6. Formas canónicas de Jordan

6.1. Polinomios aplicados a operadores

Dado un espacio vectorial V sobre un cuerpo $\mathbb{K}$, como ya vimos en la unidad de transformaciones lineales, los operadores lineales $\mathcal{L}(V)$ juegan un rol destacado. Para $T \in \mathcal{L}(V)$, $T \circ T = TT = T^2$ también es un operador en $\mathcal{L}(V)$.

Definición 6.1 *Sea $T \in \mathcal{L}(V)$ y $m \in \mathbb{Z}^+$. Entonces*

- *$T^m \in \mathcal{L}(V)$ es la composición m veces de T, i.e. $T^m = \underbrace{T \cdots T}_{m \text{ veces}}$.*
- *$T^0 = I$, operador identidad en $\mathcal{L}(V)$.*
- *Si T es inversible con inversa T^{-1}, entonces $T^{-m} \in \mathcal{L}(V)$ está definido por $T^{-m} = (T^{-1})^m$.*

Observación 6.1 *Para $T \in \mathcal{L}(V)$, se verifica: $T^mT^n = T^nT^m = T^{m+n}$, $\quad (T^m)^n = T^{mn}$, para $m, n \in \mathbb{Z}$ si T es inversible, para $m, n \in \mathbb{N}$ para T no inversible.*

Definición 6.2 *Sea $T \in \mathcal{L}(V)$ y $p \in \mathbb{K}[x]$ un polinomio dado por*

$$p(z) = a_0 + a_1 z + a_2 z^2 + \cdots + a_m z^m.$$

Luego $p(T)$ es un operador en $\mathcal{L}(V)$ definido por

$$p(T) = a_0 I + a_1 T + a_2 T^2 + \cdots + a_m T^m.$$

Fijado un operador $T \in \mathcal{L}(V)$, la aplicación de $\mathbb{K}[x]$ en $\mathcal{L}(V)$ que a un polinomio dado $p \mapsto p(T)$ es una transformación lineal (verlo).

Definición 6.3 *Sean $p, q \in \mathbb{K}[x]$, luego $pq \in \mathbb{K}[z]$ es el polinomio que se define como $(pq)(z) = p(z)q(z), \forall z \in \mathbb{K}$.*

Proposición 6.1 *Sean $p, q \in \mathbb{K}[x]$ y $T \in \mathcal{L}(V)$. Luego*

1. *$(pq)(T) = p(T)q(T), \forall T \in \mathcal{L}(V)$.*
2. *$p(T)q(T) = q(T)p(T), \forall T \in \mathcal{L}(V)$.*

Demos: Ejercicio. ■

Proposición 6.2 *Sean $T \in \mathcal{L}(V)$ y $p \in \mathbb{K}[x]$. Luego $nul(p(T))$ y $img(p(T)) = p(T)(V)$ son subespacios invariantes bajo T.*

Demos: Sea $u \in nul(p(T))$ i.e. $p(T)(u) = 0$. Luego

$$0 = T(0) = T(p(T)(u)) = (p(T)T)(u) = p(T)(Tu),$$

Por lo tanto $T(u) \in nul(p(T))$ y $nul(T)$ es invariante bajo T.

Sea $u \in img(p(T))$, luego existe un $v \in V$ tal que $u = p(T)(v)$. Así $Tu = T(p(T)v) = p(T)(Tv)$. Concluimos que $Tu \in img(p(T))$. ■

Observación 6.2 *Sean V espacio vectorial de dimensión finita sobre $\mathbb{K}$, $T \in \mathcal{L}(V)$. Se dice que un polinomio $p \in \mathbb{K}[x]$ anula a T si el operador $p(T) = 0$. La colección $M \underset{s.e.}{\subseteq} \mathbb{K}[x]$ de polinomios p que anulan a T es un ideal en el álgebra de los polinomios $\mathbb{K}[x]$ (es decir que $pq \in M, \forall p \in M, q \in \mathbb{K}[x]$.*

Proposición 6.3 *Sean V espacio vectorial de dimensión finita sobre $\mathbb{K}$, $T \in \mathcal{L}(V)$. Se tiene que, el ideal $M \underset{s.e.}{\subseteq} \mathbb{K}[x]$ de polinomios que anulan a T es $M \neq \{0\}$. Es decir que existe un polinomio distinto del polinomio nulo que anula a T.*

Demos: Consideremos las primeras (n^2+1) potencias de T

$$I, T, T^2, \cdots, T^{n^2}.$$

Esta es una colección de (n^2+1) operadores en $\mathcal{L}(V)$, cuya dimensión es n^2. Por lo tanto nuestra colección es linealmente dependiente, con lo cual

$$c_0 I + c_1 T + c_2 T^2 + \cdots + c_{n^2} T^{n^2} = 0,$$

con c_i no todos nulos. En consecuencia hay un polinomio de grado menor o igual que n^2 que anula a T. ■

Definición 6.4 *Un **polinomio mónico** es un polinomio tal que el coeficiente del término de mayor grado es igual a 1.*

Proposición 6.4 *Sean V espacio vectorial de dimensión finita sobre $\mathbb{K}$, $T \in \mathcal{L}(V)$. Existe un único polinomio mónico $p \in \mathbb{K}[x]$ de menor grado tal que $p(T) = 0$. Más aún $grad(p) \leqslant dim(V)$.*

Demos: Para $dim(V) = 0$, la identidad I es el operador nulo en V y tomamos como p el polinomio constante 1.
Ahora hacemos inducción sobre $dim(V)$. Supongamos $dim(V) > 0$ y supongamos que es cierta la tesis para todos los operadores definidos en espacios vectoriales de dimensión menor. Sea $0 \neq u \in V$. Consideremos el conjunto linealmente dependiente en V

$$\{u, Tu, \cdots, T^{dim(V)}u\}.$$

Luego existe un $m \leqslant dim(V)$, el menor, tal que $T^m u$ es una combinación lineal de $u, Tu, \cdots, T^{m-1}u$ (es más dicho conjunto es l.i.). Así existen escalares $c_0, c_-1, dots, c_{m-1} \in \mathbb{K}$ tales que

$$c_0 u + c_1 Tu + \cdots + c_{m-1} T^{m-1} u + T^m u = 0. \tag{15}$$

Definamos el polinomio mónico $q \in \mathbb{K}[x]$ como

$$q(x) = c_0 + c_1 x + \cdots + c_{m-1} x^{m-1} + x^m.$$

Entonces la ecuación (15) implica que $q(T)u = 0$.
Si $k \in \mathbb{N}$, se tiene que $q(T)(T^k u) = T^k(q(T)u) = T^k(0) = 0$. Esto más el hecho que $u, Tu, \cdots, T^{m-1}u$ es l.i. implica que $dim(nulq(T)) \geqslant m$. Así

$$rg(q(T)) = dim(V) - dim(nul(q(T)) \leqslant dim(V) - m.$$

Como $img(q(T))$ es invariante bajo T, podemos usar nuestra hipótesis de inducción en el operador $T|_{img(q(T))}$ en el espacio vectorial $img(q(T))$. Luego existe un polinomio mónico $s \in \mathbb{K}[x]$ con

$$grad(s) \leqslant dim(V) - m, \qquad s(T|_{img(q(T))} = 0.$$

Luego para todo $v \in V$ se tiene que

$$((sq)(T))(v) = s(T)(q(T)v) = 0,$$

pues $q(T)v \in img(q(T))$ y $s(T)|img(q(T)) = s(T|img(q(T))) = 0$. Así sq es un polinomio mónico tal que $grad(sq) \leqslant dim(V)$ y $sq(T) = 0$.

Teniendo la existencia de un polinomio mónico de grado menor que $dim(V)$ que anula a T podemos afirmar que existe uno de menor grado que también anula a T con lo que queda probada la existencia.

Veamos la unicidad. Supongamos que $p, r \in \mathbb{K}[x]$ sos polinomios mónicos de menor grado tal que $p(T) = r(T) = 0$. Luego $(p-r)(T) = 0$ y $grad(p-r) < grad(p)$. Así si $p - r$ no es el polinomio nulo dividiendo por el coeficiente de mayor grado de $(p-r)$ existiría un polinomio mónico de grado menor que $grad(p)$ que anularía T. Luego $p - r = 0$. ■

Definición 6.5 *Sean V espacio vectorial de dimensión finita sobre $\mathbb{K}$, $T \in \mathcal{L}(V)$. El* ***polinomio minimal de T*** *es el único polinomio mónico $p \in \mathbb{K}[x]$ de menor grado tal que $p(T) = 0$.*

De manera análoga se tiene

Definición 6.6 *Sean ł $\in \mathcal{M}_n(\mathbb{K})$. El* ***polinomio minimal de A*** *es el único polinomio mónico $p \in \mathbb{K}[x]$ de menor grado tal que $p(A) = 0$.*

Proposición 6.5 *Si un operador $T \in \mathcal{L}(V)$ tiene asociada una matriz A fijada una cierta base ordenada del espacio vectorial V entonces T y A tienen el mismo polinomio minimal. Además todas las matrices semejantes tienen el mismo polinomio minimal.*

Demos: Para cualquier polinomio $f \in \mathbb{K}[x]$, el operador $f(T)$ tiene como matriz asociada en la misma base fijada a la matriz $f(A)$. Luego $f(T) = 0$ si y solo si $f(A) = 0$.
Si B es una matriz semejante a A, existe una matriz inversible P, tal que $B = P^{-1}AP$. Luego para cualquier polinomio $f \in \mathbb{K}[x]$ se tiene que $f(B) = f(P^{-1}AP) = P^{-1}f(A)P$. ■

Teorema 6.1 *Sea V espacio vectorial de dimensión finita sobre $\mathbb{K}$ y $T \in \mathcal{L}(V)$. Las raíces del polinomio minimal de T son los autovalores de T. Si $\mathbb{K} = \mathbb{C}$, entonces el polinomio minimal de T tiene la forma*

$$p(z) = (z - \lambda_1) \cdots (z - \lambda_m)$$

siendo $\lambda_1, \cdots, \lambda_m$ los autovalores de T, posiblemente con repetición.

Demos: Supongamos que λ es un cero de p, polinomio minimal de T. Luego podemos escribir

$$p(z) = (z - \lambda)q(z),$$

con $q \in \mathbb{K}[z]$ mónico. Como $p(T) = 0$ para todo $v \in V$ se tiene que

$$0 = p(T)v = (T - \lambda I)(q(T)v).$$

Luego, ya que $grad(q) = grad(p) - 1$ y p es el polinomio minimal de T debe existir al menos un vector $v \in V$, tal que $q(T)v \neq 0$ y la ecuación anterior implica que λ es un autovalor de T.
Recíprocamente, si λ es un autovalor de T, existe algún $0 \neq v \in V$ tal que $Tv = \lambda v$. Aplicando reiteradas veces T a ambos miembros de esta igualdad obtenemos que $T^k v = \lambda^k v$ para todo k entero no negativo. Así

$$p(T)v = p(\lambda)v.$$

Como p es el polinomio minimal de T se tiene que $P(T)v = 0$ y en consecuencia, como $v \neq 0$ resulta $p(\lambda) = 0$, como queríamos ver.
La observación para $\mathbb{K} = \mathbb{C}$ surge de aplicar el teorema fundamental del álgebra. ■

Teorema 6.2 *Sea V un espacio vectorial de dimensión finita sobre $\mathbb{K}$, $T \in \mathcal{L}(V)$ y $q \in \mathbb{K}[x]$. Luego $q(T) = 0$ si y solo si q es un polinomio múltiplo del polinomio minimal de T.*

Demos: Sea $p \in \mathbb{K}[x]$ el polinomio minimal de T. Por el algoritmo de la división para polinomios sabemos que existen polinomios $s, r \in \mathbb{K}[x]$ tales que

$$q = ps + r, \qquad\qquad grad(r) < grad(p).$$

Así se tiene que

$$0 = q(T) = p(T)s(T) + r(T) = r(T).$$

Esta igualdad implica que el polinomio $r = 0$, pues sino habría una contradicción con el hecho de que p es el polinomio minimal para T. De este modo se obtiene que $q = ps$, con lo cual q es múltiplo de p.
Recíprocamente, si q es un polinomio múltiplo de p, es decir existe un $s \in \mathbb{K}[x]$ tal que $q = ps$, se tiene que

$$q(T) = p(T)s(T) = 0s(T) = 0,$$

como queríamos probar. ■

6.2. Invariancia

Definición 6.7 *Sea $T : V \to V$ una transformación lineal. Un subespacio W de V se dice* ***invariante por T*** *si T aplica a W en si mismo, i.e., si $v \in W \Rightarrow T(v) \in W$. En este caso T restringido a W define un operador lineal,*

$$\begin{array}{rccl} \hat{T}: & W & \to & W \\ & w & \mapsto & \hat{T}(w) = T(w). \end{array}$$

Ejemplo 6.1 *Sea $T : \mathbb{R}^3 \to \mathbb{R}^3$, tal que $T(x, y, z) = (x \cos\theta - y \operatorname{sen}\theta, x \operatorname{sen}\theta + y \cos\theta, z)$. Es el operador lineal que rota cada vector alrededor del eje z un ángulo θ.*
Observemos que cada vector $w = (a, b, 0) \in W$ donde W es el plano xy permanece en W al aplicarle la transformación T, luego W es invariante por T.
También resulta invariante el eje z.

Ejemplo 6.2 *Los vectores propios de un operador lineal $T : V \to V$ pueden caracterizarse como generadores de subespacios invariantes por T de dimensión 1.*
Supongamos que $T(v) = \lambda v$, $v \neq 0$, entonces $W = \{\alpha v : \alpha \in \mathbb{K}\}$ es invariante por T pues

$$T(\alpha v) = \alpha T v = \alpha(\lambda v) = (\lambda\alpha)v \in W.$$

Recíprocamente, supongamos que $\dim U = 1$, $U =< u >$ y U invariante por T. Entonces como $T(u) \in U$ resulta $T(u) = \beta u$ para algún $\beta \in \mathbb{K}$, con lo que resulta u un autovector de T.

Teorema 6.3 *Sea W un subespacio invariante de $T : V \to V$, espacio vectorial sobre $\mathbb{K}$ de dimensión finita. Entonces T tiene una representación matricial por bloques*

$$\begin{bmatrix} A & B \\ 0 & C \end{bmatrix},$$

donde A es una representación matricial de la restricción de T a W.

Demos: Sea una base ordenada $\{w_1, \cdots, w_r\}$ de W y la extendemos a una base $\{w_1, \cdots, w_r, v_1, \cdots, v_s\}$ de V. Se tiene

$$\begin{array}{rcccl}
\hat{T}(w_1) & = & T(w_1) & = & a_{11}w_1 + \cdots + a_{1r}w_r \\
& \vdots & & & \\
\hat{T}(w_r) & = & T(w_r) & = & a_{r1}w_1 + \cdots + a_{rr}w_r \\
& & T(v_1) & = & b_{11}w_1 + \cdots + b_{1r}w_r + c_{11}v_1 + \cdots + c_{1s}v_s \\
& & & \vdots & \\
& & T(v_s) & = & b_{s1}w_1 + \cdots + b_{sr}w_r + c_{s1}v_1 + \cdots + c_{ss}w_s
\end{array}$$

Pero la matriz de T en esta base es la transpuesta de la matriz de los coeficientes en el sistema anterior de ecuaciones. Por lo tanto tiene la forma

$$\begin{bmatrix} A & B \\ 0 & C \end{bmatrix}.$$

Por el mismo argumento A es la matriz de $\hat{T}$ relativa a la base $\{w_i\}$ de W. ■

6.3. DESCOMPOSICIÓN EN SUMA DIRECTA DE ESPACIOS INVARIANTES

Recordemos que

Definición 6.8 *Se dice que un espacio vectorial V es la* **suma directa** *de sus subespacios $W_1, \cdots, W_r$, si todo vector $v \in V$ puede escribirse de manera única como*

$$v = w_1 + \cdots + w_r, w_i \in W_i,\ i = 1, \cdots, r.$$

tal que $W_i \cap W_j = \{0\}, \forall i \neq j$. Se nota $V = W_1 \oplus \cdots \oplus W_r$.

Teorema 6.4 *Sean $W_1, \cdots, W_r$ subespacios de V, y supongamos que $B_1 = \{w_1^1, \cdots, w_{n_1}^1\}, \cdots, B_r = \{w_1^r, \cdots, w_{n_r}^r\}$ son bases de $W_1, \cdots, W_r$ respectivamente. Entonces V es la suma directa de los W_i si y sólo si $B = \{w_1^1, \cdots, w_{n_1}^1, \cdots, w_1^r, \cdots, w_{n_r}^r\}$ es una base de V.*

Demos: $\Leftarrow$) Supongamos que B es base de V, luego para todo $v \in V$ existen escalares tales que

$$v = a_{11}w_1^1 + \cdots + a_{1n_1}w_{n_1}^1 + \cdots + a_{r1}w_1^r + \cdots + a_{rn_1}w_{n_1}^r = w_1 + \cdots + w_r,$$

con $w_i = a_{i1}w_1^i + \cdots + a_{in_i}w_{n_1}^i \in W_i$.
Veamos que la suma es única. Supongamos que $v = \tilde{w}_1 + \cdots + \tilde{w}_r$ con $\tilde{w}_i \in W_i$. Usando la base correspondiente a cada W_i se tendrá que existen escalares tales que $\tilde{w}_i = b_{i1}w_1^i + \cdots + b_{in_i}w_{n_i}^i$, luego se tiene que

$$v = b_{11}w_1^1 + \cdots + b_{1n_1}w_{n_1}^1 + \cdots + b_{r1}w_1^r + \cdots + b_{rn_1}w_{n_1}^r.$$

Como B es una base de V, resulta $a_{ij} = b_{ij}$ para cada i y j, de ese modo los términos de la suma de v son únicos y así resulta V suma directa de $W_i, \cdots, W_r$.
$\Rightarrow$) Supongamos que V suma directa de $W_i, \cdots, W_r$. Luego todo $v \in V$ puede expresarse como $v = v_1 + \cdots + v_r$ con $w_i \in W_i$ únicos. Como B_i es base de W_i, cada v_i es combinación lineal de los vectores $\{w_1^i, \cdots, w_{n_i}^i\}$, resultando así v combinación lineal de los elementos de B por lo tanto $V =< B >$.
Veamos que los vectores en B son *l.i.*. Consideremos

$$0 = a_{11}w_1^1 + \cdots + a_{1n_1}w_{n_1}^1 + \cdots + a_{r1}w_1^r + \cdots + a_{rn_1}w_{n_1}^r.$$

Observemos que $a_{i1}w_1^i + \cdots + a_{in_i}w_{n_1}^i \in W_i$, por ser V suma directa de tales subespacios, el 0 se escribe de manera única y así se tiene que $a_{i1}w_1^i + \cdots + a_{in_i}w_{n_1}^i = 0$, $\forall i = 1, \cdots, r$. Además como $\{w_1^i, \cdots, w_{n_i}^i\}$ es base de W_i subespacio, se tiene que $a_{i1} = \cdots = a_{in_i} = 0$.
Esto completa la prueba de que B es base de V. ■

Definición 6.9 *Sean $T : V \to V$ un operador lineal sobre un espacio vectorial $V = W_1 \oplus \cdots \oplus W_r$, con W_i subespacios invariantes bajo T ($T(W_i) \subset W_i, i = 1, \cdots, r$). Sea T_i la restricción de T a W_i. Se dice que T se* ***descompone en los operadores** T_i o que T **es suma directa de los** T_i y se escribe*

$$T = T_1 \oplus T_2 \oplus \cdots \oplus T_r.$$

También se dice que los subespacios $W_1, \cdots, W_r$ reducen a T, o que forman una descomposición de V en una suma directa invariante por T.

Observación 6.3 *Consideremos los espacios U, W que reducen un operador $T : V \to V$, con $\{u_1, u_2\}$ y $\{w_1, w_2, w_3\}$ bases ordenadas de U y W respectivamente. Sean además T_1, T_2 las restricciones de T a U y a W respectivamente. Luego*

$$\begin{array}{rcl} T_1(u_1) & = & a_{11}u_1 + a_{12}u_2 \\ T_1(u_2) & = & a_{21}u_1 + a_{22}u_2 \end{array}, \qquad \begin{array}{rcl} T_2(w_1) & = & b_{11}w_1 + b_{12}w_2 + b_{13}w_3 \\ T_2(w_2) & = & b_{21}w_1 + b_{22}w_2 + b_{23}w_3 \\ T_2(w_3) & = & b_{31}w_1 + b_{32}w_2 + b_{33}w_3 \end{array}.$$

Quedan determinadas dos matrices

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}, \qquad B = \begin{bmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{bmatrix}$$

que son las matrices asociadas a las transformaciones T_1 y T_2 relativas a las respectivas bases ordenadas dadas para U y W.
Por el teorema anterior sabemos que $\{u_1, u_2, w_1, w_2, w_3\}$ es una base de V. Además como $T(u_i) = T_1(u_i)$ y $T(w_j) = T_2(w_j)$, la matriz de T relativa a esta base es la matriz diagonal por bloques

$$\begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix}.$$

Teorema 6.5 *Sea V, e.v.s/$\mathbb{K}$, $T : V \in \mathcal{L}(V)$ y V es la suma directa de subespacios invariantes invariantes por T, $W_1, \cdots, W_r$. Si A_i es la representación matricial de la restricción de T a W_i relativa a bases ordenadas dadas de W_i, entonces T tiene asociada la matriz diagonal por bloques*

$$M = \begin{bmatrix} A_1 & 0 & \cdots & 0 \\ 0 & A_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & A_r \end{bmatrix},$$

relativa a la base ordenada que resulta de concatenar en orden las bases de cada uno de los W_i.

Observación 6.4 *La matriz diagonal por bloques M con bloques diagonales $A_1, \cdots, A_r$ se llama a veces la suma directa de las matrices $A_1, \cdots, A_r$ y se representa $M = A_1 \oplus \cdots \oplus A_r$.*

6.4. Descomposición primaria

Teorema 6.6 *Sea $T : V \to V$ lineal, y $f(t) = g(t)h(t)$ polinomios tales que $f(T) = 0$ y $g(t)$ y $h(t)$ son primos relativos (es decir que su máximo común divisor es 1). Entonces V es la suma directa de los subespacios U y W invariantes por T, donde $U = nul(g(T))$ y $W = nul(h(T))$.*

Demos: Como $g(t)$ y $h(t)$ son primos relativos, existen polinomios $r(t)$ y $s(t)$ tales que

$$r(t)g(t) + s(t)h(t) = 1.$$

Luego aplicado al operador T esto nos da

$$r(T)g(T) + s(T)h(T) = \mathbb{I}. \tag{16}$$

Sea $v \in V$, luego por la igualdad anterior se tiene que $v = r(T)g(T)v + s(T)h(T)v$.
Se tiene que $r(T)g(T)v \in W = nul(h(T))$ pues

$$h(T)r(T)g(T)v = r(T)g(T)h(T)v = r(T)f(T)v = r(T)0v = 0.$$

De forma similar se tiene que $s(T)h(T)v \in U = nul(g(T))$. De este modo V es la suma de U y W.

Para ver que $V = U \oplus W$ debemos mostrar que para cada $v = u + w \in V$ la forma de escribirlo es única con $u \in U, w \in W$. Observemos que

$$r(T)g(T)v = r(T)g(T)(u + w) = r(T)g(T)u + r(T)g(T)w = r(T)g(T)w.$$

Aplicando (16) a w se tiene

$$w = r(T)g(T)w + s(T)h(T)w = r(T)g(T)w.$$

De estas últimas dos igualdades surge que $w = r(T)g(T)v$, de este modo vemos que w está determinado por v. De modo análogo se obtiene que $u = s(T)h(T)v$. Luego $V = U \oplus W$, como queríamos ver. ■

Teorema 6.7 *En el teorema 6.6, si $f(t)$ es el polinomio minimal de T y ($g(t)$ y $h(t)$ son mónicos), entonces $g(t)$ y $h(t)$ son los polinomios minimales de $T_1 = T|U$ y $T_2 = T|W$ respectivamente.*

Demos: Sean $m_1(t)$ y $m_2(t)$ los polinomios minimales de T_1 y T_2 respectivamente. Notemos que $g(T_1) = 0$ y $h(T_2) = 0$ pues $U = nul(g(T))$ y $W = nul(h(T))$.
Luego se tiene que $m_1(t)$ divide a $g(t)$ y $m_2(t)$ divide a $h(t)$. Por el teorema 6.6, $f(t)$ es el mínimo común múltiplo de $m_1(t)$ y $m_2(t)$, pero ambos son primos relativos ya que lo son $g(t)$ y $h(t)$.
En consecuencia $f(t) = m_1(t)m_2(t)$ y se tiene también que $f(t) = g(t)h(t)$. Por lo tanto

$$m_1(t)m_2(t) = g(t)h(t) \quad \Rightarrow \quad m_2(t) = \frac{g(t)}{m_1(t)}h(t) = p(t)h(t).$$

con lo cual si $p(t)$ es un polinomio de grado positivo resultaría de h un polinomio que anula a T_2 y de grado estrictamente menor que el $gr(m_2)$ lo que contradice el hecho de que m_2 sea el polinomio minimal de T_2. En consecuencia resulta $g(t) = m_1(t)$ y $h(t) = m_2(t)$. ■

Teorema 6.8 ***Teorema de Descomposición Primaria***
Sea $T : V \to V$ un operador lineal con polinomio minimal

$$m(t) = f_1(t)^{m_1} \cdots f_r(t)^{m_r},$$

donde los $f_i(t)$ son polinomios mónicos irreducibles diferentes. Entonces V es la suma directa de los subespacios invariantes por T, $W_1, \cdots, W_r$, donde W_i es el espacio nulo de $f_i(T)^{m_i}$. Además $f_i(t)^{m_i}$ es el polinomio minimal de la restricción de T a W_i.

Demos: Hacemos inducción sobre r. Para $r = 1$ es trivial.
Supongamos que el resultado es válido para $r - 1$, veamos que vale para r. Por el Teorema 6.6 sabemos que V puede escribirse como la suma directa de subespacios invariantes por T, W_1 y V_1, donde $W_1 = nul(f_1(T)^{n_1})$ y $V_1 = nul(f_2(t)^{n_2} \cdots f_r(t)^{n_r})$.
Por el Teorema 6.7, los polinomios minimales de la restricción de T a W_1 y V_1 son respectivamente

$f_1(T)^{n_1}$ y $f_2(t)^{n_2} \cdots f_r(t)^{n_r}$.
Representemos la restricción de T a V_1 por T_1. Por la hipótesis de inducción, V_1 es la suma directa de subespacios $W_2, \cdots, W_r$ tales que $W_i = nul(f_i(T)^{n_i})$ y tales que $f_i(T)^{n_i}$ es el polinomio minimal de la restricción de T_1 a W_i.
Se tiene que $nul(f_i(T)^{n_i}) \subset V_1$ para $i = 2, \cdots, r$ pues cada $f_i(t)^{n_i}$ divide a $f_2(t)^{n_2} \cdots f_r(t)^{n_r}$, luego $nul(f_i(T)^{n_i})$ es el mismo que $nul(f_i(T_i)^{n_i})$ que es W_i. También la restricción de T a W_i es la misma restricción que de T_1 a W_i, $i = 1, \cdots, r$, con lo cual $f_i(t)^{n_i}$ también es el polinomio minimal de la restricción de T a W_i.
Luego $V = W_1 \oplus W_2 \oplus \cdots \oplus W_r$ resulta la descomposición buscada. ■

Teorema 6.9 *Un operador lineal $T : V \to V$ tiene una representación matricial diagonal por bloques si y sólo si su polinomio minimal $m(t)$ es un producto de polinomios lineales diferentes.*

Demos: $\Leftarrow$) Sea $m(t) = (t-\lambda_1)(t-\lambda_2)\cdots(t-\lambda_r)$, donde los λ_i son escalares diferentes. Por el Teorema 6.8, V es la suma directa de subespacios $W_1, \cdots, W_r$ donde $W_i = nul(T-\lambda_i I)$. Luego si $v \in W_i$ entonces $(T-\lambda_i I)v = 0$, es decir que todo vector de W_i es autovector de T correspondiente a λ_i.
Por el Teorema 6.4, la unión de bases de $W_1 \cdots, W_r$ es una base de V. Esa base está formada por autovectores de T por lo tanto T es diagonalizable.
$\Rightarrow$) Sea T diagonalizable, es decir que V tiene una base formada por autovectores de T. Sean $\lambda_1, \cdots, \lambda_r$ los distintos autovalores de T. Luego el operador

$$f(T) = (T-\lambda_1 I)(T-\lambda_2 I)\cdots(T-\lambda_r I)$$

aplica cada vector de la base en 0. Resulta así $f(T) = 0$ y por lo tanto el polinomio minimal $m(t)$ de T divide al polinomio

$$f(t) = (t-\lambda_1)(t-\lambda_2)\cdots(t-\lambda_r) = m(t)$$

en consecuencia, $m(t)$ es un producto de polinomios lineales diferentes. ■

Observación 6.5 *Una forma equivalente del Teorema 6.9 es la siguiente:*
Una matriz A es semejante a una matriz diagonal si y solo si su polinomio minimal $m(t)$ es un producto de polinomios lineales diferentes.

6.5. Autovectores generalizados

Lema 6.1 *Sean V e.v.s / $\mathbb{K}$ de dimensión finita y $T \in \mathcal{L}(V)$. Luego se tiene que*

$$\{0\} = nul(T^0) \subseteq nul(T^1) \subseteq \cdots \subseteq nul(T^k) \subseteq nul(T^{k+1}) \subseteq \cdots.$$

Demos: Sea $k \in \mathbb{N}$ y sea $v \in nul(T^k)$, i.e. $T^k v = 0$. Luego $T^{k+1}v = T(T^k v) = T(0) = 0$ y se tiene que $v \in nul(T^{k+1})$. ■

Lema 6.2 *Sean V e.v.s / $\mathbb{K}$ de dimensión finita, $T \in \mathcal{L}(V)$ y $m \in \mathbb{N}_0$ tal que*

$$nul(T^m) = nul(T^{m+1}).$$

Luego $nul(T^{m+k}) = nul(T^{m+k+1})$, $\forall k \in \mathbb{N}_0$.

Demos: Sea $v \in nul(T^{m+k+1})$, luego

$$T^{m+1}(T^k v) = T^{m+k+1} = 0, \quad \Rightarrow \quad T^k v \in nul(T^{m+1}) = nul(T^m).$$

Así $T^{m+k}v = T^m(T^k v) = 0$, con lo cual $v \in nul(T^{m+k})$. En consecuencia probamos que $T^{m+k+1} \subseteq T^{m+k}$ y la otra contención es consecuencia directa del lema anterior. ■

Nos preguntamos entonces cuando existe ese $m \in \mathbb{N}$ tal que $nul(T^m) = nul(T^{m+1})$.

Teorema 6.10 *Sean V e.v.s / $\mathbb{K}$ de dimensión finita, $T \in \mathcal{L}(V)$. Entonces se tiene que*

$$nul(T^{dim(V)}) = nul(T^{dim(V)+1}) = nul(T^{dim(V)+2}) = \cdots$$

Demos: Solo hay que probar la primera igualdad. Supongamos que no es válido. Así usando los dos lemas anteriores se tiene que

$$\{0\} = nul(T^0) \subsetneq nul(T^1) \subsetneq \cdots \subsetneq nul(T^{dim(V)}) \subsetneq nul(T^{dim(V)+1}).$$

En todas las contenciones estrictas anteriores, la dimensión de los espacios aumenta al menos 1. Así resultaría $dim(nulT^{dim(V)+1}) \geqslant dim(V) + 1$, lo que es un absurdo pues un subespacio no puede tener más dimensión que el espacio que lo contiene. ■

Observación 6.6 *Notemos que para $T \in \mathcal{L}(V)$ no siempre vale que $V = nul(T) \oplus img(T)$. (Dar un contraejemplo). Sin embargo se tiene es siguiente resultado*

Teorema 6.11 *Sean V e.v.s / $\mathbb{K}$, $dim(V) = n$ y $T \in \mathcal{L}(V)$, luego*

$$V = nul(T^n) \oplus img(T^n).$$

Demos: Veamos que $nul(T^n) \cap img(T^n) = \{0\}$.
Supongamos $v \in nul(T^n) \cap img(T^n) = \{0\}$. Luego $T^n v = 0$ y existe $u \in V$ tal que $v = T^n u$. Aplicando T^n a ambos miembros de la última ecuación se tiene que $T^n v = T^{2n}u$ y así $T^{2n}u = 0$ que por el teorema anterior implica que $0 = T^n u = v$. Con esto obtenemos que $nul(T^n) + img(T^n)$ es una suma directa.
Además

$$dim(nul(T^n) \oplus img(T^n)) = dim(nul(T^n)) + dim(img(T^n)) = dim(V),$$

esto implica que $V = nul(T^n) \oplus img(T^n)$. ■

Todos estos resultados nos serán útiles para los conceptos que definiremos a continuación. Nuestro objetivo sigue siendo, dado un operador $T \in \mathcal{L}(V)$ descomponer el espacio V en suma directa de adecuados subespacios invariantes por T, es decir poder escribir

$$V = V_1 \oplus V_2 \oplus \cdots \oplus V_r, T(V_k) \subseteq V_k, k = 1, \cdots, r.$$

El mejor de los casos sería que cada uno de esos subespacios fuese de dimensión 1, y esto ya vimos es posible si y sólo si el espacio V tiene una base de autovectores y en ese caso podemos tener la descomposición

$$V = E_{\lambda_1}(T) \oplus E_{\lambda_2}(T) \cdots \oplus E_{\lambda_m}(T),$$

con $\lambda_1, \cdots, \lambda_m$ son autovalores distintos de T. Sabemos que este puede no ser el caso para todo operador T, las sumas de las dimensiones de los autoespacios puede ser menor estricto a la $dim(V)$. Necesitamos más vectores.

Definición 6.10 *Sean V e.v.s / $\mathbb{K}$, $dim(V) = n$, $T \in \mathcal{L}(V)$ y λ un autovalor de T. Un vector $v \in V$ es llamado* ***autovector generalizado de T correspondiente a λ*** *si $v \neq 0$ y*

$$(T - \lambda I)^k v = 0,$$

para algún $k \in \mathbb{N}$.

Teorema 6.12 *Sean V e.v.s / $\mathbb{C}$, $dim(V) = n$ y $T \in \mathcal{L}(V)$. Existe una base de V formada por autovectores generalizados de T.*

Demos: Haremos inducción sobre n.
Si $n = 1$ el resultado es cierto pues todo vector no nulo de V es autovector de T.
Supongamos $n > 1$ y que el resultado es válido para todas las dimensiones de V menores a n. Sea λ un autovalor de T, usando el resultado del Teorema 6.11 al operador $(T - \lambda I)$ resulta

$$V = nul((T - \lambda I)^n) \oplus img((T - \lambda I)^n).$$

Si $nul((T - \lambda I)^n) = V$, luego todo vector no nulo en V es un autovector generalizado de T, así toda base de V estará formada por autovectores generalizados.
Podemos suponer entonces que $nul((T - \lambda I)^n) \neq V$, lo que implica que $img((T - \lambda I)^n) \neq \{0\}$. Por otro lado $nul((T - \lambda I)^n) \neq \{0\}$, pues λ es autovalor de T. Así se tiene que

$$0 < rg((T - \lambda I)^n) < n.$$

Más aún $img((T - \lambda I)^n)$ es invariante bajo T (considerar $p(z) = (z - \lambda)^n \in \mathbb{C}[z]$ y Teo.6.2). Sea $S \in \mathcal{L}(img((T - \lambda I)^n))$ la restricción de T al subespacio $img((T - \lambda I)^n)$. La hipótesis inductiva aplicada al operador S implica que existe una base de autovectores generalizados de S, que obviamente son autovectores generalizados de T. Agregando esta base de $img((T - \lambda I)^n)$ a la base de $nul((T - \lambda I)^n)$ tenemos una base de autovectores generalizados de T. ■

Observación 6.7 *Si cambiamos $\mathbb{K} = \mathbb{C}$ por $\mathbb{K} = \mathbb{R}$ habrá operadores para los cuales haya una base de autovectores generalizados y otros que no.*

Proposición 6.6 *Sean V e.v.s / $\mathbb{C}$, $dim(V) = n$ y $T \in \mathcal{L}(V)$. Luego, cada autovector generalizado corresponde a un único autovalor de T.*

Demos: Supongamos que V es una autovector generalizado de T correspondiente a los autovalores α y λ de T. Sea m el menor natural tal que $(T - \alpha I)^m v = 0$. Luego

$$0 = (T - \lambda I)^n v = ((T - \alpha I) + (\alpha I - \lambda)I)^n v = \sum_{k=0}^{n} b_k (\alpha - \lambda)^{n-k} (T - \alpha I)^k v,$$

con $b_0 = 1$ y el resto no importan. Aplicando el operador $(T - \alpha I)^{m-1}$ a ambos miembros de la ecuación se llega a

$$0 = (\alpha - \lambda)^n (T - \alpha I)^{m-1} v.$$

Como $(T - \alpha I)^{m-1} v \neq 0$, luego la ecuación implica que $\alpha = \lambda$, como queríamos probar. ■

Teorema 6.13 *Sean V e.v.s / $\mathbb{C}$, $dim(V) = n$ y $T \in \mathcal{L}(V)$. Luego toda lista de autovectores generalizados de T correspondientes a autovalores distintos de T son linealmente independientes.*

Demos: Supongamos que el resultado es falso. Luego existe un menor natural m tal que existe un conjunto $\{v_1, \cdots, v_m\}$ linealmente dependiente de autovectores generalizados de T correspondientes a autovalores distintos $\lambda_1, \cdots, \lambda_m$. Luego existen $a_1, \cdots, a_m \in \mathbb{K}$, ninguno de ellos igual a 0 (pues m es mínimo) tales que

$$a_1 v_1 + \cdots + a_m v_m = 0.$$

Apliquemos el operador $(T - \lambda_m I)^n$ a ambos miembros de la igualdad anterior, se obtiene

$$a_1 (T - \lambda_m I)^n v_1 + \cdots + a_m (T - \lambda_m I)^n v_{m-1} = 0. \tag{17}$$

Supongamos $k \in \{1, 2, \cdots, m-1\}$, luego $(T - \lambda_m I)^n v_k \neq 0$, ya que si no sería v_k un autovector generalizado de T correspondiente a dos autovalores distintos λ_k y λ_m. Sin embargo

$$(T - \lambda_k I)^n ((T - \lambda_m I)^n v_k) = (T - \lambda_m I)^n ((T - \lambda_k I)^n v_k) = 0.$$

Esto último muestra que $((T - \lambda_m I)^n v_k)$ es un autovector generalizado de T correspondiente al autovalor λ_k. Así, por la igualdad (17))resutan

$$((T - \lambda_m I)^n v_1), \cdots, ((T - \lambda_m I)^n v_{m-1})$$

autovectores generalizados correspondientes a autovalores distintos linealmente dependientes, lo que contradice la minimalidad de m. Y esto concluye la prueba. ■

6.6. Operadores nilpotentes

Definición 6.11 *Un operador $T : V \to V$ se llama* ***nilpotente*** *si $T^n = 0$ para algún n natural.*

Observación 6.8 *Un operador $T \in \mathcal{L}(V)$ es nilpotente si y solo si todo vector no nulo en V es un autovector generalizado de T correspondiente al autovalor 0.*

Ejemplo 6.3

1. *El operador $T \in \mathcal{L}(\mathbb{K}^4)$ definido por*

$$T(z_1, z_2, z_3, z_4) = (0, 0, z_1, z_2),$$

es nilpotente ya que $T^2 = 0$.

2. *El operador en $\mathbb{K}^3$ cuya matriz asociada respecto de la base ordenada canónica es*

$$\begin{bmatrix} -3 & 9 & 0 \\ -7 & 9 & 6 \\ 4 & 0 & -6 \end{bmatrix},$$

es un operador nilpotente y puede verse elevando al cubo la matriz anterior para obtener la matriz nula.

3. *El operador derivación en el espacio $\mathbb{R}_m[x]$ es nilpotente ya que la derivada $(m+1)$-ésima de un polinomio a lo sumo de grado m es igual a 0.*

Proposición 6.7 *Sean V e.v.s$/\mathbb{K}$, $dim(V) = n$ y $T \in \mathcal{L}(V)$ nilpotente. Entonces $T^n = 0$.*

Demos: . Como T es nilpotente, existe un $k \in \mathbb{N}$ tal que $T^k = 0$. Así $null(T^k) = V$. Luego por (6.1) y (6.10) se tiene que $T^n = V$, y así $T^n = 0$. ■

Definición 6.12 *Se llama* ***índice de nilpotencia de T*** *a $k \in \mathbb{N}$ tal que $T^k = 0$ y $T^{k-1} \neq 0$.*

Definición 6.13 *Análogamente, una matriz $A \in \mathcal{M}_{n\times n}(\mathbb{K})$ se llama* **nilpotente** *si $A^n = 0$ para algún natural n y de* **índice de nilpotencia** k *si $A^k = 0$ y $A^{k-1} \neq 0$.*

Proposición 6.8 *Sean V e.v.s / $\mathbb{K}$, $dim(V) = n$ y $T \in \mathcal{L}(V)$.*

1. *Si T es nilpotente, luego 0 es el único autovalor de T.*
2. *Si $\mathbb{K} = \mathbb{C}$ y 0 es el único autovalor de T, luego T es nilpotente*

Demos: $\cdots$ ■

Teorema 6.14 *Sea $T : V \to V$ un operador nilpotente de índice k, entonces T tiene una representación matricial diagonal por bloques y los bloques diagonales son de la forma*

$$N = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & & & & \end{bmatrix}$$

Existe al menos un bloque diagonal N de orden k y los restantes son de orden $\leqslant k$. El número de bloques diagonales para cada orden posible está determinado únicamente por T. Además, el número total de bloques diagonales de todos los órdenes es igual a la $\dim(nul(T))$.

Observación 6.9 *Las matrices como la matriz N del teorema anterior son nilpotentes y su índice de nilpotencia es igual a su orden. Además notemos que la matriz N de orden 1 es la matriz 0.*

6.7. Autoespacios generalizados

Definición 6.14 *Sean V e.v.s / $\mathbb{K}$, $dim(V) = n$, $T \in \mathcal{L}(V)$ y $\lambda \in \mathbb{K}$. El* **autoespacio generalizado de T correspondiente a λ** *se nota como $G(,_\lambda, T) = G_\lambda(T)$ y es el subespacio de V dado por*

$$G_\lambda(T) = \{v \in V : (T - \lambda I)^k = 0, \text{ para algún } k \in \mathbb{N}\}.$$

Así $G_\lambda(T)$ es el conjunto de los autovalores generalizados de T correspondientes al autovalor λ unión $\{0\}$.

Notemos que siendo todo autovector de T correspondiente al autovalor λ un autovector generalizado de T correspondiente al autovalor λ resulta

$$E_\lambda(T) \underset{s.e.}{\subseteq} G_\lambda(T)$$

Proposición 6.9 *Sean V e.v.s / $\mathbb{K}$, $dim(V) = n$, $T \in \mathcal{L}(V)$ y $\lambda \in \mathbb{K}$. Entonces $G_\lambda(T) = null(T - \lambda I)^n$.*

Demos: $\cdots$ ■

Teorema 6.15 *Sea $\mathbb{K} = \mathbb{C}$. Sean V e.v.s / $\mathbb{C}$, $dim(V) = n$, $T \in \mathcal{L}(V)$. Sean $\lambda_1, \lambda_2, \cdots, \lambda_m$ los autovalores distintos de T. Luego*

1. *$G_{\lambda_k}(T)$ es invariante bajo T para cada $k = 1, \cdots, m$;*
2. *$(T - \lambda_k I)|_{G_{\lambda_k}(T)}$ es nilpotente para cada $k = 1, \cdots, m$;*
3. *$V = G_{\lambda_1}(T) \oplus \cdots \oplus G_{\lambda_m}(T)$.*

Demos: $\cdots$ ■

Este teorema nos está diciendo que cuando el campo de escalares es $\mathbb{C}$ y consideramos un operador lineal $T \in \mathcal{L}(V)$ podemos descomponer al espacio vectorial V en suma de subespacios invariantes, dados por los autoespacios generalizados correspondientes a cada uno de los autovalores distintos.
Si consideramos el subespacio $G_{\lambda_k}(T)$ observemos que podemos escribir $T|_{G_{\lambda_k}(T)} = (T-\lambda_k I)|_{G_{\lambda_k}(T)} + \lambda_k I|_{G_{\lambda_k}(T)}$ es decir como la suma de un operador diagonal y un operador nilpotente, es esto lo que aprovecharemos para elegir una base que nos sea conveniente.

6.8. FORMA CANÓNICA DE JORDAN

Nota 6.1 *Un operador T puede expresarse en la forma canónica de Jordan si sus polinomios minimales y característico se factorizan en polinomios lineales. Esto siempre es posible si $\mathbb{K} = \mathbb{C}$. En cualquier caso podemos extender el cuerpo $\mathbb{K}$ a uno en el cual los polinomios minimales y característicos puedan factorizarse en factores lineales, entonces en un sentido amplio cualquier operador tiene una forma canónica de Jordan. Análogamente, toda matriz es semejante a una matriz en forma canónica de Jordan.*

Teorema 6.16 *Sea $T : V \to V$ un operador lineal cuyos polinomios minimal y característico son respectivamente*

$$p(t) = \det(T - tI) = (t - \lambda_1)^{n_1} \cdots (t - \lambda_r)^{n_r}$$

$$m(t) = (t - \lambda_1)^{m_1} \cdots (t - \lambda_r)^{m_r},$$

donde los λ_i son escalares distintos. Entonces T tiene una representación matricial diagonal por bloques J cuyos elementos diagonales son de la forma

$$J_{ij} = \begin{bmatrix} \lambda_i & 1 & 0 & \cdots & 0 & 0 \\ 0 & \lambda_i & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & \lambda_i & 1 \\ 0 & 0 & 0 & \cdots & 0 & \lambda_i \end{bmatrix}.$$

Para cada λ_i los bloques correspondientes J_{ij} tienen las siguientes propiedades:
i) Existe al menos un J_{ij} de orden m_i, los demás J_{ij} son de orden $\leqslant m_i$.
ii) La suma de los órdenes de los J_{ij} es n_i.
iii) La cantidad de J_{ij} es igual a la multiplicidad geométrica de λ_i (es decir la dimensión de su autoespacio).
iv) La cantidad de J_{ij} de cada orden posible está determinado únicamente por T.
*A la matriz J se la llama **forma canónica de Jordan**.*

Demos: Por el Teorema 6.8, T se puede descomponer en operadores $T_1, \cdots, T_r$, esto es $T = T_1 \oplus \cdots \oplus T_r$, donde $(t - \lambda_i)^{m_i}$ es el polinomio minimal de T_i. Luego en particular

$$(T_1 - \lambda_1 I)^{m_1} = 0, \cdots, (T_r - \lambda_r I)^{m_r} = 0.$$

Sea $N_i = T_i - \lambda_i I$, entonces para cada $i = 1, \cdots, r$ $T_i = N_i + \lambda_i I$, con $N_i^{m_i} = 0$. Esto es, T_i es la suma del operador $\lambda_i I$ y un operador nilpotente N_i, el cual tiene índice m_i ya que $(t - \lambda_i)^{m_i}$ es el polinomio minimal de T_i.
Ahora, podemos elegir una base tal que N_i esté en su forma canónica. En esta base $T_i = N_i + \lambda_i I$ se representa por una matriz diagonal por bloques M_i, cuyos elementos de la diagonal son las matrices J_{ij}. La suma directa de las matrices M_i es una matriz J que es una forma canónica de Jordan y por el Teorema 6.5 es una representación matricial de T.
Por último, veamos que los bloques J_{ij} satisfacen las propiedades requeridas:
i) Se obtiene por ser N_i de índice m_i.
ii) Vale porque T y J tiene el mismo polinomio característico.
iii) Vale pues la $\dim(nul(N_i)) = \dim(nul(T_i - \lambda_i I))$ es igual a la dimensión del autoespacio correspondiente a λ_i.
iv) Vale por estar los T_i (y los N_i) determinados únicamente por T. ■

Observación 6.10 $J_{ij} = \lambda_i I + N_i$

Ejemplo 6.4 *Hallar la forma canónica de Jordan de la siguiente matriz*

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

Primero calculamos el polinomio característico de B y así sus autovalores.

$$p_B(\lambda) = \det(B - \lambda I) = (1-\lambda)^3(2-\lambda)^3.$$

Vemos que los autovalores son $\lambda = 2$ y $\lambda = 1$ ambos con multiplicidad algebraica 3. Por lo tanto existirán dos bloques de Jordan

$$J = \begin{bmatrix} J(1) & 0 \\ 0 & J(2) \end{bmatrix}.$$

Calculamos los rangos $rg_j((A-I)^j) = rg_j(1)$ y $rg_i((A-2I)^i) = rg_i(2)$ hasta que $rg_k(\star) = rg_{k+1}(\star)$, así tenemos:

$$\begin{array}{lcl} rg(B-I) = 4, & \Leftrightarrow & dim(nul(B-I)) = 2, \\ rg(B-I)^2 = 3 & \Leftrightarrow & dim(nul(B-I)^2) = 3, \\ rg(B-I)^3 = 3 & \Leftrightarrow & dim(nul(B-I)^3) = 3, \end{array} \qquad \begin{array}{lcl} rg(B-2I) = 5, & \Leftrightarrow & dim(nul(B-I)) = 1, \\ rg(B-2I)^2 = 4, & \Leftrightarrow & dim(nul(B-I)^2) = 2, \\ rg(B-2I)^3 = 3, & \Leftrightarrow & dim(nul(B-I)^3) = 3, \\ rg(B-2I)^4 = 3, & \Leftrightarrow & dim(nul(B-I)^4) = 3. \end{array}$$

Observamos que como $dim(nul(B-I)) = 2$ la dimensión del autoespacio correspondiente al autovalor $\lambda = 1$ es 2, en consecuencia habrá dos bloques de Jordan para este autovalor: $J_11(1)$ y $J_12(1)$.
Como $dim(nul(B-2I)) = 1$ la dimensión del autoespacio correspondiente al autovalor $\lambda = 2$ es 1, en consecuencia habrá un bloques de Jordan para este autovalor: $J_21(2)$.
Además el índice correspondiente al autovalor $\lambda = 1$ es $m_1 = 2$ y para $\lambda = 2$ se tiene $m_2 = 3$. Luego el polinomio minimal de B está dado por $m_B(\lambda) = (1-\lambda)^2(2-\lambda)^3$.
Esto nos dice que el bloque de Jordan más grande de $J(1)$ es 2×2, mientras que el bloque de Jordan de $J(2)$ es 3×3.

Más aun, como $p_B(\lambda) = (\lambda-1)^3(\lambda-2)^3$, sabemos que la suma de los órdenes de $J_{1j}(1)$ es 3 y la de $J_{2j}(2)$ también es 3. Así resulta que la forma canónica de Jordan de B es

$$J = \left[\begin{array}{cc|c|ccc} 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ \cline{1-3} 0 & 0 & 1 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 2 & 1 & 0 \\ 0 & 0 & 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 0 & 0 & 2 \end{array}\right].$$