1. Formas canónicas de Jordan

1.1. Invariancia

Definición 1 *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,*

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

Ejemplo 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 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 1 *Sean $T : V \to V$ y $p(t)$ un polinomio un polinomio cualquiera. Entonces $nul(p(T))$ es invariante por T.*

Demos: Sea $v \in nul(p(T))$, es decir $p(T)v = 0$. Debemos probar que $Tv \in nul(p(T))$, es decir si $p(T)(Tv) = 0$.
Observemos que como se verifica que $p(t)t = tp(t)$ para cualquier polinomio, se tiene que $p(T)T = Tp(T)$. Luego

$$p(T)(Tv) = (p(T)T)v = (Tp(T))v = T(p(T)v) = T(0) = 0.$$

como queríamos ver. ■

Teorema 2 *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. ■

1.2. DESCOMPOSICIÓN EN SUMA DIRECTA DE INVARIANTES

Recordemos que

Definición 2 *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 $W_i \cap W_j = \{0\}$. Se nota $V = W_1 \oplus \cdots \oplus W_r$.

Teorema 3 *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 3 *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 1 *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 4 *Supongamos que $T : V \to T$ es lineal 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 2 *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 $M = A_1 \oplus \cdots \oplus A_r$.*

1.3. Descomposición primaria

Teorema 5 *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. 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

$$r(T)g(T) + s(T)h(T) = \mathbb{I}. \tag{1}$$

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 (1) 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. ■

Observación 3 *Sea A una matriz $n \times n$ sobre $\mathbb{K}$. Observemos que existen polinomios no nulos $f(t)$ tales que $f(A) = 0$, por ejemplo el polinomio característico asociado a A. Entre todos estos polinomios consideramos los de grado mínimo y con coeficiente principal igual a 1. Este polinomio $m(t)$ existe, es único y se llama **polinomio minimal de A** . Puede probarse que dicho polinomio divide a cualquier polinomio que aplicado a la matriz A de la matriz nula, en particular divide al polinomio característico de A. También puede probarse que el polinomio minimal y el polinomio característico de una matriz A tienen los mismos factores irreducibles.*

Teorema 6 *En el teorema 5, 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 las restricciones de T a U y W respectivamente.*

Teorema 7 ***Teorema de Descomposición Primaria***
Sea $T : V \to V$ un operador lineal con polinomio minimal

$$m(t) = f_1(t)^{n_1} \cdots f_r(t)^{n_r},$$

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 nula de $f_i(T)^{n_i}$. Además $f_i(t)^{n_i}$ 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 5 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, 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 8 *Un operador lineal $T : V \to V$ tiene una representación matricial diagonal si y sólo si su polinomio minimal $m(t)$ es un producto de polinomios lineales diferentes.*

Demos: $\Leftarrow$) Sea $m(t) = (t - \lambda_1)(t - \lambda_2) \cdots (t - \lambda_n)$, donde los λ_i son escalares diferentes. Por el Teorema 7, 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 3, 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

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

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

Observación 4 *Una forma equivalente del Teorema 8 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.

1.4. Operadores nilpotentes

Definición 4 *Un operador $T : V \to V$ se llama **nilpotente** si $T^n = 0$ para algún n natural.*

Definición 5 *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 *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 $A^{k-1} \neq 0$.*

Observación 5 *Es claro que el polinomio minimal de un operador (matriz) nilpotente de índice k es $m(t) = t^k$, luego 0 es su único valor propio.*

Teorema 9 *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*

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

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

Demos: Supongamos que $\dim(V) = n$. Sean $W_1 = nul(T), W_2 = nul(T^2), \cdots, W_k = nul(T^k)$. Sean $m_i = \dim(W_i)$, $i = 1, \cdots, k$. Como T tiene índice de nilpotencia k, $W_k = V$ y $W_{k-1} \neq V$ por lo tanto $m_{k-1} < m_k = n$. Entonces $W_1 \subset W_2 \subset \cdots \subset W_k = V$.
Luego por inducción podemos elegir una base $\{u_1, \cdots, u_n\}$ de V tal que $\{u_1, \cdots, u_{m_i}\}$ sea base de W_i.
Elegimos ahora una nueva base de V respecto de la cual T tenga la forma deseada. Es conveniente señalar los miembros de esta nueva base por parejas de índices

$$v(1,k) = u_{m_{k-1}+1}, v(2,k) = u_{m_{k-1}+2}, \cdots, v(m_k - m_{k-1}, k) = u_{m_k}$$

$$\begin{array}{rl} m_k - m_{k-1} & \text{componentes diagonales de orden } k \\ (m_{k-1} - m_{k-2}) - (m_k - m_{k-1}) = 2m_{k-1} - m_k - m_{k-2} & \text{componentes diagonales de orden } k-1 \\ \vdots & \\ 2m_2 - m_1 - m_3 & \text{componentes diagonales de orden } 2 \\ 2m_1 - m_2 & \text{componentes diagonales de orden } 1. \end{array}$$

y tomando

$$v(1,k-1) = Tv(1,k), v(2,k-1) = Tv(2,k), \cdots, v(m_k - m_{k-1}, k-1) = Tv(m_k - m_{k-1}, k)$$

podemos mostrar que

$$S_1 = \{u_1, \cdots, u_{m_{k-2}}, v(1,k-1), \cdots, v(m_k - m_{k-1}, k-1)\}$$

es un subconjunto *l.i.* de W_{k-1}. Extendemos S_1 a una base de W_{k-1}, agregando nuevos elementos (si es necesario) los cuales representamos por

$$v(m_k - m_{k-1} + 1, k-1), v(m_k - m_{k-1} + 2, k-1), \cdots, v(m_{k-1} - m_{k-2}, k-1)$$

Luego establecemos

$$v(1,k-2) = Tv(1,k-1), v(2,k-2) = Tv(2,k-1), \cdots, v(m_{k-1} - m_{k-2}, k-2) = Tv(m_{k-1} - m_{k-2}, k-1).$$

Nuevamente se tiene que

$$S_2 = \{u_1, \cdots, u_{m_{k-3}}, v(1,k-2), \cdots, v(m_{k-1} - m_{k-2}, k-2)\}$$

es un subconjunto *l.i.* de W_{k-2}, el cual puede extenderse a una base de dicho subespacio agregando elementos

$$v(m_{k-1} - m_{k-2} + 1, k-2), v(m_{k-1} - m_{k-2} + 2, k-2), \cdots, v(m_{k-2} - m_{k-3}, k-2).$$

Prosiguiendo de este modo se consigue una nueva base de V, que podemos ordenar por comodidad de referencia como

$$\begin{array}{l} v(1,k), \cdots, v(m_k - m_{k-1}, k) \\ v(1,k-1), \cdots, v(m_k - m_{k-1}, k-1) \cdots, v(m_{k-1} - m_{k-2}, k-1) \\ \vdots \\ v(1,2), \cdots, v(m_k - m_{k-1}, 2), \cdots, , v(m_{k-1} - m_{k-2}, 2), v(m_2 - m_1, 2) \\ v(1,1), \cdots, v(m_k - m_{k-1}, 1), \cdots, , v(m_{k-1} - m_{k-2}, 1), \cdots, v(m_2 - m_1, 1), \cdots, v(m_1, 1) \end{array}.$$

La última fila forma una base de W_1, las últimas 2 filas forman una base de W_2, etc. Además notemos que T aplica a cada vector en el vector inmediatamente inferior en la tabla o en 0 si es un vector de la última fila. Es decir

$$Tv(i,j) = \left| \begin{array}{ll} v(i,j-1) & j > 1 \\ 0 & j = 1 \end{array} \right. .$$

Ahora, T tendrá la forma deseada si los $v(i,j)$ están ordenados lexicográficamente, empezando con $v(1,1)$ y avanzando sobre la primera columna hasta el $v(1,k)$, luego saltando a $v(2,1)$ y avanzando por la segunda columna tanto como sea posible, etc.
Además, hay exactamente como se ve directamente de la tabla. En particular, como los números $m_1, \cdots, m_k$ están determinados únicamente por T, el número de componentes diagonales en cada orden está únicamente determinado por T.
Finalmente la identidad

$$m_1 = (m_k - m_{k-1}) + (2m_{k-1} - m_k - m_{k-2} + \cdots + (2m_2 - m_1 - m_3) + (2m_1 - m_2),$$

muestra que la $\dim(nul(T))$ es la cantidad total de componentes de la diagonal de T. ■

Observación 6 *La matriz N es nilpotente y su índice de nilpotencia es igual a su orden. Además notemos que la matriz N de orden 1 es la matriz 0.*

1.5. Forma canónica de Jordan

Nota 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 10 *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 7, 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 es 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 este 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 4 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 7 $J_{ij} = \lambda_i I + N_i$

Ejemplo 3 *Hallar la forma canónica de Jordan de la siguiente matriz*

$$A = \begin{bmatrix} 5 & 4 & 0 & 0 & 4 & 3 \\ 2 & 3 & 1 & 0 & 5 & 1 \\ 0 & -1 & 2 & 0 & 2 & 0 \\ -8 & -8 & -1 & 2 & -12 & -7 \\ 0 & 0 & 0 & 0 & -1 & 0 \\ -8 & -8 & -1 & 0 & -9 & -5 \end{bmatrix}.$$

Primero calculamos los autovalores, obteniéndose $\lambda_1 = 2$ *y* $\lambda_2 = -1$*, por lo tanto existirán dos bloques de Jordan*

$$J = \begin{bmatrix} J_1(2) & 0 \\ 0 & J_2(-1) \end{bmatrix}.$$

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

$$\begin{array}{ll} rg_1(2) = rg(A-2I) = 4, & rg_1(-1) = rg(A+I) = 4, \\ rg_2(2) = rg(A-2I)^2 = 3, & rg_2(-1) = rg(A+I)^2 = 4. \\ rg_3(2) = rg(A-2I)^3 = 2, & \\ rg_4(2) = rg(A-2I)^4 = 2, & \end{array}$$

Entonces $m_1 = 3$*, el índice de* $\lambda_1 = 2$ *y* $m_2 = 1$*.*
Esto nos dice que el bloque de Jordan más grande de $J_1(2)$ *es* 3×3*, mientras que el bloque de Jordan más grande de* $J_2(-1)$ *es* 1×1*. Esto significa que* $J(-1)$ *es una matriz diagonal. Más aun, como* $\det(A - \lambda I) = (t-2)^4(t+1)^2$*, sabemos que la suma de los órdenes de* $J_{1j}(2)$ *es* 4 *y la de* $J_{2j}(-1)$ *es* 2*. Como también sabemos que existe al menos un* $J_{1j}(2)$ *de orden* $m_1 = 3$ *tenemos que la forma canónica de Jordan de* A *debe ser*

$$J = \left[\begin{array}{ccc|c|cc} 2 & 1 & 0 & 0 & 0 & 0 \\ 0 & 2 & 1 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 2 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & -1 & 0 \\ \cline{5-6} 0 & 0 & 0 & 0 & 0 & -1 \end{array}\right].$$

Referencias

[1] Axler S., *Linear Algebra done right*, (3era edición), Springer, 2015.

[2] Escalante M.S., *Notas de clases*, 2013.

[3] Friedberg S.H., Insel A.J., Spence L.E., *Linear Algebra*, Prentice Hall, New York,1989.

[4] Hoffman K., Kunze R., *Linear algebra*, Prentice Hall, México 1971.

[5] Lankham I., Nachtergaele B., Schilling A., *Linear algebra as an introduction to abstract mathematics*, World Scientific, 2016.

[6] Lay D.C., *Álgebra lineal y sus aplicaciones*, Pearson Educación, México 2007.

[7] Santillán Marcus E.A., *Notas de clases*, 2012.

[8] Strang G., *Linear algebra and its applications, 3rd ed*, Harcourt College Publishers. 1988.

[9] Strang G., *Introduction to linear algebra, 4th ed.*, Wellesley. Cambridge Press, 2009.