2

Forma canónica de Jordan (Versión 08-02-2018)

En este capítulo...

Motivación

La matriz
$$C = \begin{pmatrix} 0 & 2 & 2 \\ 2 & 0 & -1 \\ -1 & -1 & 0 \end{pmatrix}$$
no es diagonalizable (es el Ejercicio 1 de la Hoja 1 que vimos en clase). El polinomio característico es
$$p(\lambda) = (-\lambda - 2)(\lambda^2 - 2\lambda + 1) = (-\lambda - 2)(\lambda - 1)^2$$
y los autovalores son -2 y 1. Los subespacios propios de cada autovalor eran
$$\begin{cases} E_1(-2) = L[(1,-1,0)], \\ E_1(1) = L[(0,-1,1)], \end{cases}$$
y por tanto la matriz no es diagonalizable porque
$$\begin{cases} \mathrm{mult}_g(-2) = \mathrm{mult}_a(-2) = 1, \\ 1 = \mathrm{mult}_g(1) < \mathrm{mult}_a(1) = 2. \end{cases}$$

En esta sección vamos a ver que C es al menos semejante a una matriz que, aunque no es diagonal, es lo suficientemente sencilla como para poder usarla cómodamente. En este ejemplo concreto, veremos que la matriz C es semejante a la matriz
$$J = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix},$$
es decir, veremos que existe una matriz regular P tal que $C = PJP^{-1}$. La matriz J no es diagonal, pero por poco: solo tiene elementos distintos de cero en la diagonal principal, y en la diagonal justo por encima de ésta última, donde tenemos un uno. Este tipo de matrices (la definición formal la veremos en la siguiente sección) son las que van a jugar el papel que jugaban las matrices diagonales en el tema anterior. En este ejemplo concreto podemos decir que al menos la matriz J, a pesar de no ser diagonal, no es tan mala: no es difícil ver (por inducción) que
$$J^k = \begin{pmatrix} 1^k & k \cdot 1^{k-1} & 0 \\ 0 & 1^k & 0 \\ 0 & 0 & (-2)^k \end{pmatrix}$$
y que por tanto podemos calcular las potencias de C de una forma sencilla, ya que
$$C^k = (PJP^{-1})^k = PJ^kP^{-1}.$$

Ya veremos más adelante cómo encontrar esa matriz J a la cual va a ser semejante C (a priori no está claro), pero por ahora, en este ejemplo concreto, vamos a suponer que ya sabemos que la matriz J anterior es semejante a C. Nos vamos a concentrar en las siguientes líneas en encontrar la matriz P, eso nos va a dar una intuición de por donde van a ir los tiros en el futuro.

Si alguien nos asegura que la matriz C es semejante a J, entonces sabemos que lo que está ocurriendo es lo siguiente. Si denotamos por
$$f : \mathbb{R}^3 \to \mathbb{R}^3$$
al endomorfismo cuya matriz es C respecto de la base canónica, es decir, que f es
$$f(x,y,z) = (2y + 2z, 2x - z, -x - y),$$

entonces decir que C es semejante a J es lo mismo que decir que somos capaces de encontrar una base $B = \{v_1, v_2, v_3\}$ respecto de la cual la matriz de f es precisamente J.

Mirando la matriz J, ¿qué propiedades tienen los vectores v_1, v_2, v_3? Si la matriz de f respecto de la base B es J entonces sabemos que por definición las columnas de J son las coordenadas de $f(v_1), f(v_2), f(v_3)$ respecto de la base B respectivamente. Es decir,

$$\left\{ \begin{array}{lll} f(v_1) = & (1,0,0)_B & = v_1, \\ f(v_2) = & (1,1,0)_B & = v_1 + v_2, \\ f(v_3) = & (0,0,-2) & = -2v_3. \end{array} \right.$$

Es evidente que v_1 y v_3 deben ser autovectores de f asociados a los autovalores 1 y -2 respectivamente...¿pero qué propiedad especial tiene v_2? Si manipulamos un poco la expresión anterior llegamos a que

$$f(v_2) = v_1 + v_2 \Leftrightarrow f(v_2) - v_2 = v_1 \Leftrightarrow (f - I)(v_2) = v_1.$$

Es decir, que cuando calculo la imagen de v_2 por medio del endomorfismo $f - I$ me da precisamente v_1. Pero aún así, sigo sin saber de dónde sacar el vector v_2. Si seguimos manipulando un poco, como sabemos que $(f - I)(v_2) = v_1$, al aplicar el endomorfismo $f - I$ a los dos lados obtengo

$$(f - I)(v_2) = v_1 \Longrightarrow (f - I)[(f - I)(v_2)] = (f - I)(v_1) = 0$$

donde $[(f - I)(v_1) = 0$ porque ya sabía que v_1 era un autovector asociado al autovalor 1. Así que si denoto por

$$(f - I)^2$$

a la composición $(f - I) \circ (f - I)$, lo que estoy diciendo es que

$$(f - I)^2(v_2) = 0 \Longrightarrow v_2 \in \operatorname{Ker}[(f - I)^2].$$

Así que ya sé dónde buscar v_2, en el núcleo de $(f - I)^2$. La matriz de $(f - I)$ respecto de la base canónica es evidentemente $C - I$, así que la matriz de $(f - I)^2$ debe ser $(C - I)^2$ (porque la matriz de la composición de dos aplicaciones es el producto de las matrices). Por tanto, podemos ya calcular todos los vectores $(x, y, z) \in \mathbb{R}^3$ que pertenecen al núcleo de $(f - I)^2$. Se tiene que satisfacer

$$\underbrace{\begin{pmatrix} 3 & -6 & -6 \\ -3 & 6 & 6 \\ 0 & 0 & 0 \end{pmatrix}}_{(C-I)^2} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

es decir,

$$\left\{ \begin{array}{r} 3x - 6y - 6z = 0 \\ -3x + 6y + 6z = 0 \\ 0 = 0 \end{array} \right. \quad \rightarrow \quad \left\{ \begin{array}{l} x = 2\alpha + 2\beta \\ y = \alpha \\ z = \beta \end{array} \right.$$

y por tanto

$$\operatorname{Ker}((f - I)^2) = L[(2,1,0), (2,0,1)].$$

Nuestro vector v_2 debe pertenecer a $\operatorname{Ker}((f - I)^2) = L[(2,1,0), (2,0,1)]$, así que tomemos por ejemplo el vector

$$v_2 = (2,1,0).$$

La otra cosa que necesitábamos de v_2 era que cuando le aplicábamos el endomorfismo $(f-I)$ nos debía quedar un autovector v_1 asociado al autovalor 1. Si calculamos las coordenadas de $(f-I)(v_2)$ respecto de la base canónica nos queda

$$\underbrace{\begin{pmatrix} -1 & 2 & 2 \\ 2 & -1 & -1 \\ -1 & -1 & -1 \end{pmatrix}}_{(C-I)} \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 3 \\ -3 \end{pmatrix}$$

¿Es el vector $(0,3,-3)$ un autovector asociado al autovalor 1? Pues sí, porque ya sabíamos que $E_1(1) = L[(0,-1,1)]$. Así que podemos fijar $v_1 = (0,3,-3)$. ¡Qué suerte hemos tenido! (bueno, en realidad ya veremos que no ha sido suerte).

Así que finalmente, como al vector v_3 lo único que le teníamos que pedir es que fuese un autovector en $E_1(-2) = L[(1,-1,0)]$, podemos fijar la siguiente base de $\mathbb{R}^3$,

$$B = \{\underbrace{(0,3,-3)}_{v_1}, \underbrace{(2,1,0)}_{v_2}, \underbrace{(1,-1,0)}_{v_3}\}$$

Es fácil ver que son vectores linealmente independientes y que por tanto es realmente una base (de nuevo, ya veremos que esto tampoco ha sido una casualidad).

¿Cuál es la matriz de f respecto de la base B? Tenemos dos forma de calcularlo. Usando la matriz de f respecto de la canónica, para calcular la matriz del endomorfismo f en la base B debemos hacer el producto

$$P^{-1}AP$$

donde P es la matriz de cambio de base de B a la canónica, es decir,

$$P = \begin{pmatrix} 0 & 2 & 1 \\ 3 & 1 & -1 \\ -3 & 0 & 0 \end{pmatrix}.$$

Realizando el producto nos queda que

$$\underbrace{\begin{pmatrix} 0 & 2 & 1 \\ 3 & 1 & -1 \\ -3 & 0 & 0 \end{pmatrix}^{-1} \begin{pmatrix} 0 & 2 & 2 \\ 2 & 0 & -1 \\ -1 & -1 & -1 \end{pmatrix} \begin{pmatrix} 0 & 2 & 1 \\ 3 & 1 & -1 \\ -3 & 0 & 0 \end{pmatrix}}_{P^{-1}CP} = \underbrace{\begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix}}_{J}.$$

El hecho de que nos haya quedado la matriz J no es una casualidad, ha ocurrido precisamente porque los vectores v_1, v_2, v_3 los hemos buscado precisamente de forma que satisfagan

$$\left\{ \begin{array}{lll} f(v_1) = & (1,0,0)_B & = v_1, \\ f(v_2) = & (1,1,0)_B & = v_1 + v_2, \\ f(v_3) = & (0,0,-2) & = -2v_3. \end{array} \right.$$

de manera que si calculamos la matriz de f en la base B de la otra forma, es decir, poniendo como columnas las coordenadas de $f(v_1), f(v_2), f(v_3)$ en la base B nos queda precisamente la matriz J.

Observación ¿Podíamos haber tomado como v_2 cualquier vector de

$$\mathrm{Ker}((f-I)^2) = L[(2,1,0),(2,0,1)]?$$

No (pero casi). Si nos damos cuenta, el vector $(0,-1,1)$, que es el generador de $E_1(1)$, pertenece a $\mathrm{Ker}((f-I)^2) = L[(2,1,0),(2,0,1)]$ ya que

$$(0,-1,1) = -(2,1,0) + (2,0,1),$$

es decir, que $E_1(1) \subset \mathrm{Ker}((f-I)^2)$ (y de nuevo, esto no es una coincidencia). Si nos hubiese dado por escoger precisamente $v_2 = (0,-1,1)$ entonces habríamos tenido un problema, porque nosotros queríamos que $(f-I)(v_2)$ fuese un autovector v_1 asociado al autovalor 1 (pero no nulo, porque queremos usar v_1 como un vector de una base), y sin embargo

$$(f-I)(v_2) = 0$$

ya que el propio vector v_2 era un autovector asociado al autovalor 1. Como ya veremos más adelante, la única precaución que teníamos que haber tomado es precisamente la de escoger v_2 en $\mathrm{Ker}((f-I)^2)$ pero que *no* estuviese en $E_1(1)$.

Bloques de Jordan y matrices de Jordan

Definamos la matrices que van a jugar el papel que jugaban las matrices diagonales en el tema anterior.

■ **Definición** Un *bloque de Jordan* es una matriz cuadrada que tiene todos los elementos de la diagonal idénticos, la línea encima de la diagonal son unos, y el resto de entradas son ceros.

Ejemplo 2.1 Por ejemplo las matrices

$$\begin{pmatrix} 5 & 1 & 0 & 0 \\ 0 & 5 & 1 & 0 \\ 0 & 0 & 5 & 1 \\ 0 & 0 & 0 & 5 \end{pmatrix} \quad \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} \quad \begin{pmatrix} 2 \end{pmatrix}$$

son bloques de Jordan de órdenes 4, 3 y 1 respectivamente. Las matrices

$$\begin{pmatrix} 5 & 1 & 0 & 0 \\ 0 & 5 & 0 & 0 \\ 0 & 0 & 5 & 1 \\ 0 & 0 & 0 & 5 \end{pmatrix} \quad \begin{pmatrix} 5 & 1 & 0 & 0 \\ 0 & 5 & 1 & 0 \\ 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 5 \end{pmatrix}$$

no son bloques de Jordan.

■ **Definición** Una *matriz de Jordan* es una matriz cuadrada diagonal por bloques de forma que los bloques en la diagonal son de Jordan.

Ejemplo 2.2 La matriz

$$J = \left(\begin{array}{ccc:cc:cc:c} 5 & 1 & 0 & & & & & \\ 0 & 5 & 1 & & & & & \\ 0 & 0 & 5 & & & & & \\ \hdashline & & & 2 & 1 & & & \\ & & & 0 & 2 & & & \\ \hdashline & & & & & 2 & 1 & \\ & & & & & 0 & 2 & \\ \hdashline & & & & & & & -3 \end{array}\right)$$

donde los espacios en blanco representan matrices de ceros, es una matriz de Jordan compuesta por los bloques de Jordan

$$\begin{pmatrix} 5 & 1 & 0 \\ 0 & 5 & 1 \\ 0 & 0 & 5 \end{pmatrix} \quad \begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix} \quad \begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix} \quad \begin{pmatrix} -3 \end{pmatrix}$$

Ejemplo 2.3 Todas la matrices diagonales son matrices de Jordan. Por ejemplo, la matriz

$$J = \left(\begin{array}{c|c|c} 5 & 0 & 0 \\ \hline 0 & -3 & 0 \\ \hline 0 & 0 & 4 \end{array}\right)$$

es una matriz de Jordan con bloques de Jordan $\big(5\big)$, $\big(-3\big)$ y $\big(4\big)$ en la diagonal.

■ **Definición** Dada $A \in \mathcal{M}_n(\mathbb{K})$, si existe una matriz de Jordan $J \in \mathcal{M}_n(\mathbb{K})$ semejante a A, decimos que J es una *forma canónica de Jordan* de A.

Nuestro objetivo en este tema es probar el siguiente teorema:

■ **Teorema** (de Jordan) *Sea $f : V \to V$ un endomorfismo de un $\mathbb{K}$-espacio vectorial V. Supongamos que todas raíces $\lambda_1, \ldots, \lambda_r$ del polinomio característico de f pertenecen a $\mathbb{K}$. Entonces existe una base B de V respecto de la cual f tiene una matriz de Jordan.*

O dicho en otras palabras, dada una matriz $A \in \mathcal{M}_n(\mathbb{K})$, si todos sus autovalores pertenecen a $\mathbb{K}$ entonces A es semejante a una matriz de Jordan.

Subespacios propios generalizados

En esta sección V denota una $\mathbb{K}$-espacio vectorial.

■ **Definición** Sea $f : V \to V$ un endomorfismo y sea $\lambda \in \mathbb{K}$ un autovalor. Los subespacios propios generalizados de λ son

$$E_\ell(\lambda) := \operatorname{Ker}(f - \lambda I)^\ell$$

para cada $\ell \in \mathbb{N}$.

Observación 1) Por definición $E_1(\lambda)$ es el subespacio propio de λ.

2) Los subespacios propios generalizados forman una cadena

$$E_1(\lambda) \subset E_2(\lambda) \cdots \subset E_k(\lambda) \subset \cdots$$

En efecto, se tiene que dado $v \in E_k(\lambda)$ entonces $v \in E_{k+1}(\lambda)$ ya que

$$v \in E_k(\lambda) \Longrightarrow (f - \lambda I)^k(v) = 0 \Longrightarrow (f - \lambda I)(f - \lambda I)^k(v) = (f - \lambda I)(0) = 0$$

$$\Longrightarrow (f - \lambda I)^{k+1}(v) = 0,$$

y por tanto $v \in E_{k+1}(\lambda)$.

3) Esta cadena no puede ser infinita, es decir, a partir de un momento todos los subespacios propios generalizados deben ser iguales. Claro, cada $E_\ell(\lambda)$ es un subespacio de V y por tanto $\dim(E_\ell(\lambda)) \leq n$. Deducimos que sus dimensiones no pueden crecer indefinidamente, es decir que para algún ℓ vamos a tener que $\dim(E_\ell(\lambda)) = \dim(E_{\ell+1}(\lambda))$ y por tanto $E_\ell(\lambda) = E_{\ell+1}(\lambda)$. Es más, como $E_\ell(\lambda) = E_{\ell+1}(\lambda)$ vamos a tener que

$$E_\ell(\lambda) = E_{\ell+1}(\lambda) = E_{\ell+2}(\lambda) = \cdots$$

En efecto, si $v \in E_{\ell+2}(\lambda)$ entonces $0 = (f - \lambda I)^{k+2}(v) = (f - \lambda I)^{k+1} \circ (f - \lambda I)(v)$ y por tanto $(f - \lambda I)(v) \in E_{\ell+1}(\lambda) = E_\ell(\lambda)$. Es decir, $(f - \lambda I)^k \circ (f - \lambda I)(v) = (f - \lambda I)^{k+1}(v) = 0$, de donde deducimos que $v \in E^{k+1}(\lambda)$.

■ **Definición** Al número natural ℓ más pequeño tal que $E_\ell(\lambda) = E_{\ell+1}(\lambda)$ se le llama *índice de nilpotencia* de λ, se denota por $\mathrm{nil}(\lambda)$, y al subespacio propio generalizado $E^\ell(\lambda)$ se le llama *subespacio máximo* de λ y se denota por $M(\lambda) = E^\ell(\lambda)$.

Ejemplo 2.4 Sea el endomorfismo $f : \mathbb{R}^3 \to \mathbb{R}^3$ cuya matriz respecto de la base canónica es la matriz del ejercicio 1 de la Hoja 1,

$$C = \begin{pmatrix} 0 & 2 & 2 \\ 2 & 0 & -1 \\ -1 & -1 & 0 \end{pmatrix}$$

Ya sabemos que el polinomio característico es

$$p(\lambda) = (-\lambda - 2)(\lambda^2 - 2\lambda + 1) = (-\lambda - 2)(\lambda - 1)^2$$

y los autovalores son -2 y 1.

Calculemos primero los subespacios propios generalizados del autovalor $\lambda = -2$. Ya vimos en la sección 2.1 que

$$E_1(-2) = L[(1, -1, 0)].$$

Calculemos $E_2(-2)$. La matriz de $(f+2I)^2$ es $(C+2I)^2$ y por tanto debemos calcular todos los vectores $(x, y, z) \in \mathbb{R}^3$ que satisfacen

$$\underbrace{\begin{pmatrix} 6 & 6 & 6 \\ 9 & 9 & 0 \\ -6 & -6 & 3 \end{pmatrix}}_{(C+2I)^2} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

es decir,

$$\left\{ \begin{array}{rl} 6x + \ 6y + 6z &= 0 \\ 9x + \ 9y \qquad &= 0 \\ -6x + -6y + 3z &= 0 \end{array} \right. \to \left\{ \begin{array}{rl} x + \ y + z &= 0 \\ x + \ y \qquad &= 0 \\ -2x + -2y + z &= 0 \end{array} \right. \to \left\{ \begin{array}{rl} x + y + \ z &= 0 \\ + \quad -z &= 0 \\ + \ + \ 3z &= 0 \end{array} \right.$$

y por tanto

$$E_2(-2) = \mathrm{Ker}((f + 2I)^2) = L[(1, -1, 0)].$$

Como $E_1(-2) = E_2(-2)$ el índice de nilpotencia de $\lambda = -2$ es 1 y su subespacio maximal es $M(-2) = E_1(-2) = L[(1, -1, 0)]$.

Calculemos ahora los subespacios propios generalizados del autovalor $\lambda = 1$. Ya vimos en la Sección 2.1 que $E_1(1) = L[(0, -1, 1)]$ y $E_2(1) = \mathrm{Ker}((f - I)^2) = L[(2, 1, 0), (2, 0, 1)]$. Obsérvese que, como ya vimos en teóricamente, el subespacio $E_1(1)$ está contenido en el subespacio $E_2(1)$ dado que

$$(0, -1, 1) = -(2, 1, 0) + (2, 0, 1).$$

Calculemos ahora $E_3(1)$. La matriz de $(f - I)^3$ es $(C - I)^3$ y por tanto debemos calcular todos los vectores $(x, y, z) \in \mathbb{R}^3$ que satisfacen

$$\underbrace{\begin{pmatrix} -9 & 18 & 18 \\ 9 & -18 & -18 \\ 0 & 0 & 0 \end{pmatrix}}_{(C-I)^3} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

es decir,

$$\left\{ \begin{array}{r} -9x + 18y + 18z = 0 \\ 9x - 18y - 18z = 0 \\ 0 = 0 \end{array} \right. \to \left\{ \begin{array}{r} x - 2y - 2z = 0 \\ 0 = 0 \\ 0 = 0 \end{array} \right.$$

y por tanto

$$E_3(1) = \mathrm{Ker}((f-I)^3) = L[(2,1,0),(2,0,1)].$$

Como $E_2(1) = E_3(1)$ el índice de nilpotencia de $\lambda = 1$ es 2 y su subespacio maximal es $M(1) = E_2(1) = L[(2,1,0),(2,0,1)]$.

Los subespacios máximos van a jugar el papel que jugaban en el tema anterior los subespacios propios (es decir, el de los autovectores). De hecho, en esta ocasión vamos a probar que los subespacios máximos tienen *siempre* las siguientes buenas propiedades:

■ **Teorema** (de los subespacios máximos) *Sea V un $\mathbb{K}$-espacio vectorial y sea $f : V \to V$ un endomorfismo. Supongamos que todas raíces $\lambda_1, \ldots, \lambda_r$ del polinomio característico de f pertenecen a $\mathbb{K}$. Entonces se tiene que:*

a) $V = M(\lambda_1) \bigoplus \cdots \bigoplus M(\lambda_r)$.

b) $mult_a(\lambda_i) = \dim(M(\lambda_i))$ *para todo* $i = 1, \ldots, r$.

Lo primero que observamos es que si en el teorema anterior estamos trabajando sobre $\mathbb{K} = \mathbb{C}$ entonces la hipótesis de que todos los autovalores pertenecen a $\mathbb{C}$ se satisface siempre. Esa hipótesis solo supone una limitación cuando estamos trabajando sobre $\mathbb{R}$.

Probaremos este teorema en la Sección 2.6. Hasta entonces, lo que vamos a tratar de hacer es entender cómo se deduce el Teorema de Jordan del Teorema de los subespacios máximos. Es decir, dado un endomorfismo y supuesto como cierto el Teorema de los subespacios máximos, queremos construir una base respecto de la cual el endomorfismo tiene una matriz de Jordan.

Endomorfismos con un único autovalor

En esta sección vamos a empezar estudiando endomorfismos con un único autovalor. Aunque parezca un caso muy especial, pronto veremos que si sabemos cómo tratar este tipo de endomorfismo entonces sabremos cómo tratar cualquiera.

Recordemos que nuestro objetivo ahora es, suponiendo que el Teorema de los subespacios máximos es verdadero, entender cómo construir una base respecto de la cual nuestro endomorfismo tiene una matriz de Jordan. De hecho vamos a empezar al revés: vamos a suponer que ya tenemos una base respecto de la cual nuestro endomorfismo tiene una matriz de Jordan y vamos a deducir una serie de propiedades que dicha base debería tener.

Diagramas de puntos

Así pues, sea $f : V \to V$ un endomorfismo que tiene un único autovalor $\lambda \in \mathbb{K}$ (obsérvese que si $\mathbb{K} = \mathbb{R}$ entonces seguro que $\lambda \in \mathbb{R}$, porque si no fuese así, como el polinomio característico tiene coeficientes reales, entonces el conjugado de esa raíz también dería ser una raíz, es decir un autovalor, pero hemos supuesto que solo tenemos uno). Y supongamos que f tiene respecto de una cierta base $B = \{v_1, \ldots, v_n\}$ una matriz J de Jordan. Intercambiando la posición de los vectores, es evidente que podemos suponer que los bloques de Jordan que aparecen en J van de mayor a menor orden. Por ejemplo, la matriz

$$J = \left(\begin{array}{ccc|cc|cc|c}
\lambda & 1 & 0 & & & & & \\
0 & \lambda & 1 & & & & & \\
0 & 0 & \lambda & & & & & \\
\hline
 & & & \lambda & 1 & & & \\
 & & & 0 & \lambda & & & \\
\hline
 & & & & & \lambda & 1 & \\
 & & & & & 0 & \lambda & \\
\hline
 & & & & & & & \lambda
\end{array}\right)$$

A continuación vamos a mostrar cómo asociar un diagrama de puntos a la matriz J, de forma que nos permita entender mejor su comportamiento (y por tanto, el del endomorfismo f). Expliquemos el método a través de un ejemplo, por ejemplo, con la matriz J anterior. Es decir, que tenemos un endomorfismo $f : \mathbb{K}^8 \to \mathbb{K}^8$ y una base $B = \{v_1, \ldots, v_8\}$ respecto de la cual su matriz es J. El primer bloque nos indica que v_1 es un autovector, es decir $v_1 \in \operatorname{Ker}(f - \lambda I) = E_1(\lambda)$, y que se tienen las dos siguientes igualdades

$$\begin{cases} f(v_2) = v_1 + \lambda v_2 \Longleftrightarrow v_1 = (f - I\lambda)(v_2) \\ f(v_3) = v_2 + \lambda v_3 \Longleftrightarrow v_2 = (f - I\lambda)(v_3) \end{cases}$$

Es decir, que v_1 lo obtenemos aplicando al vector v_2 la aplicación $f - \lambda I$ y a su vez v_2 lo obtenemos aplicando al vector v_3 la aplicación $f - \lambda I$. Dicho esto, escribimos el siguiente diagrama de puntos

$$\bullet_{v_1} \quad \longleftarrow \quad \bullet_{v_2} \quad \longleftarrow \quad \bullet_{v_3}$$

Los puntos representan, de izquierda a derecha, los vectores v_1, v_2, v_3 y la flecha $\longleftarrow$ representa que el vector de la izquierda se calcula aplicando la función $f - \lambda I$ al vector de la derecha. Observemos que $v_2 \in E_2(\lambda) = \operatorname{Ker}(f - \lambda I)^2$ dado que

$$(f - I\lambda)^2(v_2) = (f - \lambda I)(f - \lambda I)(v_2) = (f - \lambda I)(v_1) = 0$$

pero $v_2 \notin E_1(\lambda) = \operatorname{Ker}(f - \lambda I)$, ya que en caso contrario se tendría que $v_1 = (f - \lambda I)(v_2) = 0$, lo cual es absurdo porque v_1 forma parte de una base. Razonando de forma parecida tenemos que $v_3 \in E_3(\lambda) = \operatorname{Ker}(f - \lambda I)^3$ pero $v_3 \notin E_2(\lambda) = \operatorname{Ker}(f - \lambda I)^2$. Esto lo podemos representar poniendo encima de cada punto

$$\begin{array}{ccccc} E_1(\lambda) & \hookrightarrow & E_2(\lambda) & \hookrightarrow & E_3(\lambda) \\ \bullet_{v_1} & \longleftarrow & \bullet_{v_2} & \longleftarrow & \bullet_{v_3} \end{array}$$

Sigamos construyendo el diagrama. Ahora pasamos al siguiente bloque de orden 2. Razonando de forma parecida a como hemos hecho anteriormente, tenemos que $v_4 \in E_1(\lambda)$ es un autovector y $v_4 = (f - \lambda I)(v_5)$. Y de nuevo $v_5 \in E_2(\lambda)$ pero $v_5 \notin E_1(\lambda)$. Para representar este nuevo bloque escribimos una línea de puntos debajo de la anterior,

$$\begin{array}{ccccc} E_1(\lambda) & \hookrightarrow & E_2(\lambda) & \hookrightarrow & E_3(\lambda) \\ \bullet_{v_1} & \longleftarrow & \bullet_{v_2} & \longleftarrow & \bullet_{v_3} \\ \bullet_{v_4} & \longleftarrow & \bullet_{v_5} & & \end{array}$$

Lo mismo hacemos con el último bloque de orden 2 y el bloque de orden 1, quedando finalmente el diagrama como

$$\begin{array}{ccccc} E_1(\lambda) & \hookrightarrow & E_2(\lambda) & \hookrightarrow & E_3(\lambda) \\ \bullet_{v_1} & \longleftarrow & \bullet_{v_2} & \longleftarrow & \bullet_{v_3} \\ \bullet_{v_4} & \longleftarrow & \bullet_{v_5} & & \\ \bullet_{v_6} & \longleftarrow & \bullet_{v_7} & & \\ \bullet_{v_8} & & & & \end{array}$$

De esta forma hemos llegado a un *diagrama de puntos*, tal que:

R_1) los puntos representan exactamente $\dim(V)$ vectores linealmente independientes de V, que son por tanto una base de V,

R_2) la flecha $\longleftarrow$ simboliza que el vector que representa el punto de la izquierda de la flecha se obtiene aplicando a $(f - \lambda I)$ al vector que representa el punto de la derecha de la flecha.

R_3) los puntos en la columna debajo de $E_i(\lambda)$ representan vectores que pertenecen a $E_i(\lambda)$ (en realidad solo es necesario para $i = 1$, para el resto vendrá dado por R_1 y R_2).

Y tan solo con estas dos propiedades hemos razonado también que tenemos una propiedad más fuerte que R_3:

> *Los puntos en la columna debajo de $E_i(\lambda)$ representan vectores que pertenecen a $E_i(\lambda)$, y ninguno de ellos pertenece a $E_{i-1}(\lambda)$.*

De hecho, trabajando un poquito más, podemos llegar a que:

> *Ninguna combinación lineal de los vectores que están representados por puntos que están debajo de un $E_i(\lambda)$ pertenece a $E_{i-1}(\lambda)$, excepto la combinación nula.*

En efecto, ya hemos argumentado que el punto que aparece en la columna de la derecha representa un vector en $E_3(\lambda)$ que *no* pertenece a $E_2(\lambda)$, y por tanto ningún múltiplo suyo (excepto el nulo) pertenecerá a $E_2(\lambda)$. Veamos ahora que en nuestro ejemplo los tres puntos que aparecen en la columna central representan tres vectores que pertenecen a $E_2(\lambda)$ con la propiedad de que ninguna combinación lineal de ellos no nula pertenece a $E_1(\lambda)$. Si existieran $a, b, c \in \mathbb{K}$ tales que

$$av_2 + bv_5 + cv_7 \in E_1(\lambda)$$

entonces

$$a(f - \lambda I)(v_2) + b(f - \lambda I)(v_5) + c(f - \lambda I)(v_7) = 0,$$

es decir

$$av_1 + bv_4 + cv_6 = 0,$$

de lo que deducimos $a = b = c = 0$ porque los vectores v_1, v_4, v_6 son linealmente independientes. Obsérvese que este argumento va a funcionar en general, porque dados los vectores representados por puntos en una columna debajo de $E_i(\lambda)$, si una combinación de ellos pertenece a $E_{i-1}(\lambda)$, entonces la combinación con los mismos coeficientes de sus imágenes por medio de $(f - \lambda I)^{i-1}$ nos va a dar 0, lo cual implica que los coeficientes son todos cero ya que esas imágenes sabíamos que eran linealmente independientes (porque son vectores representados por puntos debajo de $E_1(\lambda)$).

El papel que ha jugado la base B es ciertamente auxiliar: solo hemos utilizado las propiedades que los vectores de dicha base deben tener si f respecto de ella tiene matriz J, no nos importan quienes son los vectores en concreto. Así pues, escribiremos el diagrama anterior como

$$\begin{array}{ccccc}
E_1(\lambda) & \hookrightarrow & E_2(\lambda) & \hookrightarrow & E_3(\lambda) \\
\bullet & \longleftarrow & \bullet & \longleftarrow & \bullet \\
\bullet & \longleftarrow & \bullet & & \\
\bullet & \longleftarrow & \bullet & & \\
\bullet & & & &
\end{array}$$

Eso sí, puesto que cada punto corresponde a un vector de la base, lo único que sí deberíamos tener presente en adelante es cuál es el orden que ocupan dichos vectores en la base. La respuesta es fácil: debemos empezar por la primera línea y ordenar de izquierda a derecha, para luego continuar con la segunda línea, de izquierda a derecha, y así sucesivamente. Es decir, que debemos seguir el orden que usamos cuando leemos un libro.

El diagrama de puntos es útil porque nos permite calcular fácilmente las dimensiones de los subespacios propios generalizados. En efecto, si contamos los puntos es evidente que $\dim(E_3(\lambda)) = 8$. El punto que aparece en la columna de la derecha representa un vector que *no* pertenece a $E_2(\lambda)$, y dado que en $E_2(\lambda)$ tenemos al menos 7 vectores linealmente independientes, deducimos $\dim(E_2(\lambda)) = 7$. Vayamos con la dimensión de $E_1(\lambda)$. Ya sabemos que los vectores que representan los puntos en la columna debajo de $E_1(\lambda)$ son linealmente independientes y por tanto al menos sabemos que $\dim(E_1(\lambda)) \geq 4$. Queremos ver que de hecho se da la igualdad. También sabemos que los tres puntos que aparecen en la columna central representan tres vectores de $E_2(\lambda)$ con la propiedad de que ninguna combinación lineal de ellos pertenece a $E_1(\lambda)$. Denotemos con W al subespacio que generan esos tres vectores. Por tanto sabemos que $E_1(\lambda) \bigoplus W = E_2(\lambda)$, de lo que deducimos que

$$\dim(E_1(\lambda)) + \dim(W) = \dim(E_2(\lambda)) \quad \Rightarrow \quad \dim(E_1(\lambda)) + 3 = 4$$

es decir, $\dim(E_1(\lambda)) = 4$. Es decir, que *los puntos de las columnas sumados de izquierda a derecha nos da la dimensión de los subespacios generalizados*,

$$\dim(E_1(\lambda)) = 4, \qquad \dim(E_2(\lambda)) = 7, \qquad \dim(E_3(\lambda)) = 8.$$

Y esto de nuevo es general, solo hemos usado la propiedad de que ninguna combinación lineal de los vectores que están representados por puntos que están debajo de un $E_i(\lambda)$ pertenece a $E_{i-1}(\lambda)$, excepto la combinación nula. En particular, tenemos la siguiente consecuencia:

■ **Corolario** *Dado f un endomorfismo con un único autovalor, si f tiene matrices de Jordan respecto de dos bases distintas, entonces dichas matrices deben tener los mismos bloques de Jordan (salvo permutación).*

Demostración. Si ordenamos los bloques de mayor a menor de dicha matrices y hacemos sus diagramas nos deben quedar iguales. En caso contrario, tendríamos dos dimensiones distintas para algún subespacio propio generalizado de f, lo cual es absurdo. □

Pero incluso más importante para nosotros es que de igual manera que hemos asociado a una matriz de Jordan un diagrama de puntos, podemos asociar a un diagrama de puntos una matriz de Jordan. Así por ejemplo, si nos dicen que tenemos un endomorfismo $f : V \to V$, donde $\dim(V) = 8$, y que tenemos una base B de V de forma que asociando vectores a puntos podemos construir el siguiente diagrama con las reglas R_1,R_2,R_3 anteriores

$$\begin{array}{ccccc} E_1(\lambda) & \hookrightarrow & E_2(\lambda) & \hookrightarrow & E_3(\lambda) \\ \bullet & \longleftarrow & \bullet & \longleftarrow & \bullet \\ \bullet & \longleftarrow & \bullet & \longleftarrow & \bullet \\ \bullet & \longleftarrow & \bullet & & \end{array}$$

entonces f respecto de esa base B tiene matriz de Jordan

$$J = \left(\begin{array}{ccc|ccc|cc} \lambda & 1 & 0 & & & & & \\ 0 & \lambda & 1 & & & & & \\ 0 & 0 & \lambda & & & & & \\ \hline & & & \lambda & 1 & 0 & & \\ & & & 0 & \lambda & 1 & & \\ & & & 0 & 0 & \lambda & & \\ \hline & & & & & & \lambda & 1 \\ & & & & & & 0 & \lambda \end{array}\right)$$

En efecto, tenemos tantos bloques como líneas tiene nuestro diagrama, y el orden de cada bloque viene dado por el número de puntos en la línea. Además vemos que el número de puntos debajo de la columna $E_1(\lambda)$ debe ser igual a $\dim(E_1(\lambda))$, y el número de puntos debajo de la columna $E_i(\lambda)$ debe ser igual $\dim(E_{i-1}(\lambda)) - \dim(E_i(\lambda))$ para $i > 1$.

Con esta idea en la cabeza, veamos cómo construir una base B de cualquier endomorfismo $f : V \to V$ con un único autovalor respecto de la cual la matriz de f es una matriz de Jordan.

Teorema de Jordan de endomorfismos con un único autovalor

Recordar que estamos dando por cierto el Teorema de los subespacios máximos.

■ **Teorema** (de Jordan con un único autovalor) *Sea $f : V \to V$ un endomorfismo con un único autovalor $\lambda \in \mathbb{K}$. Entonces existe una base B respecto de la cual f tiene una matriz de Jordan.*

Demostración. Sea $k = \mathrm{nil}(\lambda)$ y por tanto $M(\lambda) = E_k(\lambda)$. La idea de la demostración es construir una base que corresponda a un diagrama de puntos de una matriz de Jordan como hemos visto en la sección anterior. Recuérdese que por el Teorema de los subespacios máximos se tiene que

$$\dim(M(\lambda)) = \mathrm{mult}_a(\lambda) = \dim(V).$$

Consideremos una base B_{k-1} de $E_{k-1}(\lambda)$ y sean vectores $\{v_1, \ldots, v_r\}$ de forma que la unión de todos $B_{k-1} \cup \{v_1, \ldots, v_r\}$ es una base de V. Observamos que $r = \dim(E_k(\lambda)) - \dim(E_{k-1}(\lambda))$ y que ninguna combinación lineal de los vectores

$$\{v_1, \ldots, v_r\}$$

pertenece a $E_{k-1}(\lambda)$. A continuación calculamos las sucesivas imágenes de cada $v_1, \ldots, v_r$ por medio de la aplicación $f - \lambda I$, y empezamos a construir nuestro diagrama

$$\begin{array}{ccc} E_{k-1}(\lambda) & \hookrightarrow & E_k(\lambda) \\ \bullet & \longleftarrow & \bullet \\ \vdots & & \vdots^{(r)} \\ \bullet & \longleftarrow & \bullet \end{array}$$

En efecto, en general para cualquier vector $v \in V$ y cualquier $\ell = 2, 3, \ldots$ se tiene que

$$v \in E_\ell(\lambda) \Leftrightarrow (f - \lambda I)^\ell(v) = 0 \Leftrightarrow (f - \lambda I)^{\ell-1}(f - \lambda I)(v) = 0 \Leftrightarrow (f - \lambda I)(v) \in E_{\ell-1}(\lambda) \quad (*)$$

Observamos también que si una combinación lineal de los vectores

$$(f - \lambda I)(v_1), \ldots, (f - \lambda I)(v_r)$$

pertenece a $E_{k-2}(\lambda)$, es decir,

$$a_1(f - \lambda I)(v_1) + \ldots + a_r(f - \lambda I)(v_r) \in E_{k-2}(\lambda)$$

entonces $(f - \lambda I)(a_1 v_1 + \cdots + a_r v_r) \in E_{k-2}(\lambda)$, y en particular $a_1 v_1 + \cdots + a_r v_r \in E_{k-1}(\lambda)$, lo cual quiere decir que $a_1 = \ldots = a_r = 0$. Así que ninguna combinación no nula de los vectores anteriores pertenece a $E_{k-2}(\lambda)$ (en particular, estos vectores son linealmente independientes).

A continuación, fijamos una base cualquiera B_{k-2} de $E_{k-2}(\lambda)$. Si unimos esta base a los vectores que vienen representados por los puntos debajo de $E_{k-1}(\lambda)$, es decir,

$$B_{k-2} \cup \{(f - \lambda I)(v_1), \ldots, (f - \lambda I)(v_r)\}, \qquad (\bigstar)$$

obtenemos un conjunto de vectores linealmente independientes de $E_{k-1}(\lambda)$. Claro, si existe una combinación de ellos, entonces tendremos que

$$v + w = 0,$$

donde v es la combinación de los vectores en B_{k-2} y w es la combinación de los vectores $\{(f-\lambda I)(v_1), \ldots, (f-\lambda I)(v_r)\}$. En particular, $w = -v \in E_{k-2}(\lambda)$. Como la combinación de los vectores $\{(f-\lambda I)(v_1), \ldots, (f-\lambda I)(v_r)\}$ que forman w pertenece a $E_{k-2}(\lambda)$, deducimos que esa combinación debe ser nula. Por tanto también $v = 0$, y como v era una combinación de vectores de la base B_{k-2}, también deducimos que debe ser una combinación nula.

Finalmente, añadimos vectores $\{w_1, \ldots, w_\ell\}$ a los vectores en ($\bigstar$) hasta obtener una base de $E_{k-1}(\lambda)$. Escribimos ℓ puntos en la columna de $E_{k-1}(\lambda)$ debajo de los que ya teníamos y calculamos las imágenes de todas esa columna por medio de $f - \lambda I$, es decir,

$$\begin{array}{ccccc}
E_{k-2}(\lambda) & \hookrightarrow & E_{k-1}(\lambda) & \hookrightarrow & E_k(\lambda) \\
\bullet & \longleftarrow & \bullet & \longleftarrow & \bullet \\
\vdots & & \vdots & & \vdots^{(r)} \\
\bullet & \longleftarrow & \bullet & \longleftarrow & \bullet \\
\bullet & \longleftarrow & \bullet & & \\
\vdots & & \vdots^{(\ell)} & & \\
\bullet & \longleftarrow & \bullet & &
\end{array}$$

Obsérvese que por definición ninguna combinación lineal de los vectores que representan los puntos debajo de $E_{k-1}(\lambda)$ puede pertenecer a $E_{k-2}(\lambda)$.

Aplicando repetidamente esta construcción llegamos a un diagrama de puntos de la forma

$$\begin{array}{ccccccccc}
E_1(\lambda) & \hookrightarrow & \ldots & E_{k-2}(\lambda) & \hookrightarrow & E_{k-1}(\lambda) & \hookrightarrow & E_k(\lambda) \\
\bullet & \longleftarrow & \ldots & \bullet & \longleftarrow & \bullet & \longleftarrow & \bullet \\
\vdots & & & \vdots & & \vdots & & \vdots^{(r)} \\
\bullet & \longleftarrow & \ldots & \bullet & \longleftarrow & \bullet & \longleftarrow & \bullet \\
\bullet & \longleftarrow & \ldots & \bullet & \longleftarrow & \bullet & & \\
\vdots & & & \vdots & & \vdots^{(\ell)} & & \\
\bullet & \longleftarrow & \ldots & \bullet & \longleftarrow & \bullet & & \\
\vdots & & & \vdots & & & & \\
\bullet & \longleftarrow & & \bullet & & & & \\
\bullet & & & & & & & \\
\vdots & & & & & & & \\
\bullet & & & & & & &
\end{array}$$

donde la flecha $\longleftarrow$ simboliza que el vector que representa el punto de la izquierda se obtiene aplicando a $(f - \lambda I)$ al vector que representa el vector de la derecha, y que los vectores que vienen representados por los puntos en una misma columna $E_i(\lambda)$ satisfacen que ninguna combinación no nula de ellos pertenece a $E_{i-1}(\lambda)$, y que por tanto son en particular linealmente independientes. Además, por construcción vamos a tener en total $\dim(V)$ puntos, ya que la suma de todos los puntos debe ser igual a la dimensión de $E_k(\lambda)$, que ya sabemos que es igual a $\dim(V)$.

Así que lo único que nos falta por argumentar es que los vectores que hemos encontrado y que están representados en el diagrama son linealmente independientes. Si fuese así, formarían una base para la cual tenemos un diagrama de puntos con las propiedades $\mathsf{R}_1, \mathsf{R}_2, \mathsf{R}_3$ y por tanto f respecto de esta base tendrá una matriz de Jordan.

Pero probar que estos vectores son linealmente es muy fácil. Los vectores que vienen representados por puntos debajo de la columna $E_1(\lambda)$ son linealmente independientes por construcción. De nuevo por construcción ninguna combinación no nula de los vectores representados por puntos debajo de la columna $E_2(\lambda)$ pertenece a $E_1(\lambda)$ y por tanto los vectores de las dos primeras columnas son linealmente independientes. Si seguimos este argumento ascendente llegamos a que todos los vectores son linealmente independientes. □

Observación 2.5 En realidad para probar el teorema anterior solo hemos usado el Teorema de los subespacios máximos para asegurarnos que $V = M(\lambda)$. En el caso de que ya supíesemos que $V = M(\lambda)$ por algún otro motivo, entonces igualmente tendríamos el resultado.

Ejemplo 2.6 Sea el endomorfismo $f : \mathbb{R}^5 \to \mathbb{R}^5$ cuya matriz respecto de la base canónica es

$$A = \begin{pmatrix} -2 & 0 & 3 & 4 & 5 \\ 0 & -2 & 0 & 6 & 7 \\ 0 & 0 & -2 & 0 & 8 \\ 0 & 0 & 0 & -2 & 0 \\ 0 & 0 & 0 & 0 & -2 \end{pmatrix}$$

Evidentemente su único autovalor es $\lambda = -2$. Calculemos primero las dimensiones de los subespacios propios generalizados. Para ello tenemos que calcular las distintas potencias de $A - (-2)I$, que son

$$[A-(-2)I]^2 = \begin{pmatrix} 0 & 0 & 0 & 0 & 24 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \qquad [A-(-2)I]^3 = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Así pues,

$$\begin{cases} \dim(E_1(-2)) = 5 - \mathrm{rg}(A-(-2)I) = 2 \\ \dim(E_2(-2)) = 5 - \mathrm{rg}(A-(-2)I)^2 = 4 \\ \dim(E_3(-2)) = 5 - \mathrm{rg}(A-(-2)I)^3 = 5 \end{cases}$$

Así pues, nuestro diagrama de puntos será

$$\begin{array}{ccccc} E_1(-2) & \hookrightarrow & E_2(-2) & \hookrightarrow & E_3(-2) \\ \bullet & \longleftarrow & \bullet & \longleftarrow & \bullet \\ \bullet & \longleftarrow & \bullet & & \end{array}$$

y la matriz A será semejante a la matriz de Jordan

$$J = \left(\begin{array}{ccc:cc} -2 & 1 & 0 & & \\ 0 & -2 & 1 & & \\ 0 & 0 & -2 & & \\ \hdashline & & & -2 & 1 \\ & & & 0 & -2 \end{array}\right)$$

Calculemos una base $B = \{v_1, v_2, v_3, v_4, v_5\}$ de $\mathbb{R}^5$ respecto de la cual el endomorfismo f tiene la matriz J anterior. Lo hacemos siguiendo el método de la demostración del Teorema de Jordan con un único autovalor.

Calculemos primero una base de $E_2(-2)$, es decir, buscamos $(x, y, z, t, w) \in \mathbb{R}^5$ tales que

$$\underbrace{\begin{pmatrix} 0 & 0 & 0 & 0 & 24 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}}_{[A-(-2)I]^2} \begin{pmatrix} x \\ y \\ z \\ t \\ w \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

y por tanto sacamos que

$$E_2(-2) = L[(1, 0, 0, 0, 0), (0, 1, 0, 0, 0), (0, 0, 1, 0, 0), (0, 0, 0, 1, 0)]$$

A continuación, buscamos un vector que añadido a la base anterior de $E_2(-2)$ que hemos calculado, nos de una base de $E_3(-2) = \mathbb{R}^5$. Basta tomar el vector

$$v_3 = (0, 0, 0, 0, 1).$$

Este vector será el que representa el punto debajo de $E_3(-2)$, por eso lo hemos denotado con v_3. A continuación, calculamos su imágen por medio de $f - (-2)I$, obteniendo

$$v_2 = (f - (-2)I)(v_3) = (5, 7, 8, 0, 0)$$

A continuación calculamos una base de $E_1(-2)$, es decir, buscamos $(x, y, z, t, w) \in \mathbb{R}^5$ tales que

$$\underbrace{\begin{pmatrix} 0 & 0 & 3 & 4 & 5 \\ 0 & 0 & 0 & 6 & 7 \\ 0 & 0 & 0 & 0 & 8 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}}_{[A-(-2)I]} \begin{pmatrix} x \\ y \\ z \\ t \\ w \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

y por tanto sacamos que

$$E_1(-2) = L[(1, 0, 0, 0, 0), (0, 1, 0, 0, 0)]$$

Ya sabemos que el vector v_2 es independiente de $(1, 0, 0, 0, 0)$ y $(0, 1, 0, 0, 0)$, y necesitamos un vector que unido a estos tres vectores nos de una base de $E_2(-2)$. Podemos tomar

$$v_5 = (0, 0, 0, 1, 0).$$

Por último, calculamos las imágenes de v_2 y v_5 por medio de $f - (-2)I$ para obtener los vectores v_1 y v_4 respectivamente,

$$\begin{cases} v_1 = (f - (-2)I)(v_2) = (24, 0, 0, 0, 0) \\ v_4 = (f - (-2)I)(v_5) = (4, 6, 0, 0, 0) \end{cases}$$

Finalmente, hemos llegado a que la matriz de f respecto de la base

$$B = \{\underbrace{(24, 0, 0, 0, 0)}_{v_1}, \underbrace{(5, 7, 8, 0, 0)}_{v_2}, \underbrace{(0, 0, 0, 0, 1)}_{v_3}, \underbrace{(4, 6, 0, 0, 0)}_{v_4}, \underbrace{(0, 0, 0, 1, 0)}_{v_5}\}$$

es la matriz de Jordan J.

Observación Al construir una base respecto de la cual un endomorfismo f tiene una matriz de Jordan, nos enfrentamos varias veces al problema de poseer unos cuantos vectores linealmente independientes de $E_k(\lambda)$ y tener que que buscar unos cuantos más de forma que, cuando se los añadimos a los que ya tenemos, nos quede una base de $E_k(\lambda)$. En el ejemplo anterior lo hemos podido hacer a ojo, pero lo suyo sería tener un método general. Lo mostramos con un ejemplo. Supongamos que en $\mathbb{R}^5$ tenemos un subespacio vectorial

$$W = L[(4,6,6,2,5),(1,0,-1,1,3),(0,3,5,1,1),(2,1,2,0,0)]$$

y que sabemos que

$$v_1 = (1,2,0,0,1), v_2 = (0,-2,-8,0,3)$$

son dos vectores linealmente independientes de W. Queremos encontrar otros dos vectores v_3 y v_4 tales que $\{v_1,v_2,v_3,v_4\}$ sea una base de W. Tan solo tenemos que poner los vectores v_1 y v_2 y los de la base de W que ya conocemos como filas de una matriz, y escalonar,

$$\begin{pmatrix} 1 & 2 & 0 & 0 & 1 \\ 0 & -2 & -8 & 0 & 3 \\ 4 & 6 & 6 & 2 & 5 \\ 1 & 0 & -1 & 1 & 3 \\ 0 & 3 & 5 & 1 & 1 \\ 2 & 1 & 2 & 0 & 0 \end{pmatrix} \xrightarrow{escalonamos} \begin{pmatrix} 1 & 2 & 0 & 0 & 1 \\ 0 & -2 & -8 & 0 & 3 \\ 0 & 0 & 14 & 2 & -2 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2 & \frac{9}{2} \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Ya sabemos que las operaciones elementales en filas respetan la dependencia lineal, es decir, que el subespacio que generaban las filas de la matriz de la izquierda (que evidentemente es W) es el mismo que el que generan las filas de la matriz de la derecha. Así pues, podemos tomar

$$v_3 = (0,0,14,2,-2) \qquad v_4 = (0,0,0,2,9/2)$$

Teorema de Jordan

En esta sección queremos abordar ya la prueba del Teorema de Jordan en el caso general de un endomorfismo $f : V \to V$ cualquiera (no necesariamente con un único autovalor). Recordemos que seguimos dando por cierto el Teorema de los subespacios máximos (el cual probaremos más adelante). En realidad, asumiendo este teorema y con lo visto en la sección anterior, el caso general es casi inmediato. Necesitamos antes el siguiente resultado que nos va a permitir reducirlo todo al caso de tener un único autovalor (que es el caso que ya sabemos hacer).

■ **Proposición** *Sea $f : V \to V$ un endomorfismo, y sea $\lambda \in \mathbb{K}$ un autovalor. Entonces los subespacios propios generalizados $E_k(\lambda)$ de λ son subespacios invariantes de f. En particular, el subespacio máximo $M(\lambda)$ de λ es un subespacio invariante de f.*

Demostración. Dado $v \in E_k(\lambda)$, queremos probar que $f(v) \in E_k(\lambda)$. Se tiene que

$$(f - \lambda I)(v) = f(v) - \lambda v \Rightarrow f(v) = v + (f - \lambda I)(v).$$

Por un lado tenemos que $v \in E_k(\lambda)$ pero también $(f - \lambda I)(v) \in E_{k-1}(\lambda) \subset E_k(\lambda)$, y puesto que $E_k(\lambda)$ es un subespacio vectorial de V, deducimos que $f(v) = v + (f - \lambda I)(v) \in E_k(\lambda)$. □

■ **Teorema** (de Jordan) *Sea $f : V \to V$ un endomorfismo de un $\mathbb{K}$-espacio vectorial V. Supongamos que todas raíces $\lambda_1, \ldots, \lambda_r$ del polinomio característico de f pertenecen a $\mathbb{K}$. Entonces existe una base B de V respecto de la cual f tiene una matriz de Jordan.*

Demostración. Consideremos para cada λ_i su subespacio máximo $M(\lambda_i)$. Por la proposición anterior, $M(\lambda_i)$ es un subespacio invariante de f. Por tanto, podemos considerar el endomorfismo

$$f|_{M(\lambda_i)} : M(\lambda_i) \to M(\lambda_i).$$

El único autovalor de $f|_{M(\lambda_i)}$ es λ_i, ya que si tuviese otro autovalor $\mu \neq \lambda_i$, entonces por definición existiría un vector $w \in M(\lambda_i)$ no nulo tal que $f|_{M(\lambda_i)}(w) = f(w) = \mu w$. Es decir μ debe ser un autovalor de f, por tanto $\mu = \lambda_j$ con $i \neq j$, y entonces $w \in E_1(\lambda_j) \subset M(\lambda_j)$, pero ya sabemos que $M(\lambda_i) \cap M(\lambda_j) = 0$ por el Teorema de los subespacios máximos, contradicción. Por lo visto en la sección anterior, podemos encontrar una base B_i de $M(\lambda_i)$ tal que $f|_{M(\lambda_i)}$ tiene una matriz de Jordan J_i. Ahora bien, si consideramos la unión de todas esas bases

$$B = B_1 \cup \cdots \cup B_r$$

nos proporciona una base de $M(\lambda_1) \bigoplus \cdots \bigoplus M(\lambda_r)$. Por el Teorema de los subespacios máximos también sabemos que $M(\lambda_1) \bigoplus \cdots \bigoplus M(\lambda_r) = V$ y por tanto B es una base de V. Finalmente, la matriz de f respecto de la base B es la matriz de Jordan,

$$\left(\begin{array}{c:c:c} J_1 & & \\ \hdashline & \ddots & \\ \hdashline & & J_r \end{array}\right)$$

$\square$

Ejemplo 2.7 Sea la matriz

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

El polinomio característico es

$$p(\lambda) = \det(A - \lambda I) = \begin{vmatrix} 1-\lambda & 0 & 0 & 0 \\ -1 & 2-\lambda & 1 & 0 \\ -1 & 0 & 2-\lambda & 0 \\ -1 & 0 & 1 & 1-\lambda \end{vmatrix} = (1-\lambda) \begin{vmatrix} 2-\lambda & 1 & 0 \\ 0 & 2-\lambda & 0 \\ 0 & 1 & 1-\lambda \end{vmatrix} =$$

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

Y por tanto los autovalores son $\lambda =_1$ y $\lambda_2 = 2$ ambos con multiplicidad algebraica 2.

Cálculo de los subespacios propios de $\lambda_1 = 1$. Calculamos primero $E_1(1)$, es decir, los autovectores de $\lambda_1 = 1$. Buscamos $(x, y, z, t) \in \mathbb{R}^4$ tales que

$$\begin{pmatrix} 0 & 0 & 0 & 0 \\ -1 & 1 & 1 & 0 \\ -1 & 0 & 1 & 0 \\ -1 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \to \begin{cases} 0 = 0 \\ -x + y + z = 0 \\ -x + z = 0 \\ -x + z = 0 \end{cases}$$

de lo que deducimos que

$$E_1(1) = L[(1, 0, 1, 0), (0, 0, 0, 1)]$$

Como la dimensión de $E_1(1)$ coincide con la multiplicidad algebraica de $\lambda_1 = 1$ ya sabemos entonces que $M(1) = E_1(1)$.

Cálculo de los subespacios propios de $\lambda_2 = 2$. Calculamos primero $E_1(2)$, es decir, los autovectores de $\lambda_2 = 2$. Buscamos $(x, y, z, t) \in \mathbb{R}^4$ tales que

$$\begin{pmatrix} -1 & 0 & 0 & 0 \\ -1 & 0 & 1 & 0 \\ -1 & 0 & 0 & 0 \\ -1 & 0 & 1 & -1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \to \begin{cases} -x = 0 \\ -x + z = 0 \\ -x = 0 \\ -x + z - t = 0 \end{cases}$$

de lo que deducimos que

$$E_1(2) = L[(0,1,0,0)].$$

De forma inmediata ya sabemos entonces que $\dim(E_2(2)) = 2$ y $M(2) = E_2(2)$. Pero busquemos una base de $E_2(2)$.

$$\underbrace{\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & -1 & 1 \end{pmatrix}}_{(A-2I)^2} \begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \to \begin{cases} x = 0 \\ 0 = 0 \\ x = 0 \\ x - z + t = 0 \end{cases}$$

de lo que deducimos que

$$E_2(2) = L[(0,0,1,1),(0,1,0,0)].$$

Forma canónica de Jordan de A. Con los cálculos anteriores ya sabemos que la forma canónica de Jordan de A va a ser

$$J = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 2 \end{pmatrix}$$

Cálculo de la matriz P. Para los bloques de orden 1 correspondientes a $M(1)$ podemos tomar la base que hemos calculado anteriormente

$$\{(0,0,0,1),(1,0,1,0)\}.$$

Para el bloque de Jordan de orden 2 correspondiente a $M(2)$, debemos aplicar lo que aprendimos en la Sección 2.4. Así pues, buscamos una base de $E_1(2)$ y la completamos hasta obtener una base de $M(2) = E_2(2)$. Observamos que de hecho las mismas bases que nos han quedado después de los cálculos nos sirven. Tomamos por tanto el vector $(0,0,1,1)$ y calculamos su imagen por medio de $A - 2I$, que da la casualidad que nos queda $(0,1,0,0)$. Tenemos por tanto que el endomorfismo que tiene matriz asociada A respecto de la canónica, tiene respecto de la base

$$B = \{(0,0,0,1),(1,0,1,0),(0,1,0,0),(0,0,1,1)\},$$

la matriz J anterior. Así pues, dada la matriz de cambio de base de B a canónica,

$$P = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & 0 & 1 \end{pmatrix},$$

tenemos la igualdad $A = PJP^{-1}$.

Ejemplo 2.8 Sea la matriz

$$A = \begin{pmatrix} 2 & 2 & -3 & 4 \\ -2 & 2 & 1 & 0 \\ 3 & 3 & -5 & 7 \\ 4 & 2 & -6 & 7 \end{pmatrix}$$

El polinomio característico es

$$p(\lambda) = \det(A - \lambda I) = \begin{vmatrix} 2-\lambda & 2 & -3 & 4 \\ -2 & 2-\lambda & 1 & 0 \\ 3 & 3 & -5-\lambda & 7 \\ 4 & 2 & -6 & 7-\lambda \end{vmatrix} =$$

$$= \text{cuatro hojas de cálculos después} = (1-\lambda)^2(2-\lambda)^2.$$

Y por tanto los autovalores son $\lambda_1 = 1$ y $\lambda_2 = 2$ ambos con multiplicidad algebraica 2.

Cálculo de los subespacios propios de $\lambda_1 = 1$. Calculamos primero $E_1(1)$, es decir, los autovectores de $\lambda_1 = 1$. Buscamos $(x, y, z, t) \in \mathbb{R}^4$ tales que

$$\begin{pmatrix} 1 & 2 & -3 & 4 \\ -2 & 1 & 1 & 0 \\ 3 & 3 & -6 & 7 \\ 4 & 2 & -6 & 6 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \to \begin{cases} x + 2y - 3z + 4t = 0 \\ -2x + y + z = 0 \\ 3x + 3y - 6z + 7t = 0 \\ 4x + 2y - 6z + 6t = 0 \end{cases}$$

de lo que deducimos que

$$E_1(1) = L[(1, 1, 1, 0)]$$

En particular eso nos indica que $\dim(E_2(1)) = 2$ y $M(1) = E_2(1)$. Busquemos una base de $E_2(2)$.

$$\underbrace{\begin{pmatrix} 4 & 3 & -7 & 7 \\ -1 & 0 & 1 & -1 \\ 7 & 5 & -12 & 12 \\ 6 & 4 & -10 & 10 \end{pmatrix}}_{(A-2I)^2} \begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

de lo que deducimos que

$$E_2(1) = L[(1, 1, 0, -1), (0, 0, 1, 1)].$$

Cálculo de los subespacios propios de $\lambda_2 = 2$. Calculamos primero $E_1(2)$, es decir, los autovectores de $\lambda_2 = 2$. Buscamos $(x, y, z, t) \in \mathbb{R}^4$ tales que

$$\begin{pmatrix} 0 & 2 & -3 & 4 \\ -2 & 0 & 1 & 0 \\ 3 & 3 & -7 & 7 \\ 4 & 2 & -6 & 5 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

de lo que deducimos que

$$E_1(2) = L[(1, -1, 2, 2)].$$

De forma inmediata ya sabemos entonces que $\dim(E_2(2)) = 2$ y $M(2) = E_2(2)$. Pero busquemos una base de $E_2(2)$.

$$\underbrace{\begin{pmatrix} 3 & -1 & -1 & -1 \\ 3 & -1 & -1 & -1 \\ 1 & -1 & 1 & -2 \\ -2 & 0 & 2 & -1 \end{pmatrix}}_{(A-2I)^2} \begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

de lo que deducimos que

$$E_2(2) = L[(1, 2, 1, 0), (0, 3, -1, -2)].$$

Forma canónica de Jordan de A. Con los cálculos anteriores ya sabemos que la forma canónica de Jordan de A va a ser

$$J = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 2 \end{pmatrix}$$

Cálculo de la matriz P. Busquemos la base que corresponde al bloque de orden 2 de $M(1)$. Tomamos por ejemplo la base $\{(1,1,1,0)\}$ de $E_1(1)$ que habíamos calculado, y la completamos a una de $E_2(1)$, por ejemplo, $\{(1,1,1,0),(0,0,1,1)\}$. Tomamos el vector que no está en $E_1(1)$, es decir, $(0,0,1,1)$, y le aplicamos $A-I$, lo que casualmente da como resultado $(1,1,1,0)$ (sabíamos en cualquier caso que el resultado tenía que ser un múltiplo de $(1,1,1,0)$). Así pues, debemos tomar la base

$$\{(1,1,1,0),(0,0,1,1)\}$$

para el bloque de $M(1)$.

Para el bloque de Jordan de orden 2 correspondiente a $M(2)$, tomamos por ejemplo la base $\{(1,-1,2,2)\}$ de $E_1(1)$ que habíamos calculado, y la completamos hasta obtener una base de $E_2(2)$, por ejemplo, $\{(1,-1,2,2),(1,2,1,0)\}$. Tomamos el vector que no está en $E_1(2)$, es decir, $(1,2,1,0)$, y le aplicamos $A-2I$, lo que casualmente da como resultado $(1,-1,2,2)$. Así pues, debemos tomar la base

$$\{(1,-1,2,2),(1,2,1,0)\}$$

para el bloque de $M(2)$.

Tenemos por tanto que el endomorfismo que tiene matriz asociada A respecto de la canónica, tiene respecto de la base

$$B=\{(1,1,1,0),(0,0,1,1),(1,-1,2,2),(1,2,1,0)\},$$

la matriz J anterior. Así pues, dada la matriz de cambio de base de B a canónica,

$$P=\begin{pmatrix} 1 & 0 & 1 & 1 \\ 1 & 0 & -1 & 2 \\ 1 & 1 & 2 & 1 \\ 0 & 1 & 2 & 0 \end{pmatrix},$$

tenemos la igualdad $A=PJP^{-1}$.

Polinomios anuladores y el Teoremas de los subespacios máximos

En esta sección vamos a suponer que todos los espacios vectoriales son sobre $\mathbb{K}=\mathbb{C}$. Esta sección está basada en la Sección 14.6 de Volumen 2 del Libro de *Álgebra Lineal* de Fernando, Gamboa, Ruiz.

Polinomios anuladores de un endomorfismo

Sea $f:V\to V$ un endomorfismo de un $\mathbb{C}$-espacio vectorial. Denotamos como es habitual

$$f^0=\mathrm{I}, \quad \text{y} \quad f^k=f\circ\overset{k)}{\cdots}\circ f \quad \text{para } k\geq 1.$$

Dado un polinomio

$$P(T)=\sum_{k=0}^{p} a_kT^k\in\mathbb{C}[T]$$

queremos darle sentido a sustituir $T=f$ en $P(T)$. Simplemente interpretamos T^k como la composición f^k, de forma que queda

$$P(f)=\sum_{k=0}^{p} a_kf^k$$

el cual también es un endomorfismo de V bien definido.

Observación 2.9 1) Si f respecto de una base B tiene matriz A, entonces la matriz del endomorfismo $P(f)$ respecto de B es evidentemente

$$P(A) = \sum_{k=0}^{p} a_k A^k.$$

2) La sustitución anterior respeta el producto de polinomios, es decir, dados dos polinomios $P(T), Q(T) \in \mathbb{C}[T]$, se cumple

$$(PQ)(f) = P(f) \circ Q(f).$$

Por ejemplo, si $P(T) = T^2 - 5$ y $Q(T) = T^3 - T^2$ se tiene que

$$\begin{aligned} P(T)Q(T) =& (T^2-5)(T^3-T^2) = T^2(T^3-T^2) - 5(T^3-T^2) = \\ =& T^{2+3} - T^{2+2} - 5T^3 + 5T^2 = T^5 - T^4 - 5T^3 + 5T^2 \\ \Rightarrow (PQ)(f) =& f^5 - f^4 - 5f^3 + 5f^2. \end{aligned}$$

Por otro lado $P(f) = f^2 - 5I$ y $Q(f) = f^3 - f^2$ y se tiene que

$$\begin{aligned} P(f) \circ Q(f) =& (f^2 - 5I) \circ (f^3 - f^2) = f^2 \circ (f^3 - f^2) - 5I \circ (f^3 - f^2) = \\ =& f^{2+3} - f^{2+2} - 5f^3 + 5f^2 = f^5 - f^4 - 5f^3 + 5f^2. \end{aligned}$$

Vemos que la demostración en general se reduce a la igualdad evidente $f^k \circ f^\ell = f^{k+\ell}$, después de hacer cuidadosamente todas las operaciones. En particular, como el producto entre polinomios es conmutativo, deducimos que

$$P(f) \circ Q(f) = (PQ)(f) = (QP)(f) = Q(f) \circ P(f).$$

3) Supongamos que tenemos polinomios $P_1(T), \ldots, P_\ell(T)$ y que sabemos que para cierto vector $v \in V$ se tiene que $P_i(f)(v) = 0$. Entonces para el polinomio $P(T) = P_1(T) \cdots P_\ell(T)$ también se tiene que $P(f)(v) = 0$. En efecto,

$$\begin{aligned} P(f)(v) &= (P_1 \cdots P_\ell)(f)(v) = \\ &= (P_1 \cdots P_{i-1} P_{i+1} \cdots P_\ell P_i)(f)(v) = \\ &= (P_1 \cdots P_{i-1} P_{i+1} \cdots P_\ell) \circ P_i(f)(v) = \\ &= (P_1 \cdots P_{i-1} P_{i+1} \cdots P_\ell)(0) = 0. \end{aligned}$$

4) Si $\lambda \in \mathbb{C}$ es un autovalor de f, y u un autovector asociado, es decir, $f(u) = \lambda u$, entonces se tiene que

$$P(f)(u) = P(\lambda)u.$$

En efecto, observamos primero que

$$f^2(u) = f(f(u)) = f(\lambda u) = \lambda f(u) = \lambda(\lambda u) = \lambda^2 u.$$

y repitiendo el cálculo se ve que $f^k(u) = \lambda^k u$ para cualquier k. En particular,

$$P(f)(u) = \sum_{k=0}^{p} a_k f^k(u) = \sum_{k=0}^{p} a_k \lambda^k u = (\sum_{k=0}^{p} a_k \lambda^k)u = P(\lambda)u.$$

■ **Definición** Dado un endomorfismo $f : V \to V$, decimos que un polinomio $P(T) \in \mathbb{C}[T]$ es un *polinomio anulador* de f si $P(f) = 0$.

Observación 2.10 (1) Todo endomorfismo tiene un polinomio anulador. En efecto, supongamos que hemos fijado una base B de V y que la matriz de f es A respecto de esta base. Ya sabemos que el espacio de matrices $\mathcal{M}_n(\mathbb{C})$ tiene dimensión n^2, así que entre las siguientes n^2+1 matrices,

$$\mathrm{I}, A, A^2, \ldots, A^{n^2},$$

debe haber una dependencia lineal, es decir, existen $a_0, \ldots, a_{n^2}$ tales que

$$a_0 I + a_1 A + \ldots + a_{n^2} A^{n^2} = 0.$$

En particular estamos diciendo que $P(T) = \sum_{k=0}^{n^2} a_k T^k$ es una polinomio anulador de f.

(2) Si $P(T)$ es un producto de polinomios $Q_1, \ldots, Q_\ell$ y alguno de ellos es anulador de f, entonces P también es anulador de f. Claro, siempre podemos reordenarlos y podemos suponer que Q_ℓ es el anulador de f, de forma que $P(f) = Q_1(f) \circ \cdots Q_{\ell-1}(f) \circ Q_\ell(f) = Q_1(f) \circ \cdots Q_{\ell-1} \circ 0 = 0$.

¿Y al revés? ¿Si P es un polinomio anulador de f y $P = Q_1, \ldots, Q_\ell$, entonces algún Q_i es un anulador de f? No, eso no es cierto. Por ejemplo, tomemos $Q_1(T) = T$ y $Q_2(T) = T$, y consideremos el producto $P(T) = Q_1(T)Q_2(T) = T^2$. Se tiene entonces que el endomorfismo $f : \mathbb{R}^2 \to \mathbb{R}^2 : (x, y) \mapsto (x - y, x - y)$ se anula en $P(T)$ pero no se anula en Q_1 ni en Q_2.

(3) Los autovalores de f son raíces de cualquier polinomio anulador $P(T)$ de f. En efecto, si $f(u) = \lambda u$ con $u \neq 0$, entonces por la Observación 2.9 (4)

$$0 = P(f)(u) = P(\lambda)u,$$

y como $u \neq 0$, es $P(\lambda) = 0$.

(4) Todo endomorfismo f tiene un polinomio cuyas raíces en $\mathbb{C}$ son exactamente los autovalores de f. En efecto, sea $P(T)$ un polinomio anulador de f y μ una raíz de P que no es autovalor de f. Entonces $P(T) = (T - \mu)Q(T)$, y para $u \in V$ se tiene:

$$0 = P(f)(u) = (f - \mu\mathrm{I}) \circ Q(f)(u) = f\big(Q(f)(u)\big) - \mu Q(f)(u),$$

luego si denotamos por $v = Q(f)(u)$ hemos deducido que $f(v) = \mu v$. Como μ no es autovalor, el vector $v = Q(f)(u)$ debe ser nulo. Es decir, hemos probado que para todo $u \in V$ se tiene que $Q(f)(u) = 0$, o dicho de otra forma, $Q(f) = 0$. Por tanto $Q(T)$ es también un polinomio anulador de f. De esta manera, podemos eliminar todas las raíces que no son autovalores.

■ **Proposición 2.11** *Sea $f : V \to V$ un endomorfismo, y sean $\lambda_1, \ldots, \lambda_r \in \mathbb{C}$ todos sus autovalores. Supongamos que el polinomio $P(T) = (T-\lambda_1)^{m_1} \cdots (T-\lambda_r)^{m_r} = \prod_i (T-\lambda_i)^{m_i}$ es un polinomio anulador de f. Entonces*

1) $E_{m_j}(\lambda_j) \cap \sum_{i \neq j} E_{m_i}(\lambda_i) = \{0\}$,

2) $V = E_{m_1}(\lambda_1) + \cdots + E_{m_r}(\lambda_r)$,

3) $M(\lambda_i) = E_{m_i}(\lambda_i)$ *y por tanto* $\mathrm{nil}(\lambda_i) \leq m_i$.

Demostración. 1) Para cada $j = 1, \ldots, r$ consideramos los polinomios

$$P_j(T) = (T - \lambda_j)^{m_j} \quad \text{y} \quad Q_j(T) = \prod_{i \neq j} P_i(T) = \prod_{i \neq j} (T - \lambda_i)^{m_i}.$$

Observamos que estos polinomios tienen la siguiente propiedad:

$$P_j(f)(u) = 0, \quad \text{si } u \in E_{m_j}(\lambda_j),$$
$$Q_j(f)(v) = 0, \quad \text{si } v \in E_{m_i}(\lambda_i), i \neq j.$$

En efecto, dado $u \in E_{m_j}(\lambda_j)$ se tiene que $P_j(f)(u) = (f - \lambda_j I)^{m_j}(u) = 0$. Por otro lado, si $v \in E_{m_i}(\lambda_i)$ con $i \neq j$, como $P_i(f)(v) = 0$, se tiene por la Observación 2.9.(3) que $Q_j(f)(v) = 0$.

Observamos también que $P_j(T)$ y $Q_j(T)$ no tiene raíces comunes y por tanto podemos encontrar ciertos polinomios $A(T)$ y $B(T)$ de forma que

$$1 = A(T)P_j(T) + B(T)Q_j(T)$$

y sustituyendo $T = f$,

$$\mathrm{I} = A(f) \circ P_j(f) + B(f) \circ Q_j(f).$$

Por tanto, para cada $u \in V$ resulta

$$u = A(f)(P_j(f)(u)) + B(f)(Q_j(f)(u)) \tag{$\star$}$$

Finalmente, tomemos

$$u \in E_{m_j}(\lambda_j) \cap \sum_{i \neq j} E_{m_i}(\lambda_i)$$

y veamos que $u = 0$. Por una parte, como $u \in E_{m_j}(\lambda_j)$ ya sabemos que $P_j(f)(u) = 0$, y por otra parte, como $u = \sum_{i \neq j} u_i$ para ciertos $u_i \in E_{m_i}(\lambda_i)$, se tiene que

$$Q_j(f)(u) = Q_j(f)\big(\sum_{i \neq j} u_i\big) = \sum_{i \neq j} Q_j(f)(u_i) = 0.$$

En consecuencia, aplicando esto a la ecuación $(\star)$,

$$u = A(f)(P_j(f)(u)) + B(f)(Q_j(f)(u)) = A(f)(0) + B(f)(0) = 0.$$

es decir, $u = 0$ que es lo que queríamos probar.

2) Como los polinomios $Q_1(T), \ldots, Q_r(T)$ no tienen ninguna raíz en común, existen polinomios $A_1(T), \ldots, A_r(T)$ tales que

$$1 = A_1(T)Q_1(T) + \cdots + A_r(T)Q_r(T),$$

y substituyendo $T = f$,

$$\mathrm{I} = A_1(f) \circ Q_1(f) + \cdots + A_r(f) \circ Q_r(f).$$

Si aplicamos los endomorfismos de la izquierda y de la derecha a un vector $u \in V$ arbitrario deducimos que

$$u = A_1(f) \circ Q_1(f)(u) + \cdots + A_r(f) \circ Q_r(f)(u).$$

Denotemos $u_i = A_i(f) \circ Q_i(f)(u)$ para cada $i = 1, \ldots, r$, es decir,

$$u = \underbrace{A_1(f) \circ Q_1(f)(u)}_{u_1} + \cdots + \underbrace{A_r(f) \circ Q_r(f)(u)}_{u_r}.$$

Evidentemente, si probamos que

$$u_i \in E_{m_i}(\lambda_i) = \ker(f - \lambda_i \mathrm{I})^{m_i} = \ker(P_i(f)),$$

habremos terminado. Pero lo anterior es inmediato ya que

$$\begin{aligned} P_i(f)(u_i) &= P_i(f) \circ A_i(f) \circ Q_i(f)(u) = \\ &= A_i(f) \circ (Q_i P_i)(f)(u) = A_i(f) \circ P(f)(u) = A_i(f)(0) = 0. \end{aligned}$$

3) Los dos puntos anteriores nos dicen que tenemos la descomposición en suma directa

$$(*) \qquad V = E_{m_1}(\lambda_1) \oplus \cdots \oplus E_{m_r}(\lambda_r).$$

Evidentemente el polinomio

$$Q(T) = P(T)(T-\lambda_1)\cdots(T-\lambda_r) = (T-\lambda_1)^{m_1+1}\cdots(T-\lambda_r)^{m_r+1}$$

también es un polinomio anulador de f. Luego también

$$(**) \qquad V = E_{m_1+1}(\lambda_1) \oplus \cdots \oplus E_{m_r+1}(\lambda_r).$$

Como $(*)$ y $(**)$ son dos sumas directas, cada sumando de la primera está contenido en el correspondiente de la segunda, tiene que ser

$$E_{m_i}(\lambda_i) = E_{m_i+1}(\lambda_i).$$

En otras palabras la sucesión de espacios invariantes de λ_i se estabiliza al menos a partir de $E_{m_i}(\lambda_i)$, así que éste es el subespacio invariante maximal $M(\lambda_i)$. □

La demostración del Teorema de los subespacios máximos es ahora muy fácil:

Demostración del Teorema de los subespacios máximos. Probemos primero el caso en el que $\mathbb{K} = \mathbb{C}$. Por la Observación 2.10 sabemos que f tiene al menos un polinomio anulador cuyas raíces son los autovalores de f. Por la proposición anterior deducimos directamente que $V = M(\lambda_1) \bigoplus \cdots \bigoplus M(\lambda_r)$. Por otro lado, por el Teorema de Jordan con un único autovalor (véase la Observación 2.5) tenemos que cada

$$f|_{M(\lambda_i)} : M(\lambda_i) \to M(\lambda_i)$$

tiene una base respecto de la cual su matriz es de Jordan J_i y cuyo orden es evidentemente $\dim(M(\lambda_i))$. Denotemos $d_i = \dim(M(\lambda_i))$. Si juntamos todas las bases obtendremos una base de V respecto de la cual f tiene matriz

$$\left(\begin{array}{c|c|c} J_1 & & \\ \hline & \ddots & \\ \hline & & J_r \end{array}\right)$$

Así pues, el polinomio característico de f es

$$\begin{aligned} P(T) &= \det(J - T\,I_n) = \det(J_1 - T\,I_{d_1})\cdots\det(J_r - T\,I_{d_r}) \\ &= (\lambda_1 - T)^{d_1}\cdots(\lambda_r - T)^{d_r}, \end{aligned}$$

En particular, vemos que d_i debe ser exactamente la multiplicidad algebraica de λ_i, es decir, se tiene que $\operatorname{mult}_a(\lambda_i) = \dim(M(\lambda_i))$.

Probemos por último el caso $\mathbb{K} = \mathbb{R}$, es decir, tenemos un endomorfismo $f : V \to V$ de un $\mathbb{R}$-espacio vectorial V cuyas raíces del polinomio característico pertenecen todas a $\mathbb{R}$. Si fijamos una base de B de V podemos trabajar en coordenadas, y por tanto podemos asumir que nuestro endomorfismo es de la forma $f : \mathbb{R}^n \to \mathbb{R}^n$. Consideremos ahora el complexificado de f, es decir, el endomorfismo $\widetilde{f} : \mathbb{C}^n \to \mathbb{C}^n$ cuya matriz es la misma que la matriz de f. Sea λ es un autovalor de f, el cual por hipótesis debe ser real. Si denotamos por $E_i^{\mathbb{C}}(\lambda) = \operatorname{Ker}(\widetilde{f} - \lambda I)^i \subset \mathbb{C}^n$ y por $E_i(\lambda) = \operatorname{Ker}(f - \lambda I)^i \subset \mathbb{R}^n$ observamos que

$$E_i^{\mathbb{C}}(\lambda) \cap \mathbb{R}^n = E_i(\lambda)$$

ya que $\widetilde{f}|_{\mathbb{R}^n} = f$. También sabemos que

$$\begin{aligned} \dim_{\mathbb{R}}(E_i(\lambda)) &= \dim_{\mathbb{R}}(\operatorname{Ker}(f - \lambda I)^i) = n - \operatorname{rg}_{\mathbb{R}}(A - \lambda I)^i = \\ &= n - \operatorname{rg}_{\mathbb{C}}(A - \lambda I)^i = \dim_{\mathbb{C}}(\operatorname{Ker}(\widetilde{f} - \lambda I)^i) = \dim_{\mathbb{C}}(E_i^{\mathbb{C}}(\lambda)). \end{aligned}$$

En la igualdad anterior hemos usado que una matriz real tiene el mismo rango como matriz compleja que como matriz real, es decir, que el número máximo de filas o columnas linealmente independientes de dicha matriz sobre $\mathbb{R}$ o sobre $\mathbb{C}$ es el mismo (lo cual resulta obvio si pensamos en que el rango coincide con el mayor de los órdenes de los menores no nulos). Si juntamos todo lo anterior llegamos a que el índice de nilpotencia de λ respecto de $\widetilde{f}$ y respecto de f coinciden y que

$$M^{\mathbb{C}}(\lambda) \cap \mathbb{R}^n = E^{\mathbb{C}}_{\mathrm{nil}(\lambda)}(\lambda) \cap \mathbb{R}^n = E_{\mathrm{nil}(\lambda)}(\lambda) = M(\lambda) \quad \& \quad \dim_{\mathbb{C}}(M^{\mathbb{C}}(\lambda)) = \dim_{\mathbb{R}}(M(\lambda)).$$

Si $\lambda_1, \ldots, \lambda_r$ son los autovalores de f, que sabemos que son reales, evidentemente también serán los de $\widetilde{f}$ y por el Teorema de Jordan complejo la familia $M^{\mathbb{C}}(\lambda_1), \ldots, M^{\mathbb{C}}(\lambda_r)$ es independiente sobre $\mathbb{C}$, y por tanto $M(\lambda_1), \ldots, M(\lambda_r)$ es también una familia independiente sobre $\mathbb{R}$. Además,

$$\begin{aligned} n &= \dim_{\mathbb{C}}(M^{\mathbb{C}}(\lambda_1)) + \cdots + \dim_{\mathbb{C}}(M^{\mathbb{C}}(\lambda_r)) = \\ &= \dim_{\mathbb{R}}(M(\lambda_1)) + \cdots + \dim_{\mathbb{R}}(M(\lambda_r)). \end{aligned}$$

es decir, que tenemos que $\mathbb{R}^n = M(\lambda_1) \bigoplus \cdots \bigoplus M(\lambda_r)$. Por último, observamos que se tiene que $\mathrm{mult}_a(\lambda) = \dim_{\mathbb{C}}(M^{\mathbb{C}}(\lambda_i)) = \dim_{\mathbb{R}}(M(\lambda_i))$. □

■ **Definición** Sean $\lambda_1, \ldots, \lambda_r \in \mathbb{C}$ todos los autovalores de $f : V \to V$. Entonces llamamos al polinomio

$$P_{\min}(T) = \prod\nolimits_i (T - \lambda_i)^{\mathrm{nil}(\lambda_i)}$$

el *polinomio mínimo* de f.

■ **Proposición** *El polinomio mínimo es un polinomio anulador de f.*

Demostración. Sea $v \in M(\lambda_i) = E_{\mathrm{nil}(\lambda_i)}(\lambda_i)$. Evidentemente

$$(f - \lambda_i I)^{\mathrm{nil}(\lambda_i)}(v) = 0,$$

y por la Observación 2.9.(3) se tiene entonces que $P_{\min}(f)(v) = 0$. Por tanto dado cualquier vector $v \in V$, como lo podemos escribir

$$v = v_1 + \cdots + v_r$$

con $v_i \in M(\lambda_i)$, deducimos que

$$P_{\min}(f)(v) = P_{\min}(f)(v_1) + \cdots + P_{\min}(f)(v_r) = 0$$

y por tanto $P_{\min}(f) = 0$. □

■ **Proposición** *Si $Q(T)$ es un polinomio anulador de f, entonces $P_{min}(T)$ divide a $Q(T)$.*

Demostración. Ya hemos visto que los autovalores deben ser raíces de $Q(T)$, por tanto

$$Q(T) = (T - \lambda_1)^{m_1} \cdots (T - \lambda_r)^{m_r} R(T)$$

para algún polinomio $R(T)$ que no se anula en $T = \lambda_i$ con $i = 1, \ldots, r$. Ya vimos en la Observación 2.10.(3) que si μ es una raíz de $Q(T)$ que no es un autovalor (es por tanto una raíz de $R(T)$), entonces si dividimos $Q(T)$ entre $T - \mu$ nos sigue quedando un polinomio anulador de f. Así que si hacemos esto repetidamente llegamos a que $(T - \lambda_1)^{m_1} \cdots (T - \lambda_r)^{m_r}$ es un polinomio anulador de f. Y por la Proposición 2.11.(3) deducimos que $m_i \geq \mathrm{nil}(\lambda_i)$. En particular,

$$Q(T) = P_{\min}(T)(T - \lambda_1)^{m_1 - \mathrm{nil}(\lambda_1)} \cdots (T - \lambda_r)^{m_r - \mathrm{nil}(\lambda_r)} R(T),$$

es decir, $P_{\min}(T)$ divide a $Q(T)$. □

■ **Corolario** (Teorema de Cayley-Hamilton) *El polinomio característico $P(T)$ de f es divisible entre $P_{min}(T)$. En particular, se tiene que $P(f) = 0$.*

Demostración. Si $\lambda_1, \ldots, \lambda_r$ son todos los autovalores de f, lo que queremos probar es que

$$P(T) = (T - \lambda_1)^{\mathrm{mult}_a(\lambda_1)} \cdots (T - \lambda_r)^{\mathrm{mult}_a(\lambda_r)}$$

es divisible entre

$$P_{\min}(T) = (T - \lambda_1)^{\mathrm{nil}(\lambda_1)} \cdots (T - \lambda_r)^{\mathrm{nil}(\lambda_r)},$$

es decir, necesitamos probar que $\mathrm{nil}(\lambda_i) \leq \mathrm{mult}_a(\lambda_i)$ para cualquier $i = 1, \ldots, r$. Por definición

$$E_1(\lambda_i) \subsetneq E_2(\lambda_i) \subsetneq \cdots \subsetneq E_{\mathrm{nil}(\lambda_i)-1}(\lambda_i) \subsetneq E_{\mathrm{nil}(\lambda_i)}(\lambda_i) = M(\lambda_i),$$

y como los contenidos son estrictos, cada vez que pasamos de un subespacio propio generalizado al siguiente aumentamos al menos en 1 la dimensión, por tanto

$$\dim(M(\lambda_i)) =$$

$$= \dim(E_1(\lambda_i)) + \big(\dim(E_2(\lambda_i)) - \dim(E_1(\lambda_i))\big) + \cdots + \big(\dim(E_{\mathrm{nil}(\lambda_i)}(\lambda_i)) - \dim(E_{\mathrm{nil}(\lambda_i)-1}(\lambda_i))\big) \geq$$

$$\geq \dim(E_1(\lambda_i)) + \big(\mathrm{nil}(\lambda_i) - 1\big) \geq 1 + \big(\mathrm{nil}(\lambda_i) - 1\big) = \mathrm{nil}(\lambda_i),$$

y puesto que $\dim(M(\lambda_i)) = \mathrm{mult}_a(\lambda_i)$, deducimos que $\mathrm{nil}(\lambda_i) \leq \mathrm{mult}_a(\lambda_i)$. □

Aplicaciones del Teorema de Jordan

En esta sección seguimos suponiendo que $\mathbb{K} = \mathbb{C}$.

Endomorfismos y matrices nilpotentes

■ **Definición** Decimos que una matriz $A \in \mathcal{M}_n(\mathbb{C})$ es nilpotente si existe $k \in \mathbb{N}$ tal que $A^k = 0$. Decimos que un endomorfismo $f : V \to V$ es nilpotente si existe $k \in \mathbb{N}$ tal que $f^k = 0$.

Observación Evidentemente, si $f : V \to V$ es un endomorfismo cualquiera que tiene matriz asociada A respecto de una cierta base de V, se tiene que f es nilpotente si y solo si A es nilpotente.

■ **Proposición** *Un endomorfismo f es nilpotente si y solo si su único autovalor es 0.*

Demostración. Supongamos que $f^k = 0$ y sea $\lambda \in \mathbb{C}$ un autovalor. Entonces existe un $v \in V$ no nulo tal que $f(v) = \lambda v$, y por tanto $f^2(v) = f(\lambda v) = \lambda f(v) = \lambda^2 v$, y si hacemos esto repetidamente llegamos a que

$$f^k(v) = \lambda^k v \Rightarrow 0 = \lambda^k v \Rightarrow \lambda^k = 0 \Rightarrow \lambda = 0.$$

Supongamos ahora que el único autovalor de f es 0. Entonces por el Teorema A sabemos que $M(0) = V$, es decir, que

$$\mathrm{Ker}(f - 0I)^{\mathrm{nil}(0)} = V,$$

es decir, $f^{\mathrm{nil}(0)} = 0$ y por tanto f es nilpotente. □

Observación Si $f : V \to V$ es un endomorfismo con un único autovalor λ entonces $f - \lambda I$ es un endomorfismo cuyo único autovalor es 0 y por tanto $f - \lambda I$ es nilpotente.

■ **Corolario** *Sea $f : V \to V$ un endomorfismo. Entonces existen endomorfismos $f_d : V \to V$ diagonalizable y $f_n : V \to V$ nilpotente tales que $f = f_d + f_n$ y $f_d \circ f_n = f_n \circ f_d$.*

Demostración. Por el Teorema de Jordan existe una base B de V respecto de la cual f tiene una matriz J de Jordan. La matriz J la podemos escribir de la forma

$$J = D + N$$

donde D es la matriz diagonal cuya diagonal coincide con la de J, y donde $N = J - D$. La matriz N es triangular superior con ceros en la diagonal, y por tanto es nilpotente puesto que su único autovalor es el cero. Por tanto, si denotamos por $f_d : V \to V$ y $f_n : V \to V$ a los endomorfismos cuyas matrices respecto de B son D y N respectivamente, entonces tenemos que f_d es diagonalizable, f_n es nilpotente y $f_d \circ f_n = f_n \circ f_d$ (ya que $DN = ND$ por el ejercicio 1 de la Hoja 2). □

■ **Proposición** *Sea $f : V \to V$ un endomorfismo y sea $A \in \mathcal{M}_n(\mathbb{C})$ su matriz respecto de cierta base. Entonces f es nilpotente (y por tanto A) si y solo si*

$$\operatorname{tr}(A) = \operatorname{tr}(A^2) = \cdots \operatorname{tr}(A^n) = 0.$$

Demostración. Debemos probar que 0 es el único autovalor de f. Denotemos por $\lambda_1, \ldots, \lambda_r \in \mathbb{C}$ todos los autovalores distintos de f. Por el Teorema de Jordan, tenemos que A es semejante a su forma canónica de Jordan J. En la diagonal de J aparece cada autovalor tantas veces como su multiplicidad algebraica. Es decir, la traza de J es

$$0 = \operatorname{tr}(A) = \operatorname{tr}(J) = \operatorname{mult}_a(\lambda_1)\lambda_1 + \cdots + \operatorname{mult}_a(\lambda_r)\lambda_r.$$

De forma general, para cada $k = 1, \ldots, n$, la matriz A^k es semejante a J^k, y en la diagonal de J^k aparece λ_i^k tantas veces como su multiplicidad algebraica de λ_i, es decir,

$$0 = \operatorname{tr}(A^k) = \operatorname{tr}(J^k) = \operatorname{mult}_a(\lambda_1)\lambda_1^k + \cdots + \operatorname{mult}_a(\lambda_r)\lambda_r^k.$$

En particular, como $r \leq n$ ya que como mucho podemos tener n autovalores distintos (que es el grado del polinomio característico) si tomamos las primeras r ecuaciones anteriores obtenemos el sistema,

$$\left\{\begin{array}{ccccccc} \operatorname{mult}_a(\lambda_1)\lambda_1 & + & \cdots & + & \operatorname{mult}_a(\lambda_r)\lambda_r & = & 0, \\ & & \vdots & & & & \\ \operatorname{mult}_a(\lambda_1)\lambda_1^r & + & \cdots & + & \operatorname{mult}_a(\lambda_r)\lambda_r^r & = & 0, \end{array}\right.$$

que matricialmente podemos escribir

$$\underbrace{\begin{pmatrix} \lambda_1 & \cdots & \lambda_r \\ \vdots & & \vdots \\ \lambda_1^r & \cdots & \lambda_r^r \end{pmatrix}}_{Q} \begin{pmatrix} \operatorname{mult}_a(\lambda_1) \\ \vdots \\ \operatorname{mult}_a(\lambda_r) \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}$$

Supongamos primero que ninguno de los autovalores $\lambda_1, \ldots, \lambda_r$ es 0. Entonces hemos llegado a una contradicción puesto que Q es una matriz de Vandermonde, y como los autovalores son distintos,

$$\det(Q) = \lambda_1 \cdots \lambda_r \prod_{i<j} (\lambda_j - \lambda_i) \neq 0,$$

así que Q es invertible, y por tanto

$$\begin{pmatrix} \operatorname{mult}_a(\lambda_1) \\ \vdots \\ \operatorname{mult}_a(\lambda_r) \end{pmatrix} = Q^{-1} \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}$$

lo cual es absurdo ya que las multiplicidades son números naturales y por tanto son mayores o iguales que 1.

Por tanto uno de los autovalores es 0, y reordenando podemos asumir que $\lambda_r = 0$. Si no hay más autovalores, hemos terminado. En caso contrario, eliminando la última ecuación en el sistema anterior obtenemos que

$$\left\{ \begin{array}{ccccccc} \mathrm{mult}_a(\lambda_1)\lambda_1 & + & \cdots & + & \mathrm{mult}_a(\lambda_{r-1})\lambda_{r-1} & = & 0, \\ & & \vdots & & & & \\ \mathrm{mult}_a(\lambda_1)\lambda_1^{r-1} & + & \cdots & + & \mathrm{mult}_a(\lambda_r)\lambda_{r-1}^{r-1} & = & 0, \end{array} \right.$$

y llegamos de nuevo a una contradicción. $\square$

Ejemplo 2.12 Demostrar que el sistema de ecuaciones $\left\{ \begin{array}{ccccccc} x_1 & + & \cdots & + & x_n & = & 0 \\ x_1^2 & + & \cdots & + & x_n^2 & = & 0 \\ \vdots & & & & \vdots & & \vdots \\ x_1^n & + & \cdots & + & x_n^n & = & 0 \end{array} \right.$ no tiene más solución en $\mathbb{C}$ que la trivial. Supongamos que $x_1, \ldots, x_n \in \mathbb{C}$ es una solución del sistema. Consideremos la matriz diagonal

$$D = \begin{pmatrix} x_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & x_n \end{pmatrix}$$

Evidentemente se tiene que $\mathrm{tr}(D^i) = x_1^i + \cdots + x_n^i = 0$ para todo $1 \leq i \leq n$. Por tanto D es nilpotente, es decir, existe $k \leq 1$ tal que $D^k = 0$, lo cual implica que $x_1^k = \cdots = x_n^k = 0$ y por tanto $x_1 = \cdots = x_n = 0$.

Potencias de matrices

Sea $A \in \mathcal{M}_n(\mathbb{C})$ una matriz. Queremos calcular sus potencias A^k de forma sencilla aplicando el Teorema de Jordan. Por dicho teorema, sabemos que existe una matriz de Jordan $J \in \mathcal{M}_n(\mathbb{C})$ y una matriz $P \in \mathcal{M}_n(\mathbb{C})$ regular tales que $A = PJP^{-1}$ y por tanto $A^k = PJ^kP^{-1}$. Así que todos se reduce a calcular las potencias de matrices de Jordan. De hecho, como dichas matrices son diagonales por bloques de la forma

$$J = \left(\begin{array}{c:c:c} J_1 & & \\ \hdashline & \ddots & \\ \hdashline & & J_r \end{array} \right)$$

donde cada J_i es un cierto bloque Jordan, se tiene que

$$J^k = \left(\begin{array}{c:c:c} J_1^k & & \\ \hdashline & \ddots & \\ \hdashline & & J_r^k \end{array} \right)$$

así que todo se reduce de nuevo a calcular las potencias de los bloques de Jordan. Un bloque de Jordan es de la forma

$$M = \begin{pmatrix} \lambda & 1 & 0 & \cdots & 0 \\ 0 & \lambda & 1 & \cdots & 0 \\ \vdots & & \ddots & \ddots & \vdots \\ 0 & \cdots & \cdots & & 1 \\ 0 & \cdots & \cdots & & \lambda \end{pmatrix} \in \mathcal{M}_m(\mathbb{C})$$

el cual podemos escribir de la forma

$$M = \underbrace{\begin{pmatrix} \lambda & 0 & 0 & \cdots & 0 \\ 0 & \lambda & 0 & \cdots & 0 \\ \vdots & & \ddots & \ddots & \vdots \\ 0 & \cdots & \cdots & & 0 \\ 0 & \cdots & \cdots & & \lambda \end{pmatrix}}_{\lambda I_m} + \underbrace{\begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & & \ddots & \ddots & \vdots \\ 0 & \cdots & \cdots & & 1 \\ 0 & \cdots & \cdots & & 0 \end{pmatrix}}_{N_m}$$

En Matemáticas Básicas probamos el binomio de Newton, es decir, vimos que dados dos números reales $a, b \in \mathbb{R}$ se tiene que

$$(a+b)^k = \sum_{j=0}^{k} \binom{k}{j} a^{k-j} b^j.$$

Si uno recupera la demostración y la vuelve a leer pensando que a y b son dos matrices cuadradas que conmutan, entonces descubrirá que la prueba sigue siendo válida. En nuestro caso, como evidentemente la matrices λI_m y N_m conmutan, se tiene que

$$M^k = (\lambda I_m + N_m)^k = \sum_{j=0}^{k} \binom{k}{j} \lambda^{k-j} I_m N_m^j = \sum_{j=0}^{k} \binom{k}{j} \lambda^{k-j} N_m^j.$$

Además, la matriz N_m^j es fácil de calcular: es una matriz con todas las entradas nulas excepto en la diagonal j-ésima por encima de la diagonal principal, cuyas entradas son todas 1. En particular, para $j \geq m$ se tiene que $N_m^j = 0$.

Ejemplo 2.13 Sea

$$M = \begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix} = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix} + \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$$

Entonces se tiene que

$$\begin{aligned} M^k =& \sum_{j=0}^{k} \binom{k}{j} 2^{k-j} \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}^j = \\ =& \binom{k}{0} 2^{k-0} \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}^0 + \binom{k}{1} 2^{k-1} \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = \\ =& 2^k \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + k2^{k-1} \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 2^k & k2^{k-1} \\ 0 & 2^k \end{pmatrix}. \end{aligned} \tag{2.1}$$

Ejemplo 2.14 Calcular $\sum_{k=0}^{2000} (-1)^k M^k$ para la matriz

$$M = \begin{pmatrix} 1 & 2 & -2 \\ 1 & 1 & 0 \\ 1 & 2 & -1 \end{pmatrix}$$

Haciendo las cálculos llegamos a que la forma canónica de Jordan de la matriz anterior es

$$J = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix} \text{ con matriz de paso } P = \begin{pmatrix} 2 & 0 & 1 \\ -1 & 1 & 0 \\ 1 & 1 & 0 \end{pmatrix}$$

Se tiene que

$$J^k = \begin{pmatrix} (-1)^k & 0 & 0 \\ 0 & 1 & k \\ 0 & 0 & 1 \end{pmatrix}$$

y por tanto

$$\sum_{k=0}^{2000} (-1)^k J^k = \begin{pmatrix} \sum_{k=0}^{2000} (-1)^k (-1)^k & 0 & 0 \\ 0 & \sum_{k=0}^{2000} (-1)^k & \sum_{k=0}^{2000} (-1)^k k \\ 0 & 0 & \sum_{k=0}^{2000} (-1)^k \end{pmatrix} =$$

$$= \begin{pmatrix} 2001 & 0 & 0 \\ 0 & 1 & \sum_{k=0}^{2000} (-1)^k k \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 2001 & 0 & 0 \\ 0 & 1 & 1000 \\ 0 & 0 & 1 \end{pmatrix}$$

ya que agrupando de dos en dos nos damos cuenta de que $\sum_{k=0}^{2000} (-1)^k k = 1000$. Así que finalmente

$$\sum_{k=0}^{2000} (-1)^k M^k = P\big(\sum_{k=0}^{2000} (-1)^k J^k\big) P^{-1} = \begin{pmatrix} 1 & -2000 & 2000 \\ 1000 & 2001 & -2000 \\ 1000 & 0 & 1 \end{pmatrix}.$$

Ejemplo 2.15 Queremos calcular

$$M = \begin{pmatrix} 2 & -1 \\ 1 & 0 \end{pmatrix}^{15}$$

Haciendo los cálculos llegamos a que su forma canónica de Jordan es

$$J = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \text{ con matriz de paso } P = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}$$

Se tiene entonces que

$$A^{15} = PJ^{15}P^{-1} = P \begin{pmatrix} 1 & 15 \\ 0 & 1 \end{pmatrix} P^{-1} = \begin{pmatrix} 16 & -15 \\ 15 & -14 \end{pmatrix}$$

Ejemplo 2.16 ¿Es diagonalizable un endomorfismo f de $\mathbb{C}^2$ tal que f^k es la identidad de $\mathbb{C}^2$ para cierto entero $k \geq 1$? Lo primera observación obvia es que si $k = 1$ entonces la respuesta es sí, f es diagonalizable. Así que podemos suponer que $k > 1$. Por el Teorema de Jordan sabemos que existe una base de $\mathbb{C}^2$ respecto de la cual f tiene una matriz J de Jordan. Si f no es diagonalizable entonces la única posibilidad es que J sea de la forma

$$J = \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix}$$

Por otro lado sabemos que f^k es la identidad, es decir, que

$$J^k = I \quad \Rightarrow \quad \begin{pmatrix} \lambda^k & k\lambda^{k-1} \\ 0 & \lambda^k \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

y por tanto $k\lambda^{k-1} = 0$ de lo que deducimos que $\lambda = 0$. Pero entonces también tendríamos que $0 = \lambda^k = 1$, absurdo. Por tanto la respuesta final es sí, f debe ser diagonalizable.

Semejanza de matrices y el Teorema de Jordan real

Semejanza de matrices

El Teorema de Jordan no solo nos asegura que una matriz compleja es semejante a una matriz de Jordan, sino que nos proporciona también un criterio para decidir cuando dos matrices van a ser semejantes.

■ **Teorema** *Dos matrices $A, C \in \mathcal{M}_n(\mathbb{C})$ son semejantes si y solo si sus formas canónicas de Jordan (salvo permutación de los bloques) son iguales.*

Demostración. Si A y C tienen la misma forma de Jordan J entonces existen matrices regulares $P, Q \in \mathcal{M}_n(\mathbb{C})$ tales que $A = PJP^{-1}$ y $C = QJQ^{-1}$. En particular se tiene que $A = PJP^{-1} = P(Q^{-1}CQ)P^{-1} = (PQ^{-1})C(PQ^{-1})^{-1}$ y por tanto A y C son semejantes. Por el contrario, si A y C son semejantes entonces sus autovalores son iguales y las dimensiones de los subespacios propios generalizados de cada autovalor también son iguales: en particular sus formas de Jordan deben ser iguales (porque ya vimos que dichas formas de Jordan vienen determinadas precisamente por los autovalores y por las dimensiones de los subespacios propios generalizados de cada autovalor). □

¿Y si nuestras matrices no eran complejas, y si eran reales? ¿Tenemos un criterio para saber cuando van a ser semejante sobre los reales?

■ **Teorema** *Dos matrices $A, C \in \mathcal{M}_n(\mathbb{R})$ son semejantes en $\mathcal{M}_n(\mathbb{R})$ si y solo si son semejantes en $\mathcal{M}_n(\mathbb{C})$. En particular, A y C son semejantes en $\mathcal{M}_n(\mathbb{R})$ si y solo si tienen la misma forma de Jordan.*

Demostración. Si A y C son semejantes en $\mathcal{M}_n(\mathbb{R})$ entonces existe $P \in \mathcal{M}_n(\mathbb{R})$ regular tal que $A = PCP^{-1}$. En particular, A y C son evidentemente semejantes en $\mathcal{M}_n(\mathbb{C})$.

Supongamos ahora que A y C son semejantes en $\mathcal{M}_n(\mathbb{C})$, es decir, que existe $P \in \mathcal{M}_n(\mathbb{C})$ regular tal que $A = PCP^{-1}$. Queremos mostrar que podemos encontrar una matriz $Q \in \mathcal{M}_n(\mathbb{R})$ regular tal que $A = QCQ^{-1}$. La matriz P la podemos escribir de la forma $P = P_1 + iP_2$ donde $P_1, P_2 \in \mathcal{M}_n(\mathbb{R})$. Puesto que $AP = PC$, deducimos que

$$A(P_1 + iP_2) = (P_1 + iP_2)C \Rightarrow AP_1 + iAP_2 = P_1C + iP_2C \Rightarrow$$

$$\Rightarrow AP_1 = P_1C \quad \text{y} \quad AP_2 = P_2C$$

Sin embargo debemos tener cuidado, ya que no sabemos si P_1 y P_2 son regulares. En cualquier caso, dado un número cualquiera $\beta \in \mathbb{R}$, tenemos que

$$A(P_1 + \beta P_2) = AP_1 + \beta AP_2 = P_1C + \beta P_2C = (P_1 + \beta P_2)C,$$

así que nuestro objetivo ahora es probar que podemos encontrar $\beta \in \mathbb{R}$ tal que $P_1 + \beta P_2$ es regular, o lo que es lo mismo, tal que $\det(P_1 + \beta P_2) \neq 0$. Evidentemente, la función $F(T) = \det(P_1 + TP_2)$ es un polinomio con coeficientes reales. Supongamos que lo que queremos probar es falso, es decir, que

$$F(\beta) = \det(P_1 + \beta P_2) = 0 \quad \text{para todo } \beta \in \mathbb{R}.$$

Entonces el polinomio $F(T)$ debe ser el polinomio nulo, todos sus coeficientes deben ser 0. Ahora bien, este polinomio lo podemos evaluar en la raíz imaginaria i, y puesto que todos los coeficientes hemos dicho que eran 0, nos quedará $F(i) = 0$. Pero por otro lado, como la matriz P es regular obtenemos que

$$F(i) = \det\det(P_1 + iP_2) = \det(P) \neq 0$$

lo cual es una contradicción. Así pues, debe existir $\beta \in \mathbb{R}$ tal que la matriz $Q = P_1 + \beta P_2$ es regular y en particular A y C son semejantes en $\mathcal{M}_n(\mathbb{R})$ puesto que ya vimos que

$$A(P_1 + \beta P_2) = (P_1 + \beta P_2)C \Rightarrow AQ = QC \Rightarrow A = QCQ^{-1}.$$

□

El Teorema de Jordan en el caso real

Dado un endomorfismo $f : \mathbb{R}^n \to \mathbb{R}^n$, *si todas las raíces del polinomio característico son reales*, entonces el Teorema de Jordan nos asegura que podemos encontrar una base B de $\mathbb{R}^n$ respecto de la cual la matriz de f es una matriz de Jordan $J \in \mathcal{M}_n(\mathbb{R})$. Es decir, que si la matriz de f respecto de la base canónica es A, y la matriz de paso de B a la base canónica es $P \in \mathcal{M}_n(\mathbb{R})$ entonces tendremos que $A = PJP^{-1}$.

¿Pero y si nuestro endomorfismo $f : \mathbb{R}^n \to \mathbb{R}^n$ no tiene la propiedad de que todas las raíces del polinomio característico son reales? Bueno, al menos siempre nos podemos imaginar que su matriz $A \in \mathcal{M}_n(\mathbb{R})$ respecto de la base canónica es la matriz de su complexificado $\widetilde{f} : \mathbb{C}^n \to \mathbb{C}^n$ y por tanto, como todas las raíces del polinomio característico $\det(A - \lambda I)$ sí que pertenecen a $\mathbb{C}$, seremos capaces de encontrar una matriz $J \in \mathcal{M}_n(\mathbb{C})$ de Jordan y una matriz $P \in \mathcal{M}_n(\mathbb{C})$ regular tales que $A = PJP^{-1}$. Eso sí, no podemos escapar al hecho de que tanto J como P *son* complejas. Sin embargo, esto sigue siendo suficientemente bueno para las aplicaciones: por ejemplo, podemos seguir usando la igualdad $A = PJP^{-1}$ para calcular las potencias de A, incluso aunque J y P sean complejas.

¿Y no podemos encontrar una base de $\mathbb{R}^n$ respecto de la cual f tenga una matriz real lo más sencilla posible? Sí, lo vemos con un ejemplo. Aunque antes, hagamos una observación útil:

Observación Sea $f : \mathbb{C}^n \to \mathbb{C}^n$ un endomorfismo que tiene respecto de la base canónica una matriz A real. Ya vimos en su día que si $\lambda \in \mathbb{C}$ es un autovalor de f, entonces su conjugado $\overline{\lambda} \in \mathbb{C}$ también lo es. De hecho, vimos que si $v = (x_1, \ldots, x_n) \in E_1(\lambda)$ entonces

$$f(v) = \lambda v \Rightarrow (A - \lambda I) \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}$$

y dado que el conjugado de un producto y el conjugado de una suma es el producto y la suma de los conjugados respectivamente, y el conjugado de un número real es el mismo,

$$(A - \overline{\lambda} I) \begin{pmatrix} \overline{x}_1 \\ \vdots \\ \overline{x}_n \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}$$

así que deducimos que $\overline{\lambda}$ también es autovalor y que $\overline{v} = (\overline{x_1}, \ldots, \overline{x_n})$ es un autovector asociado a $\overline{\lambda}$. Haciendo el cálculo del revés podemos probar que $E_1(\overline{\lambda}) = \overline{E_1(\lambda)}$. Es más, de forma muy parecida podemos mostrar que $E_k(\overline{\lambda}) = \overline{E_k(\lambda)}$, y en particular, $M(\overline{\lambda}) = \overline{M(\lambda)}$.

Por último observamos que si tenemos una base B de $M(\lambda)$ respecto de la cual $f|_{M(\lambda)}$ tiene matriz de Jordan, entonces si tomamos conjugados de los vectores de esa base obtendremos una base $\overline{B}$ de $M(\overline{\lambda})$ respecto de la cual $f|_{M(\overline{\lambda})}$ tendrá matriz de Jordan (que encima será la conjugada de la anterior). Claro, para la base B tendremos un diagrama de puntos como el que aprendimos un su momento, pero si tomamos conjugados, volveremos a obtener una base de $M(\overline{\lambda})$ con el mismo diagrama de puntos, dado que la flecha que en el diagrama de B representaba aplicar $A - \lambda I$, ahora significará aplicar $\overline{A - \lambda I} = A - \overline{\lambda} I$.

Ejemplo 2.17 Sea $f : \mathbb{R}^4 \to \mathbb{R}^4$ el endomorfismo con matriz

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

respecto de la canónica. Calculemos su matriz de Jordan compleja. El polinomio característico es

$$\det(A - \lambda I) = \text{ ejercicio para casa } = (\lambda^2 - 2x + 2)^2 = \big(\lambda - (1+i)\big)^2 \big(\lambda - (1-i)\big)^2$$

Así que tenemos dos autovalores $\lambda = 1 + i$ y $\overline{\lambda} = 1 - i$. Recordamos que no ha sido una casualidad que uno sea el conjugado del otro. Si calculamos los subespacios propios generalizados del autovalor $\lambda = 1 + i$ nos va a quedar

$$\begin{cases} E_1(1+i) = L[(0,0,1,-1-i)], \\ M(1+i) = E_2(1+i) = L[(0,0,1,-1-i),(1,-1-i,0,-1+i)] \end{cases}$$

y si aplicamos el método aprendido para calcular una base del subespacio máximo $M(1+i)$ respecto de la cual nos quede una matriz de Jordan, llegamos a que debemos tomar los vectores

$$\{(0,0,1,-1-i),(1,-1-i,0,-1+i)\}.$$

Ahora, para calcular los subespacios propios generalizados del autovalor $\lambda = 1 - i$ nos basta con hacer conjugados a los anteriores

$$\begin{cases} E_1(1-i) = L[(0,0,1,-1+i)], \\ M(1-i) = E_2(1-i) = L[(0,0,1,-1+i),(1,-1+i,0,-1-i)] \end{cases}$$

y de nuevo si aplicamos el método aprendido para calcular una base del subespacio máximo $M(1-i)$ respecto de la cual nos quede una matriz de Jordan, llegamos a que debemos tomar los vectores (o bien tomando conjugados a la base de $M(1+i)$ que calculamos anteriormente)

$$\{(0,0,1,-1+i),(1,-1+i,0,-1-i)\}.$$

Por tanto la matriz de Jordan de A respecto de la base

$$B = \{\underbrace{(0,0,1,-1-i)}_{w_1}, \underbrace{(1,-1-i,0,-1+i)}_{w_2}, \underbrace{(0,0,1,-1+i)}_{\overline{w_1}}, \underbrace{(1,-1+i,0,-1-i)}_{\overline{w_2}}\}$$

es

$$\begin{pmatrix} 1+i & 1 & 0 & 0 \\ 0 & 1+i & 0 & 0 \\ 0 & 0 & 1-i & 1 \\ 0 & 0 & 0 & 1-i \end{pmatrix}.$$

Pero nuestro objetivo ahora es conseguir una base de $\mathbb{R}^4$ respecto de la cual la matriz A tenga una matriz lo más sencilla posible.

La idea es encontrar una base especial, no para $M(\lambda)$ y $M(\overline{\lambda})$ por separado, sino para su suma $M(\lambda) \bigoplus M(\overline{\lambda})$. Podemos adelantar que nuestra intención es que la matriz de f respecto de esta base que queremos encontrar sea

$$\begin{pmatrix} \mathrm{Re}(\lambda) & \mathrm{Im}(\lambda) & 1 & 0 \\ -\mathrm{Im}(\lambda) & \mathrm{Re}(\lambda) & 0 & 1 \\ 0 & 0 & \mathrm{Re}(\lambda) & \mathrm{Im}(\lambda) \\ 0 & 0 & -\mathrm{Im}(\lambda) & \mathrm{Re}(\lambda) \end{pmatrix} = \begin{pmatrix} 1 & 1 & 1 & 0 \\ -1 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & -1 & 1 \end{pmatrix}$$

De hecho, vamos a fijarnos en la parte real y la parte imaginaria de los vectores w_1 y w_2. En efecto, podemos escribir

$$\underbrace{(0,0,1,-1-i)}_{w_1} = \underbrace{(0,0,1,-1)}_{v_1} + i\,\underbrace{(0,0,0,-1)}_{v_2}$$

$$\underbrace{(1,-1-i,0,-1+i)}_{w_2} = \underbrace{(1,-1,0,-1)}_{v_3} + i\,\underbrace{(0,-1,0,1)}_{v_4}$$

Observamos que seguro que los vectores v_1, v_2, v_3, v_4 pertenecen a $M(\lambda) \bigoplus M(\overline{\lambda})$ ya que podemos escribir

$$\begin{array}{lcl} v_1 = \frac{1}{2}(w_1 + \overline{w_1}) & \in & M(\lambda) \bigoplus M(\overline{\lambda}) \\ v_2 = \frac{1}{2i}(w_1 - \overline{w_1}) & \in & M(\lambda) \bigoplus M(\overline{\lambda}) \\ v_3 = \frac{1}{2}(w_2 + \overline{w_2}) & \in & M(\lambda) \bigoplus M(\overline{\lambda}) \\ v_4 = \frac{1}{2i}(w_2 + \overline{w_2}) & \in & M(\lambda) \bigoplus M(\overline{\lambda}) \end{array}$$

Además, v_1, v_2, v_3, v_4 generan todo $M(\lambda) \bigoplus M(\overline{\lambda})$ ya que evidentemente w_1, w_2, w_3, w_4 son combinación de ellos (así que en particular son linealmente independientes puesto que $M(\lambda) \bigoplus M(\overline{\lambda})$ tiene dimensión 4). Deducimos por tanto que

$$B' = \{v_1, v_2, v_3, v_4\}$$

es una base de $M(\lambda) \bigoplus M(\overline{\lambda})$. Veamos que esta base es la que estábamos buscando, calculemos la matriz de f respecto de B'. Observamos que

$$f(w_1) = \lambda w_1 \Rightarrow f(v_1 + iv_2) = \lambda(v_1 + iv_2) \Rightarrow f(v_1 + iv_2) = \big(\mathrm{Re}(\lambda) + i\mathrm{Im}(\lambda)\big)(v_1 + iv_2)$$

$$\Rightarrow f(v_1) + i(v_2) = \big(\mathrm{Re}(\lambda)v_1 - \mathrm{Im}(\lambda)v_2\big) + i\big(\mathrm{Im}(\lambda)v_1 + \mathrm{Re}(\lambda)v_2\big)$$

y por tanto

$$\begin{cases} f(v_1) = \mathrm{Re}(\lambda)v_1 - \mathrm{Im}(\lambda)v_2 = (\mathrm{Re}(\lambda), -\mathrm{Im}(\lambda), 0, 0)_{B'}, \\ f(v_2) = \mathrm{Im}(\lambda)v_1 + \mathrm{Re}(\lambda)v_2 = (\mathrm{Im}(\lambda), \mathrm{Re}(\lambda), 0, 0)_{B'}. \end{cases}$$

Por último,

$$f(w_2) = \lambda w_2 + w_1 \Rightarrow f(v_3 + iv_4) = \big(\mathrm{Re}(\lambda) + i\mathrm{Im}(\lambda)\big)(v_3 + iv_4) + (v_1 + iv_2) \Rightarrow$$

$$\Rightarrow f(v_3) + if(v_4) = \big(\mathrm{Re}(\lambda)v_3 - \mathrm{Im}(\lambda)v_4\big) + i\big(\mathrm{Im}(\lambda)v_3 + \mathrm{Re}(\lambda)v_4\big) + (v_1 + iv_2)$$

$$\Rightarrow f(v_3) + if(v_4) = \big(\mathrm{Re}(\lambda)v_3 - \mathrm{Im}(\lambda)v_4 + v_1\big) + i\big(\mathrm{Im}(\lambda)v_3 + \mathrm{Re}(\lambda)v_4 + v_2\big)$$

y por tanto

$$\begin{cases} f(v_3) = \mathrm{Re}(\lambda)v_3 - \mathrm{Im}(\lambda)v_4 + v_1 = (1, 0, \mathrm{Re}(\lambda), -\mathrm{Im}(\lambda))_{B'}, \\ f(v_4) = \mathrm{Im}(\lambda)v_3 + \mathrm{Re}(\lambda)v_4 + v_2 = (0, 1, \mathrm{Im}(\lambda), \mathrm{Re}(\lambda))_{B'}. \end{cases}$$