} p_f(X) &= \begin{vmatrix} -1.-X & 3. & 6. \\ -6. & 8.-X & 16. \\ 2. & -2. & -4.-X \end{vmatrix} \\ &= (-1-X)(8-X)(-4-X)+96+72 \\ &\quad -12(8-X)+18(-4-X)+32(-1-X) \\ &= -X^3+3X^2-2X \\ &= -X(X^2-3X+2) \end{aligned}$$

de raíces 0, 1 y 2. Utilizando el corolario (6.3.8) se puede llegar también a la conclusión de que f es diagonalizable.

6.3.13 Ejemplo. Los endomorfismos f y g del ejemplo (6.2.12) se comportan de manera diferente a pesar de venir dados por la misma matriz

$$A = \begin{bmatrix} 0 & -1. \\ 1. & 0 \end{bmatrix}.$$

El endomorfismo $g \in \mathcal{L}(\mathbb{C}^2)$ es diagonalizable; en la base $(\,(1,-i),(1,i)\,)$ de $\mathbb{C}^2$, la matriz de g es

$$\begin{bmatrix} i & \\ & -i \end{bmatrix}.$$

Sin embargo, el endomorfismo $f \in \mathcal{L}(\mathbb{R}^2)$ no es diagonalizable.

La matriz A es diagonalizable en $\mathbb{C}$, puesto que

$$\begin{bmatrix} 1. & 1. \\ -i & i \end{bmatrix}^{-1} \begin{bmatrix} 0 & -1. \\ 1. & 0 \end{bmatrix} \begin{bmatrix} 1. & 1. \\ -i & i \end{bmatrix} = \begin{bmatrix} i & \\ & -i \end{bmatrix}$$

(compruébese este hecho). Sin embargo, A no es diagonalizable en $\mathbb{R}$, puesto que no existe ninguna matriz $P \in M_{\mathbb{R}}(2)$ (¡con elementos reales!) inversible y tal que $P^{-1}A\,P$ sea diagonal con elementos reales.

6.3.14 Ejercicios.

1 Sea E un e.v. de dimensión $n \neq 0$ y h_λ una homotecia de E (véase el ejercicio 5 de la sección 3.4). Pruébese que, cualquiera que sea la base de E, f se representa por una matriz diagonal.

2 Demuéstrese que si A es una matriz diagonalizable, entonces A^2 lo es también; si además A es inversible, pruébese que A^{-1} es también diagonalizable.

3 Demuéstrese que la matriz de $M(2)$

$$A = \begin{bmatrix} 1. & 1. \\ 0 & 1. \end{bmatrix}$$

no es diagonalizable ni en $\mathbb{R}$ ni en $\mathbb{C}$.

4 a) Sea A una matriz $n \times n$ con un solo valor propio λ de multiplicidad n. Demuéstrese que A es diagonalizable si y sólo si A es la matriz escalar

$$A = \begin{bmatrix} \lambda & & \\ & \ddots & \\ & & \lambda \end{bmatrix}$$

(en cuyo caso ya es diagonal).

b) Sea $f \in \mathcal{L}(E)$, donde E es un e.v. de dimensión $n \neq 0$. Demuéstrese que, si todo subespacio de E es invariante para f, entonces f es una homotecia (véase el ejercicio 5 de la sección 3.4).

5 Demuéstrese que la matriz de $M_{\mathbb{R}}(3)$

$$A = \begin{bmatrix} 7. & -10. & 0 \\ 3. & -4. & 0 \\ 1. & -2. & 2. \end{bmatrix}$$

es diagonalizable en $\mathbb{R}$. Búsquese una matriz inversible $P \in M_{\mathbb{R}}(3)$ tal que $A' = P^{-1} A\, P$ sea diagonal y dígase cuánto vale A'.

6 Demuéstrese que la matriz de $M(3)$

$$A = \begin{bmatrix} 4. & 1. & -4. \\ -3. & 0 & 3. \\ 3. & 1. & -3. \end{bmatrix}$$

no es diagonalizable ni en $\mathbb{R}$ ni en $\mathbb{C}$.

7 Consideremos la matriz

$$A = \begin{bmatrix} 4. & 0 & -20. \\ 2. & 0 & -10. \\ 1. & -1. & -2. \end{bmatrix}.$$

Demuéstrese que no existe ninguna matriz inversible $P \in M_{\mathbb{R}}(3)$ tal que $P^{-1} A\, P$ sea diagonal. Calcúlese, sin embargo, una matriz inversible $P \in M_{\mathbb{C}}(3)$ tal que $A' = P^{-1} A\, P$ sea diagonal y dígase cuánto vale A'.

8 Sea E un e.v. de dimensión $n \neq 0$ y $(a_1, \ldots, a_n)$ una base de E; sean f y g dos endomorfismos de E.

a) Demuéstrese que si los vectores $a_1, \ldots, a_n$ son vectores propios a la vez de f y de g, entonces $f \circ g = g \circ f$.

b) Demuéstrese que si A y B son matrices $n \times n$, diagonalizables mediante una misma matriz inversible P (o sea, tales que $P^{-1} A\, P$ y $P^{-1} B\, P$ son diagonales), entonces $AB = BA$.

c) Supóngase que los vectores $a_1, \ldots, a_n$ son vectores propios de f correspondientes a valores propios $\lambda_1, \ldots, \lambda_n$, todos diferentes entre sí, y que $f \circ g = g \circ f$. Pruébese que, entonces, los vectores $a_1, \ldots, a_n$ son también vectores propios de g.

d) Demuéstrese que si A y B son matrices $n \times n$ tales que $AB = BA$, y si A tiene n valores propios diferentes, existe una matriz inversible P tal que $P^{-1} A\, P$ y $P^{-1} B\, P$ son diagonales.

9 Pruébese que toda matriz $A \in M_{\mathbb{R}}(2)$, simétrica, es diagonalizable en $\mathbb{R}$.

6.4 Forma triangular de endomorfismos y matrices.

6.4.1 Cuando no es posible diagonalizar un endomorfismo se recurre a representarlo por una matriz sencilla de otro tipo. La forma alternativa más importante es la llamada forma canónica o forma de Jordan; las matrices de este tipo son triangulares y casi-diagonales. Trataremos este tema en la sección 6.6. En algunas ocasiones basta con reducir el endomorfismo o la matriz a la forma triangular; esta será la posibilidad que estudiaremos ahora.

6.4.2 Sea E un e.v. sobre $\mathbb{K}$ ($\mathbb{R}$ o $\mathbb{C}$) de dimensión $n \neq 0$ y $f \in \mathcal{L}(E)$; sea $(a_1, \ldots, a_n)$ una base de E y $[f, (a_i)] = [\alpha_i^j]$ la matriz de f en dicha base. Si la matriz es triangular, los elementos diagonales, $\alpha_1^1, \ldots, \alpha_n^n$, son los valores propios de f. En efecto

$$p_f(X) = (\alpha_1^1 - X)(\alpha_2^2 - X) \cdots (\alpha_n^n - X)\,,$$

y basta aplicar (6.2.9).

Sea $A \in M_{\mathbb{K}}(n)$ y supongamos que existe una matriz inversible $P \in M_{\mathbb{K}}(n)$ tal que la matriz $A' = P^{-1} A\, P$ es triangular con elementos en $\mathbb{K}$. Entonces los elementos diagonales de A' son los valores propios de A, como se prueba con un razonamiento idéntico al anterior.

6.4.3 PROPOSICIÓN

---

a) Sea E un e.v. sobre $\mathbb{K}$ ($\mathbb{R}$ o $\mathbb{C}$) de dimensión $n \neq 0$ y $f \in \mathcal{L}(E)$. Para que exista una base $(a_1, \ldots, a_n)$ de E tal que la matriz $[f, (a_i)]$ es triangular superior, es condición necesaria y suficiente que se verifique

(d1) $p_f(X)$ posee n raíces en $\mathbb{K}$, iguales o distintas,

propiedad que siempre se verifica cuando $\mathbb{K} = \mathbb{C}$.

b) Sea $A \in M_{\mathbb{K}}(n)$, $\mathbb{K} = \mathbb{R}$ o $\mathbb{C}$. Para que A sea semejante en $\mathbb{K}$ a una matriz triangular superior, es condición necesaria y suficiente que se verifique

(d1) $p_A(X)$ posee n raíces en $\mathbb{K}$, iguales o distintas,

propiedad que siempre se verifica cuando $\mathbb{K} = \mathbb{C}$. Toda matriz cuadrada compleja es semejante a una matriz triangular superior de números complejos.

c) Los resultados precedentes son también ciertos cuando substituimos 'triangular superior' por 'triangular inferior'.

Veamos primero el resultado para matrices.

El razonamiento del apartado (6.4.2) prueba que la condición **(d1)** es necesaria.

Recíprocamente, supongamos que A verifica **(d1)**. Vamos a probar que entonces A es semejante a una matriz triangular superior. Lo probaremos por inducción sobre la dimensión n de la matriz. Para $n=1$ el resultado es evidente, puesto que toda matriz 1×1 es triangular superior. Supongamos cierto el resultado para n (hipótesis de recurrencia) y sea $A\in M_{\mathbb{K}}(n+1)$ una matriz tal que $p_A(X)$ posee $n+1$ raíces en $\mathbb{K}$. Llamemos λ a una de estas raíces; λ es un valor propio de A y del endomorfismo $f\in\mathcal{L}(\mathbb{K}^{n+1})$ dado por A, luego existe $a_1\in\mathbb{K}^{n+1}$, $a_1\neq 0$, tal que

$$f(a_1)=\lambda a_1\,.$$

Existen vectores $a_2,\ldots,a_{n+1}\in\mathbb{K}^{n+1}$ tales que $(a_1,a_2,\ldots,a_{n+1})$ es una base de $\mathbb{K}^{n+1}$. La matriz $A'=[f,(a_i)]$ es

$$A'=Q^{-1}A\,Q\,,$$

donde $Q=[(a_i),(e_i)]$, y es de la forma

$$A'=\left[\begin{array}{c|ccc}\lambda & \alpha_1' & \cdots & \alpha_n'\\ \hline 0 & & & \\ \vdots & & A_1 & \\ 0 & & & \end{array}\right]$$

con $A_1\in M_{\mathbb{K}}(n)$. Calculando su polinomio característico, que coincide con el de A, resulta

$$p_A(X)=p_{A'}(X)=(\lambda-X)\det(A_1-X\,I_n)=(\lambda-X)p_{A_1}(X)\,.$$

El polinomio $p_{A_1}(X)$ posee entonces n raíces en $\mathbb{K}$. Utilizando la hipótesis de recurrencia, existe una matriz inversible $Q_1\in M_{\mathbb{K}}(n)$ tal que la matriz

$$A_1'=Q_1^{-1}A_1\,Q_1$$

es triangular superior. Pongamos ahora

$$P=Q\left[\begin{array}{c|c}1&0\\ \hline 0&Q_1\end{array}\right]\;;$$

la matriz P es inversible y su inversa es

$$P^{-1}=\left[\begin{array}{c|c}1&0\\ \hline 0&Q_1^{-1}\end{array}\right]Q^{-1}\,.$$

Además,

$$\begin{aligned}P^{-1}A\,P\;&=\;\left[\begin{array}{c|c}1&0\\ \hline 0&Q_1^{-1}\end{array}\right]Q^{-1}A\,Q\left[\begin{array}{c|c}1&0\\ \hline 0&Q_1\end{array}\right]\\ &=\;\left[\begin{array}{c|c}1&0\\ \hline 0&Q_1^{-1}\end{array}\right]A'\left[\begin{array}{c|c}1&0\\ \hline 0&Q_1\end{array}\right]\end{aligned}$$

$$
\begin{aligned}
&= \left[\begin{array}{c|c} 1 & 0 \\ \hline 0 & Q_1^{-1} \end{array}\right] \left[\begin{array}{c|ccc} \lambda & \alpha'_1 & \cdots & \alpha'_n \\ \hline 0 & & A_1 & \end{array}\right] \left[\begin{array}{c|c} 1 & 0 \\ \hline 0 & Q_1 \end{array}\right] \\
&= \left[\begin{array}{c|c} 1 & 0 \\ \hline 0 & Q_1^{-1} \end{array}\right] \left[\begin{array}{c|ccc} \lambda & \beta_1 & \cdots & \beta_n \\ \hline 0 & & A_1\, Q_1 & \end{array}\right] \\
&= \left[\begin{array}{c|ccc} \lambda & \beta_1 & \cdots & \beta_n \\ \hline 0 & & Q_1^{-1} A_1\, Q_1 & \end{array}\right] \\
&= \left[\begin{array}{c|ccc} \lambda & \beta_1 & \cdots & \beta_n \\ \hline 0 & & A'_1 & \end{array}\right]
\end{aligned}
$$

y esta matriz es triangular superior, puesto que A'_1 lo es. Esto termina la recurrencia y la demostración de b).

Utilizando b) y el mismo argumento que en (6.3.4) se prueba sin dificultad el apartado a).

Finalmente, si $[f, (a_1, a_2, \ldots, a_n)]$ es triangular superior (inferior), entonces $[f, (a_n, \ldots, a_2, a_1)]$ es triangular inferior (superior). Se obtiene así el resultado c) para endomorfismos. El correspondiente resultado para matrices se sigue de éste con el mismo argumento que en (6.3.3).

6.4.4 Ejemplo. La matriz

$$
A = \begin{bmatrix} 3. & 2. & 4. \\ 0 & 1. & 0 \\ -2. & 0 & -3. \end{bmatrix}
$$

de $M_{\mathbb{R}}(3)$ del ejemplo (6.3.11) no es diagonalizable (ni en $\mathbb{R}$ ni en $\mathbb{C}$), pero verifica **(d1)**; es entonces semejante en $\mathbb{R}$ a una matriz triangular superior. Sus valores propios son, como vimos, 1 y -1 (el primero con multiplicidad 2); vimos también que

$$
(-2, 0, 1) \in V(1) .
$$

$((-2, 0, 1), (0, 1, 0), (0, 0, 1))$ es una base de $\mathbb{R}^3$; la matriz

$$
Q = \begin{bmatrix} -2. & 0 & 0 \\ 0 & 1. & 0 \\ 1. & 0 & 1. \end{bmatrix}
$$

es inversible, con

$$
Q^{-1} = \frac{1}{-2.} \begin{bmatrix} 1. & 0 & 0 \\ 0 & -2. & 0 \\ -1. & 0 & -2. \end{bmatrix} ,
$$

y se tiene

$$
A' = Q^{-1} A\, Q = \left[\begin{array}{c|cc} 1. & -1. & -2. \\ \hline 0 & 1. & 0 \\ 0 & 1. & -1. \end{array}\right] .
$$

Ponemos

$$
A_1 = \begin{bmatrix} 1. & 0 \\ 1. & -1. \end{bmatrix} ;
$$

como

$$p_A(X) = p_{A'}(X) = (1 - X)p_{A_1}(X)\,,$$

ya sabemos que los valores propios de A_1 son 1 y -1 (ahora ambos con multiplicidad 1). El subespacio $V(1)$ de $\mathbb{R}^2$, correspondiente a la matriz A_1 , es

$$V(1) = \langle (2,1) \rangle\,.$$

$(\,(2,1),(0,1)\,)$ es una base de $\mathbb{R}^2$; la matriz

$$Q_1 = \begin{bmatrix} 2. & 0 \\ 1. & 1. \end{bmatrix}$$

es inversible, con

$$Q_1^{-1} = \frac{1}{2.}\begin{bmatrix} 1. & 0 \\ -1. & 2. \end{bmatrix}\,,$$

y se tiene que

$$A_1' = Q_1^{-1} A_1\, Q_1 = \begin{bmatrix} 1. & 0 \\ 0 & -1. \end{bmatrix}\,.$$

Calculamos entonces

$$\begin{aligned} P \;=\; Q & \left[\begin{array}{c|cc} 1. & 0 & 0 \\ \hline 0 & \multicolumn{2}{c}{\multirow{2}{*}{Q_1}} \\ 0 & & \end{array}\right] \\ = & \begin{bmatrix} -2. & 0 & 0 \\ 0 & 1. & 0 \\ 1. & 0 & 1. \end{bmatrix} \begin{bmatrix} 1. & 0 & 0 \\ 0 & 2. & 0 \\ 0 & 1. & 1. \end{bmatrix} \\ = & \begin{bmatrix} -2. & 0 & 0 \\ 0 & 2. & 0 \\ 1. & 1. & 1. \end{bmatrix}. \end{aligned}$$

La matriz P es inversible y

$$P^{-1} A\, P = \begin{bmatrix} 1. & -4. & -2. \\ 0 & 1. & 0 \\ 0 & 0 & -1. \end{bmatrix}\,,$$

como se comprueba fácilmente.

En el tratamiento de este ejemplo hemos seguido paso a paso el desarrollo realizado en la demostración de la proposición anterior. De hecho se puede proceder de forma más rápida si en cada paso se reduce lo más posible el tamaño de la siguiente matriz a triangularizar. Esto es lo que haremos en el ejemplo que sigue.

6.4.5 Ejemplo. Consideremos la matriz

$$A = \begin{bmatrix} 2. & 1. & 1. & 2. \\ 2. & 1. & -1. & 0 \\ 1. & 0 & 0 & -1. \\ 1. & -2. & 0 & 1. \end{bmatrix}$$

de $M_{\mathbb{R}}(4)$. Su polinomio característico

$$p_A(X) = (X-2)^3(X+2)$$

posee raíces 2 (con multiplicidad 3) y -2.

Es fácil comprobar que $\dim V(2) = 1$ y que A no es diagonalizable (ni en $\mathbb{R}$ ni en $\mathbb{C}$); sin embargo, A es triangularizable en $\mathbb{R}$.

Tenemos que

$$(1,1,1,-1) \in V(2) \quad \text{y} \quad (1,-1,-1,-1) \in V(-2)\,.$$

$(\,(1,1,1,-1),(1,-1,-1,-1),(0,0,1,0),(0,0,0,1)\,)$ es una base de $\mathbb{R}^4$; la matriz

$$Q = \begin{bmatrix} 1. & 1. & 0 & 0 \\ 1. & -1. & 0 & 0 \\ 1. & -1. & 1. & 0 \\ -1. & -1. & 0 & 1. \end{bmatrix}$$

es inversible, con

$$Q^{-1} = \frac{1}{2.}\begin{bmatrix} 1. & 1. & 0 & 0 \\ 1. & -1. & 0 & 0 \\ 0 & -2. & 2. & 0 \\ 2. & 0 & 0 & 2. \end{bmatrix},$$

y se tiene

$$A' = Q^{-1}A\,Q = \begin{bmatrix} 2. & 0 & 0 & 1. \\ 0 & -2. & 1. & 1. \\ 0 & 0 & 1. & -1. \\ 0 & 0 & 1. & 3. \end{bmatrix}.$$

Ponemos

$$A_1 = \begin{bmatrix} 1. & -1. \\ 1. & 3. \end{bmatrix};$$

como

$$p_A(X) = p_{A'}(X) = (2-X)(-2-X)p_{A_1}(X)\,,$$

el único valor propio de A_1 es 2, ahora con multiplicidad 2. Para el subespacio $V(2)$ de $\mathbb{R}^2$, correspondiente a la matriz A_1, se tiene que

$$(1,-1) \in V(2)\,.$$

$(\,(1,-1),(0,1)\,)$ es una base de $\mathbb{R}^2$; la matriz

$$Q_1 = \begin{bmatrix} 1. & 0 \\ -1. & 1. \end{bmatrix}$$

es inversible, con

$$Q_1^{-1} = \begin{bmatrix} 1. & 0 \\ 1. & 1. \end{bmatrix}$$

y se tiene

$$A_1' = Q_1^{-1} A_1 \, Q_1 = \begin{bmatrix} 2. & -1. \\ 0 & 2. \end{bmatrix} .$$

Ponemos entonces

$$\begin{aligned} P &= Q \left[\begin{array}{cc|c} 1. & & \\ & 1. & 0 \\ \hline & 0 & Q_1 \end{array} \right] \\ &= \begin{bmatrix} 1. & 1. & 0 & 0 \\ 1. & -1. & 0 & 0 \\ 1. & -1. & 1. & 0 \\ -1. & -1. & 0 & 1. \end{bmatrix} \begin{bmatrix} 1. & 0 & 0 & 0 \\ 0 & 1. & 0 & 0 \\ 0 & 0 & 1. & 0 \\ 0 & 0 & -1. & 1. \end{bmatrix} \\ &= \begin{bmatrix} 1. & 1. & 0 & 0 \\ 1. & -1. & 0 & 0 \\ 1. & -1. & 1. & 0 \\ -1. & -1. & -1. & 1. \end{bmatrix} . \end{aligned}$$

La matriz P es inversible y

$$P^{-1} A \, P = \begin{bmatrix} 2. & 0 & -1. & 1. \\ 0 & -2. & 0 & 1. \\ 0 & 0 & 2. & -1. \\ 0 & 0 & 0 & 2. \end{bmatrix} ,$$

como puede comprobarse.

6.4.6 Ejercicios.

1 Sea E un e.v. de dimensión $n \neq 0$, $f \in \mathcal{L}(E)$ y $(a_1, \ldots, a_n)$ una base de E. Demuéstrese que la matriz $[f, (a_i)]$ es triangular superior si y sólo si los subespacios

$$\langle a_1 \rangle , \quad \langle a_1, a_2 \rangle , \quad \ldots \quad , \langle a_1, a_2, \ldots, a_r \rangle , \quad \ldots \quad , \langle a_1, a_2, \ldots, a_n \rangle$$

son todos invariantes.

2 Demuéstrese que la matriz

$$A = \begin{bmatrix} 4. & 1. & -4. \\ -3. & 0 & 3. \\ 3. & 1. & -3. \end{bmatrix}$$

de $M_{\mathbb{R}}(3)$ del ejercicio 6 de la sección 6.3 es semejante en $\mathbb{R}$ a una matriz triangular. Búsquese una matriz inversible $P \in M_{\mathbb{R}}(3)$ tal que $A' = P^{-1} A \, P$ sea triangular superior, y dígase cuanto vale A'.

Demuéstrese que la matriz

$$A = \begin{bmatrix} 1. & 0 & 0 & 1. \\ 2. & 0 & 1. & 2. \\ -1. & -1. & 2. & 1. \\ -1. & 0 & 0 & 3. \end{bmatrix}$$

de $M_{\mathbb{R}}(4)$ no es diagonalizable ni en $\mathbb{R}$ ni en $\mathbb{C}$. Pruébese que es triangularizable en $\mathbb{R}$ y búsquese una matriz inversible $P \in M_{\mathbb{R}}(4)$ tal que $A' = P^{-1} A\, P$ sea triangular superior, y dígase cuanto vale A'.

3 Pruébese que la matriz

$$A = \begin{bmatrix} 1. & 1. \\ -1. & 1. \end{bmatrix}$$

no es semejante a ninguna matriz triangular real. Pruébese que A es semejante a una matriz compleja A', triangular superior, y calcúlese A'.

4 Pruébese que, si $A \in M_{\mathbb{C}}(2)$, entonces es, o bien diagonalizable, o bien semejante a una matriz de la forma

$$\begin{bmatrix} \lambda & 1. \\ & \lambda \end{bmatrix},$$

con $\lambda \in \mathbb{C}$.

6.5 Polinomios que anulan una matriz.

6.5.1 Vamos a dedicar esta sección a desarrollar brevemente una idea que tendremos que utilizar en la construcción de las formas canónicas. Conviene que el lector recuerde lo que significan $p(f)$ y $p(A)$ cuando p es un polinomio, f un endomorfismo y A una matriz cuadrada (v. 3.4.8 y 3.4.29).

6.5.2 Si A y A' son matrices semejantes, o sea, si $A' = P^{-1} A\, P$ para una matriz inversible P, y si $p(X)$ es un polinomio, entonces $p(A)$ y $p(A')$ son también matrices semejantes. Más exactamente, se tiene

$$p(A') = P^{-1} p(A)\, P\,.$$

La demostración de este hecho no presenta problemas. Se procede por inducción (v. 1.4.5) para probar que

$$A'^n = P^{-1} A^n\, P$$

para todo $n \in \mathbb{N}$, y luego se extiende el resultado a los polinomios.

6.5.3 Si $p(X)$ es un polinomio y

$$A = \begin{bmatrix} \boxed{A_1} & & & \\ & \boxed{A_2} & & \\ & & \ddots & \\ & & & \boxed{A_k} \end{bmatrix}$$

una matriz cuadrada diagonal por bloques (se entiende que los bloques A_i son cuadrados), entonces

$$p(A)=\begin{bmatrix} \boxed{p(A_1)} & & & \\ & \boxed{p(A_2)} & & \\ & & \ddots & \\ & & & \boxed{p(A_k)} \end{bmatrix}.$$

Esto se puede probar también por inducción para potencias cualesquiera de A, pasando luego a polinomios.

6.5.4 Sea A una matriz $n \times n$ con elementos en $\mathbb{K}$ ($\mathbb{R}$ o $\mathbb{C}$). No es evidente que A sea 'raiz' de algún polinomio, esto es, que

$$p(A) = 0$$

para algún polinomio $p \in \mathbb{K}[X]$ que no sea, claro está, el polinomio 0. Sin embargo existen infinitos polinomios que anulan A; si $p(A) = 0$ es claro que tambiés será $q(A) = 0$ para todo polinomio, $q(X) = r(X)\,p(X)$, que sea múltiplo de p. Basta pues encontrar un polinomio que anule A para tener una infinidad de ellos.

Veamos una primera forma de probar la existencia de un polinomio que anula A. Las $n^2 + 1$ matrices $I, A, A^2, \ldots, A^{n^2}$, constituyen un sistema ligado de $M(n)$, puesto que la dimensión de este espacio es n^2. En consecuencia, existen escalares $\alpha_0, \alpha_1, \alpha_2, \ldots, \alpha_{n^2}$, no todos nulos y tales que

$$\alpha_0 I + \alpha_1 A + \alpha_2 A^2 + \cdots + \alpha_{n^2} A^{n^2} = 0\,,$$

esto es, el polinomio

$$p(X) = \alpha_0 + \alpha_1 X + \alpha_2 X^2 + \cdots + \alpha_{n^2} X^{n^2}$$

verifica que $p(A) = 0$. Nótese que, como los α_i no son todos nulos, el polinomio $p(X)$ no es idénticamente nulo.

Este resultado puede ser un punto de partida, pero en sí mismo no es muy útil. En primer lugar, p puede ser de grado n^2 y veremos que se pueden encontrar polinomios que anulen A y tengan menor grado. Pero, además, no sabemos cuál es el polinomio p, ni aún conociendo la matriz A. Veremos inmediatamente un resultado que permite rebajar el grado del polinomio que anula A y calcular este polinomio.

Conviene recordar que, si A y A' son semejantes, lo son $p(A)$ y $p(A')$ para todo polinomio p. En consecuencia, dos matrices semejantes son anuladas por los mismos polinomios.

Se pueden hacer idénticas consideraciones para un endomorfismo $f \in \mathcal{L}(E)$ de un espacio E de dimensión finita. Para ello basta recordar que, si A representa a f en una base, $p(A)$ representa en la misma base a $p(f)$. Esto significa en particular que un endomorfismo y la matriz que lo representa son anulados por los mismos polinomios.

6.5.5 Ejemplo. Para la matriz

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

del ejemplo (6.3.9), el lector podrá comprobar que

$$A^2 - 3A + 2\,I = 0\,,$$

o sea, $p(A) = 0$ para

$$p(X) = X^2 - 3X + 2 = (X-1)(X-2)\,.$$

Para llegar a esta conclusión no es siquiera necesario realizar el cálculo de $A^2 - 3A + 2\,I$. Basta recordar de (6.3.9) que A es semejante a

$$D = \begin{bmatrix} 1. & & \\ & 1. & \\ & & 2. \end{bmatrix}\,,$$

observar que $p(D) = (D - I)(D - 2\,I) = 0$, y utilizar el hecho de que $p(A)$ y $p(D)$ son semejantes.

Nótese que el polinomio p es un divisor del polinomio característico de A ya que

$$p_A(X) = -(X-1)\,p(X)\,.$$

En consecuencia

$$p_A(A) = -(A-I)\,p(A) = 0\,.$$

Probaremos ahora que esto es lo que ocurre para todas las matrices.

6.5.6 TEOREMA (de Cayley-Hamilton[1])

Toda matriz $A \in M_{\mathbb{K}}(n)$ ($\mathbb{K} = \mathbb{R}$ o $\mathbb{C}$) es raíz de su polinomio característico, es decir, $p_A(A) = 0$.

Todo endomorfismo $f \in \mathcal{L}(E)$, de un espacio finito-dimensional real o complejo, cumple que $p_f(f) = 0$, donde p_f es el polinomio característico de f.

La demostración para el caso real es indirecta y se basa en el caso complejo. Comenzaremos entonces por este último, suponiendo que f es un endomorfismo del espacio complejo E de dimensión $n \neq 0$. La prueba servirá al mismo tiempo para el enunciado con una matriz compleja A, denotando por f en ese caso el endomorfismo de $\mathbb{C}^n$ dado por A.

[1] Así llamado en honor del matemático inglés Arthur Cayley (1821-1895) y del astrónomo y matemático irlandés William Rowan Hamilton (1805-1865).

Sea $p_f(X)$ el polinomio característico de f y $\lambda_1, \lambda_2, \ldots, \lambda_n$ sus raíces (no necesariamente distintas). Entonces

$$p_f(X) = (-1)^n (X - \lambda_1)(X - \lambda_2) \cdots (X - \lambda_n) .$$

Por lo tanto (v. 3.4.8),

$$p_f(f) = (-1)^n (f - \lambda_1\, id_E) \circ (f - \lambda_2\, id_E) \circ \cdots \circ (f - \lambda_n\, id_E) ,$$

lo que abreviaremos, poniendo $g_i = f - \lambda_i\, id_E$, como

$$p_f(f) = (-1)^n g_1 \circ g_2 \circ \cdots \circ g_n .$$

Los endomorfismos g_i conmutan entre sí, puesto que todos ellos son polinomios de f. (Además, esta conmutatividad se puede comprobar sin dificultad en este caso concreto.) Denotemos por $(a_1, a_2, \ldots, a_n)$ una base de E en la que la matriz de f sea triangular superior (v. 6.4.3), o sea,

$$[f, (a_i)] = \begin{bmatrix} \lambda_1 & \alpha_2^1 & \cdots & \alpha_n^1 \\ & \lambda_2 & \cdots & \alpha_n^2 \\ & & \ddots & \vdots \\ & & & \lambda_n \end{bmatrix} .$$

Vamos a probar por inducción (v. 1.4.12) que, para todo $k = 1, 2, \ldots, n$, el endomorfismo $g_1 \circ g_2 \circ \cdots \circ g_k$ se anula en los vectores $a_1, a_2, \ldots, a_k$. En primer lugar

$$g_1(a_1) = (f - \lambda_1\, id)(a_1) = f(a_1) - \lambda_1 a_1 = \lambda_1 a_1 - \lambda_1 a_1 = 0 .$$

Si ahora suponemos que $g_1 \circ g_2 \circ \cdots \circ g_k$ (con $k \leq n-1$) se anula en $a_1, a_2, \ldots, a_k$, entonces, para $i = 1, 2, \ldots, k$,

$$g_1 \circ \cdots \circ g_k \circ g_{k+1}(a_i) = g_{k+1} \circ g_1 \circ \cdots \circ g_k(a_i) = g_{k+1}(0) = 0 ,$$

mientras que para a_{k+1} , teniendo en cuenta que $f(a_{k+1}) = \alpha_{k+1}^1 a_1 + \cdots + \alpha_{k+1}^k a_k + \lambda_{k+1} a_{k+1}$, se tiene también que

$$\begin{aligned} g_1 \circ \cdots \circ g_k \circ g_{k+1}(a_{k+1}) &= g_1 \circ \cdots \circ g_k(\, (f - \lambda_{k+1}\, id)(a_{k+1})\,) \\ &= g_1 \circ \cdots \circ g_k(f(a_{k+1}) - \lambda_{k+1} a_{k+1}) \\ &= g_1 \circ \cdots \circ g_k(\alpha_{k+1}^1 a_1 + \cdots + \alpha_{k+1}^k a_k) \\ &= 0 . \end{aligned}$$

Esto termina la prueba por inducción.

Para $k = n$, lo que acabamos de ver significa que $g_1 \circ g_2 \circ \cdots \circ g_n$ se anula en todos los vectores de la base $(a_1, a_2, \ldots, a_n)$; lo mismo ocurre con el endomorfismo $p_f(f) = (-1)^n g_1 \circ g_2 \circ \cdots \circ g_n$, que, por lo tanto, es nulo. O sea, $p_f(f) = 0$.

Si A es una matriz compleja $n \times n$ y f es el endomorfismo de $\mathbb{C}^n$ dado por A, se tiene que $p_A(f) = p_f(f) = 0$, luego $p_A(A) = 0$.

Ocupémonos ahora del caso real. Si A es una matriz real $n \times n$, es también una matriz compleja y, por lo tanto, $p_A(A) = 0$. Finalmente, si f es un endomorfismo de un espacio real E y A es la matriz de f en una base cualquiera, tenemos que $p_f(A) = p_A(A) = 0$, luego $p_f(f) = 0$.

6.5.7 Obsérvese que, en el caso real, p_A y p_f son polinomios con coeficientes reales. Lo que impide hacer en el caso real la misma demostración directa del caso complejo es que p_A y p_f pueden no admitir una descomposición

$$p_f(X) = (-1)^n (X - \lambda_1)(X - \lambda_2) \cdots (X - \lambda_n)$$

con números λ_i reales, porque son polinomios que pueden tener raíces complejas no reales.

6.5.8 Al margen de la utilización que muy pronto haremos del teorema precedente, existen algunas aplicaciones inmediatas que es interesante citar.

Si A es una matriz $n \times n$, $p_A(A) = 0$, o sea,

$$(-1)^n A^n + (-1)^{n-1} \operatorname{tr} A \cdot A^{n-1} + \cdots + (\det A)\, I = 0\,.$$

De la igualdad anterior se puede despejar A^n, con lo que se obtiene A^n como un polinomio de A de grado igual o menor a $n-1$. Lo mismo se puede hacer a continuación con las potencias $A^{n+1}, A^{n+2}, \ldots$

Resulta así que cualquier potencia de A y, más generalmente, cualquier polinomio de A coincide con un polinomio de A de grado igual o menor que $n-1$.

6.5.9 **Ejemplo.** La matriz

$$A = \begin{bmatrix} 5. & 4. \\ -2. & -1. \end{bmatrix}$$

tiene por polinomio característico

$$p_A(X) = X^2 - 4X + 3$$

luego

$$A^2 - 4A + 3\, I = 0$$

y

$$A^2 = 4A - 3\, I\,.$$

En consecuencia, cualquier potencia de A es de la forma

$$A^n = \alpha_n A + \beta_n\, I\,.$$

Por ejemplo,

$$A^3 = A^2 A = (4A - 3\, I)A = 4A^2 - 3A = 4(4A - 3\, I) - 3A = 13A - 12\, I\,.$$

Es fácil obtener una relación de recurrencia entre los coeficientes α_n y β_n y los α_{n+1} y β_{n+1}. Como

$$\begin{aligned} A^{n+1} &= A^n A = (\alpha_n A + \beta_n\, I)A = \alpha_n A^2 + \beta_n A \\ &= \alpha_n(4A - 3\, I) + \beta_n A = (4\alpha_n + \beta_n)A - 3\alpha_n\, I\,, \end{aligned}$$

resulta que

$$\alpha_{n+1} = 4\alpha_n + \beta_n \quad \text{y} \quad \beta_{n+1} = -3\alpha_n \,.$$

Sabemos además que $\alpha_0 = 0$, $\beta_0 = 1$, $\alpha_1 = 1$ y $\beta_1 = 0$. Se puede probar fácilmente por inducción que, para todo n,

$$\alpha_n = \frac{3^n - 1}{2} \quad \text{y} \quad \beta_n = \frac{-3^n + 3}{2} \,.$$

Por lo tanto, para todo $n \in \mathbb{N}$,

$$A^n = \frac{3^n - 1}{2} A + \frac{-3^n + 3}{2} I \,,$$

o sea,

$$A^n = \begin{bmatrix} 2 \cdot 3^n - 1 & 2 \cdot 3^n - 2 \\ -3^n + 1 & -3^n + 2 \end{bmatrix} .$$

6.5.10 Supongamos ahora que A es una matriz $n \times n$ inversible; denotemos por $p_A(X) = \alpha_n X^n + \alpha_{n-1} X^{n-1} + \cdots + \alpha_1 X + \alpha_0$ el polinomio característico de A. Sabemos que

$$\alpha_n A^n + \alpha_{n-1} A^{n-1} + \cdots + \alpha_1 A + \alpha_0 \, I = 0 \,.$$

Multiplicando por A^{-1} obtenemos

$$\alpha_n A^{n-1} + \alpha_{n-1} A^{n-2} + \cdots + \alpha_1 \, I + \alpha_0 A^{-1} = 0$$

y, como $\alpha_0 = \det A \neq 0$, podemos despejar A^{-1}

$$A^{-1} = -\frac{1}{\alpha_0} (\alpha_n A^{n-1} + \alpha_{n-1} A^{n-2} + \cdots + \alpha_1 \, I) \,.$$

Resulta así la inversa de A como un polinomio de A de grado $n - 1$.

6.5.11 Ejemplo. Para la matriz

$$A = \begin{bmatrix} 2. & 2. & 1. \\ 4. & 4. & 1. \\ 1. & 0 & 0 \end{bmatrix}$$

(es la misma que en el ejemplo 3.4.36) el polinomio característico vale

$$p_A(X) = -X^3 + 6X^2 + X - 2 \,,$$

luego

$$-A^3 + 6A^2 + A - 2\, I = 0 \,,$$
$$-A^2 + 6A + I - 2A^{-1} = 0 \,,$$

y, por tanto,

$$A^{-1} = \frac{1}{2} (-A^2 + 6A + I) \,.$$

Se puede comprobar que, efectivamente, se obtiene así la inversa que ya fue calculada en (3.4.36).

6.5.12 Sea A una matriz $n \times n$. El polinomio característico $p_A(X)$ anula A, pero puede no ser el polinomio de menor grado que lo hace.

Por ejemplo, si A es diagonalizable y tiene valores propios repetidos se produce esta situación. Es lo que ocurre con la matriz de los ejemplos (6.3.9) y (6.5.5). Su polinomio característico es

$$p_A(X) = -X^3 + 4X^2 - 5X + 2 = (1-X)^2(2-X)$$

y p_A anula A, pero también anula A el polinomio

$$p(X) = X^2 - 3X + 2 = (1-X)(2-X)\,,$$

como ya vimos.

La descripción de los polinomios que anulan una matriz cuadrada A se puede resumir en los dos puntos siguientes:

— Existe un único polinomio (salvo multiplicación por constantes) de entre los que anulan A que posee grado mínimo. Se denomina *polinomio minimal* de A.

— Los polinomios que anulan A son exactamente los múltiplos del polinomio minimal.

A pesar de que no es particularmente difícil, no nos detendremos en justificar estas afirmaciones, ya que no vamos a utilizarlas en el posterior desarrollo del tema.

Naturalmente, si $p(X)$ es un polinomio minimal de A, también lo es $\alpha\, p(X)$ para cualquier $\alpha \neq 0$. A veces se acuerda tomar como polinomio minimal el que tiene coeficiente 1 en el término de mayor grado, pero, generalmente, la expresión 'polinomio minimal' sirve para designar a cualquiera de ellos.

Dos matrices semejantes poseen el mismo polinomio minimal, puesto que son anuladas por los mismos polinomios (v. 6.5.2).

Como el polinomio característico anula la matriz, resulta que es un múltiplo del polinomio minimal (coincidente a veces con él).

El principal inconveniente del polinomio minimal es la inexistencia de un método sencillo y general para el cálculo de este polinomio, al contrario de lo que ocurre con el polinomio característico.

En los casos sencillos se puede encontrar por tanteo, sabiendo que divide al polinomio característico y que posee sus mismas raíces (véase el ejercicio 14 de la sección 6.1). Es decir, si el polinomio característico de A es

$$p_A(X) = (\lambda_1 - X)^{k_1}(\lambda_2 - X)^{k_2} \cdots (\lambda_p - X)^{k_p}$$

con $\lambda_1, \lambda_2, \ldots, \lambda_p$ distintos, entonces el polinomio minimal de A es de la forma

$$p(X) = (\lambda_1 - X)^{l_1}(\lambda_2 - X)^{l_2} \cdots (\lambda_p - X)^{l_p}$$

con $1 \leq l_i \leq k_i$.

6.5.13 Ejemplo. La matriz de los ejemplos (6.3.9) y (6.5.5) tiene por polinomio característico

$$p_A(X) = -X^3 + 4X^2 - 5X + 2 = (1-X)^2(2-X)\,,$$

luego el polinomio minimal es o bien $p_A(X)$ o bien $p(X) = (1-X)(2-X)$. Como ya hemos comprobado, este último anula A, luego es el polinomio minimal de A.

6.5.14 Ejercicios.

1 Siendo $p(X)$ un polinomio y

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

una matriz cuadrada triangular por bloques, con bloques A y D cuadrados, compruébese que la matriz $p(M)$ tiene la forma

$$p(M) = \left[\begin{array}{c|c} p(A) & B' \\ \hline 0 & p(D) \end{array}\right].$$

2 Búsquese una fórmula que proporciones las sucesivas potencias de la matriz

$$A = \begin{bmatrix} 2. & 2. \\ -2. & -3. \end{bmatrix}$$

en función de las matrices A e I.

3 Utilícese el procedimiento basado en el teorema de Cayley-Hamilton (v. 6.5.10) para calcular la inversa de la matriz

$$\begin{bmatrix} 4. & 2. & 5. \\ 2. & 2. & -1. \\ 0 & -1. & 4. \end{bmatrix}.$$

4 a) Calcúlese el polinomio minimal de una matriz diagonal arbitraria. (Véase también el ejercicio 13 de la sección 6.6.)

b) Búsquense dos matrices que posean el mismo polinomio minimal y no sean semejantes.

5 Calcúlese el polinomio minimal de la matriz $n \times n$

$$\begin{bmatrix} \lambda & 1 & & & \\ & \lambda & \ddots & & \\ & & \ddots & \ddots & \\ & & & \lambda & 1 \\ & & & & \lambda \end{bmatrix}.$$

(Véase también el ejercicio 13 de la sección 6.6.)

6.6 Forma canónica de endomorfismos y matrices.

6.6.1 Lo que pretendemos en esta sección es conseguir representar los endomorfismos mediante matrices diagonales por bloques

$$A = \begin{bmatrix} \boxed{A_1} & & & \\ & \boxed{A_2} & & \\ & & \ddots & \\ & & & \boxed{A_p} \end{bmatrix},$$

en las que cada bloque diagonal tenga la forma

$$\begin{bmatrix} \lambda & 1 & & & \\ & \lambda & \ddots & & \\ & & \ddots & \ddots & \\ & & & \lambda & 1 \\ & & & & \lambda \end{bmatrix}.$$

Una matriz diagonal por bloques con bloques de este tipo se llama *matriz de Jordan*[2]. Las matrices de la forma

$$\begin{bmatrix} \lambda & 1 & & & \\ & \lambda & \ddots & & \\ & & \ddots & \ddots & \\ & & & \lambda & 1 \\ & & & & \lambda \end{bmatrix}$$

se denominan *cajas elementales de Jordan.*

Así pues, lo que llamaremos matriz de Jordan es una matriz diagonal por bloques cuyos bloques diagonales sean cajas elementales.

En (6.1.3) vimos un ejemplo de matriz de Jordan con dos cajas elementales en la diagonal.

Nótese que una matriz de Jordan es triangular superior. Además los únicos elementos que pueden no ser nulos son los diagonales y los de la superdiagonal (los α_i^j con $i = j + 1$); estos últimos serán 1 o 0 según correspondan al interior de una caja elemental o a la separación entre dos de estas cajas. Como las matrices de Jordan son triangulares, sus elementos diagonales serán los valores propios de la matriz.

Ya vimos en (6.1.2) la idea esencial para conseguir matrices diagonales por bloques; estudiaremos ahora la forma de conseguir que los bloques diagonales sean cajas elementales de Jordan.

[2] del ingeniero francés Camille Jordan (1838-1921).

6.6.2 Concentremos nuestra atención en una caja elemental con diagonal nula, es decir en una matriz $n \times n$ de la forma

$$A = \begin{bmatrix} 0 & 1 & & & \\ & 0 & \ddots & & \\ & & \ddots & \ddots & \\ & & & 0 & 1 \\ & & & & 0 \end{bmatrix}.$$

Si f es un endomorfismo que se representa por esta matriz, entonces $f^n = 0$, o sea $A^n = 0$. En efecto, si $[f, (a_i)] = A$,

$$f(a_1) = 0\,, \quad f(a_2) = a_1\,, \ldots, f(a_n) = a_{n-1}\,,$$

luego

$$f(a_1) = 0\,, \quad f^2(a_2) = 0\,, \ldots, f^n(a_n) = 0\,,$$

y, en consecuencia,

$$f^n(a_i) = 0\,, \quad i = 1, 2, \ldots, n\,.$$

Por lo tanto $f^n = 0$.

Nótese que la base $(a_1, \ldots, a_n)$ en la que $[f, (a_i)] = A$ es una base formada por las imágenes sucesivas

$$f^{n-1}(a_n)\,, \ldots, f^2(a_n), f(a_n), a_n\,,$$

del vector a_n.

6.6.3 DEFINICIÓN. Sea E un e.v. sobre $\mathbb{K}$ y $f \in \mathcal{L}(E)$; decimos que f es un endomorfismo *nilpotente* cuando $f^p = 0$ para algún $p \in \mathbb{N}$. Obviamente, si $f^p = 0$, entonces $f^{p+1} = f^{p+2} = \cdots = 0$. Si f es nilpotente, existe un único $p \in \mathbb{N}$ tal que

$$f^{p-1} \neq 0 \quad \text{y} \quad f^p = 0\,;$$

se dice entonces que f es *nilpotente de orden* p. Si $f^k = 0$, el orden de f es igual o menor que k.

Si A es una matriz cuadrada y $A^p = 0$ para algún $p \in \mathbb{N}$, se dice que A es una matriz *nilpotente*. Si p es tal que

$$A^{p-1} \neq 0 \quad \text{y} \quad A^p = 0\,,$$

se dice que A es *nilpotente de orden* p.

Cuando A representa a f en alguna base, A^k representa en la misma base a f^k, luego f es nilpotente si y sólo si lo es A, y, en ese caso, lo son del mismo orden.

6.6.4 Hemos visto en (6.6.2) que toda caja elemental $n \times n$ de diagonal nula (y todo endomorfismo representado por una matriz de ese tipo) es nilpotente; y no es difícil ver que lo es de orden n (v. 6.6.24).

La afirmación más o menos recíproca de ésta será el primer resultado importante de esta sección (v. 6.6.8). Vamos a anteponer algunos resultados sencillos que tendremos que utilizar.

6.6.5 PROPOSICIÓN

Si $f \in \mathcal{L}(E)$, $x \in E$, $f^{k-1}(x) \neq 0$ y $f^k(x) = 0$, el sistema

$$(x, f(x), \ldots, f^{k-1}(x))$$

es libre.

En efecto, los vectores $x, f(x), \ldots, f^{k-1}(x)$ son no nulos y si

$$\alpha_0 x + \alpha_1 f(x) + \cdots + \alpha_{k-1} f^{k-1}(x) = 0\,,$$

aplicando sucesivamente $f^{k-1}, f^{k-2}, \ldots, f$ a ambos miembros de la igualdad, se obtiene que $\alpha_0 = 0$, $\alpha_1 = 0$, etc.

6.6.6 PROPOSICIÓN

Si $f \in \mathcal{L}(E)$, entonces

$$\{0\} = \operatorname{Ker} f^0 \subset \operatorname{Ker} f \subset \operatorname{Ker} f^2 \subset \cdots$$

y

$$E = \operatorname{Im} f^0 \supset \operatorname{Im} f \supset \operatorname{Im} f^2 \supset \cdots$$

La demostración es un ejercicio sencillo. Para $\operatorname{Ker} f^0$ e $\operatorname{Im} f^0$, recuérdese que $f^0 = id_E$.

6.6.7 Si f es nilpotente de orden p se tiene además

$$\operatorname{Ker} f^{p-1} \subsetneqq \operatorname{Ker} f^p = E = \operatorname{Ker} f^{p+1} = \operatorname{Ker} f^{p+2} = \cdots$$

y

$$\operatorname{Im} f^{p-1} \supsetneqq \operatorname{Im} f^p = \{0\} = \operatorname{Im} f^{p+1} = \operatorname{Im} f^{p+2} = \cdots$$

6.6.8 Obtención de la matriz de Jordan de un endomorfismo nilpotente.

Sea E un espacio de dimensión $n \neq 0$ y $f \in \mathcal{L}(E)$ un endomorfismo nilpotente de orden p.

Pongamos $K_i = \operatorname{Ker} f^i$, $i = 0, 1, \ldots, p$; sabemos que

$$\{0\} = K_0 \subset K_1 \subset K_2 \subset \cdots \subset K_{p-1} \subset K_p = E\,.$$

Los números $n_i = \dim K_i$, $i = 0, 1, \ldots, p$ verifican que

$$0 = n_0 \leq n_1 \leq n_2 \leq \cdots \leq n_{p-1} \leq n_p = n\,;$$

sus diferencias $d_i = n_i - n_{i-1}$, $i = 1, \ldots, p$, cumplen $d_1 + d_2 + \cdots + d_p = n$. Nótese que $d_1 = n_1$. Por otra parte, como $f^{p-1} \neq 0$, $K_{p-1} \neq E$ y $n_{p-1} < n$, luego $d_p \geq 1$.

Consideremos un suplementario, G_p, de K_{p-1} en $K_p = E$, o sea,

$$E = K_p = G_p \oplus K_{p-1} \,;$$

G_p será de dimensión d_p. Tomemos una base $(a_1, \ldots, a_{d_p})$ de G_p. Los vectores

$$f(a_1), \ldots, f(a_{d_p})$$

pertenecen a K_{p-1}, forman un sistema libre y

$$\langle f(a_1), \ldots, f(a_{d_p}) \rangle \cap K_{p-2} = \{0\} \,.$$

La primera afirmación se comprueba sin dificultad. Para la segunda, supongamos que

$$\alpha_1 f(a_1) + \cdots + \alpha_{d_p} f(a_{d_p}) = 0 \,;$$

entonces

$$f^{p-1}(\alpha_1 a_1 + \cdots + \alpha_{d_p} a_{d_p}) = f^{p-2}(\alpha_1 f(a_1) + \cdots + \alpha_{d_p} f(a_{d_p})) = f^{p-2}(0) = 0 \,,$$

luego $\alpha_1 a_1 + \cdots + \alpha_{d_p} a_{d_p} \in K_{p-1}$ y entonces $\alpha_1 a_1 + \cdots + \alpha_{d_p} a_{d_p} = 0$, lo que significa que $\alpha_1 = \cdots = \alpha_{d_p} = 0$. Para probar la tercera de las afirmaciones, supongamos que

$$\alpha_1 f(a_1) + \cdots + \alpha_{d_p} f(a_{d_p}) = x \in K_{p-2} \,;$$

se obtiene como en el caso anterior que

$$f^{p-1}(\alpha_1 a_1 + \cdots + \alpha_{d_p} a_{d_p}) = 0$$

y que $\alpha_1 = \cdots = \alpha_{d_p} = 0$, lo que significa que $x = 0$.

Consideremos ahora el subespacio $\langle f(a_1), \ldots, f(a_{d_p}) \rangle \oplus K_{p-2}$ de K_{p-1}; nótese que forzosamente $d_p + n_{p-2} \leq n_{p-1}$, o sea, $d_p \leq d_{p-1}$. Denotemos por G_{p-1} un suplementario en K_{p-1} de dicho subespacio, es decir

$$K_{p-1} = G_{p-1} \oplus \langle f(a_1), \ldots, f(a_{d_p}) \rangle \oplus K_{p-2} \,.$$

Ahora es posible que $G_{p-1} = \{0\}$, si es que $d_p = d_{p-1}$. Tomemos una base $(a_{d_p+1}, \ldots, a_{d_{p-1}})$ de G_{p-1}, que estará formada por $d_{p-1} - d_p$ vectores. Los vectores

$$f^2(a_1), \ldots, f^2(a_{d_p}), f(a_{d_p+1}), \ldots, f(a_{d_{p-1}})$$

pertenecen a K_{p-2}, forman un sistema libre y

$$\langle f^2(a_1), \ldots, f(a_{d_{p-1}}) \rangle \cap K_{p-3} = \{0\} \,.$$

La comprobación de estos hechos es muy parecida a la que hemos detallado más arriba.

El proceso continúa de la misma manera hasta que se obtiene el último suplementario G_1,

$$K_1 = G_1 \oplus \langle f^{p-1}(a_1), \ldots, f(a_{d_2})\rangle$$

(recuérdese que $K_0 = \{0\}$), suplementario que tendrá dimensión $d_1 - d_2$. Se elige finalmente una base $(a_{d_2+1}, \ldots, a_{d_1})$ de G_1.

Consideremos ahora los $d_p + d_{p-1} + \cdots + d_2 + d_1 = n$ vectores que hemos ido obteniendo

$$\begin{array}{ccccccccccc}
a_1 & \cdots & a_{d_p} \\
f(a_1) & \cdots & f(a_{d_p}) & a_{d_p+1} & \cdots & a_{d_{p-1}} \\
\vdots & & \vdots & \vdots & & \vdots \\
f^{p-2}(a_1) & \cdots & f^{p-2}(a_{d_p}) & f^{p-3}(a_{d_p+1}) & \cdots & f^{p-3}(a_{d_{p-1}}) & \cdots & a_{d_2} \\
f^{p-1}(a_1) & \cdots & f^{p-1}(a_{d_p}) & f^{p-2}(a_{d_p+1}) & \cdots & f^{p-2}(a_{d_{p-1}}) & \cdots & f(a_{d_2}) & a_{d_2+1} & \cdots & a_{d_1}
\end{array}$$

organizados en p filas y en $d_1 = n_1$ columnas. En cada columna figura un vector y sus imágenes sucesivas; nótese que, en cada caso, la imagen del vector por la siguiente potencia de f ya es nula. El número de columnas de las diferentes alturas es d_p, $d_{p-1}-d_p$, $\ldots$, d_1-d_2; sabemos que $d_p \geq 1$, pero las otras cantidades pueden ser nulas todas ellas.

En este conjunto de n vectores, la última fila constituye una base de K_1; al reunirla con la penúltima se obtiene una base de K_2, y así sucesivamente. La reunión de todas ellas forma una base de $K_p = E$.

Los vectores $a_i, f(a_i), f^2(a_i), \ldots, f^k(a_i)$, de una misma columna, forman una base de un subespacio que es invariante para f. Si se considera la restricción de f a este subespacio, y la base formada por los vectores de la columna en el orden

$$(f^k(a_i), \ldots, f^2(a_i), f(a_i), a_i),$$

la matriz en esta base es una caja elemental de la forma

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

En consecuencia, si se toma la base de E que hemos obtenido, en el orden

$$(f^{p-1}(a_1), \ldots, f(a_1), a_1, \ldots, f(a_{d_2}), a_{d_2}, a_{d_2+1}, \ldots, a_{d_1}),$$

la matriz del endomorfismo nilpotente f resulta ser

$$\begin{bmatrix} \boxed{A_1} & & \\ & \ddots & \\ & & \boxed{A_{d_1}} \end{bmatrix},$$

diagonal por bloques, con bloques diagonales de tamaño decreciente que son cajas elementales con ceros en la diagonal.

Nótese que la base puede ser costosa de obtener, pero el número de las cajas y los tamaños de las cajas resultan de un simple estudio de los rangos de las potencias de f, que proporcionan las dimensiones n_i y las diferencias d_i.

6.6.9 Ejemplo. El endomorfismo f de $\mathbb{R}^4$ dado por la matriz

$$A = \begin{bmatrix} -1. & 1. & 0 & 0 \\ -1. & 0 & 1. & 0 \\ -1. & 0 & 1. & 0 \\ -1. & 0 & 1. & 0 \end{bmatrix}$$

es nilpotente de orden 3, ya que se tiene que

$$A^2 = \begin{bmatrix} 0 & -1. & 1. & 0 \\ 0 & -1. & 1. & 0 \\ 0 & -1. & 1. & 0 \\ 0 & -1. & 1. & 0 \end{bmatrix}$$

y que $A^3 = 0$. Como $\operatorname{rg} A = 2$ y $\operatorname{rg} A^2 = 1$, tenemos (utilizando las notaciones del apartado precedente) que

$$\begin{array}{cccc} n_0 = 0 & n_1 = 2 & n_2 = 3 & n_3 = 4 \\ & d_1 = 2 & d_2 = 1 & d_3 = 1 \ . \end{array}$$

La base de $\mathbb{R}^4$ que buscamos tendrá la estructura

$$\begin{array}{ll} a_1 & \\ f(a_1) & \\ f^2(a_1) & a_2 \, , \end{array}$$

y al ordenarla como

$$(f^2(a_1), f(a_1), a_1, a_2)$$

la matriz de f es esta base será

$$\left[\begin{array}{ccc|c} 0 & 1 & & \\ & 0 & 1 & \\ & & 0 & \\ \hline & & & 0 \end{array}\right] .$$

Busquemos ahora los vectores a_1 y a_2 apropiados. $K_2 = \operatorname{Ker} f^2$ es el subespacio de dimensión 3

$$K_2 = \{(x, y, z, t) \in \mathbb{R}^4 \mid -y + z = 0\} .$$

Basta tomar como a_1 cualquier vector no nulo que no pertenezca a K_2, ya que entonces

$$\mathbb{R}^4 = \langle a_1 \rangle \oplus K_2 .$$

Tomaremos, por ejemplo,

$$a_1 = (0,1,0,0)\,.$$

Este vector tiene por imagen

$$f(a_1) = (1,0,0,0)\,.$$

$K_1 = \operatorname{Ker} f$ es el subespacio de dimensión 2 con ecuación

$$\begin{array}{rcrcrcl} -x & + & y & & & = & 0 \\ -x & & & + & z & = & 0\ . \end{array}$$

En este nivel no hay que escoger ningún vector ya que

$$K_2 = \langle f(a_1)\rangle \oplus K_1\,,$$

como sabemos por el estudio de las dimensiones, aunque el lector puede comprobar directamente este hecho. La columna se completa con el vector

$$f^2(a_1) = f(f(a_1)) = (-1,-1,-1,-1)\,.$$

El vector a_2 es cualquier vector que forme con $f^2(a_1)$ una base de K_1 . Tomaremos, por ejemplo,

$$a_2 = (0,0,0,1)\,.$$

En la base

$$(\,(-1,-1,-1,-1),(1,0,0,0),(0,1,0,0),(0,0,0,1)\,)$$

el endomorfismo f tendrá la matriz que hemos descrito antes. El lector podrá comprobar que, para

$$P = \begin{bmatrix} -1 & 1. & 0 & 0 \\ -1 & 0 & 1. & 0 \\ -1 & 0 & 0 & 0 \\ -1 & 0 & 0 & 1. \end{bmatrix},$$

se tiene que

$$P^{-1}A\,P = \left[\begin{array}{ccc|c} 0 & 1 & & \\ & 0 & 1 & \\ & & 0 & \\ \hline & & & 0 \end{array}\right].$$

6.6.10 Obtención de la matriz de Jordan de un endomorfismo con un sólo valor propio de multiplicidad igual a la dimensión del espacio.

Sea E un espacio sobre $\mathbb{K}$ ($\mathbb{R}$ o $\mathbb{C}$) de dimensión $n \neq 0$ y $f \in \mathcal{L}(E)$ un endomorfismo. Supondremos que f posee un solo valor propio $\lambda \in \mathbb{K}$ de multiplicidad n. El polinomio característico de f es entonces

$$p_f(X) = (-1)^n (X-\lambda)^n\,.$$

De acuerdo con el teorema de Cayley-Hamilton se tiene que $p_f(f) = 0$, o sea,

$$(f - \lambda\, id_E)^n = 0\,.$$

El endomorfismo $g = f - \lambda\, id_E$ es nilpotente de orden igual o menor que n.

El método que hemos explicado en (6.6.8) permite entonces obtener una base $(a_1, \ldots, a_n)$ en la que la matriz de g es diagonal por bloques

$$[g, (a_i)] = \begin{bmatrix} \boxed{A_1} & & \\ & \ddots & \\ & & \boxed{A_p} \end{bmatrix},$$

con bloques diagonales A_i de la forma

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

Como $f = g + \lambda\, id_E$, es inmediato comprobar que la matriz de f en la misma base es diagonal por bloques,

$$[f, (a_i)] = \begin{bmatrix} \boxed{B_1} & & \\ & \ddots & \\ & & \boxed{B_p} \end{bmatrix},$$

con bloques diagonales B_i de la forma

$$\begin{bmatrix} \lambda & 1 & & & \\ & \lambda & \ddots & & \\ & & \ddots & \ddots & \\ & & & \lambda & 1 \\ & & & & \lambda \end{bmatrix}.$$

La matriz $[f, (a_i)]$ es pues una matriz de Jordan.

6.6.11 Ejemplo. El endomorfismo f de $\mathbb{R}^4$ dado por la matriz

$$A = \begin{bmatrix} 0 & 1. & 0 & 0 \\ -1. & 1. & 1. & 0 \\ -1. & 0 & 2. & 0 \\ -1. & 0 & 1. & 1. \end{bmatrix}$$

tiene como polinomio característico

$$p_f(X) = (X - 1)^4\,,$$

luego un solo valor propio, 1, de multiplicidad 4.

El endomorfismo $g = f - id$ viene dado por la matriz

$$B = \begin{bmatrix} -1. & 1. & 0 & 0 \\ -1. & 0 & 1. & 0 \\ -1. & 0 & 1. & 0 \\ -1. & 0 & 1. & 0 \end{bmatrix},$$

o sea, es el mismo endomorfismo nilpotente que fue estudiado en el ejemplo (6.6.9).

Para la base

$$(\,(-1,-1,-1,-1),(1,0,0,0),(0,1,0,0),(0,0,0,1)\,)$$

de $\mathbb{R}^4$, la matriz de g es

$$\left[\begin{array}{ccc|c} 0 & 1 & & \\ & 0 & 1 & \\ & & 0 & \\ \hline & & & 0 \end{array}\right],$$

luego la matriz de $f = g + id$ en esa misma base es

$$\left[\begin{array}{ccc|c} 1 & 1 & & \\ & 1 & 1 & \\ & & 1 & \\ \hline & & & 1 \end{array}\right].$$

6.6.12 Nos queda por ver lo que ocurre en el caso general. Para ello necesitamos un resultado esencial, que es el que presentamos en (6.6.14).

Sea E un espacio sobre $\mathbb{K}$ de dimensión $n \neq 0$ y $f \in \mathcal{L}(E)$ un endomorfismo. Supongamos que $p_f(X)$ posee n raíces en $\mathbb{K}$, iguales o distintas. Denotemos por $\lambda_1,\ldots,\lambda_p \in \mathbb{K}$ las raíces de p_f y por $k_1,\ldots,k_p$ sus multiplicidades, que sumarán $k_1 + \cdots + k_p = n$.

Si $V(\lambda_1),\ldots,V(\lambda_p)$ son los correspondientes subespacios propios, la suma directa (v. 6.1.16)

$$V(\lambda_1) \oplus \cdots \oplus V(\lambda_p)$$

puede ser todo E cuando f es diagonalizable (v. 6.3.6) pero, cuando f no es diagonalizable, es un subespacio diferente de E.

Lo que se puede hacer es substituir cada subespacio propio $V(\lambda_i)$, que es $\mathrm{Ker}(f - \lambda_i\, id_E)$, por alguno de los núcleos $\mathrm{Ker}(f - \lambda_i\, id_E)^r$, que son mayores (v. 6.6.6); se suele dar el nombre de *subespacios propios generalizados* a estos núcleos. Veremos que los que sirven son justamente los

$$F_i = \mathrm{Ker}(f - \lambda_i\, id_E)^{k_i}\,,$$

donde el exponente k_i es la multiplicidad de λ_i como raíz.

Comenzaremos por algún resultado sencillo, para terminar probando la proposición (6.6.14), que es el resultado de demostración más difícil de entre los que se proponen en esta sección.

Los subespacios generalizados, $\operatorname{Ker}(f-\lambda_i\, id_E)^r$, son invariantes para f. En efecto, si $x \in \operatorname{Ker}(f-\lambda_i\, id)^r$, entonces, teniendo en cuenta que f y $f-\lambda_i\, id$ conmutan, resulta para $f(x)$ que

$$(f-\lambda_i\, id)^r(f(x)) = (f-\lambda_i\, id)^r \circ f(x) = f \circ (f-\lambda_i\, id)^r(x) = f(0) = 0\,,$$

por lo que $f(x) \in \operatorname{Ker}(f-\lambda_i\, id)^r$.

El resultado que anunciamos ahora es una generalización de (6.2.13).

6.6.13 PROPOSICIÓN

Sea E un e.v. sobre $\mathbb{K}$ de dimensión $n \neq 0$, $f \in \mathcal{L}(E)$; sea $\lambda \in \mathbb{K}$ un valor propio de f. Denotemos por k el orden de multiplicidad de λ como raíz de $p_f(X)$. Si $r \in \mathbb{N}$, entonces el subespacio $\operatorname{Ker}(f-\lambda\, id_E)^r$ verifica que

$$1 \leq \dim \operatorname{Ker}(f-\lambda\, id_E)^r \leq k\,.$$

Pongamos $F = \operatorname{Ker}(f-\lambda\, id)^r$. Como

$$V(\lambda) = \operatorname{Ker}(f-\lambda\, id) \subset \operatorname{Ker}(f-\lambda\, id)^r = F$$

y λ es un valor propio, $F \neq \{0\}$ y $\dim F \geq 1$.

Denotemos por h la dimensión de F; vamos a probar que $h \leq k$. El endomorfismo inducido por f en F verifica que $(f-\lambda\, id)^r = 0$, a causa de la propia definición de F. Si g es el endomorfismo de F dado por $g = f-\lambda\, id$, g es nilpotente y existe una base $(a_1,\ldots,a_h)$ de F tal que

$$[g,(a_1,\ldots,a_h)] = \begin{bmatrix} 0 & \alpha_1 & & & \\ & 0 & \ddots & & \\ & & \ddots & \ddots & \\ & & & 0 & \alpha_{h-1} \\ & & & & 0 \end{bmatrix},$$

donde los α_i valen 0 o 1. Para esta misma base,

$$[f,(a_1,\ldots,a_h)] = \begin{bmatrix} \lambda & \alpha_1 & & & \\ & \lambda & \ddots & & \\ & & \ddots & \ddots & \\ & & & \lambda & \alpha_{h-1} \\ & & & & \lambda \end{bmatrix}.$$

Completemos $(a_1, \ldots, a_h)$ hasta formar una base $(a_1, \ldots, a_h, \ldots, a_n)$ de E. La matriz de f en esta base de E es de la forma

$$\left[\begin{array}{ccccc|c} \lambda & \alpha_1 & & & & \\ & \lambda & \ddots & & & \\ & & \ddots & \ddots & & A' \\ & & & \lambda & \alpha_{h-1} & \\ & & & & \lambda & \\ \hline & & 0 & & & A'' \end{array}\right],$$

por lo que, siguiendo el mismo razonamiento que en (6.2.13),

$$p_f(X) = (\lambda - X)^h q(X)\,,$$

lo que demuestra que k (multiplicidad de la raíz λ) es igual o mayor que h.

6.6.14 PROPOSICIÓN

Sea E un e.v. sobre $\mathbb{K}$ ($\mathbb{R}$ o $\mathbb{C}$) de dimensión $n \neq 0$ y $f \in \mathcal{L}(E)$. Supongamos que el polinomio característico $p_f(X)$ posee n raíces en $\mathbb{K}$, iguales o distintas. Denotemos por $\lambda_1, \ldots, \lambda_p$ las raíces diferentes de $p_f(X)$ y por $k_1, \ldots, k_p$ sus multiplicidades. Pongamos

$$F_i = \operatorname{Ker}(f - \lambda_i\, id_E)^{k_i}\,, \quad i = 1, \ldots, p\,.$$

Entonces se tiene que

$$\dim F_i = k_i\,, \quad i = 1, \ldots, p$$

y que los subespacios $F_1, \ldots, F_p$ son invariantes y verifican

$$E = F_1 \oplus \cdots \oplus F_p\,.$$

El polinomio característico de f será

$$p_f(X) = (-1)^n (X - \lambda_1)^{k_1} \cdots (X - \lambda_p)^{k_p}\,,$$

y el teorema de Cayley-Hamilton nos garantiza que $p_f(f) = 0$.

Recordemos la conmutatividad entre dos polinomios cualesquiera del endomorfismo f (v. 3.4.8); tendremos que utilizarla varias veces a lo largo de la demostración.

Denotemos por $p_i(X)$ los polinomios

$$p_i(X) = \frac{(-1)^n p_f(X)}{(X - \lambda_i)^{k_i}}\,, \quad i = 1, \ldots, p\,,$$

resultantes de suprimir el factor $(X-\lambda_i)^{k_i}$ en el polinomio característico. Se tiene para todo $i=1,\ldots,p$ que

$$(X-\lambda_i)^{k_i} p_i(X) = (-1)^n p_f(X)\,,$$

luego

$$\textbf{(1)} \qquad (f-\lambda_i\, id)^{k_i} \circ p_i(f) = 0\,.$$

Observemos también que, si $x \in F_i = \operatorname{Ker}(f-\lambda_i\, id)^{k_i}$, entonces

$$\textbf{(2)} \qquad p_j(f)(x) = 0\,, \quad \text{para } j \neq i\,,$$

ya que $p_j(f) = q(f) \circ (f-\lambda_i\, id)^{k_i}$, donde q es un polinomio.

Como $\lambda_1,\ldots,\lambda_p$ son diferentes, los polinomios $p_1(X),\ldots,p_p(X)$ son primos entre sí (v. 1.13.14); en consecuencia, existen polinomios $q_1(X),\ldots,q_p(X)$ tales que

$$q_1(X)p_1(X) + \cdots + q_p(X)p_p(X) = 1\,.$$

Resulta así que

$$\textbf{(3)} \qquad q_1(f)\circ p_1(f) + \cdots + q_p(f)\circ p_p(f) = id_E\,.$$

Comencemos por probar que

$$E = F_1 + \cdots + F_p\,.$$

Sea $x \in E$. Por la igualdad **(3)**

$$x = q_1(f)\circ p_1(f)(x) + \cdots + q_p(f)\circ p_p(f)(x) = x_1 + \cdots + x_p\,,$$

donde hemos puesto

$$x_i = q_i(f)\circ p_i(f)(x)\,, \quad i=1,\ldots,p\,.$$

La imagen de x_i por $(f-\lambda_i\, id)^{k_i}$ vale

$$\begin{aligned}
(f-\lambda_i\, id)^{k_i}(x_i) &= (f-\lambda_i\, id)^{k_i}\circ q_i(f)\circ p_i(f)(x)\\
&= q_i(f)\circ (f-\lambda_i\, id)^{k_i}\circ p_i(f)(x)\\
&= 0\,,
\end{aligned}$$

como consecuencia de **(1)**. Así pues

$$x_i \in \operatorname{Ker}(f-\lambda_i\, id)^{k_i} = F_i\,.$$

Esto demuestra que la suma de los subespacios F_i es E.

Para probar que la suma es directa bastará con que mostremos que si

$$0 = x_1 + \cdots + x_p$$

con $x_1 \in F_1, \ldots, x_p \in F_p$, necesariamente $x_1 = \cdots = x_p = 0$. Veamos que eso es lo que ocurre efectivamente. Tomemos uno cualquiera de los x_i y apliquemos el endomorfismo $p_i(f)$ a la igualdad $0 = x_1 + \cdots + x_p$; el resultado es

$$0 = p_i(f)(x_1) + \cdots + p_i(f)(x_i) + \cdots + p_i(f)(x_p)$$

y, utilizando **(2)** y la igualdad precedente, se obtiene que

$$0 = p_i(f)(x_i)\,.$$

Utilizando ahora **(3)**,

$$x_i = q_1(f) \circ p_1(f)(x_i) + \cdots + q_i(f) \circ p_i(f)(x_i) + \cdots + q_p(f) \circ p_p(f)(x_i) = 0\,,$$

puesto que todos los términos son 0 a consecuencia de **(2)** y de la igualdad que acabamos de obtener. Por lo tanto, la suma es directa.

Ya hemos visto en (6.6.12) que los subespacios F_i son invariantes.

Las dimensiones de los F_i verifican

$$\dim F_1 + \cdots + \dim F_p = n\,;$$

además $\dim F_i \leq k_i$ para todo i, y, por hipótesis, $k_1 + \cdots + k_p = n$. En consecuencia debe ser

$$\dim F_i = k_i$$

para todo i.

6.6.15 El que la dimensión de $\operatorname{Ker}(f - \lambda_i\, id)^{k_i}$ coincida con el orden k_i de multiplicidad de λ_i significa que F_i es el mayor de todos los subespacios propios generalizados $\operatorname{Ker}(f - \lambda_i\, id)^r$. Por lo tanto no se puede aumentar el subespacio aumentando la potencia r por encima de la multiplicidad de λ_i.

Por otra parte, es posible que F_i se obtenga como $\operatorname{Ker}(f - \lambda_i\, id)^r$ para $r < k_i$; de hecho es algo que sucede con frecuencia.

6.6.16 Obtención de la matriz de Jordan de un endomorfismo (caso general).

Sea E un espacio sobre $\mathbb{K}$ ($\mathbb{R}$ o $\mathbb{C}$) de dimensión $n \neq 0$ y $f \in \mathcal{L}(E)$ un endomorfismo. Supongamos que el polinomio característico $p_f(X)$ posee n raíces en $\mathbb{K}$, iguales o distintas. Denotemos por $\lambda_1, \ldots, \lambda_p$ las raíces diferentes de $p_f(X)$ y por $k_1, \ldots, k_p$ sus multiplicidades. Hemos visto que los subespacios

$$F_i = \operatorname{Ker}(f - \lambda_i\, id_E)^{k_i}\,, \quad i = 1, \ldots, p\,,$$

son invariantes, de dimensión k_i, y $E = F_1 \oplus \cdots \oplus F_p$.

Eligiendo una base de E que sea reunión de bases de los subespacios F_i, obtenemos para f una matriz diagonal por bloques,

$$A = \begin{bmatrix} \boxed{A_1} & & \\ & \ddots & \\ & & \boxed{A_p} \end{bmatrix}\,,$$

donde cada bloque A_i es la matriz del endomorfismo que f induce en F_i.

Ahora bien, en cada subespacio F_i podemos emplear el procedimiento de (6.6.8) para el endomorfismo $g = f - \lambda_i\, id$, que es nilpotente ya que $g^{k_i} = 0$ sobre F_i. Obtendremos así una base de F_i en la que la matriz de g será una matriz de Jordan con 0 en la diagonal, mientras que la matriz de f será igual pero con el valor λ_i en la diagonal.

Si son éstas las bases que se reúnen, cada uno de los bloques A_i es una matriz de Jordan con el valor λ_i en todos los lugares de la diagonal,

$$A = \begin{bmatrix} \begin{array}{ccccc|} \lambda_i & 1 & & & \\ & \lambda_i & \ddots & & \\ & & \ddots & \ddots & \\ & & & \lambda_i & 1 \\ & & & & \lambda_i \\ \hline \end{array} & & \\ & \ddots & \\ & & \boxed{\lambda_i} \end{bmatrix}.$$

Por lo tanto, la matriz A será una matriz de Jordan con los valores $\lambda_1, \ldots, \lambda_p$ en la diagonal.

Hemos obtenido así el resultado que enunciamos seguidamente:

6.6.17 TEOREMA

---

a) Sea E un espacio sobre $\mathbb{K}$ ($\mathbb{R}$ o $\mathbb{C}$) de dimensión $n \neq 0$ y $f \in \mathcal{L}(E)$. Para que exista una base $(a_1, \ldots, a_n)$ de E en la que la matriz $[f, (a_i)]$ sea una matriz de Jordan es condición necesaria y suficiente que se verifique

(d1) $p_f(X)$ posee n raíces en $\mathbb{K}$, iguales o distintas,

propiedad que siempre se verifica cuando $\mathbb{K} = \mathbb{C}$.

b) Sea $A \in M_{\mathbb{K}}(n)$ ($\mathbb{K} = \mathbb{R}$ o $\mathbb{C}$). Para que A sea semejante a una matriz de Jordan es condición necesaria y suficiente que se verifique

(d1) $p_A(X)$ posee n raíces en $\mathbb{K}$, iguales o distintas,

propiedad que siempre se verifica cuando $\mathbb{K} = \mathbb{C}$. Toda matriz cuadrada compleja es semejante a una matriz de Jordan.

---

a) Ya hemos visto que la condición es suficiente, a la vez que explicábamos un método para buscar la base y la matriz de Jordan.

La condición es también necesaria, puesto que las matrices de Jordan son matrices triangulares.

b) El resultado para una matriz se obtiene considerando el endomorfismo de $\mathbb{K}^n$ dado por la matriz.

6.6.18 DEFINICIÓN. Si J es una matriz de Jordan que representa al endomorfismo f en una base, se dice que J es la *forma canónica de Jordan* de f.

Si J es una matriz de Jordan semejante a la matriz A, se dice que J es la *forma canónica de Jordan* de A.

Todas las matrices complejas (y, por lo tanto, todas las matrices reales) admiten una forma canónica. Ahora bien, para una matriz real es posible que su forma canónica no sea real, cuando las raíces de su polinomio característico no son todas reales; es ese caso, la matriz de paso tampoco será real.

Si un endomorfismo, o una matriz, es diagonalizable, entonces la correspondiente matriz diagonal es su forma canónica. El proceso de construcción que hemos visto conduce en ese caso a la matriz diagonal (véase el ejercicio 14 de esta sección).

6.6.19 Ejemplo. El endomorfismo f de $\mathbb{R}^4$ dado por la matriz

$$A = \begin{bmatrix} 0 & -4. & 0 & -1. \\ 0 & 2. & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 4. & 8. & -12. & 4. \end{bmatrix}$$

tiene como polinomio característico

$$p_f(X) = X(X-2)^3\,,$$

con raíces 0 (de multiplicidad 1) y 2 (de multiplicidad 3). El rango de $A - 2\,I$ es 2, luego $\dim V(2) = 2$ y el endomorfismo no es diagonalizable. Pondremos

$$F_1 = \operatorname{Ker}(f - 2\,id)^3 \quad \text{y} \quad F_2 = \operatorname{Ker} f\,.$$

El endomorfismo $g = f - 2\,id$ viena dado por la matriz $B = A - 2\,I$,

$$B = \begin{bmatrix} -2. & -4. & 0 & -1. \\ 0 & 0 & 0 & 0 \\ 0 & 0 & -2. & 0 \\ 4. & 8. & -12. & 2. \end{bmatrix}.$$

Tenemos que

$$B^2 = \begin{bmatrix} 0 & 0 & 12. & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 4. & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \quad B^3 = \begin{bmatrix} 0 & 0 & -24. & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & -8. & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix};$$

por lo tanto, $F_1 = \operatorname{Ker} g^3 = \operatorname{Ker} g^2$ es el subespacio de dimensión 3 formado por los $(x, y, z, t) \in \mathbb{R}^4$ tales que $z = 0$. El endomorfismo g es nilpotente de orden 2 sobre F_1. Utilizando las notaciones de (6.6.8), el núcleo $K_1 = \operatorname{Ker} g$ tiene dimensión 2 y ecuación

$$\begin{aligned} 2x + 4y \quad + t &= 0 \\ z \quad\quad &= 0\,. \end{aligned}$$

En consecuencia,

$$\begin{array}{lll} n_0 = 0 & n_1 = 2 & n_2 = 3 \\ & d_1 = 2 & d_2 = 1 \end{array}$$

y la base de F_1 que buscamos tendrá la estructura

$$\begin{array}{ll} a_1 & \\ g(a_1) & a_2 \end{array}$$

y habrá que ordenarla como

$$(g(a_1), a_1, a_2)\,.$$

Como a_1 se puede tomar cualquier vector de F_1 que no esté en K_1 ; pondremos

$$a_1 = (0,0,0,1)\,.$$

Entonces

$$g(a_1) = (-1,0,0,2)\,.$$

Ahora basta elegir a_2 de manera que complete con $g(a_1)$ una base de K_1 ; tomaremos, por ejemplo,

$$a_2 = (2,-1,0,0)\,.$$

Esto termina el trabajo con F_1 ; la matriz del endomorfismo inducido por $g = f - 2\,id$ en F_1 , para la base $(g(a_1), a_1, a_2)$, es

$$\left[\begin{array}{cc|c} 0 & 1 & \\ & 0 & \\ \hline & & 0 \end{array}\right],$$

luego la matriz del endomorfismo indicido por f en F_1 , para esta misma base, es

$$\left[\begin{array}{cc|c} 2 & 1 & \\ & 2 & \\ \hline & & 2 \end{array}\right].$$

El subespacio $F_2 = \operatorname{Ker} f$ tiene dimensión 1 y ecuación

$$\begin{array}{rcl} t & = & 0 \\ y & = & 0 \\ x - 3z & = & 0\,. \end{array}$$

En este caso bastará tomar un vector cualquiera, no nulo, de F_2 . Por ejemplo, el vector $(3,0,1,0)$.

En la base de $\mathbb{R}^4$

$$(\,(-1,0,0,2),(0,0,0,1),(2,-1,0,0),(3,0,1,0)\,)\,,$$

reunión de las bases que hemos elegido en F_1 y en F_2 , la matriz de f es

$$\left[\begin{array}{cc|c|c} 2 & 1 & & \\ & 2 & & \\ \hline & & 2 & \\ \hline & & & 0 \end{array}\right],$$

como puede comprobar el lector.

6.6.20 Vamos a dar una alternativa al procedimiento que hemos desarrollado en el ejemplo precedente. No presenta ninguna diferencia de fondo con lo que ya hemos hecho; simplemente evita tener que trabajar constantemente con las ecuaciones de los subespacios $F_i = \mathrm{Ker}(f - \lambda_i\, id)^{k_i}$, lo que en ocasiones puede resultar molesto.

La variante consiste en reducir, en un primer paso, la matriz de f a la forma diagonal por bloques (con bloques cualesquiera). Para ello basta tomar una base del espacio que sea reunión de bases (cualesquiera) de los subespacios F_i. Se puede así obtener una matriz Q que transforma la matriz A de f en

$$Q^{-1} A\, Q = \begin{bmatrix} \boxed{A_1} & & & \\ & \boxed{A_2} & & \\ & & \ddots & \\ & & & \boxed{A_p} \end{bmatrix},$$

sin que las matrices A_i sean de ninguna forma específica.

En una segunda fase, lo que se hace es reducir cada A_i a la forma canónica, buscando una matriz Q_i, inversible y del mismo tamaño que A_i, tal que $A_i' = Q_i^{-1} A_i\, Q_i$ sea una matriz de Jordan. Para ello basta seguir las indicaciones de (6.6.10), puesto que la matriz A_i posee un solo valor propio.

Poniendo entonces

$$P = Q \begin{bmatrix} \boxed{Q_1} & & & \\ & \boxed{Q_2} & & \\ & & \ddots & \\ & & & \boxed{Q_p} \end{bmatrix},$$

se tiene que

$$P^{-1} A\, P = \begin{bmatrix} \boxed{A_1'} & & & \\ & \boxed{A_2'} & & \\ & & \ddots & \\ & & & \boxed{A_p'} \end{bmatrix},$$

como es fácil comprobar. La matriz obtenida ahora es la forma canónica de f y de A, y la matriz P proporciona la base de E en la que f se representa por dicha forma canónica.

6.6.21 Ejemplo. Veamos cómo se utilizan las recomendaciones del apartado precedente para el endomorfismo del ejemplo (6.6.19), que venía dado por la matriz

$$A = \begin{bmatrix} 0 & -4. & 0 & -1. \\ 0 & 2. & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 4. & 8. & -12. & 4. \end{bmatrix}.$$

El subespacio correspondiente al valor propio 2, $F_1 = \mathrm{Ker}(f - 2\, id)^3$, tiene por ecuación $z = 0$. El sistema

$$(\,(1,0,0,0),(0,1,0,0),(0,0,0,1)\,)$$

es una base de F_1.

El subespacio correspondiente al valor propio 0, $F_2 = \operatorname{Ker} f$, tiene por ecuación

$$\begin{aligned} t &= 0 \\ y \qquad\quad &= 0 \\ x \quad - 3z \qquad &= 0\,, \end{aligned}$$

y el vector $(3,0,1,0)$ forma una base de F_2.

Tomando entonces

$$Q = \begin{bmatrix} 1. & 0 & 0 & 3. \\ 0 & 1. & 0 & 0 \\ 0 & 0 & 0 & 1. \\ 0 & 0 & 1. & 0 \end{bmatrix}$$

se obtiene la matriz diagonal por bloques

$$Q^{-1} A\, Q = \left[\begin{array}{ccc|c} 0 & -4. & -1. & \\ 0 & 2. & 0 & \\ 4. & 8. & 4. & \\ \hline & & & 0 \end{array}\right] .$$

Nos ocuparemos ahora de la matriz

$$A_1 = \begin{bmatrix} 0 & -4. & -1. \\ 0 & 2. & 0 \\ 4. & 8. & 4. \end{bmatrix}$$

que tiene el único valor propio 2 con multiplicidad 3. Consideremos el endomorfismo nilpotente g de $\mathbb{R}^3$ dado por la matriz

$$B_1 = A_1 - 2\,I = \begin{bmatrix} -2. & -4. & -1. \\ 0 & 0 & 0 \\ 4. & 8. & 2. \end{bmatrix} .$$

Como $(B_1)^2 = 0$, g es nilpotente de orden 2. Con las notaciones habituales, tenemos que $K_1 = \operatorname{Ker} g$ es el subespacio de $\mathbb{R}^3$ de dimensión 2 y ecuación

$$2x + 4y + z = 0\,.$$

Las dimensiones que interesan y sus diferencias son

$$\begin{array}{ccc} n_0 = 0 & n_1 = 2 & n_2 = 3 \\ & d_1 = 2 & d_2 = 1 \end{array} ,$$

y la base de R^3 que buscamos tendrá la estructura

$$\begin{array}{cc} a_1 & \\ g(a_1) & a_2 \end{array}$$

y habrá que ordenarla como

$$(g(a_1), a_1, a_2)\,.$$

Como a_1 se puede tomar cualquier vector de $\mathbb{R}^3$ que no esté en K_1 ; pondremos

$$a_1 = (1,0,0)\,.$$

Entonces

$$g(a_1) = (-2,0,4)\,.$$

Ahora, elegimos a_2 de manera que complete con $g(a_1)$ una base de K_1 ; por ejemplo,

$$a_2 = (2,-1,0)\,.$$

Poniendo

$$Q_1 = \begin{bmatrix} -2. & 1. & 2. \\ 0 & 0 & -1. \\ 4. & 0 & 0 \end{bmatrix}$$

se obtiene que

$$Q_1^{-1} A_1 Q_1 = \left[\begin{array}{cc|c} 2 & 1 & 0 \\ 0 & 2 & 0 \\ \hline 0 & 0 & 2 \end{array}\right]\,.$$

Pongamos finalmente

$$\begin{aligned} P &= Q \left[\begin{array}{c|c} Q_1 & \\ \hline & 1 \end{array}\right] \\ &= \begin{bmatrix} 1. & 0 & 0 & 3. \\ 0 & 1. & 0 & 0 \\ 0 & 0 & 0 & 1. \\ 0 & 0 & 1. & 0 \end{bmatrix} \left[\begin{array}{ccc|c} -2. & 1. & 2. & \\ 0 & 0 & -1. & \\ 4. & 0 & 0 & \\ \hline & & & 1. \end{array}\right] \\ &= \begin{bmatrix} -2. & 1. & 2. & 3. \\ 0 & 0 & -1. & 0 \\ 0 & 0 & 0 & 1. \\ 4. & 0 & 0 & 0 \end{bmatrix}\,. \end{aligned}$$

Con esta matriz se obtiene que

$$P^{-1} A\, P = \left[\begin{array}{cc|c|c} 2 & 1 & & \\ & 2 & & \\ \hline & & \boxed{2} & \\ & & & \boxed{0} \end{array}\right]\,,$$

es decir, se obtiene la forma canónica de f, como en el ejemplo (6.6.19). Nótese que existen diferentes bases de $\mathbb{R}^4$ en las que f tiene su forma canónica. Por ejemplo, ahora hemos obtenido la base

$$(\,(-2,0,0,4), (1,0,0,0), (2,-1,0,0), (3,0,1,0)\,)\,,$$

diferente de la que obtuvimos en (6.6.19). Esto no depende del método empleado, sino de la elección que se haga en cada caso de los vectores arbitrarios.

6.6.22 Ejemplo. La matriz compleja

$$A = \begin{bmatrix} 2i & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}$$

tiene como polinomio característico

$$p_A(X) = -(X-1)(X-i)^2 \,,$$

con raíces 1 (simple) e i (doble). Denotaremos por f el endomorfismo de $\mathbb{C}^3$ dado por A y pondremos

$$F_1 = \mathrm{Ker}(f - id) \quad \text{y} \quad F_2 = \mathrm{Ker}(f - i \cdot id)^2 \,.$$

El vector $(0,1,0)$ forma una base de F_1.

F_2 es el subespacio

$$F_2 = \{(x,y,z) \in \mathbb{C}^3 \mid y = 0\} \,,$$

mientras que $K_1 = \mathrm{Ker}(f - i \cdot id)$ tiene dimensión 1 (por eso A no es diagonalizable) y ecuación

$$\begin{array}{rcl} ix \qquad + z &=& 0 \\ y \qquad &=& 0 \,. \end{array}$$

Denotemos por g el endomorfismo $g = f - i \cdot id$. La base de F_2 que interesa será de la forma $(g(a), a)$, donde a debe ser un vector de F_2 que no pertenezca a K_1. Tomaremos el vector $(1,0,0)$, cuya imagen por g es el vector $(i,0,1)$. En la base

$$(\,(0,1,0),(i,0,1),(1,0,0)\,)$$

de $\mathbb{C}^3$, f tendrá la forma canónica; o sea, para

$$P = \begin{bmatrix} 0 & i & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \,,$$

la forma canónica de A se obtendrá como $P^{-1}A\,P$. El lector comprobará que, efectivamente,

$$P^{-1}A\,P = \left[\begin{array}{c|cc} 1 & & \\ \hline & i & 1 \\ & & i \end{array}\right] .$$

6.6.23 Vamos a terminar esta sección estudiando la manera en que la forma canónica de Jordan sirve para el cálculo de polinomios de matrices (y, consecuentemente, de endomorfismos).

Sea A una matriz compleja $n \times n$, de la que denotaremos por J su forma de Jordan; sea $p(X)$ un polinomio con coeficientes complejos. Como $J = P^{-1}A\,P$ para una matriz inversible P, entonces

$$P(A) = P\,p(J)\,P^{-1}$$

(v. 6.5.2). Esto permite calcular $p(A)$ a través de $p(J)$.

La ventaja estriba en que, como J tiene una estructura sencilla, sus potencias son más fáciles de calcular, y lo mismo ocurre con los polinomios. Pero la sencillez de J no proviene sólo de que posee bastantes ceros, sino del tipo y disposición de sus elementos no nulos; esto permite realizar el cálculo de $p(J)$ aplicando una fórmula sencilla que vamos a obtener.

Si J es una matriz diagonal

$$J = \begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{bmatrix},$$

entonces

$$p(J) = \begin{bmatrix} p(\lambda_1) & & & \\ & p(\lambda_2) & & \\ & & \ddots & \\ & & & p(\lambda_n) \end{bmatrix},$$

como es fácil comprobar. (Para todas las comprobaciones de este tipo conviene empezar con potencias y pasar luego al caso de polinomios cualesquiera.)

Cuando J no es diagonal, la cosa es algo menos sencilla. Hay que tener en cuenta que

$$J = \begin{bmatrix} \boxed{J_1} & & & \\ & \boxed{J_2} & & \\ & & \ddots & \\ & & & \boxed{J_k} \end{bmatrix},$$

donde los bloques J_i son cajas elementales de Jordan. Por lo tanto (v. 6.5.3)

$$p(J) = \begin{bmatrix} \boxed{p(J_1)} & & & \\ & \boxed{p(J_2)} & & \\ & & \ddots & \\ & & & \boxed{p(J_k)} \end{bmatrix},$$

y la dificultad queda reducida a saber calcular $p(J)$ cuando J es una caja elemental.

6.6.24 Si

$$H = \begin{bmatrix} 0 & 1 & & & \\ & 0 & \ddots & & \\ & & \ddots & \ddots & \\ & & & 0 & 1 \\ & & & & 0 \end{bmatrix}$$

es una caja elemental $n \times n$ con la particularidad de tener ceros en la diagonal, es fácil ver que sus potencias vienen dadas por

$$H^k = \begin{bmatrix} 0 & \cdots & 1 & & & \\ & 0 & \cdots & 1 & & \\ & & \ddots & & \ddots & \\ & & & \ddots & & 1 \\ & & & & \ddots & \vdots \\ & & & & & 1 \end{bmatrix},$$

es decir, que su término general α_i^j es 1 cuando $i = j + k$ y 0 en los restantes casos. La idea para probar esto se puede ver en (6.6.2). En particular,

$$H^{n-1} = \begin{bmatrix} 0 & \cdots & \cdots & 1 \\ & \ddots & & \vdots \\ & & \ddots & \vdots \\ & & & 0 \end{bmatrix}$$

y $H^n = H^{n+1} = \cdots = 0$.

6.6.25 Sea ahora

$$J = \begin{bmatrix} \lambda & 1 & & & \\ & \lambda & \ddots & & \\ & & \ddots & \ddots & \\ & & & \lambda & 1 \\ & & & & \lambda \end{bmatrix}$$

una caja elemental cualquiera, de tamaño $n \times n$, y sea $p(X) = \alpha_0 + \alpha_1 X + \cdots + \alpha_r X^r$ un polinomio. La *fórmula de Taylor*[3] significa para este polinomio que

$$p(X) = p(\lambda) + \frac{p'(\lambda)}{1!}(X-\lambda) + \frac{p''(\lambda)}{2!}(X-\lambda)^2 + \cdots + \frac{p^{r)}(\lambda)}{r!}(X-\lambda)^r ,$$

donde $p', p'', \ldots, p^{r)}$ representan las derivadas sucesivas de p. Entonces

$$p(J) = p(\lambda)\, I + \frac{p'(\lambda)}{1!}(J - \lambda\, I) + \frac{p''(\lambda)}{2!}(J-\lambda\, I)^2 + \cdots + \frac{p^{r)}(\lambda)}{r!}(J - \lambda\, I)^r .$$

[3] del matemático inglés Brook Taylor (1685-1731); la fórmula será sin duda conocida del lector en su versión con resto, pero le hacemos notar que, en el caso de un polinomio, se trata de una fórmula exacta siempre que se llegue en el desarrollo por lo menos hasta el grado del polinomio.

Basta ahora tener en cuenta cómo son las potencias de $H = J - \lambda I$, que han sido descritas en (6.6.24), para obtener la fórmula

$$p(J) = \begin{bmatrix} p(\lambda) & p'(\lambda) & \dfrac{p''(\lambda)}{2!} & \cdots & \cdots & \dfrac{p^{n-1)}(\lambda)}{(n-1)!} \\ & p(\lambda) & p'(\lambda) & \cdots & \cdots & \dfrac{p^{n-2)}(\lambda)}{(n-2)!} \\ & & \ddots & \ddots & & \vdots \\ & & & \ddots & \ddots & \vdots \\ & & & & \ddots & p'(\lambda) \\ & & & & & p(\lambda) \end{bmatrix}.$$

6.6.26 Ejemplo. Calculemos $I + A - A^8$ para la matriz

$$A = \begin{bmatrix} 0 & 1. & 0 & 0 \\ -1. & 1. & 1. & 0 \\ -1. & 0 & 2. & 0 \\ -1. & 0 & 1. & 1. \end{bmatrix}$$

del ejemplo (6.6.11). Vimos en aquel ejemplo que para

$$P = \begin{bmatrix} -1. & 1. & 0 & 0 \\ -1. & 0 & 1. & 0 \\ -1. & 0 & 0 & 0 \\ -1. & 0 & 0 & 1. \end{bmatrix}$$

se tenía que

$$J = P^{-1} A\, P = \left[\begin{array}{ccc|c} 1 & 1 & & \\ & 1 & 1 & \\ & & 1 & \\ \hline & & & 1 \end{array}\right].$$

Pongamos $p(X) = 1 + X - X^8$,

$$J_1 = \begin{bmatrix} 1 & 1 & \\ & 1 & 1 \\ & & 1 \end{bmatrix} \quad \text{y} \quad J_2 = \begin{bmatrix} 1 \end{bmatrix}.$$

Los cálculos que hay que realizar son

$$p(J_1) = \begin{bmatrix} p(1) & p'(1) & p''(1)/2 \\ & p(1) & p'(1) \\ & & p(1) \end{bmatrix} = \begin{bmatrix} 1. & -7. & -28. \\ & 1. & -7. \\ & & 1. \end{bmatrix},$$

$$p(J_2) = \begin{bmatrix} p(1) \end{bmatrix} = \begin{bmatrix} 1. \end{bmatrix},$$

$$p(J)=\left[\begin{array}{c|c} p(J_1) & \\ \hline & p(J_2) \end{array}\right]=\left[\begin{array}{ccc|c} 1. & -7. & -28. & \\ & 1. & -7. & \\ & & 1. & \\ \hline & & & 1. \end{array}\right],$$

y, finalmente,

$$I+A-A^8=p(A)=P\,p(J)\,P^{-1}=\begin{bmatrix} 8. & 21. & -28. & 0 \\ 7. & 29. & -35. & 0 \\ 7. & 28. & -34. & 0 \\ 7. & 28. & -35. & 1. \end{bmatrix}.$$

6.6.27 La fórmula de (6.6.25) que proporciona $p(J)$ cuando p es un polinomio y J una caja elemental $n\times n$, tiene una consecuencia inmediata: si J es

$$J=\begin{bmatrix} \lambda & 1 & & & \\ & \lambda & \ddots & & \\ & & \ddots & \ddots & \\ & & & \lambda & 1 \\ & & & & \lambda \end{bmatrix}$$

y $p(X)$ y $q(X)$ son dos polinomios tales que

$$p(\lambda)=q(\lambda),p'(\lambda)=q'(\lambda),\ldots,p^{n-1)}(\lambda)=q^{n-1)}(\lambda),$$

entonces $p(J)=q(J)$.

Este resultado se puede extender a una matriz cualquiera. Sea A una matriz compleja $n\times n$, $\lambda_1,\ldots,\lambda_p$ los valores propios distintos de A,

$$p_A(X)=(-1)^n(X-\lambda_1)^{k_1}\cdots(X-\lambda_p)^{k_p}$$

su polinomio característico y $p(X)$ y $q(X)$ dos polinomios. Si coinciden los siguientes valores de los polinomios y sus derivadas

$$\begin{array}{l} p(\lambda_1)=q(\lambda_1),p'(\lambda_1)=q'(\lambda_1),\ldots,p^{k_1-1)}(\lambda_1)=q^{k_1-1)}(\lambda_1), \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ p(\lambda_p)=q(\lambda_p),p'(\lambda_p)=q'(\lambda_p),\ldots,p^{k_p-1)}(\lambda_p)=q^{k_p-1)}(\lambda_p), \end{array}$$

entonces $p(A)=q(A)$.

Para probar este hecho se puede razonar sobre la forma canónica de A con ayuda del resultado para las cajas elementales.

Hay, sin embargo, una demostración directa muy breve, que es la que vamos a exponer. Denotemos por $r(X)=p(X)-q(X)$ el polinomio diferencia de ambos. Como para todo $i=1,\ldots,p$

$$r(\lambda_i)=0,r'(\lambda_i)=0,\ldots,r^{k_i-1)}(\lambda_i)=0,$$

resulta de la fórmula de Taylor (v. 6.6.25) en potencias de $X - \lambda_i$ que

$$r(X) = (X - \lambda_i)^{k_i} s_i(X)\,,$$

donde $s_i(X)$ es un polinomio; esto, para todo $i = 1, \ldots, p$. En consecuencia, $r(X)$ es divisible por $(X - \lambda_1)^{k_1}$, por $(X - \lambda_2)^{k_2}$, etc.; por lo tanto

$$r(X) = (X - \lambda_1)^{k_1}(X - \lambda_2)^{k_2} \cdots (X - \lambda_p)^{k_p} s(X)\,,$$

donde $s(X)$ es un polinomio. Como, por el teorema de Cayley-Hamilton, $p_A(A) = 0$, resulta que

$$r(A) = (-1)^n p_A(A) s(A) = 0\,.$$

Así pues

$$p(A) - q(A) = r(A) = 0$$

y se obtiene la igualdad.

Nótese que si A es una matriz $n \times n$ con valores propios $\lambda_1, \ldots, \lambda_n$ todos simples, la condición suficiente para la igualdad de $p(A)$ y $q(A)$ se reduce a que

$$p(\lambda_1) = q(\lambda_1), \ldots, p(\lambda_n) = q(\lambda_n)\,.$$

En el fondo, las ideas expuestas en este apartado son otra versión (más cómoda en su forma actual) de lo que ya vimos en (6.5.8).

6.6.28 Ejemplo. Para calcular cualquier potencia de la matriz

$$A = \begin{bmatrix} 5. & 4. \\ -2. & -1. \end{bmatrix},$$

cuyos valores propios son 3 y 1, podemos buscar para la potencia A^n un polinomio de la forma $\alpha_n A + \beta_n I$ que coincida con A^n. Es suficiente para ello que los polinomios

$$p(X) = X^n \quad \text{y} \quad q(X) = \alpha_n X + \beta_n$$

tomen los mismos valores en 3 y en 1. Resulta sencillo ver que esto equivale a que

$$\alpha_n = \frac{3^n - 1}{2} \quad \text{y} \quad \beta_n = \frac{-3^n + 3}{2}\,.$$

Por lo tanto,

$$\begin{aligned} A^n &= \frac{3^n - 1}{2} A + \frac{-3^n + 3}{2} I \\ &= \begin{bmatrix} 2 \cdot 3^n - 1 & 2 \cdot 3^n - 2 \\ -3^n + 1 & -3^n + 2 \end{bmatrix}. \end{aligned}$$

Compárese este procedimiento con el que desarrollamos en el ejemplo (6.5.9).

6.6.29 Sólo un breve comentario final sobre la definición y el cálculo de funciones $h(A)$ de una matriz $n \times n$, A, cuando h es una función diferente de un polinomio. El ejemplo más importante entre las funciones de una matriz es e^A (que es $h(A)$ para la función exponencial $h(z) = e^z$), a causa de su utilidad en la resolución de sistemas diferenciales lineales.

La generalización más obvia de los polinomios son las series de potencias. El lector sabe que una serie de potencias

$$\sum_{n=0}^{\infty} \alpha_n z^n$$

no converge generalmente para cualquier valor de $z \in \mathbb{C}$. Los mismos problemas surgen cuando se pretende definir la matriz suma de la serie

$$\sum_{n=0}^{\infty} \alpha_n A^n$$

como límite de los polinomios

$$p_r(A) = \sum_{n=0}^{r} \alpha_n A^n$$

cuando r tiende a infinito.

El lector podrá ver en cursos posteriores que, en el caso de la exponencial, no existe problema alguno para definir

$$e^A = \sum_{n=0}^{\infty} \frac{1}{n!} A^n ,$$

como no lo hay tampoco para la convergencia de la serie

$$e^z = \sum_{n=0}^{\infty} \frac{z^n}{n!}$$

para todo $z \in \mathbb{C}$. La función e^A, definida como acabamos de decir, verifica buena parte de las propiedades de la exponencial de números complejos. Por ejemplo, si $AB = BA$, se tiene que

$$e^{A+B} = e^A e^B ;$$

también

$$e^0 = I ;$$

asimismo, e^A es inversible y

$$(e^A)^{-1} = e^{-A} .$$

Una de las maneras de calcular la exponencial de una matriz (y también otras funciones) es utilizar la forma canónica de la matriz y las fórmulas de (6.6.23) y (6.6.25), que son también válidas cuando p es, en lugar de un polinomio, la función $p(z) = e^z$.

También es válido para la exponencial de una matriz (y para otras funciones) el comentario del apartado (6.6.27). Es decir, $e^A = q(A)$ para todo polinomio q cuyo valor, y el de las correspondientes derivadas, coincida con los de la función e^z, y sus derivadas (recuérdese que la exponencial es su propia derivada), en los valores propios de A.

El objetivo de estas consideracioneses únicamente introducir al lector en el 'ambiente' de ciertas ideas que tendrá que estudiar en cursos posteriores. Naturalmente, no hemos justificado las distintas afirmaciones, ya que ello nos llevaría a introducir y estudiar la convergencia de matrices, tema que cae fuera del propósito de este libro.

6.6.30 Ejercicios.

1 Siendo f un endomorfismo, nilpotente de orden p, de un e.v. E, se considera para todo vector $x \neq 0$ el número $r \in \mathbb{N}$ para el que

$$f^{r-1}(x) \neq 0 \quad \text{y} \quad f^r(x) = 0\,.$$

Se dice que r es la *altura* del vector x. La altura del vector 0 es 0, por definición. Compruébese que la altura de cualquier vector es igual o menor que p. Pruébese que, si $x_1, \ldots, x_k$ son vectores no nulos y de alturas diferentes, el sistema $(x_1, \ldots, x_k)$ es libre.

2 Denotemos por D el endomorfismo de derivación de los polinomios sobre el espacio $\mathbb{R}_n[X]$ (véase el ejercicio 9 de la sección 2.1 y también el ejercicio 2 de la sección 6.1). Compruébese que D es nilpotente. Calcúlese la forma canónica de Jordan para D, y búsquese una base del espacio en la que D se represente por dicha forma canónica.

3 Búsquense matrices 3×3, con polinomio característico $(1 - X)^3$, para las que $A - I$ sea una matriz nilpotente de orden 1, 2 y 3.

4 Búsquese una base de $\mathbb{R}^5$ en la que el endomorfismo dado por la matriz

$$\begin{bmatrix} -1. & 3. & -5. & 2. & 0 \\ 0 & 1. & -3. & 1. & 0 \\ 0 & 1. & -2. & 0 & 0 \\ 0 & 0 & 1. & -2. & 0 \\ 0 & -1. & 3. & -2. & -1. \end{bmatrix}$$

se represente por una matrix de Jordan. Calcúlese también la matriz de Jordan que resulte.

5 Calcúlese la forma de Jordan de la matriz

$$\begin{bmatrix} 6. & 0 & 0 & 3. \\ 0 & 6. & 0 & 0 \\ 0 & 3. & 6. & 0 \\ 0 & 0 & 0 & 6. \end{bmatrix}$$

y la matriz de paso que permite obtenerla.

6 Calcúlese la forma canónica de Jordan de la matriz $n \times n$ $(n \geq 3)$

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

7 Pruébese que todas las matrices $n \times n$ de la forma

$$\begin{bmatrix} \lambda & \alpha_2^1 & \alpha_3^1 & \cdots & \cdots & \alpha_n^1 \\ & \lambda & \alpha_3^2 & \cdots & \cdots & \alpha_n^2 \\ & & \lambda & \ddots & \cdots & \alpha_n^3 \\ & & & \ddots & \ddots & \vdots \\ & & & & \lambda & \alpha_n^{n-1} \\ & & & & & \lambda \end{bmatrix},$$

con $\alpha_2^1, \alpha_3^2, \ldots, \alpha_n^{n-1}$ no nulos, poseen la misma forma canónica

$$\begin{bmatrix} \lambda & 1 & & & \\ & \lambda & \ddots & & \\ & & \ddots & \ddots & \\ & & & \lambda & 1 \\ & & & & \lambda \end{bmatrix}.$$

8 Sabiendo que el endomorfismo f se representa en una base $(a_1, \ldots, a_n)$ por la matriz

$$[f, (a_i)] = \begin{bmatrix} \lambda & & & & \\ 1 & \lambda & & & \\ & \ddots & \ddots & & \\ & & \ddots & \lambda & \\ & & & 1 & \lambda \end{bmatrix},$$

búsquese una base del espacio en la que f se represente por su forma canónica.

9 Pruébese que, si f es un endomorfismo de E y λ un escalar que no es valor propio, entonces los subespacios $\operatorname{Ker}(f - \lambda\, id_E)^r$ se reducen a $\{0\}$ para todo $r \in \mathbb{N}$.

10 Sea f un endomorfismo del e.v. E y μ un escalar cualquiera; pongamos $g = f - \mu\, id_E$. Pruébese que, si λ es un valor propio de f, entonces $\lambda - \mu$ es un valor propio de g; compruébese también que los subespacios $\operatorname{Ker}(f - \lambda\, id)^r$ y $\operatorname{Ker}(g - (\lambda - \mu)\, id)^r$ coinciden.

11 Sea f un automorfismo (un endomorfismo biyectivo) del e.v. E.

a) Compruébese que, para $\lambda \neq 0$, se tiene la igualdad

$$(-\lambda f) \circ (f^{-1} - (1/\lambda)\, id_E) = f - \lambda\, id_E\,.$$

b) Pruébese que, si λ es un valor propio de f, entonces $1/\lambda$ lo es de f^{-1}. Pruébese además que, para todo $r \in \mathbb{N}$, los subespacios propios generalizados $\operatorname{Ker}(f - \lambda\, id)^r$ y $\operatorname{Ker}(f^{-1} - (1/\lambda)\, id)^r$ coinciden.

12 Sea E un e.v. y $f, g \in \mathcal{L}(E)$ dos endomorfidmos tales que $g \circ f = f \circ g$. Pruébese que, para todo $\lambda \in \mathbb{K}$ y todo $r \in \mathbb{N}$ el subespacio $\operatorname{Ker}(f - \lambda\, id)^r$ (que ya era invariante para f) es también invariante para g.

13 Sea A una matriz compleja $n \times n$.

a) Pruébese que el polinomio minimal y el polinomio característico de A coinciden si y sólo si para cada valor propio de A existe una sola caja elemental diagonal en la forma canónica de A.

b) Pruébese que el polinomio minimal de A posee únicamente raíces simples si y sólo si A es diagonalizable.

14 Comprébese con un estudio de las dimensiones de los subespacios que, en el caso de una matriz diagonalizable, el proceso de obtención de la forma de Jordan lleva a una matriz diagonal.

15 Calcúlese la forma canónica de la matriz

$$\begin{bmatrix} 2. & 1. & 1. & 2. \\ 2. & 1. & -1. & 0 \\ 1. & 0 & 0 & -1. \\ 1. & -2. & 0 & 1. \end{bmatrix}$$

del ejemplo (6.4.5), y la matriz de paso que proporciona dicha forma canónica.

16 Calcúlese la forma canónica de la matriz

$$\begin{bmatrix} 6. & -10. & 0 \\ 3. & -5. & 0 \\ 1. & -2. & 1. \end{bmatrix}$$

y la matriz de paso que proporciona dicha forma canónica.

17 Calcúlese la forma canónica de la matriz

$$\begin{bmatrix} 3. & -1. & 0 & 1. \\ 9. & -5. & 0 & 9. \\ -3. & 2. & 1. & -3. \\ 4. & -3. & 0 & 6. \end{bmatrix}$$

y la matriz de paso que proporciona dicha forma canónica.

18 Calcúlese la forma canónica de

$$\begin{bmatrix} 0 & -1. & 1. & -1. & 1. \\ 2. & 3. & -1. & 1. & -1. \\ -2. & -1. & 3. & -1. & 1. \\ 3. & 3. & 0 & 3. & -1. \\ 6. & 9. & -6. & 7. & -5. \end{bmatrix}$$

y la matriz de paso que permite obtenerla.

19 Pruébese que, si J es la forma de Jordan de una matriz A, entonces $J - \lambda I$ es la forma de Jordan de la matriz $A - \lambda I$.

20 a) Siendo J la caja elemental de Jordan

$$J = \begin{bmatrix} \lambda & 1 & & & \\ & \lambda & \ddots & & \\ & & \ddots & \ddots & \\ & & & \lambda & 1 \\ & & & & \lambda \end{bmatrix},$$

con $\lambda \neq 0$, búsquese la forma canónica de la matriz J^{-1}. (*Sugerencia*: es muy sencillo con la ayuda del problema 7 de esta sección.)

b) Siendo J una matriz de Jordan con valores diagonales no nulos, búsquese la forma canónica de la matriz J^{-1}.

c) Siendo A una matriz inversible $n \times n$, descríbase la forma canónica de A^{-1} en función de la de A. (*Nota*: si bien se puede responder a esta apartado aprovechando los dos anteriores, tambiés se puede dar una solución sencilla usando el problema 11 de esta sección.)

21 a) Pruébese que toda caja elemental es semejante a su traspuesta.

b) Pruébese que toda matriz de Jordan es semejante a su traspuesta.

c) Pruébese que toda matriz cuadrada compleja es semejante a su traspuesta.

22 Sea $A \in M(n)$ una matriz cuadrada real o compleja y $\lambda_1, \ldots, \lambda_n$ las raíces complejas (iguales o distintas) de su polinomio característico. Pruébese que, para todo polinomio $p(X)$,

$$\operatorname{tr} p(A) = p(\lambda_1) + \cdots + p(\lambda_n) \quad \text{y} \quad \det p(A) = p(\lambda_1) \cdots p(\lambda_n)\,.$$

23 Descríbanse todas las matrices $n \times n$ que conmutan con la caja elemental

$$J = \begin{bmatrix} \lambda & 1 & & & \\ & \lambda & \ddots & & \\ & & \ddots & \ddots & \\ & & & \lambda & 1 \\ & & & & \lambda \end{bmatrix}.$$

Pruébese que dichas matrices son exactamente los polinomios de J.

24 Compruébese que, si $\lambda_1, \dots, \lambda_n$ son n números diferentes y $p(X)$ un polinomio, el polinomio

$$q(X) = \sum_{k=1}^{n} \left[\frac{p(\lambda_k)}{\prod\limits_{\substack{i=1 \\ i \neq k}}^{n} (\lambda_k - \lambda_i)} \prod_{\substack{i=1 \\ i \neq k}}^{n} (X - \lambda_i) \right]$$

(el símbolo $\prod$ sirve para abreviar un producto, como $\sum$ hace con las sumas) es de grado igual o menor que $n-1$, y vale lo mismo que $p(X)$ en $\lambda_1, \dots, \lambda_n$.

25 Para la matriz

$$A = \begin{bmatrix} 5. & 2. & -4. \\ 3. & 0 & -3. \\ 6. & 2. & -5. \end{bmatrix},$$

búsquese un polinomio $q(X)$, de grado igual o menor que 2, para el que $q(A) = I + A^7$. Calcúlese así la matriz $I + A^7$.

26 a) Siendo $\lambda_1, \lambda_2, \mu_1, \mu_2, \mu_3, \mu_4$ y μ_5 nùmeros reales o complejos arbitrarios, con $\lambda_1 \neq \lambda_2$, pruébese que, entre los polinomios de la forma

$$p(X) = \alpha_1 + \alpha_2(X - \lambda_1) + \alpha_3(X - \lambda_1)^2 + (X - \lambda_1)^3(\alpha_4 + \alpha_5(X - \lambda_2)) \,,$$

existe uno y sólo uno que cumple

$$\begin{aligned} p(\lambda_1) &= \mu_1 \,, \quad & p'(\lambda_1) &= \mu_2 \,, \quad & p''(\lambda_1) &= \mu_3 \,, \\ p(\lambda_2) &= \mu_4 \,, \quad & p'(\lambda_2) &= \mu_5 \,. \end{aligned}$$

b) Búsquese un polinomio $p(X)$, de grado igual o menor que 4, que cumpla

$$\begin{aligned} p(1) &= 1 \,, \quad & p'(1) &= 0 \,, \quad & p''(1) &= 2 \,, \\ p(2) &= 0 \,, \quad & p'(2) &= 1 \,. \end{aligned}$$

27 Calcúlense las potencias, A^n, de la matriz

$$A = \begin{bmatrix} 1 & 1 & -1 \\ 0 & 2-i & -2+2\,i \\ 0 & 1-i & -1+2\,i \end{bmatrix}.$$

28 Calcúlese la matriz e^A para

$$A = \begin{bmatrix} 2. & 1. \\ -1. & 0 \end{bmatrix},$$

a) mediante el cálculo de la forma canónica J de A y el posterior cálculo de e^J;
b) buscando un polinomio $p(X)$ tal que $p(A) = e^A$ y calculando entonces $p(A)$.

29 Siendo A una matriz cuadrada y $p(X)$ y $q(X)$ dos polinomios, se considera la función racional $r(X) = p(X)/q(X)$. Cuando $q(A)$ es inversible, el símbolo $r(A)$ representa $p(A)q(A)^{-1}$.

a) Compruébese la existencia de $r(A)$ y $r(B)$ para la función racional

$$r(X) = \frac{X^2+1}{X+1}$$

y las matrices

$$A = \begin{bmatrix} -1. & 4. \\ -1. & 3. \end{bmatrix} \quad \text{y} \quad B = \begin{bmatrix} 4. & 2. \\ -4. & -2. \end{bmatrix}$$

Calcúlese $r(A)$ y $r(B)$.

b) Compruébese la existencia de dos polinomios $a(X)$ y $b(X)$ tales que

$$r(A) = a(A) \quad \text{y} \quad r(B) = b(B)\,.$$