Álgebra I - Curso 2005/06 - Grupos M1 y M2

# Tema 9: **REDUCCIÓN POR SEMEJANZA DE UNA MATRIZ. DIAGONALIZACIÓN**

por Mario López Gómez

## 1. Valores y vectores propios.

**Definición.-** Dada una matriz **cuadrada** $A \in \mathbb{K}^{n\times n}$, se dice que $\lambda \in \mathbb{K}$ es un **valor propio** o **autovalor** de $A$ si existe algún vector **no nulo** $u \in \mathbb{K}^n$ tal que $Au = \lambda u$.

El vector $u$ anterior se dice **vector propio** o **autovector** de $A$, asociado al autovalor $\lambda$.

Así pues, los vectores propios de una matriz $A$ son aquellos vectores (no nulos) que se transforman mediante $A$ en proporcionales a sí mismos, siendo los valores propios las correspondientes constantes de proporcionalidad.

**Definición.-** Dados un espacio vectorial $E$ sobre $\mathbb{K}$ y $f$ un endomorfismo de $E$, se define valor propio o autovalor de $f$ como aquel valor $\lambda \in \mathbb{K}$ para el cual existe algún $v \in E$, $v \neq \mathbf{0}$, tal que $f(v) = \lambda v$. Para cada valor propio $\lambda$, los vectores propios o autovectores asociados a $\lambda$ son aquellos $v \in E$ tales que $f(v) = \lambda v$ (excluyendo el vector nulo).

**Observación.-** Si $E$ es de dimensión finita y $\mathcal{B}$ es una base de $E$, $\lambda$ es autovalor de $f$ si y sólo si lo es de la matriz de $f$ respecto a $\mathcal{B}$.

En esta asignatura restringiremos nuestro estudio casi exclusivamente a los valores y vectores propios en espacios de dimensión finita, es decir, valores o vectores propios de matrices (recuérdese que éstas han de ser necesariamente cuadradas). Todos los resultados que veamos para valores y vectores propios de una matriz cuadrada, tendrán su traducción inmediata para endomorfismos en dimensión finita.

**Definición.-** Dada una matriz cuadrada $A \in \mathbb{K}^{n\times n}$, se define su **espectro**, y se denota por $\sigma(A)$, como el conjunto de todos los autovalores de $A$, es decir,

$$\sigma(A) := \{\lambda \in \mathbb{K} : \lambda \text{ es autovalor de } A\}.$$

Idéntica definición se tiene para el espectro de un endomorfismo

**Observación.-** $Au = \lambda u \Leftrightarrow (A - \lambda I)u = \mathbf{0} \Leftrightarrow u \in \ker(A - \lambda I)$. Así pues, dada una matriz $A \in \mathbb{K}^{n \times n}$, **el conjunto de los vectores propios asociados a un mismo valor propio $\lambda$, junto con el vector nulo, es un subespacio vectorial de $\mathbb{K}^n$.**

**Observación.-** Una matriz $A$ es invertible si y solo si 0 no es autovalor de $A$.

**Definición.-** Dada $A \in \mathbb{K}^{n \times n}$ y $\lambda$ autovalor de $A$, llamamos **subespacio propio** o **subespacio característico** de $A$ asociado a $\lambda$ al subespacio formado por todos los vectores propios asociados a $\lambda$, además del vector nulo, es decir,

$$\ker(A - \lambda I) = \{u \in \mathbb{K}^n : Au = \lambda u\}.$$

# 2. El polinomio característico.

Por lo visto anteriormente, un escalar $\lambda \in \mathbb{K}$ es autovalor de una matriz $A$ de orden $n$ si y sólo si $\dim(\ker(\lambda I - A)) \geq 1$, es decir, si $r(\lambda I - A) < n$. Esta última condición equivale a su vez a que

$$\det(\lambda I - A) = 0,$$

que es una ecuación en la incógnita $\lambda$ cuyas soluciones son precisamente los autovalores de $A$.

Si consideramos

$$\chi_A(\lambda) := |\lambda I - A| = \begin{vmatrix} \lambda - a_{11} & -a_{12} & \cdots & -a_{1n} \\ -a_{21} & \lambda - a_{22} & \cdots & -a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ -a_{n1} & -a_{n2} & \cdots & \lambda - a_{nn} \end{vmatrix},$$

observamos que el desarrollo de tal determinante nos da un polinomio en la indeterminada $\lambda$, cuyo término de mayor grado, que se obtiene a partir del producto de los elementos de la diagonal, es $\lambda^n$; se trata, pues, de un polinomio mónico de grado igual al orden de la matriz.

**Definición.-** Dada $A \in \mathbb{K}^{n \times n}$, se llama **polinomio característico** de $A$, y se denota por $\chi_A$, al polinomio con coeficientes en $\mathbb{K}$, mónico de grado $n$, definido por

$$\chi_A(\lambda) := |\lambda I - A|,$$

cuyas raíces son los autovalores de $A$.

**Observación.-** Evaluando $\chi$ en el 0, obtenemos $\chi(0) = \det(-A) = (-1)^n \det A$; sabemos que $\chi(0)$ es el término independiente de dicho polinomio, que coincidirá por tanto con el valor de $(-1)^n \det A$.

**Definición.-** Dada $A \in \mathbb{K}^{n \times n}$, se define su *traza* como $\mathrm{tr} A = a_{11} + a_{22} + \ldots + a_{nn}$, es decir, la suma de los elementos de su diagonal principal.

**Observación.-** El término de grado $n-1$ de $\chi_A$ se obtiene también a partir del producto de todos los elementos de la diagonal principal:

$$(\lambda - a_{11})(\lambda - a_{22})\dots(\lambda - a_{nn}) = \lambda^n - (a_{11} + a_{22} + \dots + a_{nn})\lambda^{n-1} + \dots$$

Se observa, pues, que **el coeficiente de grado $n-1$ de $\chi_A$ es el opuesto a la traza de $A$.**

**Proposición.- Matrices semejantes entre sí tienen el mismo polinomio característico.**

**Demostración.-** Sean $A, B$ matrices cuadradas del mismo orden tales que $B = P^{-1}AP$ para una cierta matriz de paso $P$; entonces

$$\chi_B(\lambda) = |\lambda I - B| = |\lambda I - P^{-1}AP| = |P^{-1}\lambda IP - P^{-1}AP| = |P^{-1}(\lambda I - A)P| =$$
$$= |P^{-1}||(\lambda I - A)||P| = |P|^{-1}|(\lambda I - A)||P| = |(\lambda I - A)| = \chi_A(\lambda).$$

Así pues, el polinomio característico es un dato intrínseco de cada endomorfismo, es decir, no depende de la base concreta (de la expresión matricial concreta) en que el endomorfismo se representa. Por lo visto anteriormente, también se conservan mediante semejanza la traza de la matriz y el determinante, es decir:

**Corolario.-** Matrices semejantes entre sí tienen el mismo determinante y la misma traza.

**Definición.-** Dada $A \in \mathbb{K}^{n\times n}$, un autovalor $\lambda$ de $A$ se dice que es autovalor múltiple de orden $m$ (**tiene multiplicidad** $m$) si tiene multiplicidad $m$ como raíz de $\chi_A$.

**Nota.-** La multiplicidad antes definida también se llama a veces *multiplicidad algebraica* de un autovalor, mientras que $\dim(\ker(A - \lambda I))$ se llama *multiplicidad geométrica* de $\lambda$.

Obsérvese que si $A$ es una matriz real, puede tener algunos o todos sus autovalores complejos no reales; si $A$ es una matriz compleja de orden $n$, por el teorema de factorización en $\mathbb{C}$ tiene $n$ autovalores, contando sus multiplicidades.

**Observación.-** Si $A \in \mathbb{R}^{n\times n}$ y $\lambda$ es un autovalor de $A$, entonces su conjugado $\overline{\lambda}$ también lo es (y de la misma multiplicidad). Así pues, **los autovalores complejos no reales de una matriz real son conjugados dos a dos.** En particular, **toda matriz real y de orden impar tiene algún autovalor real**.

**Observación.-** Si $A$ es una matriz real y $\lambda$ un autovalor de $A$, entonces

$$Au = \lambda u \Leftrightarrow \overline{Au} = \overline{\lambda u} \Leftrightarrow A\overline{u} = \overline{\lambda}\overline{u},$$

es decir, que el subespacio propio $\ker(A - \overline{\lambda}I)$ está formado por los conjugados de los vectores de $\ker(A-\lambda I)$. Por lo tanto, los subespacios propios $\ker(A-\lambda I)$ y $\ker(A-\overline{\lambda}I)$ tienen la misma dimensión.

# 3. Diagonalización.

Nuestro objetivo es, dada una matriz cuadrada $A$, encontrar, cuando sea posible, una matriz **diagonal** semejante a $A$; esto es equivalente a encontrar una base de $\mathbb{K}^n$ en la que el endomorfismo determinado por $A$ se exprese de forma diagonal, es decir, los vectores de la base se transformen mediante $A$ en proporcionales a sí mismos; diagonalizar una matriz $A$ equivale, por tanto, a encontrar una base de vectores propios de $A$ (y los correspondientes valores propios).

El primer resultado importante sobre diagonalización es que **vectores propios asociados a valores propios distintos son linealmente independientes**, es decir:

**Proposición.-** Sea $A \in \mathbb{K}^{n\times n}$, sean $u_1, u_2, \ldots, u_k$ vectores propios de $A$ asociados, respectivamente, a los valores propios $\lambda_1, \lambda_2, \ldots, \lambda_k$, siendo éstos distintos dos a dos. Entonces $u_1, u_2, \ldots, u_k$ son linealmente independientes.

**Demostración.-** Lo demostraremos por inducción en $k$, el número de valores propios distintos. Para $k = 1$, el resultado es trivial. Supongamos que es cierto para $k$ valores propios. Si tenemos ahora $k+1$ valores propios distintos, $\lambda_1, \ldots, \lambda_{k+1}$, cuyos vectores propios correspondientes son $u_1, \ldots, u_{k+1}$, supongamos que, para unos escalares $\alpha_1, \ldots, \alpha_{k+1} \in \mathbb{K}$, se tiene que

$$\sum_{j=1}^{k+1} \alpha_j u_j = \mathbf{0}.$$

Aplicándole a la combinación lineal anterior la matriz $A - \lambda_{k+1} I$, y puesto que $Au_j = \lambda_j u_j$, se tiene que

$$\mathbf{0} = (A - \lambda_{k+1} I) \sum_{j=1}^{k+1} \alpha_j u_j = \sum_{j=1}^{k+1} \alpha_j (A - \lambda_{k+1} I) u_j = \sum_{j=1}^{k} \alpha_j (\lambda_j - \lambda_{k+1}) u_j.$$

La hipótesis de inducción (el resultado cierto para $k$) nos dice que los vectores propios $u_1, \ldots, u_k$ son linealmente independientes, luego los escalares $\alpha_j(\lambda_j - \lambda_{k+1})$, $j = 1, 2, \ldots, k$ de la última combinación lineal han de ser todos nulos. Pero, puesto que $\lambda_{k+1}$ es distinto de todos los restantes $\lambda_j$, esto implica que $\alpha_j = 0$ para $j = 1, 2, \ldots, k$. Por tanto, nos queda que $\alpha_{k+1} u_{k+1} = \mathbf{0}$ que, al ser $u_{k+1}$ no nulo, nos permite concluir que también es $\alpha_{k+1} = 0$. Así pues, todos los coeficientes $\alpha_j$ son forzosamente nulos, lo que demuestra la independencia lineal de los $u_j$.

**Corolario.- Los subespacios propios de una matriz dan suma directa.**

**Demostración.-** Sean $\lambda_1, \ldots, \lambda_k$ valores propios de una misma matriz $A$, distintos dos a dos. Para cada $i \in \{1, 2, \ldots, k\}$, consideramos el subespacio $\ker(A - \lambda_i I) \cap \sum_{j\neq i} \ker(A - \lambda_j I)$, y tenemos que comprobar que este subespacio se reduce al nulo.

Sea un vector $u \in \ker(A - \lambda_i I) \cap \sum_{j\neq i} \ker(A - \lambda_j I)$; por pertenecer a $\sum_{j\neq i} \ker(A - \lambda_j I)$, se podrá expresar como $\sum_{j\neq i} u_j$, en donde $u_j \in \ker(A - \lambda_j I)$ para cada $j \neq i$; pero, por otro lado $u \in \ker(A - \lambda_i I)$, con lo que

$$\mathbf{0} = -u + \sum_{j \neq i} u_j$$

no puede tener ningún sumando no nulo, pues los sumandos no nulos darían una combinación lineal de vectores propios asociados a valores propios distintos, y por tanto linealmente independientes, igualada al vector nulo. Luego nuestro vector $u$ de partida no puede ser otro que el nulo.

**Definición.-** Sea $A \in \mathbb{K}^{n \times n}$; se dice que $A$ es **diagonalizable** si es semejante a una matriz diagonal con elementos en $\mathbb{K}$, es decir,

$$\exists P \in \mathbb{K}^{n \times n} \text{ invertible tal que } P^{-1}AP = \operatorname{diag}(\lambda_1, \ldots, \lambda_n).$$

**Observación.-** Los elementos $\lambda_j$ de la expresión anterior son forzosamente autovalores de $A$, mientras que las columnas de la matriz $P$ son autovectores a ellos asociados.

**Observación.-** Si $A \in \mathbb{K}^{n \times n}$ tiene $n$ autovalores distintos, entonces es diagonalizable.

**Observación.-** $A \in \mathbb{K}^{n \times n}$ es diagonalizable si y solo si existe una base de autovectores de $A$, lo cual es equivalente a que la suma de las dimensiones de los subespacios propios de $A$ sea como mínimo $n$; pero el importante resultado que enunciamos a continuación nos dice que estas dimensiones no pueden sumar más de $n$.

**Teorema.-** Sea $A \in \mathbb{K}^{n \times n}$ y sea $\lambda \in \mathbb{K}$ un autovalor de $A$ de multiplicidad $m$. Entonces

$$\dim(\ker(A - \lambda I)) \leq m.$$

Es decir, que la *multiplicidad geométrica* de un autovalor no puede exceder de su *multiplicidad algebraica.*

**Demostración.-** Consideramos un autovalor $\lambda_0$ de una matriz $A \in \mathbb{K}^{n \times n}$; llamamos $p = \dim(\ker(A - \lambda I))$, y vamos a ver que $\lambda_0$ es raíz de $\chi_A$ de multiplicidad mayor o igual que $p$.

Tenemos, pues, $p$ vectores propios $u_1, \ldots, u_p$ linealmente independientes, asociados a $\lambda$. Por el teorema de ampliación de la base podemos ampliar ese sistema libre con $n - p$ vectores hasta completar una base de $\mathbb{K}^n$, digamos

$$\mathcal{B} = (u_1, \ldots, u_n).$$

Si consideramos la matriz $P$ de paso cuyas columnas son los vectores de dicha base, tenemos que $P^{-1}AP$ es una matriz de la forma

$$\tilde{A} = \left( \begin{array}{ccc|c} \lambda_0 & \cdots & 0 & \\ \vdots & \ddots & \vdots & B \\ 0 & \cdots & \lambda_0 & \\ 0 & \cdots & 0 & \\ \vdots & \ddots & \vdots & C \\ 0 & \cdots & 0 & \end{array} \right),$$

en donde $B$ y $C$ son submatrices, siendo $C$ cuadrada de orden $n-p$. Como el polinomio característico se conserva mediante semejanza, es $\chi_A = \chi_{\tilde{A}}$. Pero resulta evidente que

$$\chi_{\tilde{A}}(\lambda) = \det(\lambda I - \tilde{A}) = (\lambda - \lambda_0)^p \det(\lambda I - C) = (\lambda - \lambda_0)^p \chi_C(\lambda),$$

con lo que queda claro que $\lambda_0$ tiene como mínimo multiplicidad $p$.

**Corolario.-** Caracterización de las matrices diagonalizables.
Sea $A \in \mathbb{K}^{\times n}$; $A$ es diagonalizable (en $\mathbb{K}$) si y solo si se satisfacen:

i) $\chi_A$ tiene, contando sus multiplicidades, $n$ raíces en $\mathbb{K}$.

ii) Para todo autovalor $\lambda$ de $A$, $\dim(\ker(A - \lambda I))$ coincide con la multiplicidad algebraica de $\lambda$.

En particular: una matriz **compleja** es diagonalizable si y solo si para todo autovalor $\lambda$ de $A$, $\dim(\ker(A - \lambda I))$ coincide con la multiplicidad algebraica de $\lambda$.

Una matriz **real** es diagonalizable (en $\mathbb{R}$) si y sólo si todos sus autovalores son reales y, además, para todos ellos se cumple la condición de la dimensión.

**Observaciones:** Todas las propiedades siguientes se enuncian para una matriz cuadrada de orden $\mathbb{K}^{n \times n}$:

1. **Si $A$ es diagonalizable, entonces $A^t$ también lo es** (y la matriz de paso que la diagonaliza es la inversa de la transpuesta de la de $A$).

   En efecto, si $P^{-1}AP = D$ diagonal, entonces

   $$D = D^t = (P^{-1}AP)^t = P^t A^t (P^{-1})^t = P^t A^t (P^t)^{-1}.$$

2. **Si $A$ es diagonalizable e invertible, entonces $A^{-1}$ también es diagonalizable** (y la matriz de paso que la diagonaliza es la misma que la de $A$)..

   En efecto, si $P^{-1}AP = D = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$ con todos los $\lambda_j$ no nulos, entonces

   $$\operatorname{diag}(\lambda_1^{-1}, \ldots, \lambda_n^{-1}) = D^{-1} = (P^{-1}AP)^{-1} = P^{-1}A^{-1}P.$$

3. **Si $A$ es diagonalizable, entonces cualquier potencia de $A$ también lo es** (y la matriz que la diagonaliza es la misma).

   En efecto, $P^{-1}AP = D = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$ implica que

   $$\operatorname{diag}(\lambda_1^k, \ldots, \lambda_n^k) = D^k = (P^{-1}AP)^k = (P^{-1}AP)(P^{-1}AP)\cdots(P^{-1}AP) = P^{-1}A^kP.$$

# 4. Reducción por semejanza de una matriz a forma triangular.

**Teorema.-** Toda matriz cuadrada compleja es semejante a una triangular superior.

**Demostración.-** Razonaremos por inducción sobre el orden $n$ de la matriz. Para $n = 1$ es trivialmente cierto. Supongamos que el resultado es válido para $n - 1$.

Sea $A \in \mathbb{C}^{n\times n}$,

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}.$$

Por el teorema fundamental del Álgebra, $\chi_A$ tiene alguna raíz, es decir, existe algún valor propio de $A$, digamos $\lambda_1 \in \mathbb{C}$. Sea $u_1 \in \mathbb{C}^n$ un vector propio no nulo asociado a $\lambda_1$. Por el teorema de ampliación de la base, existen $u_2, \ldots, u_n \in \mathbb{C}^n$ tales que $\mathcal{B} = (u_1, u_2, \ldots, u_n)$ es una base de $\mathbb{C}^n$.

Sea $P$ la matriz de paso de la canónica a esta base, es decir, la matriz cuyas columnas son los $n$ vectores $u_1, \ldots, u_n$. La matriz $P^{-1}AP$, semejante a $A$, que es la que expresa en la base $\mathcal{B}$ el endomorfismo de $A$, es de la forma

$$P^{-1}AP = \begin{pmatrix} \lambda_1 & a'_{12} & \cdots & a'_{1n} \\ 0 & a'_{21} & \cdots & a'_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & a'_{n2} & \cdots & a'_{nn} \end{pmatrix},$$

ya que la primera columna de dicha matriz es el vector de coordenadas de $Au_1$ respecto a la propia base $(u_1, \ldots, u_n)$.

Ahora podemos considerar la submatriz

$$\tilde{A} = \begin{pmatrix} a'_{12} & \cdots & a'_{1n} \\ \vdots & \ddots & \vdots \\ a'_{n2} & \cdots & a'_{nn} \end{pmatrix} \in \mathbb{C}^{(n-1)\times(n-1)}$$

y aplicarle la hipótesis de inducción, es decir, que existe una matriz $\tilde{P} \in \mathbb{C}^{(n-1)\times(n-1)}$ invertible tal que $\tilde{T} = \tilde{P}^{-1}\tilde{A}\tilde{P}$ es triangular superior.

Si ahora consideramos la matriz invertible

$$Q = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & & \tilde{P} & \\ 0 & & & \end{pmatrix} \in \mathbb{C}^{n\times n},$$

es fácil ver multiplicando por cajas que el producto

$$\begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & & \tilde{P} & \\ 0 & & & \end{pmatrix}^{-1} \begin{pmatrix} \lambda_1 & a'_{12} & \cdots & a'_{1n} \\ 0 & & & \\ \vdots & & \tilde{A} & \\ 0 & & & \end{pmatrix} \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & & \tilde{P} & \\ 0 & & & \end{pmatrix} =$$

$$= \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & & \tilde{P}^{-1}\tilde{A}\tilde{P} & \\ 0 & & & \end{pmatrix} = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & & \tilde{T} & \\ 0 & & & \end{pmatrix} = T,$$

es triangular superior. O lo que es lo mismo:

$$(PQ)^{-1}A(PQ) = Q^{-1}(P^{-1}AP)Q = T$$

luego nuestra matriz $A$ de partida es semejante a una triangular superior.

# 5. Nociones sobre la forma canónica de Jordan.

Cuando una matriz $A \in \mathbb{C}^{n\times n}$ no sea diagonalizable, intentaremos expresar el endomorfismo de $A$ de la forma más sencilla posible; podrá encontrarse una base de $\mathbb{C}^n$ en la que el endomorfismo se represente mediante una matriz que solamente tendrá elementos no nulos en la diagonal principal y en la línea inmediatamente superior a ésta, cuyos elementos sólo podrán ser unos y ceros.

En particular, será una matriz **diagonal por bloques**, siendo estos bloques de la forma

$$\begin{pmatrix} \lambda_0 & 1 & & & & \\ & \lambda_0 & 1 & & & \\ & & \lambda_0 & & & \\ & & & \ddots & & \\ & & & & \lambda_0 & 1 \\ & & & & & \lambda_0 \end{pmatrix}$$

en donde $\lambda_0$ es un autovalor de $A$.

Un caso particular es el bloque de orden 1, en que sólo aparece el autovalor en la diagonal. Puede haber varios bloques asociados a un mismo autovalor, y cada autovalor debe aparecer en la diagonal un número de veces igual a su orden de multiplicidad. Este número de veces coincide con la suma de los órdenes de los bloques asociados a dicho autovalor.

El número de unos "encima"de $\lambda_0$ es la diferencia entre su orden de multiplicidad y $\dim(\ker(A-\lambda_0 I))$. Obsérvese que este número coincide con la suma de los órdenes de los bloques asociados a $\lambda_0$ menos el número de dichos bloques.

El vector de la nueva base asociado a la primera columna de cada bloque es un vector propio ($Au_j = \lambda_0 u_j$). Los restantes vectores $u_j$ verifican

$$Au_j = u_{j-1} + \lambda_0 u_j,$$

cada uno se transforma en combinación lineal del anterior y de sí mismo. Cada bloque nos da, pues, una cadena de vectores de la base así relacionados, el primero de los cuales es un vector propio. El subespacio engendrado por todos ellos es invariante mediante $A$.

Una matriz semejante a $A$ con las propiedades dichas se denomina **forma canónica de Jordan** de $A$, y es única salvo el orden de los bloques. Obviamente, si $A$ es diagonalizable, su forma canónica de Jordan es la diagonal semejante a $A$.

# 6. Matriz asociada a un polinomio.

**Definición.-** Sea $P$ un polinomio mónico de grado $n$,

$$P(x) = a_0 + a_1 x + \ldots + a_{n-1}x^{n-1} + x^n.$$

Se define su matriz *compañera* como

$$\begin{pmatrix} 0 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1 & \cdots & 0 & 0 \\ 0 & 0 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & 1 \\ -a_0 & -a_1 & -a_2 & \cdots & -a_{n-2} & -a_{n-1} \end{pmatrix}.$$

**Proposición.-** Para cada polinomio $P$, el polinomio característico de la matriz *compañera* de $P$ es el propio $P$.

**Demostración.-** Véase el ejercicio 9.5 de la colección.

Así pues, todo polinomio mónico es el polinomio característico de alguna matriz.

**Proposición.-** Si $C$ es una matriz *compañera* y $\lambda_0$ es un valor propio de $C$, entonces $\dim(\ker(C - \lambda_0 I)) = 1$.

**Demostración.-** Si $C$ es de orden $n$, $C - \lambda_0 I$ tiene al menos $n-1$ columnas linealmente independientes, debido a los unos y ceros por encima de la diagonal. Luego $r(C - \lambda_0 I) = n - 1$.

**Corolario.-** Una matriz *compañera* es diagonalizable si y sólo si todos sus valores propios son simples.

# 7. Polinomios anuladores. El polinomio mínimo.

## 7.1. Evaluación de un polinomio en una matriz.

### 7.1.1. Definición y propiedades.

Una matriz **cuadrada** $A$ con elementos en un cuerpo $\mathbb{K}$, admite la potenciación para cualquier exponente natural (definiendo $A^0 = I$). De este modo, dado cualquier polinomio $p(x) = a_0 + a_1 x + a_2 x^2 + \ldots + a_k x^k \in \mathbb{K}[x]$, podemos definir

$$p(A) := a_0 I + a_1 A + a_2 A^2 + \ldots + a_k A^k,$$

que es otra matriz cuadrada del mismo orden que $A$.

Por las propiedades distributivas del producto de matrices respecto a la suma, y del producto de escalar por matriz respecto a la suma de matrices, tenemos que, si $p, q \in \mathbb{K}[x]$,

$$\begin{aligned} (p+q)(A) &= p(A) + q(A); \\ (pq)(A) &= p(A)q(A), \end{aligned}$$

de modo que podemos tratar los polinomios evaluados en matrices como si estuvieran evaluados en escalares.

Obsérvese, además, que cualesquiera dos **matrices de la forma** $p(A), q(A)$ **conmutan**.

### 7.1.2. Valores propios de una matriz polinomial.

**Proposición.-** Dada una matriz $A \in \mathbb{K}^{n\times n}$, para cualquier polinomio $p \in \mathbb{K}[x]$ se cumple:

Si $\lambda \in \mathbb{K}$ es un valor propio de $A$ , entonces, $p(\lambda)$ es un valor propio de $p(A)$.

Además, si $v \in \mathbb{K}^n$ es un vector propio de $A$ asociado a $\lambda$, entonces es vector propio de $p(A)$ asociado a $p(\lambda)$.

**Demostración.-** $Av = \lambda v \Rightarrow A^j v = \lambda^j v \Rightarrow \sum_{j=0}^{k} a_j A^j v = \left(\sum_{j=0}^{k} a_j \lambda^j\right) v.$

Además, cuando el cuerpo es $\mathbb{C}$, se tiene el siguiente recíproco:

**Proposición.-** $A \in \mathbb{C}^{n\times n}$, $p \in \mathbb{C}[x]$; si $\lambda \in \mathbb{C}$ es un valor propio de $p(A)$, entonces existe $\alpha \in \mathbb{C}$ tal que $p(\alpha) = \lambda$.

**Demostración.-** Consideramos el polinomio $p(z) - \lambda$ que, salvo un caso trivial, será no constante, y por el teorema de factorización se podrá escribir como

$$p(z) - \lambda = c\prod_{j=1}^{k}(z - \alpha_j),$$

con $c \neq 0$ y los $\alpha_j$ números complejos.

Evaluando el polinomio anterior en la matriz $A$, obtenemos

$$p(A) - \lambda I = c\prod_{j=1}^{k}(A - \alpha_j I),$$

y, tomando determinantes,

$$\det(p(A) - \lambda I) = c^n \prod_{j=1}^{k} \det(A - \alpha_j I).$$

Ahora bien, el primer miembro es nulo por ser $\lambda$ valor propio de $p(A)$, de suerte que algún factor del segundo miembro debe anularse también, es decir, $\det(A - \alpha_s I) = 0$, lo que significa que $\alpha_s$ es valor propio de $A$. Pero este valor es raíz del polinomio $p(z) - \lambda$, es decir, satisface que $p(\alpha_s) = \lambda$.

## 7.2. Ejemplo: las matrices circulantes.

### 7.2.1. Definición de matriz circulante.

Dado un número natural $n$ y un vector $\mathbf{a} = (a_1, a_2, \ldots .a_n)$ , se define la **matriz circulante asociada** $C_n$ como

$$C_n(\mathbf{a}) = \begin{pmatrix} a_1 & a_2 & a_3 & \cdots & a_n \\ a_n & a_1 & a_2 & \cdots & a_{n-1} \\ a_{n-1} & a_n & a_1 & \cdots & a_{n-2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_2 & a_3 & a_4 & \cdots & a_1 \end{pmatrix}.$$

Obsérvese que las filas (y columnas) de una tal matriz contienen permutaciones de los elementos de la primera. Además, sobre la diagonal principal y cada línea paralela a ella, todos los elementos son iguales. Asimismo, los elementos de la línea que "comienza" en el $(1, j)$ son iguales a los de la línea que "termina" en el $(j-1, n)$.

### 7.2.2. La matriz circulante $S_n$.

Cuando, fijado $n$, consideramos el vector $\mathbf{e}_2$ (segundo vector canónico), la matriz circulante asociada se llama $S_n$ :

$$S_n = (\mathbf{e}_2 |\, \mathbf{e}_3 |\, \mathbf{e}_4 | \ldots |\, \mathbf{e}_n |\, \mathbf{e}_1) = \begin{pmatrix} 0 & 1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 0 & 1 & \cdots & 0 & 0 \\ 0 & 0 & 0 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & 0 & 1 \\ 1 & 0 & 0 & 0 & \cdots & 0 & 0 \end{pmatrix}$$

### 7.2.3. Propiedades de $S_n$.

Obsérvese que $S_n$ es una matriz de permutación, concretamente la de la permutación que lleva cada columna a la siguiente (llevando la última a la primera), es decir, rota las columnas hacia la derecha.

Las potencias de $S_n$ son matrices de permutación, $S_n^k$ da la permutación que desplaza cada columna $k$ lugares hacia la derecha:

$$S_n^2 = (\mathbf{e}_3 |\, \mathbf{e}_4 |\, \mathbf{e}_5 | \ldots |\, \mathbf{e}_1 |\, \mathbf{e}_2)\,; \; S_n^3 = (\mathbf{e}_4 |\, \mathbf{e}_5 |\, \mathbf{e}_6 | \ldots |\, \mathbf{e}_2 |\, \mathbf{e}_3)\,; \; \ldots$$

Así pues, $S_n^n = I$, es decir, la permutación es de orden $n$.

También puede interpretarse, por supuesto, como matriz de permutación por filas (un lugar hacia arriba).

$S_n$ es, como todas las de permutación, una **matriz ortogonal** (es decir, $S_n^t = S_n^{-1}$.

Obsérvese, finalmente, que $S_n$ es una matriz *compañera* asociada al polinomio (de grado $n$) $q(\lambda) = \lambda^n - 1$, siendo éste, por tanto, su **polinomio característico**.

### 7.2.4. Caracterización de las matrices circulantes.

Una matriz cuadrada de orden $n$ es circulante si y sólo si es resultado de evaluar un polinomio en $S_n$, es decir:

$$A \in \mathbb{K}^{n \times n};\; A \text{ es circulante } \Leftrightarrow \exists p \in \mathbb{K}[x] \text{ tal que } p(S_n) = A.$$

**Demostración.-** $\boxed{\Rightarrow}$

Si $A = \begin{pmatrix} a_1 & a_2 & a_3 & \cdots & a_n \\ a_n & a_1 & a_2 & \cdots & a_{n-1} \\ a_{n-1} & a_n & a_1 & \cdots & a_{n-2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_2 & a_3 & a_4 & \cdots & a_1 \end{pmatrix}$, se descompone fácilmente como

$$A = a_1 I + a_2 S_n + a_3 S_n^2 + \ldots + a_n S_n^{n-1}.$$

$\boxed{\Leftarrow}$ Sea un polinomio $q(\lambda) = b_0 + b_1\lambda + \ldots + b_m\lambda^m$. Para $k \geq n,\ k = nl + r$, con $0 \leq r \leq n-1$, se tiene que $A^k = A^r$, de modo que

$$\begin{aligned} q(A) &= b'_0 I + b'_1 A + \ldots + b'_{n-1}A^{n-1} = \begin{pmatrix} b'_0 & b'_1 & b'_2 & \cdots & b'_{n-1} \\ b'_{n-1} & b'_0 & b'_1 & \cdots & b'_{n-2} \\ b'_{n-2} & b'_{n-1} & b'_0 & \cdots & b'_{n-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ b'_1 & b'_2 & b'_3 & \cdots & b'_0 \end{pmatrix} = \\ &= C_n\left(b'_0, b'_1, \ldots, b'_{n-1}\right). \end{aligned}$$

### 7.2.5. Valores y vectores propios de la matriz $S_n$.

Como se ha visto antes, el polinomio característico de $S_n$ es $\lambda^n - 1$, de modo que los valores propios de esta matriz son las raíces $n$-ésimas de la unidad, a saber:

$$\lambda_k = \xi^k,\ k = 0, 1, \ldots, n-1,\ \text{siendo } \xi = e^{i\frac{2\pi}{n}}.$$

Al tener $n$ valores propios distintos, $S_n$ es, evidentemente, diagonalizable.

**Nota:** Al ser una matriz ortogonal, lo cual es un caso particular de **matriz normal** (concepto que se estudiará en Álgebra II), $S_n$ será **diagonalizable unitariamente**, es decir, a través de una matriz de paso unitaria (extensión al campo complejo del concepto de matriz ortogonal).

Todos los valores propios tienen módulo unidad.

Los vectores propios de una matriz *compañera* vienen dados por las columnas de la matriz de Vandermonde construida a partir de los valores propios, es decir:

$$\ker\left(S_n - \xi^k I\right) = L\left\{\left(1, \xi^k, \xi^{2k}, \ldots, \xi^{n-k}\right)\right\},$$

obteniéndose la matriz de paso:

$$P = \begin{pmatrix} 1 & 1 & 1 & 1 & \cdots & 1 \\ 1 & \xi & \xi^2 & \xi^3 & \cdots & \xi^{n-1} \\ 1 & \xi^2 & \xi^4 & \xi^6 & \cdots & \xi^{n-2} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & \xi^{n-1} & \xi^{n-2} & \xi^{n-3} & \cdots & \xi \end{pmatrix},$$

llamada **matriz de Fourier** de orden $n$.

Así, $P^{-1}S_nP = \bar{P}^tS_nP = \text{diag}(1, \xi, \xi^2, \ldots, \xi^{n-1})$.

### 7.2.6. Valores y vectores propios de una matriz circulante.

Como

$$\begin{aligned} C_n(\mathbf{a}) &= \sum_{j=1}^n a_j S_n^j = \sum_{j=1}^n a_j \left(P\text{diag}(1, \xi, \xi^2, \ldots, \xi^{n-1})P^{-1}\right)^j = \\ &= \sum_{j=1}^n a_j P\text{diag}(1, \xi, \xi^2, \ldots, \xi^{n-1})^j P^{-1} = \\ &= P\left(\sum_{j=1}^n a_j \text{diag}(1, \xi^j, \xi^{2j}, \ldots, \xi^{n-j})\right)P^{-1} = \\ &= P\text{diag}\left(\sum_{j=1}^n a_j, \sum_{j=1}^n a_j\xi^j, \ldots, \sum_{j=1}^n a_j\xi^{n-j}\right)P^{-1}, \end{aligned}$$

queda claro que los valores propios de $C_n(\mathbf{a})$ son

$$\lambda_k = \sum_{j=1}^{n} a_j \xi^{kj} \ ; \ k = 0, 1, \ldots, n-1.$$

Así, **toda matriz circulante es diagonalizable**, y la matriz de paso es la matriz de Fourier.

## 7.3. Polinomios anuladores de una matriz.

Dada $A \in \mathbb{K}^{n\times n}$, diremos que un polinomio $q \in \mathbb{K}[x]$ es un **anulador** de $A$ si $q(A) = 0$ (matriz nula).

Como el espacio $\mathbb{K}^{n\times n}$ tiene dimensión $n^2$, es obvio que, si $k \geq n^2$, las matrices $I, A, A^2, \ldots, A^k$ serán linealmente dependientes cualquiera que sea la matriz $A \in \mathbb{K}^{n\times n}$.

Así, existirá algún polinomio no nulo $q,$ de grado menor o igual que $n^2$, anulador de $A$. (Más adelante veremos que, de hecho, el propio polinomio característico de la matriz, que tiene grado $n$, es un anulador).

### 7.3.1. Propiedades de los polinomios anuladores.

En lo que sigue, suponemos **fijada la matriz** $A$.

El polinomio nulo es un anulador. La suma de anuladores, el producto de anuladores y el producto de un escalar por un anulador, son anuladores. Así, el conjunto de anuladores tiene las estructuras de subanillo y subespacio vectorial de $\mathbb{K}[x]$. (Esto puede expresarse diciendo que es una subálgebra del álgebra $\mathbb{K}[x]$).

Además, **si multiplicamos un anulador por un polinomio cualquiera, obtenemos un anulador** (esto se puede expresar diciendo que el conjunto de anuladores de $A$ es un **ideal** del anillo $\mathbb{K}[x]$).

**Matrices semejantes tienen los mismos polinomios anuladores:** en efecto, si $B = P^{-1}AP$, $\forall q \in \mathbb{K}[x]$, $q(B) = P^{-1}q(A)P$.

## 7.4. El polinomio mínimo.

### 7.4.1. Definición de polinomio mínimo.

Como existen anuladores no nulos de grado menor o igual que $n^2$, existirá entre ellos alguno **de grado mínimo, y mónico.** Éste será el **polinomio mínimo de la matriz** $A$, que denotaremos por $m_A$.

**El polinomio mínimo es único.** En efecto, si existieran dos polinomios mínimos distintos $m_A$ y $m'_A$, ambos tendrían que ser del mismo grado, digamos $r$. Pero entonces, su diferencia $m_A - m'_A$ sería un polinomio anulador de $A$, no nulo y de grado menor que $r$, contradiciendo la hipótesis de que $m_A$ y $m'_A$ fueran polinomios mínimos.

**El polinomio mínimo genera todos los anuladores**, es decir:

$$\text{Si } p(A) = 0, \text{ entonces } p(x) = q(x)m_A(x), \text{ para algún } q \in \mathbb{K}[x].$$

En efecto, puesto que el grado de $p$ ha de ser mayor o igual que el de $m_A$, podemos efectuar la división euclídea $p(x) = q(x)m_A(x) + r(x)$, siendo $\mathrm{gr}(r) < \mathrm{gr}(m_A)$. Pero como $r(x) = p(x) - q(x)m_A(x)$ también es un anulador, no puede ser más que $r(x) \equiv 0$.

### 7.4.2. Los ceros del polinomio mínimo.

**Proposición.- Todo valor propio de $A$ es un cero del polinomio mínimo.**

**Demostración.-** Si $\lambda_0$ es valor propio y $\mathbf{v}_0$ un vector propio asociado a él se tiene que

$$m_A(\lambda_0)\,\mathbf{v}_0 = m_A(A)\mathbf{v}_0 = 0\mathbf{v}_0 = \mathbf{0},$$

de donde se deduce, puesto que $\mathbf{v}_0$ es no nulo, que $m_A(\lambda_0) = 0$.

**Proposición.- Todo cero de $m_A$ es valor propio de $A$.**

**Demostración.-** Como ya sabemos que todos los valores propios $\lambda_1, \ldots, \lambda_p$ de $A$ son ceros de $m_A$, podemos escribir

$$m_A(\lambda) = (\lambda - \lambda_1)^{r_1} \ldots (\lambda - \lambda_p)^{r_p} q(\lambda).$$

(La factorización anterior es posible en $\mathbb{C}[\lambda]$). Supongamos que $m_A$ tiene algún cero $\lambda_0$ que no es valor propio: $m_A(\lambda) = (\lambda - \lambda_1)^{r_1} \ldots (\lambda - \lambda_p)^{r_p}(\lambda - \lambda_0)r(\lambda)$. Puesto que $\lambda_0$ no es valor propio, es $(A - \lambda_0 I)$ invertible, de donde

$$0 = (A - \lambda_0 I)^{-1} m_A(A) = (A - \lambda_1 I)^{r_1} \ldots (A - \lambda_p I)^{r_p} r(A),$$

lo cual significa que $(\lambda - \lambda_1)^{r_1} \ldots (\lambda - \lambda_p)^{r_p} r(\lambda)$ es un anulador de $A$ de grado menor que el de $m_A$, que es una contradicción.

Hemos concluido que **los ceros del polinomio mínimo son los valores propios de la matriz,** pero ¿cuáles son sus multiplicidades? El teorema de Cayley-Hamilton nos da una cota para éstas, al afirmar que, como máximo, son las mismas que en el polinomio característico.

## 7.5. El teorema de Cayley-Hamilton.

**Teorema.-** (de Cayley-Hamilton)
**El polinomio mínimo de $A$ divide al polinomio característico de $A$.**
O, equivalentemente:
**El polinomio característico de $A$ es un anulador de $A$.**

**Demostración.-** Sabemos que, sobre el cuerpo $\mathbb{C}$, toda matriz cuadrada es semejante a una triangular superior. También sabemos que matrices semejantes tienen el mismo polinomio característico y los mismos polinomios anuladores. Por todo ello, basta demostrar el teorema para matrices triangulares superiores.

Para ello, procederemos por inducción en el orden $n$ de la matriz. Para $n = 1$, el resultado es trivial. Supongamos, pues, que es cierto para las matrices triangulares de orden $n$ y sea $T_{n+1}$ una matriz triangular superior de orden $n + 1$, que escribiremos de la forma

$$T_{n+1} = \begin{pmatrix} a & u^t \\ \mathbf{0} & T_n \end{pmatrix},$$

en donde $a \in \mathbb{K}$, $u \in \mathbb{K}^n$ y $T_n$ es una matriz del mismo tipo pero de orden $n$, cuyos elementos diagonales (es decir, sus valores propios) denotamos por $a_1, \ldots, a_n$.

El polinomio característico de una tal matriz puede escribirse como

$$\chi_{T_{n+1}}(\lambda) = (a-\lambda)\chi_{T_n}(\lambda) = (a-\lambda)\prod_{j=1}^{n}(a_j - \lambda).$$

Evaluando el polinomio anterior en la matriz $T_{n+1}$,

$$\chi_{T_{n+1}}(T_{n+1}) = (aI_{n+1} - T_{n+1})\chi_{T_n}(T_{n+1}) = (aI_{n+1} - T_{n+1})\prod_{j=1}^{n}(a_j I_{n+1} - T_{n+1}),$$

que es la expresión que queremos ver que se anula.

La primera matriz de este producto se escribe como

$$aI_{n+1} - T_{n+1} = \begin{pmatrix} 0 & -u^t \\ \mathbf{0} & aI_n - T_n \end{pmatrix}$$

Cada uno de los restantes factores del producto anterior es de la forma

$$a_j I_{n+1} - T_{n+1} = \begin{pmatrix} a_j - a & -u^t \\ \mathbf{0} & a_j I_n - T_n \end{pmatrix},$$

de modo que el producto de todos ellos se puede escribir como

$$\begin{pmatrix} \prod_{j=1}^{k}(a_j - a) & v^t \\ \mathbf{0} & \prod_{j=1}^{k}(a_j I_n - T_n) \end{pmatrix} = \begin{pmatrix} \prod_{j=1}^{k}(a_j - a) & v^t \\ \mathbf{0} & \chi_{T_n}(T_n) \end{pmatrix} = \begin{pmatrix} \prod_{j=1}^{k}(a_j - a) & v^t \\ \mathbf{0} & O \end{pmatrix},$$

en donde se ha aplicado la hipótesis de inducción para obtener la última igualdad.

Por lo tanto, al multiplicar finalmente

$$\begin{pmatrix} 0 & -u^t \\ \mathbf{0} & aI_n - T_n \end{pmatrix} \begin{pmatrix} \prod_{j=1}^{k}(a_j - a) & v^t \\ \mathbf{0} & O \end{pmatrix},$$

obtenemos la matriz nula, como queríamos.

## 7.6. Caracterización de las matrices diagonalizables en términos de los ceros del polinomio mínimo.

**Teorema.-** Sea $A \in \mathbb{C}^{n\times n}$; $A$ **es diagonalizable si y sólo si todos los ceros de** $m_A$ **son simples.**

**Demostración.-** $\boxed{\Rightarrow}$ $A$ tiene el mismo polinomio mínimo que $D = P^{-1}AP$, y obviamente $\prod_{\lambda j \ \in\sigma(A)} (\lambda - \lambda_j)$ es un anulador de $D$.

$\boxed{\Leftarrow}$ Sea $m_A = (\lambda - \lambda_1) \dots (\lambda - \lambda_r)$, con $\lambda_i \neq \lambda_j$ si $i \neq j$. Consideramos los polinomios

$$p_j(\lambda) = \frac{m_A(\lambda)}{\lambda - \lambda_j} \ , \ j = 1, \dots, r.$$

Estos polinomios son primos entre sí $(\text{m.c.d.}\,(p_1, \dots, p_r) = 1)$, de modo que, por la identidad de Bézout generalizado, existen $q_1, \dots, q_n \in \mathbb{K}\,[x]$ tales que

$$\sum_{j=1}^{r} q_j(\lambda) p_j(\lambda) \equiv 1.$$

Evaluando el polinomio anterior en $A$, tenemos que

$$\sum_{j=1}^{r} q_j(A) p_j(A) = I,$$

lo cual quiere decir que, para todo $\mathbf{v} \in \mathbb{K}^n$,

$$\sum_{j=1}^{r} q_j(A) p_j(A) \mathbf{v} = \mathbf{v}.$$

Ahora, llamando $\mathbf{v}_j = q_j(A) p_j(A) \mathbf{v}$, se tiene que

$$(A - \lambda_j I)\mathbf{v}_j = q_j(A)p_j(A)(A - \lambda_j I)\mathbf{v} = q_j(A)m_A(A)\mathbf{v} = \mathbf{0},$$

es decir, que $\mathbf{v}_j \in \ker(A - \lambda_j I)$.

Hemos concluido que todo $\mathbf{v} \in \mathbb{K}^n$ se descompone como

$$\mathbf{v} = \mathbf{v}_1 + \dots + \mathbf{v}_r \ , \quad \text{con } \mathbf{v}_j \in \ker(A - \lambda_j I),$$

con lo que $\mathbb{K}^n = \ker(A - \lambda_1 I) \oplus \dots \oplus \ker(A - \lambda_r I)$, lo que equivale a que $A$ sea diagonalizable.

# 8. Cálculo del polinomio mínimo.

## 8.1. Cálculo a partir del polinomio característico.

Dado el polinomio característico $p_A(\lambda) = \prod_{j=1}^{r} (\lambda - \lambda_j)^{s_j}$, sabemos que el polinomio mínimo será de la forma $m_A(\lambda) = \prod_{j=1}^{r} (\lambda - \lambda_j)^{r_j}$ con $1 \leq r_j \leq s_j$. Por lo tanto, basta con ir probando con todos los polinomios de esta forma, en orden creciente de grado, hasta encontrar uno que sea anulador de $A$.

## 8.2. Cálculo a partir de la forma canónica de Jordan.

Si se conoce la forma canónica de Jordan de $A$, podemos calcular fácilmente su polinomio mínimo, que coincide con el de $A$.

Está claro que el polinomio mínimo de un bloque de Jordan del tipo

$$\begin{pmatrix} \lambda_0 & 1 & & & & \\ & \lambda_0 & 1 & & & \\ & & \lambda_0 & & & \\ & & & \ddots & & \\ & & & & \lambda_0 & 1 \\ & & & & & \lambda_0 \end{pmatrix}$$

es precisamente $(\lambda - \lambda_0)^s$, siendo $s$ el orden del bloque. Si existen varios bloques asociados a un mismo valor propio, y tomamos el polinomio mínimo del mayor de ellos, ese polinomio es anulador de todos ellos y, consecuentemente, el polinomio mínimo de toda la "submatriz asociada a $\lambda_0$"; por ejemplo:

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

tiene como polinomio mínimo $(\lambda - 2)^3$.

Si este procedimiento lo repetimos para las cajas asociadas a cada uno de los valores propios, y multiplicamos todos los polinomios mínimos obtenidos, habremos hallado el polinomio mínimo de $A$.