LA FORMA DE JORDAN

11 de diciembre de 2006

1. Ideas preliminares

Definición *Sea A $n \times n$ y S un subespacio de $\mathbb{R}^n$ (ó $\mathbb{C}^n$). Decimos que S es invariante por A si,*

$$A\,u \in S, \qquad \forall u \in S.$$

Ejemplo. Si λ es un valor propio de A, $S(\lambda) := \{u \in \mathbb{R}^n \,|\, A\,u = \lambda\,u\}$ es invariante por A.

Definición *Sea A una matriz $n \times n$. Si $\mathbb{R}^n = S \oplus T$, con S y T invariantes por A, decimos que $\mathbb{R}^n = S \oplus T$ es una descomposición invariante por A. La misma definición se extiende a*

$$\mathbb{R}^n = S_1 \oplus \ldots \oplus S_k$$

con S_j invariante por A para todo j.

Ejemplo. Si A es diagonalizable y $\lambda_1, \ldots, \lambda_k$ son sus valores propios, entonces

$$\mathbb{R}^n = S(\lambda_1) \oplus \ldots \oplus S(\lambda_k)$$

es una descomposición invariante por A.

Aplicación. Si $\mathbb{R}^n = S \oplus T$ es una descomposición invariante por A, consideramos una base de $\mathbb{R}^n$ construida de la forma

$$\underbrace{s_1, \ldots, s_\ell}_{\text{base de } S}, \underbrace{t_{\ell+1}, \ldots, t_n}_{\text{base de } T}$$

y

$$P = \left[\begin{array}{c|c|c|c|c|c} s_1 & \ldots & s_\ell & t_{\ell+1} & \ldots & t_n \end{array}\right]$$

entonces

$$P^{-1}\,A\,P = \begin{bmatrix} A_{\ell\times\ell} & 0 \\ 0 & A_{n-\ell\times n-\ell} \end{bmatrix}$$

Las descomposiciones invariantes producen matrices **diagonales por bloques**. Nótese que si

$$P^{-1}\,A\,P = \begin{bmatrix} A_1 & & & \\ & A_2 & & \\ & & \ddots & \\ & & & A_k \end{bmatrix}$$

con A_j cuadrada para todo j, entonces: (a) la partición correspondiente de columnas de P genera una descomposición invariante en k subespacios; (b) el polinomio característico de A es el producto de los polinomios característicos de las matrices A_j.

Objetivo. En un primer paso, el objetivo de llegar a una forma canónica de Jordan es lograr una transformación P tal que

$$P^{-1}\,A\,P = \begin{bmatrix} A_1 & & & \\ & A_2 & & \\ & & \ddots & \\ & & & A_k \end{bmatrix}$$

donde el polinomio característico de cada A_j contiene una única raíz, distinta de la de los demás bloques. Por tanto, si el polinomio característico de A es

$$(x-\lambda_1)^{m(\lambda_1)}\,(x-\lambda_2)^{m(\lambda_2)}\,\ldots\,(x-\lambda_k)^{m(\lambda_k)}$$

el bloque A_j tendrá como polinomio característico

$$(x-\lambda_j)^{m(\lambda_j)}.$$

Así, estamos realizando una descomposición invariante por A

$$\mathbb{C}^n = E_1 \oplus E_2 \oplus \ldots \oplus E_k$$

donde $\dim E_j = m(\lambda_j)$.

2. Núcleos iterados

Definición *Sea A una matriz $n\times n$ y λ un valor propio de A. Consideramos los subespacios*

$$E_j(\lambda) := N(\,(A-\lambda\,I)^j\,), \qquad j\geq 1,$$

a los que llamamos núcleos iterados. Denotamos $E_0(\lambda) = N(\,I\,) = 0$.

Propiedades elementales.

(1) $E_1(\lambda) = S(\lambda)$.

(2) $E_j(\lambda) \subseteq E_{j+1}(\lambda)$ para todo j.

Dem. Es consecuencia inmediata de que

$$(A - \lambda I)^j u = 0 \qquad \Longrightarrow \qquad (A - \lambda I)^{j+1} u = 0.$$

$\square$

(3) Si $E_\ell(\lambda) = E_{\ell+1}(\lambda)$, entonces $E_{\ell+1}(\lambda) = E_{\ell+2}(\lambda)$. En consecuencia

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

Dem. Se deduce de la siguiente cadena de implicaciones

$$\begin{aligned} (A - \lambda I)^{\ell+2} u = 0 \qquad &\Longrightarrow \qquad v := (A - \lambda I)u \in E_{\ell+1}(\lambda) = E_\ell(\lambda) \\ &\Longrightarrow \qquad (A - \lambda I)^{\ell+1} u = (A - \lambda I)^\ell v = 0, \end{aligned}$$

luego $E_{\ell+2}(\lambda) \subseteq E_{\ell+1}(\lambda) \subseteq E_{\ell+2}(\lambda)$. $\square$

(4) Los subespacios $E_j(\lambda)$ son invariantes por A.

Dem. Notemos que si $u \in E_j(\lambda)$, entonces

$$\begin{aligned} (A - \lambda I)^j u = 0 \qquad &\Longrightarrow \qquad 0 = A\,(A - \lambda I)^j u \overset{(*)}{=} (A - \lambda I)^j A\, u \\ &\Longrightarrow \qquad Au \in E_j(\lambda). \end{aligned}$$

Para ver (*) basta notar que

$$(A - \lambda I)^j = \sum_{i=0}^{j} (-\lambda)^{j-i} \binom{j}{i} A^i,$$

luego $A\,(A - \lambda I)^j = (A - \lambda I)^j\, A$. $\square$

Propiedades no tan elementales.

(5) Para todo $j \geq 1$

$$\dim E_{j+1}(\lambda) - \dim E_j(\lambda) \leq \dim E_j(\lambda) - \dim E_{j-1}(\lambda).$$

(6) Sean $\lambda_1, \dots, \lambda_k$ los valores propios de A. Sean $j_1, \dots, j_k \geq 1$. Entonces la siguiente suma de subespacios es directa:

$$E_{j_1}(\lambda_1) \oplus E_{j_2}(\lambda_2) \oplus \dots \oplus E_{j_k}(\lambda_k).$$

(7) La dimensión máxima de los núcleos iterados es la multiplicidad del valor propio, esto es, para cada λ, existe ℓ tal que

$$E_\ell(\lambda) = E_{\ell+1}(\lambda), \qquad \dim E_\ell(\lambda) = m(\lambda).$$

Sobre las dimensiones de los núcleos iterados. Sean

$$n_j := \dim E_j(\lambda) = \dim N(\,(A - \lambda\, I)^j\,), \qquad j \geq 0$$

(nótese que $n_0 = 0$). Entonces, las propiedades anteriores implican que esta sucesión de enteros cumple las tres siguientes propiedades:

$$\begin{array}{ll} (1) & 0 = n_0 < n_1 < \ldots < n_\ell = n_{\ell+1} = \ldots, \\ (2) & n_\ell - n_{\ell-1} \leq n_{\ell-1} - n_{\ell-2} \leq \ldots \leq n_2 - n_1 \leq n_1 - n_0 = n_1 \\ (3) & n_\ell = m(\lambda). \end{array}$$

La sucesión de núcleos iterados queda entonces:

$$\begin{array}{rccccccccccccc} & E_0(\lambda) & \subset & E_1(\lambda) & \subset & E_2(\lambda) & \subset & \ldots & \subset & E_\ell(\lambda) & = & E_{\ell+1}(\lambda) & = & \ldots \\ \dim: & 0 = n_0 & < & n_1 & < & n_2 & < & \ldots & < & n_\ell & = & n_{\ell+1} & = & \ldots \end{array}$$

Una observación. Si $B := P^{-1}AP$, entonces

$$u \in N(\,(B - \lambda I)^j\,) \qquad \Longleftrightarrow \qquad P\,u \in N(\,(A - \lambda I)^j\,).$$

Por tanto, las dimensiones de los núcleos iterados son invariantes por semajanza.

Corolario (Teorema de descomposición de Jordan). Si el polinomio característico de A es

$$(x - \lambda_1)^{m(\lambda_1)}\,(x - \lambda_2)^{m(\lambda_2)}\,\ldots\,(x - \lambda_k)^{m(\lambda_k)}$$

entonces existen ℓ_j de forma que

$$\mathbb{C}^n = E_{\ell_1}(\lambda_1) \oplus \ldots \oplus E_{\ell_k}(\lambda_k)$$

y un cambio de base de forma que

$$P^{-1}\,A\,P = \begin{bmatrix} A_1 & & \\ & \ddots & \\ & & A_k \end{bmatrix}$$

donde para cada j el polinomio característico de A_j es $(x - \lambda_j)^{m(\lambda_j)}$.

Además, las dimensiones de los núcleos iterados de las submatrices A_i coinciden con las de los de A.

Dem. La primera afirmación es consecuencia directa de las propiedades (6) y (7). Tomando conjuntamente una base de todos los $E_{\ell_j}(\lambda)$ y poniéndola por columnas en P se llega a la matriz diagonal por bloques.

Veamos que las dimensiones de los núcleos iterados de A_1 coinciden con las dimensiones de $E_j(\lambda_1)$. Sea $B := P^{-1}\,A\,P$. Como

$$u \in N(\,(B - \lambda_1 I)^j\,) \qquad \Longleftrightarrow \qquad P\,u \in E_j(\lambda_1) \subseteq E_{\ell_1}(\lambda_1) = \mathbb{C}\langle p_1, p_2, \ldots, p_{m(\lambda_1)}\rangle$$

(p_j son las columnas de P), entonces

$$N(\,(B-\lambda_1 I)^j\,) \subseteq \mathbb{C}\langle e_1, e_2, \ldots, e_{m(\lambda_1)}\rangle$$

(e_j son los vectores canónicos). Esto quiere decir que los elementos de $N(\,(B-\lambda_1 I)^j\,)$ sólo pueden tener las primeras $m(\lambda_1)$ componentes no nulas. A partir de allí, es obvio que

$$u \in N(\,(B-\lambda_1 I)^j\,) \qquad \Longleftrightarrow \qquad u = \begin{bmatrix} v \\ 0 \end{bmatrix}, \quad v \in N(\,(A_1-\lambda_1 I)^j\,).$$

Lo mismo se puede hacer con los demás valores propios. $\square$

Conclusión. A partir de ahora podemos concentrarnos en matrices con un único valor propio, ya que el teorema precedente parte el trabajo en submatrices con un único valor propio.

3. Demostraciones

Propiedad (5). Para todo $j \geq 1$

$$\dim E_{j+1}(\lambda) - \dim E_j(\lambda) \leq \dim E_j(\lambda) - \dim E_{j-1}(\lambda).$$

Dem. Sea $m := \dim E_{j+1}(\lambda) - \dim E_j(\lambda)$ y tomemos $v_1, \ldots, v_m$ una base de un suplementario de $E_j(\lambda)$ en $E_{j+1}(\lambda)$, esto es, vectores independientes tales que

$$E_{j+1}(\lambda) = E_j(\lambda) \oplus \mathbb{C}\langle v_1, \ldots, v_m\rangle.$$

Nótese que esto quiere decir que

$$\begin{array}{l} v_1, \ldots, v_m \in E_{j+1}(\lambda), \\ \xi_1 v_1 + \ldots + \xi_m v_m \in E_j(\lambda) \end{array} \qquad \Longleftrightarrow \qquad \xi_1 = \ldots = \xi_m = 0.$$

Definimos $w_i := (A - \lambda\, I)v_i$. Entonces es fácil comprobar que

$$\begin{array}{l} w_1, \ldots, w_m \in E_j(\lambda), \\ \xi_1 w_1 + \ldots + \xi_m w_m \in E_{j-1}(\lambda) \end{array} \qquad \Longleftrightarrow \qquad \xi_1 = \ldots = \xi_m = 0,$$

luego

$$\begin{array}{l} \dim \mathbb{C}\langle w_1, \ldots, w_m\rangle = m, \\ \mathbb{C}\langle w_1, \ldots, w_m\rangle \subset E_j(\lambda), \\ \mathbb{C}\langle w_1, \ldots, w_m\rangle \cap E_{j-1}(\lambda) = 0, \end{array}$$

esto es, el resultado buscado. $\square$

Lema. Si $p(A)u = 0 = q(A)u$ y $r(x) := \text{m.c.d.}\{p(x), q(x)\}$, entonces $r(A)\,u = 0$.

Dem. Sea $m(x)$ el polinómico mónico de menor grado tal que $m(A)\,u = 0$. Por hipótesis, m existe (ver algunos comentarios después). Si m no fuera único, entonces tendríamos m_1 y m_2 mónicos del mismo grado tales que $m_1(A)\,u = 0 = m_2(A)\,u$, luego $m_3 := m_1 - m_2$ tiene grado estrictamente menor y $m_3(A)\,u = 0$, lo que contradice la elección de m.

Si $p(A)\,u = 0$, entonces p es múltiplo de m, ya que en caso contrario el resto de dividir p por m (no nulo y de grado estrictamente menor que m) contradiría la elección de m.

Por esta razón, como $p(A)u = 0 = q(A)u$ y $\text{m.c.d.}(p(x), q(x)) = 1$, necesariamente $m(x) \equiv 1$, luego $u = 0$. $\square$

Comentario. El resultado anterior se puede ver aplicando un sencillo resultado de álgebra de polinomios en una variable. Consideramos el conjunto

$$\mathcal{P}_u := \{p(x) \,|\, p(A)u = 0\},$$

que es un ideal en $\mathbb{P}$. Como $\mathbb{P}$ es un dominio de ideales principales, $\mathcal{P}_u$ tiene un único generador mónico (el $m(x)$ de la demostración). De hecho, es la demostración de que $\mathbb{P}$ es dominio de ideales principales no difiere mucho de la realizada.

Que el conjunto $\mathcal{P}_u$ no es vacío se sigue, tal y como estamos trabajando, de las hipótesis. No obstante, se puede demostrar que el polinomio característico de A está en $\mathcal{P}_u$ para todo u (Teorema de Cayley–Hamilton). Sin usar este resultado, se puede ver directamente que existe A^k que es combinación lineal de las potencias de menor grado de A, ya que el espacio de las matrices $n \times n$ tiene dimensión n^2. $\square$

Propiedad (6). Sean $\lambda_1, \ldots, \lambda_k$ los valores propios de A. Sean $j_1, \ldots, j_k \geq 1$. Entonces la siguiente suma de subespacios es directa:

$$E_{j_1}(\lambda_1) \oplus E_{j_2}(\lambda_2) \oplus \ldots \oplus E_{j_k}(\lambda_k).$$

Dem. Sean $u_1, \ldots, u_k$ tales que

$$u_1 + u_2 + \ldots + u_k = 0, \qquad u_i \in E_{j_i}(\lambda_i).$$

Tenemos que

$$\begin{aligned} &(A - \lambda_1 I)^{j_1} u_1 = 0, \\ &(A - \lambda_2 I)^{j_2} \ldots (A - \lambda_k I)^{j_k} u_1 = -(A - \lambda_2 I)^{j_2} \ldots (A - \lambda_k I)^{j_k} (u_2 + \ldots + u_k) = 0 \end{aligned}$$

luego por el Lema, $u_1 = 0$. De la misma forma se ve que todos los demás u_i son nulos. $\square$

Propiedad (7a). Para todo j

$$\dim E_j(\lambda) \leq m(\lambda).$$

Dem. Sea P invertible de forma que sus p primeras columnas formen una base de $E_j(\lambda)$ ($p := \dim E_j(\lambda)$). Entonces, con el cambio de base

$$P^{-1}\,A\,P = \begin{bmatrix} A_p & C_p \\ 0 & B_p \end{bmatrix}$$

(con A $p \times p$). Si $A_p\,v = \mu\,v$, entonces

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

Sea

$$w := P \begin{bmatrix} v \\ 0 \end{bmatrix} \in S(\mu) \cap E_j(\lambda).$$

Si $\mu \neq \lambda$, $S(\mu) \cap E_j(\lambda) = E_1(\mu) \cap E_j(\lambda) = 0$, luego $w = 0$ y, por tanto, $v = 0$. Entonces, hemos visto que el polinomio característico de A_p es $(x-\lambda)^p$. Puesto que el polinomio característico de A_p divide al de A, $\dim E_j(\lambda) = p \leq m(\lambda)$. $\square$

Propiedad (7b). Si $E_{\ell+1}(\lambda) = E_\ell(\lambda)$, entonces

$$\dim E_\ell(\lambda) \geq m(\lambda).$$

Dem. Sea $p := \dim E_\ell(\lambda)$. Tomamos P como en la demostración del teorema anterior, de forma que

$$P^{-1}\,A\,P = \begin{bmatrix} A_p & C_p \\ 0 & B_p \end{bmatrix} =: \widetilde{A},$$

donde A_p es $p \times p$. Vamos a ver que B_p no tiene el valor propio λ.

Notemos de nuevo que para todo j

$$u \in N(\,(\widetilde{A} - \lambda\,I)^j\,) \qquad \Longleftrightarrow \qquad P\,u \in N(\,(A - \lambda\,I)^j\,),$$

luego se deduce que:

$$N(\,(\widetilde{A} - \lambda\,I)^\ell\,) = N(\,(\widetilde{A} - \lambda\,I)^{\ell+k}\,), \qquad \forall j$$
$$N(\,(\widetilde{A} - \lambda\,I)^\ell\,) = \mathbb{R}\langle e_1, \ldots, e_p\rangle,$$

y de esta última igualdad, que las primeras p columnas de $(\widetilde{A} - \lambda\,I)^\ell$ son nulas Consideramos la matriz

$$(\widetilde{A} - \lambda\,I)^\ell = \begin{bmatrix} (A_p - \lambda\,I)^\ell & \widehat{C}_p \\ 0 & (B_p - \lambda\,I)^\ell \end{bmatrix} = \begin{bmatrix} 0 & \widehat{C}_p \\ 0 & (B_p - \lambda\,I)^\ell \end{bmatrix}$$

Si $(B_p - \lambda I)v = 0$, entonces

$$(\widetilde{A} - \lambda\,I)^{2\ell} \begin{bmatrix} 0 \\ v \end{bmatrix} = \begin{bmatrix} 0 & \widehat{C}_p \\ 0 & (B_p - \lambda\,I)^\ell \end{bmatrix} \begin{bmatrix} \widehat{C}_p v \\ (B_p - \lambda\,I)^\ell v \end{bmatrix} = \begin{bmatrix} 0 & \widehat{C}_p \\ 0 & (B_p - \lambda\,I)^\ell \end{bmatrix} \begin{bmatrix} \widehat{C}_p v \\ 0 \end{bmatrix} = 0$$

lo que implica

$$(\widetilde{A} - \lambda\,I)^\ell \begin{bmatrix} 0 \\ v \end{bmatrix} = 0$$

y por tanto $v = 0$. $\square$

Nota. La Propiedad (7b) es suficiente para garantizar que $\dim E_\ell(\lambda) = m(\lambda)$. La razón es que por la Propiedad (6),

$$n \geq \dim E_{\ell_1}(\lambda_1) + \ldots + \dim E_{\ell_k}(\lambda_k) \geq m(\lambda_1) + \ldots + m(\lambda_k) = n$$

lo que fuerza a que las desigualdades sean igualdades.

4. Cajas de Jordan y matrices de Jordan

Definición *Llamamos caja de Jordan de orden k asociada a un valor propio λ a una matriz $k \times k$ de la forma*

$$J_k(\lambda) := \begin{bmatrix} \lambda & 1 & & & \\ & \lambda & 1 & & \\ & & \lambda & \ddots & \\ & & & \ddots & 1 \\ & & & & \lambda \end{bmatrix}.$$

Propiedades.

(1) La matriz $J_k(\lambda)$ tiene como polinomio característico $(x-\lambda)^k$ y $\dim S(\lambda) = 1$.

(2) Las dimensiones de los núcleos iterados de $J_k(\lambda)$ son:

$$\dim N(\,(J_k(\lambda) - \lambda I)^j\,) = \left\{ \begin{array}{ll} j, & j = 1, \ldots, k \\ k, & j \geq k \end{array} \right\} = \min(k, j).$$

Dem. Se puede hacer directamente notando que

$$J_k(\lambda) - \lambda I = \begin{bmatrix} 0 & 1 & & & \\ & 0 & 1 & & \\ & & 0 & \ddots & \\ & & & \ddots & 1 \\ & & & & 0 \end{bmatrix}$$

y viendo cómo decrece de uno en uno el rango de las potencias de esta matriz. También se puede ver notando que, como $n_1 = 1$, por la propiedad (5) de las dimensiones de los núcleos iterados $n_j = j$ hasta que $j = k$. □

Definición *Una matriz de Jordan asociada a un único valor propio λ es una matriz diagonal por bloques cuyos bloques son cajas de Jordan asociadas a λ*

$$\begin{bmatrix} J_{k_1}(\lambda) & & & \\ & J_{k_2}(\lambda) & & \\ & & \ddots & \\ & & & J_{k_r}(\lambda) \end{bmatrix}$$

Una matriz de Jordan general es una matriz diagonal por bloques cuyos bloques son matrices de Jordan asociadas a distintos valores propios, o bien (equivalentemente), una matriz diagonal por bloques cuyos bloques son cajas de Jordan.

5. Dos problemas combinatorios

Objetivo. Vamos a mostrar seguidamente cómo se establece una relación uno a uno entre las distintas posibilidades de dimensiones de los núcleos iterados de una matriz de Jordan $m \times m$ con un único valor propio y las posibilidades de configurar las cajas de Jordan hasta llenar el espacio $m \times m$.

Problema (P). Hallar una sucesión de enteros que cumpla las tres siguientes propiedades:

(a) $0 = n_0 < n_1 < \ldots < n_\ell = n_{\ell+1} = \ldots,$

(b) $n_\ell - n_{\ell-1} \leq n_{\ell-1} - n_{\ell-2} \leq \ldots \leq n_2 - n_1 \leq n_1 - n_0 = n_1$

(c) $n_\ell = m.$

Problema (Q). Hallar un conjunto de valores $q_1, q_2, \ldots, q_\ell, \ldots$ de forma que

(a) $q_{\ell+1} = q_{\ell+2} = \ldots = 0$

(b) $q_1 + 2\,q_2 + 3\,q_3 + \ldots + \ell\, q_\ell = m$

Nota. Antes de mostrar cómo se relacionan, indicamos qué van a representar las soluciones de estos problemas y la de un problema intermedio:

- $\{n_j\}$ serán las dimensiones de los núcleos iterados;
- $\{p_j\}$ será el número de cajas de Jordan de tamaño $j \times j$ o superior;
- $\{q_j\}$ será el número de cajas de Jordan de tamaño $j \times j$.

El trabajo se podrá hacer sin tener en cuenta cuál es el valor de ℓ en el que la solución de (P) se detiene y la de (Q) toma su último valor no nulo, valor que de hecho será el mismo al relacionar ambas sucesiones. $\Box$

Propiedad A. Dada una solución del Problema (P), definimos

$$p_j := n_j - n_{j-1}, \qquad j \geq 1.$$

Nótese que esta sucesión cumple

$$0 = \ldots = p_{\ell+1} < p_\ell \leq p_{\ell+1} \leq \ldots \leq p_1 = n_1.$$

Seguidamente definimos

$$q_j := p_j - p_{j+1}, \qquad j \geq 1,$$

de modo que $q_\ell = p_\ell$ y $q_j = 0$ para todo $j > \ell$. Entonces esta sucesión es una solución del Problema (Q).

Propiedad B. Dada una solución del Problema (Q), definimos

$$p_j := q_j + q_{j+1} + \ldots + q_\ell, \qquad j = 1, \ldots, \ell$$

($p_j = 0$ para $j > \ell$) y por recurrencia

$$\left| \begin{array}{l} n_0 := 0 \\ n_j := p_j + n_{j-1}, \qquad j \geq 1. \end{array} \right.$$

Entonces esta sucesión es solución del Problema (P). Por consiguiente, hay una correspondencia biunívoca entra las soluciones de los (P) y (Q).

Teorema. Dada una solución de (P) y $\{q_j\}$ la solución asociada del Problema (Q), entonces la matriz de Jordan $m \times m$ asociada al valor propio λ construida con

$$q_j \textit{ bloques } j \times j \textit{ para cada } j$$

cumple

$$\dim E_j(\lambda) = n_j, \qquad j \geq 0.$$

Dem. Tomemos una matriz de Jordan construida de la forma anterior. Entonces, dado que la matriz es diagonal por bloques,

$$\begin{aligned}
\dim E_j(\lambda) &= q_\ell \dim N(\,(J_\ell(\lambda) - \lambda I)^j\,) + q_{\ell-1} \dim N(\,(J_{\ell-1}(\lambda) - \lambda I)^j\,) + \\
&\quad + \ldots + q_1 \dim N(\,(J_1(\lambda) - \lambda I)^j\,) = \\
&= q_\ell \min(j, \ell) + q_{\ell-1} \min(j, \ell - 1) + \ldots + q_1 \min(j, 1) = \\
&= j\,(q_\ell + q_{\ell-1} + \ldots + q_{j+1}) + j q_j + (j-1) q_{j-1} + \ldots + 2q_2 + q_1 = \\
&= j(p_\ell + p_{\ell-1} - p_\ell + \ldots + p_{j+1} - p_{j+2}) + \\
&\quad + j(p_j - p_{j+1}) + \ldots + 2(p_2 - p_3) + (p_1 - p_2) = \\
&= j\,p_{j+1} - j\,p_{j+1} + p_j + p_{j-1} + \ldots + p_1 = \\
&= n_j
\end{aligned}$$

□

Conclusiones.

(1) Para cualquier solución del Problema (P), existe una matriz (de hecho una matriz de Jordan) $m \times m$ con un único valor propio de forma que

$$\dim E_j(\lambda) = n_j, \qquad \forall j.$$

(2) Por consiguiente, dadas posibles configuraciones de las dimensiones de los núcleos iterados asociados a todos los valores propios, existe una matriz (de hecho, matriz de Jordan) que los tiene como núcleos iterados.

(3) Dos matrices de Jordan esencialmente distintas (que no se obtengan por permutación del orden de los bloques) no son semejantes.

Dem. La configuración en bloques (cuántos bloques hay de cada tamaño) da unívocamente una posibilidad de dimensiones de los núcleos iterados, pero ésta es un invariante por semejanza. □

6. Existencia de forma de Jordan

Primera observación. Si

$$P^{-1}\,A\,P = \begin{bmatrix} J_k(\lambda) & \times \\ 0 & \times \end{bmatrix}$$

y $u_1, \ldots, u_k$ son las primeras columnas de P, entonces

$$\begin{aligned} Au_1 &= \lambda u_1 \\ Au_2 &= \lambda u_2 + u_1 \\ &\vdots \\ Au_k &= \lambda u_k + u_{k-1} \end{aligned}$$

luego para todo $j \geq 2$

$$u_{j-1} = (A - \lambda I)u_j.$$

Segunda observación. Si S es un subespacio que cumple

$$E_{j-1}(\lambda) \oplus S \subseteq E_j(\lambda)$$

y $T := (A - \lambda I)S = \{(A - \lambda I)\,u \,|\, u \in S\}$, entonces

$$E_{j-2}(\lambda) \oplus T \subseteq E_{j-1}(\lambda), \qquad \dim T = \dim S.$$

Dem. Hay que ver tres cosas: (a) $T \subset E_{j-1}(\lambda)$, lo cual es obvio, ya que

$$u \in E_j(\lambda) \qquad \Longrightarrow \qquad (A - \lambda I)u \in E_{j-1}(\lambda);$$

(b) $T \cap E_{j-2}(\lambda) = 0$, ya que si $v \in T$ (luego $v = (A - \lambda I)u$, con $u \in S$), entonces

$$(A - \lambda I)^{j-2}v = 0 \qquad \Longrightarrow \qquad (A - \lambda I)^{j-1}u = 0 \qquad \Longrightarrow \qquad u \in S \cap E_{j-1}(\lambda) = 0;$$

(c) si $u \in S$, entonces

$$(A - \lambda I)u = 0 \qquad \Longrightarrow \qquad u \in E_1(\lambda) \cap S \subseteq E_{j-1}(\lambda) \cap S = 0,$$

luego la independencia lineal se preserva al multiplicar por $(A - \lambda I)$. $\square$

Teorema. Toda matriz $m \times m$ con un único valor propio λ es semejante a una matriz de Jordan asociada al valor propio λ. Además, las dimensiones de las cajas de Jordan vienen dadas por los coeficientes $\{q_j\}$ ligados a las dimensiones de los núcleos iterados.
Dem. Esa $E_\ell(\lambda)$ el primer núcleo iterado de tamaño máximo. Sea $S_\ell = S_\ell^0$ tal que

$$E_\ell(\lambda) = E_{\ell-1}(\lambda) \oplus S_\ell^0,$$

de modo que $\dim S_\ell^0 = n_\ell - n_{\ell-1} = p_\ell = q_\ell$. Por la observación anterior podemos poner

$$E_{\ell-1}(\lambda) = E_{\ell-2}(\lambda) \oplus \underbrace{(A-\lambda I)S_\ell^0 \oplus S_{\ell-1}^0}_{S_{\ell-1}}$$

con

$$\dim S_{\ell-1}^0 = (n_{\ell-1} - n_{\ell-2}) - \dim S_\ell^0 = p_{\ell-1} - p_\ell = q_{\ell-1} \geq 0.$$

De la misma forma

$$\begin{aligned} E_{\ell-2}(\lambda) &= E_{\ell-3}(\lambda) \oplus \underbrace{(A-\lambda I)S_{\ell-1} \oplus S_{\ell-2}^0}_{S_{\ell-2}} = \\ &= E_{\ell-3}(\lambda) \oplus (A-\lambda I)^2 S_\ell^0 \oplus (A-\lambda I)S_{\ell-1}^0 \oplus S_{\ell-2}^0 \end{aligned}$$

con

$$\dim S_{\ell-2}^0 = q_{\ell-2}.$$

Así sucesivamente se construye una partición de $E_\ell(\lambda)$

$$\begin{array}{ccccccccccc} E_\ell(\lambda) & = & S_\ell^0 & \oplus & (A-\lambda I)S_\ell^0 & \oplus & \dots & \oplus & (A-\lambda I)^{\ell-2}S_\ell^0 & \oplus & (A-\lambda I)^{\ell-1}S_\ell^0 \\ & & & \oplus & S_{\ell-1}^0 & \oplus & \dots & \oplus & (A-\lambda I)^{\ell-3}S_{\ell-1}^0 & \oplus & (A-\lambda I)^{\ell-2}S_{\ell-1}^0 \\ & & & & & & & \vdots & & \vdots & \\ & & & & & & & \oplus & S_2^0 & \oplus & (A-\lambda I)S_2^0 \\ & & & & & & & & & \oplus & S_1^0 \\ & & & & & & & & & & \\ & = & S_\ell & \oplus & S_{\ell-1} & \oplus & \dots & \oplus & S_2 & \oplus & S_1 \end{array}$$

de forma que

$$\dim S_j^0 = q_j, \qquad j = 1, \dots, \ell.$$

Construimos ahora una base de $E_\ell(\lambda)$ de la siguiente manera: se toma una base de cada S_j^0 (es posible que alognos sean nulos y no haya que tomar base)

$$u_{j,1}, \dots, u_{j,q_j}$$

y se construyen las secuencias de vectores para cada uno de ellos (digamos u)

$$\begin{array}{rcccl} v_1 & = & (A-\lambda I)^{j-1}u & = & (A-\lambda I)v_2 \\ v_2 & = & (A-\lambda I)^{j-2}u & = & (A-\lambda I)v_3 \\ & \vdots & & & \\ v_{j-1} & = & (A-\lambda I)u & = & (A-\lambda I)v_j \\ v_j & = & u, & & \end{array}$$

de manera que

$$Av_1 = \lambda v_1, \qquad Av_k = v_{k-1} + \lambda v_k, \quad k = 2, \dots, j.$$

La matriz obtenida por el cambio de base obtenido mediante la unión de estas secuencias de vectores es una matriz de Jordan con q_j cajas $j \times j$ para todo j. $\square$

Teorema. Toda matriz es semejante a una matriz de Jordan (su forma de Jordan). La forma de Jordan (única, salvo permutaciones de bloques) es la única matriz de Jordan en la que coinciden las dimensiones de los núcleos iterados con los de A.

Teorema. En la misma clase de semejanza de matrices están todas las matrices con el mismo polinomio característico e idénticas dimensiones de núcleos iterados y ninguna más.

7. Detalles adicionales

7.1. Teorema de Cayley–Hamilton

Teorema (Cayley–Hamilton). Sea $p(x)$ el polinomio característico de A. Entonces

$$p(A) = 0.$$

Dem. Una forma elemental de demostrar esto es hacerlo con la forma de Jordan. Puesto que $p(J) = P^{-1}p(A)P$, basta ver que $p(J) = 0$. Ahora bien, es elemental que

$$p(J) = \begin{bmatrix} p(J_{k_1}(\lambda_1)) & & & \\ & p(J_{k_2}(\lambda_2)) & & \\ & & \ddots & \\ & & & p(J_{k_r}(\lambda_r)) \end{bmatrix}$$

y por otro lado

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

luego el teorema ya está probado.

Sin llegar a emplear la forma de Jordan, el resultado es también una consecuencia del Teorema de Descomposición de Jordan. Por este teorema, y dado que

$$p(P^{-1}AP) = \begin{bmatrix} p(A_1) & & & \\ & p(A_2) & & \\ & & \ddots & \\ & & & p(A_k) \end{bmatrix},$$

bastaría con demostrar el resultado para matrices con un único valor propio. Ahora bien, si el polinomio característico de A es $(x - \lambda)^n$, entonces existe $\ell \le n$ tal que $E_\ell(\lambda) = N(\,(A - \lambda I)^\ell\,) = \mathbb{C}^n$, lo que implica que $(A - \lambda I)^\ell = 0$ con $\ell \le n$. $\square$

Comentarios. Este tipo de resultados motiva la introducción de los conceptos de polinomio mínimo, divisores elementales, factores invariantes, etc. Vamos simplemente a listar la relación entre este enfoque y el de polinomios:

(1) El polinomio mínimo de una matriz es el polinomio mónico de menor grado tal que $q(A) = 0$. Su existencia y unicidad se deducen de argumentos similares a los dados en el Lema de la Sección 3. Es fácil ver que si para cada λ_j, ℓ_j es el mínimo valor tal que

$$\dim E_{\ell_j}(\lambda_j) = m(\lambda_j),$$

entonces

$$m(x) := (x - \lambda_1)^{\ell_1} \dots (x - \lambda_k)^{\ell_k}$$

es el polinomio mínimo de la forma de Jordan de A y, por tanto, de A.

(2) Los divisores elementales de A son polinomios de la forma $(x - \lambda_j)^r$ construidos según la siguiente regla:

por cada bloque de Jordan $J_{k_j}(\lambda_j)$ se toma un polinomio $(x - \lambda_j)^{k_j}$.

Esto quiere decir que hay tantos divisores elementales como cajas de Jordan y que sus grados marcan los tamaños de las mismas.

(3) Una matriz es diagonalizable si y sólo si su forma de Jordan es diagonal. Equivalentemente, si y sólo si

$$E_2(\lambda) = E_1(\lambda), \qquad \forall\lambda \text{ valor propio de } A.$$

De nuevo, esto equivale a que todos los divisores elementales sean de grado uno y a que el polinomio mínimo no tenga raíces múltiples.

7.2. Forma traspuesta

En bastantes contextos es común encontrarse con la forma de Jordan compuesta de cajas de la forma

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

es decir, con los 1s debajo de la diagonal. Si se ha sabido encontrar la base que lleva a la forma de Jordan habitual, basta cambiar el orden de los vectores para pasar a esta forma. En concreto si los vectores

$$u_1, u_2, \dots, u_p$$

crean una caja $p \times p$ asociada a λ, los vectores

$$u_p, u_{p-1}, \dots, u_1$$

dan la misma con los unos por debajo de la diagonal.

Consecuencia. Toda matriz es semejante a su traspuesta. Esto también se puede ver notando que los núcleos iterados de A y $A^\top$ tienen las mismas dimensiones.

7.3. Forma de Jordan real

Si A es real, $\lambda = \alpha + \beta\imath$ (con $\beta \neq 0$) y ya se tiene la base calculada para la parte de la forma de Jordan referida a λ, entonces para $\overline{\lambda} = \alpha - \beta\imath$ basta conjugar los vectores obtenidos,esto es,

$$u \in E_j(\lambda) \qquad \Longleftrightarrow \qquad \overline{u} \in E_j(\overline{\lambda})$$

Por tanto, el número y tipo de cajas de Jordan es el mismo.

Si $u_1, \ldots, u_p$ son los vectores de la base que lleva a forma de Jordan relacionados con el valor propio $\lambda = \alpha + \beta\imath$ (con $\beta \neq 0$), luego $\overline{u_1}, \ldots, \overline{u_p}$ son los relacionados con $\overline{\lambda}$, entonces se toman

$$\begin{aligned} v_1 &= \operatorname{Re} u_1, & v_2 &= \operatorname{Im} u_1 \\ v_3 &= \operatorname{Re} u_2, & v_4 &= \operatorname{Im} u_2 \\ &\vdots \\ v_{2p-1} &= \operatorname{Re} u_p, & v_{2p} &= \operatorname{Im} u_p \end{aligned}$$

(Re e Im denotan la parte real e imaginaria del vector). Con estos vectores, las cajas de λ y $\overline{\lambda}$ se mezclan, dando lugar a cajas más grandes, pero ya reales. Por ejemplo

$$\left.\begin{array}{c} \begin{bmatrix} \alpha+\beta\imath & 1 \\ 0 & \alpha+\beta\imath \end{bmatrix} \\ \\ \begin{bmatrix} \alpha-\beta\imath & 1 \\ 0 & \alpha-\beta\imath \end{bmatrix} \end{array}\right\} \longmapsto \begin{bmatrix} \alpha & \beta & 1 & 0 \\ -\beta & \alpha & 0 & 1 \\ 0 & 0 & \alpha & \beta \\ 0 & 0 & -\beta & \alpha \end{bmatrix}$$

De esta forma, cuando hay valores propios complejos en una matriz real, se puede trabajar con cajas de Jordan por bloques

$$J_k(\lambda, \overline{\lambda}) := \begin{bmatrix} \Lambda & I_2 & & \\ & \Lambda & \ddots & \\ & & \ddots & I_2 \\ & & & \Lambda \end{bmatrix}, \qquad \Lambda := \begin{bmatrix} \alpha & \beta \\ -\beta & \alpha \end{bmatrix}, \qquad I_2 := \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

7.4. Matrices simétricas

Proposición. Toda matriz simétrica (hermitiana) es diagonalizable.
Dem. Vamos a limitarnos al caso simétrico. Si $Au = \lambda u$, con $u \in \mathbb{C}^n$ y $\lambda \in \mathbb{C}$. entonces

$$\lambda \|u\|^2 = \overline{u}^\top A\, u = \overline{(A\, u)}^\top u = \overline{\lambda}\|u\|^2,$$

luego λ es real. Basta ver ahora que para todo valor propio $E_2(\lambda) = E_1(\lambda)$, lo cual se concluye de esta cadena de implicaciones:

$$u \in E_2(\lambda) \qquad \Longrightarrow \qquad (A - \lambda I)^2 u = 0$$

$$
\begin{aligned}
\Longrightarrow \qquad & u^\top (A-\lambda I)^2 u = 0\\
\Longrightarrow \qquad & ((A-\lambda I)u)^\top ((A-\lambda I)u) = 0\\
\Longrightarrow \qquad & \|(A-\lambda I)u\|^2 = 0\\
\Longrightarrow \qquad & u \in E_1(\lambda).
\end{aligned}
$$

$\square$

by F.–J. Sayas
Mathematics, Ho!, October, 2002

Fuentes y otras lecturas

Eugenio Hernández. Álgebra y Geometría. Addison–Wesley.

Rainer Kress. Linear Integral Operators. Springer Verlag.