TEMA 16: FORMAS CANONICAS DE JORDAN

1. MODULOS SOBRE ANILLOS.

DEFINICION: Sea A un anillo conmutativo y unitario. Un A-módulo es un grupo abeliano M con una ley externa definida así

$$\cdot : (\lambda, x) \in A \times M \longmapsto \lambda \cdot x \in M$$

verificando las propiedades:

① $\forall \lambda, \mu \in A, \forall x \in M, \ (\lambda+\mu)x = \lambda \cdot x + \mu \cdot x$

② $\forall \lambda \in A, \forall x, y \in M, \ \lambda \cdot (x+y) = \lambda \cdot x + \lambda \cdot y$

③ $(\lambda \mu) \cdot x = \lambda \cdot (\mu \cdot x)$; $\forall \lambda, \mu \in A, \forall x \in M$

④ $1 \cdot x = x, \ \forall x \in M$

Nota: Si A no fuese conmutativo, ésta sería la definición de A-módulo por la izquierda. Si se sustituye ③ por

③ $\forall \lambda, \mu \in A, \forall x \in M, \ (\lambda \mu) \cdot x = \mu \cdot (\lambda \cdot x)$

se dice que M es un A-módulo por la derecha.

DEFINICION: Un submódulo es un subgrupo de M cerrado respecto a la ley externa.

Ejemplo: Todo grupo abeliano es un $\mathbb{Z}$-módulo con la ley externa $n \cdot x = \overbrace{x + \dots + x}^{n}$.

DEFINICION: (Orden de un elemento de un A-módulo).

Sea M un A-módulo y $x \in M$. Definimos el orden de x como el conjunto de elementos de A que anulan a x, es decir

$$O_x = \{a \in A / a \cdot x = 0\}.$$

Es trivial la siguiente

1.1. PROPOSICION: El orden de un elemento de un A-módulo es un ideal de A.

DEFINICION: Diremos que el orden de un elemento x de un A-módulo es finito si $O_x \neq \{0\}$.

1.2. TEOREMA: Sea A un dominio principal y M un A-módulo de generación finita cuyos elementos son de orden finito. Entonces M es suma directa de submódulos cíclicos, es decir $M = \langle x_1 \rangle \oplus \dots \oplus \langle x_s \rangle$ donde los submódulos pueden tomarse de manera que verifiquen una de las dos condiciones siguientes:

1) $\forall i \in \{1, \dots, s\}$, $\langle x_i \rangle$ es primario, es decir, cada elemento a_i es potencia de un elemento irreducible de A, siendo a_i tal que $O_{x_i} = (a_i)$.

2) Si $O_{x_i} = (a_i)$, entonces, $a_i \mid a_{i+1}$, $i = 1, \dots, s-1$.

Además, dos descomposiciones de M del tipo 1) o del tipo 2) tienen el mismo número de sumandos de cada orden.

La demostración es análoga a la hecha para grupos abelianos en el Tema 6º (Teoremas 2.6, 2.7, 2.13, 2.15).

Observaciones: · Al hablar de submódulos cíclicos de M se entenderá que son de la forma $\langle w_i \rangle = \{a \cdot w_i / a \in A\}$.

· Se puede escribir $O_{x_i} = (a_i)$ pues hemos supuesto que Δ es un dominio principal, es decir, todos sus ideales son principales.

DEFINICION a) Si $\langle x_1 \rangle \oplus \dots \oplus \langle x_s \rangle$ es una descomposición de M del tipo 1) se dice que es una descomposición primaria de M.

b) Si $\langle x_1 \rangle \oplus \dots \oplus \langle x_s \rangle$ es una descomposición de M del tipo 2) se dice que es una descomposición canónica de M.

2. Espacio vectorial "considerado" como K[x]-módulo

Sea V un espacio vectorial sobre un cuerpo K y $T: V \to V$ un endomorfismo de V.

Podemos considerar V como módulo sobre K[x] con la ley externa

$$f(x) \cdot v = f(T)(v) \quad , \forall f(x) \in K[x], \forall v \in V$$

es decir, si $f(x) = a_0 + a_1 x + \dots + a_n x^n$ entonces

$$f(x) \cdot v = (a_0 + a_1 T + \dots + a_n T^n)(v) = a_0 \cdot v + a_1 T(v) + \dots + a_n \cdot T^n(v).$$

Se demuestra que esta ley externa verifica las propiedades de los módulos. (*)

Por tanto, dado un endomorfismo T sobre V, podemos considerar a V como un K[x]-módulo. Cuando nos refiramos a V como K[x]-módulo escribiremos V^T.

El orden de un vector $v \in V$ es

$$O_v = \{f(x) \in K[x] / f(x) \cdot v = 0\}$$

Siendo K[x] un dominio de ideales principales, para cada $v \in V$ existe un polinomio $f_v(x)$, que podemos tomar mónico, tal que $O_v = (f_v(x))$.

DEFINICION: A $f_v(x)$ se le llama polinomio mínimo o anulador de v.

Es claro que $f_v(x)$ es el polinomio mónico de menor grado que anula a v en el sentido de la definición, es decir, que $f_v(x) \cdot v = 0$.

2.1. PROPOSICION: W es un submódulo de V^T si, y solo si, W es un subespacio vectorial invariante por T, es decir, tal que $T(W) \subset W$.

Demostr: $\Rightarrow$ Supongamos que W es submódulo de V^T. Entonces W es subgrupo de V. Falta ver que es cerrado respecto a la ley externa de V y que es invariante por T.
Dado $\lambda \in K$, hacemos $f(x) = \lambda \in K$. Entonces, dado $w \in W$ tenemos que $f(x) \cdot w \in W$, por ser W submódulo de V^T.
Como $f(x) \cdot w = \lambda w$ se deduce que $\lambda w \in W$.
Veamos que W es invariante por T. Dado $w \in W$, haciendo $f(x) = x$ tenemos que $f(x) \cdot w \in W$. Pero $f(x) \cdot w = T(w)$. Luego $T(w) \in W$.
$\Leftarrow$ Puesto que $T(W) \subset W$, es trivial que $T^n(W) \subset W$ cualquiera que sea $n \in \mathbb{N}$. Además $\forall \lambda \in K$, $\lambda \cdot W \subset W$.
Luego, trivialmente, $\forall f(x) \in K[x]$, $f(x) \cdot W \subset W$. c.s.q.d.

2.2. PROPOSICION: Si V es un espacio vectorial de dimensión finita n entonces V^T es un $K[x]$-módulo de generación finita cuyos elementos son de orden finito.

Demostr: Sea $\mathcal{B} = \{v_1, \dots, v_n\}$ una base de V sobre K.
Dado $v \in V$, existen $\lambda_1, \dots, \lambda_n \in K$ tal que $v = \sum_{i=1}^{n} \lambda_i \cdot v_i$. Puesto que $\lambda_i \in K[x]$ se deduce que todo elemento v de V^T está engendrado por $\mathcal{B}$. Luego V^T es de generación finita. (*)
Además, dado $v \in V$, los vectores $v, T(v), \dots, T^n(v)$ son linealmente dependientes, pues $\dim V = n$ y tenemos $n+1$ vectores.
Luego $\exists \{\lambda_i\}_{i=0}^{n} \subset K$ tal que $\sum_{i=0}^{n} \lambda_i T^i(v) = 0$.
Además no todos los λ_i son nulos.
Haciendo $f(x) = \sum_{i=0}^{n} \lambda_i x^i$ tenemos que $f(x) \neq 0$ y $f(x) \cdot v = 0$.
Por tanto, $\forall v \in V$, $0_v \neq (0)$, es decir, todos los elementos de V son de orden finito. c.s.q.d.

Esta proposición nos permite aplicar el teorema 1.2. al $K[x]$-módulo V^T que, como hemos probado, es de generación finita y cuyos elementos son de orden finito, y obtenemos así el siguiente teorema fundamental:

2.3. TEOREMA: Sea V un espacio vectorial de dimensión finita, $T: V \to V$ un endomorfismo y V^T el $K[x]$-módulo asociado. Entonces V es suma directa de submódulos cíclicos e invariantes, es decir

$$V = \langle w_1 \rangle \oplus \dots \oplus \langle w_s \rangle$$

donde $\langle w_i \rangle = \{ f(x) \cdot w_i / f(x) \in K[x] \}$. Además, si llamamos $f_i(x)$ al polinomio mínimo de w_i podemos tomar estos submódulos cíclicos de manera que satisfagan una de las condiciones siguientes:

1) Los $f_i(x)$ son potencia de polinomios irreducibles de $K[x]$.
2) $f_i(x) \mid f_{i+1}(x)$, $i = 1, \dots, s-1$.

Además, dos descomposiciones de V^T del tipo 1) o del tipo 2) tienen el mismo número de sumandos de cada orden.

2.4. PROPOSICION: Sea V un espacio vectorial y W un subespacio vectorial de V de dimensión finita. Si T es un endomorfismo de V y W es invariante por T las proposiciones siguientes son equivalentes:

a) W es un submódulo cíclico de V^T.
b) $\exists w_0 \in W$ y $K > 0$ tal que $\{w_0, Tw_0, \dots, T^{K-1}w_0\}$ es base de W y $T^K w_0$ es linealmente dependiente con los $T^i w_0$ con $i < K$.

Demostr: $\Rightarrow$ Supongamos que W es un submódulo cíclico de V^T. Entonces existe $w_0 \in W$ tal que $W = \langle w_0 \rangle$. Sea $f_{w_0}(x)$ el polinomio mínimo de w_0. Supongamos que

$$f_{w_0}(x) = a_0 + a_1 x + a_2 x^2 + \dots + a_K x^K, \quad a_K \neq 0.$$

Entonces el sistema $S = \{w_0, Tw_0, \dots, T^{k-1}w_0\}$ es libre, pues si existiesen $\lambda_0, \dots, \lambda_{k-1}$ no todos nulos tales que $\sum_{i=0}^{k-1} \lambda_i T^i w_0 = 0$, considerando el polinomio no nulo $g(x) = \sum_{i=0}^{k-1} \lambda_i x^i$ tendríamos que $g(x) \cdot w_0 = 0$. Siendo $\partial^\circ(g(x)) < \partial^\circ(f_{w_0}(x))$ y $g(x) \neq 0$ llegamos a una contradicción. Luego S es libre.

Veamos que también es sistema de generadores.

Dado $w \in W$, $\exists g(x) \in K[x] / w = g(x) \cdot w_0$, pues $W = \langle w_0 \rangle$.

Dado $g(x)$, existen $q(x), r(x) \in K[x]$ tal que

$$g(x) = f_{w_0}(x) \cdot q(x) + r(x) \quad \text{con } \partial^\circ(r(x)) < \partial^\circ(f_{w_0}(x)).$$

Entonces $g(x) \cdot w_0 = r(x) \cdot w_0$, pues $f_{w_0}(x) \cdot q(x) \in O_{w_0}$.

Entonces $w = g(x) \cdot w_0 = \sum_{i=0}^{l} \lambda_i T^i w_0$ y $l < K$ pues $\partial^\circ(r(x)) < \partial^\circ(f_{w_0}(x))$.

Que prueba que S es un sistema de generadores de W.

$\Leftarrow$ Supongamos que existen $w_0 \in W$ y $K>0$ tales que
$\{w_0, Tw_0, \dots, T^{K-1}w_0\}$ es base de W
y $T^K w_0 \in W$.
Veamos que $W = \langle w_0 \rangle$.
Que $W \subset \langle w_0 \rangle$ es trivial, pues si $v \in W$, existen $\lambda_0, \dots, \lambda_{K-1} \in K$ tales que $w = \sum_{i=0}^{K-1} \lambda_i T^i w_0 = (\sum_{i=0}^{K-1} \lambda_i x^i) \cdot w_0 \in \langle w_0 \rangle$.

Probemos que $\langle w_0 \rangle \subset W$. Dado $g(x) \cdot w_0 \in \langle w_0 \rangle$, sean $q(x), r(x) \in K[x]$, tales que $g(x) = f_{w_0}(x) q(x) + r(x)$ con $\partial^\circ(r(x)) < \partial^\circ(f_{w_0}(x))$. (I)
Veamos que $\partial^\circ(f_{w_0}(x)) = K$.
Como $T^K w_0 \in W$, $\exists \{\mu_i\}_{i=0}^{K-1} \subset K$ / $T^K w_0 = \sum_{i=0}^{K-1} \mu_i T^i w_0$.
Entonces $T^K w_0 - \sum_{i=0}^{K-1} \mu_i T^i w_0 = 0$.
Sea $x^K - \sum_{i=0}^{K-1} \mu_i x^i \in K[x]$. Entonces $(x^K - \sum_{i=0}^{K-1} \mu_i x^i) w_0 = 0$.
Por tanto existe un polinomio de grado K que anula a w_0.
Además, ningún polinomio no nulo de grado menor que K anula a w_0, pues si $(a_0 + a_1 x + \dots + a_l x^l) w_0 = 0$ con $l < K$ se tendría que $a_0 w_0 + a_1 T w_0 + \dots + a_l T^l w_0 = 0$ con algún $a_i \neq 0$ lo cual contradice que $\{w_0, Tw_0, \dots, T^{K-1}w_0\}$ es base de W. Entonces $f_{w_0}(x) = x^K - \sum_{i=0}^{K-1} \mu_i x^i$ y, por tanto, $\partial^\circ(f_{w_0}(x)) = K$.
A la vista de (I) tenemos fue $g(x) \cdot w_0 = r(x) \cdot w_0$ y $\partial^\circ(r(x)) < K$.
Luego $g(x) \cdot w_0 \in W$. c.s.q.d.

Observación: Trivialmente se prueba a partir de b) que W es invariante. Por tanto, podemos decir que:
(W es un submódulo cíclico e invariante de V^T) $\Longleftrightarrow$ ($\exists w_0 \in W$ y $K>0$ tales que $\{w_0, Tw_0, \dots, T^{K-1}w_0\}$ es base de W y $T^K w_0 \in W$).

3. Diagonalización por bloques de una matriz.

* Sea V un espacio vectorial de dimensión finita. Entonces (teorema 7.3), V admite una descomposición como suma directa de submódulos cíclicos e invariantes

$$V = W_1 \oplus \dots \oplus W_s \qquad (I)$$

Sea W uno de estos submódulos. Entonces existen $w_0 \in W$ y $K>0$ tales que $\{w_0, Tw_0, \dots, T^{K-1}w_0\}$ es base de W y $T^K w_0 \in W$.
Puesto que W es invariante por T, la restricción de T a W es un endomorfismo de W. Sabemos fue el grado del polinomio mínimo de w_0 es K. Sea $f_{w_0}(x) = b_0 + b_1 x + \dots + b_{K-1} x^{K-1} + x^K$

Entonces $T^K w_0 = -b_0 w_0 - b_1 T w_0 - \dots - b_{K-1} T^{K-1} w_0$.

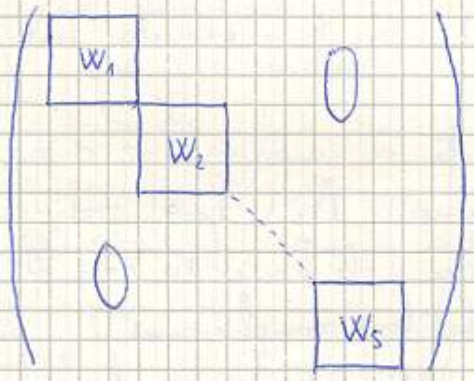
El endomorfismo $T_{|W}$ está representado por una matriz, una vez fijada la base $\{w_0, Tw_0, \dots, T^{K-1}w_0\}$ de W, cuyas columnas están formadas por los coeficientes de los transformados por T de los vectores de la base. Si hacemos $e_i = T^i w_0$, $i = 0, \dots, K-1$ tenemos

$$Te_0 = Tw_0 = 0e_0 + 1e_1 + 0e_2 + \dots + 0e_{K-1}$$
$$Te_1 = T^2 w_0 = 0e_0 + 0e_1 + 1e_2 + \dots + 0e_{K-1}$$
$$\dots\dots\dots\dots\dots\dots\dots\dots$$
$$Te_{K-2} = T^{K-1} w_0 = 0e_0 + 0e_1 + 0e_2 + \dots + 1e_{K-1}$$
$$Te_{K-1} = T^K w_0 = -b_0 e_0 - b_1 e_1 - b_2 e_2 - \dots - b_{K-1} e_{K-1}$$

Luego $T_{|W}$ viene representado por la matriz

$$\begin{pmatrix} 0 & 0 & \dots & 0 & -b_0 \\ 1 & 0 & \dots & 0 & -b_1 \\ 0 & 1 & \dots & 0 & -b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & 0 & -b_{K-2} \\ 0 & 0 & \dots & 1 & -b_{K-1} \end{pmatrix} \qquad (II)$$

Teniendo en cuenta (I), si $\mathcal{B}_i = \{w_i, Tw_i, \dots, T^{K_i-1} w_i\}$ es base de W_i, $\mathcal{B} = \bigcup_{i=1}^{s} \mathcal{B}_i$ es base de V, y en esta base la matriz de T es de la forma



(III)

matriz que se llama diagonal por bloques, y donde cada bloque es una matriz de la forma (II).

3.1. PROPOSICION: Sea W uno de los sumandos cíclicos e invariantes de V, $W = \langle w_0 \rangle$ y $f(x) = x^K + b_{K-1}x^{K-1} + \dots + b_1 x + b_0$ el polinomio mínimo de w_0. Sea C la matriz de $T_{|W}$ (II). Entonces, el polinomio característico de C es

$$\Delta(\lambda) = \det(C - \lambda I) = f(\lambda) \quad \text{(salvo el signo)}$$

1ª Demostr: Por el teorema de Cayley-Hamilton (Geometría I), $\Delta(T) = 0$.

Luego $\Delta(T)(w_0) = 0$ y, por tanto, $\Delta(x) \cdot w_0 = 0$.
Entonces $\Delta(x)$ es múltiplo del polinomio mínimo de w_0.
Como $\partial^\circ(\Delta(x)) = K$ (= orden de C) y $\partial^\circ(f(x)) = K$ debe ser $\Delta(x) = f(x)$, salvo quizás el signo.

2ª Demostr.: Queremos calcular

$$\Delta_k(\lambda) = \begin{vmatrix} -\lambda & 0 & 0 & \cdots & 0 & -b_0 \\ 1 & -\lambda & 0 & \cdots & 0 & -b_1 \\ 0 & 1 & -\lambda & \cdots & 0 & -b_2 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & -\lambda & -b_{k-2} \\ 0 & 0 & 0 & \cdots & 1 & -b_{k-1}-\lambda \end{vmatrix}$$

Desarrollando por los elementos de la primera fila tenemos

$$\Delta_k(\lambda) = (-\lambda)\Delta_{k-1}(\lambda) - (-1)^{k+1} b_0 = (-\lambda)\Delta_{k-1}(\lambda) + (-1)^k b_0 =$$
$$= (-\lambda)^2 \Delta_{k-2}(\lambda) + (-1)^k (\lambda b_1 + b_0) = \cdots =$$
$$= (-\lambda)^{k-1} \Delta_1(\lambda) + (-1)^k (\lambda^{k-2} b_{k-2} + \cdots + \lambda b_1 + b_0) =$$
$$= (-\lambda)^{k-1} [-b_{k-1} - \lambda] + (-1)^k (\lambda^{k-2} b_{k-2} + \cdots + \lambda b_1 + b_0) =$$
$$= (-1)^k (\lambda^k + \lambda^{k-1} b_{k-1} + \lambda^{k-2} b_{k-2} + \cdots + \lambda b_1 + b_0) = (-1)^k f(\lambda). \text{ c.q.d.}$$

* Sea A una matriz de orden n sobre un cuerpo K. A puede considerarse como la matriz de un endomorfismo T de K^n, una vez prefijada una base $\mathcal{B} = \{e_i\}_{i=1}^n$ en K^n, siendo $T(e_i) = C_i$ (columna i-ésima de A).

Esta matriz es semejante a una matriz diagonal por bloques de la forma (III). Si llamamos $\Delta_i(\lambda)$ al polinomio característico del bloque i-ésimo se verifica, de acuerdo con el teorema 2.3 y la proposición anterior, una de las condiciones

1) $\Delta_i(\lambda)$ es potencia de un polinomio irreducible de $K[\lambda]$.
2) $\Delta_i(\lambda) \mid \Delta_{i+1}(\lambda)$, $i = 1, \ldots, s-1$.

Si se verifica la condición 2), de la matriz obtenida se dice que es una forma canónica racional de A.

Dado el espacio vectorial $V = K^n$ y el endomorfismo T, V admite una descomposición como suma directa de submódulos cíclicos e invariantes: $V = W_1 \oplus \cdots \oplus W_s$.

Sea $W_i = \langle w_i \rangle$. Según el teorema 2.3, los órdenes de los elementos $w_1, \ldots, w_s$ están determinados de manera única, es

tiene que $O_{W_i} = (f_{W_i}(\lambda))$. Como $f_{W_i}(\lambda) = \pm \Delta_i(\lambda)$, por la proposición 3.1, se tiene que $O_{W_i} = (\Delta_i(\lambda))$.

Si observamos que el endomorfismo T definido por la matriz Δ depende de la base $\mathcal{B} = \{e_i\}_{i=1}^n$ elegida, notaremos que para otra base $\mathcal{B}' = \{e_i'\}_{i=1}^n$ de $V = K^n$ obtendríamos otro endomorfismo T' y otro módulo $V^{T'}$. Podemos preguntarnos si al considerar la base $\mathcal{B}'$ los polinomios característicos de los bloques $(\Delta_i(\lambda))$ cambian. La respuesta es no, como probaremos después.

Si probamos que V^T y $V^{T'}$ son módulos isomorfos, dos descomposiciones de los mismos como suma directa de submódulos cíclicos e invariantes tendrán el mismo número de sumandos de cada orden y, puesto que los órdenes son ideales generados por los polinomios característicos, se deduciría que dichos polinomios característicos no varían cualquiera que sea la base considerada. Basta ver entonces que V^T y $V^{T'}$ son isomorfos.

Si P es la matriz de cambio de base de $\mathcal{B}$ a $\mathcal{B}'$, y Δ' es la matriz asociada a T' en $\mathcal{B}$ tenemos que $\Delta' = P\Delta P^{-1}$ pues Δ es la matriz de T' en $\mathcal{B}'$. Como Δ es la matriz de T en $\mathcal{B}$ se deduce que las matrices de T y T' en $\mathcal{B}$ son semejantes. Si probamos que (a) $\Rightarrow$ (b) donde

(a) $\equiv$ Δ y Δ' son semejantes

(b) $\equiv$ V^T y $V^{T'}$ son isomorfos

quedaría probado que los polinomios característicos de los bloques no varían, como fueremos ver. Es más, probaremos que (a) $\Leftrightarrow$ (b); pero antes veamos un lema preliminar.

3.2. Lema: Sea V un espacio vectorial de dimensión finita y T un endomorfismo de V. Sea W otro espacio vectorial de dimensión finita y S un endomorfismo de W. Sean V^T y W^S los $K[x]$-módulos asociados. Entonces una aplicación $f: V^T \to W^S$ es un $K[x]$-homomorfismo si, y sólo si, se satisfacen las condiciones

i) f es homomorfismo de espacios vectoriales.

ii) $f(T(v)) = S(f(v))$, $\forall v \in V$.

Demostr: $\Rightarrow$] Supongamos que $f: V^T \to W^S$ es un $K[x]$-homomorfismo. Puesto que $K \subset K[x]$ y f es $K[x]$-homomorfismo se tiene

i) $\forall \lambda \in K$, $f(\lambda v) = \lambda f(v)$, $\forall v \in V = V^T$ (esta es una igualdad de espacios);

es "la misma" para los $K[x]$-homomorfismos entre los módulos V^T y W^S que para los homomorfismos de los espacios vectoriales V y W).

ii) Dado $x \in K[x]$, $\forall v \in V^T$, $f(x \cdot v) = x \cdot f(v)$.

Pero $x \cdot v = T(v)$ y $x \cdot f(v) = S(f(v))$. Luego $f(T(v)) = S(f(v))$, $\forall v \in V^T$.

$\Leftarrow$ Por i), $\forall \lambda \in K$, $\forall v \in V$, $f(\lambda v) = \lambda f(v)$.

Por ii), $\forall v \in V^T$, $f(x \cdot v) = x \cdot f(v)$

Por inducción, $f(x^n \cdot v) = x^n f(v)$.

Luego, $\forall \varphi(x) \in K[x]$, $f(\varphi(x) \cdot v) = \varphi(x) \cdot f(v)$. c.s.q.d.

3.3. PROPOSICION: Sea $V = K^n$ y $\mathcal{B}$ una base de V. Sean T y S endomorfismos de V y A y B las matrices de orden n asociadas a T y S en $\mathcal{B}$. Entonces, A y B son semejantes si, y solo si, los $K[x]$-modulos asociados V^T y V^S son $K[x]$-isomorfos.

Demostr.: $\Rightarrow$ Si A y B son semejantes, existe una matriz $P \in M_n$ con $\det P \neq 0$ tal que $B = PAP^{-1}$.

En la base $\mathcal{B}$ de V podemos considerar P como la matriz de un isomorfismo $f: V \to V$. Veamos que f induce un $K[x]$-isomorfismo entre V^T y V^S. Puesto que f es biyectiva, falta ver que f satisface la proposición ii) del lema, es decir, que $\forall v \in V$, $f(T(v)) = S(f(v))$, o bien que $f \circ T = S \circ f$. Basta ver, en términos de matrices, que $P \cdot A = B \cdot P$. Pero esto es cierto por hipótesis. Luego $f: V^T \to V^S$ es un $K[x]$-isomorfismo.

$\Leftarrow$ Supongamos que $f: V^T \to V^S$ es un $K[x]$-isomorfismo. Queremos ver que A y B son semejantes.

Podemos considerar f como un isomorfismo de V en V. Sea P la matriz de f en $\mathcal{B}$; entonces P es no singular. Además, en virtud del lema $f \circ T = S \circ f$. Luego $P \cdot A = B \cdot P$, que prueba que A y B son semejantes. c.s.q.d.

Visto esto, tenemos que dada una matriz A de orden n y dos bases $\mathcal{B}$ y $\mathcal{B}'$ de $V = K^n$, A es la matriz de dos endomorfismos T y T' en $\mathcal{B}$ y $\mathcal{B}'$, respectivamente; endomorfismos que en la base $\mathcal{B}$ vienen dados por matrices semejantes y, por tanto, V^T y $V^{T'}$ son $K[x]$-isomorfos. Luego, como se indicó anteriormente, los polinomios característicos Δ_i de los bloques quedan determinados de manera única. Pasos: Entonces:

DEFINICIONES: Sea A una matriz de orden n. Los polinomios característicos $\Delta_i(\lambda)$ de los bloques que son potencias de polinomios irreducibles de $K[x]$ se llaman divisores elementales de A. Si estos polinomios característicos son del tipo 2) $(\Delta_i(\lambda) | \Delta_{i+1}(\lambda))$ se llamarán factores invariantes de A.

3.4. COROLARIO: Dos matrices A y B son semejantes si, y solo si, tienen los mismos divisores elementales (resp. los mismos factores invariantes).

Demostr.: $\Rightarrow$] Sean T y S los endomorfismos de V determinados por A y B en una determinada base. Los divisores elementales (resp. factores invariantes) quedan univocamente determinados por los módulos V^T y V^S, los cuales son isomorfos por ser A y B semejantes. Por tanto, A y B tienen los mismos divisores elementales (resp. factores invariantes).

$\Leftarrow$] Los divisores elementales (resp. factores invariantes) de A y B determinan matrices diagonales por bloques $\tilde{A}$ y $\tilde{B}$ semejantes a A y a B, respectivamente. Puesto que los divisores elementales de A y B son los mismos se tiene que $\tilde{A} \sim \tilde{B}$. Luego, como $A \sim \tilde{A}$ y $B \sim \tilde{B}$ se tiene que A y B son semejantes: $A \sim B$. c.q.d.

NOTA: Observar que, siendo los factores invariantes de una matriz únicos y no pudiendose permutar, una matriz dada solo tiene asociada una forma canónica racional. Sin embargo, una matriz dada es semejante a varias matrices diagonales por bloques obtenidas a partir de los divisores elementales, pues éstos sí se pueden permutar.

4 Polinomio mínimo de una matriz. Teorema de Cayley-Hamilton.

DEFINICION: Sea A una matriz cuadrada de orden n sobre un cuerpo K. Definimos el polinomio mínimo de A, $m(x) = \text{Min}(A, K)$, como el polinomio mónico de menor grado tal que $m(A) = 0$. (*)

Vamos a probar la existencia del polinomio mínimo de una matriz dada.

Sea T un endomorfismo de $V = K^n$ asociado a A en

Entonces $u(A)$ es la matriz nula de orden n si, y solo si, $u(T)$ es el endomorfismo nulo de V. Sea

$$V = W_1 \oplus \dots \oplus W_s \qquad (I)$$

la descomposición de V como suma directa de submódulos cíclicos e invariantes $W_i = \langle w_i \rangle$, y sea $f_i(x)$ el polinomio mínimo de w_i, verificándose $f_i(x) \mid f_{i+1}(x)$, $i = 1, \dots, s-1$.

Probemos que $u(x) = f_s(x)$.

Veamos en primer lugar que $f_s(T)$ es el endomorfismo nulo de V, es decir, que $f_s(T)\cdot v = 0$, $\forall v \in V$.

De acuerdo con (I), una base de V será

$$\mathcal{B} = \{w_1, Tw_1, \dots, w_2, Tw_2, \dots, w_s, Tw_s, \dots\}.$$

Basta ver entonces que $f_s(T)(T^r w_i) = 0$, $r \geq 0$.

Por definición, $f_s(T)\cdot w_s = 0$ y, siendo $f_s(T)$ multiplo de $f_i(T)$ si $i \leq s$, se tiene que $f_s(T)\cdot w_i = 0$, $1 \leq i \leq s$.

Además, $f_s(T)(T^r w_i) = T^r(f_s(T)\cdot w_i) = T^r(0) = 0$

Luego $f_s(T) = 0$.

Además, $f_s(T)$ es el polinomio de menor grado que anula todos los vectores de V, pues es el polinomio mínimo de $w_s \in V$. Luego $f_s(x) = u(x)$.

De lo anterior sale como consecuencia el teorema de Cayley-Hamilton: "Toda matriz "satisface" su polinomio característico".

Dada una matriz A su polinomio característico es $\Delta(\lambda) = |A - \lambda I|$.

Puesto que dos matrices semejantes tienen el mismo polinomio característico, sea $\tilde{A}$ la forma canonica racional asociada a A, es decir, $\tilde{A}$ es una matriz diagonal por bloques semejante a A y tal que si $\Delta_1(\lambda), \dots, \Delta_s(\lambda)$ son los polinomios característicos de los bloques entonces $\Delta_i(\lambda) \mid \Delta_{i+1}(\lambda)$, $i = 1, \dots, s-1$.

Por el desarrollo de Laplace de un determinante por bloques tenemos que $|\tilde{A} - \lambda I| = \Delta(\lambda) = \Delta_1(\lambda)\cdot\Delta_2(\lambda)\cdot\dots\cdot\Delta_s(\lambda)$.

Siendo $\Delta_s(A) = 0$, se tiene que $\Delta(A) = 0$, como queríamos ver.

5. Formula de Binet-Cauchy

Vamos a utilizar la siguiente notación: Si C es una matriz de orden n, denotamos

$$C\begin{pmatrix} 1 & \dots & n \\ 1 & \dots & n \end{pmatrix} = \det \begin{pmatrix} c_{11} & \dots & c_{1n} \\ \vdots & & \vdots \\ c_{n1} & \dots & c_{nn} \end{pmatrix}$$

Con esta notación, el menor de orden p de la matriz C formado por las filas $i_1,\dots,i_p$ y las columnas $j_1,\dots,j_p$ es

$$C\begin{pmatrix} i_1 \dots i_p \\ j_1 \dots j_p \end{pmatrix} = \det \begin{pmatrix} c_{i_1 j_1} \dots c_{i_1 j_p} \\ \dots \\ c_{i_p j_1} \dots c_{i_p j_p} \end{pmatrix}$$

5.1. PROPOSICION: (fórmula de Binet-Cauchy)

Sean $A=(a_{ik})$ y $B=(b_{kj})$ matrices de ordenes $m\times n$ y $n\times m$, respectivamente, con $n \geq m$. Sea $C=(c_{ij}) \in \mathcal{M}_m$ la matriz $C = A\times B$, y por tanto, $c_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj}$. Entonces

$$C\begin{pmatrix} 1\,2 \dots m \\ 1\,2 \dots m \end{pmatrix} = \sum_{1\leq k_1 < \dots < k_m \leq n} A\begin{pmatrix} 1\,2 \dots m \\ k_1 k_2 \dots k_m \end{pmatrix} \cdot B\begin{pmatrix} k_1 k_2 \dots k_m \\ 1\,2 \dots m \end{pmatrix}$$

Si fuese $n<m$, entonces $C\begin{pmatrix} 1\,2 \dots m \\ 1\,2 \dots m \end{pmatrix} = 0$.

Demostr: Se tiene por definición de $A\times B$ que

$$C\begin{pmatrix} 1\,2 \dots m \\ 1\,2 \dots m \end{pmatrix} = \det \begin{pmatrix} \sum_{\alpha_1=1}^{n} a_{1\alpha_1} b_{\alpha_1 1} & \sum_{\alpha_2=1}^{n} a_{1\alpha_2} b_{\alpha_2 2} & \dots & \sum_{\alpha_m=1}^{n} a_{1\alpha_m} b_{\alpha_m m} \\ \sum_{\alpha_1=1}^{n} a_{2\alpha_1} b_{\alpha_1 1} & \sum_{\alpha_2=1}^{n} a_{2\alpha_2} b_{\alpha_2 2} & \dots & \sum_{\alpha_m=1}^{n} a_{2\alpha_m} b_{\alpha_m m} \\ \dots & \dots & \dots & \dots \\ \sum_{\alpha_1=1}^{n} a_{m\alpha_1} b_{\alpha_1 1} & \sum_{\alpha_2=1}^{n} a_{m\alpha_2} b_{\alpha_2 2} & \dots & \sum_{\alpha_m=1}^{n} a_{m\alpha_m} b_{\alpha_m m} \end{pmatrix} =$$

$$= \sum_{1\leq \alpha_1,\alpha_2,\dots,\alpha_m \leq n} \det \begin{pmatrix} a_{1\alpha_1} b_{\alpha_1 1} & a_{1\alpha_2} b_{\alpha_2 2} & \dots & a_{1\alpha_m} b_{\alpha_m m} \\ a_{2\alpha_1} b_{\alpha_1 1} & a_{2\alpha_2} b_{\alpha_2 2} & \dots & a_{2\alpha_m} b_{\alpha_m m} \\ \dots & \dots & \dots & \dots \\ a_{m\alpha_1} b_{\alpha_1 1} & a_{m\alpha_2} b_{\alpha_2 2} & \dots & a_{m\alpha_m} b_{\alpha_m m} \end{pmatrix} =$$

$$= \sum_{1\leq \alpha_1,\alpha_2,\dots,\alpha_m \leq n} A\begin{pmatrix} 1\,2 \dots m \\ \alpha_1 \alpha_2 \dots \alpha_m \end{pmatrix} \cdot b_{\alpha_1 1} \cdot b_{\alpha_2 2} \cdot \dots \cdot b_{\alpha_m m}$$

Si $n<m$, para cada colección de índices $(\alpha_1,\dots,\alpha_m)$ deben existir $i,j \in \{1,\dots,m\}$ tales que $\alpha_i = \alpha_j$, y por tanto, en el determinante correspondiente $A\begin{pmatrix} 1\,2 \dots m \\ \alpha_1 \alpha_2 \dots \alpha_m \end{pmatrix}$ hay dos columnas iguales, que por tanto será nulo. Luego $C\begin{pmatrix} 1\,2 \dots m \\ 1\,2 \dots m \end{pmatrix} = 0$ si $n<m$.

Supongamos $n \geq m$. Como $\alpha_1,\dots,\alpha_m$ varían independientemente entre 1 y n, van a aparecer sumandos con dos α iguales y estos los podemos suprimir, es decir:

$$C\begin{pmatrix} 1\,2 \dots m \\ 1\,2 \dots m \end{pmatrix} = \sum_{1\leq \alpha_1 \neq \dots \neq \alpha_m \leq n} A\begin{pmatrix} 1 \dots m \\ \alpha_1 \dots \alpha_m \end{pmatrix} b_{\alpha_1 1} \dots b_{\alpha_m m} \quad . \quad (I)$$

Sea $(\alpha_1,\dots,\alpha_m)$ una colección de índices tales que $\alpha_i \neq \alpha_j$ si $i \neq j$. Para esta colección de índices tenemos $m!$ sumandos en (I). Vamos a calcular cuanto vale la suma de estos $m!$ términos. Supongamos que $k_1 < \dots < k_m$ son los índices $\alpha_1,\dots,\alpha_m$ ordenados.

Si σ es la permutación que lleva k_i en α_i, $i=1,\dots,m$, tenemos que $A\begin{pmatrix}1&2&\dots&m\\ \alpha_1&\alpha_2&\dots&\alpha_m\end{pmatrix} = \pi(\sigma)\cdot A\begin{pmatrix}1&2&\dots&m\\ k_1&k_2&\dots&k_m\end{pmatrix}$

donde $\pi(\sigma)$ es la signatura de σ.

Entonces la suma de los $m!$ sumados es

$$\sum_{\sigma\in S_m} \pi(\sigma)\, A\begin{pmatrix}1&2&\dots&m\\ k_1&k_2&\dots&k_m\end{pmatrix} b_{\sigma(k_1)1}\cdot\dots\cdot b_{\sigma(k_m)m} =$$

$$= A\begin{pmatrix}1&2&\dots&m\\ k_1&k_2&\dots&k_m\end{pmatrix} \sum_{\sigma\in S_m} \pi(\sigma) b_{\sigma(k_1)1}\cdot\dots\cdot b_{\sigma(k_m)m} = A\begin{pmatrix}1&2&\dots&m\\ k_1&k_2&\dots&k_m\end{pmatrix}\cdot B\begin{pmatrix}k_1&k_2&\dots&k_m\\ 1&2&\dots&m\end{pmatrix}$$

Por tanto $C\begin{pmatrix}1&2&\dots&m\\ 1&2&\dots&m\end{pmatrix} = \sum_{1\le k_1<\dots<k_m\le n} A\begin{pmatrix}1&2&\dots&m\\ k_1&k_2&\dots&k_m\end{pmatrix}\cdot B\begin{pmatrix}k_1&k_2&\dots&k_m\\ 1&2&\dots&m\end{pmatrix}$. c.s.q.d.

Veamos una aplicación de la fórmula de Binet-Cauchy. Se trata de expresar los menores de orden p de una matriz $C=A\times B$ en función de los menores de orden p de las matrices A y B. Sea $A=(a_{ik})\in \mathcal{M}_{m\times n}$ y $B=(b_{kj})\in\mathcal{M}_{n\times q}$ y $C=A\times B$. Sea $p\le \min(m,q)$. Consideremos el menor de orden p

$$C\begin{pmatrix}i_1&\dots&i_p\\ j_1&\dots&j_p\end{pmatrix}$$

Por la fórmula de Binet-Cauchy tenemos:

$$C\begin{pmatrix}i_1&\dots&i_p\\ j_1&\dots&j_p\end{pmatrix} = \sum_{1\le k_1<\dots<k_p\le n} A\begin{pmatrix}i_1&\dots&i_p\\ k_1&\dots&k_p\end{pmatrix}\cdot B\begin{pmatrix}k_1&\dots&k_p\\ j_1&\dots&j_p\end{pmatrix}$$

si $p\le n$. Si $p>n$, el menor es nulo.

6. Polinomios invariantes de una matriz.

Sea A una matriz cuadrada de orden n. Llamaremos

$$D_n(\lambda) = \det(A-\lambda I)$$

y si $p\le n$ llamamos $D_p(\lambda)$ al máximo común divisor de los menores de orden p de la matriz $A-\lambda I$.

Probemos que $D_p(\lambda) \mid D_{p+1}(\lambda)$, $1\le p\le n-1$.

Basta ver que $D_p(\lambda)$ divide a todo menor de orden $p+1$.

Sea $M_{p+1}(\lambda)$ un menor de orden $p+1$ de la matriz $A-\lambda I$.

Desarrollando por los elementos de una fila de este menor podemos escribir

$$M_{p+1}(\lambda) = \sum_{i=1}^{p+1} a_i M_p^i(\lambda),$$ donde $M_p^i(\lambda)$ son menores de orden p. Como $D_p(\lambda)\mid M_p^i(\lambda)$, $\forall i$, se tiene que $D_p(\lambda)\mid M_{p+1}(\lambda)$.

DEFINICION: Llamamos polinomios invariantes de una matriz A de orden n a los siguientes

$$i_1(\lambda)=\frac{D_n(\lambda)}{D_{n-1}(\lambda)},\ \dots,\ i_p(\lambda)=\frac{D_{n-p+1}(\lambda)}{D_{n-p}(\lambda)},\ \dots,\ i_n(\lambda)=\frac{D_1(\lambda)}{D_0(\lambda)}$$

con el convenio $D_0(\lambda)=1$.

Probaremos después que los polinomios invariantes coinciden con los factores invariantes. Veamos primero un lema preliminar.

6.1. LEMA: Dos matrices semejantes tienen los mismos polinomios invariantes.

Demostr: Sean A y $\hat{A}$ dos matrices semejantes. Entonces existe una matriz P no singular tal que $\hat{A}=PAP^{-1}$

Si denotamos por $D_p(\lambda)$ y $\hat{D}_p(\lambda)$ los polinomios (definidos al principio de este apartado) para A y $\hat{A}$ respectivamente, tenemos

$$\hat{D}_n(\lambda)=|\hat{A}-\lambda I|=|PAP^{-1}-\lambda I|=|PAP^{-1}-\lambda PIP^{-1}|=$$
$$=|P(A-\lambda I)P^{-1}|=|A-\lambda I|=D_n(\lambda).$$

Veamos ahora, en general, $D_p(\lambda)=\hat{D}_p(\lambda)$.

Un menor de orden p cualquiera de la matriz $\hat{A}_\lambda=\hat{A}-\lambda I$ es de la forma

$$\hat{A}_\lambda\begin{pmatrix} i_1 & i_2 & \dots & i_p \\ j_1 & j_2 & \dots & j_p \end{pmatrix}$$

Si hacemos $A_\lambda=A-\lambda I$ tenemos que $\hat{A}_\lambda=PA_\lambda P^{-1}$. Entonces por la fórmula de Binet-Cauchy tenemos

$$\hat{A}_\lambda\begin{pmatrix} i_1 & \dots & i_p \\ j_1 & \dots & j_p \end{pmatrix}=\sum_{\alpha_i} P\begin{pmatrix} i_1 & \dots & i_p \\ \alpha_1 & \dots & \alpha_p \end{pmatrix}\cdot(A_\lambda P^{-1})\begin{pmatrix} \alpha_1 & \dots & \alpha_p \\ j_1 & \dots & j_p \end{pmatrix}=$$

$$=\sum_{\alpha_i,\beta_i} P\begin{pmatrix} i_1 & \dots & i_p \\ \alpha_1 & \dots & \alpha_p \end{pmatrix}\cdot A_\lambda\begin{pmatrix} \alpha_1 & \dots & \alpha_p \\ \beta_1 & \dots & \beta_p \end{pmatrix}\cdot P^{-1}\begin{pmatrix} \beta_1 & \dots & \beta_p \\ j_1 & \dots & j_p \end{pmatrix}. \quad (I)$$

Como $D_p(\lambda)$ divide a todos los menores de orden p de A_λ se tiene que $D_p(\lambda)$ divide a todos los sumandos de (I). Luego $D_p(\lambda)$ divide a cualquier menor de orden p de $\hat{A}_\lambda$ y, por tanto, $D_p(\lambda)\mid\hat{D}_p(\lambda)$.

Analogamente se prueba que $\hat{D}_p(\lambda)\mid D_p(\lambda)$. Luego, y como queríamos probar, $D_p(\lambda)=\hat{D}_p(\lambda)$. ■ (*)

6.2 TEOREMA: Los polinomios invariantes de una matriz A coinciden con los factores invariantes de la misma.

Demostr.: Sea $\hat{A}$ la matriz diagonal por bloques semejante a A. Siempre podemos reordenar estos bloques de manera fija, si llamamos $\Delta_1(\lambda), \dots, \Delta_s(\lambda)$ a los determinantes de los bloques correspondientes de la matriz $\hat{A}-\lambda I$, tengamos $\Delta_p(\lambda) \mid \Delta_{p-1}(\lambda)$, $p=2,\dots,s$. (I)

Probaremos que $i_1(\lambda)=\Delta_1(\lambda), i_2(\lambda)=\Delta_2(\lambda), \dots, i_s(\lambda)=\Delta_s(\lambda)$, $i_{s+1}(\lambda)=1, \dots, i_n(\lambda)=1$, siendo n el orden de la matriz cuadrada A.

Sabemos que $D_n(\lambda)=\Delta_1(\lambda)\cdot\dots\cdot\Delta_s(\lambda)$. Para probar que $i_1(\lambda)=\Delta_1(\lambda)$ basta ver entonces que $D_{n-1}(\lambda)=\Delta_2(\lambda)\cdot\dots\cdot\Delta_s(\lambda)$.

En general, para ver que $i_p(\lambda)=\Delta_p(\lambda)$ basta probar que

$$D_{n-p}(\lambda)=\Delta_{p+1}(\lambda)\cdot\dots\cdot\Delta_s(\lambda) \quad \text{si } p<s$$

y $D_{n-p}(\lambda)=1$, si $p\geq s$.

* Supongamos $p<s$. Un menor de orden $n-p$ de $\hat{A}-\lambda I$ se obtiene suprimiendo p filas y p columnas en $\hat{A}-\lambda I$. Sea $M_{n-p}(\lambda)$ un menor cualquiera de orden $n-p$ de $\hat{A}-\lambda I$.

Vamos a probar: ① $\Delta_{p+1}(\lambda)\cdot\dots\cdot\Delta_s(\lambda) \mid M_{n-p}(\lambda)$ y ② existe un menor de orden $n-p$ de $\hat{A}-\lambda I$ tal que $M_{n-p}(\lambda)=\Delta_{p+1}(\lambda)\cdot\dots\cdot\Delta_s(\lambda)$ con lo cual quedará visto que $D_{n-p}(\lambda)=\Delta_{p+1}(\lambda)\cdot\dots\cdot\Delta_s(\lambda)$.

① En la formación del menor $M_{n-p}(\lambda)$ pueden ocurrir dos casos:

1.a) Existe un bloque diagonal de $\hat{A}-\lambda I$ al cual se le suprimen distinto número de filas que de columnas. Sea $\Delta_j(\lambda)$ uno de ellos y supongamos que en la formación de $M_{n-p}(\lambda)$ este bloque aporta r filas y t columnas con $r\neq t$. Supongamos $r>t$.

Para calcular $M_{n-p}(\lambda)$ desarrollamos este determinante por los menores de orden r de las r filas que están en $\Delta_j(\lambda)$. Entonces cualquier menor de orden r así formado es nulo pues tiene una columna formada por ceros. En este caso sería $M_{n-p}(\lambda)=0$.

1.b) En la formación del menor $M_{n-p}(\lambda)$ cada bloque aporta el mismo número de filas que de columnas. Entonces $M_{n-p}(\lambda)$ es diagonal por bloques. Como $p<s$ y $M_{n-p}(\lambda)$ se obtiene suprimiendo p filas y p columnas de $\hat{A}-\lambda I$, entonces

$s-p$ bloques (al menos) a los que no se ha suprimido ninguna fila ni ninguna columna. Supongamos que estos bloques son $\Delta_{j_1}(\lambda), \dots, \Delta_{j_{s-p}}(\lambda)$ y que $j_1 < \dots < j_{s-p}$.
Entonces $j_{s-p} \leq s$, $j_{s-p-1} \leq s-1$, $\dots$, $j_1 \leq p+1$.
Luego, por (I), $\Delta_s(\lambda) \mid \Delta_{j_{s-p}}(\lambda), \dots, \Delta_{p+1}(\lambda) \mid \Delta_{j_1}(\lambda)$. (II)

Puesto que $M_{n-p}(\lambda) = \chi_1(\lambda) \cdots \chi_p(\lambda) \cdot \Delta_{j_1}(\lambda) \cdots \Delta_{j_{s-p}}(\lambda)$
donde $\Delta_{j_k}(\lambda)$ son los determinantes de los bloques a los que no se ha quitado ninguna "línea" y $\chi_i(\lambda)$ los restantes bloques, a la vista de (II) podemos escribir

$$M_{n-p}(\lambda) = \chi_1'(\lambda) \cdots \chi_p'(\lambda) \cdot \Delta_{p+1}(\lambda) \cdots \Delta_s(\lambda)$$

siendo $\chi_1'(\lambda) = \chi_1(\lambda) \cdot \dfrac{\Delta_{j_1}(\lambda)}{\Delta_{p+1}(\lambda)}, \dots, \chi_p'(\lambda) = \chi_p \dfrac{\Delta_{j_{s-p}}(\lambda)}{\Delta_s(\lambda)}$.

Luego $\Delta_{p+1}(\lambda) \cdots \Delta_s(\lambda) \mid M_{n-p}(\lambda)$, cualquiera que sea el menor de orden $n-p$; por tanto

$$\Delta_{p+1}(\lambda) \cdots \Delta_s(\lambda) \mid D_{n-p}(\lambda).$$

② Veamos ahora que existe un menor de orden $n-p$ que coincide con $\Delta_{p+1}(\lambda) \cdots \Delta_s(\lambda)$.
Los bloques diagonales de la matriz $\hat{A} - \lambda I$ son de la forma

$$\begin{vmatrix} -\lambda & 0 & 0 & \dots & 0 & -b_0 \\ 1 & -\lambda & 0 & \dots & 0 & -b_1 \\ 0 & 1-\lambda & \dots & \dots & 0 & -b_2 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & -\lambda & -b_{k-2} \\ 0 & 0 & 0 & \dots & 1 & -b_{k-1}-\lambda \end{vmatrix}$$

Si a cada uno de los p bloques primeros quitamos la primera fila y la última columna obtenemos un menor de orden $n-p$ que es diagonal por bloques. Luego

$$M_{n-p}(\lambda) = 1 \overset{p}{\cdots} 1 \cdot \Delta_{p+1}(\lambda) \cdots \Delta_s(\lambda).$$

** Supongamos $p \geq s$: Se tiene entonces que $D_{n-p}(\lambda) \mid D_{n-s}(\lambda)$.
Si probamos que $D_{n-s}(\lambda) = 1$ quedará visto que $D_{n-p}(\lambda) = 1$.
Basta probar para ello que existe un menor de orden $n-s$ que es igual a 1. Si quitamos a cada bloque la primera fila y la última columna obtenemos un menor de orden $n-s$ igual a 1.

$i_1(\lambda) = \frac{\Delta_1(\lambda)\cdot\Delta_2(\lambda)\cdot\ldots\cdot\Delta_s(\lambda)}{\Delta_2(\lambda)\cdot\ldots\cdot\Delta_s(\lambda)} = \Delta_1(\lambda), \ldots, i_s(\lambda) = \frac{D_{n-s+1}(\lambda)}{D_{n-s}(\lambda)} = \frac{\Delta_s(\lambda)}{1} = \Delta_s(\lambda)$

e $i_{s+k}(\lambda) = 1$ para $k \geq 1$. c.s.q.d.

7. Cálculo de los divisores elementales de una matriz.

Sea A una matriz de orden n e $i_1(\lambda), \ldots, i_s(\lambda)$ sus factores invariantes, verificando que $i_p(\lambda) \mid i_{p-1}(\lambda)$, $p = 2, \ldots, s$. Sea $i_1(\lambda) = \varphi_1(\lambda)^{c_1^1} \cdot \varphi_2(\lambda)^{c_2^1} \cdot \ldots \cdot \varphi_t(\lambda)^{c_t^1}$ la descomposición de $i_1(\lambda)$ como producto de factores irreducibles. Entonces

$$i_2(\lambda) = \varphi_1(\lambda)^{c_1^2} \cdot \varphi_2(\lambda)^{c_2^2} \cdot \ldots \cdot \varphi_t(\lambda)^{c_t^2} \quad 0 \leq c_i^2 \leq c_i^1$$

$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$

$$i_s(\lambda) = \varphi_1(\lambda)^{c_1^s} \varphi_2(\lambda)^{c_2^s} \cdot \ldots \cdot \varphi_t(\lambda)^{c_t^s} \quad 0 \leq c_i^s \leq c_i^{s-1}.$$

Vamos a probar que los divisores elementales de A son $\varphi_1(\lambda)^{c_1^1}, \varphi_2(\lambda)^{c_2^1}, \ldots, \varphi_t(\lambda)^{c_t^1}, \varphi_1(\lambda)^{c_1^2}, \ldots, \varphi_t(\lambda)^{c_t^2}, \ldots\ldots, \varphi_1(\lambda)^{c_1^s}, \ldots, \varphi_t(\lambda)^{c_t^s}$.

Sea $V = K^n$, siendo K el cuerpo al que pertenecen los elementos de A. Fijada una base de V la matriz A tiene asociado un endomorfismo de V, y V admite una descomposición como suma directa de submódulos cíclicos e invariantes

$$V = W_1 \oplus \ldots \oplus W_s$$

Sea $W = \langle w \rangle$ uno de estos submódulos cíclicos e invariantes, e $i(x) = \varphi_1(x)^{c_1} \cdot \ldots \cdot \varphi_t(x)^{c_t}$ el polinomio mínimo de w. Vamos a probar que W se puede descomponer como suma directa de submódulos cíclicos cuyos polinomios mínimos son los $\varphi_i(x)^{c_i}$. Haciendo esto con todos los W_i obtendremos una descomposición primaria de V y, puesto que dos descomposiciones primarias de V tienen el mismo número de sumandos de cada orden, quedará visto que $\{\varphi_i(\lambda)^{c_i^j}\}_{\substack{i=1,\ldots,t \\ j=1,\ldots,s}}$ es el conjunto de divisores elementales de A. Antes probamos el siguiente:

7.1. LEMA. Sea W un submódulo cíclico de V y $g(x)$ el polinomio mínimo de W. Supongamos que $g(x) = g_1(x) \cdot g_2(x)$, siendo $g_1(x)$ y $g_2(x)$ primos entre sí. Entonces $W = \langle w_1 \rangle \oplus \langle w_2 \rangle$ donde g_1 es el polinomio mínimo de w_1 y g_2 el polinomio mínimo de w_2.

Demostr: Puesto que $K[x]$ es un dominio principal, por la identidad de Bezout, existen $\alpha_1(x), \alpha_2(x) \in K[x]$ tales que $\alpha_1(x) g_1(x) + \alpha_2(x) g_2(x) = 1$

Sean $W_1 = g_2(x)\cdot W$ y $W_2 = g_1(x)\cdot W$.
Entonces el polinomio mínimo de W_1 es $g_1(x)$, pues $g(x)$ es el polinomio mínimo de W y $g = g_1 \cdot g_2$. Análogamente, $g_2(x)$ es el polinomio mínimo de W_2.
Falta probar que $W = \langle W_1 \rangle \oplus \langle W_2 \rangle$.
De (I) se deduce que $W = \alpha_1(x)\cdot g_1(x)\cdot W + \alpha_2(x)\cdot g_2(x)\cdot W =$
$= \alpha_1(x)\cdot W_2 + \alpha_2(x)\cdot W_1 \in \langle W_1 \rangle + \langle W_2 \rangle$. Luego $W \subset \langle W_1 \rangle + \langle W_2 \rangle$.
Además, por definición, $W_1, W_2 \in \langle W \rangle$. Luego $\langle W_1 \rangle + \langle W_2 \rangle \subset W$.
Queda probar que la suma es directa.
Puesto que $\langle W_1 \rangle + \langle W_2 \rangle = W$, la unión de una base de $\langle W_1 \rangle$ y otra de $\langle W_2 \rangle$ es un sistema de generadores de W. Si probamos que dicho sistema de generadores tiene el mismo número de vectores que una base de $\langle W \rangle$, quedaría visto que la suma es directa.
Una base de $\langle W \rangle$ es de la forma $\{w, Tw, \dots, T^{r-1}w\}$ siendo $r = \partial^\circ(g(x))$; es decir, $\dim \langle W \rangle = \partial^\circ(g)$.
Análogamente, $\dim \langle W_1 \rangle = \partial^\circ(g_1)$ y $\dim \langle W_2 \rangle = \partial^\circ(g_2)$.
Entonces el sistema de generadores obtenido al "unir" las bases de $\langle W_1 \rangle$ y $\langle W_2 \rangle$ tiene $\partial^\circ(g_1) + \partial^\circ(g_2)$ vectores.
Puesto que $\partial^\circ(g) = \partial^\circ(g_1) + \partial^\circ(g_2)$ (consideramos g, g_1 y g_2 mónicos) queda probado el lema. ■

Volviendo al problema planteado tenemos que si $W = \langle w \rangle$ e $i(x) = \varphi_1(x)^{c_1} \cdots \varphi_t(x)^{c_t}$ es el polinomio mínimo de w, entonces utilizando el lema anterior y, por recurrencia simple, se prueba que $W = \langle u_1 \rangle \oplus \langle u_2 \rangle \oplus \dots \oplus \langle u_t \rangle$ siendo $\varphi_i(x)^{c_i}$ el polinomio mínimo de u_i. Luego, los divisores elementales de A son los $\varphi_i(x)^{d_j}$, $i \in \{1, \dots, t\}$, $j \in \{1, \dots, s\}$.

8. Forma canónica de Jordan de una matriz.

Sea A una matriz de orden n con coeficientes en un cuerpo K.
Supongamos que sus divisores elementales son $\varphi_1(x), \dots, \varphi_r(x)$, que no tienen porque ser distintos, en principio.
Supondremos que el cuerpo K es algebraicamente cerrado o, al menos, que contiene todos los autovalores de A (raíces del polinomio característico de A: $|A - \lambda I|$).
Puesto que los divisores elementales son potencias de polinomios

rístico de A, los divisores elementales son de la forma

$$\varphi_i(x) = (x-\lambda)^{c_i}$$

donde λ es un autovalor de A.

Fijada una base de $V = K^n$, sea T el endomorfismo asociado a la matriz A. Entonces V admite una descomposición

$$V = W_1 \oplus \dots \oplus W_r$$

donde los $W_i = \langle w_i \rangle$ son submódulos cíclicos invariantes por T de manera que el polinomio mínimo de w_i es $\varphi_i(x)$.

Sea $W = \langle w \rangle$ uno de estos submódulo cíclicos. Entonces si k es el grado del polinomio mínimo de w tenemos que $\{w, Tw, \dots, T^{k-1}w\}$ es base de W (como subespacio vectorial de V). El polinomio mínimo de w es entonces $(x-\lambda)^k$, donde λ es un autovalor de A.

Consideremos entonces el siguiente sistema de vectores

$$S = \{w, (T-\lambda i)w, \dots, (T-\lambda i)^{k-1}w\}$$

donde i es la identidad de V.

Veamos que S es base de W.

Por supuesto, $S \subset W$ pues $B = \{w, Tw, \dots, T^{k-1}w\}$ es base de W y $(T-\lambda i)^t \cdot w = (x-\lambda)^t \cdot w \in W$, por ser submódulo.

Veamos que S es sistema de generadores de W, con lo cual será base, por ser minimal. Basta probar que todo elemento de B es combinación lineal de los vectores de S.

Puesto que $w \in S$ queda ver que $T^j w$ está generado por los elementos de S con $j \in \{1, \dots, k-1\}$ (para $j=0$ ya lo hemos probado, pues $T^0 w = w$). Hagamos la hipótesis de inducción: $T^{j-1}w = \sum_{l=0}^{k-1} \alpha_l (T-\lambda i)^l \cdot w$.

Entonces $T^j w = T\left(\sum_{l=0}^{k-1} \alpha_l (T-\lambda i)^l \cdot w\right) = \sum_{l=0}^{k-1} \alpha_l T(T-\lambda i)^l \cdot w =$

$$= \sum_{l=0}^{k-1} \alpha_l (T-\lambda i)^{l+1} \cdot w + \sum_{l=0}^{k-1} \lambda \alpha_l (T-\lambda i)^l w$$

Si $l < k-1$, $T^j w$ es combinación lineal de los vectores de S.

Si $l = k-1$, tenemos que $(T-\lambda i)^{l+1} w = (T-\lambda i)^k w = 0$, pues $(x-\lambda)^k$ es el polinomio mínimo de w.

Luego, en cualquier caso, S es base de W.

Consideremos entonces, para cada subespacio W_i de S.

se de V. En esta base el endomorfismo T viene representado por una matriz diagonal por bloques (semejante a A) y estos bloques son de la forma:

$$\begin{array}{c|cccc} T_w & T(T-\lambda i)w & \cdots & T(T-\lambda i)^{k-2}w & T(T-\lambda i)^{k-1}w \\ \hline \lambda & 0 & \cdots & 0 & 0 \\ 1 & \lambda & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ 0 & 0 & \cdots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & \lambda & 0 \\ 0 & 0 & \cdots & 1 & \lambda \end{array} \qquad (I)$$

pues haciendo $e_1 = w$, $e_2 = (T-\lambda i)w$, ---, $e_k = (T-\lambda i)^{k-1}w$ tenemos

$$Tw = (T-\lambda i)w + \lambda w = \lambda e_1 + e_2$$
$$T(T-\lambda i)w = (T-\lambda i)^2 w + \lambda(T-\lambda i)w = \lambda e_2 + e_3$$
$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$
$$T(T-\lambda i)^{k-2}w = (T-\lambda i)^{k-1}w + \lambda(T-\lambda i)^{k-2}w = \lambda e_{k-1} + e_k$$
$$T(T-\lambda i)^{k-1}w = (T-\lambda i)^k w + \lambda(T-\lambda i)^{k-1}w = \lambda e_k, \text{ pues } (T-\lambda i)^k w = 0.$$

Por tanto, toda matriz cuadrada A es semejante a una matriz diagonal por bloques de la forma (I), donde λ es un autovalor de la matriz.

A esta matriz se le llama FORMA CANONICA DE JORDAN de A.

Se prueba que la forma canónica de Jordan saldrá diagonal cuando, y solo cuando, los divisores elementales son de grado 1 (en cuyo caso los bloques solo tienen un elemento).

Que los divisores elementales sean de grado 1 equivale a que exista una base de V formada por autovectores.