Universitat de Barcelona

Facultat de Matemàtiques i Informàtica

GRAU DE MATEMÀTIQUES

Treball final de grau

Forma canònica de Weyr i aplicacions

Autor: Héctor Pablo Fernàndez Zacarías

Director: Dra. Mª Eulàlia Montoro López
Realitzat a: Departament de Matemàtiques i Informàtica
Barcelona, 18 de gener de 2020

Abstract

Matrix canonical forms (with respect to similarity) provide exemplars for each similarity class, and let us study in a simpler way some properties of the square matrices. We will focus on the Weyr canonical form. We will see that although it's the best known, the Jordan form it's not always the most suitable form for the study of a given matrix. In general, we can't say there is a best canonical form, it will depend on each case. We will see differents proofs of the existence and uniqueness of the canonical forms, how to calculate them, the relationships between them, and some applications.

Resum

Les formes canòniques, respecte a una relació de semblança, ens donen representants per a cada classe, i ens permeten estudiar de manera més senzilla certs aspectes sobre les matrius quadrades. Ens centrarem en la forma canònica de Weyr. Veurem que tot i ser la més coneguda, la forma canònica de Jordan no sempre resulta la més adient per l'estudi d'una matriu donada. Tampoc podem dir que hi hagi una forma canònica millor en general, ja que dependrà de les circumstàncies en què ens trobem. Veurem diferents demostracions de l'existència i unicitat d'ambdues formes canòniques, com calcular-les, de quina manera es relacionen i algunes aplicacions.

2010 Mathematics Subject Classification. 15A21, 15A27

Agraïments

Vull agrair primer de tot a la meva família per haver-me aguantat durant tots aquests anys, a l'Andrea per ser-hi sempre, i sobretot a la meva tutora Eulàlia ja que aquest treball no hagués estat possible sense el seu suport.

Índex

1 Introducció

Una forma canònica d'una matriu, respecte a la relació de semblança, ens prové d'un representant amb una forma relativament simple per a cada classe i ens permet contestar de manera més senzilla certes qüestions respecte aquest representant.
Bàsicament es coneixen tres formes canòniques. De les tres, la més coneguda és la forma canònica de Jordan i la menys coneguda és la forma canònica de Weyr (ambdues formes es poden obtenir quan treballem sobre un cos algebraicament tancat). La tercera forma canònica és la racional (que es pot treballar sobre qualsevol cos).
Coneguts els valors propis de la matriu, la forma canònica de Weyr es calcula de manera algorítmicament més simple que la de Jordan, i a més presenta millors propietats a l'hora d'interaccionar amb altres matrius. Tot i que a nivell "estètic" la forma de Jordan resulta més atractiva. Per tant, no es pot dir que una forma canònica és millor ja que depenent de les circumstàncies serà més convenient una forma o altra.

La forma canònica de Jordan fou descoberta pel matemàtic francès *Camille Jordan* l'any 1870 i publicada per primera vegada a [9], on oferia un estudi exhaustiu de la Teoria de Galois, a més de ser la primera publicació en tractar la Teoria de grups com a tal. És aquí on va aparèixer per primera vegada la Forma canònica de Jordan (sobre cossos finits), i on sembla ser desconeixia els resultats publicats prèviament per *Karl Weierstrass*, on es definia una forma normal equivalent a la de Jordan, però en el cos dels complexos.

La forma canònica de Weyr va ser descrita per primera vegada, de manera molt breu, l'any 1885 ([10]) pel propi *Eduard Weyr*, i no va ser fins l'any 1890 ([11]) que va desenvolupar de manera més extensa. Tot i que l'any 1932 *Herbert Turnbull* va fer referència a la característica de Weyr, aquesta forma canònica ha estat totalment oblidada fins fa realment poc (De fet, al *Handbook* d'àlgebra lineal del 2007 no es menciona enlloc). No resulta sorprenent doncs, que s'hagi redescobert i reanomenat diverses vegades durant el darrer segle (s'ha arribat a dir que era una forma modificada de Jordan, una segona forma de Jordan, etc.).

La forma de Weyr té diverses aplicacions, una d'elles és el Teorema de Gerstenhaber, que ens diu que la subàlgebra $F[A, B]$ generada per dues matrius $A, B \in M_n(\mathbb{K})$ que commuten, té dimensió com a màxim n. Aquest teorema va ser demostrat inicialment utilitzant tècniques de geometria algebraica, però posteriorment *Barría i Halmos* a [12], i *Laffey i Lazarus* a [13], van obtenir demostracions fent servir únicament conceptes d'àlgebra lineal i la forma canònica de Jordan. En [1] la demostració de *Barría i Halmos* es simplifica fent servir la forma canònica de Weyr.

El treball està distribuït de la següent manera. En la Secció 2 es recorden definicions i s'introdueix notació que apareixerà al llarg del treball. En la Secció 3 es mostren demostracions de la forma de Jordan no vistes al grau. En la Secció 4 s'introdueix i es demostra la forma canònica de Weyr, on a més s'inclou un algoritme per calcular-la. En la Secció 5 veurem de quina manera es relacionen les formes canòniques de Jordan i de Weyr. En la Secció 6 estudiem els centralitzadors d'una matriu donada; Quan el cos és algebraicament tancat, el càlcul del centralitzador es pot reduir a considerar que la matriu està en forma canònica. Hem considerat important incloure aquesta secció degut a que aparentment la forma de Weyr presenta un centralitzador amb aparença més simple que el de Jordan. Finalment, en la Secció 7 es mostren alguns resultats on s'ha demostrat que la utilització de la forma canònica de Weyr simplifica les demostracions envers la utilització de la forma canònica de Jordan.

2 Preliminars

Suposarem durant tot el treball que $\mathbb{K}$ és un cos algebraicament tancat per tal de garantir l'existència de zeros dels diferents polinomis que aniran apareixent. Denotarem per $M_n(\mathbb{K})$ el conjunt de les matrius quadrades de dimensió n amb coeficients sobre el cos $\mathbb{K}$, i $M_n = M_n(\mathbb{C})$ quan $\mathbb{K} = \mathbb{C}$ sigui el cos dels complexos. Donada una matriu A, ens referirem als valors i vectors propis de A com VAPs i VEPs, respectivament, i $Spec(A)$ denotarà el conjunt dels seus valors propis.

Recordem ara algunes definicions:

Una base, d'un $\mathbb{K}$-espai vectorial E de dimensió finita, és un conjunt de vectors $\mathcal{B} = \{v_1, \ldots, v_n\}$ linealment independents i que generen l'espai E.

Donat $f : E \to E$ un endomorfisme, si fixem una base de l'espai E podem representar l'endomorfisme amb una única matriu $A \in M_n(\mathbb{C})$, que dependrà de la base escollida anteriorment.

Donat un espai vectorial tenim infinites bases i per tant infinites matrius associades a un cert endomorfisme. Volem trobar aquelles representacions que ens facilitin, per exemple, realitzar operacions amb la seva matriu associada o l'estudi del nostre endomorfisme (saber si existeixen altres que estiguin relacionats amb el nostre però que tinguin una representació més senzilla, la seva estructura, etc.).

Donat un endomorfisme $f : E \to E$, un subespai $U \subset E$ es diu **invariant** sobre f si $f(U) \subseteq U$. Un vector $v \in E$ és un VEP de la matriu associada a l'aplicació f si $U = < v >$ és invariant sobre f. De fet, aquest és el subespai invariant més petit. Es defineix $Im(f) = \{v \in E \mid v = f(u), u \in E\}$ i $ker(f) = \{v \in E \mid f(v) = 0\}$. A més, denotarem per $rang(f)$ a la dimensió del subespai $Im(f)$, i $nul(f)$ com la dimensió del $ker(f)$. Els subespais invariants són claus a l'hora d'obtenir representacions "senzilles"de l'endomorfisme.

Dues matrius A i $B \in M_n(\mathbb{K})$ són **semblants** si existeix $S \in M_n(\mathbb{K})$ invertible de manera que $B = S^{-1}AS$. En altres paraules, si dues matrius A i B són semblants, aleshores són representacions de la mateixa transformació lineal però en bases diferents. Quan $\mathbb{K} = \mathbb{C}$, diem que són **unitàriament semblants** si existeix $U \in M_n$ unitària tal que $A = UBU^*$. Si $\mathbb{K} = \mathbb{R}$, aleshores diem que A és **ortogonalment semblant** a B.

El fet que dues matrius siguin semblants, dóna lloc a una relació d'equivalència definida per

$$A \sim B \Leftrightarrow \ \exists \ S \ invertible \ tal \ que \ S^{-1}AS = B.$$

Veiem-ho:

Propietat reflexiva: Si prenem $S = Id$, tenim que $A = Id^{-1}AId$, i per tant $A \sim A$.

Propietat simètrica: suposem que $A \sim B$, per tant existeix S invertible tal que $S^{-1}AS = B$, aleshores si fem el producte per l'esquerra per la matriu S a les dues bandes de la igualtat, obtenim $AS = SB$, i fent el producte per la dreta per la matriu S^{-1} a la nova igualtat acabem obtenint que $A = SBS^{-1}$, i.e., $B \sim A$.

Propietat transitiva: Si $A \sim B$ i $B \sim C$, aleshores existeixen dues matrius invertibles T i R tals que $B = T^{-1}AT$, $C = R^{-1}BR$, ajuntant les dues igualtats obtenim que $C = R^{-1}(T^{-1}AT)R = (TR)^{-1}A(TR) = S^{-1}AS \Rightarrow A \sim C$.

El fet que la semblança entre matrius indueixi una relació d'equivalència ens permet establir classes d'equivalència i triar els representants d'aquestes, amb l'objectiu que siguin tan "senzills" com puguem. Aquests representants els anomenarem **formes canòniques**.

Per saber si dues matrius són semblants, una manera de procedir seria establir un conjunt de matrius amb unes certes característiques i veure si a través d'alguna transformació per semblances, les dues matrius resultants coincideixen amb algun dels representats que hem establert abans. Si és així, com que la relació per semblances és transitiva i reflexiva, han de ser semblants, en cas contrari, no ho podrien ser.

Sabem que totes les matrius quadrades, sobre un cos $\mathbb{K}$ algebraicament tancat, són semblants a una matriu triangular superior (Lema de Schur), malgrat això, el representant de les matrius triangulars superiors no és un bon candidat a representant. Podríem pensar que si dues matrius poden ser transformades via semblances en dues matrius triangulars superiors amb la mateixa diagonal però elements diferents fóra de la diagonal, aleshores són semblants, però això no és cert. Per exemple,

$$\begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix} \not\sim \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Una primera forma canònica són matrius triangulars per blocs amb unes certes característiques, concretament **matrius de Jordan**. Si, via semblances, transformem les nostres dues matrius en matrius de Jordan i resulta que tenen els mateixos blocs a la diagonal (llevat de permutacions), aleshores podem dir que són semblants. De fet, el recíproc també és cert.

Definició 2.1. *Una* ***matriu bàsica de Jordan amb VAP*** λ *és una matriu*

$$J = \begin{bmatrix} \lambda & 1 & & \\ & \lambda & & \\ & & \ddots & 1 \\ & & & \lambda \end{bmatrix}$$

amb el VAP λ a la diagonal, on la diagonal superior són tot 1's i la resta d'elements de la matriu són 0. Observem que

$$J = \lambda Id + N, \text{ on } N = \begin{bmatrix} 0 & 1 & & \\ & 0 & & \\ & & \ddots & 1 \\ & & & 0 \end{bmatrix} \text{ és nilpotent}$$

*Una **matriu de Jordan** es defineix com*

$$J = \begin{bmatrix} J_1 & & & \\ & J_2 & & \\ & & \ddots & \\ & & & J_k \end{bmatrix} = diag(J_1, \dots, J_k),$$

on J és suma directa de k matrius bàsiques de Jordan amb els seus VAPs corresponents.

*En cas que existeixen blocs amb el mateix VAP, els agruparem i ordenarem de manera creixent segons les dimensions de cada bloc. Aleshores si els blocs, ordenats, tenen mida $m_1 \geq \dots \geq m_s$, direm que $(m_1, \dots, m_s)$ és la seva **estructura de Jordan associada al VAP** λ, també coneguda com **característica de Segre**.*

Exemple 2.2.

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

té característica de Segre $(3, 2)$ pel VAP $\lambda = 3$ i característica de Segre (1) pel VAP $\lambda = 2$.

Recordem que l'**espai propi generalitzat** de A, corresponent al VAP λ, es defineix com

$$G(\lambda) = \{x \in E \mid (A - \lambda Id)^m x = 0, m \in \mathbb{N}\} \supseteq E(\lambda),$$

on $E(\lambda) = ker(A - \lambda Id)$.

Quan $\mathbb{K}$ és algebraicament tancat, i $p(x) = (x - \lambda_1)^{m_1} \dots (x - \lambda_k)^{m_k}$ és la factorització en factors lineals del polinomi mínim de A, llavors

$$G(\lambda_i) = \{x \in \mathbb{K}^n \mid (A - \lambda_i Id)^{m_i} x = 0\} = ker(A - \lambda_i Id)^{m_i}.$$

Una primera descomposició en subespais invariants de E, que ens permet trobar una representació "senzilla"de l'endomorfisme, és la que ve donada pel següent teorema.

Teorema 2.3. ***[1] (Primer Teorema de descomposició).*** *Siguin E un $\mathbb{K}$-espai vectorial de dimensió n, l'endomorfisme $f : E \to E$, i $p(x) \in \mathbb{K}[x]$ un cert polinomi mònic, no constant, tal que $p(f) = 0$. Sigui $p = p_1^{m_1} \dots p_k^{m_k}$ la factorització de p en polinomis mònics irreductibles diferents, $W_i = ker(p_i(f))^{m_i}$, per $i = 1, \dots, k$, aleshores els subespais W_i són invariants per f, i*

$$E = W_1 \oplus \dots \oplus W_k.$$

En particular, quan $\mathbb{K}$ és algebraicament tancat tenim el següent corol·lari.

Corol·lari 2.4. *Per qualsevol matriu $A \in M_n(\mathbb{K})$, tenim que*

$$E = G(\lambda_1) \oplus \dots \oplus G(\lambda_k),$$

on $\lambda_1, \ldots, \lambda_k$ són els diferents VAPs de A, i $G(\lambda_i)$ és l'espai propi generalitzat associat al VAP λ_i.

Recordem que el **generador d'una cadena de Jordan** és un vector propi generalitzat v_r tal que $(A - \lambda Id)^r v_r = 0$, on $r =$ dimensió del bloc de Jordan. Així, podem generar una **cadena de Jordan** formada per vectors v_i tals que $v_i = (A - \lambda Id)^{-1} v_{i-1}$.

Pel Corol·lari 2.4, tota matriu $A \in M_n(\mathbb{K})$ admet una representació de la forma

$$A = \begin{bmatrix} A_1 & & \\ & \ddots & \\ & & A_k \end{bmatrix}.$$

Més concretament, agafant bases corresponents a cada $G(\lambda_i)$, es deriva el resultat següent.

Teorema 2.5. *[2] Tota matriu $A \in M_n$ es pot representar, de manera única, en la forma canònica de Jordan.*

3 Altres demostracions de la forma canònica de Jordan

A continuació detallem dues demostracions diferents corresponents a l'existència i unicitat de la forma de Jordan.

3.1 Primera demostració de l'existència i unicitat de la forma canònica de Jordan

La forma canònica de Jordan també es pot demostrar sense utilitzar el primer teorema de descomposició.

Veiem en primer lloc una demostració basada en semblances de matrius. Aquesta demostració la dividirem en tres passos.

Primer pas: Veurem que tota matriu és semblant a una matriu triangular superior, tal que a la seva diagonal principal hi apareixen els seus VAPs (agrupats en cas que la seva multiplicitat sigui més gran que 1).

Segon pas: Demostrarem que una matriu amb la forma descrita anteriorment és semblant a una matriu diagonal per blocs, on cada bloc de la diagonal és triangular superior.

Tercer pas: Finalment, veurem que una matriu triangular superior tal que els seus elements a la diagonal són tots iguals, és semblant a una matriu de Jordan.

Teorema 3.1. ***[8] (Forma de Schur / Triangularització de Schur)*** *Sigui $A \in M_n$ amb VAPs $\lambda_1, \dots, \lambda_n$, i sigui $x \in \mathbb{C}^n$ un vector unitari tal que $Ax = \lambda_1 x$. Aleshores es compleix que:*

(a) Existeix una matriu unitària $U = (x, u_2, \dots, u_n) \in M_n$ tal que $U^*AU = T = (t_{ij})$, on T és triangular superior amb $t_{ii} = \lambda_i, i = 1, \dots, n$.
(b) Si $A \in M_n(\mathbb{R})$ és tal que els seus VAPs són tots reals, aleshores podem escollir que $x \in \mathbb{R}$ tal que $Q = (x, q_2, \dots, q_n) \in M_n(\mathbb{R})$ i $Q^TAQ = T = (t_{ij})$, on T és triangular superior amb $t_{ii} = \lambda_i, i = 1, \dots, n$.

Demostració. Sigui x el VEP normalitzat de la matriu A associat al VAP λ_1, i.e., $x^*x = 1$ i $Ax = \lambda_1 x$. Sigui $U_1 = [x, u_2, \dots, u_n]$ la matriu unitària on la seva primera columna és el vector x. Aleshores

$$U_1^*AU_1 = U_1^*[Ax, Au_2, \dots, Au_n] = U_1^*[\lambda_1 x, Au_2, \dots, Au_n] = \begin{bmatrix} \lambda_1 x^*x & x^*Au_2 & \dots & x^*Au_n \\ \lambda_1 u_2^*x & & & \\ \vdots & & A_1 & \\ \lambda_1 u_n^*x & & & \end{bmatrix}$$

$= \begin{bmatrix} \lambda_1 & * \\ 0 & A_1 \end{bmatrix}$, ja que les columnes de la matriu U_1 són ortonormals.

Els VAPs de la submatriu $A_1 = [u_i^*Au_j]_{i,j=2}^n \in M_{n-1}$ són $\lambda_2, \dots, \lambda_n$. En el cas $n = 2$, hem aconseguit la triangularització desitjada. Si $n \neq 2$, sigui $\xi \in \mathbb{C}^{n-1}$ un VEP de la submatriu A_1 associat al VAP λ_2, ara podem aplicar sobre A_1 el mateix procediment que hem aplicat abans. Si $U_2 \in M_{n-1}$ és una matriu unitària qualsevol tal que la seva primera columna és ξ, aleshores

$$U_2^* A_1 U_2 = \begin{bmatrix} \lambda_2 & * \\ 0 & A_2 \end{bmatrix}.$$

Sigui $V_2 = [1] \oplus U_2$, aleshores $(U_1V_2)^* A U_1 V_2 = V_2^* u_1^* A U_1 V_2 = \begin{bmatrix} \lambda_1 & * & * \\ 0 & \lambda_2 & * \\ 0 & 0 & A_2 \end{bmatrix}$.

Així construïm matrius unitàries $U_i \in M_{n-1}, i = 1, \ldots, n-1$ i $V_i \in M_n, i = 2, \ldots, n-2$. La matriu $U = U_1 V_2 V_3 \ldots V_{n-2}$ és unitària i $U^* A U$ és triangular superior.

Si tots els VAPs de la matriu $A \in M_n(\mathbb{R})$ són reals, aleshores tots els VEPs i matrius unitàries escollides en la demostració es podem escollir que siguin reals. $\square$

A continuació veurem el segon pas, on demostrarem que una matriu triangular superior es pot transformar mitjançant semblances en una matriu diagonal per blocs, on cada bloc és triangular superior.

Teorema 3.2. *[8] Siguin $A \in M_n(\mathbb{K})$, i $\lambda_1, \ldots, \lambda_d \in M_n(\mathbb{K})$ els corresponents VAPs (amb multiplicitats $n_1, \ldots, n_d$, respectivament). El Teorema 3.1 ens assegura que A és unitàriament semblant a una matriu triangular superior per blocs, $T = (T_{ij})_{i,j=1}^d \in M_d$, on cada bloc $T_{ij} \in M_{n_i,n_j}$ és tal que $T_{ij} = 0, i > j$. A més, cada bloc diagonal $T_{ii} = \lambda_i Id_{n_i} + R_i$, on $R_i \in M_{n_i}$ és estrictament triangular superior. Aleshores A és semblant a*

$$T = \begin{bmatrix} T_{11} & 0 & \ldots & 0 \\ 0 & T_{22} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \ldots & 0 & T_{dd} \end{bmatrix}$$

Si $A \in M_n(\mathbb{R})$ i tots els seus VAPs són reals, aleshores tant la semblança unitària que redueix A a una matriu triangular superior T i la semblança matricial que redueix T a una forma diagonal per blocs, es poden prendre reals.

Demostració. Triem la partició de T tal que $T = \begin{bmatrix} T_{11} & Y \\ 0 & S_2 \end{bmatrix}$, on $S_2 = (T_{ij})_{i,j=2}^d$.

Observem que l'únic VAP de T_{11} és λ_1 i que els VAPs de S_2 són $\lambda_2, \ldots, \lambda_d$. Sabem pel Teorema de Sylvester que l'equació $T_{11}X - XS = -Y$ té una solució X, per tant, considerem $M = \begin{bmatrix} Id_{n_1} & X \\ 0 & Id \end{bmatrix} \Rightarrow M^{-1} = \begin{bmatrix} Id_{n_1} & -X \\ 0 & Id \end{bmatrix}$. Aleshores,

$$M^{-1}TM = \begin{bmatrix} Id_{n_1} & -X \\ 0 & Id \end{bmatrix} \begin{bmatrix} T_{11} & Y \\ 0 & S_2 \end{bmatrix} \begin{bmatrix} Id_{n_1} & X \\ 0 & Id \end{bmatrix} = \begin{bmatrix} T_{11} & T_{11}X - XS_2 + Y \\ 0 & S_2 \end{bmatrix} = \begin{bmatrix} T_{11} & 0 \\ & S_2 \end{bmatrix}$$

Si $d = 2$, tenim la diagonalització per blocs que desitgem.
Si $d > 2$, repetim el procés de reducció i veiem que S_2 és semblant a la matriu $T_{22} \oplus S_3$, on $S_3 = (T_{ij})_{i,j=3}^d$. Després de $d-1$ reduccions, tenim que T és semblant a la matriu

$T_{11} \oplus \cdots \oplus T_{dd}$.

En cas que A sigui real i tots els seus VAPs també ho siguin, A és ortogonalment real a una matriu real triangular superior per blocs, ja que cadascun dels passos anteriors (fets en el cas complex) es pot fer també en el cas real. $\square$

En aquest darrer pas, veurem que si A és estrictament triangular superior, la podem transformar en una matriu diagonal per blocs, on cada bloc tindrà l'estructura d'un bloc bàsic de Jordan.

Teorema 3.3. *[8] Sigui $A \in M_n$ estrictament triangular superior, aleshores existeix una matriu invertible $S \in M_n$ i existeixen $m_1, \dots, m_r \in \mathbb{Z} \mid m_1 \geq m_2 \geq \cdots \geq m_r \geq 1$ tals que si $m_1 + \cdots + m_r = n$,*

$$A = S(J_{m_1}(0) \oplus \cdots \oplus J_{m_r}(0))S^{-1}, \text{ on } J_{m_i}(\lambda) = \begin{bmatrix} \lambda & 1 & & \\ & \lambda & \ddots & \\ & & \ddots & 1 \\ & & & \lambda \end{bmatrix}.$$

Observació: si $A \in M_n(\mathbb{R})$, podem triar S real.

Demostració. Veiem-ho per inducció sobre n:
Si $n = 1, A = [0]$ i és trivial. Fem-ho ara per inducció sobre $n > 1$, suposem que és cert per totes les matrius estrictament triangular superiors amb dimensió menor que n. Triem la partició de la matriu

$$A = \begin{bmatrix} 0 & a^T \\ 0 & A_1 \end{bmatrix},$$

on $a \in \mathbb{C}^{n-1}$ i $A_1 \in M_{n-1}(\mathbb{C})$ és estrictament triangular superior. Per hipòtesis d'inducció, existeix una matriu invertible $S_1 \in M_{n-1} \mid S_1^{-1} A_1 S_1$ té l'estructura desitjada, i.e.,

$$S_1^{-1} A_1 S_1 = \begin{bmatrix} J_{k_1} & & \\ & \ddots & \\ & & J_{k_s} \end{bmatrix} = \begin{bmatrix} J_{k_1} & 0 \\ 0 & J \end{bmatrix},$$

on cada $k_1 \geq \cdots \geq k_s \geq 1, \sum_{i=1}^{s} k_i = n-1, J_{k_i} = J_{k_i}(0), i\ J = J_{k_2} \oplus \cdots \oplus J_{k_s} \in M_{n-k_1-1}$. Cap bloc diagonal de Jordan en J, J_{k_i}, té dimensió més gran que k_1, per tant $J^{k_1} = 0$. Es pot observar que

$$\begin{bmatrix} 1 & 0 \\ 0 & s_1^{-1} \end{bmatrix} A \begin{bmatrix} 1 & 0 \\ 0 & S_1 \end{bmatrix} = \begin{bmatrix} 0 & a^T S_1 \\ 0 & S_1^{-1} A_1 S_1 \end{bmatrix}.$$

Fem $a^T S_1 = [a_1^T a_2^T]$, amb $a_1 \in \mathbb{C}^{k_1}, i\ a_2 \in \mathbb{C}^{n-k_1-1}$, i escriurem

$$\begin{bmatrix} 1 & 0 \\ 0 & s_1^{-1} \end{bmatrix} A \begin{bmatrix} 1 & 0 \\ 0 & S_1 \end{bmatrix} = \begin{bmatrix} 0 & a_1^T & a_2^T \\ 0 & J_{k_1} & 0 \\ 0 & 0 & J \end{bmatrix}.$$

Si considerem la semblança
$\begin{bmatrix} 1 & -a_1^T J_{k_1}^T & 0 \\ 0 & Id & 0 \\ 0 & 0 & Id \end{bmatrix} \begin{bmatrix} 0 & a_1^T & a_2^T \\ 0 & J_{k_1} & 0 \\ 0 & 0 & J \end{bmatrix} \begin{bmatrix} 1 & a_1^T J_{k_1}^T & 0 \\ 0 & Id & 0 \\ 0 & 0 & Id \end{bmatrix} = \begin{bmatrix} 0 & a_1^T(Id - J_{k_1}^T J_{k_1}) & a_2^T \\ 0 & J_{k_1} & 0 \\ 0 & 0 & J \end{bmatrix} =$
$\begin{bmatrix} 0 & (a_1^T e_1) e_1^T & a_2^T \\ 0 & J_{k_1} & 0 \\ 0 & 0 & J \end{bmatrix}$, on hem fet servir que $(Id - J_k^T J_k)x = (x^T e_1)e_1$.

Suposem que $a_1^T e_1 \neq 0$, aleshores

$$\begin{bmatrix} 1/a_1^T e_1 & 0 & 0 \\ 0 & Id & 0 \\ 0 & 0 & 1/a_1^T e_1 Id \end{bmatrix} \begin{bmatrix} 0 & (a_1^T e_1) e_1^T & a_2^T \\ 0 & J_{k_1} & 0 \\ 0 & 0 & J \end{bmatrix} \begin{bmatrix} a_1^T e_1 & 0 & 0 \\ 0 & Id & 0 \\ 0 & 0 & a_1^T e_1 Id \end{bmatrix} = \begin{bmatrix} 0 & e_1^T & a_2^T \\ 0 & J_{k_1} & 0 \\ 0 & 0 & J \end{bmatrix}$$

$= \begin{bmatrix} \bar{J} & a_2^T \\ 0 & J \end{bmatrix}$ on $\tilde{\mathrm{J}} = \begin{bmatrix} 0 & e_1^T \\ 0 & J_{k_1} \end{bmatrix} = J_{k_1+1}(0)$. Com que $\bar{J} e_{i+1} = e_i,\ i = 1, \ldots, k_1$, tenim que

$$\begin{bmatrix} Id & e_2 a_2^T \\ 0 & Id \end{bmatrix} \begin{bmatrix} \bar{J} & e_1 a_2^T \\ 0 & J \end{bmatrix} \begin{bmatrix} Id & -e_2 a_2^T \\ 0 & Id \end{bmatrix} = \begin{bmatrix} \bar{J} & -\bar{J} e_2 a_2^T + e_1 a_2^T + e_2 a_2^T J \\ 0 & J \end{bmatrix} = \begin{bmatrix} \bar{J} & e_2 a_2^T J \\ 0 & J \end{bmatrix}$$

De manera recursiva calculem les semblances
$\begin{bmatrix} Id & e_2 a_2^T \\ 0 & Id \end{bmatrix} \begin{bmatrix} \bar{J} & e_1 a_2^T \\ 0 & J \end{bmatrix} \begin{bmatrix} Id & -e_2 a_2^T \\ 0 & Id \end{bmatrix} = \begin{bmatrix} \bar{J} & e_{i+1} a_2^T J^i \\ 0 & J \end{bmatrix}$, per $i = 2, 3, \ldots$.

Com que $J^{k_1} = 0$, després d'almenys k_1 iterades, aconseguim eliminar l'últim element fora de la diagonal que no era 0. Així, A és semblant a $\begin{bmatrix} \bar{J} & 0 \\ 0 & J \end{bmatrix}$, que és una matriu de Jordan estrictament triangular superior per blocs, com volíem.

Suposem ara que $a_1^T e_1 = 0$, aleshores A és semblant a la matriu $\begin{bmatrix} 0 & 0 & a_2^T \\ 0 & \bar{J}_{k_1} & 0 \\ 0 & 0 & J \end{bmatrix}$, que és semblant per permutacions a la matriu $\begin{bmatrix} J_{k_1} & 0 & 0 \\ 0 & 0 & a_2^T \\ 0 & 0 & J \end{bmatrix}$.

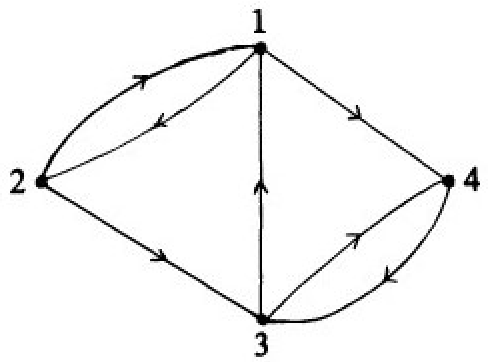
Per hipòtesis d'inducció, existeix una matriu invertible $S_2 \in M_{n-k_1}$ tal que $S_2^{-1} \begin{bmatrix} 0 & a_2^T \\ 0 & J \end{bmatrix} S_2 =$ $\hat{J} \in M_{n-k_1}$ és una matriu de Jordan amb zeros a la diagonal. Per tant, la matriu $\begin{bmatrix} J_{k_1} & 0 & 0 \\ 0 & 0 & a_2^T \\ 0 & 0 & J \end{bmatrix}$, i per tant A també, és semblant a la matriu $\begin{bmatrix} J_{k_1} & 0 \\ 0 & \hat{J} \end{bmatrix}$, que és una matriu de Jordan amb l'estructura que volíem, excepte perquè els blocs de Jordan de la diagonal poden no estar en ordre de dimensió creixent, però amb una matriu de permutacions per blocs podem canviar l'ordre dels blocs com ens convingui.

Observem que si la matriu A és real, totes les semblances també ho són, així que A és semblant, via semblances reals, a una matriu de Jordan amb la forma enunciada a les hipòtesis. □

3.2 Segona demostració de l'existència i unicitat de la forma canònica de Jordan

Richard A. Brualdi a [3], reescriu al llenguatge actual una antiga demostració de *Herbert Turnbull* fent servir llenguatge combinatori i teoria de grafs. Tota matriu $A \in M_n$ té associat un graf dirigit, $D(A)$, el qual té com a *vèrtexs* (numerats) els enters $1, 2, \ldots, n$ (que corresponen simultàniament a files i columnes de la matriu). Direm que existeix una *aresta (i,j)*, amb origen el vèrtex $\{i\}$ i destí el vèrtex $\{j\}$, si l'element $a_{ij} \neq 0$. Si $\{1, \ldots, k\}$ és un conjunt de vèrtexs diferents del graf tals que, per exemple, $(1,2), (2,3), \ldots, (k-1, k)$ són arestes del graf, aleshores diem que la seqüència $(1, \ldots, k)$ és un *camí* del graf.

Exemple 3.4. Si $A = \begin{bmatrix} 3 & 1 & 0 & 2 \\ -4 & 0 & 3 & 0 \\ 1 & 0 & 2 & 1 \\ 0 & 0 & 1 & 1 \end{bmatrix}$, aleshores el seu graf associat és



Observem a l'exemple 3.4 que el vèrtex $\{1\}$ és l'origen de dues arestes (corresponents als elements a_{12} i a_{14} de A), i el final d'altres dues (les corresponents als elements a_{21} i a_{31}). Un camí que uneix els vèrtexs $\{1\}$ i $\{2\}$ és, per exemple, $(1,2)$; Però també podria ser $(1,4,3,1,2)$. Aquest darrer camí, té un *cicle*, que és una seqüència $(1, \ldots, k, 1)$ de $k+1$ vèrtexs, on els vèrtexs per $i = 1, \ldots, k$ són diferents dos a dos, i les parelles $(1,2), (2,3), \ldots, (k-1,k), (k,1)$ són arestes del graf.
Si una matriu A té associat un graf sense cap cicle, i.e., és *no-cíclic*, aleshores A és semblant a una matriu triangular superior. De fet, el graf associat a una matriu triangular superior T, és no-cíclic.

Com que per a cada matriu tenim un graf dirigit associat, el nostre objectiu és expressar la matriu de manera que el seu graf associat sigui el més simple possible. El graf més simple de tots és aquell format únicament per vèrtexs, i sense cap aresta, el qual té associada una matriu diagonal (donat un element a_{ii} de la matriu, no existirà cap element diferent de zero ni en la seva fila ni en la seva columna). Veure exemple 3.5.

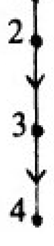
Exemple 3.5. Si $A = \begin{bmatrix} 3 & 0 & 0 & 2 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 2 \end{bmatrix}$, aleshores el seu graf associat és

Ara bé, com que no totes les matrius són semblants a una matriu diagonal, no sempre serà possible obtenir un graf d'aquest tipus. Com hem vist al Teorema 3.1, tota matriu A és semblant a una matriu triangular superior, i aquesta tindrà un graf associat no-cíclic. Així, aplicant permutacions per matrius, podem ordenar els vèrtexs del graf en ordre descendent en una columna, a dalt el vèrtex 1, després el vèrtex 2, i així successivament fins el vèrtex n, de manera que totes les arestes del graf sortiran d'un vèrtex $\{i\}$, i aniran a parar a un vèrtex $\{j\}$ tal que $j = i + d \leq n,\ d > 0$. Veure exemple 3.6.

Exemple 3.6. Si $T = \begin{bmatrix} 3 & -1 & 2 & 8 \\ 0 & 2 & 0 & 3 \\ 0 & 0 & 0 & 5 \\ 0 & 0 & 0 & 2 \end{bmatrix}$, aleshores el seu graf associat és

Dins d'aquests tipus de grafs existeixen els que anomenarem *grafs camí*, tals que si $(1, \ldots, k)$ són els seus vèrtexs, aleshores totes les parelles $(1,2), (2,3), \ldots, (k-1, k)$ són arestes del graf. Així, donat un graf camí, cada aresta sortirà d'un vèrtex $\{i\}$ i anirà a parar al vèrtex consecutiu $\{i+1\}$, i.e., si la seva matriu associada és $T = (t_{ij}) \in M_n$, aleshores $t_{i,j} \neq 0$, si $1 \leq i \leq n, j \in \{i, i+1\}$. Per tant, mitjançant semblances diagonals, T és similar a una matriu de Jordan, com es pot veure a l'exemple 3.7.

Exemple 3.7. Sigui $T = \begin{bmatrix} 3 & 2 & 0 & 0 \\ 0 & 1 & 7 & 0 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 4 \end{bmatrix} \sim \begin{bmatrix} 3 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 4 \end{bmatrix}$, aleshores $D(T)$ és



Així doncs, donada una matriu $A \in M_n$, el seu graf associat $D(A)$ serà del tipus graf camí, si i només si, A és semblant a una matriu B de la forma

$$B = \begin{bmatrix} \lambda_1 & 1 & & \\ & \lambda_2 & \ddots & \\ & & \ddots & 1 \\ & & & \lambda_n \end{bmatrix}.$$

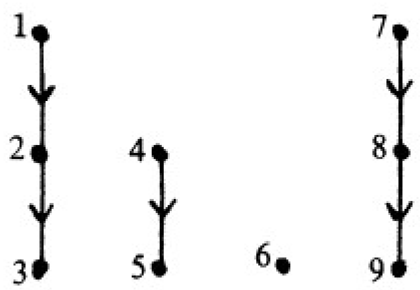
Ara bé, el graf associat a una matriu de Jordan, en general, no és un graf camí. A continuació presentem el Teorema de Jordan en llenguatge combinatori.

Teorema 3.8. *[3] (Forma canònica de Jordan en llenguatge combinatori)*

Tota matriu $A \in M_n$ és semblant a una matriu J (que és suma directa de matrius bàsiques de Jordan), tal que $D(J)$ és un conjunt de grafs camí sense vèrtexs en comú, i on cada graf camí està associat a un únic VAP de A.

Exemple 3.9. Veiem ara el graf associat a una matriu de Jordan.

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



4 Forma canònica de Weyr

Ens centrarem ara en la forma canònica de Weyr, i veurem als Teoremes 4.6 i 4.16 dues maneres de trobar-la de manera independent de la forma de Jordan.

Definició 4.1. *Una* ***matriu bàsica de Weyr amb VAP*** λ *és una matriu* $W \in M_n$ *tal que existeix una partició* $\omega_1 + \cdots + \omega_r = n$*, on* $\omega_1 \geq \cdots \geq \omega_r \geq 1$ *de manera que si mirem la matriu* W *com una matriu per blocs* $r \times r$*, els blocs* $W_{ij} \in M_{\omega_i,\omega_j}$ *han de complir:*

a) $W_{ii} = \lambda Id,\ i = 1, \ldots, r$
b) La primera diagonal superior, $W_{i,i+1} \in M_{\omega_i,\omega_{i+1}}$*, té rang complet per columnes i està formada per la matriu identitat* $Id \in M_{\omega_i}$*, seguida per* $\omega_i - \omega_{i+1}$ *files de zeros.*
c) La resta de blocs $W_{ij},\ j \neq i, i+1$ *són zero.*

En aquest cas, diem que la ***característica de Weyr*** *de la matriu* W *és* $(\omega_1, \ldots, \omega_r)$*.*

Exemple 4.2.

$$W = \left[\begin{array}{ccc|cc|c|c} \lambda & 0 & 0 & 1 & 0 & & \\ & \lambda & 0 & 0 & 1 & & \\ & & \lambda & 0 & 0 & & \\ \hline & & & \lambda & 0 & 1 & \\ & & & & \lambda & 0 & \\ \hline & & & & & \lambda & 1 \\ \hline & & & & & & \lambda \end{array}\right], \text{ té estructura } (3,2,1,1).$$

Observacions:
a) Una matriu bàsica de Weyr amb estructura $(1, \ldots, 1)$ és una matriu bàsica de Jordan.
b) Diem que l'estructura de Weyr d'una matriu és **homogènia** si $\omega_1 = \cdots = \omega_r$.
c) Per definició, una matriu de Weyr amb un únic VAP és el mateix que una matriu bàsica de Weyr.

D'acord amb la definició que hem donat, no és cert que una matriu de Weyr sigui igual a una suma directa de matrius bàsiques de Weyr amb un sol VAP, com succeïa amb les matrius de Jordan. Per tant passem a definir què és una matriu de Weyr.

Definició 4.3. *Siguin* $W \in M_n(\mathbb{K})$ *i* $\lambda_1, \ldots, \lambda_k \in \mathbb{K}$*, els diferents VAPs de* W*. Aleshores, diem que* W *és una* ***matriu de Weyr*** *si* W *és suma directa de matrius bàsiques de Weyr (cadascuna amb un VAP* λ *diferent), i.e.,*

$$W = \begin{bmatrix} W_1 & & & \\ & W_2 & & \\ & & \ddots & \\ & & & W_k \end{bmatrix},$$

on cada bloc $W_i,\ i = 1, \ldots, k$ *és una matriu bàsica de Weyr amb VAP* λ_i*.*

Exemple 4.4.

$$W = \left[\begin{array}{cc|c|c|cc|cc}
2 & 0 & 1 & & & & & \\
 & 2 & 0 & & & & & \\
\hline
 & & 2 & 1 & & & & \\
\hline
 & & & 2 & 1 & 0 & & \\
\hline
 & & & & -1 & 0 & 1 & 0 \\
 & & & & 0 & -1 & 0 & 1 \\
\hline
 & & & & & & -1 & 0 \\
 & & & & & & 0 & -1
\end{array}\right]$$

4.1 Primera demostració de l'existència i unicitat de la forma canònica de Weyr

La demostració d'aquest teorema és constructiva i consta de dues etapes. A la primera, trobarem bases del nucli i després les ampliarem a tot l'espai. A la segona etapa, utilitzant operacions elementals per files, reduirem les submatrius a la forma desitjada.

Observació 4.5. Com a conseqüència del Teorema 2.3 sabem que tota matriu $A \in M_n(\mathbb{K})$ es pot reduir a una matriu diagonal per blocs de la forma

$$A = \begin{bmatrix} A_1 & & \\ & \ddots & \\ & & A_k \end{bmatrix} = diag(A_1, \ldots, A_k),$$

amb $Spec(A_i) \cap Spec(A_j) = \emptyset,\ i \neq j$. Per tant, com a conseqüència del teorema de Sylvester, ens podem reduir a calcular la forma canònica associada a cada bloc per separat, i en particular, com que $S^{-1}(A_i - \lambda_i Id)S = S^{-1}A_iS$, podem suposar que A és nilpotent. Recordem a més que la característica de Weyr $(\omega_1, \ldots, \omega_r)$ d'una matriu $A \in M_n$ es calcula com $\omega_i = ker(A - \lambda Id)^i - ker(A - \lambda Id)^{i-1}$.

Per la observació 4.5, podem suposar que A és nilpotent.

Teorema 4.6. *Tota matriu quadrada $A \in M_n(\mathbb{K})$ és semblant a una matriu en forma de Weyr.*

Demostració. Demostrem l'existència.
Siguin $d = nul(A) = dim\ ker(A)$, $V = \mathbb{K}^n$ i $\mathcal{B}$ la base canònica de V. Podem veure A com la matriu associada a la base $\mathcal{B}$ de la transformació lineal sobre V, donada per la multiplicació per l'esquerra. Així, escollim una base del nucli de A i l'estenem a una base de l'espai V, que anomenarem $\mathcal{B}'$. Fent el canvi de base, obtenim que A és semblant a la matriu

$$P^{-1}AP = \left[\begin{array}{c|c} 0 & B_2 \\ \hline 0 & A_2 \end{array}\right],$$

on P és la matriu de canvi de base corresponent, i les dimensions dels blocs de la diagonal són $d \times d$ i $(n-d) \times (n-d)$. Sigui $m = n - d$, observem que A_2 és una matriu nilpotent amb dimensions $m \times m,\ m < n$. Fem inducció sobre n: suposem que A_2 està en forma de Weyr, sigui $Q \in GL_m(K) \mid Q^{-1}A_2Q = W$ és una matriu de Weyr nilpotent. Per tant, conjugant $P^{-1}AP$ i $diag(Id, Q)$, obtenim la matriu per blocs

$$X = \left[\begin{array}{c|c} 0 & Y \\ \hline 0 & W \end{array}\right],$$

on Y i W són matrius de Weyr nilpotents. Sigui $\omega_1 = d$, i sigui $(\omega_2, \ldots, \omega_r)$ l'estructura de Weyr de W, observem també que $\omega_2 = nul(W)$. Com que $rang(X) \leq rang(W) + rang(Y) \Rightarrow \omega - \omega_1 \leq (\omega - \omega_1 - \omega_2) + \omega_1 \Rightarrow \omega_1 \geq \omega_2$. Partim ara les nostres matrius $n \times n$ amb la partició $n = \omega_1 + \cdots + \omega_r$, així X té la forma

$$\begin{bmatrix} 0 & X_{12} & X_{13} & \ldots & X_{1r} \\ & 0 & I_3 & 0 & \ldots \\ & & 0 & \ddots & \\ & & & \ddots & I_r \\ & & & & 0 \end{bmatrix},$$

on cada I_j representa la matriu $\omega_{j-1} \times \omega_j$ formada per la matriu identitat ω_j-dimensional seguida en la seva part superior per $\omega_{j-1} - \omega_j$ files de zeros, i on $Y = [X_{12}, \ldots, X_{1r}]$. Fent servir que les primeres ω_1 columnes de X són zero, podem convertir en zero els elements $X_{1i},\ i = 1, \ldots, r$ conjugant X amb matrius elementals adequades que no modifiquin el bloc W. Però encara hem de transformar X_{12} en I_2: $n - \omega_1 = rang(A) = rang(X) = rang(X_{12}) + rang(I_3) + \cdots + rang(I_r) = rang(X_{12}) + \omega_3 + \omega_4 + \cdots + \omega_r = rang(X_{12}) + n - \omega_1 - \omega_2 \Rightarrow rang(X_{12}) = \omega_2$, i.e., X_{12} és una matriu amb dimensions $\omega_1 \times \omega_2$ amb rang complet per columnes, això implica que la podem reduir a la matriu I_2. Com que les primeres ω_1 columnes de la matriu X, transformada, són zero podem conjugar X amb una matriu elemental $n \times n$ adequada per convertir X en I_2 sense modificar els altres elements de X. Així, hem transformat la nostra matriu X en una matriu en forma de Weyr amb estructura $(\omega_1, \ldots, \omega_r)$.

Demostrem ara la unicitat.
De fet és suficient demostrar que dues matrius bàsiques de Weyr amb un mateix VAP i que són semblants, han de ser iguals. Suposem que tenim dues matrius de Weyr, $W = diag(W_1, \ldots, W_k)$ i $W' = diag(W'_1, \ldots, W'_k)$, formades per k blocs bàsics de Weyr cadascun, i semblants a la mateixa matriu A. Permutant els blocs, si fos necessari, tenim que cada parella de blocs W_i i W'_i tenen associat el mateix VAP λ_i. Volem veure que $W = W'$, però com que són matrius per blocs semblants, on cada bloc té associat el mateix VAP, W_i és semblant a W'_i. Per finalitzar, és suficient veure que dos matrius bàsiques de Weyr que són semblants, amb mateix VAP λ, han de tenir la mateixa estructura de Weyr, i.e., han de ser iguals. Això és el que veurem a la Proposició 4.7, on veurem que l'estructura d'una matriu bàsica de Weyr amb VAP λ està completament determinada per les nul·litats de les potències de $W - \lambda Id$. Això és anàleg a determinar l'estructura de Jordan d'una matriu amb un únic VAP. En canvi, la demostració fent servir les nul·litats de matrius bàsiques de Weyr són molt més senzilles d'establir. □

Proposició 4.7. *Sigui $W \in M_n$ una matriu bàsica de Weyr, amb VAP associat λ, i estructura $(\omega_1, \ldots, \omega_r)$, aleshores*

r = *índex de nilpotència de la matriu* $W - \lambda Id = mín\{r \in \mathbb{Z}$ *t.q.* $A^r = 0\}$,

$\omega_1 = nul(W - \lambda Id)$,

$\omega_i = nul(W - \lambda Id)^i - nul(W - \lambda Id)^{i-1}, \ \forall \ i = 2, \ldots, r.$

Per tant, dues matrius bàsiques de Weyr que són semblants han de ser iguals.

Demostració. Sigui $N = W - \lambda Id$ aleshores veiem la matriu N i les seves potències com matrius per blocs $r \times r$ amb partició $n = \omega_1 + \cdots + \omega_r$. Sigui I_j una matriu amb ω_j columnes formada per la matriu identitat, $Id \in M_{\omega_j \times \omega_j}$, a la part superior seguida de files de zeros. Observem que la dimensió de I_j dependrà del bloc en què estem treballant. Així,

$$N = \begin{bmatrix} 0 & I_2 & 0 & \ldots & 0 \\ & 0 & I_3 & \ldots & 0 \\ & & \ddots & & \\ & & & 0 & I_r \\ & & & & 0 \end{bmatrix}, N^i = \begin{bmatrix} 0 & \ldots & 0 & I_{i+1} & 0 & \ldots & 0 \\ & & & 0 & I_{i+2} & 0 & \vdots \\ & & & & \ddots & \ddots & \\ & & & & & 0 & I_r \\ & & & & & & 0 \\ & & & & & & \vdots \\ & & & & & & 0 \end{bmatrix}, \ \forall i = 1, \ldots, r-1.$$

Òbviament N té índex de nilpotència r. Observem que per $i = 1, \ldots, r-1$ tenim que $rang(N^i) = \omega_{i+1} + \omega_{i+2} + \cdots + \omega_r$, i.e., $\omega_i = rang(N^{i-1}) - rang(N^i) = nul(N^i) - nul(N^{i-1})$, que es satisfà també quan $i = r$. Ara bé, com que matrius semblants tenen la mateixa nul·litat, i sabem que si dos matrius X *i* Y són semblants, aleshores $(X - \lambda Id)^i$ i $(Y - \lambda Id)^i$ també són semblants (sota la mateixa transformació de semblances), i ja tenim el que volíem. $\square$

Com a conseqüència de la Proposició 4.7, tenim el següent corol·lari:

Corol·lari 4.8. *Dues matrius* A *i* $B \in M_n(\mathbb{K})$ *són semblants* $\Leftrightarrow$ *tenen els mateixos VAPs i la mateixa característica (o estructura) de Weyr, associades als VAPs corresponents.*

Així doncs, com a conseqüència del Teorema 4.6 i de la Proposició 4.7, tenim que tota matriu quadrada sobre un cos $\mathbb{K}$ és **semblant a una única matriu de Weyr, W**, llevat de permutacions, que anomenarem *forma canònica de Weyr de la matriu A*.

Si definim $\omega_i = nul(A - \lambda Id)^i - nul(A - \lambda Id)^{i-1}$, per $i = 1, 2, \ldots$, aleshores ens referirem a la partició $\omega_1, \omega_2, \ldots$ com la **característica de Weyr de A associada al VAP** λ, a més el conjunt finit de termes de la característica de Weyr, $\{\omega_1, \ldots, \omega_r\}, \ \omega_i \neq 0$, on $r = min\{j \mid nul(A - \lambda Id)^j = nul(A - \lambda Id)^{j+1}\}$, coincideix amb l'estructura de Weyr de la matriu A associada al VAP λ (per la proposició vista anteriorment).

Corol·lari 4.9. *Sigui* $A \in M_n(\mathbb{K})$ *nilpotent, amb* $nul(A) = d$ *i* $m = n - d$. *Per tant, podem escriure* A *com*

$$\left[\begin{array}{c|c} 0 & B \\ \hline 0 & C \end{array}\right],$$

on $C \in M_m(\mathbb{K})$, *i* $0 \in M_d(\mathbb{K})$.

Llavors, si $(\omega_2, \dots, \omega_r)$ *és l'estructura de Weyr de la matriu nilpotent* C, *aleshores l'estructura de Weyr de la matriu* A *és* $(d, \omega_2, \dots, \omega_r)$.

Proposició 4.10. *Dues matrius* $A, B \in M_n$ *són semblants, si i només si, tenen els mateixos VAPs i* $nul(A - \lambda Id)^j = nul(B - \lambda Id)^j$, *per a cada VAP* λ *i per* $j = 1, ..., n$.

Demostració. La implicació cap a la dreta és immediata ja que són condicions necessàries per tal que les dues matrius siguin semblants. Veiem ara l'altra implicació: Suposem que es compleixen les hipòtesis i anem a demostrar que aleshores han de tenir la mateixa forma de Weyr. Siguin $\lambda_1, ..., \lambda_k$ els diferents VAPs de A i B. Aleshores sabem que

$$A = diag(A_1, ..., A_k), B = diag(B_1, ..., B_k),$$

on A_i i B_i tenen com a únic VAP λ_i. Per tant, és suficient veure que A_i i B_i tenen la mateixa estructura de Weyr. Com que $(A - \lambda Id)^j = diag((A_1 - \lambda_1 Id)^j, *, ..., *)$, on * són blocs invertibles (per $i \geq 2$, el bloc i-èssim té el terme $(\lambda_i - \lambda_1)^j \neq 0$ com a únic VAP), veiem que $nul(A - \lambda_1 Id)^j = nul(A_1 - \lambda_1 Id)^j$, per tota j. Igualment, $nul(B - \lambda_1 Id)^j = nul(B_1 - \lambda_1 Id)^j$. Per tant, per les nostres hipòtesis tenim que $nul(A_1 - \lambda_1 Id)^j = nul(B_1 - \lambda_1 Id)^j$, $j = 1, ..., n$. En particular, com que $A_1 - \lambda_1 Id$ i $B_1 - \lambda_1 Id$ són matrius nilpotents, implica que A_1 i B_1 han de tenir les mateixes dimensions. (Si $N \in M_m$ és nilpotent, sabem pel Teorema de Cayley-Hamilton que $N^m = 0$ i per tant $N^k = 0, k \geq m$). Així doncs, A_1 i B_1 han de tenir la mateixa estructura de Weyr. □

Observació 4.11. La transposada d'una matriu quadrada $A \in M_n$ i A són semblants, ja que ambdues tenen els mateixos VAPs i la mateixa nul·litat (rang per columnes és igual al rang per files). A més, $(A^T - \lambda Id)^i = ((A - \lambda Id))^i)^T$, ja que les potències $(A - \lambda Id)^i$ i $(A^T - \lambda Id)^i$ tenen la mateixa nul·litat.

4.2 Segona demostració de l'existència i unicitat de la forma canònica de Weyr

A [4] trobem una demostració de l'existència i unicitat de la forma de Weyr diferent. Aquesta demostració és independent de la forma de Jordan. En primer lloc, veurem que ens podem reduir al cas nilpotent i a partir de la característica de Weyr obtenim la forma canònica de Weyr.

Denotarem les matrius triangulars per blocs amb diagonal $A_1, \dots, A_k$ com

$$\mathfrak{J}(A_1, \dots, A_k) = \begin{bmatrix} A_1 & \dots & A_{1k} \\ & \ddots & \vdots \\ & & A_k \end{bmatrix}.$$

A continuació presentem dos lemes que ens permeten obtenir la característica de Weyr d'una matriu nilpotent $A \in M_n$, sense necessitat de calcular les successives potències d'aquesta.

Lema 4.12. *Sigui $A \in M_n(\mathbb{K})$ una matriu de la forma*

$$A = \begin{bmatrix} 0 & A_{12} \\ 0 & A_2 \end{bmatrix} = \mathfrak{J}(0_{\omega_1}, A_2), \text{ on } \omega_1 = nul(A).$$

Sigui $X \in \mathbb{K}^n$ tal que $X = \begin{bmatrix} X_1 \\ X_2 \end{bmatrix}$, on $X_1 \in \mathbb{K}^{\omega_1}$, i $X_2 \in \mathbb{K}^{n-\omega_1}$. Aleshores, per a qualsevol enter r, $A^r X = 0$, si i només si, $A_2^{r-1} X_2 = 0$.

Demostració. Observem que

$A^r = \begin{bmatrix} 0 & A_{12}A_2^{r-1} \\ & A_2^r \end{bmatrix} \Rightarrow A^r X = \begin{bmatrix} A_{12} \\ A_2 \end{bmatrix} \begin{bmatrix} A_2^{r-1} & X_2 \end{bmatrix}$. Com que $rang(A) = n - \omega_1$, la matriu $\begin{bmatrix} A_{12} \\ A_2 \end{bmatrix}$ té columnes linealment independents, i per tant

$$\begin{bmatrix} A_{12} \\ A_2 \end{bmatrix} Y = 0 \Leftrightarrow Y = 0.$$

Si escrivim $Y = A_2^{r-1} X_2$, tenim que $A^r X = 0 \Leftrightarrow A_2^{r-1} X_2 = 0$.

□

Lema 4.13. *Sigui $A = \mathfrak{J}(0_{\omega_1}, A_2) \in M_n(\mathbb{K})$ una matriu nilpotent, amb característica de Weyr $\omega(A) = (\omega_1, \ldots, \omega_k)$. Aleshores, $\omega(A_2) = (\omega_2, \ldots, \omega_k)$, amb $\omega_1 \geq \omega_2 \geq \cdots \geq \omega_k$.*

Demostració. Pel Lema 4.12, $nul(A^i) = \omega_1 + nul(A_2^{i-1})$, així que per a cada $i \geq 2$ tenim que $nul(A_2^{i-1}) - nul(A_2^{i-2}) = nul(A^i) - nul(A^{i-1}) = \omega_i$. Per tant, $\omega(A_2) = (\omega_2, \ldots, \omega_k)$. Provem ara que $\omega_{k+1} \leq \omega_k$ per inducció sobre k. Per $k = 2$, $n - \omega_1 = rang(A) \leq rang(A_{12}) + rang(A_2) = [(n - \omega_1) - nul(A_2)] + rang(A_2) \Rightarrow nul(A_2) \leq rang(A_{12})$ Però com que $\omega_2 = nul(A_2)$, i $rang(A_{12}) \leq \omega_1$, tenim que $\omega_2 \leq \omega_1$.

Per hipòtesis d'inducció, és cert per A_2 i per tant $\omega_{i+1} \leq \omega_i$, $i \geq 2$. □

Per tant, dels lemes 4.12 i 4.13 es deriva el següent algoritme recursiu pel càlcul de la característica de Segre d'una matriu nilpotent.

Lema 4.14. *Sigui f un endomorfisme nilpotent sobre E, amb $\omega(f) = (\omega_1, \ldots, \omega_r)$. Aleshores, f pot ser representat per la matriu $A = \mathfrak{J}(0_{\omega_1}, 0_{\omega_2}, \hat{A})$, on $rang(A_{12}) = \omega_2$ i A_{12} té rang complet per columnes.*

Demostració. Com que $\omega_1 = nul(A)$, podem representar la nostra matriu A com una matriu $B = \mathfrak{J}(0_{\omega_1}, 0_{\omega_2}, \hat{A})$. Pel segon lema d'abans, $\omega_2 = nul(B_2)$ així que existeix una matriu quadrada Q de dimensió $n - \omega_1$ tal que $Q^{-1} B_2 Q = \mathfrak{J}(0_{\omega_2}, \hat{A})$. Sigui ara $P = diag(Id_{\omega_1}, Q)$, aleshores tenim que $P^{-1} B P = \mathfrak{J}(0_{\omega_1}, 0_{\omega_2}, \hat{A})$, així que $A = \mathfrak{J}(0_{\omega_1}, 0_{\omega_2}, \hat{A})$ és la matrius associada a f i és de la forma

$$A = \begin{bmatrix} 0_{\omega_1} & A_{12} & A_{13} \\ & 0_{\omega_2} & A_{23} \\ & & \hat{A} \end{bmatrix},$$

i com que $rang(A) = n - \omega_1$, les darreres $n - \omega_1$ columnes de A han de ser linealment independents, i per tant el bloc A_{12} que té dimensió $\omega_1 \times \omega_2$ té rang complet per columnes. $\square$

Observem que si $\mathbb{K} = \mathbb{C}$, aleshores fent servir una base ortonormal per $\mathbb{C}^n$, on els primers ω_1 vectors siguin una base del nucli de f, podem fer servir una matriu unitària Q i obtenir una representació de f en la forma abans descrita utilitzant una base ortonormal adient.

Teorema 4.15. *Sigui $f : E \to E$ un endomorfisme nilpotent. Aleshores, $\omega(f) = (\omega_1, \ldots, \omega_r)$, si i només si, f es pot representar com una matriu triangular per blocs de la forma $A = \mathfrak{J}(0_{\omega_1}, \ldots, 0_{\omega_r})$, on cada bloc a la diagonal superior té rang complet per columnes, i.e., $rang(A_{i,i+1}) = \omega_{i+1}$.*

Demostració. Ho demostrarem per inducció sobre k. Sigui $\omega(f) = (\omega_1, \ldots, \omega_k)$, aleshores si $k = 1$, f té matriu 0. Si $k = 2$, pel Lema 4.14 tenim que és cert, i pel cas general, aplicant també el Lema 4.14 veiem que f té matriu associada $B = \mathfrak{J}(0_{\omega_1}, 0_{\omega_2}, \hat{B})$, on el bloc B_{12} té rang complet per columnes. Sigui B_2 la submatriu quadrada en les darreres $n - \omega_1$ files i columnes, aleshores $B_2 = \mathfrak{J}(0_{\omega_2}, \hat{B})$. Pel Lema 4.13, $\omega(B_2) = (\omega_2, \ldots, \omega_r)$, i per hipòtesis d'inducció, existeix una matriu regular Q amb dimensió $n - \omega_1$ tal que $Q^{-1} B_2 Q = \mathfrak{J}(0_{\omega_2}, \ldots, 0_{\omega_r})$ on cada bloc per sobre de la diagonal principal té rang complet per columnes. Si apliquem la semblança $P = diag(Id_{\omega_1}, Q)$ a B, tenim la matriu A amb la forma que volíem.

Demostrem ara la implicació contrària, és suficient veure que $A = \mathfrak{J}(0_{\omega_1}, \ldots, 0_{\omega_r})$ amb blocs a la primera diagonal superior amb rang complet per columnes té característica de Weyr $(\omega_1, \ldots, \omega_k)$. Ho farem per inducció sobre k. Observem que les darreres $n - \omega_1$ columnes d'aquesta matriu són linealment independents, per tant $nul(A) = \omega_1$. Si $k = 1$, $A = 0$ i ja ho tenim. Per altra banda, $A = \mathfrak{J}(0_{\omega_1}, A_2)$ i el Lema 4.13 ens diu que la seva característica de Weyr és $(\omega_1, \omega'_2, \ldots, \omega'_r)$, on $(\omega'_2, \ldots, \omega'_r) = \omega(A_2)$, i per hipòtesis d'inducció tenim que $\omega_i = \omega'_i$, per $i \geq 2$. $\square$

Finalment, anem a obtenir la forma canònica de Weyr pel cas nilpotent.

Teorema 4.16. ***[4] (Forma canònica de Weyr)*** *Sigui $f : E \to E$ un endomorfisme nilpotent. Aleshores, $\omega(f) = (\omega_1, \ldots, \omega_r)$, si i només si, f es pot representar per la matriu triangular per blocs $W = \mathfrak{J}(0_{\omega_1}, \ldots, 0_{\omega_r})$, on els únics blocs diferents de zero són els de la primera diagonal superior i són de la forma $W_{i,i+1} = Id_{\omega_i, \omega_{i+1}}$, $i = 1, \ldots, r-1$.*

Demostració. Fent servir el Teorema 4.15, és suficient veure que la matriu $B = \mathfrak{J}(0_{\omega_1}, \ldots, 0_{\omega_r})$, on cada bloc a la primera diagonal superior té rang complet per columnes, és semblant a W. Fem-ho per inducció sobre k. Si $k = 1$, tenim que $B = 0$ i ja ho tindríem. Sigui ara $k > 1$, aleshores la submatriu formada per les darreres $n - \omega_1$ files i columnes de B té característica de Weyr $(\omega_2, \ldots, \omega_r)$, i per hipòtesis d'inducció tenim que és semblant a una matriu amb la forma desitjada, i.e., existeix una matriu quadrada invertible Q, amb dimensió $n - \omega_1$, tal que $C = diag(Id_{\omega_1}, Q^{-1})\ B\ diag(Id_{\omega_1}, Q)$ té la forma que volíem excepte, pot ser, en la primera fila de blocs, i.e.,

$$C = \begin{bmatrix} 0_{\omega_1} & C_{12} & \dots & C_{1r} \\ & 0_{\omega_2} & Id_{\omega_2,\omega_3} & 0 \\ & & \ddots & Id_{\omega_{r-1},\omega_r} \\ & & & 0_{\omega_r} \end{bmatrix}.$$

Així, $nul(B) = nul(C) = \omega_1$, i per tant, C_{12} té rang complet per columnes. Anem a transformar en 0 els blocs $C_{13}, \dots, C_{1k}$.

El bloc C_{1r} té dimensions $\omega_1 \times \omega_r$, sigui doncs $\tilde{C}_{1r}$ la matriu $\omega_1 \times \omega_{r-1}$ obtinguda afegint $\omega_{r-1} - \omega_r$ columnes de zeros a la matriu C_{1r}. Per tant, $\tilde{C}_{1r} = \begin{bmatrix} C_{1r} & 0_{\omega_1,\omega_{r-1}-\omega_r} \end{bmatrix}$, i $\tilde{C}_{1r} Id_{\omega_{r-1},\omega_r} = C_{1r}$. Sigui ara P la matriu de la forma $\mathfrak{J}(Id_{\omega_1}, Id_{n-\omega_1})$, on les primeres ω_1 files són els blocs $(Id_{\omega_1}, \tilde{C}_{13}, \dots, \tilde{C}_{1k}, 0_{\omega_1,\omega_r})$, i.e.,

$$P = \left[\begin{array}{c|c} Id_{\omega_1} & \tilde{C}_{13}\tilde{C}_{14} \dots \tilde{C}_{1r} 0_{\omega_1,\omega_r} \\ \hline & Id_{n-\omega_1} \end{array}\right].$$

Aleshores, P^{-1} té la mateixa forma però els blocs de les primeres ω_1 files són els blocs $(Id_{\omega_1}, -\tilde{C}_{13}, -\tilde{C}_{14}, \dots, -\tilde{C}_{1r}, 0_{\omega_1,\omega_r})$. Operant per blocs es pot veure que $P^{-1}CP$ té la forma desitjada però té la matriu C_{12} en el seu bloc $(1,2)$. Ara bé, com que C_{12} té rang complet per columnes, existeix una matriu invertible $W \in M_{\omega_1}$ tal que $WC_{12} = Id_{\omega_1,\omega_2}$, i si anomenem $S = diag(W^{-1}, Id_{\omega_2}, \dots, Id_{\omega_r})$, la matriu $S^{-1}P^{-1}CPS$ té la forma desitjada. $\square$

4.3 Calculem la forma de Weyr d'una matriu

Sigui $A \in M_n$, per calcular la seva forma canònica de Weyr seguirem els següents passos.

Pas 1. Anomenem $A_1 = A$.

Pas 2. Si $A_1 \neq 0$, trobem una base del nucli de A_1 (amb operacions elementals sobre les files) i la estenem a una base de $\mathbb{K}^n$. Sigui $P_1 \in M_n$ tal que té els últims vectors de la base com a columnes. Aleshores,

$$P_1^{-1} A_1 P_1 = \left[\begin{array}{c|c} 0 & B_2 \\ \hline 0 & A_2 \end{array}\right],$$

on A_2 és una matriu quadrada de dimensió $n - nul(A_1)$.

Pas 3. Si $A_2 \neq 0$, repetim el pas 2 amb la matriu A_2 i obtenim una matriu invertible P_2 de dimensió $n - nul(A_1)$ tal que

$$P_2^{-1} A_2 P_2 = \left[\begin{array}{c|c} 0 & B_3 \\ \hline 0 & A_3 \end{array}\right],$$

on A_3 és una matriu quadrada de dimensió $n - nul(A_1) - nul(A_2)$. Successivament, construïm matrius quadrades $A_1, A_2, \dots$ cada vegada amb dimensió més petita i les matrius invertibles associades $P_1, P_2, \dots$ (amb dimensions adients) fins que aparegui la primera matriu zero (que ha de ser així perquè A és nilpotent). Aleshores l'estructura de Weyr de la matriu A és $(\omega_1, ..., \omega_r)$, on $\omega_i = nul(A_i)$, per $i = 1, ..., r$.

Pas 4. Conjugant A pel producte de les matrius $n \times n$ invertibles: $P_1, diag(Id, P_2), ...,$ $diag(Id, P_{r-1})$, i triant dimensions adients per les matrius Id, tenim la següent matriu $r \times r$ per blocs

$$\begin{bmatrix} 0 & X_{12} & X_{13} & \dots & X_{1r} \\ & 0 & X_{23} & \dots & X_{2r} \\ & & \ddots & \ddots & \\ & & & 0 & X_{r-1,r} \\ & & & & 0 \end{bmatrix},$$

on cada $X_{ij} \in M_{\omega_i,\omega_j}$ és una matriu tal que els seus blocs zero que són a la diagonal tenen dimensions $\omega_1 \times \omega_1, \dots, \omega_r \times \omega_r$. La primera diagonal superior té rang complet per columnes, i.e., rang$(X_{i,i+1}) = \omega_{i+1}$, per $i = 1, ..., r-1$.

Pas 5. Fent servir operacions elementals sobre les files, obtenim $Y_{r-1} \in GL_{\omega_{r-1}}(\mathbb{K})$ tal que

$$Y_{r-1}X_{r-1,r} = \begin{bmatrix} Id \\ 0 \end{bmatrix},$$

on $Id \in M_{\omega_r}$. Conjugant X per $Q_1 = diag(Id, Id, ..., Y_{r-1}^{-1}, Id)$ per aconseguir convertir $X_{r-1,r}$ en

$$Id_{r-1,r} = \begin{bmatrix} Id \ 0 \end{bmatrix},$$

la matriu $\omega_{r-1} \times \omega_r$ que té la matriu identitat amb dimensions $\omega_r \times \omega_r$ a la part superior i zeros a la resta de files.

Així preservem la forma de X (els únics blocs que s'han transformat són a la columna $r-1$, i fent servir conjugacions per un producte R_1 de matrius elementals per eliminar els blocs anteriors $Id_{r-1,r}$ en la darrera columna de blocs. Així preservem la forma de X (on només canvia la columna r-èssima). És possible escriure R_1 de manera explícita:

$$R_1 = Id + \begin{bmatrix} 0 & \dots & 0 & T_{1r} & 0 \\ 0 & \dots & 0 & T_{2r} & 0 \\ \vdots & & \vdots & \vdots & \vdots \\ 0 & \dots & 0 & T_{r-2,r} & 0 \\ 0 & \dots & 0 & 0 & 0 \end{bmatrix},$$

on T_{ir} és el bloc X_{ir} afegint $\omega_{r-1} - \omega_r$ columnes zero.

Pas 6. Repetim el pas 5 en la columnes $r-1$ convertint el bloc $(r-2, r-1)$ en $Id_{r-2,r-1}$ via alguna conjugació amb $Q_2 \in GL_n(\mathbb{K})$. Fent servir aquest bloc ïdentitat"per eliminar els blocs anteriors, via conjugació pel producte R_2 de matrius elementals. Això no canvia la darrera columna dels blocs i preserva la forma de X. Repetint aquest procés en les columnes $r-2, r-3, ..., 3, 2$ fent servir conjugacions per $Q_3, R_3, ..., Q_{r-2}, R_{r-2}, Q_{r-1}$. La matriu resultant W està en la forma de Weyr. Sigui

$$S = P_1 diag(Id, P_2) \dots diag(Id, P_{r-1}) Q_1 R_1 Q_2 \dots R_{r-2} Q_{r-1}.$$

Aleshores, $W = S^{-1}AS$ ens dona explícitament la transformació per semblances que desitgem. Com en el pas 5, podem escriure explícitament les matrius $R_2, R_3, ...$ ■

Observació 4.17. A la pràctica, depenent de les propietats de la nostra matriu pot ser no caldrà realitzar tots els passos del primer fins el sisè. Només hi ha dos fases que ens caldrà recordar:

(1) Transformar la matriu A en una matriu estrictament triangular superior per blocs $X = (X_{ij})$ tal que els blocs de la diagonal decreixen (la seva dimensió), i la primera diagonal superior per blocs $X_{12}, ..., X_{r-1,r}$ tingui rang complet per columnes.

(2) Començant per la darrera columna de blocs i treballant çap enrereçonvertim la primera diagonal superior per blocs en una matriu ïdentitat", i eliminem els blocs superiors.

Exemple 4.18. Sigui la matriu nilpotent

$$A = \begin{bmatrix} 2 & -1 & 3 & 1 \\ 1 & -1 & 2 & 1 \\ -2 & 1 & -3 & -1 \\ 3 & -2 & 5 & 2 \end{bmatrix}$$

i volem calcular la seva forma de Weyr.

1.- $A_1 = A$.

2.- Amb operacions elementals a les files tenim que:

$$A_1 \to \begin{bmatrix} 1 & -1 & 2 & 1 \\ 0 & 1 & -1 & -1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

Així doncs, $nul(A_1) = 2$, i el conjunt $\{(-1, 1, 1, 0), (0, 1, 0, 1), (0, 0, 1, 0), (0, 0, 0, 1)\}$ és una base de $\mathbb{K}^4$ on els dos primers vectors formen una base del nucli de A_1, i posant aquests vectors en columnes obtenim la matriu invertible

$$P_1 = \begin{bmatrix} -1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{bmatrix}.$$

Per tant, $P_1^{-1} A_1 P_1 = \left[\begin{array}{cc|cc} 0 & 0 & -3 & -1 \\ 0 & 0 & 5 & 2 \\ \hline 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right].$

Com que el bloc inferior dretà A_2 de la transformació de A_1 és el bloc zero, podem deduir que l'estructura de Weyr de la matriu A és $(2, 2)$ i així podem escriure la forma de Weyr de A.

Si ens fixem en el nostre algoritme, els passos 3 i 4 no són necessaris, per tant passem al 5.

5.- Sigui X_{12} el bloc $(1, 2)$ de la matriu $X = P_1^{-1} A_1 P_1$. Com que $rang(X_{12}) = 2$, podem convertir $X_{12} \rightarrow Id$ fent servir la conjugació per

$$Q_1 = diag(X_{12}, Id) = \left[\begin{array}{cc|cc} -3 & -1 & 0 & 0 \\ 5 & 2 & 0 & 0 \\ \hline 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{array}\right],$$

Per tant, conjugant A per la matriu invertible

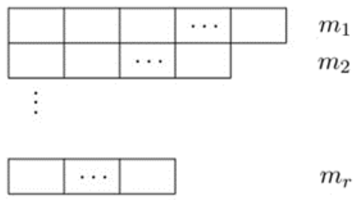
$$S = P_1 Q_1 = \begin{bmatrix} 3 & 1 & 0 & 0 \\ 2 & 1 & 0 & 0 \\ -3 & -1 & 1 & 0 \\ 5 & 2 & 0 & 1 \end{bmatrix}.$$

ens dona la forma de Weyr de la matriu A:

$$W = S^{-1} A S = \left[\begin{array}{cc|cc} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ \hline 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right].$$

5 Relació entre la forma canònica de Jordan i la forma canònica de Weyr

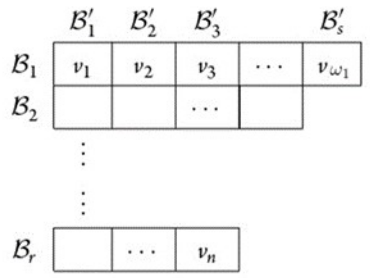
Definició 5.1. *Donada una partició* $(m_1, \ldots, m_r)$ *d'una matriu tal que* $n = m_1 + \cdots + m_r$*, es defineix el* ***Diagrama de Ferrer*** *com*



Sigui $N \in M_r$ una matriu tal que a la seva primera diagonal superior té uns i a la resta zeros. Aleshores, és nilpotent amb índex de nilpotència r. Si $J = diag(J_1, \ldots, J_r)$ és la seva forma canònica de Jordan, i $(m_1, \ldots, m_s)$ és la seva característica de Segre, es compleix que per a cada bloc J_i la seva nul·litat és igual a 1, i per tant, $nul(A) = t$. Així, si $\omega(A) = (\omega_1, \ldots, \omega_s)$, aleshores $\omega_1 = s$ és el nombre de blocs de la forma de Jordan de A. $nul(J^2) = t + m_2$, on m_2 és el nombre de blocs de dimensió 2 o superior. Iterant, tenim que calculant $nul(J^k)$, ω_k és exactament el nombre de blocs de Jordan amb dimensió k o superior. Si viem la les característiques de Weyr i de Segre com particions de n, tenim que la característica de Weyr és la partició conjugada de la característica de Segre, i per tant podem obtenir una en funció de l'altra.

Teorema 5.2. *Les estructures de Weyr i de Jordan d'una matriu nilpotent* $A \in M_n$ *(més en general, una matriu amb un únic VAP) són les particions conjugades de* n*. A més, la forma de Weyr i la forma de Jordan són matrius quadrades conjugades sota una certa transformació per permutacions.*

Demostració. Suposem que A és una matriu de Weyr nilpotent amb estructura $(\omega_1, ..., \omega_r)$. Si veiem la matriu com la transformació $f : \mathbb{K}^n \to \mathbb{K}$ corresponent a una certa base ordenada $\mathcal{B} = \{v_1, ..., v_n\}$. Escrivim aquesta base com $\mathcal{B} = \mathcal{B}_1 \cup \mathcal{B}_2 \cup \cdots \cup \mathcal{B}_r$, on $\mathcal{B}_1 = \{v_1, ..., v_{\omega_1}\}$ són els primers ω_1 vectors de la base, $\mathcal{B}_2$ són els ω_2 següents, etc. Per la forma de A, l'acció de f sobre $\mathcal{B}$ és eliminar $\mathcal{B}_1$ i desplaçar (en el corresponent ordre) els ω_i vectors de $\mathcal{B}_i$ als ω_i vectors en $\mathcal{B}_{i-1}$ corresponents, per $i = 2, ..., r$. Ara reordenem la base $\mathcal{B}$ com $\mathcal{B}' = \mathcal{B}'_1 \cup \cdots \cup \mathcal{B}'_s$, on $\mathcal{B}'_1$ està format pels primers elements de $\mathcal{B}_1, ..., \mathcal{B}_r$ (en l'ordre determinat per $\mathcal{B}$), mentre que $\mathcal{B}'_2$ està format pels segons elements de $\mathcal{B}_1, ..., \mathcal{B}_r$ (òbviament d'aquells B_i que tinguin més d'un element),..., $\mathcal{B}'_s$ està format pels últims elements d'aquells $\mathcal{B}_i$ tals que tenen $|\mathcal{B}_i| = \omega_1$. Així, tenim el diagrama de Ferrer on cada bloc conté els vectors de la base $\mathcal{B}$ distribuïts en les seves files i els vectors de la base $\mathcal{B}'$ distribuïts en les seves columnes.



Observant com actua f sobre els vectors de $\mathcal{B}$, veiem que f actua cíclicament sobre cada $\mathcal{B}_i'$, desplaçant cada vector cap al seu predecessor i eliminant el primer de tots. Per tant, la matriu J corresponent a f relativa a la base $\mathcal{B}'$ és la forma de Jordan de A, quan la característica de Segre de A és $(m_1, ..., m_s)$ amb $m_i = |\mathcal{B}_i'|$, per $i = 1, ..., s$. Per tant, pel diagrama de Ferrer anterior, les estructures de Weyr i Jordan són particions duals. A més, $J = P^{-1}AP$, on $P = [\mathcal{B}', \mathcal{B}]$ és la matriu de canvi de base corresponent, que és la matriu de permutacions corresponent a reordenar els vectors de les bases.
En el cas general en que A és suma directa de matrius de Weyr corresponents a diferents VAPs, es fa aquesta permutació en cada matriu bàsica de Weyr i obtenim la forma de Jordan desitjada. $\square$

Observació 5.3. Com a resultat del Teorema 5.2, en la teoria, tot el que es pot fer amb la forma de Jordan es pot fer amb la forma de Weyr, i a l'inrevés. Ara bé, a la pràctica poden existir situacions en què no sigui "fàcil"formular la forma de Weyr en termes de la de Jordan.

D'altra banda, l'existència de la forma de Weyr està directament relacionada amb l'existència de la forma de Jordan.

Corol·lari 5.4. *Siguin $(\omega_1, \omega_2, \ldots, \omega_r)$ la característica de Weyr de la matriu $A \in M_n(\mathbb{K})$ amb VAP λ, i sigui J la forma de Jordan de A. Aleshores, per a tot enter positiu k:*

(1) El nombre de blocs bàsics de Jordan en la matriu J amb VAP λ i amb dimensió almenys $k \times k$ és ω_k.
(2) El nombre de blocs $k \times k$ bàsics de Jordan en J amb VAP λ és $\omega_k - \omega_{k+1}$.

Demostració. És suficient demostrar-ho pel cas en què A té com a únic VAP λ. Siguin $(\omega_1, ..., \omega_r)$ i $(m_1, ..., m_s)$ la característica de Weyr i la característica de Segre, respectivament, de la matriu A (que com bé sabem són particions duals de n). Sigui r l'índex de nilpotència de la matriu $A - \lambda Id$, que alhora és m_1, la dimensió del bloc de Jordan més gran de la seva forma de Jordan. Per tant, (1) es compleix quan $k > r$.
Suposem ara $k \leq r$, i suposem el diagrama de Ferrer amb $(\omega_1, ..., \omega_r)$. El nombre d'$m_i$ que són al menys k han de ser ω_k perquè m_i és la mida de la columna i-èssima del diagrama de Ferrer i ω_k és la mida de la fila k-èssima. Per últim, per la Proposició 4.7 tenim que $\omega_k = m_k$. $\square$

Corol·lari 5.5. *Sigui J una matriu de Jordan amb un únic VAP λ i característica de Segre $(m_1, ..., m_s)$. Aleshores,*

$s = nul(J - \lambda Id)$,

$m_1 =$ l'índex de nilpotència de $J - \lambda Id$,
$m_i =$ el número d'enters j, $1 < j < s$, tals que $nul(J - \lambda Id)^j - nul(J - \lambda Id)^{j-1} \geq i$, per $i = 2, ..., s$.

Demostració. Per una banda, la Proposició 4.7 ens diu que la característica de Weyr de J és $(\omega_1, ..., \omega_r)$, on r és l'índex de nilpotència de $J - \lambda Id$ i $\omega_j = nul(J - \lambda Id)^j - nul(J - \lambda Id)^{j-1}$, per $j = 1, ..., r$. Per altre banda, el Teorema 5.2 ens diu que $(m_1, ..., m_s)$ és la partició dual de $(\omega_1, ..., \omega_r)$, així doncs $s = \omega_1, m_1 = r$, i $m_i =$ quantitat d'enters j entre 1 i s tals que $\omega_j \geq i$, per $i = 1, ..., s$. $\square$

Una característica interessant de la forma de Weyr és com afecta a les files i columnes d'una certa matriu A el producte per l'esquerra o per la dreta amb una matriu de Jordan nilpotent J. Per exemple,

$$AJ = \begin{bmatrix} a & b & c & d \\ e & f & g & h \\ i & j & k & l \\ m & n & o & p \end{bmatrix} \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & a & b & c \\ 0 & e & f & g \\ 0 & i & j & k \\ 0 & m & n & o \end{bmatrix}$$

$$JA = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} a & b & c & d \\ e & f & g & h \\ i & j & k & l \\ m & n & o & p \end{bmatrix} = \begin{bmatrix} e & f & g & h \\ i & j & k & l \\ m & n & o & p \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Suposem ara que J i W són matrius nilpotents de Jordan i Weyr, respectivament. Aleshores, si tenen la mateixa estructura, per exemple $(3, 3, 3)$, tenim que si

$$A = \begin{bmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \end{bmatrix}, A_{ij} \in M_3 \Rightarrow AW = \begin{bmatrix} 0 & A_{11} & A_{12} \\ 0 & A_{21} & A_{22} \\ 0 & A_{31} & A_{32} \end{bmatrix}$$

En canvi, no és possible expressar AJ com una matriu per blocs els quals només estiguin en funció dels blocs A_{ij} perquè J només trasllada les seves columnes o files de manera local, no global.

Observació 5.6. Per tant una matriu de Jordan nilpotent J actua sobre el producte per la dreta traslladant les columnes de A cap a la dreta i introduint a la primera columna un vector de zeros. En canvi, si W és una matriu nilpotent de Weyr el producte per la dreta actuarà de la mateixa manera només si $\omega_j > \omega_{j+1}$, on $(\omega_1, ..., \omega_r)$ és la seva estructura de Weyr. En el cas del producte per l'esquerra, necessitarem que $\omega_i > \omega_{i+1}$.

6 Centralitzador d'una matriu

Una de les claus en la simplificació de certes demostracions quan treballem amb la forma de Weyr enlloc de amb la forma de Jordan és que la forma de Weyr presenta un centralitzador format per matrius triangulars superiors per blocs.

Definició 6.1. *Donada una matriu $A \in M_n(\mathbb{K})$, el conjunt de totes les matrius $B \in M_n(\mathbb{K})$ que commuten amb A formen una subàlgebra de $M_n(\mathbb{K})$ que anomenarem el* ***centralitzador de** A, $\mathcal{C}(A)$.*

Proposició 6.2. *Sigui $A \in M_n(\mathbb{K})$ una matriu diagonal per blocs on cada bloc $A_i \in M_{m_i}(\mathbb{K})$ té un únic VAP λ_i, $\lambda_i \neq \lambda_j$, si $i \neq j$. Aleshores, les matrius de $\mathcal{C}(A)$ són de la forma $B = diag(B_1, ..., B_k)$, on cada bloc $B_i \in M_{m_i}(\mathbb{K})$ pertany a $\mathcal{C}(A_i)$. Així doncs,*

$$\mathcal{C}(\mathcal{A}) \cong \prod_{i=1}^{k} \mathcal{C}(A_i).$$

A més, $\mathcal{C}(A_i) = \mathcal{C}(A_i - \lambda_i Id)$.

Demostració. Suposem que $B = diag(B_1, ..., B_k)$, on cada bloc $B_i \in M_{m_i}(\mathbb{K})$. Si cada B_i centralitza A_i, clarament B centralitzarà A.

D'altra banda, suposem que $B \in M_n(\mathbb{K})$ centralitza A. Escrivim $B = (B_{ij})$ com una matriu per blocs $k \times k$ amb la mateixa estructura que A. Fixem índexs $i \neq j$, i com que A i B commuten, $A_i B_{ij} = B_{ij} A_j$, i $B_{ij} = 0$ pel Teorema de Sylvester (A_i i A_j no tenen VAPs en comú), per tant B és una matriu diagonal per blocs. Notem la forma en què cada bloc diagonal multiplica i veiem que el bloc diagonal i-èssim de B ha de centralitzar el de A. □

Per tant, podem suposar que A és nilpotent i està en forma canònica.

6.1 Centralitzador d'una matriu en la seva forma canònica de Jordan

A continuació veurem que les matrius que commuten amb una matriu de Jordan són matrius de Toeplitz per blocs.

Proposició 6.3. *Siguin $J \in M_n$ una matriu de Jordan nilpotent amb característica de Segre $(m_1, ..., m_s)$, i $T = (T_{ij}) \in M_n$ una matriu per blocs tal que $T_{ij} \in M_{m_i, m_j}$, per $i, j = 1, \ldots, s$. Aleshores, $TJ = JT$ si i només si, cadascun dels s^2 blocs T_{ij} és triangular superior i tal que si $i \geq j$,*

$$T_{ij} = \begin{bmatrix} 0 & \dots & 0 & a & \dots & x & y & z \\ & & & & a & \dots & x & y \\ & & & & & \ddots & & x \\ & & & & & & & \vdots \\ & & & & & & & a \end{bmatrix},$$

i si $i \leq j$,

$$T_{ij} = \begin{bmatrix} a & b & c & & \dots & z \\ 0 & a & b & c & & \\ & 0 & a & b & & \\ \vdots & & & & \ddots & \\ & & & & & a \\ 0 & & & \dots & & 0 \end{bmatrix}$$

Notem que per abreujar notació, no hem indicat la dependència en i, j dels elements $a, b, c, ...$

Demostració. Podem escriure $J = diag(J_1, ..., J_s)$, on cada $J_i \in M_{m_i}$ és una matriu bàsica de Jordan, per $i = 1, ..., s$ de manera que

$$J_i = \begin{bmatrix} 0 & 1 & & \\ & 0 & \ddots & \\ & & \ddots & 1 \\ & & & 0 \end{bmatrix}.$$

La condició que T commuta amb J es veu fàcilment del fet que

$$(*) \quad J_i T_{ij} = T_{ij} J_j,$$

per $i, j = 1, ...s$. Ara ens fixem en l'efecte desplaçament cap a munt en les files d'una matriu sota l'acció del producte per l'esquerra per J_i, i en l'efecte desplaçament cap a la dreta en les columnes d'una matriu quan fem el producte per la dreta per J_j. Per tant, $(*)$ implica que totes les diagonals d'amunt a l'esquerra cap a baix a la dreta (superiors i inferiors) de T_{ij} han de ser constants. Però en el cas en què $i \leq j$, també ens diu que la primera columna de T_{ij}, a banda de l'element $(1, 1)$, han de ser zero, i de fet totes les diagonals per sota són també zero. En el cas en què $i \geq j$, la condició $(*)$ implica que els darreres elements de les files de T_{ij}, a banda dels elements (m_i, m_j) han de ser zero. Això implica que els elements de les diagonals per sota han de ser zero. En aquest cas, T_{ij} és triangular superior amb elements constants a la diagonal i superdiagonals.

D'altra banda, és directe demostrar que aquesta condició implica que es satisfà la condició $(*)$. □

Proposició 6.4. ***(Formula de Frobenius)*** *Sigui $A \in M_n$ una matriu nilpotent i sigui $(m_1, ..., m_s)$ la seva la característica de Segre. Aleshores,*

$$dim\ \mathcal{C}(A) = m_1 + 3m_2 + 5m_3 + \cdots + (2s - 1)m_s.$$

Demostració. Podem suposar que A està en la seva forma de Jordan perquè una transformació per semblances no alteraria la dimensió del centralitzador. (Si $B = S^{-1}AS$, aleshores $\mathcal{C}(B) = S^{-1}\mathcal{C}(A)S \cong \mathcal{C}(A)$, i per tant $dim\ \mathcal{C}(B) = dim\ \mathcal{C}(A)$). Per la Proposició 6.3, sabem la forma de la matriu $s \times s$ per blocs $K = (K_{ij})$ que centralitza A. Comptem ara el nombre d'opcions diferents independents que tenim per les entrades de K considerades com a matriu $n \times n$. Aquesta serà la dimensió del centralitzador. Les entrades de cadascuna de les matrius T_{ij} amb dimensions $m_i \times m_i$ pot ser escollida independentment de les altres. El nombre d'opcions per les entrades de T_{ij} serà $m = min\{m_i, m_j\}$, això és, la m tal que $k = max\{i, j\}$. Donat $1 \leq k \leq s$, existeixen $2k-1$ parelles d'índexs (i, j) per cada $max\{i, j\} = k$. Per a cada parella, el corresponent T_{ij} pot tenir m_k opcions independents de les diagonals que van d'amunt a l'esquerra a baix a la dreta. Per tant, el nombre total d'opcions independents per les entrades de T_{ij} amb $max\{i, j\} = k$ és $(2k-1)m_k$. Així, el nombre total d'opcions a triar per les entrades de T és

$$dim\, \mathcal{C}(A) = m_1 + 3m_2 + \cdots + (2k-1)m_k + \cdots + (2s-1)m_s.$$

□

6.2 Centralitzador d'una matriu en la seva forma canònica de Weyr

A la següent proposició veurem que les matrius que commuten amb una matriu de Weyr són matrius triangulars superiors per blocs.

Proposició 6.5. *Siguin $W \in M_n$ una matriu bàsica de Weyr amb un únic VAP i estructura $(\omega_1, ..., \omega_r), r \geq 2$,i $K = (K_{ij}) \in M_n$ una matriu per blocs tal que $n = \omega_1 + \cdots + \omega_r$, on $K_{ij} \in M_{\omega_i, \omega_j}$, $i, j = 1, ..., r$. Aleshores, $WK = KW$ si i només si,*

$$K_{ij} = \begin{bmatrix} K_{i+1,j+1} & * \\ 0 & * \end{bmatrix}, \ 1 \leq i \leq j \leq r-1$$

Així, escrivim K_{ij} com una matriu per blocs on el bloc 0 té dimensions $(\omega_i - \omega_{i+1}) \times \omega_{j+1}$, els blocs $*$ no tenen cap restricció (però tenen la particularitat que desapareixen si $\omega_j = \omega_{j+1}$, i la fila $[0*]$ desapareix si $\omega_i = \omega_{i+1}$.

Demostració. Si eliminem la diagonal principal de la matriu W, podem suposar que és nilpotent sense que el centralitzador canviï. Per $j = 2, ..., r$, sigui $I_j \in M_{\omega_{j-1}, \omega_j}$ formada per la matriu $Id \in M_{\omega_j}$ a la seva part superior, seguida per $\omega_{j-1} - \omega_j$ files de zeros, i.e.,

$$W = \begin{bmatrix} 0 & I_2 & & \\ & 0 & \ddots & \\ & & \ddots & I_r \\ & & & 0 \end{bmatrix}.$$

Si ara suposem que $KW = WK$, mirant els blocs podem concloure que $K_{21} = K_{31} = \cdots = K_{r1} = 0$, i a més, $K_{ij} I_{j+1} = I_{i+1} K_{i+1,j+1}$, $1 \leq i, j \leq r-1$; I per tant,

$$K_{ij} = \begin{bmatrix} K_{i+1,j+1} & * \\ 0 & * \end{bmatrix}, 1 \leq i, j \leq r-1.$$

Si ajuntem aquest resultat amb el fet que $K_{j1} = 0, j = 2, ..., r$ tenim que $K_{ij} = 0, i > j$, i ja tenim la primera implicació.

Suposem ara que K satisfà les hipòtesis i, amb uns pocs càlculs, arribem a la conclusió que

$$K_{ij}I_{j+1} = I_{i+1}K_{i+1,j+1}, 1 \leq i \leq j \leq r-1.$$

Així doncs, tenim que els elements del bloc $(i, j+1)$ de les matrius KW i WK coincideixen amb $1 \leq i \leq j \leq r-1$. Així tenim que K és una matriu triangular superior per blocs i per tant els blocs de la diagonal de les matrius KW i WK han de ser zero. Com que les dues són triangulars superiors per blocs, els blocs de les dues matrius han de coincidir. □

Proposició 6.6. *Sigui* $A \in M_n$ *una matriu nilpotent, i sigui* $(\omega_1, ..., \omega_r)$ *la seva estructura de Weyr. Aleshores,*

$$dim\ \mathcal{C}(A) = \omega_1^2 + \omega_2^2 + \cdots + \omega_r^2.$$

Demostració. Podem suposar que A està en la seva forma de Weyr ja que no alterarà la dimensió del centralitzador. Les matrius K que centralitzen A són matrius triangulars superiors per blocs $r \times r$, amb la mateixa estructura que A, de la forma

$$K = \begin{bmatrix} K_{11} & K_{12} & \dots & K_{1r} \\ & K_{22} & \dots & K_{2r} \\ & & \ddots & \vdots \\ & & & K_{rr} \end{bmatrix},$$

on $K_{ij} = \begin{bmatrix} K_{i+1,j+1} & * \\ 0 & * \end{bmatrix}$, per $1 \leq i \leq j \leq r-1$.

Anem a calcular el nombre de tries independents que podem fer per als elements d'una matriu $n \times n$. Començant per la darrera fila de blocs i si anem pujant, cada fila i ens dóna exactament ω_i^2 tries independents.
Observem que a mesura que ascendim per les files, cada fila i-èssima ens afegeix ω_i^2 noves possibles tries.

Argumentem recursivament, per la matriu $K_{rr} \in M_{\omega_r}$ tenim ω_r^2 possibles tries per als seus elements. Si fixem $1 \leq i < r$ i suposem que hem triat els elements de les darreres $r - i$ files de blocs, per la Proposició 6.5, les nostres tries d'elements per la fila i-èssima de blocs estan condicionades per la relació entre K_{ij} i $K_{i+1,j+1}$, per $i \leq j \leq r-1$, i.e., només podem triar els elements que són asteriscs de la matriu

$$K_{ij} = \begin{bmatrix} K_{i+1,j+1} & * \\ 0 & * \end{bmatrix}.$$

Així doncs, per a cada bloc en la fila i-èssima, tenim

$$\omega_i[(\omega_i - \omega_{i+1}) + (\omega_{i+1} - \omega_{i+2}) + \cdots + (\omega_{r-1} - \omega_r) + \omega_r] = \omega_i^2$$

tries addicionals possibles, i per inducció ja ho tenim. Per tant, tenim un total de $\omega_r^2 + \omega_{r-1}^2 + \cdots + \omega_2^2 + \omega_1^2$ tries possibles per els elements de K, el que ens dona la dimensió de $\mathcal{C}(A)$. □

Exemple 6.7. Sigui $A \in M_7$ una matriu nilpotent amb característica de Segre $(3,2,2)$ i per tant, estructura de Weyr $(3,3,1)$ (Teorema 5.2) . Si posem la matriu A en la seva forma de Jordan, $T \in \mathcal{C}(A)$ és de la forma

$$T = \left[\begin{array}{ccc|cc|cc} a & b & c & d & e & f & g \\ 0 & a & b & 0 & d & 0 & f \\ 0 & 0 & a & 0 & 0 & 0 & 0 \\ \hline 0 & h & i & j & k & l & m \\ 0 & 0 & h & 0 & j & 0 & l \\ \hline 0 & n & p & q & r & s & t \\ 0 & 0 & n & 0 & q & 0 & s \end{array}\right].$$

A més, aplicant la fórmula de Frobenius, tenim que

$$dim\ \mathcal{C}(A) = 1 \times 3 + 3 \times 2 + 5 \times 2 = 19.$$

Suposem ara que A està en forma de Weyr, per la Proposició 6.5, una matriu $K \in \mathcal{C}(A)$, és de la forma

$$K = \left[\begin{array}{ccc|ccc|c} a & b & c & h & k & l & r \\ 0 & d & e & i & m & n & s \\ 0 & f & g & j & p & q & t \\ \hline & & & a & b & c & h \\ & & & 0 & d & e & i \\ & & & 0 & f & g & j \\ \hline & & & & & & a \end{array}\right],$$

i per la Proposició 6.5, tenim que

$$dim\ \mathcal{C}(A) = 3^2 + 3^2 + 1^2 = 19.$$

7 Aplicacions de la forma canònica de Weyr

7.1 Resultat de combinatòria

Proposició 7.1. *Siguin* $(m_1, ..., m_s)$ *una partició d'un cert enter* n*, i sigui* $(\omega_1, ..., \omega_r)$ *la seva partició dual. Aleshores,*

$$m_1 + 3m_2 + 5m_3 + \cdots + (2s-1)m_s = \omega_1^2 + \omega_2^2 + \cdots + \omega_r^2.$$

Demostració. Sigui $A \in M_n$ una matriu nilpotent amb característica de Segre $(m_1, ..., m_s)$, i partició dual $(\omega_1, ..., \omega_r)$. Per les Proposicions 6.4 i 6.6, tenim que

$$m_1 + 3m_2 + 5m_3 + \cdots + (2s-1)m_s = dim\ \mathcal{C}(A) = \omega_1^2 + \cdots + \omega_r^2.$$

□

7.2 Triangularització simultània

La proposició següent és un resultat clàssic en la teoria de matrius que utilitzarem en la demostració del Teorema 7.3

Proposició 7.2. *[1] Siguin* $A_1, ..., A_k \in M_n$ *matrius que commuten dos a dos. Aleshores,* $A_1, ..., A_k$ *poden ser triangularitzades simultàniament, i.e, existeix una matriu invertible* $S \in M_n(\mathbb{K})$ *tal que* $S^{-1}A_iS$ *és una matriu triangular superior, per* $i = 1, ..., k$*.*

Demostració. Sigui $E = \mathbb{K}^n$, sigui $\mathcal{B}$ la base canònica de E, i pensem cada matriu A_j com la matriu de l'aplicació producte per l'esquerra de V per A_j, associada a la base $\mathcal{B}$. Podem suposar que no totes les nostres matrius són matrius escalars, així si les reordenem tenim que A_1 és no escalar. (*Recordem que una matriu escalar és una matriu de la forma* λId). Com que $\mathbb{K}$ és algebraicament tancat, podem triar un VAP $\lambda \in Spec(A_1)$ i sigui $U = ker(\lambda Id - A_1)$ el corresponent espai propi. Així tenim que U és invariant sota cada A_j, ja que A_j commuta amb $\lambda Id - A_1$. Si diem $m = dim(U)$, tenim que $1 \leq m < n$. Triem ara una base $\mathcal{B}'$ de V tal que estén algunes bases per U. La matriu associada a l'aplicació producte per A_j associada a la base $\mathcal{B}'$ té la forma

$$P^{-1}A_jP = \begin{bmatrix} B_j & C_j \\ 0 & D_j \end{bmatrix},$$

on P és la matriu de canvi de base, $B_j \in M_m$ i $D_j \in M_{n-j}$. Com que $P^{-1}A_jP$ commuta, també ho faran B_j i D_j, $j = 1, ..., k$. Per inducció sobre n, podem suposar que les matrius B_j poden ser simultàniament triangularitzades per $R \in GL_m(\mathbb{K})$, i les matrius D_j amb S invertible. Aleshores, conjugant per

$$Q = \begin{bmatrix} R & 0 \\ 0 & S \end{bmatrix},$$

tenim que totes les matrius $P^{-1}A_jP$ són triangulars superiors. Per tant, conjugant $A_1, ..., A_k$ per $S = PQ$, tenim el resultat desitjat. □

El següent teorema ens permet transformar simultàniament k matrius que commuten en $k-1$ matrius triangulars superiors i una en forma de Weyr. Això no és possible en forma de Jordan.

Teorema 7.3. *[1] Siguin $A_1, \ldots, A_k \in M_n$ matrius que commuten, aleshores existeix una transformació per semblances que transforma la matriu A_1 en la seva forma de Weyr i simultàniament transforma les matrius $A_2, \ldots, A_k$ en matrius triangulars superiors.*

Demostració. Sabem que podem conjugar simultàniament $A_1, ..., A_k$ per transformar A_1 en una matriu diagonal per blocs, on cada bloc té un únic VAP i els VAPs de diferents blocs són diferents dos a dos. Per la commutativitat i pel fet que cada bloc té un únic VAP i és diferent al dels altres blocs, la partició dels altres A_i serà la mateixa que la de A_1. Així, podem suposar que A_1 té un únic VAP. També podem suposar que les matrius $A_1, ..., A_k$ commuten, amb A_1 essent una matriu bàsica de Weyr, amb estructura de Weyr $(\omega_1, ..., \omega_r)$. Així, aplicant la Proposició 6.5 i fent $W = A_1$, veiem que cada A_i és una matriu triangular superior per blocs amb l'estructura per blocs de A_1. Completem la demostració construint inductivament una matriu diagonal per blocs invertible $S = diag(S_1, ..., S_r)$ amb estructura per blocs $(\omega_1, ..., \omega_r)$ tal que

(i) S centralitza A_1,
(ii) S conjuga $A_2, ..., A_k$ simultàniament a matrius triangulars superiors.

Conjugant les matrius $A_1, ..., A_k$ per S deixa A_1 en forma de Weyr i simultàniament transforma les matrius $A_2, ..., A_k$ en triangulars superiors. Construirem les matrius S_i recursivament en l'ordre $S_r, ..., S_1$. Els blocs diagonals (r, r) de les matrius $A_2, ..., A_k$ commuten, així que per la Proposició 7.2 existeix una matriu S_r invertible $n_r \times n_r$ que conjuga simultàniament aquests blocs a matrius triangulars superiors. Suposem que hem construït S_i per algun $i > 1$, construïm ara el bloc S_{i-1}:
Si $\omega_{i-1} = \omega_i$, podem agafar $S_{i-1} = S_i$.
Suposem $\omega_{i-1} > \omega_i$, com que $A_2, ..., A_k$ centralitzen A_1, per la Proposició 6.5 el bloc $(i-1, i-1)$ de la matriu A_j té la forma

$$\begin{bmatrix} Y_j & * \\ 0 & Z_j \end{bmatrix}, \ per\ j = 2, ..., k,$$

on Y_j és el bloc (i, i) de A_j i $Z_j \in M_{\omega_{i-1}-\omega_i}$ (tot i que les $Y's$ i $Z's$ depenen de i, el veurem com fix per evitar dobles índexs). Els Z_j commuten, perquè $A_2, ..., A_k$ commuten, així escollim una matriu $D_{i-1} \in M_{\omega_{i-1}-\omega_i}$ invertible que conjugui simultàniament $Z_2, ..., Z_k$ a matrius triangulars superiors. Per tant,

$$C_{i-1} = \begin{bmatrix} C_i & 0 \\ 0 & D_{i-1} \end{bmatrix}.$$

Això completa la construcció de la matriu $S = diag(S_1, \ldots, S_k)$. La importància de les propietats (i) i (ii) és que la primera garanteix que $SA_1 = A_1S$ centralitza A_1, per la Proposició 6.5. La segona, observem que inductivament per cada $i = r, r-1, ..., 1$, S_i conjuga els blocs (i, i) de $A_2, ..., A_k$ a matrius triangulars superiors. Observem també que per qualsevol matriu triangular superior per blocs $X = (X_{ij}) \in M_r$, amb la mateixa estructura per blocs que C, tenim que

$$S^{-1}XS = \begin{bmatrix} S_1^{-1}X_{11}S_1 & * & \dots & * \\ 0 & S_2^{-1}X_{22}S_2 & & \vdots \\ \vdots & & \ddots & * \\ 0 & \dots & 0 & S_r^{-1}X_{rr}S_r \end{bmatrix}.$$

En particular, deixant que X s'estengui sobre $A_2, .., A_k$ tenim la propietat (ii). $\square$

7.3 Teorema de Gerstenhaber

Si $A_1, ..., A_k \in M_n$, denotem per $F[A_1, ..., A_k]$ la subàlgebra de M_n generada per $A_1, ..., A_k$, i.e., la subàlgebra més petita de M_n que conté $A_1, ..., A_k$. Direm que $F[A_1, ..., A_k]$ és una **subàlgebra k-generada**. En general, les $k-$àlgebres poden ser complicades però si les matrius $A_1, \dots, A_k$ commuten, aquesta descripció es simplifica.
La forma canònica de Weyr es pot utilitzar per trobar fites de subàlgebres $F[A, B]$ de matrius. En el cas d'una matriu $A \in M_n$, el teorema de Cayley-Hamilton ens diu que $dim\ F[A] \leq n$. En el cas de dues matrius $A, B \in M_n$, ens podem trobar que $dim\ F[A, B] = n^2 = dim\ M_n$ (per exemple, si A és una matriu bàsica de Jordan nilpotent i $B = A^T$); Ara bé, si A i B commuten, sorprenentment tenim que $dim\ F[A, B] \leq n$.

Aquest és el resultat al que va arribar *Murray Gerstenhaber* l'any 1961. El teorema que porta el seu nom, ha sigut demostrat per diversos matemàtics amb el pas dels anys, alguns fent servir tècniques de geometria algebraica, i d'altres aplicant conceptes matricials relacionats amb la forma canònica de Jordan.

Abans d'enunciar el teorema vegem un exemple de subàlgebra generada per una matriu.

Sigui

$$A = \left[\begin{array}{ccc|cc|cc} 0 & 1 & 0 & & & & \\ 0 & 0 & 1 & & & & \\ 0 & 0 & 0 & & & & \\ \hline & & & 0 & 1 & & \\ & & & 0 & 0 & & \\ \hline & & & & & 0 & 1 \\ & & & & & 0 & 0 \end{array}\right],$$

amb característica de Segre $(3, 2, 2)$, i per tant, índex de nilpotència $r = 3$.

Tenim que Id, A, A^2 són una base de $F[A]$, i un element de la subàlgebra tindrà el següent aspecte:

$$aId + bA + cA = \left[\begin{array}{ccc|cc|cc} a & b & c & & & & \\ 0 & a & b & & & & \\ 0 & 0 & a & & & & \\ \hline & & & a & b & & \\ & & & 0 & a & & \\ \hline & & & & & a & b \\ & & & & & 0 & a \end{array}\right].$$

I si posem A en la seva forma de Weyr, tenim que

$$A = \left[\begin{array}{ccc|ccc|c} 0 & 0 & 0 & 1 & 0 & 0 & \\ 0 & 0 & 0 & 0 & 1 & 0 & \\ 0 & 0 & 0 & 0 & 0 & 1 & \\ \hline & & & 0 & 0 & 0 & 1 \\ & & & 0 & 0 & 0 & 0 \\ & & & 0 & 0 & 0 & 0 \\ \hline & & & & & & 0 \end{array}\right],$$

té estructura de Weyr $(3, 3, 1)$ i un element genèric de la subàlgebra $F[A]$ serà

$$aId + bA + cA^2 = \left[\begin{array}{ccc|ccc|c} a & 0 & 0 & b & 0 & 0 & c \\ 0 & a & 0 & 0 & b & 0 & 0 \\ 0 & 0 & a & 0 & 0 & b & 0 \\ \hline & & & a & 0 & 0 & b \\ & & & 0 & a & 0 & 0 \\ & & & 0 & 0 & a & 0 \\ \hline & & & & & & a \end{array}\right].$$

De les dues representacions de A, la forma de Jordan seria millor ja que genera matrius diagonals per blocs en $F[A]$. A continuació veurem que en el cas de dues matrius la forma de Jordan ja no funciona tan bé.

Estudiarem el cas $k = 2$, i.e., $F[A, B]$. Tot i imposar que A estigui en la seva forma de Jordan, B només estarà restringida pel centralitzador descrit en la Proposició 6.3, així que en general els elements de $F[A, B]$ no seran ni tan sols triangulars superiors per blocs amb l'estructura de A. És la matriu triangular superior per blocs del centralitzador d'una matriu de Weyr, descrita en la Proposició 6.5, la que sembla donar-nos una de les raons de perquè les matrius en forma de Weyr són les més adients en l'estudi de les subàlgebres commutatives $k-$generades, quan $k > 1$.

La proposició següent ens permetrà suposar que les matrius $A_1, \ldots, A_k$ són nilpotents.

Proposició 7.4. ***(Reducció nilpotent)*** *Siguin $A_1, ..., A_k \in M_n$ matrius que commuten. Aleshores, existeix una semblança simultània de $A_1, ..., A_k$ tal que:*

1) Totes les A_i es transformen en matrius diagonals per blocs en què cada bloc té un únic VAP (amb la multiplicitat que correspongui), i.e.,

$$A_i = diag(A_{i1}, ..., A_{it}), \text{ per } i = 1, ..., k,$$

on cada $A_{ij} \in M_{m_j}$ té un únic VAP.

2) Com a àlgebres,

$$F[A_1, ..., A_k] \cong \prod_{j=1}^{t} F[A_{1j}, ..., A_{kj}].$$

3) Si fixem j, després d'eliminar les matrius escalars que calgui, les k matrius commutatives $A_{1j}, ..., A_{kj} \in M_{m_j}$ es transformen en matrius nilpotents (que encara generen *la mateixa subàlgebra).*

Demostració. Siguin $\lambda_1, ..., \lambda_t$ els diferents VAPs de A_1 i sigui $p(x) = (x - \lambda_1)^{m_1} \dots (x - \lambda_t)^{m_t}$ el polinomi característic de A_1. Aleshores sabem que existeix una transformació per semblances tal que $A_1 = diag(A_{11}, ..., A_{1t})$, on A_{1j} és una matriu $m_j \times m_j$ que té com a únic VAP λ_j. Si apliquem la mateixa transformació a les matrius $A_2, ..., A_k$, les noves matrius han de centralitzar encara la nova A_1. Per tant, per la Proposició 6.2, $A_2, ..., A_k$ són també matrius diagonals per blocs amb la mateixa estructura per blocs que A_1, i escrivim $A_i = diag(A_{i1}, ..., A_{it})$, per $i = 1, ..., k$.

Per cada $i = 1, ..., t$ sigui $f_i(x) = \prod_{j \neq i}(x - \lambda_j)^{m_j}$. Aquest polinomis són relativament primers entre ells, per tant existeixen polinomis $g_1(x), ..., g_t(x)$ tals que:

$$f_1(x)g_1(x) + \cdots + f_t(x)g_t(x) = 1. \text{ (\textit{Identitat de Bezout})}$$

Sigui $E_i = f_i(A_1)g_i(A_1)$, per $i = 1, ..., t$. Observem que $E_i \in F[A_1]$. Així doncs,

$$E_1 + \cdots + E_t = Id.$$

Com a matrius diagonal per blocs, escrivim $E_i = diag(E_{i1}, ..., E_{it})$. Com que A_{1j} té com a únic VAP λ_j, $(A_{1j} - \lambda_j Id)^{m_j} = 0$ (Teorema de Cayley-Hamilton). Per tant, $f_i(A_{1j}) = 0$, per cada $j \neq i$, i $E_{ij} = f_i(A_{1j})g_i(A_{1j}) = 0$, per cada $j \neq i$. Això implica que E_i és una matrius diagonal per blocs formada per la matriu identitat amb dimensions $m_i \times m_i$ en el bloc i-èssim de la diagonal, i a la resta de blocs el bloc zero. Això és important perquè hem creat $E_1, ..., E_t$ tal que són idempotents, ortogonals i la seva suma és igual a la identitat en $F[A_1]$. Per tant, existeix una àlgebra associada descomposició directa com a producte de qualsevol àlgebra commutativa que contingui A_1, en particular, una descomposició de $\mathcal{A} = F[A_1, ..., A_k]$. Així doncs, tenim un isomorfisme

$$\xi : \mathcal{A} \to \prod_{j=1}^{t} E_j\mathcal{A}, \;\; X \to (E_1X, ..., E_tX).$$

Si identifiquem cada E_jX amb el bloc diagonal j-èssim, podem veure que ξ indueix un isomorfisme

$$F[A_1, ..., A_k] \cong \prod_{j=1}^{t} F[A_{1j}, ..., A_{kj}].$$

Pràcticament ja hem acabat la demostració, però encara hem de veure que cada A_{ij} té un únic VAP per $i \neq 1$. Suposem que un d'aquests blocs té dos VAPs diferents, podem repetir la partició en aquest bloc amb el mateix argument que hem fet servir pel bloc A_1. Per inducció si cal, aconseguim una partició en la que cada A_{ij} té un únic VAP, i per tant és igual a la suma d'una matriu escalar i una nilpotent. (Poden existir alguns A_{ij} i A_{im} que tinguin algun VAP en comú quan $j \neq m$.) $\square$

Teorema 7.5. ***(L'equació generalitzada de Cayley-Hamilton).*** *Siguin $A, B \in M_n$ dues matrius que commuten, sigui d el màxim de les dimensions dels blocs bàsics de Jordan que tenen el VAP λ a la matriu A. Aleshores,*

$$B^d = A_0 + A_1B + A_2B^2 + \cdots + A_{d-1}B^{d-1},$$

per algunes matrius $A_i \in F[A]$, $i = 0, ..., d-1$, i.e., les matrius A_i són polinomials en A amb coeficients al cos $\mathbb{K}$. L'equació clàssica de Cayley-Hamilton s'obté si considerem $A = Id$ i $d = n$.

Demostració. Cas 1: A és nilpotent d'estructura homogènia de Jordan.

Podem suposar, perquè una transformació per semblances no afectarà el teorema, que A és una matriu de Jordan nilpotent amb índex de nilpotència r amb característica de Segre $(r, ..., r)$ formada per d blocs, on d és la seva nul·litat. Per la definició de centralitzador, sabem que $B \in M_d$ és una matriu de la forma

$$B = \begin{bmatrix} B_{11} & \dots & B_{1d} \\ \vdots & & \vdots \\ B_{d1} & \dots & B_{dd} \end{bmatrix},$$

on cada $B_{ij} \in M_r$ és una matriu de la forma

$$B_{ij} = \begin{bmatrix} a & b & c & \dots & z \\ & a & b & \dots & y \\ & & \ddots & \vdots & \\ & & & a & b \\ & & & & a \end{bmatrix},$$

on els elements $a, b, ..., z$ depenen de la posició (i, j). A més, B la podem veure com una matriu sobre un anell *commutatiu* $\mathcal{R}$. El sentit de la demostració està en observar que el Teorema clàssic de Cayley-Hamilton funciona no només sobre cossos sinó també sobre qualsevol anell commutatiu. Per tant, existeixen polinomis $p_0(J), ..., p_{d-1}(J)$ en J amb coeficients al cos $\mathbb{K}$ tals que

$$B^d = p_0(J)Id + p_1(J)(B) + p_2(J)B^2 + \cdots + p_{d-1}(J)B^{d-1}.$$

Entenem que els escalars $p_i(J)$ actuen sota la multiplicació escalar, i.e., pel producte per la matriu diagonal per blocs $diag(p_i(J), p_i(J), ..., p_i(J))$, que té exactament d blocs. Però la darrera matriu és $p_i(A)$ perquè $A = diag(J, ..., J)$, així doncs tenim que

$$B^d = p_0(A) + p_1(A)B + \cdots + p_{d-1}(A)B^{d-1},$$

com volíem veure.

Cas 2: A és una matriu nilpotent d'estructura de Weyr no homogènia.

El que volem és reduir aquest cas a l'anterior Amb una transformació per semblances podem suposar que A és una matriu de Weyr nilpotent amb estructura $(\omega_1, ..., \omega_r)$. Observem que $\omega_1 = d$. Sigui $p = dr$ i sigui $W \in M_p(\mathbb{K})$, considerem ara l'aplicació, definida en la notació per files superiors (*), tal que

$$\xi : \mathcal{C}(A) \to \mathcal{C}(W),\ [X_1, ..., X_r] \to [\overline{X_1}, ..., \overline{X_r}],$$

on els centralitzadors pertanyen a $M_n(\mathbb{K})$ i $M_p(\mathbb{K})$, respectivament, i on $\overline{X_j}$ és la matriu $d \times d$ $[X_j 0]$, i.e., cadascun dels blocs X_j amb dimensions $d \times \omega_j$ s'ha transformat en una matriu quadrada $\overline{X_j}$ afegint-hi $d - \omega_j$ columnes de zeros.

Observació 7.6. *Notació per files superiors:* (*)

Sigui K una matriu de la forma que es descriu a continuació, aleshores ens pot convenir expressar-la en termes dels seus blocs superiors, i.e.,

$$K = \left[\begin{array}{ccc|cc|cc} 1 & 1 & 4 & 4 & 6 & 2 & 1 \\ 2 & 3 & 6 & 7 & 5 & 8 & 6 \\ 0 & 0 & 9 & 0 & 0 & 3 & 9 \\ \hline & & & 1 & 1 & 4 & 6 \\ & & & 2 & 3 & 7 & 5 \\ \hline & & & & & 1 & 1 \\ & & & & & 2 & 3 \end{array}\right] \text{"} = \text{"} \left[\begin{array}{ccc|cc|cc} 1 & 1 & 4 & 4 & 6 & 2 & 1 \\ 2 & 3 & 6 & 7 & 5 & 8 & 6 \\ 0 & 0 & 9 & 0 & 0 & 3 & 9 \end{array}\right].$$

En $\xi(X)$, les n columnes de la matriu $X \in \mathcal{C}(A)$ ocuparan noves posicions, les direm $c_1, ..., c_n$, però el més important és que només tenien zeros en les posicions (i, c_j) per $i \notin \{c_1, ..., c_n\}$ (cadascun desplaçant els elements de la fila inferior una fila cap a baix) per convertir-los en vectors columna amb dimensions $p \times 1$. Per exemple, si A té estructura de Weyr $(3, 2, 1)$, aleshores

$$\xi : \left[\begin{array}{ccc|cc|c} a & b & d & g & i & l \\ 0 & c & e & h & j & m \\ 0 & 0 & f & 0 & k & n \\ \hline & & & a & b & g \\ & & & 0 & c & h \\ \hline & & & & & a \end{array}\right] \to \left[\begin{array}{ccc|ccc|ccc} a & b & d & g & i & 0 & l & 0 & 0 \\ 0 & c & e & h & j & 0 & m & 0 & 0 \\ 0 & 0 & f & 0 & k & 0 & n & 0 & 0 \\ \hline & & & a & b & d & g & i & 0 \\ & & & 0 & c & e & h & j & 0 \\ & & & 0 & 0 & f & 0 & k & 0 \\ \hline & & & & & & a & b & d \\ & & & & & & 0 & c & e \\ & & & & & & 0 & 0 & f \end{array}\right],$$

així que $c_i = 0$, per $i = 1, ...5$ i $c_6 = 7$. Sigui $\mathcal{S}$ el conjunt de totes les matrius $M = (m_{ij}) \in M_p$ tals que $m_{ij} = 0$ sempre que j però no i pertanyi al conjunt $\{c_1, ..., c_n\}$. Aleshores $\mathcal{S}$ és una subàlgebra de $M_p(\mathbb{K})$ i l'aplicació $\pi : \mathcal{S} \to M_n(\mathbb{K})$ que elimina les files i columnes i-èssimes d'una matriu quan i no pertany al conjunt $\{c_1, ..., c_n\}$ és un homomorfisme. Per veure-ho, ho farem en termes de transformacions lineals: $\mathcal{S}$ es correspon amb l'àlgebra de les transformacions de $\mathbb{K}^p$ que deixen invariant el subespai format pels vectors de la base estàndard en les posicions $c_1, ..., c_n$. π correspon a l'aplicació que restringeix aquestes transformacions al subespai, tenim doncs que $\xi(\mathcal{C}(A)) \subseteq \mathcal{S}$, i π restringida a la imatge de ξ és l'aplicació inversa de ξ.

Sigui ara $K = \xi(B)$, com que $K \in \mathcal{C}(W)$, existeixen matrius W i K commutatives $p \times p$ en $\mathcal{S}$ tals que W és nilpotent amb una estructura homogènia i la seva nul·litat continua sent d. Pel cas anterior,

$$K^d = Y_0 + Y_1 K + \cdots + Y_{d-1} K^{d-1},$$

per algunes matrius $Y_i \in F[W]$. Aplicant l'homomorfisme π a aquesta equació, observem que $\pi(W) = A$ i $\pi(K) = B$, obtenint així el resultat que volíem.

Cas 3: A és una matriu qualsevol.

Siguin $\lambda_1, ..., \lambda_k$ els diferents VAPs de A. Podem suposar que $A = diag(A_1, ..., A_k)$ i $B = diag(B_1, ..., B_k)$, amb les mateixes estructures per blocs, els blocs (corresponents) commuten i cada A_i té com a únic VAP λ_i. Suposem $k = 2$, però el resultat general és molt semblant.
Siguin d i e les multiplicitats geomètriques de λ_1 i λ_2, respectivament, i suposem $d \geq e$. Observem que qualsevol polinomi en $A_1 - \lambda_1 Id$ es pot expressar com a polinomi en A_1, com en el cas $A_2 - \lambda_2 Id$. Ara, fent servir els dos casos anteriors, tenim que

$$B_1^d = s_0(A_1) + s_1(A_1)B_1 + s_2(A_1)B_1^2 + \cdots + s_{d-1}(A_1)B_1^{d-1},$$
$$B_2^d = s_0(A_2) + t_1(A_2)B_2 + t_2(A_2)B_2^2 + \cdots + t_{d-1}(A_2)B_2^{d-1}.$$

Siguin ara $m_1(x)$ i $m_2(x)$ els polinomis minimals de A_1 i A_2, respectivament. No tenen factors en comú perquè són potències de $x - \lambda_1$ i $x - \lambda_2$, respectivament, així que pel Teorema xinès del residu, existeixen polinomis $p_0(x), ..., p_{d-1}(x)$ tals que

$$p_i(x) \equiv s_i(x) \ (mod\ m_1(x)),$$
$$p_i(x) \equiv t_i(c) \ (mod\ m_2(x)),$$

per $i = 0, ..., d-1$. Com que $m_1(A_1) = 0 = m_2(A_2)$, tenim que

$$B_1^d = p_0(A_1) + p_1(A_1)B_1 + p_2(A_1)B_1^2 + \cdots + p_{d-1}(A_1)B_1^{d-1},$$
$$B_2^d = p_0(A_2) + p_1(A_2)B_2 + p_2(A_2)B_2^2 + \cdots + p_{d-1}(A_2)B_2^{d-1}.$$

Pel simple fet que les matrius diagonals per blocs interactuen algebraicament, per les dues darreres equacions tenim que $B = diag(B_1, B_2)$ i $A = diag(A_1, A_2)$, i per tant

$$B^d = p_0(A) + P_1(A)B + p_2(A)B^2 + \cdots + p_{d-1}(A)B^{d-1}.$$

$\square$

Teorema 7.7. *Siguin $A, B \in M_n$ matrius nilpotents que commuten, i sigui $(\omega_1, ..., \omega_r)$ la característica de Weyr de la matriu A. Aleshores, el conjunt de matrius*

$$Id, B, B^2, \ldots, B^{\omega_1 - 1}$$
$$A, BA, B^2A, \ldots, B^{\omega_2 - 1}A$$

$$A^2, BA^2, B^2A^2, \ldots, B^{\omega_3 - 1}A^2$$
$$\vdots$$
$$A^{r-1}, BA^{r-1}, \ldots, B^{\omega_r - 1}A^{r-1}$$

que anomenarem $\mathcal{B}$ genera la subàlgebra (com a espai vectorial) $F[A,B]$ de $M_n(\mathbb{K})$. En particular, $dim F[A,B] \leq n$.

Demostració. Ho farem per inducció sobre r (l'índex de nilpotència de la matriu A). Si $r = 1$, aleshores $A = 0$, $\omega_1 = n$, i aplicant el Teorema clàssic de Cayley-Hamilton aplicat a la matriu B tenim el resultat desitjat.
Suposem ara que $r > 1$ i que A està en la seva forma de Weyr. Sigui $\mathcal{T}$ la subàlgebra de $M_n(\mathbb{K})$ formada per totes les matrius triangular superiors per blocs amb la mateixa estructura que A, i.e., els blocs de les seves diagonals tenen dimensions $\omega_1, ..., \omega_r$, respectivament. Observem que $\mathcal{A} \subseteq \mathcal{T}$, i sigui la projecció

$$\pi : \mathcal{A} \to \mathcal{T}$$

tal que

$$\begin{bmatrix} X_{11} & X_{12} & \ldots & X_{1r} \\ & x_{22} & \ldots & X_{2r} \\ & & \ddots & \vdots \\ & & & X_{rr} \end{bmatrix} \to \begin{bmatrix} 0 & 0 & \ldots & 0 \\ & x_{22} & x_2 & \ldots \\ & & \ddots & \\ & & & x_{rr} \end{bmatrix},$$

que és un homomorfisme que no preserva la identitat. A més, π s'ajusta naturalment a l'àlgebra de les matrius $(n - \omega_1) \times (n - \omega_1)$ sobre $\mathbb{K}$. Si veiem la veiem dins de $M_{n-\omega_1}(\mathbb{K})$, la matriu $\pi(A)$ està encara en la seva forma de Weyr amb la mateixa estructura que abans. Aplicant el Teorema a les matrius nilpotents que commuten $\pi(A)$ (amb índex de nilpotència $r - 1$) i $\pi(B)$, com que π és un homomorfisme, per inducció en l'índex de nilpotència tenim que $\pi(\mathcal{A}) = F[\pi(A), \pi(B)]$ és generat (com a espai vectorial) per $\pi(\mathcal{B}')$, on $\mathcal{B}'$ és el conjunt de matrius

$$Id, B, B^2, \ldots, B^{\omega_2 - 1}$$
$$\vdots$$
$$A^{r-2}, BA^{r-2}, \ldots, B^{\omega_r - 1}A^{r-2}.$$

Observem que $ker(\pi) = \{X \in \mathcal{C} \mid XA = 0\}$, ja que $X \in \mathcal{A}$ pertany al $ker(\pi)$ quan les files de blocs $2, ..., r$ són zero. En canvi, com que X centralitza A, les primeres files de la matriu X han de ser de la forma

$$[X_{11}, ...X_{1r}],$$

on X_{1j} és una matriu $\omega_1 \times \omega_j$ amb les seves ω_{j+1} primeres columnes iguals a zero, per $j = 1, ..., r-1$.

$$x_{1j} = \begin{bmatrix} X_{2,j+1} & * \\ 0 & * \end{bmatrix}.$$

Però pel desplaçament i eliminació parcial de les columnes de la dreta que es produeix en X pel producte per la dreta per A, veiem que són precisament les condicions necessàries per tal que X pertanyi a $\{X \in \mathcal{A} \mid XA = 0\}$.

Sigui $X \in \mathcal{A}$, com que $\pi(X)$ és generador de $\pi(\mathcal{B}')$, podem escriure $\pi(X) = \pi(Y)$, per alguna Y dels generadors de $\mathcal{B}'$. Per l'observació d'abans, $X - Y \in \{X \in \mathcal{A} \mid XA = 0\}$. Per tant, com que YA pertany a l'espai generat $< \mathcal{B}'A >$, tenim que XA pertany als generadors de $\mathcal{B}'A$, per qualsevol $X \in \mathcal{A} = F[A, B]$.

En altres paraules, qualsevol producte en $F[A, B]$ que tingui com a factor la matriu A està automàticament en $< \mathcal{B}'A >$, i per tant en el generador $< \mathcal{B} >$ de $\mathcal{B}$ ja que $\mathcal{B}'A \subseteq \mathcal{B}$. Observem també que $\omega_1 = nul(A)$, ja que és la dimensió del primer bloc de la seva forma de Weyr. A més,

$$\mathcal{B} = \{Id, B, B^2, ..., B^{\omega_1 - 1}\} \cup \mathcal{B}'A.$$

Per l'equació de Cayley-Hamilton generalitzada amb $d = \omega_1$, $< \mathcal{B} >$ conté totes les potències de B, i és invariant sota el producte per B i també és invariant sota el producte per A. Més encara, com que $< \mathcal{B} >$ conté Id, A i B, aleshores $F[A, B]$ està també continguda.

La inclusió en sentit contrària també la tenim, per tant $F[A, B]$ està generat per $\mathcal{B}$, com volíem veure, i això completa la inducció. $\square$

Ja tenim els 3 resultats previs que volíem per establir el *Teorema de Gerstenhaber*:

Teorema 7.8. ***Teorema de Gerstenhaber***

Si $A, B \in M_n$ són dues matrius que commuten, aleshores dim $F[A, B] \leq n$.

Per la dualitat entre la forma de Jordan y la de Weyr, el nostre conjunt generador de $F[A, B]$ del Teorema 7.7 té un dual. Si transposem el nostre conjunt generador (escrivint les columnes com files) i fent servir el fet que les estructures de Jordan i Weyr són particions duals, ja ho tenim.

Corol·lari 7.9. *Siguin $A, B \in M_n$ dues matrius nilpotents que commuten, i sigui $(m_1, ..., m_s)$ la característica de Segre de la matriu A. Aleshores el següent conjunt genera la subàlgebra $F[A, B]$ de $M_n(\mathbb{K})$*

$$\begin{gathered} Id, A, \dots, A^{m_1 - 1} \\ B, BA, BA^2, \dots, BA^{m_2 - 1} \\ \vdots \\ B^{s-1}, B^{s-1}A, \dots, B^{s-1}A^{m_s - 1}. \end{gathered}$$

Observació 7.10. Es pot veure que per $n \geq 4$, existeixen 4 matrius, A_1, A_2, A_3 i A_4, $n \times n$ tals que commuten i $dim\ F[A_1, A_2, A_3, A, 4] = n + 1$.

7.4 Relació entre les formes canòniques de Jordan de AB i BA

La relació entre la forma canònica de Jordan de les matrius producte AB i BA, amb $A, B \in M_n$, va ser descrita per primera vegada l'any 1951 per *Harley Flanders* ([7]). En el cas dels seus blocs regulars, els seus VAPs diferents de zero i les seves estructures de Jordan coincideixen. En canvi, als seus blocs singulars les seves dimensions difereixen en una unitat. Donades $A, B^T \in M_{n,m}$, volem estudiar la relació entre les formes de Jordan de les matrius producte AB i BA. Si A ó B són invertibles, la relació es trivial ja que si suposem que A és la matriu invertible, tenim que $AB = A(BA)A^{-1}$, i.e., les matrius AB i BA són semblants. Per tant, tenen els mateixos VAPs amb les mateixes multiplicitats, i la mateixa forma canònica de Jordan.

En cas que les dues matrius A i B fossin singulars, un argument semblant fent servir matrius del tipus $A + \epsilon Id$ ens assegura que les matrius AB i BA tindrien els mateixos VAPs i multiplicitats. Tot i això, en general AB i BA no serien semblants, i les seves formes canòniques de Jordan podrien ser diferents (més concretament, en les dimensions dels seus blocs associats al VAP $\lambda = 0$).

Aquest problema, va ser resolt per *Harley Flanders* a [7] però posteriorment es va presentar una demostració alternativa que detallarem a continuació, on fa servir la característica de Weyr ([5]).

Teorema 7.11. *Donades dues matrius* $A, B^T \in M_{n,m}$*,*

(i) *Els blocs de Jordan regulars de les matrius* AB *i* BA *són semblants, i.e., les seves característiques de Weyr són iguals:*

$$\omega_i(AB - \lambda Id) = \omega_i(BA - \lambda Id), \forall\ i,\ \lambda \neq 0.$$

(ii) *Per al VAP* $\lambda = 0$*, les característiques de Weyr de* AB *i* BA *satisfan*

$$\omega_{i-1}(AB) \geq \omega_i(BA) \geq \omega_{i+1}(AB),\ \forall\ i.$$

(iii) *O equivalentment en termes de la característica de Segre,*

$$|m_i(AB) - m_i(BA)| \leq 1,\ \forall\ i.$$

(iv) *A més, si* $M \in M_n, N \in M_m$ *són tals que* $\omega_i(M - \lambda Id) = \omega_i(N - \lambda Id),\ \lambda \neq 0,$ *i* $\omega_{i-1}(M) \leq \omega_i(N) \leq \omega_{i+1}(M)$*. Aleshores existeixen matrius* $A, B^T \in M_{n,m}$ *tals que* $M = AB$ *i* $N = BA$*.*

La demostració d'aquest teorema la realitzarem ajuntant els resultats dels Teoremes 7.12, 7.14 i del Lema 7.13.

En el següent teorema s'estableix la relació entre les característiques de Weyr de AB i BA. Els apartats (i) i (ii) del Teorema 7.11 són conseqüència del Teorema 7.12.

Teorema 7.12. *Donades dues matrius* $A, B^T \in M_{n,m}$*, per a tot* $i > 0$ *se satisfà*

$$\omega_i(AB - \lambda Id) = \omega_i(BA - \lambda Id), \ \forall \lambda \in \mathbb{K}, \lambda \neq 0,$$
$$\omega_i(BA) \geq \omega_{i+1}(AB), \ quan \ \lambda = 0.$$

Demostració. Suposem que $\lambda \neq 0$, aleshores per a qualsevol polinomi $p(x)$, $p(BA)B = Bp(AB)$. Així doncs, $p(AB)v = 0 \Rightarrow p(BA)Bv = 0$. Com que $Bv = 0 \Rightarrow p(AB)v = p(0)v$, tenim que $dim\ ker(p(AB)) = dim\ ker(p(BA))$, quan $p(0) \neq 0$. Per tant,

$$\omega_i(AB - \lambda Id) = \omega_i(BA - \lambda Id), \ \forall \lambda \in \mathbb{K}, \lambda \neq 0.$$

Suposem ara que $\lambda = 0$, i definim els següents nuclis, per $i \geq 0$:

$$\mathcal{R}_i := \{v \in \mathbb{K}^n \mid B(AB)^i v = 0\}, \ \mathcal{R}'_i := \{v \in \mathbb{K}^n \mid (AB)^i v = 0\},$$

$$\mathcal{L}_i := \{v \in \mathbb{K}^m \mid v^T(BA)^i = 0\}, \ \mathcal{L}'_i := \{v \in \mathbb{K}^m \mid v^T(BA)^i B = 0\}.$$

Observem que $\mathcal{R}_i \subset \mathcal{R}'_{i+1}$, $\mathcal{L}_i \subset \mathcal{L}'_{i+1}$, i que $dim\ \mathcal{R}_{i+1} - dim\ \mathcal{R}_i = dim\ \mathcal{L}'_{i+1} - dim\ \mathcal{L}'_i$.

Siguin $v_1, \ldots, v_k \in \mathcal{R}'_{i+2}$ un conjunt de vectors linealment independents mòdul $\mathcal{R}_{i+1}$. Així, $c_1v_1 + \cdots + c_kv_k \in \mathcal{R}_{i+1}$ només si $c_1 = \cdots = c_k = 0$. Aleshores els vectors $ABv_1, \ldots, ABv_k \in \mathcal{R}'_{i+1}$, són linealment independents mòdul $\mathcal{R}_i$. Per tant, $dim(\mathcal{R}'_{i+1}/\mathcal{R}_{i+1}) \geq dim(\mathcal{R}'_{i+2}/\mathcal{R}_{i+1})$. Si $v_1, \ldots, v_k \in \mathcal{L}'_{i+2}$ és un conjunt de vectors linealment independent mòdul $\mathcal{L}_{i+1}$, aleshores els vectors $(BA)^T v_1, \ldots, (BA)^T v_k \in \mathcal{L}'_{i+1}$ són linealment independents mòdul $\mathcal{L}_i$. Així doncs, $dim(\mathcal{L}'_{i+1}/\mathcal{L}_i) \geq dim(\mathcal{L}'_{i+2}/\mathcal{L}_{i+1})$. Observem que

$$dim(\mathcal{R}'_{i+2}/\mathcal{R}_{i+1}) = v_{i+2}(AB) - dim\ \mathcal{R}_{i+1}, \ \ dim(\mathcal{L}'_{i+2}/\mathcal{L}_{i+1}) = dim\ \mathcal{L}'_{i+2} - v_{i+1}(BA).$$

Aleshores, si $dim(\mathcal{R}'_{i+1}/\mathcal{R}_i) \geq dim(\mathcal{R}'_{i+2}/\mathcal{R}_{i+1})$, tenim que $dim\ \mathcal{R}_{i+2} - dim\ \mathcal{R}_{i+1} \geq v_{i+2}(AB) - v_{i+1}(AB)$,

i si $dim(\mathcal{L}'_{i+1}/\mathcal{L}_i) \geq dim(\mathcal{L}'_{i+2}/\mathcal{L}_{i+1})$, tenim que $v_{i+1}(BA) - v_i(BA) \geq dim\ \mathcal{L}'_{i+2} - dim\ \mathcal{L}'_{i+1}$.

I finalment, $\omega_{i+1}(BA) \geq \omega_{i+2}(AB)$, ja que $\omega_{i+1} = v_{i+1} - v_i$.

□

Ara volem traduir les relacions obtingudes al Teorema 7.12 a relacions de la característica de Segre i obtenir el resultat (*iii*) del Teorema 7.11. Per això presentem el següent lema.

Lema 7.13. *Siguin $p_1 \geq p_2 \geq \ldots$ una partició de n, i $p'_1 \geq p'_2 \geq \ldots$ una partició de n', amb particions conjugades $q_1 \geq q_2 \geq \ldots$, i $q'_1 \geq q'_2 \geq \ldots$, respectivament. Si fixem $d \in \mathbb{N}$, aleshores*

$$q'_i \geq q_{i+d}, q_i \geq q'_{i+d} \Leftrightarrow |p_i - p'_i| \leq d, \ \forall i > 0.$$

Demostració. Si $p'_i > d \Rightarrow q'_{p'_i} \geq i > q_{p_i+1}$. Per hipòtesis, $q_{p'_i-d} \geq q'_{p'_i} > q_{p_i+1}$, i llavors, $p'_i - d < p_i + 1$, ja que $\{q\}_j$ és monòtonament decreixent en j. Per tant, $p'_i \leq p_i + d$. Amb un argument simètric, tenim que $p_i \leq p'_i + d$.
Veiem ara l'altra implicació, si $q_{i+d} > 0 \Rightarrow p'_{q_i+d} \geq p_{q_i+d} - d \geq (i+d) - d = i > p'_{q'_i+1}$, on la primera desigualtat bé donada per les hipòtesis i les dos següents les obtenim conjugant les condicions. Com que p'_j és monòtonament decreixent, tenim que $q_{i+d} < q'_i + 1$, i per tant, $q_{i+d} \leq q'_i$, $\forall i > 0$. Anàlogament trobem que $q'_{i+d} \leq q_i$. □

Per la nostra demostració, nosaltres considerarem $\omega(AB) = \{p_i\}$, $\omega(BA) = \{p'_i\}$, $m_i(AB) = q_i$, i $m_i(BA) = \{q'_i\}$. Finalment, la part (iv) del Teorema 7.11 resulta del Teorema 7.14.

Teorema 7.14. *Siguin $m_1 \geq m_2 \geq \ldots$ i $m'_1 \geq m'_2 \geq \ldots$ particions de n i m, respectivament. Si $|m_i - m'_i| \leq 1$, aleshores existeixen matrius $A, B^T \in M_{n,m}(\mathbb{K})$ tals que $m_j(AB) = m_j$, i $m_j(BA) = m'_j$.*

Demostració. Per a cada j tal que $m_j, m'_j \geq 1$, construïm matrius $A_j, B_j^T \in M_{m_j,m'_j}(\mathbb{K})$ tals que $A_jB_j = J_{m_j}(0)$ i $B_jA_j = J_{m'_j}(0)$, d'acord amb els següents tres casos:

(1) $m_j = m'_j$: aleshores, $A_j = J_{m_j}(0), B_j = Id_{m_j}$,

(2) $m_j + 1 = m'_j$: aleshores, $A_j = [0\ Id_{m_j}], B_j = \begin{bmatrix} Id_{m_j} \\ 0 \end{bmatrix}$,

(3) $m_j = m'_j + 1$: aleshores, $A_j = \begin{bmatrix} Id_{m'_j} \\ 0 \end{bmatrix}$, $B_j = [0 Id_{m'_j}]$.

Això defineix $k := min\{\omega_1(AB), \omega_1(BA)\}$, per a cada parell de matrius (A_j, B_j). Considerem $\{m_j\}$ com la partició per n files i $\{m'_j\}$ com la partició per m columnes. Aleshores, construïm la matriu diagonal per blocs $A = diag(A_1, \ldots, A_k, 0, \ldots, 0)$, omplint amb zeros la part inferior a la dreta. Aleshores, amb particions $\{m'_j\}$ per m files i $\{m_j\}$ per n columnes tenim que $B = diag(B_1, \ldots, B_k, 0, \ldots, 0)$. □

Referències

[1] O'Meara, Kevin; Clark, John; Vinsonhaler, Charles: *Advanced Topics in Linear Algebra*. Oxford University Press, 2011.

[2] Strang, Gilbert: *Linear Algebra and its Applications*, 3a edició, Harcourt, Brace, Jovanovich, Publishers, 1988

[3] Brualdi, Richard A.; The Jordan canonical form: an old proof, *American Mathematical*, 94:257-267, 1987

[4] Helene Shapiro: The Weyr Characteristic, *The American Mathematical Monthly*, 106(10):919-929, 1999.

[5] Lippert, Ross A.; Strang, Gilbert: The Jordan forms of AB and BA, *Electron. J. Linear Algebra*, 18:281-288, 2009.

[6] A. F. Filippov: *A short proof of the theorem on the reduction of a matrix to Jordan form*, Vestnik, Moscow University, no. 2 (1971) 18-19.

[7] Harley Flanders. Elementary divisors of AB and BA. Proc. Amer. Math. Soc., 2(6):871–874, 1951. MR0044493.

[8] Horn, Roger A. Johnson; Charles R.: *Matrix Analysis*, 2a edició, Cambridge University Press, Cambridge, 2013.

[9] Jordan, Camille: *Traité des substitutions et des équations algébriques*, 1870.

[10] E. Weyr: Répartition des matrices en espèces et formation de toutes les espèces, *C. R. Acad. Sci. Paris*, 100:966-969, 1885.

[11] E. Weyr: Zur Theorie der bilinearen Formen, *Monatsh. Math. und Physik*, 1: 163-236, 1890.

[12] Barría, J; Halmos, P.: *Vector bases for two commuting matrices*, Linear and Multilinear Algebra, 27:147-157, 1990.

[13] Laffey, T.J.; Lazarus, S.: *Two-generated commutative matrix subalgebras*, Linear Algebra, 147:249-273, 1991.