CURSUL 12: FORMA JORDAN A UNEI MATRICE DIN $\mathcal{M}_n(\mathbb{C})$

G. MINCU

În acest curs, k va fi un corp comutativ, V un k-spaţiu vectorial de dimensiune $n \in \mathbb{N}^*$, matricele cu care lucrăm vor fi din $\mathcal{M}_n(k)$, iar bazele considerate vor fi ordonate.

1. MATRICI, VECTORI ŞI POLINOAME

În această primă secţiune, A va fi o matrice arbitrară, dar fixată, din $\mathcal{M}_n(k)$.

Observaţia 1. Exemplele de la cursul precedent au arătat că, dacă λ este o valoare proprie a matricei A, atunci se poate întâmpla să existe vectori $v \in k^n$ pentru care $(A - \lambda I_n)v \neq 0$, dar $(A - \lambda I_n)^2 v \neq 0$. Această observaţie conduce la următoarele chestiuni:

Definiţia 1. Dacă $\lambda \in \sigma_k(A)$, vom numi **vector propriu generalizat asociat lui** λ orice vector nenul $v \in k^n$ pentru care există $r \in \mathbb{N}^*$ astfel încât $(A - \lambda I_n)^r v = 0$.

Tot observaţia 1 conduce şi la constatarea că, notând $P = (X - \lambda)^2 \in k[X]$, avem $P(A)v = 0$ (pentru o corectă interpretare a expresiei $P(A)$ ţinem cont de identificarea canonică a elementului $\alpha \in k$ cu matricea $\alpha I_n \in \mathcal{M}_n(k)$). Aceste consideraţii ne sugerează să ne îndreptăm atenţia către expresiile de tipul $P(A)v$, unde $P \in k[X]$, $A \in \mathcal{M}_n(k)$, iar $v \in k^n$.

Propoziţia 1. Pentru orice $v \in k^n$ există polinoame nenule $P \in k[X]$ cu proprietatea $P(A)v = 0$.

Propoziţia 2. Dacă $v \in k^n$, atunci $I_v^A = \{P \in k[X] : P(A)v = 0\}$ este un ideal al lui $k[X]$.

Definiţia 2. Dacă $v \in k^n$, idealul I_v^A se numeşte **anulatorul** lui v.

Dacă matricea A este fixată în context, vom folosi pentru anulatorul lui v notaţia I_v. Dealtfel, o **notaţie** mai frecvent folosită pentru anulatorul lui v este $\mathrm{Ann}(v)$

Întrucât inelul $k[X]$ are toate idealele principale, anulatorul unui $v \in k[X]$ este ideal principal.

Definiţia 3. Dacă $v \in k^n$, vom numi **ordin** al lui v orice generator al lui I_v.

Observaţia 2. Dacă $v \in k^n$, ordinul său este unic până la o asociere în divizibilitate în $k[X]$. Din acest motiv, atunci când proprietăţile la care ne referim nu sunt afectate de o asociere în divizibilitate, vom spune, pe scurt, ,,ordinul lui v" în loc de ,,un ordin al lui v"

Vom nota ordinul lui $v \in k^n$ cu μ_v^A. Dacă matricea A este subînţeleasă din context, atunci vom folosi notaţia mai succintă μ_v.

Observaţia 3. Dacă $v \in k^n$, μ_v este caracterizat de proprietăţile:
(i) Este polinomul nenul P de grad minim pentru care $P(A)v = 0$.
(ii) Dacă $P \in k[X]$ şi $P(A)v = 0$, atunci $\mu_v | P$.

2. teorema Hamilton-Cayley

Până acum ne-am referit la expresii de tipul $P(A)v$ cu $P \in k[X]$, $A \in \mathcal{M}_n(k)$, iar $v \in k^n$. Dacă ne concentrăm asupra expresiilor de tip $P(A)$, cu P şi A ca mai sus, constatăm că:

Propoziţia 3. Dacă $A \in \mathcal{M}_n(k)$, atunci $I_A = \{P \in k[X] : P(A) = 0\}$ este un ideal al lui $k[X]$.

Nu este foarte clar a priori că acest ideal nu este trivial. Un argument în acest sens este teorema Hamilton-Cayley, pe care o prezentăm mai jos. Începem cu câteva chestiuni tehnice:

Fie $f = B_m X^m + \cdots + B_1 X + B_0 \in M_n(K)[X]$ şi $A \in M_n(K)$.

Definiţia 4. Matricea $f_d(A) = B_m A^m + \cdots + B_1 A + B_0$ se numeşte **valoarea la dreapta a lui f în** A.
Matricea $f_s(A) = A^m B_m + \cdots + AB_1 + B_0$ se numeşte **valoarea la stânga a lui f în** A.

Observaţia 4. $(If)(A)_d = (If)(A)_s = f(A)$.

Teorema lui Bézout generalizată. Fie $f \in M_n(K)[X]$ şi $A \in M_n(K)$. Atunci există $q \in M_n(K)[X]$ astfel încât $f = q(IX - A) + f_d(A)$ şi aceasta este unica scriere a lui f sub forma $f = q'(IX - A) + r$ cu $q' \in M_n(K)[X]$ şi $r \in M_n(K)$.

Analog, există $q \in M_n(K)[X]$ astfel încât $f = (IX - A)q + f_s(A)$ şi aceasta este unica scriere a lui f sub forma $f = (IX - A)q' + r$ cu $q' \in M_n(K)[X]$ şi $r \in M_n(K)$.

Demonstraţie: Fie $f = B_m X^m + \cdots + B_1 X + B_0$. Atunci $f_d(A) = B_m A^m + \cdots + B_1 A + B_0$. Deci

$$f - f_d(A) = B_m(IX^m - A^m) + \cdots + B_1(IX - A)$$

şi e suficient să observăm că

$$IX^k - A^k = (IX - A)(IX^{k-1} + \cdots + A^{k-1}).$$

Unicitatea. Fie $f = q(IX - A) + r = q'(IX - A) + r'$ cu $q, q' \in M_n(K)[X]$, $r, r' \in M_n(K)$. Atunci $(q - q')(IX - A) + r - r' = 0$, deci $q = q'$ şi $r = r'$, altfel polinomul $(q - q')(IX - A)$ are gradul ≥ 1. $\square$

Teorema Hamilton-Cayley. Fie $A \in M_n(K)$. Atunci $P_A(A) = 0$.

Demonstraţie: Fie $(IX - A)^*$ matricea adjunctă a matricei caracteristice $IX - A$. Atunci, $IP_A = I|IX - A| = (IX - A)^*(IX - A)$. Din teorema lui Bézout generalizată deducem că $0 = (IP_A)_d(A) = P_A(A)$. $\square$

3. Polinomul minimal al unei matrici

Ca de obicei, A desemnează o matrice fixată din $\mathcal{M}_n(k)$. Întrucât inelul $k[X]$ are toate idealele principale, I_A este ideal principal. Generatorii săi sunt asociaţi în divizibilitate, ceea ce înseamnă că I_A are exact un generator monic.

Definiţia 5. Generatorul monic al lui I_A se numeşte **polinomul minimal** al matricei A.

Vom nota polinomul minimal al lui $A \in \mathcal{M}_n(k)$ cu μ_A.

Propoziţia 4. Fie $A \in M_n(k)$ şi $f \in K[X]$. Următoarele afirmaţii sunt echivalente:
(i) $f = \mu_A$.
(ii) f este monic, $f(A) = 0$ şi f este de grad minim între polinoamele cu aceste proprietăţi.
(iii) f este monic, $f(A) = 0$ şi f divide toate polinomele $g \in K[X]$ cu proprietatea $g(A) = 0$.

4. Relaţii între tipurile de polinoame apărute până acum

Observaţia 5. Ca o consecinţă imediată a propoziţiei 4, deducem că pentru orice matrice $A \in \mathcal{M}_n(k)$ avem $\mu_A | P_A$.

Observaţia 6. Dacă $v \in k^n$, iar $A \in \mathcal{M}_n(k)$, atunci se constată, utilizând observaţia 3, că $\mu_v^A | \mu_A$. Aplicând şi observaţia 5, rezultă că are loc şi relaţia $\mu_v^A | P_A$.

Teorema 1. Dacă $A \in \mathcal{M}_n(\mathbb{C})$, atunci μ_A şi P_A au aceleaşi rădăcini.

Această teoremă este de fapt cazul particular relevant pentru discuţia noastră al teoremei lui Frobenius:

Teorema lui Frobenius. Dacă $A \in \mathcal{M}_n(k)$, atunci polinomul minimal şi polinomul caracteristic al lui A au aceiaşi factori ireductibili.

5. Subspaţii invariante

În acest paragraf vom lucra peste corpul $\mathbb{C}$ al numerelor complexe.

Notăm cu $(e_1, e_2, \ldots, e_n)$ baza canonică a lui $\mathbb{C}^n$.

Fie matricea $A \in M_n(\mathbb{C})$ care are polinomul caracteristic $P_A = (X-\lambda_1)^{m_1}(X-\lambda_2)^{m_2}\cdots(X-\lambda_r)^{m_r}$. Notăm cu F_i, $i \in \{1, 2, \ldots, r\}$, polinomul obţinut din P_A prin eliminarea factorului $(X-\lambda_i)^{m_i}$. Punem $w_{ij} = F_i(A)e_j$.

Observaţia 7. $\mu_{w_{ij}} = (X-\lambda_i)^{m_i}$

Propoziţia 5. Mulţimea $\bigcup\limits_{i,j} \{w_{ij}, (A-\lambda_i I_n)w_{ij}, \ldots, (A-\lambda_i I_n)^{m_i-1}w_{ij}\}$ constituie un sistem de generatori pentru ${}_{\mathbb{C}}\mathbb{C}^n$.

Demonstraţie: Întrucât $F_1, F_2, \ldots, F_r$ definite mai sus sunt prime între ele, există $G_1, G_2, \ldots, G_r \in \mathbb{C}[X]$ cu proprietatea $G_1F_1 + G_2F_2 + \ldots + G_rF_r = 1$. Obţinem

$$G_1(A)F_1(A)e_i + G_2(A)F_2(A)e_i + \ldots + G_r(A)F_r(A)e_i = e_i \tag{1}$$

pentru orice $i \in \{1, 2, \ldots, n\}$. Deoarece $(A-\lambda_i)^{m_i}e_i = 0$ conform observaţiei 7 putem înlocui în relaţia (1) fiecare G_j, $j \in \{1, 2, \ldots, r\}$ cu restul împărţirii sale la $(X-\lambda_i)^{m_i}$. Constatăm în acest mod că fiecare vector e_i, $i \in \{1, 2, \ldots, n\}$ este combinaţie liniară de elementele mulţimii $\bigcup\limits_{i,j} \{w_{ij}, (A-\lambda_i I_n)w_{ij}, \ldots, (A-\lambda_i I_n)^{m_i-1}w_{ij}\}$, ceea ce încheie demonstraţia. $\square$

Corolarul 1. ${}_{\mathbb{C}}\mathbb{C}^n$ admite o bază alcătuită din vectori proprii generalizaţi pentru A.

Fie $\lambda \in \sigma(A)$. **Notăm** $W_\lambda = \{v \in \mathbb{C}^n : \exists j \in \mathbb{N}^* \ (A-\lambda I_n)^j \cdot v = 0\}$.

Propoziţia 6. Fie $\lambda \in \sigma(A)$. Atunci, $W_\lambda \leq_{\mathbb{C}} \mathbb{C}^n$.

Definiţia 6. Fie $\lambda \in \sigma(A)$. W_λ se numeşte **subspaţiul vectorilor proprii generalizaţi asociaţi lui** λ.

Definiţia 7. Fie $A \in \mathcal{M}_n(k)$. k-subspaţiul vectorial V al lui k^n se numeşte A**- invariant** dacă pentru orice $v \in V$ avem $Av \in V$.

Propoziţia 7. Fie $A \in \mathcal{M}_n(\mathbb{C})$ şi $\lambda \in \sigma(A)$. Atunci, W_λ este subspaţiu vectorial A-invariant al lui ${}_{\mathbb{C}}\mathbb{C}^n$.

Propoziţia 8. Dacă $v_1, v_2, \ldots, v_r \in \mathbb{C}^n$ sunt vectori proprii generalizaţi corespunzători la valori proprii distincte ale matricei $A \in \mathcal{M}_n(\mathbb{C})$, atunci ei sunt liniar independenţi.

Din corolarul 1 şi din propoziţia 8 obţinem

Teorema 2. Dacă $A \in \mathcal{M}_n(\mathbb{C})$, iar $\sigma_{\mathbb{C}}(A) = \{\lambda_1, \lambda_2, \ldots, \lambda_r\}$, atunci $W_{\lambda_1} \dot{+} W_{\lambda_2} \dot{+} \ldots \dot{+} W_{\lambda_r}$ este o descompunere a lui $\mathbb{C}^n$ sub formă de sumă directă de subspaţii A-invariante.

Deşi descompunerea prezentată în teorema 2 generalizează rezultatul corespunzător din cazul matricilor diagonalizabile, vom căuta să o mai rafinăm, pentru a obţine matrici asemenea cu $A \in \mathcal{M}_n(\mathbb{C})$ de o formă cât mai apropiată de cea diagonală.

Fie $A \in \mathcal{M}_n(\mathbb{C})$ şi λ o valoare proprie a sa. Notăm $V_\lambda^{(s)} = \{v \in \mathbb{C}^n : (A - \lambda I_n)^s \cdot v = 0\}$.

Propoziţia 9. Fie $A \in \mathcal{M}_n(\mathbb{C})$ şi λ o valoare proprie a sa. Atunci, $V_\lambda^{(s)} \subset V_\lambda^{(s+1)}$ pentru orice $s \in \mathbb{N}$.

Propoziţia 10. În condiţiile propoziţiei 9, aplicaţia $V_\lambda^{(s+1)} \stackrel{(A-\lambda I_n)\cdot}{\longrightarrow} V_\lambda^{(s)}$ este (corect definită şi) liniară pentru orice $s \in \mathbb{N}$.

Propoziţia 11. Pentru orice $s \in \mathbb{N}$, aplicaţia liniară din propoziţia 10 induce o aplicaţie liniară şi injectivă $\dfrac{V_\lambda^{(s+2)}}{V_\lambda^{(s+1)}} \stackrel{(A-\lambda I_n)\cdot}{\longrightarrow} \dfrac{V_\lambda^{(s+1)}}{V_\lambda^{(s)}}$.

Vom identifica spaţiile $\dfrac{V_\lambda^{(s+2)}}{V_\lambda^{(s+1)}}$ cu imaginile lor prin morfismele de înmulţire cu $(A - \lambda I_n)$. Făcând acest lucru, putem vorbi de spaţiile vectoriale factor $U_s \stackrel{\text{not}}{=} \dfrac{\frac{V_\lambda^{(s)}}{V_\lambda^{(s-1)}}}{\frac{V_\lambda^{(s+1)}}{V_\lambda^{(s)}}}$ pentru orice $s \in \mathbb{N}^*$.

Fie acum $A \in \mathcal{M}_n(\mathbb{C})$, iar λ o valoare proprie cu multiplicitatea aritmetică m a lui A. Folosind notaţiile anterioare, alegem pentru fiecare $s \in \{1, 2, \ldots, m\}$ câte o mulţime $\{v_1, v_2, \ldots, v_{t_s}\} \subset \mathbb{C}^n$ astfel încât clasele vectorilor din aceste mulţimi să constituie baze în U_s. Punem $\mathcal{B}_s \overset{\text{not}}{=} \bigcup_{j=1}^{t_s}((A-\lambda I_n)^{(s-1)}v_j, (A-\lambda I_n)^{(s-2)}v_j, \ldots, v_j)$ - notaţia neobişnuită semnificând faptul că, deşi termenii reuniunii pot fi scrişi în ce ordine dorim, elementele pe care le-am scris în acel s-uplu trebuie să apară în reuniune la rând şi în acea ordine.

Teorema 3. În condiţiile anterioare, $\mathcal{B} \overset{\text{not}}{=} \bigcup_{s=1}^{m} \mathcal{B}_s$ este o bază a lui ${}_{\mathbb{C}}W_\lambda$.

Propoziţia 12. În condiţiile anterioare, $W_\lambda(s,j) \overset{\text{not}}{=}{}_{\mathbb{C}}< (A-\lambda I_n)^{(s-1)}v_j, (A-\lambda I_n)^{(s-2)}v_j, \ldots, v_j >$ este subspaţiu A-invariant al lui ${}_{\mathbb{C}}W_\lambda$ şi al lui ${}_{\mathbb{C}}\mathbb{C}^n$.

Teorema 4. În condiţiile propoziţiei 12, $W_\lambda = \dot{+}_{s,j} W_\lambda(s,j)$.

Teorema 5. În condiţiile teoremei 4, $\mathbb{C}^n = \dot{+}_{\lambda,s,j} W_\lambda(s,j)$.

Observaţia 8. Descompunerea lui $\mathbb{C}^n$ în sumă directă de subspaţii A-invariante pe care o dă teorema 5 este cea mai fină posibil. Baza $\mathcal{B}$ din teorema 3 este cea în care matricea are una dintre cele mai simple forme. Această formă se numeşte forma Jordan, şi va fi introdusă în paragraful următor.

6. Forma Jordan a unei matrice din $\mathcal{M}_n(\mathbb{C})$

Definiţia 8. Fie $\lambda \in \mathbb{C}$ şi $r \in \mathbb{N}^*$. Matricea

$$J_r(\lambda) = \begin{pmatrix} \lambda & 1 & 0 & 0 & \ldots & 0 & 0 & 0 \\ 0 & \lambda & 1 & 0 & \ldots & 0 & 0 & 0 \\ 0 & 0 & \lambda & 1 & \ldots & 0 & 0 & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & 0 & 0 & \ldots & \lambda & 1 & 0 \\ 0 & 0 & 0 & 0 & \ldots & 0 & \lambda & 1 \\ 0 & 0 & 0 & 0 & \ldots & 0 & 0 & \lambda \end{pmatrix}$$

se numeşte **celula Jordan de ordin** r asociată lui λ.

Definiţia 9. Numim matrice canonică Jordan orice matrice de forma

$$\begin{pmatrix} J_{r_1}(\lambda_1) & 0 & 0 & \ldots & 0 & 0 \\ 0 & J_{r_2}(\lambda_2) & 0 & \ldots & 0 & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & 0 & \ldots & 0 & J_{r_t}(\lambda_t) \end{pmatrix}.$$

O consecinţă imediată a teoremei 3 este

Teorema 6. Dacă $A \in \mathcal{M}_n(\mathbb{C})$, atunci există o matrice canonică Jordan J_A astfel încât $A \approx J_A$. Matricea J_A este unic determinată de A, abstracţie făcând de ordinea celulelor de pe diagonală.

Definiţia 10. J_A ca în teorema anterioară se numeşte forma canonică Jordan a lui A.

Observaţia 9. Matricea $A \in \mathcal{M}_n(\mathbb{C})$ este diagonalizabilă dacă şi numai dacă forma sa canonică Jordan este diagonală.

References

[1] T. Dumitrescu, *Algebra*, Ed. Universităţii din Bucureşti, 2006.
[2] I. D. Ion, N. Radu, *Algebra*, Ed. Universităţii din Bucureşti, 1981.
[3] C. Năstăsescu, C. Niţă, C. Vraciu, *Bazele algebrei*, Ed. Academiei, Bucureşti, 1986.