Universitat de Barcelona

Apunts

Segon Semestre

Àlgebra Lineal (AL)

Autor:
Mario Vilar

Professor:
Dr. Luis Dieulefait

3 de juny de 2021

Segona part del temari

Introducció

This silence for my sin you did impute,
Which shall be most of my glory, being dumb;
For I impair not beauty, being mute,
When other would give life and bring a tomb.

WILLIAM SHAKESPEARE

Abans de començar, una petita introducció a aquests apunts pels quals, a causa de la seva extensió, es pot veure fàcilment que he ampliat més que no pas reduït el temari...

Primer de tot, es trobarà que hi ha un índex, on hi distingim els diferents apartats ordenats d'una manera certament poc satisfactòria pel que fa a l'ordre cronològic del curs, sinó que he seguit més aviat el meu propi criteri. Hi ha capítols, seccions, subseccions (i fins i tot subsubseccions). Pel que fa al pes d'aquestes estructures en els encapçalaments:

1. el número de l'última secció/subsecció figurarà en cada cantonada superior de pàgina parella (per exemple, *1.2*);
2. el nom del capítol es trobarà a la part dreta de la capçalera de les pàgines parelles (per exemple, "Divisibilitat i nombres primers");
3. el nom de l'última secció/subsecció de la pàgina, a la cantonada dreta superior de les pàgines parelles (per exemple, "Polinomis: algorisme d'Euclides").

A més, hi ha una taula en què es veu fàcilment que s'ha seguit una mena de *sorting-by-color* per poder treballar de manera més eficient amb els diferents tipus d'enunciats matemàtics. En els encapçalaments, aleshores, tenim:

- el número de l'últim teorema, definició... de la pàgina en qüestió es trobarà a les pàgines senars, a la cantonada superior dreta (per exemple, **1.2.3**).

Teorema de prova. *Aquest és un teorema de prova. Els teoremes, les proposicions, els lemes, els corol·laris, les propietats, les conjectures i els processos tindran aquest format.*

Definició de prova. Aquesta és una definició de prova. Les definicions, els exemples i les notacions tindran aquest format.

Remarca de prova. Aquesta és una remarca de prova. Les remarques tindran aquest format.

Figura 1: Els diferents formats d'enunciats.

Índex

Taula

Capítol 1

Diagonalització

Donat un endomorfisme $f : E \longrightarrow E$ ens proposem trobat una base de E de manera que la matriu de f en aquesta base sigui el més senzilla possible. En alguns casos, veurem que podem trobar una base en què la matriu de f és diagonal. En aquest cas, diem que la matriu de f és diagonalitzable.

1.1 Valors i vectors propis

Definició 1.1.1 (Matriu diagonal). Donats escalars $\lambda_1, \ldots, \lambda_n$, on $n \geq 1$ és un enter, denotarem per $D(\lambda_1, \ldots, \lambda_n)$ la matriu diagonal de mida $n \times n$ donada per

$$D(\lambda_1, \ldots, \lambda_n) := \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & \lambda_n \end{pmatrix}. \tag{1.1.1}$$

Definició 1.1.2 (Vector propi). Diem que un vector no nul $u \in E$ és un *vector propi de* f si existeix un escalar $\lambda \in \mathbb{K}$ tal que $f(u) = \lambda u$.

Definició 1.1.3 (Valor propi). Si u és un vector propi de f, l'escalar λ s'anomena *valor propi* de u.

Observació 1.1.4.

1. La definició de vector propi i de valor propi d'un endomorfisme no utilitza bases.
2. En la definició que hem donat de vector propi s'exigeix que el vector sigui diferent de zero. El vector $0 \in E$ satisfà l'equació $f(0) = \lambda \cdot 0$ per a tot $\lambda \in \mathbb{K}$, però 0 no se sol considerar com a valor propi de f.
3. Si u és un vector propi λ aleshores, per a tot $a \in \mathbb{K}, a \neq 0$, el vector au també és un vector propi de f del mateix valor propi (unicitat del valor propi).
4. Els vectors no nuls del nucli de f són vectors propis de valor propi $\lambda = 0$.

Exemple 1.1.5. Si A és una matriu diagonal,

$$A = D(\lambda_1, \lambda_2) = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}, \tag{1.1.2}$$

aleshores,

$$A \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \lambda_1 \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad A \cdot \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ \lambda_2 \end{pmatrix} = \lambda_2 \cdot \begin{pmatrix} 0 \\ 1 \end{pmatrix}. \tag{1.1.3}$$

Definició 1.1.6 (Subespai propi de f de valor propi λ). Si $\lambda \in \mathbb{K}$, denotarem per E_λ el subespai vectorial d'E donat per

$$E_\lambda := \ker(f - \lambda I). \tag{1.1.4}$$

$u \neq 0$ és vector propi de f si, i només si, $f(u) - \lambda u = 0$. Això equival a demanar que $u \in E_\lambda$. D'altra banda, un escalar $\lambda \in \mathbb{K}$ és valor propi si, i només si, $E_\lambda \neq \{0\}$.

Definició 1.1.7 (Multiplicitat geomètrica i algebraica). Sigui λ un valor propi de f, Definim la *multiplicitat geomètrica* de λ com la dimensió de l'espai vectorial E_λ. La *multiplicitat algebraica* és la multiplicitat de λ com a arrel del polinomi característic. $MG \leq MA$.

Proposició 1.1.8. *Si r és la multiplicitat del valor propi k, és a dir, si es té $r = \dim(\ker(f - k\mathbb{I}))$ i s és la multiplicitat del zero k del polinomi característic, aleshores $r \leq s$.*

Proposició 1.1.9. *Un escalar $\lambda \in \mathbb{K}$ és valor propi de f si, i només si,* $\det(f - \lambda I) = 0$.

Demostració. Sabem que el subespai $E_\lambda := \ker(f - \lambda I)$ és trivial si, i només si, la imatge de $f - \lambda I$ té rang màxim. Això passa si, i només si, $\det(f - \lambda I) \neq 0$, tal i com volíem provar. ■

Definició 1.1.10 (Polinomi característic). El polinomi característic d'una matriu quadrada A és $p_A(\lambda) := \det(A - \lambda I)$, on

$$\begin{pmatrix} a_1^1 - \lambda & a_2^1 & \dots & a_n^1 \\ a_1^2 & a_2^2 - \lambda & \dots & a_n^2 \\ \vdots & \vdots & \ddots & \vdots \\ a_1^n & a_2^n & \dots & a_n^n - \lambda \end{pmatrix} \tag{1.1.5}$$

correspon a la matriu $A - \lambda I$. Aquest polinomi no depèn de la base escollida, sinó de l'endomorfisme en si: parlem del polinomi característic d'un *endomorfisme*, $p_f(\lambda)$.

Propietat 1.1.11 (Resum dels passos per al càlcul de valors i vectors propis).

1. *Buscar les arrels $\{\lambda_1, \dots, \lambda_n\}$ del polinomi característic $p_A(\lambda) := \det(A - \lambda I)$. Notem que poden haver-hi diverses arrels amb el mateix valor.*
2. *Per a cada arrel λ_i: buscar les solucions del sistema homogeni*

$$(A - \lambda_i I) \cdot \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}. \tag{1.1.6}$$

3. *Escollir vectors representant d'aquestes solucions per a descriure els subespais propis i la multiplicitat de cada valor propi λ_i.*

1.2

Teorema de diagonalització

Sigui E un espai vectorial de dimensió finita $n = \dim E$. Sigui $f : E \longrightarrow E$ un endomorfisme i suposem que hem trobat una base $\mathfrak{B} = \{u_1, \ldots, u_n\}$ de E tal que la matriu de f en aquesta base és diagonal $D(\lambda_1, \ldots, \lambda_n)$. Aleshores, és evident que els vectors u_i són vectors propis de valor propi λ_i. Recíprocament, suposem una base $\mathfrak{B}' = \{v_1, \ldots, v_n\}$ de vectors propis de f. Aleshores, la matriu de f en la base $\mathfrak{B}'$ és una matriu diagonal $D(\lambda_1, \ldots, \lambda_n)$ on λ_i és el valor propi associat al vector propi v_i.

Definició 1.2.1 (Diagonalitzable). Direm que un endomorfisme és diagonalitzable si existeix una base de E formada per vectors propis de f.

Observació 1.2.2. Diagonalitzar un endomorfisme f vol dir trobar una base de vectors propis de f, mentre que diagonalitzar una matriu A vol dir trobar una matriu diagonal equivalent a A. En altres paraules, l'endomorfisme f és diagonalitzable si, i només si, l'endomorfisme associat és, en efecte, diagonalitzable.

Proposició 1.2.3. *Siguin $u, u' \in E$ vectors propis de f amb valors propis $\lambda, \lambda' \in \mathbb{K}$ respectivament. Si $\lambda \neq \lambda'$, aleshores u, u' són linealment independents.*

Demostració. Suposem que $au + a'u' = 0$ per a certs $a, a' \in \mathbb{K}$. Aplicant la funció f a banda i banda ens queda que $af(u) + a'f(u') = f(0) = 0$. Com que per hipòtesi ($u, u' \in E$ *vectors propis*) tenim que $f(u) = \lambda u$ i $f(u') = \lambda' u'$, obtenim que $a\lambda u + a'\lambda' u' = 0$. Podem suposar que $\lambda \neq 0$ (ja que $\lambda \neq \lambda'$). Ara, multiplicant la primera equació per λ obtenim que $\lambda au + \lambda a'u' = 0$. Notem que tenim les dues expressions:

$$\begin{aligned} \lambda au + \lambda a'u' &= 0, \\ a\lambda u + a'\lambda' u' &= 0. \end{aligned} \tag{1.2.1}$$

Operant-les, obtenim que $a'(\lambda - \lambda')u' = 0$. Com que $u' \neq 0$ i $\lambda \neq \lambda'$, tenim que $a' = 0$. Per tant, també tenim $a = 0$ i els vectors u, u' són linealment independents. D'aquí extraurem els següents corol·laris. ■

Corol·lari 1.2.4. *Un conjunt de vectors propis $\{v_1, \ldots, v_r\}$ amb valors propis $\{\lambda_1, \ldots, \lambda_r\}$ tal que $\lambda_i \neq \lambda_j$, aleshores els vectors $\{v_1, \ldots, v_r\}$ són linealment independents.*

Demostració. Per inducció simple sobre r. ■

Corol·lari 1.2.5. *El nombre de valors propis diferents és $\leq n$. Si hi ha exactament n, f és diagonalitzable.*

Exemple 1.2.6. Considerem $f : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ amb matriu

$$A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \tag{1.2.2}$$

en la base usual $(1, 0), (0, 1)$. El polinomi característic és $(1 - x)^2$ i, per tant, l'únic valor propi és 1. Si f diagonalitzés, la seva matriu diagonal seria $A = \mathbb{I}_2$, la qual cosa no és certa. De fet, el subespai de vectors propis és de dimensió 1: $\ker(f - \mathbb{I}) = \langle (1, 0) \rangle$.

Corol·lari 1.2.7. *La suma de subespais propis és directa: si* $\{\lambda_1, \dots, \lambda_r\}$ *són tals que* $\lambda_i \neq \lambda_j$ *aleshores*

$$E_{\lambda_1} + E_{\lambda_2} + \cdots + E_{\lambda_r} = E_{\lambda_1} \oplus E_{\lambda_2} \oplus \cdots \oplus E_{\lambda_r}. \tag{1.2.3}$$

Corol·lari 1.2.8. f *és diagonalitzable si, i només si,* $E = E_{\lambda_1} \oplus E_{\lambda_2} \oplus \cdots \oplus E_{\lambda_r}$, *on* $\lambda_i, \dots, \lambda_r$ *són tals que* $\lambda_i \neq \lambda_j$, *per a* $i \neq j$.

Definició 1.2.9 (Matrius equivalents). A, B són equivalents si són dues matrius de l'endomorfisme f associades a diferents bases d'E, sabem que estan relacionades per una matriu invertible (la matriu de canvi de base), mitjançant la fórmula $A = U^{-1}BU$.

Lema 1.2.10. *Sigui* A *una matriu de mida* $n \times n$. *Per a tota matriu invertible* U *de mida* $n \times n$ *es té que* $\det(U^{-1}AU - \lambda I) = \det(A - \lambda I)$.

Demostració. En primer lloc, se satisfà que

$$U^{-1}AU - \lambda I = U^{-1}AU - U^{-1}\lambda IU = U^{-1}(A - \lambda I)U, \tag{1.2.4}$$

i per les propietats multiplicatives del determinant, $\det(U^{-1}AU - \lambda I) = \det(U^{-1})\det(A - \lambda I)\det(U) = \det(A - \lambda I)$. ■

Definició 1.2.11. Sigui $f : E \longrightarrow E$ un endomorfisme en E. El polinomi característic de f és $p_f(\lambda) := p_A(\lambda) = p_B(\lambda) = \det(A - \lambda I)$. Notem, com havíem dit anteriorment, que dues matrius són equivalents els seus polinomis característics són iguals, ja que no depenen de la base escollida.

Teorema 1.2.12 (Teorema de diagonalització). *Un endomorfisme* f *és diagonalitzable si, i només si, les dues condicions següents es verifiquen:*

1. *El polinomi característic de* f *descompon com a producte de* n *factors de grau 1 sobre* $\mathbb{K}$, *iguals o repetits. En altres paraules:*

$$p_f(x) = (-1)^n(x - \lambda_1)^{m_1} \cdots (x - \lambda_r)^{m_r},\ \lambda_i \neq \lambda_j,\ m_1 + \cdots + m_r = n. \tag{1.2.5}$$

   Per tant, $p_f(\lambda)$ *té* n *arrels a* $\mathbb{K}$ *comptades amb multiplicitat.*
2. *Per a tot valor propi* λ_i *es té que* $\dim E_{\lambda_i} = m_i$.

Demostració. Suposem que f és diagonalitzable. Aleshores, hi ha una base $\{v_1, \dots, v_n\}$ de E on v_i és vector propi de valor propi λ_i. La matriu de f en aquesta base és $D(\lambda_1, \dots, \lambda_n)$. Per tant, tenim

$$p_f(\lambda) = (\lambda_1 - \lambda) \cdots (\lambda_n - \lambda) = (-1)^n(\lambda - \lambda_1) \cdots (\lambda - \lambda_n). \tag{1.2.6}$$

$\Longrightarrow$ Reagrupant els termes amb λ iguals s'obtindria una expressió com la de (1.2.6), ja que, notem, els λ_i no tenen per què ser diferents (una vegada reagrupats sí). A més, és clar que per a tot $\lambda \in \mathbb{K}$ el subespai propi E_λ està generat pels vectors v_i de la base tals que $\lambda_i = \lambda$. Per tant, la dimensió de E_λ és la multiplicitat de λ com a arrel de $p_f(\lambda)$, provant així 1.2.12.2.

$\Longleftarrow$ Suposem que ambdues condicions del teorema són certes. Aleshores,

$$\sum \dim E_{\lambda_i} = m_1 + \cdots + m_r = n \tag{1.2.7}$$

i, per tant, $E = E_{\lambda_1} \oplus \cdots \oplus E_{\lambda_r}$. Així doncs, f diagonalitza.

■

Observació 1.2.13. Hem d'anar amb molt de compte amb el cos sobre el qual està definit E, en particular per a la primera condició.

- Notem que en el cas que $\mathbb{K} = \mathbb{R}$, la condició se satisfà si, i només si, les arrels són totes reals. En canvi, en el cas que $\mathbb{K} = \mathbb{C}$, la primera condició sempre se satisfà.
- Pel que fa la segona condició, no se satisfà sempre, ni sobre $\mathbb{R}$ ni sobre $\mathbb{C}$. La segona condició del teorema demana que multiplicitat algebraica i multiplicitat geomètrica coincideixin.
    1. En el cas que la multiplicitat algebraica de totes les arrels sigui 1 (tot arrels simples), aleshores la segona condició se satisfà automàticament (ja que si λ és valor propi aleshores $E_\lambda \neq \{0\}$ i $\dim E_\lambda > 0$).
    2. Si la multiplicitat algebraica d'alguna de les arrels és $m > 1$ aleshores cal comprovar la segona condició buscant m vectors propis linealment independents de valor propi λ, per verificar que la multiplicitat geomètrica de λ és m.

Exemple 1.2.14. Considerem la matriu

$$A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}. \tag{1.2.8}$$

Aleshores, $p_B(\lambda) = (1-\lambda)^2$, així que $p_A(\lambda)$ té una única arrel $\lambda = 1$ amb multiplicitat 2. La primera condició de 1.2.12 se satisfà. Anem a veure la segona. Calculem el subespai propi E_1. Per fer-ho cal solucionar el sistema

$$\begin{pmatrix} 1-\lambda & 1 \\ 0 & 1-\lambda \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \implies y = 0. \tag{1.2.9}$$

Aleshores, el subespai propi E_1 està format per vectors de la forma $\{x, 0\}$, amb $x \neq 0$. Podem escriure $E_1 = \langle (1,0) \rangle$ i, per tant, $\dim E_1 = 1$. Observem, però, que la multiplicitat de $\lambda = 1$ és 2: no es compleix la segona condició de 1.2.12 i A no diagonalitza.

Capítol 2

Formes bilineals i espais vectorials euclidians

2.1 Formes bilineals i producte escalar

Considerarem un espai vectorial E sobre un cos $\mathbb{K}$ de dimensió finita n.

Definició 2.1.1 (Forma bilineal en E)**.** És una aplicació

$$\begin{array}{rccc} \varphi: & E \times E & \longrightarrow & \mathbb{K} \\ & (u, v) & \longmapsto & \varphi(u, v) := \lambda \end{array} \tag{2.1.1}$$

Podem veure, doncs, que φ assigna un escalar a cada parell de vectors i que és lineal en les dues variables.

Propietat 2.1.2 (Propietats de les formes bilineals)**.** *Com que és lineal en les dues variables, podem dir que satisfà, donats $u, v, u_1, u_2, v_1, v_2 \in E$ i $\lambda \in \mathbb{K}$:*

1. $\varphi(u_1 + u_2, v) = \varphi(u_1, v) + \varphi(u_2, v)$,
2. $\varphi(u, v_1 + v_2) = \varphi(u, v_1) + \varphi(u, v_2)$,
3. $\varphi(\lambda u, v) = \lambda\varphi(u, v)$,
4. $\varphi(u, \lambda v) = \lambda\varphi(u, v)$.

Definició 2.1.3 (Forma bilineal simètrica)**.** Diem que φ és simètrica si $\varphi(u, v) = \varphi(v, u)$ per a tot $u, v \in E$.

Definició 2.1.4 (Forma bilineal definida positiva)**.** φ és definida positiva si $\varphi(u, u) \geq 0$ per a tot $u \neq 0$ i $\varphi(u, u) = 0$ si, i només si, $u = 0$.

Sigui $\varphi : E \times E \longrightarrow K$ una aplicació bilineal. Donada una base $\mathfrak{B} = \{e_1, \ldots, e_n\}$ de E denotem $g^i_j := \varphi(e_i, e_j)$.

Definició 2.1.5 (Matriu de Gram)**.** Anomenem *Matriu de Gram* en la base $\mathfrak{B}$ la matriu

$$\mathcal{G}_{\mathfrak{B}}(\varphi) := \begin{pmatrix} g^1_1 & \cdots & g^1_n \\ \vdots & \ddots & \vdots \\ g^n_1 & \cdots & g^n_n \end{pmatrix} \tag{2.1.2}$$

Observem que gràcies a la bilinealitat, per a vectors u, v amb coordenades $(a_1, \dots, a_n)$ i $(b_1, \dots, b_n)$ respectivament, tenim que

$$\varphi(u, v) = (a_1, \dots, a_n) \cdot \mathcal{G}_{\mathfrak{B}}(\varphi) \cdot \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}; \tag{2.1.3}$$

aquesta expressió se sol escriure com $\varphi(u, v) = u^t_{\mathfrak{B}} \cdot \mathcal{G}_{\mathfrak{B}}(\varphi) \cdot v_{\mathfrak{B}}$, on $v_{\mathfrak{B}} = \mathcal{M}_{\mathfrak{B}}(v)$ denota el vector columna de coordenades de v en la base $\mathfrak{B}$ i $u^t_{\mathfrak{B}} = \mathcal{M}_{\mathfrak{B}}(u)^t$ denota el vector de coordenades de u en la base $\mathfrak{B}$. Quan la base queda clara pel context, simplifiquem l'expressió

Observació 2.1.6. La matriu de Gram és la única matriu que satisfà $\varphi(u, v) = u^t_{\mathfrak{B}} \cdot \mathcal{G}_{\mathfrak{B}}(\varphi) \cdot v_{\mathfrak{B}}$, per a tot $u, v \in E$, atès que $\varphi(e_i, e_j) = g^i_j = e^t_i \cdot \mathcal{G}_{\mathfrak{B}}(\varphi) \cdot e_{\mathfrak{B}}$.

Proposició 2.1.7. *Sigui φ una forma bilineal en E. Són equivalents:*

1. *φ és simètrica,*
2. *per a tota base de E la matriu de Gram és simètrica,*
3. *existeix una base tal que la matriu de Gram en aquesta base és simètrica.*

Demostració.

$\boxed{1 \Rightarrow 2}$ Si φ és simètrica aleshores $g^i_j = g(e_i, e_j) = g(e_j, e_i) = g^j_i$. Així doncs, φ és simètrica, tal i com volíem provar

$\boxed{2 \Rightarrow 3}$ És trivial.

$\boxed{3 \Rightarrow 1}$ Escollim una base $\mathcal{B}$ per la qual $\mathcal{G}_{\mathfrak{B}}(\varphi)$ és simètrica. Aleshores,

$$\varphi(u, v) = u^t \cdot \mathcal{G}_{\mathfrak{B}}(\varphi) \cdot v = (u^t \cdot \mathcal{G}_{\mathfrak{B}}(\varphi) \cdot v)^t = v^t \cdot \mathcal{G}_{\mathfrak{B}}(\varphi)^t \cdot u = \varphi(v, u), \tag{2.1.4}$$

on en la segona igualtat hem usat que $u^t \cdot \mathcal{G}_{\mathfrak{B}}(\varphi) \cdot v$ és una matriu 1×1. En la tercera, hem aplicat les propietats de la transposició i en la quarta, la definició de forma bilineal.

■

Definició 2.1.8 (Matriu de canvi de base). Donades dues bases $\mathfrak{B} = \{e_1, \dots, e_n\}$ i $\mathfrak{B}' = \{e'_1, \dots, e'_n\}$, denotem per $\mathcal{P} = \mathcal{M}_{(\mathfrak{B}, \mathfrak{B}')}(Id)$ la matriu de canvi de base de $\mathfrak{B}$ a $\mathfrak{B}'$. Amb la notació introduïda abans, tenim que $u_{\mathfrak{B}'} = \mathcal{P} \cdot u_{\mathfrak{B}}$.

Proposició 2.1.9 (Canvi de base). *Sigui φ una forma bilineal en E. Denotem per $\mathcal{G}_{\mathfrak{B}}(\varphi)$ i $\mathcal{G}_{\mathfrak{B}'}(\varphi)$ les matrius de Gram de φ en les bases $\mathfrak{B}$ i $\mathfrak{B}'$, respectivament. Aleshores, es té*

$$\mathcal{G}_{\mathfrak{B}}(\varphi) = \mathcal{P}^t \cdot \mathcal{G}_{\mathfrak{B}'}(\varphi) \cdot \mathcal{P}, \tag{2.1.5}$$

on $\mathcal{P} = \mathcal{M}_{(\mathfrak{B}, \mathfrak{B}')}$ és la matriu de canvi de base.

Demostració. Donats $u, v \in E$ tenim

$$\varphi(u, v) = u^t_{\mathfrak{B}'} \cdot \mathcal{G}_{\mathfrak{B}'}(\varphi) \cdot u_{\mathfrak{B}'} = (P \cdot u_{\mathfrak{B}})^t \cdot \mathcal{G}_{\mathfrak{B}'}(\varphi) \cdot (P \cdot u_{\mathfrak{B}}) = u^t_{\mathfrak{B}} \cdot \mathcal{P}^t \cdot \mathcal{G}_{\mathfrak{B}'}(\varphi) \cdot P \cdot u_{\mathfrak{B}}. \tag{2.1.6}$$

Com que $\mathcal{G}_{\mathfrak{B}}(\varphi)$ és la única matriu tal que $\varphi(u, v) = u^t_{\mathfrak{B}} \cdot \mathcal{G}_{\mathfrak{B}'}(\varphi) \cdot u_{\mathfrak{B}}$, obtenim la igualtat

$$\mathcal{G}_{\mathfrak{B}}(\varphi) = \mathcal{P}^t \cdot \mathcal{G}_{\mathfrak{B}'}(\varphi) \cdot P. \tag{2.1.7}$$

■

Observació 2.1.10. Notem que, en contrast amb els canvis de base per endomorfismes, ens apareix la transposada $\mathcal{P}^t$ en comptes de la inversa $\mathcal{P}^{-1}$.

2.2 Espais vectorials Euclidians

Considerarem $\mathbb{K} = \mathbb{R}$ i més endavant comentarem les particularitats de $\mathbb{K} = \mathbb{C}$ o d'un altre cos arbitrari. Escriurem $u \cdot v$ com uv, però no uu en lloc d'$u \cdot u$. Anàlogament amb v.

Definició 2.2.1 (Producte escalar)**.** Un *producte escalar* en un espai vectorial E és una forma bilineal $\varphi : E \times E \longrightarrow \mathbb{R}$ que és simètrica i definida positiva. Un espai vectorial Euclidià és aquell $\mathbb{R}$-espai vectorial dotat d'un producte escalar.

Notació 2.2.2. Donat un producte escalar $\varphi : E\times E \longrightarrow E$, també se sol denotar $\varphi(u, v) := uv$.

Sigui $(E, \cdot)$ un espai vectorial Euclidià. Tenim el següent conjunt de definicions, força importants.

Definició 2.2.3 (Vectors ortogonals)**.** Dos vectors u, v són ortogonals si $uv = 0$.

Definició 2.2.4 (Vector unitari)**.** Un vector u és unitari si $u \cdot u = 1$.

Definició 2.2.5 (Base ortogonal)**.** Una base $\mathfrak{B} = \{e_1, \dots, e_n\}$ és *ortogonal* si $e_ie_j = 0, \forall i \neq j$ (és a dir, els vectors de la base són ortogonals dos a dos). Això equival a demanar que la matriu de Gram $\mathcal{G}$ sigui diagonal en aquesta base.

Definició 2.2.6 (Base ortonormal)**.** Una base $\mathfrak{B} = \{e_1, \dots, e_n\}$ és *ortonormal* si és ortogonal i està formada per vectors unitaris. Equivalentment, si la matriu de Gram $\mathcal{G}$ és la identitat en aquesta base.

Definició 2.2.7 (Norma)**.** Sigui E un espai vectorial sobre $\mathbb{R}$ o $\mathbb{C}$. Una norma a E és una aplicació

$$\begin{array}{rccl} ||\quad||: & E & \longrightarrow & \mathbb{R} \text{ (sempre a } \mathbb{R}!) \\ & v & \longmapsto & ||v|| \end{array} \tag{2.2.1}$$

que compleix

1. $||v|| = 0 \iff v = \vec{0}$,
2. $||kv|| = |k| \cdot ||v||$,
3. $||u + v|| \leq ||u|| + ||v||$ (desigualtat triangular).

$|k|$ indica el valor absolut si $k \in \mathbb{R}$ o bé indica el mòdul si $k \in \mathbb{C}$.

Lema 2.2.8 (Desigualtat de Cauchy-Schwarz)**.** *Es compleix que*

$$|uv|^2 \leq (u \cdot u)(v \cdot v), \forall u, v \in E \tag{2.2.2}$$

Demostració. Si $v = \vec{0}$, aleshores la desigualtat és certa. Suposem $v \neq \vec{0}$ i considerem $k = \frac{uv}{v \cdot v}$. Aleshores:

$$\begin{aligned} 0 \leq (u - kv)(u - kv) &= u \cdot u - k(vu) - k(uv) + k\overline{k}(v \cdot v) \\ &= u \cdot u - \frac{(uv)(vu)}{v \cdot v} - \frac{\overline{(uv)}(uv)}{v \cdot v} + \frac{(uv)\overline{(uv)}}{v \cdot v} \\ &= u \cdot u - \frac{(uv)\overline{(uv)}}{v \cdot v} = u \cdot u - \frac{|uv|^2}{v \cdot v}, \end{aligned} \tag{2.2.3}$$

d'on $|uv|^2 \leq (u \cdot u)(v \cdot v), \forall u, v \in E$. ∎

Proposició 2.2.9. *Sigui E un espai vectorial amb un producte escalar. L'aplicació*

$$\begin{array}{rccl} || \quad || : & E & \longrightarrow & \mathbb{R} \\ & v & \longmapsto & ||v|| := \sqrt{v \cdot v} \end{array} \tag{2.2.4}$$

és una norma.

Demostració. La primera condició de norma resulta del fet que el producte escalar és definit positiu. Per provar la segona, observem que $(kv)(kv) = k\overline{k}(v \cdot v) = |k|^2(v \cdot v)$. Per demostrar l'última, la tercera, fem el següent:

$$\begin{aligned} (u+v)(u-v) &= u \cdot u + uv + vu + v \cdot v \\ &= u \cdot u + v \cdot v + (uv + \overline{uv}) \\ &\leq u \cdot u + v \cdot v + 2|uv| \leq \text{ aplicant la proposició anterior} \\ &\leq u \cdot u + v \cdot v + 2\sqrt{u \cdot u}\sqrt{v \cdot v} = (\sqrt{u \cdot u} + \sqrt{v \cdot v})^2, \end{aligned} \tag{2.2.5}$$

d'on $\sqrt{(u+v)(u+v)} \leq \sqrt{u \cdot u} + \sqrt{v \cdot v}$. ■

Observació 2.2.10 (Producte escalar estàndard)**.**

- Siguin $u = (x_1, \ldots, x_n)$ i $v = (y_1, \ldots, y_n)$ dos vectors de l'espai $\mathbb{R}^n$. El *producte escalar estàndard* es defineix com
$$uv := x_1 y_1 + \ldots + x_n y_n. \tag{2.2.6}$$
Per a tot vector $u = (z_1, \ldots, z_n) \in \mathbb{R}^n$ la *norma de u* és
$$||u|| := \sqrt{u \cdot u} = \sqrt{z_1^2 + \ldots + z_n^2}. \tag{2.2.7}$$
- Geomètricament es pot provar que
$$v \cdot \omega = ||v|| \cdot ||\omega|| \cdot \cos\theta, \tag{2.2.8}$$
on θ és l'angle que formen els vectors v i ω. Això explica la terminologia (v i ω són ortogonals si, i només si, v i ω són perpendiculars).
- Si (E, φ) és un espai vectorial Euclidià i $\mathfrak{B} = \{e_1, \ldots, e_n\}$ és una base ortonormal, aleshores $\mathcal{G}$ és la identitat i el càlcul en coordenades del producte φ té la mateixa expressió que el producte escalar estàndard de $\mathbb{R}^n$.

Exemple 2.2.11. L'espai vectorial $E = \mathbb{R}[x]_{\leq 2}$ amb el producte escalar

$$\varphi(p(x), q(x)) := \int_0^1 p(x)q(x)dx. \tag{2.2.9}$$

2.3 Ortogonalització de Gram-Schmidt

Donat un espai vectorial Euclidià $(E, \cdot)$ ens proposem trobar una base de E que sigui ortonormal. Per fer-ho, ens servirem del mètode de Gram-Schmidt.

Definició 2.3.1 (Projecció)**.** Si $u, v \in E$ amb $u \neq 0$ definim la projecció de v en u com el vector

$$\mathrm{pr}_u(v) = \left(\frac{v \cdot u}{u \cdot u}\right) u. \tag{2.3.1}$$

per a $u = 0$ definim $\mathrm{pr}_u(v) = 0$ per a tot v. Un parell de consideracions:

1. L'expressió entre parèntesi és un escalar i per tant $\mathrm{pr}_u(v)$ és un múltiple del vector u.
2. Si representem u, v en uns eixos coordenats de $\mathbb{R}^n$, geomètricament $\mathrm{pr}_u(v)$ correspon a la projecció ortogonal de v en la direcció marcada per u.

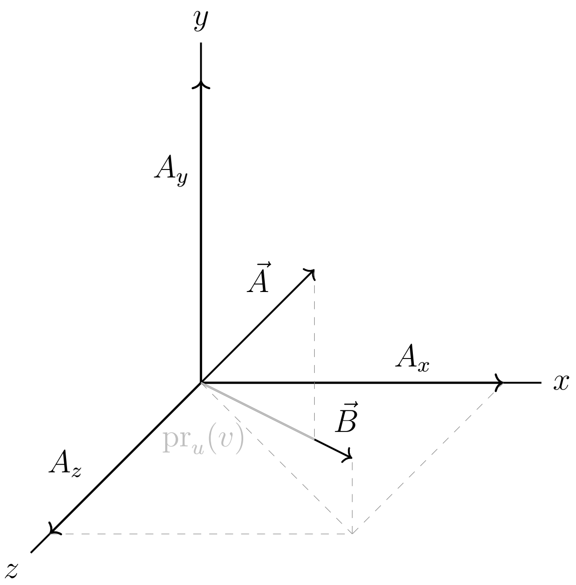

Figura 2.1: Representació del producte escalar $\mathrm{pr}_u(v)$.

3. Observem que $\mathrm{pr}_{\lambda u} = \mathrm{pr}_u(v)$ per a tot $\lambda \in \mathbb{R} \setminus \{0\}$, de manera que en efecte $\mathrm{pr}_u(v)$ depèn només de v i de la direcció de u.

Procés 2.3.2 (Procés d'ortogonalització de Gram-Schmidt)**.** *El procés d'ortogonalització de Gram-Schmidt consisteix en, donada una base $\{v_1, \dots, v_n\}$ d'E, trobar una base ortogonal $\{u_1, \dots, u_n\}$ d'E, de la següent manera:*

$$\begin{aligned} u_1 &= v_1, \\ u_2 &= v_2 - \mathrm{pr}_{u_1}(v_2), \\ u_3 &= v_3 - \mathrm{pr}_{u_1}(v_3) - \mathrm{pr}_{u_2}(v_3), \\ &\vdots \\ u_n &= v_n - \sum_{i=1}^{n-1} \mathrm{pr}_{u_i}(v_n). \end{aligned} \tag{2.3.2}$$

Proposició 2.3.3. *Els vectors $\{u_1, \dots, u_n\}$ obtinguts pel procés pel procés de Gram-Schmidt formen una base ortogonal d'E.*

Demostració. Observem que per tot $1 \leq i \leq n$, el vector u_i és de la forma

$$u_i = v_i + \sum_{1 \leq j < i} a_j v_j. \tag{2.3.3}$$

En particular, la matriu de coordenades de $\{u_1, \ldots, u_n\}$ en la base $\{v_1, \ldots, v_n\}$ té tot uns a la diagonal i zeros per sota de la diagonal, de la forma

$$\begin{pmatrix} 1 & * & \cdots & * \\ 0 & 1 & \cdots & * \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix}, \tag{2.3.4}$$

on $*$ indiquen escalars arbitraris. Com que la matriu té determinant 1, és invertible i els vectors formen una base. La ortogonalitat dos a dos dels vectors se segueix per construcció. ■

Definició 2.3.4 (Procés d'ortonormalització de Gram-Schmidt)**.** Si després d'aplicar 2.3.2 normalitzem els vectors u_i posant

$$e_i := \frac{u_i}{||u_i||} = \frac{u_i}{\sqrt{u_i \cdot u_i}} \tag{2.3.5}$$

parlem de *procés d'ortonormalització de Gram-Schmidt.*

Exemple 2.3.5. Vegem un exemple de com aplicar el procés d'ortogonalització de Gram-Schmidt en $\mathbb{R}^2$. Considerem els vectors $v_1 = (2, 1)$ i $v_2 = (1, 4)$. Aquests dos vectors són linealment independents i per tant formen una base de $\mathbb{R}^2$, però $v_1 \cdot v_2 = 6 \neq 0$ i per tant no són ortogonals. Prenem $u_1 := v_1 = (2, 1)$ i $u_2 := v_2 - \mathrm{pr}_{u_1}(v_2) = v_2 - \mathrm{pr}_{v_1}(v_2)$. La projecció de v_2 en la direcció de v_1 és

$$\mathrm{pr}_{v_1}(v_2) = \frac{v_2 \cdot v_1}{v_1 \cdot v_1} v_1 = \frac{6}{5}(2, 1). \tag{2.3.6}$$

Per tant, $u_2 = \frac{7}{5}(-1, 2)$. És immediat comprovar que $\{u_1, u_2\}$ és una base ortogonal. Si volem definir ara una base ortonormal, només cal calcular

$$\begin{aligned} ||u_1|| &= \sqrt{(2, 1) \cdot (2, 1)} = \sqrt{5} \\ ||u_2|| &= \sqrt{\frac{7}{5}(-1, 2) \cdot \frac{7}{5}(-1, 2)} = \frac{7}{\sqrt{5}}. \end{aligned} \tag{2.3.7}$$

Aleshores, si $e_1 = \frac{u_1}{||u_1||}$ i $e_2 = \frac{u_2}{||u_2||}$, la base $\{e_1, e_2\}$ és ortonormal.

Observació 2.3.6. El mètode de Gram-Schmidt funciona sobre qualsevol cos de característica $\neq 2$. Recordem que la característica d'un cos és el menor nombre de vegades que s'ha de sumar la identitat multiplicativa per a obtenir la identitat additiva.

Corol·lari 2.3.7. *Sigui $\varphi : E \times E \longrightarrow \mathbb{R}$ una forma bilineal.*

1. *φ és un producte escalar si, i només si, hi ha una base en què la matriu de Gram de φ és la identitat.*
2. *Si φ és un producte escalar aleshores la matriu de Gram $\mathcal{G}(\varphi)$ de φ en qualsevol base satisfà* $\det(\mathcal{G}(\varphi)) > 0$.

Demostració. Si φ és un producte escalar, aleshores el mètode de Gram-Schmidt ens proporciona una base en què la matriu de Gram és la identitat. Recíprocament, si $\mathcal{G}_{\mathfrak{B}}(\varphi)$ és la identitat en certa base $\mathfrak{B}$, aleshores és clar que la forma bilineal φ és simètrica (per ser-ho la matriu de Gram). A més, per a tot parell $u, v \in E$, tenim

$$\varphi(u, v) = u^t \cdot v = \sum_i a_i b_i, \tag{2.3.8}$$

on $(a_1, \dots, a_n)$ i $(b_1, \dots, b_n)$ són les coordenades de u, v respectivament. En particular,

$$\varphi(u,u) = \sum_i a_i^2 \geq 0 \tag{2.3.9}$$

i $\varphi(u,u) = 0$ si, i només si, $a_i = 0$ per tot i, la qual cosa és equivalent a demanar que $u = 0$. Per tant, φ és definida positiva.

Per provar el segon punt. Sigui $\mathfrak{B}$ una base qualsevol i sigui $\mathfrak{B}'$ una base ortonormal. Aleshores, per la fórmula del canvi de base, existeix una matriu $\mathcal{P}$ amb $\det(\mathcal{P}) \neq 0$ tal que $\mathcal{G}_{\mathfrak{B}}(\varphi) = \mathcal{P}^t \cdot I \cdot \mathcal{P}$ (ja que la matriu de Gram en la base ortonormal és la identitat). Prenent determinants obtenim

$$\det(\mathcal{G}_{\mathfrak{B}}(\varphi)) = \det(\mathcal{P})^2 > 0. \tag{2.3.10}$$

■

Exemple 2.3.8. Considerem la forma bilineal $\varphi : \mathbb{R}^3 \times R^3 \longrightarrow \mathbb{R}$ que en certa base $B = \{e_1, e_2, e_3\}$ té matriu de Gram

$$\mathcal{G}_{\mathfrak{B}}(\varphi) = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix}. \tag{2.3.11}$$

Aquesta matriu té determinant positiu (fins i tot és simètrica), però no representa un producte escalar ja que $\varphi(e_1, e_1) = 0$. Observem que els termes de la diagonal de la matriu de Gram d'un producte escalar $g_{ii} = \varphi(e_i, e_i)$ són sempre positius i això no es compleix en cap terme de la diagonal de la matriu de l'exemple. Ens podríem preguntar si les condicions $g_{ii} > 0$ són suficients, però resulta que no: per exemple, en $\mathbb{R}^4$ la forma bilineal simètrica φ que en la base canònica té per matriu de Gram

$$\mathcal{G}_{\mathfrak{B}}(\varphi) := \begin{pmatrix} 1 & 2 & 0 & 0 \\ 2 & 1 & 0 & 0 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 2 & 1 \end{pmatrix} \tag{2.3.12}$$

satisfà que el determinant és positiu i també que ho són tots els termes de la diagonal. En canvi, no representa un producte escalar: $\varphi((1,-1,0,0),(1,-1,0,0)) = -2 < 0$.

Definició 2.3.9 (Submatriu $\mathcal{G}_{\mathfrak{B}}(\varphi)_r$)**.** Sigui $\mathcal{G}_{\mathfrak{B}}(\varphi)$ la matriu de Gram d'una forma bilineal $\varphi : E \times E \longrightarrow \mathbb{R}$ en una base $\mathfrak{B} = \{v_1, \dots, v_n\}$ de E. Per a $r \geq 1$, denotarem per $\mathcal{G}_{\mathfrak{B}}(\varphi)_r$ la matriu obtinguda a partir de les primeres r files i r columnes de la matriu $\mathcal{G}_{\mathfrak{B}}(\varphi)$. En altres paraules:

$$\mathcal{G}_{\mathfrak{B}}(\varphi)_r := \begin{pmatrix} g_{11} & \cdots & g_{1r} \\ \vdots & \ddots & \vdots \\ g_{r1} & \dots & g_{rr} \end{pmatrix}. \tag{2.3.13}$$

La matriu $\mathcal{G}_{\mathfrak{B}}(\varphi)_r$ també es pot interpretar com la matriu de la restricció de φ al subespai generat per $\{v_1, \dots, v_r\}$.

Si φ és definida positiva també ho serà la seva restricció a un subespai; per tant, es complirà que $\det(\mathcal{G}_{\mathfrak{B}}(\varphi)_r) > 0$ per a tot $r \in \{1, \dots, n\}$.

Proposició 2.3.10. *Sigui $\varphi : E \times E \longrightarrow \mathbb{R}$ una forma bilineal simètrica en E. Són equivalents:*

1. *La forma bilineal φ és definida positiva.*
2. *Per a tota base de E, la matriu de Gram $\mathcal{G}$ en aquesta base satisfà $\det(\mathcal{G}_r) > 0,\ \forall 1 \leq r \leq n$.*
3. *Existeix una base tal que la matriu de Gram $\mathcal{G}$ en aquesta base satisfà $\det(\mathcal{G}_r) > 0,\ \forall 1 \leq r \leq n$.*

2.4 Subespais ortogonals i relació amb el dual

Fixem un espai euclidià $(E, \cdot)$. Donat un conjunt S de vectors d'E, en definirem el subespai ortogonal $S^{\perp}$. Les propietats de $S^{\perp}$ són anàlogues a les propietats dels anul·ladors que vam definir per a l'espai dual.

Definició 2.4.1 (Subespai ortogonal)**.** Donat un subconjunt S de E definim l'ortogonal de S com el conjunt de vectors ortogonals a tots els elements de S

$$S^{\perp} := \{x \in E \mid u \cdot x = 0, \forall u \in S\}. \tag{2.4.1}$$

Proposició 2.4.2. *Sigui S un subconjunt de E, on $(E, \cdot)$ és un espai vectorial Euclidià. Tenim que*

1. *$S^{\perp}$ és subespai vectorial d'E.*
2. *$S^{\perp} = \langle S \rangle^{\perp}$.*
3. *Si $T \subseteq S$, aleshores $S^{\perp} \subseteq T^{\perp}$.*

Proposició 2.4.3. *Sigui $(E, \cdot)$ un espai vectorial Euclidià. Si F és un subespai vectorial d'E, aleshores:*

1. *$E = F \oplus F^{\perp}$.*
2. *$\dim F^{\perp} = \dim E - \dim F$.*
3. *$(F^{\perp})^{\perp} = F$.*

Demostració. És immediat veure que $F \cap F^{\perp} = \{0\}$. Prenem una base ortonormal $\{v_1, \ldots, v_k\}$ de F i la completem a una base de E tal que $\{v_1, \ldots, v_k e_{k+1}, \ldots, e_n\}$. Apliquem ara el mètode de Gram-Schmidt per convertir aquesta base en una base ortonormal $\{v_1, \ldots, v_k, v_{k+1}, \ldots, v_n\}$. Observem que $u_j \in F^{\perp}$ quan $j \in \{k+1, \ldots, n\}$. Per tant, qualsevol $v \in E$ es pot escriure de forma única com

$$v = (a_1 u_1 + \cdots + a_k u_k) + (a_{k+1} u_{k+1} + \cdots + a_n u_n) \tag{2.4.2}$$

on el primer parèntesi pertany a F i el segon a $F^{\perp}$. Això prova que $E = F \oplus F^{\perp}$. La fórmula de les dimensions se segueix de la suma directa. És fàcil verificar que $F \subseteq (F^{\perp})^{\perp}$. Per dimensions obtenim la igualtat d'espais vectorials. ■

Proposició 2.4.4. *Si F, G són dos subespais d'un subespai euclidià $(E, \cdot)$, aleshores*

$$(F + G)^{\perp} = F^{\perp} \cap G^{\perp} \ i \ (F \cap G)^{\perp} = F^{\perp} + G^{\perp}. \tag{2.4.3}$$

Demostració. Per a la primera igualtat, provarem les dues inclusions. Si $x \in (F+G)^{\perp}$, aleshores $x(u + v) = 0$ per tot $u \in F$ i $v \in G$. En particular, si prenem $u = 0$ tenim que $xv = 0$ per tot $v \in G$, i si $v = 0$ obtenim que $xu = 0$ per a tot $u \in F$. Per tant, $x \in F^{\perp} \cap G^{\perp}$. Per l'altra inclusió, suposem que $x \in F^{\perp} \cap G^{\perp}$. Aleshores, $xu = 0$ per tot $u \in F$ i $xv = 0$ per a tot $v \in G$, la qual cosa implica que $x(u+v) = 0$ i, per tant, $x \in (F+G)^\perp$. Per provar la segona igualtat cal simplement aplicar la primera igualtat a $F^\perp$ i $G^\perp$:

$$(F^\perp + G^\perp)^\perp = (F^\perp)^\perp \cap (G^\perp)^\perp = F \cap G. \tag{2.4.4}$$

Aplicant l'ortogonal a ambdós costats de la igualtat obtenim la igualtat buscada. ■

Proposició 2.4.5. *Si $S = \{u_1, \ldots, u_s\}$ és un conjunt de vectors d'un espai euclidià, ortogonals dos a dos, aleshores aquest conjunt és linealment independent.*

Demostració. En efecte, si S és un conjunt de vectors diferents de $\vec{0}$ i ortogonals dos a dos, S és linealment independent. Si $\sum \lambda^i v_i = \vec{0}$ amb $v_i \in S$, per a cada v_k tenim

$$0 = \varphi(\sum \lambda^i v_i, v_k) = \sum \lambda^i \varphi(v_i, v_k) = \lambda^k \varphi(v_k, v_k). \tag{2.4.5}$$

Però $\varphi(v_k, v_k) \neq 0$, ja que $v_k \neq \vec{0}$. Per tant, $\lambda^k = 0, \forall k$, tal i com volíem. ■

Proposició 2.4.6. *Si (E, φ) és un espai vectorial Euclidià i $\{e_1, \ldots, e_n\}$ és una base ortonormal $\mathfrak{B}$, aleshores les coordenades d'un vector $u \in E$ són $(\varphi(u, e_1), \ldots, \varphi(u, e_n))$. En altres paraules,*

$$u = \sum_i a_i e_i, \ \textit{amb } a_i = \varphi(u, e_i). \tag{2.4.6}$$

Demostració. Si $\mathfrak{B} = \{e_1, \ldots, e_n\}$ és una base ortonormal, aleshores $u_1, \ldots, u_n$ són unitaris i ortogonals dos a dos. De fet, la matriu del producte escalar φ és la matriu identitat tal que ens queda:

$$\varphi(u_i, u_j) = \begin{cases} 0, & \text{si } i \neq j, \\ 1, & \text{si } i = j. \end{cases} \tag{2.4.7}$$

Sabent tot això, sigui $e_j \in \mathfrak{B}, j \in \{1, \ldots, n\}$ i $v = \sum_i^n b_i e_i$. Aleshores, multipliquem a banda i banda per e_j i ens queda que $v \cdot e_j = \varphi(v, e_j) = e_j \cdot \sum_i^n b_i e_i$. Tenint en compte, per (2.4.7), que $\varphi(e_i, e_j) = 1 \iff i = j$, ens queda que $\varphi(v, e_j) = b_j$, tal i com volíem demostrar. ■

2.5 Producte escalar i espai dual

Sigui $(E, \cdot)$ un espai vectorial Euclidià de dimensió finita n. Per a tot $v \in E$ definim una aplicació lineal

$$\begin{aligned} \omega_v : \ E &\longrightarrow \mathbb{R} \\ u &\longmapsto \omega_v(u) := uv. \end{aligned} \tag{2.5.1}$$

Això ens permet definir una aplicació

$$\begin{aligned} \gamma : \ E &\longrightarrow E^* \\ u &\longmapsto \omega_v \end{aligned} \tag{2.5.2}$$

Proposició 2.5.1. *L'aplicació γ és un isomorfisme.*

Lema 2.5.2. *Per a tot $F \subseteq E$ es té que $\gamma(F^\perp) = F^0$.*

Demostració. Si $u \in F^\perp$ aleshores $\gamma(u) = \omega_u$ és una forma que s'anul·la sobre F. Per tant, $\gamma(F^\perp) \subseteq F^0$. Com que tots dos espais tenen dimensió $n - \dim F$, en resulta la igualtat. ■

2.6 Espais vectorials hermítics

Molts dels resultats que hem vist per a espais vectorials euclidians (que en particular són espais definits sobre els reals) també són vàlids per a espais vectorials definits sobre els complexos amb un producte escalar hermític.

Definirem una forma sesquilineal en un espai vectorial complex copiant la definició de forma bilineal, però ara la propietat de producte per escalar portarà un *complex conjugat* associat en una de les dues variables (recordem que si $z = x + iy$ és un nombre complex, aleshores $\overline{z} = x - iy$ és el seu complex conjugat).

Considerarem E un espai vectorial aquesta vegada definit sobre el cos $\mathbb{C}$.

Definició 2.6.1 (Forma sesquilineal)**.** Una forma sesquilineal en E és una aplicació

$$\begin{array}{rccc} \varphi : & E \times E & \longrightarrow & \mathbb{C} \\ & (u, v) & \longmapsto & \varphi(u, v) := \lambda \end{array} \tag{2.6.1}$$

que assigna un nombre complex a cada parell de vectors i que satisfà

1. $\varphi(u_1 + u_2, v) = \varphi(u_1, v) + \varphi(u_2, v)$,
2. $\varphi(u, v_1 + v_2) = \varphi(u, v_1) + \varphi(u, v_2)$,
3. $\varphi(\lambda u, v) = \lambda\varphi(u, v)$,
4. $\varphi(u, \lambda v) = \overline{\lambda}\varphi(u, v)$,

on $u, v, u_1, u_2, v_1, v_2 \in E$ i $\lambda \in \mathbb{C}$.

Observació 2.6.2. L'únic canvi respecte la definició de forma bilineal és el complex conjugat $\overline{\lambda}$ que apareix a la quarta propietat.

Definició 2.6.3 (Matriu de Gram d'una forma sesquilineal)**.** La matriu de Gram $\mathcal{G}_{\mathfrak{B}}(\varphi) := \varphi(e_i, e_j)$ d'una forma sesquilineal φ aleshores es té $\varphi(u, v) = u^t\mathcal{G}_{\mathfrak{B}}(\varphi)\overline{v}$. Tenim que $\overline{v_j} = \alpha_j - i\beta_j$ i el vector $\overline{v}$ té coordenades $(\overline{v_1}, \cdots, \overline{v_n})$.

Definició 2.6.4 (Forma hermítica)**.** Direm que una forma sesquilineal és *hermítica* si $\varphi(u, v) = \overline{\varphi(v, u)}$ per a tot $u, v \in E$. De la mateixa manera que una forma bilineal, una forma sesquilineal és hermítica si, i només si, la seva matriu de Gram és *hermítica.*

Definició 2.6.5 (Matriu de Gram hermítica)**.** Una matriu de Gram $\mathcal{G}_{\mathfrak{B}}(\varphi)$ és hermítica si $\mathcal{G}_{\mathfrak{B}}(\varphi)^t = \overline{\mathcal{G}_{\mathfrak{B}}(\varphi)}$, on $\overline{\mathcal{G}_{\mathfrak{B}}(\varphi)}$ és la matriu on hem fet el complex conjugat a tots els seus coeficients.

Definició 2.6.6 (Producte escalar hermític)**.** Un *producte escalar hermític en E* és una forma sesquilineal i hermítica φ que és definida positiva: $\varphi(u, u) \in \mathbb{R}_{>0}$ per a tot $u \in E$ i $\varphi(u, u) = 0 \iff u = 0$.

Definició 2.6.7 (Espai vectorial hermític)**.** Un espai vectorial hermític (o espai vectorial unitari) és un espai vectorial complex dotat d'un producte escalar hermític, és a dir, una forma sesquilineal hermítica definida positiva.

Definició 2.6.8. Diem que un endomorfisme $f : E \longrightarrow E$ és *autoadjunt* si

$$v \cdot f(u) = f(v) \cdot u, \ \forall u, v \in E. \tag{2.6.2}$$

Proposició 2.6.9. *Si A és la matriu de f en una base ortonormal de E, aleshores f és autoadjunt si, i només si:*

Eu: *A és simètrica, és a dir, $A = A^t$;*
He: *A és hermítica, és a dir, $A = \overline{A}^t$.*

En altres paraules, tota matriu simètrica o hermítica és la matriu d'una aplicació autoadjunta en una base ortonormal.

Teorema 2.6.10. *Sigui $(E, \cdot)$ un espai vectorial hermític i $f : E \longrightarrow E$ un endomorfisme autoadjunt. Aleshores, existeix una base de vectors propis ortonormal.*

Demostració. Ho demostrarem per inducció sobre la dimensió de E. Si $n = 1$ ja queda demostrat, hem resolt el cas inicial. Si $\dim E = n$, aleshores el polinomi característic de f:

$$p_f(x) = \det(f - Ix) \in \mathbb{C}[x] \tag{2.6.3}$$

sempre té, almenys, una arrel complexa λ. Sigui v un vector unitari de valor propi d'aquesta arrel λ, o sigui $f(v) = \lambda(v)$. Aleshores, el subespai

$$F := \langle v \rangle^{\perp} = \{u \in E \mid uv = 0\} \tag{2.6.4}$$

és invariant per f (és a dir, $f(F) \subseteq F$). En efecte, si $u \in F$ aleshores

$$f(u)v = uf(v) = u(\lambda v) = \overline{\lambda}(uv) = \overline{\lambda}0 = 0, \tag{2.6.5}$$

d'on veiem que $f(u) \in F$. La propietat $f(F) \subseteq F$ ens permet definir l'endomorfisme autoadjunt $f|_F$ de f restringit al subespai F. Apliquem la hipòtesi d'inducció a F (que té dimensió $n-1$), per a trobar una base de vectors propis $\{u_2, \ldots, u_n\}$ ortonormals de F (respecte l'aplicació $f|_F$). Aleshores, el conjunt de vectors $\{v, u_2, \ldots, u_n\}$ és una base ortonormal de vectors propis de f, ja que $E = F \oplus F^{\perp}$. ■

Corol·lari 2.6.11. *Tota matriu hermítica A diagonalitza: existeix una matriu invertible C tal que*

$$D = C^{-1}AC, \tag{2.6.6}$$

amb D una matriu diagonal. A més, si la base de vectors propis és ortonormal, aleshores la matriu C és hermítica ($C^{-1} = \overline{C}^t$). En tal cas, tenim que

$$D = \overline{C}^t AC. \tag{2.6.7}$$

Per assegurar-nos que el vector propi λ que hem agafat és un valor real, definim el següent lema.

Lema 2.6.12. *Si A és una matriu simètrica real, aleshores els seus valors propis són reals.*

Demostració. Si λ és un valor propi i v un vector propi de valor propi λ, aleshores:

$$||Av||^2 = Av \cdot Av = (Av)^t(Av) = v^tA^tAv = v^tA^2v = v^t\lambda^2v = \lambda^2 v \cdot v = \lambda^2||v||^2. \tag{2.6.8}$$

Es dedueix que $\lambda^2 = \frac{||Av||^2}{||v||^2} \in \mathbb{R}_{\geq 0}$ i, per tant, $\lambda \in \mathbb{R}$. ■

Teorema 2.6.13. *Sigui* $(E, \cdot)$ *un espai vectorial euclidià i* $f : E \longrightarrow E$ *un endomorfisme autoadjunt. Aleshores, existeix una base de vectors propis ortonormal.*

Corol·lari 2.6.14. *Tota matriu simètrica real* A *diagonalitza: hi ha una matriu invertible* C *tal que*

$$D = C^{-1}AC, \tag{2.6.9}$$

i D *és diagonal. A més, si la base de vectors propis que construïm és ortonormal, aleshores la matriu* C *és ortogonal:* $C^{-1} = C^t$. *Per tant, en aquest cas tenim que*

$$D = C^tAC. \tag{2.6.10}$$

Observació 2.6.15. Tota matriu real simètrica és, en particular una matriu hermítica on la part imaginària de cada coeficient és zero, i per tant fer complexe conjugat equival a no fer res.

Capítol 3

Formes canòniques dels endomorfismes

3.1

Preliminars

Definició 3.1.1 (Matrius conjugades). Dues matrius quadrades A i B són conjugades, denotat per $A \sim B$ si existeix una matriu invertible U tal que

$$B = U^{-1}AU. \tag{3.1.1}$$

En altres paraules, podem passar d'una matriu a l'altra mitjançant un canvi de base.

Definició 3.1.2 (Endomorfismes conjugats). Diem que dos endomorfismes f i g són conjugats si hi ha un isomorfisme $\varphi : E \longrightarrow E$ tal que

$$g = \varphi^{-1} \circ f \circ \varphi. \tag{3.1.2}$$

Intentarem donar una base d'E en la qual la matriu de f en aquesta base sigui el més senzilla possible. Per exemple, veurem que tota matriu de mida 2×2 amb coeficients complexos és conjugada a una, i només una, d'aquestes dues matrius.

3.2

Polinomi mínim

E és un espai vectorial sobre un cos K. Fixem també un endomorfisme $f : E \longrightarrow E$. Ens interessa principalment la situació en què E té dimensió finita n. A partir del nostre endomorfisme f i d'un polinomi arbitrari $P(t)$ definirem un nou endomorfisme $P(f)$. Per això, utilitzarem potències de f:

$$f^0 = Id,\ f^1 = f,\ f^2 = f \circ f, \dots, f^r = f \circ f^{r-1}. \tag{3.2.1}$$

Notem que si $n = \dim E$, aleshores $\dim \operatorname{End}(E) = n^2$ i per tant, les potències d'un endomorfisme f no poden ser totes linealment independents.

Definició 3.2.1 (Endomorfisme $P(f)$). Sigui $P(t) = a_0 + a_1 t + \cdots + a_m t^m \in \mathbb{K}[t]$ un polinomi amb coeficients en $\mathbb{K}$. Definim l'endomorfisme $P(f) : E \longrightarrow E$ com

$$P(f) := a_0 f^0 + a_1 f^1 + \cdots + a_m f^m. \tag{3.2.2}$$

Exemple 3.2.2. Considerem $P(t) = t - \lambda$. Aleshores, $P(f) = f - \lambda \mathbb{I}$. Si $Q(t) = (t-2)(t-1) = t^2 - 3t + 2$, aleshores $Q(f) = f^2 - 3f + 2\mathbb{I}$.

Proposició 3.2.3. *Si $P(t), Q(t) \in \mathbb{K}[t]$ són dos polinomis, aleshores:*

1. $(P+Q)(f) = P(f) + Q(f)$, $(PQ)(f) = P(f) \circ Q(f)$.
2. *Els endomorfismes $P(f)$ i $Q(f)$ commuten: $P(f) \circ Q(f) = Q(f) \circ P(f)$.*

Definició 3.2.4. L'aplicació Φ_f

$$\begin{array}{rccl} \Phi_f : & \mathbb{K}[t] & \longrightarrow & \mathrm{End}(E) \\ & P(t) = a_0 + a_1 t + \cdots + a_m t^m & \longmapsto & \Phi_f(P(t)) := P(f) = a_0 f^0 + a_1 f^1 + \cdots + a_m f^m \end{array} \tag{3.2.3}$$

En aquesta aplicació, com veiem, s'envia tot polinomi $P(t)$ a l'endomorfisme $P(f)$ definit anteriorment.

Proposició 3.2.5. *L'aplicació Φ_f és lineal i hi ha un polinomi $M(t) \in \mathbb{K}[t]$ tal que* $\ker(\Phi_f) = \{P(t) \in \mathbb{K}[t] \mid P(t) = Q(t)M(t), Q(t) \in \mathbb{K}[t]\}$. *En altres paraules, el nucli ve donat pel conjunt de múltiples d'$M(t)$. A més, si E té dimensió finita es té $M(t) \neq 0$.*

Demostració.

$\Longrightarrow$ La linealitat de Φ_f se segueix de la proposició anterior. Per tant, el seu nucli és un subespai vectorial. Si $\ker(\Phi_f) \neq \{0\}$, escollim un polinomi $M \neq 0$ de $\ker(\Phi_f)$ de manera que sigui de mínim grau entre tots els polinomis no nuls de $\ker(\Phi_f)$. Vegem que $\ker(\Phi_f) = (M)$, on (M) és l'ideal generat pel conjunt de múltiples d'M. En efecte, $(M) \subset \ker(\Phi_f)$ per la proposició anterior.

$\Longleftarrow$ Si $P(f) =)0$, sigui $P(t) = Q(t)M(t) + R(t)$ la divisió entera de P entre M. Aleshores, $R(f) = P(f) - Q(f) \circ M(f) = 0$ i, per tant, $R(t) \in \ker(\Phi_f)$ i té grau menor que $M(t)$, d'on deduïm que $R = 0$ i $P(t) \in (M)$.

Si E té dimensió finita, aleshores $\mathrm{End}(E)$ també és de dimensió finita, i com que l'espai vectorial $\mathbb{K}[t]$ no és de dimensió finita, l'aplicació Φ_f no pot ser injectiva, la qual cosa implica que $\ker(\Phi_f) \neq \{0\}$. ■

Observació 3.2.6 (Ideal). Com ja esmentàvem abans, la notació $(M(t))$ on $M(t)$ és un polinomi indica l'ideal generat per aquest polinomi; és a dir, tots els polinomis de la forma $Q(t)M(t)$, amb $Q(t)$ polinomis qualssevol. Observem que si $R(t)$ és un polinomi qualsevol i $P(t) \in (M(t))$ aleshores $R(t)P(t) \in (M(t))$.

Definició 3.2.7 (Polinomi anul·lador). Els elements de $\ker(\Phi_f)$ s'anomenen polinomis anul·ladors d'f.

Definició 3.2.8 (Polinomi mínim). Si E té dimensió finita, el polinomi mínim de f és l'únic polinomi anul·lador $\mu_f(t)$ que satisfà $\ker(\Phi_f) = (\mu_f(t))$ i tal que el coeficient del terme de major grau és 1.

Exemple 3.2.9. Suposem que $E = \{0\}$ i $f = 0$ és l'únic endomorfisme d'E. Aleshores, tot polinomi $P(t)$ va a parar a l'endomorfisme trivial i per tant, $\ker(\Phi_f) = \mathbb{K}[t]$. D'aquí, deduïm que el polinomi mínim és $\mu_f(t) = 1$.

El polinomi mínim està estretament relacionat amb el polinomi característic.

Proposició 3.2.10. *Si E té dimensió finita, tota arrel del polinomi característic d'f és una arrel del polinomi mínim.*

Demostració. Sigui $\mu_f(t) = a_0 + a_1 t + \cdots + a_m t^m$. Si λ és una arrel de $p_f(t)$, aleshores hi ha $0 \neq v \in E$ tal que $f(v) = \lambda v$. Per tant, tenim que $f^i(v) = \lambda^i v$ per a tot i. Deduïm, doncs, que

$$0 = \mu_f(t)v = a_0 v + a_1 f(v) + \cdots + a_m f^m(v) = a_0 v + a_1 \lambda v + \cdots + \lambda^m v = \mu_f(\lambda)v. \tag{3.2.4}$$

Com que $v \neq 0$, resulta $\mu_f(\lambda) = 0$ i, per tant, λ és arrel de $\mu_f(t)$. ■

3.3 Subespais invariants

Definició 3.3.1 (Invariant per l'acció d'A)**.** Si A és una matriu de mida $n \times n$ que descompon en blocs de la forma

$$A = \begin{pmatrix} A_1 & 0 \\ 0 & A_2 \end{pmatrix}, \tag{3.3.1}$$

on A_1 és la matriu $n_1 \times n_1$ i A_2 és una matriu $n_2 \times n_2$ amb $n = n_1 + n_2$ els subespais E_1 i E_2 de $\mathbb{K}^n$, generats respectivament pels n_1 primers vectors de la base canònica i pels últims n_2 vectors de la base canònica de $\mathbb{K}^n$ són *invariants de l'acció d'A*. És a dir, es té $Av \in E_i, \forall v \in E_i$.

Definició 3.3.2 (Subespai f-invariant)**.** Diem que un subespai F d'E és invariant per f, o f-invariant, si $f(F) \subset F$.

Definició 3.3.3 (Restricció de f en F)**.** En un subespai f-invariant, f indueix un endomorfisme d'F que denotarem per $f|_F$. Es té: $f|_F(v) := f(v), \forall v \in F$. En particular, $f|_F$ és un endomorfisme tal que

$$\begin{array}{rccc} f' = f|_F : & F & \longrightarrow & F \\ & v & \longmapsto & f(v) \end{array} \tag{3.3.2}$$

i anomenarem *restricció de f en F*.

Proposició 3.3.4. *Sigui $F \subseteq E$ un subespai f-invariant i sigui $f' := f|_F$ la restricció de f a F. Aleshores, el polinomi mínim $\mu_{f'}(t)$ de f' divideix el polinomi mínim $\mu_f(t)$ de f.*

Demostració. Recordem que $\mu_f(t)$ és un element del nucli de l'aplicació Φ_f que envia un polinomi $P(t)$ a l'endomorfisme $P(f)$ d'E. En particular, sabem que $\mu_f(f) = 0$. Això implica que $\mu_f(f)(u) = 0, \forall u \in E$ i, per tant, es té que $\mu_f(f')(v) = \mu_f(f)(v) = 0, \forall v \in F$. Per tant, el polinomi $\mu_f(x)$ és un element del nucli de l'aplicació que envia un polinomi $P(t)$ a l'endomorfisme $P(f')$ de F. Per tant, $\mu_f(t)$ és un múltiple de $\mu_{f'}(t)$. ■

Corol·lari 3.3.5. *Siguin $F, G \subseteq E$ dos subespais f-invariants. Si les restriccions a F i a G tenen polinomis mínims primers entre ells, aleshores $F \cap G = \{0\}$.*

Demostració. Si F, G són f-invariants, aleshores també ho és $F \cap G$ i, per la proposició anterior, 3.3.4, el seu polinomi mínim és 1 (ja que per la proposició, sabem que el polinomi mínim de la restricció a $F \cap G$ divideix els polinomis mínims de les restriccions a F i G, respectivament), però per hipòtesi aquests són primers entre ells. Anem a veure que H és un espai vectorial i g és un endomorfisme tal que $\mu_g(t) = 1$, aleshores $H = \{0\}$. En efecte, en aquest cas la proposició

3.2.5 ens diu que $\ker(\Phi_g) = (1) = \mathbb{K}[t]$ i, per tant, Φ_g és l'aplicació lineal $\Phi_g = 0$. En paritcular, $0 = \Phi_g(1) = \mathbb{I}_H$ i, per tant, $H = \{0\}$. Aplicant aquest resultat a $H = F \cap G$ obtenim el corol·lari. ■

Proposició 3.3.6. *Si $P(t)$ és un polinomi, aleshores els subespais $\ker(P(f))$ i $\operatorname{im}(P(f))$ d'E són subespais f-invariants.*

Demostració. Això es pot reescriure com: *si $f : V \longrightarrow V$ és lineal, aleshores, per cada polinomi $P[x]$ els subespais $\operatorname{im} P(f)$ i $\ker P(f)$ són f-invariants.* Per a aquesta demostració usarem una observació, la següent:

Observació 3.3.7. Per a cada polinomi P tenim que $f \circ P(f) = P(f) \circ f$.

Suposem $\lambda \in \ker P(f)$. Considerem l'expressió $P(f)(f(\lambda))$. D'aquí és fàcil veure que $P(f)(f(\lambda)) = f(P(f)(\lambda))$, i $P(f)(\lambda) = 0$, concretament $f(\lambda) \in \ker P(f)$, per hipòtesi. Així doncs, $f(\vec{0}) = \vec{0}$. Hem deduït que $\ker P(f) \subset \ker f$. Anàlogament, suposem $\lambda = p(f)(\mu) \in \operatorname{im} P(f)$. Aleshores, apliquem f a ambdós costats de la igualtat i ens queda $f(\lambda) = f(p(f)(\mu)) = p(f)(f(\mu))$ per linealitat, d'on $f(\lambda) \in \operatorname{im} P(f)$. ■

Proposició 3.3.8. *Sigui $\mu_f(t) = P(t)Q(t)$ amb P, Q primers entre ells. Aleshores,*

$$E = \ker(P(f)) \oplus \ker(Q(f)). \tag{3.3.3}$$

A més, el polinomi mínim de la restricció d'f a $\ker(P(f))$ és $P(t)$, i el polinomi mínim de la restricció a f a $\ker(Q(f))$ és $Q(t)$. [CL09]

Demostració. Els polinomis $P(t)$ i $Q(t)$ són anul·ladors de $\ker(P(f))$ i $\ker(Q(f))$, respectivament. Per tant, per la proposició 3.2.5 i com que $P(t)$ i $Q(t)$ són primers entre ells, sabem que els polinomis mínims de les restriccions a $\ker(P(f))$ i $\ker(Q(f))$, respectivament, són primers entre ells. Per tant, podem aplicar el corol·lari 3.3.5 i afirmar que $\ker(P(f)) \cap \ker(Q(f)) = \{0\}$. D'altra banda, és fàcil veure que $\operatorname{im}(Q(f)) \subset \ker(P(f))$ i $\operatorname{im}(P(f)) \subset \ker(Q(f))$. Comprovem la primera inclusió: si $u = Q(f)(v) \in \operatorname{im}(Q(f))$ aleshores es té $P(f)(u) = P(f)Q(f)(v) = \mu_f(f)(v) = 0$ i, per tant, $u \in \ker(P(f))$. Suposant $n = \dim E$, aquestes inclusions ens indiquen que

$$\begin{aligned} n = \dim\ker(P(f)) + \dim\operatorname{im}(P(f)) &\leq \dim\ker(P(f)) + \dim\ker(Q(f)) \\ &= \dim(\ker(P(f)) \oplus \ker(Q(f))) \leq n. \end{aligned} \tag{3.3.4}$$

Les dues inclusions són, doncs, igualtats i ens queda $E = \ker(P(f)) \oplus \ker(Q(f))$. Sigui ara f' la restricció d'f a $\ker(P(f))$ i f'' la restricció a $\ker(Q(f))$. Com que $P(t)$ i $Q(t)$ són anul·ladors de f' i f'', respectivament, sabem que $\mu_{f'}(t)$ divideix $P(t)$ i que $\mu_{f''}(t)$ divideix $q(t)$. D'altra banda, el polinomi $\mu_{f'}(t) \cdot \mu_{f''}(t)$ és un polinomi anul·lador de f. En efecte, usant la descomposició $E = \ker(P(f)) \oplus \ker(Q(f))$ podem escriure tot vector d'E com $u = u_1 + u_2$ de forma única, on $u_1 \in \ker(P(f))$ i $u_2 \in \ker(Q(f))$ i obtenim que

$$\mu_{f'}(f)\mu_{f''}(f)(u) = \mu_{f''}(f)\mu_{f'}(f)(u_1) + \mu_{f'}(f)\mu_{f''}(f)(u_2) = 0 + 0 = 0. \tag{3.3.5}$$

Per tant, sabem que $\mu_{f'}(t)\mu_{f''}(t)$ és múltiple de $P(t)Q(t)$. Juntament amb el fet que $\mu_{f'}(t) \mid P(t)$ i $\mu_{f'}(t) \mid Q(t)$, deduïm que $\mu_f(t)\mu_{f''}(t) = \mu_f(t)$. Així doncs, ens queda que $\mu_{f'}(t) = P(t)$ i $\mu_{f''}(t) = Q(t)$. ■

Naturalment, si ara $P(t)$ o $Q(t)$ descompon en factors primers, podem descompondre $\ker(P(f))$ o $\ker(Q(f))$ en suma de subespais invariants i procedir iterativament amb aquest mètode tantes vegades com sigui possible. Ho enunciem de la següent manera:

Teorema 3.3.9 (Primer teorema de descomposició). *Si el polinomi mínim de f és*

$$\mu_f(t) = \mu_1(t)^{n_1} \cdots \mu_r(t)^{n_r} \tag{3.3.6}$$

on $\mu_i(t)$ són factors irreductibles de $\mu_f(t)$ aleshores es té una descomposició en subespais f-invariants tal que

$$E = E_1 \oplus \cdots \oplus E_r, \tag{3.3.7}$$

on $E_i = \ker(\mu_i(f)^{n_i})$ i el polinomi mínim de la restricció de f a E_i és $\mu_i(t)^{n_i}$. A més, aquesta és la única descomposició d'E com a suma directa de subespais f-invariants que satisfà aquesta propietat.

Demostració. Per la proposició 3.3.8 obtenim una descomposició com la que queda enunciada en aquest teorema. Per tant, només cal provar la unicitat de la descomposició. Suposem que donada la següent descomposició en subespais invariants $E = E^1 \oplus \cdots \oplus E^r$ de forma que el polinomi mínim de la restricció de f a F_i té polinomi mínim $\mu_i(t)^{n_i}$. Aquesta darrera condició ens implica que $F_i \subseteq \ker(\mu_i(f)^{n_i})$. Mirant les dimensions, tenim que

$$n = \dim F_1 + \cdots + \dim F_r \leq \dim\ker(\mu_1(f)^{n_1}) + \cdots + \dim\ker(\mu_r(f)^{n_r}) = n. \tag{3.3.8}$$

Totes les desigualtats de (3.3.8) han de ser igualtats i, finalment, obtenim que $F_i = \ker(\mu_i(f)^{n_i})$, $i = 1, \ldots, r$. ■

Observació 3.3.10. Si escrivim la matriu A d'f en una base E formada per les bases de cadascun dels subespais E_i, aleshores obtenim una matriu diagonal a blocs, de la forma

$$A = \begin{pmatrix} A_1 & 0 & \cdots & 0 \\ 0 & A_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & A_r \end{pmatrix}, \tag{3.3.9}$$

on A_i és la matriu de $f|_{E_i}$ en la base d'E_i (aquí, les matrius A_i són matrius quadrades de mida $\dim E_i \times \dim E_i$ i els zeros indiquen matrius trivials de la mida que els correspongui). Per tant, veiem que l'estudi d'A es redueix a l'estudi individual de cada bloc A_i.

Exemple 3.3.11. Sigui $f : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ l'endomorfisme d'$\mathbb{R}^2$ en què la base canònica té per matriu

$$A = \begin{pmatrix} 2 & 1 \\ -3 & -2 \end{pmatrix}. \tag{3.3.10}$$

Calculant iterativament les potències d'A tenim que $A^2 = \mathbb{I}_2$. Així doncs, $A^3 = A$ i $A^4 = \mathbb{I}_2$, i així successivament. Per tant, podem deduir fàcilment que el polinomi $P(t) = t^2 - 1 = (t+1)(t-1)$ és un polinomi anul·lador d'f. Suposem diversos casos:

1. $\mu_f(t) = (t-1)$: aleshores, $\mu_f(f) = f - \mathbb{I}_2 = 0$, d'on deduïm que $f = \mathbb{I}_2$.
2. $\mu_f(t) = (t+1)$: aleshores, $f = -\mathbb{I}_2$.

Com que, clarament, aquestes conclusions no són certes, cal que $P(t) \mid \mu_f(t)$. Per tant, tenim que $\mu_f(t) = P(t) = (t-1)(t+1)$. Usant 3.3.9 (1TD), podem escriure $E = E_1 \oplus E_2$, on $E_1 = \ker(f - \mathbb{I}) = \langle(1,-1)\rangle$ i $E_2 = \ker(f + \mathbb{I}) = \langle(1,-3)\rangle$. La restricció de f a E_1 és $\mathbb{I}$, mentre que la restricció de f a E_2 és $-\mathbb{I}$. Per tant, la matriu de f en la base $\{(1,-1),(1,-3)\}$ és

$$\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{3.3.11}$$

Veiem que, en aquest cas, E_1 i E_2 no són més que els subespais propis 1 i -1, respectivament.

Definició 3.3.12 (Involució)**.** Sigui un endomorfisme $f : E \longrightarrow E$ d'un espai vectorial E. f és una *involució* si $f^2 = \mathbb{I}$.

Exemple 3.3.13. Com a generalització de l'exemple anterior, considerem una involució f d'un espai E, tal que $\dim E = n$. Aleshores, com en l'exemple anterior, $\mu_f(t) \mid (t^2-1) = (t+1)(t-1)$ és un polinomi anul·lador d'f i dividirem en els següents casos:

1. $\mu_f(t) = (t-1)$: aleshores, $f = \mathbb{I}$.
2. $\mu_f(t) = (t+1)$: aleshores, $f = -\mathbb{I}$.
3. $\mu_f(t) = (t+1)(t-1)$: aleshores, $E = E_1 \oplus E_2$, on el polinomi mínim de la restricció $f|_{E_1}$ és $(t-1)$ i el polinomi mínim de la restricció $f|_{E_2}$ és $(t+1)$. Per tant, la restricció $f|_{E_1}$ és $\mathbb{I}$ i la restricció $f|_{E_2}$ és $-\mathbb{I}$. Prenent la unió de les bases de E_1 i E_2, la matriu de f en aquesta base és

$$\begin{pmatrix} \mathbb{I}_{k_1} & 0 \\ 0 & -\mathbb{I}_{k_2} \end{pmatrix}, \tag{3.3.12}$$

on k_1, k_2 són les dimensions d'E_1 i E_2, respectivament.

$$\begin{pmatrix} 1 & & & & & \\ & \ddots & & & 0 & \\ & & 1 & & & \\ & & & -1 & & \\ & 0 & & & \ddots & \\ & & & & & -1 \end{pmatrix} \begin{array}{l} \left.\vphantom{\begin{matrix}1\\1\\1\end{matrix}}\right\} r \text{ files} \\ \left.\vphantom{\begin{matrix}1\\1\\1\end{matrix}}\right\} n-r \text{ files.} \end{array}$$

Figura 3.1: Descomposant $\mathbb{I}_{k_2}$ i $\mathbb{I}_{k_1}$ ens quedaria una cosa tal que així. Font: [CL09]

Definició 3.3.14 (Estructura complexa en E)**.** Sigui un espai vectorial E complex. Un endomorfisme f d'E és una *estructura complexa en* E si f compleix que $f \circ f = -\mathbb{I}$ i, a més, hi ha una base tal que la matriu d'f en aquesta base és diagonal amb $\pm i$ a la diagonal.

Proposició 3.3.15. *Si k és un valor propi de f, aleshores $(t-k) \mid \mu_f(t)$.*

Teorema 3.3.16 (Teorema de diagonalització)**.** *Un endomorfisme és diagonalitzable si, i només si, el seu polinomi mínim descompon en factors lineals no repetits.*

Demostració.

$\Longrightarrow$ Provem primer que si $\mu_f(t) = (t-a_1)\cdots(t-a_r)$ amb $a_i \neq a_j, \forall i \neq j$, aleshores f és diagonalitzable. Recordem primer que per la proposició 3.2.10, tot valor propi és arrel de $\mu_f(t)$. Per 3.3.9 (PTD), podem escriure $E = E_1 \oplus \cdots \oplus E_r$, on $E_i = \ker(f - a_i\mathbb{I})$. Observem que E_i no és més que el subespai propi de valor propi a_i (a més, si a_i no és valor propi, aleshores per definició de valor propi $E_i = \{0\}$). Per tant, sabem que existeix una base de vectors propis d'E formada per la unió de les bases d'E_i.

$\Longleftarrow$ Ara ens toca suposar que f és diagonalitzable. Així, siguin $\lambda_1, \ldots, \lambda_s$ els valors propis d'f. Sabem que E és suma directa d'espais propis tal que $E = E_1 \oplus \cdots \oplus E_s$, on E_i és el subespai propi de valor propi λ_i. El polinomi mínim d'f restringida a E_i és justament $(t - \lambda_i)$. Això implica que el polinomi $P(t) := (t - \lambda_1) \cdots (t - \lambda_s) \mid \mu_f(t)$. Per veure que $\mu_f(t) = P(t)$ solament cal demostrar que $P(t)$ és anul·lador de f. Donat $u \in E$, prenem la descomposició formada per la suma directa $u = u_1 + \cdots + u_s, u_s \in E_i$ i prenem $(f - \lambda_1 \mathbb{I}) \cdots (f - \lambda_s \mathbb{I})(u)$.

$$\begin{aligned} &(f - \lambda_1 \mathbb{I}) \cdots (f - \lambda_s \mathbb{I})(u) \\ &= (f - \lambda_2 \mathbb{I}) \cdots (f - \lambda_s \mathbb{I})(f - \lambda_1 \mathbb{I})(u_1) \\ &+ (f - \lambda_1 \mathbb{I})(f - \lambda_3 \mathbb{I}) \cdots (f - \lambda_s \mathbb{I})(f - \lambda_2 \mathbb{I})(u_2) + \cdots + \\ &+ (f - \lambda_1 \mathbb{I}) \cdots (f - \lambda_s \mathbb{I})(u_s) = \vec{0} + \cdots + \vec{0} = \vec{0}. \end{aligned} \tag{3.3.13}$$

Com a nota final, es pot donar un cop d'ull a la demostració oferta a [CL09]. ■

3.4 Grau del polinomi mínim

Si E té dimensió n, aleshores el grau del polinomi mínim sempre és $\leq \dim \operatorname{End}(E) = n^2$. Definim un nou polinomi mínim associat a un vector d'E com se segueix:

Definició 3.4.1 (ϕ_f^u). Sigui $u \in E$ un vector arbitrari d'E. Definim un endomorfisme $P(f)$ el qual avaluarem en u i obtindrem una nova aplicació:

$$\begin{aligned} \phi_f^u : \quad \mathbb{K}[t] &\longrightarrow E \\ P(t) &\longmapsto P(f)(u). \end{aligned} \tag{3.4.1}$$

Definició 3.4.2 (Polinomi mínim de f en u). Resulta que el nucli d'aquesta aplicació també és un ideal (recordar 3.2.6) i, per tant, tot element de $\ker(\phi_f)$ és múltiple d'un cert polinomi, que denotem $\mu_f^u(t)$ amb coeficient 1 en el terme de major grau. Direm que aquest és el *polinomi mínim de f en u.*

Lema 3.4.3. *Si $\mu_f^u(t) = a_0 + a_1 t + \cdots + a_s t^s$, aleshores els vectors d'E $u, f(u), \ldots, f^{s-1}(u)$ són linealment independents.*

Demostració. Suposem que tenim una combinació lineal $b_0 u + b_1 f(u) + \cdots b_{s-1} f^{s-1}(u) = 0$, on no tots els b_i són zero. Definim el polinomi $Q(t) := b_0 + b_1 t + \cdots + b_{s-1} t^{s-1}$. Aleshores, $q(f)(u) = 0$ i, per tant, $q(t)$ és el nucli de ϕ_f^u. Però això és una contradicció, ja que $q(t)$ compleix que $\operatorname{gr}(q(t)) < s$, mentre que el polinomi mínim de f en u té grau s. ■

Proposició 3.4.4. *Si la dimensió d'E és n, aleshores el grau del polinomi mínim de qualsevol endomorfisme E és $\leq n$.*

Demostració. Per 3.3.9 (PTD) podem escriure $E = E_1 \oplus \cdots \oplus E_s$, on $\mu_f(t) = \mu_1(t)^{n_1} \cdots \mu_s(t)^{n_s}$ i $E_i = \ker(\mu_i(f)^{n_i})$. Com que la dimensió d'$E$ és la suma de dimensions dels E_i i el grau de $\mu_f(t)$ és la suma dels graus de $\mu_i(t)$, és suficient provar-ho per la restricció a cada subespai invariant. O sigui, suposem que $\mu_f(t) = p(t)^k$, on $p(t)$ és un polinomi irreductible i sigui r el grau de $\mu_f(t)$.

Per la definició de polinomi mínim, hi ha com a mínim un vector $u \in E$ tal que $p(f)^k(u) = 0$, però $p(f)^{k-1}(u) \neq 0$. Llavors, el polinomi mínim de f en u és $p(t)^k$ i, pel lema anterior, sabem que el conjunt $u, f(u), \dots, f^{r-1}(u)$ és un conjunt de vectors linealment independents, on recordem que r és el grau de $\mu_f(t)$. Per tant, com que són r vectors d'E linealment independents, queda provat que $r \leq n$. ■

3.5 Cayley-Hamilton

Anteriorment hem vist que els zeros del polinomi característic són també zeros del polinomi mínim. Demostrarem el recíproc i el teorema de Cayley-Hamilton.

Proposició 3.5.1. *a és un zero de $\mu_f(t)$ si, i només si, és un zero de $p_f(t)$.*

Demostració. Si a és un zero de $\mu_f(t)$, $\mu_f(t) = (t-\lambda)\mu_1(t)$. Existeix un $u \in E$ tal que $\mu_1(f)(u) \neq \vec{0}$, ja que, en cas contrari, el polinomi seria $\mu_1(t)$. Aleshores, el vector $\omega = \mu_1(f)(u)$ té valor propi λ:

$$(f - \lambda \mathbb{I})\omega = (f - \lambda \mathbb{I})\mu_1(f)(u) = \mu_f(f)(u) = \vec{0}. \tag{3.5.1}$$

Per tant, λ és un zero de $p_f(t)$. ■

Teorema 3.5.2. *El polinomi mínim sempre divideix el polinomi característic.*

Demostració. Provarem primer que si el polinomi mínim i el polinomi característic descomponen en factors lineals, aleshores el polinomi mínim divideix el polinomi característic. Suposem que $\mu_f(t) = (t-\lambda_1)^{k_1} \cdots (t-\lambda_s)^{k_s}$ i sigui $E = E_1 \oplus \cdots \oplus E_s$ la descomposició primària d'E respecte f. Aleshores sabem que la matriu de f en la base formada ajuntant les bases d'$E_1, \dots, E_s$ és de la forma

$$A = A_1 \oplus \cdots \oplus A_s = \begin{pmatrix} A_1 & 0 & \cdots & 0 \\ 0 & A_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & A_s \end{pmatrix}, \tag{3.5.2}$$

on A_i són matrius de mida $n_i \times n_i$, on $n_i = \dim E_i$. Observem, doncs, que el polinomi característic d'f es pot escriure com $p_f(t) = p_1(t) \cdots p_s(t)$, on $p_i(t)$ és el polinomi característic de la matriu A_i. Sabem per hipòtesi que cada $p_i(t)$ descompon en factors lineals, o sigui que és de la forma $p_i(t) = (t-\beta_1)^{r_1} \cdots (t-\beta_h)^{r_h}$. Però recordem que la proposició 3.2.10 ens diu que tota arrel $p_i(t)$ és arrel de $(t-\lambda_i)^{k_i}$. Això implica que $p_i(t) = (t-\lambda_i)^{n_i}$ on recordem que $n_i = \dim E_i$. O sigui que tenim

$$\mu_i(t) = (t-\lambda_i)^{k_i} \text{ i } p_i(t) = (t-\lambda_i)^{n_i}. \tag{3.5.3}$$

Per la proposició 3.4.4 sabem que $k_i \leq n_i$, la qual cosa implica que $\mu_f(t) \mid p_f(t)$. Ara, hem de generalitzar el teorema i veure que es compleix en tot cos $\mathbb{K}$:

1. Suposem $\mathbb{K} = \mathbb{C}$. Recordem que sobre el cos $\mathbb{C}$ tot polinomi divideix en factors lineals i, per tant, ja hauríem acabat.
2. Suposem $\mathbb{K} = \mathbb{R}$. En el cas d'un $\mathbb{R}$-espai vectorial, pot ser que els nostres polinomis tinguin arrels complexes i, per tant, no hi tinguem una descomposició en factors lineals. Si més no, recordem que si tenim una matriu A amb coeficients reals, en particular és una matriu

complexa i el seu polinomi característic és el mateix mirat com a matriu real que com a matriu complexa. També podem parlar del polinomi mínim de la matriu A tal com ho fèiem per a endomorfismes. Com que sobre $\mathbb{C}$ tenim que el polinomi característic és un polinomi anul·lador, evidentment això també és cert sobre $\mathbb{R}$. I per tant ja hem acabat.

■

Observació 3.5.3. El raonament sobre els cosses $\mathbb{R}$ i $\mathbb{C}$ que hem fet en la demostració del teorema anterior, 3.5.2, val per a endomorfismes definits sobre un cos arbitrari $\mathbb{K}$. En aquest cas, hem de considerar la *clàusula algebraica de* $\mathbb{K}$: si un polinomi en $\mathbb{K}[t]$ no té totes les arrels en $\mathbb{K}$, aleshores sempre hi ha un cos $\mathbb{K}'$ tal que $\mathbb{K} \subseteq \mathbb{K}'$ on aquest polinomi hi descompon en factors lineals.

Observació 3.5.4. Fixem-nos de la importància d'haver demostrat que els polinomis característic i mínim tenen les mateixes arrels: si λ és una arrel, aleshores la multiplicitat de λ com a arrel de $\mu_f(t)$ és més petita o igual que la multiplicitat algebraica de λ, $\text{mult}_\mu(\lambda) \leq m_a(\lambda)$.

En el cas que els polinomis descomponen en factors lineals tenim que

$$\begin{aligned} \mu_f(t) &= (t-\lambda_1)^{k_1} \cdots (t-\lambda_s)^{k_s}, \\ p_f(t) &= (t-\lambda_1)^{r_1} \cdots (t-\lambda_s)^{r_s}, \end{aligned} \tag{3.5.4}$$

amb $k_i \leq r_i$ i $r_1 + \cdots + r_s = n$. Aquests resultats ens donen un mètode pràctic per a calcular el polinomi mínim d'un endomorfisme f.

Procés 3.5.5 (Procés per calcular polinomi mínim d'f)**.**

1. *Sigui A la matriu d'f en una base donada i calculem el polinomi característic d'f usant la matriu A.*
2. *Descomponem $p_A(t)$ en factors irreductibles i busquem el més petit dels seus divisors $Q(t)$ tals que $Q(f) = 0$.*
3. *Aquest serà $\mu_f(t)$, sempre que arreglem les constants de tal manera que el coeficient de major grau sigui 1.*

Exemple 3.5.6. Per exemple, si sabem que $p_A(t) = (t-1)^3(t-2)^2$. Pels exponents, sabem que aquest polinomi té grau 5. Calculem, successivament, les matrius següents:

$$\begin{gathered} (A-\mathbb{I}_5)(A-2\mathbb{I}_5), \\ (A-\mathbb{I}_5)^2(A-2\mathbb{I}_5), \\ (A-\mathbb{I}_5)(A-2\mathbb{I}_5)^2, \\ (A-\mathbb{I}_5)^3(A-2\mathbb{I}_5), \\ (A-\mathbb{I}_5)^2(A-\mathbb{I}_5)^2. \end{gathered} \tag{3.5.5}$$

I ens quedarem amb la primera que s'anul·li. Si cap d'aquestes matrius s'anul·la, aleshores el polinomi mínim coincideix amb el polinomi característic.

Exemple 3.5.7. Sigui $f : \mathbb{R}^3 \longrightarrow \mathbb{R}^3$ l'endomorfisme que en la base canònica té per matriu

$$A = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 4 \end{pmatrix}. \tag{3.5.6}$$

Calculeu el polinomi característic $p_A(t) := \det(A - t\mathbb{I}_3)$ i utilitzeu la seva factorització per a trobar el polinomi mínim de f.

Demostració. Calculem el polinomi característic $p_A(t) := \det(A - t\mathbb{I}_3)$. Resolent aquest determinant d'una matriu 3×3, ens queda que el polinomi característic és $p_A(t) = (t-1)^2(4-t)$. Aleshores, donat que els valors propis són $t_1 = 1, t_2 = 4$.

$$(A + \mathbb{I})(A + 4\mathbb{I}) = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 5 \end{pmatrix} \begin{pmatrix} 5 & 1 & 0 \\ 0 & 5 & 1 \\ 0 & 0 & 8 \end{pmatrix} = \begin{pmatrix} 10 & 7 & 1 \\ 0 & 10 & 10 \\ 0 & 0 & 40 \end{pmatrix} \neq 0. \tag{3.5.7}$$

Com que no és zero ja sabem que $m_f(t) = p_f(t)$. ■

3.6
Subespais f-cíclics

El teorema de descomposició en subespais invariants permet reduir l'estudi d'un endomorfisme f a l'estudi de les seves restriccions a certs subespais invariants E^i. En aquest apartat, estudiarem les restriccions d'f als E^i en el cas general. Concretament, anem a descompondre cada subespai E^i en suma de subespais invariants sobre els quals l'actuació d'f és molt clara: els subespais f-cíclics. Ens centrarem en el cas en què el polinomi mínim descompon en factors lineals de grau 1, no necessàriament diferents (per exemple, això passa sempre si treballem sobre espais vectorials definits sobre el cos dels complexos $\mathbb{C}$).

En primer lloc, notem que la forma més senzilla de matriu d'un endomorfisme primari amb polinomi mínim és

$$\mu_f(t) = (t - \lambda)^s, \ s \leq n. \tag{3.6.1}$$

La recerca de subespais per als quals la matriu de la restricció d'f a aquests espais sigui el més senzilla possible ens porta a la definició d'espais f-cíclics.

Definició 3.6.1 (f-generador)**.** Sigui F un subespai invariant d'E. Un vector $u \in E$ és f-generador d'F si està generat per $\{u, f(u), f^2(u), \dots\}$.

Definició 3.6.2 (f-cíclic)**.** Sigui f un endomorfisme d'E i F un subespai invariant d'E. F és f-cíclic si existeix un vector $u \in E$ que sigui f-generador d'F.

Observació 3.6.3. Observem que si $u \in E$, aleshores el subespai vectorial E_u generat per $\{u, f(u), f^2(u), \dots\}$ és el subespai vectorial f-cíclic més petit d'E que conté a u.

Notació 3.6.4 (Polinomi mínim d'u)**.** Denotarem per μ_u el polinomi mínim d'$f|_{E_u}$ (la restricció d'f a E_u) i en direm el polinomi mínim d'u.

Proposició 3.6.5. *Suposem que $\mu_f(t) = (t - \lambda)^s$. Siguin $u \in E$ i $d \dim E_u$. Aleshores,*

1. $\mu_u(t) = (t - \lambda)^d, \ d \leq s$.
2. $\{u, f(u), \dots, f^{d-1}(u)\}$ *és una base d'E_u.*

Demostració. Com que E té dimensió n existeix un enter c mínim tal que $c > 0, c \leq n$ i $f^c(u)$ és combinació lineal d'$\{u, f(u), \dots, f^{c-1}(u)\}$. Aleshores, $E_u = \langle u, f(u), \dots, f^{c-1}(u)\rangle$ és un subespai invariant mínim que conté a u. Els vectors $\{u, f(u), \dots, f^{d-1}(u)\}$ són linealment independents, per definició de c. Per tant, $c = d = \dim E_u$. A més, si

$$f^d(u) = \sum_{0 \leq i < d} a_i f^i(u), \tag{3.6.2}$$

aleshores, el polinomi

$$a(t) = t^d - \sum_{0 \le i < s} a_i t^i \tag{3.6.3}$$

anul·la E_u. Recíprocament, si un polinomi de grau $r < d$ anul·lés E_u, aleshores $f^r(u)$ seria combinació lineal de $f^i(u)$, $0 \le i < r$, la qual cosa entraria amb contradicció amb la definició de c. Com que $\mu_f(f)(u) = 0$, resulta que $\mu_u \mid \mu_f$ i, per tant, $\mu_u(t) = (t-\lambda)^d, d \le s$. ■

Definició 3.6.6 (Base f-cíclica)**.** Si E és f-cíclic de dimensió d, una base que sigui de la forma $\{u, f(u), \dots, f^{d-1}(u)\}$ s'anomena *base f-cíclica*.

Proposició 3.6.7. *Siguin f un endomorfisme d'E amb polinomi mínim $\mu_f(t) = (t-\lambda)^s$. Siguin $u \in E$ i $g = f - \lambda$. Aleshores:*

1. *u és un generador d'E si, i només si, u és un g-generador d'E.*
2. *Si E és f-cíclic, la matriu de f en una base g-cíclica és*

$$\mathcal{J}(\lambda; s) = \begin{pmatrix} \lambda & & & & & \\ 1 & \lambda & & & & \\ & 1 & \lambda & & & \\ & & & \ddots & & \\ & & & & \lambda & \\ & & & & 1 & \lambda \end{pmatrix} \tag{3.6.4}$$

Demostració. La primera afirmació es dedueix de la relació entre les famílies de vectors $\{f^i(u)\}_i$ i $\{g^i(u)\}_i$. En efecte, del desenvolupament del binomi es dedueixen les següents relacions:

$$\begin{aligned} g^i(u) = (f-\lambda)^i(u) &= \sum_{0 \le j \le i} (-1)^{i-j} \lambda^{i-j} \binom{i}{j} f^j(u) = f^i(u) + \sum_{0 \le j < i} (-1)^{i-j} \lambda^{i-j} \binom{i}{j} f^j(u). \\ f^i(u) = (g+\lambda)^i(u) &= \sum_{0 \le j \le i} \lambda^{i-j} \binom{i}{j} g^j(u) = g^i(u) + \sum_{0 \le j < i} \lambda^{i-j} \binom{i}{j} g^j(u). \end{aligned} \tag{3.6.5}$$

Per tant, els subespais generats per $\{u, f(u), f^2(u), \dots\}$ i $\{u, g(u), g^2(u), \dots\}$ coincideixen. La segona afirmació se segueix de

$$f(g^i(u)) = (g+\lambda)(g^i(u)) = g^{i+1}(u) + \lambda g^i(u). \tag{3.6.6}$$

■

Observació 3.6.8. Si F és f-cíclic, és invariant per f, podem aplicar el teorema de la divisió entera per deduir que la seva dimensió és el grau del polinomi mínim de f a u. Si aquest polinomi és $a_0 + a_1 x + \cdots + a_{s-1} x^{s-1} + x^s$, aleshores $\{u, f(u), \dots, f^{s-1}(u)\}$ és una base de F i, en aquesta base, la matriu de la restricció d'f és

$$\begin{pmatrix} 0 & 0 & \cdots & 0 & -a_0 \\ 1 & 0 & \cdots & 0 & -a_1 \\ 0 & 1 & \cdots & 0 & -a_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & -a_{s-1} \end{pmatrix}. \tag{3.6.7}$$

Exemple 3.6.9. Per entendre bé la demostració anterior ajuda escriure'ns algun cas de dimensió baixa. Per exemple, $\dim E = 2$. Suposem, per tant, que $E = \langle u, f(u)\rangle$ és f-cíclic i amb polinomi mínim $\mu(t) = (f-\lambda)^2$. Sigui $g = f - \lambda$. Per escriure la matriu de f en la base $\{e_1, e_2\}$ on $e_1 = u$ i $e_2 = g(u)$, recordem que ens cal calcular les imatges d'aquests vectors per a f i escriure-les en termes d'aquests vectors. Tenim

$$\begin{gathered} f(e_1) = f(u) = g(u) + \lambda u = e_2 + \lambda e_1, \\ f(e_2) = f(g(u)) = (g+\lambda)(g(u)) = g^2(u) + \lambda g(u) = \lambda g(u) = \lambda e_2, \end{gathered} \tag{3.6.8}$$

on hem usat que $g^2(u) = 0$ si ens fixem en $\mu(t)$. Per tant, la matriu de f en aquesta base és

$$(f(e_1), f(e_2)) = \begin{pmatrix} \lambda & 0 \\ 1 & \lambda \end{pmatrix}. \tag{3.6.9}$$

Teorema 3.6.10 (Segon teorema de descomposició)**.** *Si $f \in \operatorname{End}(E)$, E és suma directa de subespais f-cíclics.*

Demostració. La demostració d'aquest teorema és força complicada i se suporta d'un lema auxiliar llarg de demostrar, també. Ho deixarem com a exercici, del qual es pot consultar la solució a [CL09, pàg. 154]. ■

Capítol 4

Jordan

4.1 Matrius de Jordan

Com ja sabem, hi ha endomorfismes que no són diagonalitzables, és a dir, la matriu que el determina no pot ser diagonal. Tot i això, les matrius de Jordan ens permetran trobar una que sigui quasi diagonal.

Suposem que el polinomi mínim de f descompon en factors lineals. Sigui $E = F_0 \oplus \cdots \oplus F_r$ la descomposició en subespais f-cíclics i $(t - \lambda_i)^{s_i}$ el polinomi mínim de la restricció d'f a F_i. Triem per a cada F_i un vector u_i amb polinomi mínim $(x - \lambda_i)^{s_i}$. Aleshores, $u_i, (f - \lambda_i I)(u_i), \ldots, (f - \lambda_{i-1} I)(u_i)$ són linealment independents.

A més, $\dim F_i = s_i$ i, per tant, aquests vectors formen una base. Escollint, en particular, $\lambda = \lambda_i$ i $s = s_i$, podem definir el següent.

Definició 4.1.1 (Matriu de Jordan elemental de valor propi λ i alçada s). La matriu de mida $s \times s$ donada per

$$\mathcal{J}(\lambda; s) = \begin{pmatrix} \lambda & & & & & \\ 1 & \lambda & & & & \\ & 1 & \lambda & & & \\ & & & \ddots & & \\ & & & & \lambda & \\ & & & & 1 & \lambda \end{pmatrix}, \tag{4.1.1}$$

és a dir, aquella matriu amb λ a la diagonal, 1 per sota la diagonal i la resta d'entrades 0.

Notació 4.1.2. En el cas d'alçada $s = 1$, posem $\mathcal{J}(\lambda; 1) = \lambda$.

Definició 4.1.3 (Matriu de Jordan de valor propi λ). S'anomena matriu de Jordan de valor propi λ a tota suma per blocs de matrius elementals de Jordan del mateix valor propi λ. Si $(s_1, \ldots, s_m)$ és una família d'enters positius, denotarem $\mathcal{J}(\lambda; s_1, \ldots, s_m) = \mathcal{J}(\lambda; s_1) \oplus \cdots \oplus \mathcal{J}(\lambda; s_m)$.

Definició 4.1.4 (Matriu de Jordan $\mathcal{J}$). Una *matriu de Jordan* $\mathcal{J}$ és una suma per blocs de matrius de Jordan de diferents valors propis $\mathcal{J} = \mathcal{J}(\lambda_1, s_1^1, \ldots, s_m^1) \oplus \mathcal{J}(\lambda_1, s_1^1, \ldots, s_m^1)$

Definició 4.1.5 (Base de Jordan). Si B és una base d'E tal que la matriu de f en la base B és una matriu de Jordan, direm que B és una base de Jordan d'E (respecte f).

Exemple 4.1.6. Sigui A una matriu 3×3 amb coeficients en $\mathbb{C}$. Veurem que sempre hi ha una matriu invertible C tal que $\mathcal{J} = C^{-1}AC$, on $\mathcal{J}$ és una de les següents matrius:

$$\begin{aligned}
\mathcal{J}(\lambda;3) &= \begin{pmatrix} \lambda & 0 & 0 \\ 1 & \lambda & 0 \\ 0 & 1 & \lambda \end{pmatrix}, \\
\mathcal{J}(\lambda;2)\oplus\mathcal{J}(\lambda;1) &= \begin{pmatrix} \lambda & 0 & 0 \\ 1 & \lambda & 0 \\ 0 & 0 & \lambda \end{pmatrix}, \\
\mathcal{J}(\lambda;1)\oplus\mathcal{J}(\lambda;1)\oplus\mathcal{J}(\lambda;1) &= \begin{pmatrix} \lambda & 0 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \lambda \end{pmatrix}, \\
\mathcal{J}(\lambda_1,2)\oplus\mathcal{J}(\lambda_2,1) &= \begin{pmatrix} \lambda_1 & 0 & 0 \\ 1 & \lambda_1 & 0 \\ 0 & 0 & \lambda_2 \end{pmatrix} \mid \lambda_1 \neq \lambda_2, \\
\mathcal{J}(\lambda_1,1)\oplus\mathcal{J}(\lambda_1,1)\oplus\mathcal{J}(\lambda_2,1) &= \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_1 & 0 \\ 0 & 0 & \lambda_2 \end{pmatrix} \mid \lambda_1 \neq \lambda_2, \\
\mathcal{J}(\lambda_1,1)\oplus\mathcal{J}(\lambda_2,1)\oplus\mathcal{J}(\lambda_3,1) &= \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix} \mid \lambda_1 \neq \lambda_2 \neq \lambda_3.
\end{aligned} \tag{4.1.2}$$

Algorisme per trobar bases de Jordan

En aquesta subsecció veurem que si E és un espai vectorial de dimensió finita sobre un cos $\mathbb{K}$ i f un endomorfisme d'E, de manera que el polinomi mínim d'f descompon completament en factors lineals, aleshores sempre hi ha una base de Jordan d'E. La matriu obtinguda serà única, sempre que obviem les permutacions que es poden donar entre blocs. Ho enunciem:

Teorema 4.1.7. *Si el polinomi mínim de $f \in \mathrm{End}(E)$ descompon en factors lineals, existeix una base d'E en la qual la matriu de F és de la forma*

$$\mathcal{J} = \begin{pmatrix} \boxed{\mathcal{J}(\lambda_0,s_0)} & & & 0 \\ & \boxed{\mathcal{J}(\lambda_1,s_1)} & & \\ & & \ddots & \\ 0 & & & \boxed{\mathcal{J}(\lambda_r,s_r)} \end{pmatrix} \tag{4.1.3}$$

on la matriu $\mathcal{J}$ és la matriu canònica de Jordan d'f.

Demostració. Consultar la demostració del corol·lari 4.1.14. ■

Observació 4.1.8. Aquesta condició en el polinomi mínim no és gens estranya: tinguem en compte, per exemple, per a espais vectorials sobre $\mathbb{C}$ aquesta condició se satisfà sempre i, per tant, tot endomorfisme sobre $\mathbb{C}$ té una base de Jordan. Pels endomorfismes d'espais vectorials sobre $\mathbb{R}$ no és sempre així. Els elements λ_i que apareixen a la diagonal de la matriu són els valors propis de f (possiblement repetits).

La base que busquem és unió de bases convenients dels subespais invariants E_i. Restringint-nos a aquests espais, podem suposar que el polinomi característic de $f \in \mathrm{End}(E)$ és $(t-\lambda)^n$ i el polinomi mínim, $(t-\lambda)^s$, $s \leq n = \dim E$ (així, fixem que f té un únic valor propi λ). Denotarem $g := f - \lambda\mathbb{I}$.

Algorisme 4.1.9.

1. *Considerem els subespais* $F_j := \ker((f-\lambda\mathbb{I})^j) = \ker(g^j)$ *i en determinem les dimensions* $n_j := \dim F_j$. *Obtenim una cadena de subespais*
$$0 = F_0 \subset F_1 \subset \cdots \subset F_{s-1} \subset F_s = E, \tag{4.1.4}$$
la qual ve donada per $\{\vec{0}\} \subset \ker(f-\lambda_i\mathbb{I}) \subset \ker(f-\lambda_i\mathbb{I})^2 \subset \cdots \subset \ker(f-\lambda_i\mathbb{I})^{s_i-1} \subsetneqq \ker(f-\lambda_i\mathbb{I})^{s_i}$ *i, per tot* $t \geq s_i$, $\ker(f-\lambda_i\mathbb{I})^{s_i} = \ker(f-\lambda_i\mathbb{I})^t = E_i$.
2. *Amb aquesta informació, ja podem dir com serà la matriu de Jordan: el nombre de blocs de Jordan elementals* $\mathcal{J}(\lambda; m)$ *de valor propi* λ *i alçada* m *és*
$$\nu_m := 2n_m - n_{m-1} - n_{m+1}. \tag{4.1.5}$$
3. *Trobem una base de* F_s/F_{s-1}: *busquem una base de* $\mathfrak{B}_{s-1}$ *de* F_{s-1} *i l'ampliem a una base* $\mathfrak{B}_s = \mathfrak{B}_{s-1} \cup \mathfrak{B}'_s$ *de* F_s. *Ens guardem els vectors de* $\mathfrak{B}'_s$ *i la resta els descartem.*
4. *Sigui* $\mathfrak{B}'_s = \{u_1^1, \ldots, u_{k_1}^1\}$ *el conjunt de vectors que ens hem guardat del pas anterior. Ara considerem els vectors* $g(\mathfrak{B}'_s) = \{g(u_1^1), \ldots, g(u_{k_1}^1)\}$, *que per construcció viuen a* F_{s-1}. *A més, determinen classes linealment independents en el quocient* F_{s-1}/F_{s-2} *(ho demostrem a 4.1.12). Completem aquesta família de classes de vectors a una base de* F_{s-1}/F_{s-2} *i n'escollim representants, o sigui: cal trobar una família de vectors* $\{u_1^2, \ldots, u_{k_2}^2\}$ *de* F_{s-1} *de manera que les classes* $[g(u_1^1)], \ldots, [g(u_{k_1}^1)], [u_1^2], \ldots, [u_{k_2}^2]$ *formin una base de* F_{s-1}/F_{s-2}.
5. *Aquest procés es va repetint: considerem les famílies de vectors* $g^2(u_i^1)$ *i* $g(u_i^2)$ *que viuen a* F_{s-2} *i completem aquestes famílies amb vectors* $\{u_i^3\}$ *de* F_{s-2} *per tal que les classes defineixin una base de* F_{s-2}/F_{s-3}*... fins a trobar una base de* F_1.
6. *Obtenim una taula de la forma:*

u_1^1	$\cdots$	$u_{k_1}^1$							
$g(u_1^1)$	$\cdots$	$g(u_{k_1}^1)$	u_1^2	$\cdots$	$u_{k_2}^2$				
$\vdots$	$\cdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$				
$g^{s-1}(u_1^1)$	$\cdots$	$g^{s-1}(u_{k_1}^1)$	$g^{s-2}(u_1^2)$	$\cdots$	$g^{s-2}(u_{k_2}^2)$	$\cdots$	u_1^s	$\cdots$	$u_{k_s}^s$

, (4.1.6)

on l'última fila correspon a la base d'F_1. *El conjunt de tots els vectors de la taula formen una base de Jordan. De fet:*

   1. *El nombre de files,* s, *és el grau del polinomi mínim.*
   2. *El nombre de columnes és la multiplicitat del valor propi* λ.
   3. *El nombre de vectors de cada fila* m *és* $\ell_m := n_m - n_{m-1} = \dim F_m - \dim F_{m-1}$.
   4. *El nombre de matrius* $\mathcal{J}(\lambda; m)$ *és* $\ell_m - \ell_{m+1} = 2n_m - n_{m-1} - n_{m+1}$, *que és la diferència entre el nombre de vectors de files consecutives.*
   5. *Cada columna és base d'un subespai* f*-cíclic de la descomposició d'*E.

Observació 4.1.10.

- $n_s \geq 1$ sempre, ja que $p_{s+1} \geq 1$. Això vol dir que sempre apareix, com a mínim, una matriu $\mathcal{J}(\lambda; s)$ de la mateixa dimensió.
- El problema de trobar la matriu canònica de Jordan d'un endomorfisme f queda, doncs, resolt si el polinomi mínim de f descompon en factors lineals $\mu_f(t) = (t-\lambda_1)^{s_1}\cdots(t-\lambda_r)^{s_r}$. Només ens cal conèixer per a cada valor propi λ_i els nombres n_t corresponents $(1 \leq t \leq s_i)$, els quals ens indiquen quants subespais f-cíclics de dimensió t apareixen (submatrius amb

el valor propi λ_i a la diagonal i uns a sota). En particular, si $s_i = 1$, el subespai invariant corresponent és un subespai de vectors propis. Així doncs, de manera general, podem dir que si l'endomorfisme té diferents valors propis, l'algorisme s'aplica per separat per a cada valor propi i després s'uneixen les bases trobades.

- L'endomorfisme és diagonalitzable si, i només si, $s_1 = \cdots = s_r = 1$.

Exemple 4.1.11. Considerem l'endomorfisme f d'$\mathbb{R}^4$ donat per la matriu

$$A = \begin{pmatrix} 3 & -1 & 1 & -7 \\ 9 & -3 & -7 & -1 \\ 0 & 0 & 4 & -8 \\ 0 & 0 & 2 & -4 \end{pmatrix}. \tag{4.1.7}$$

Resolent $\det(A)$ obtenim el polinomi característic, el qual ens dona $p(t) = t^4$. Per tant, hi ha un sol valor propi $\lambda = 0$ i tenim $B = A$, on B és la matriu de $g = f - \lambda\mathbb{I} = f$. Reduint la matriu B veiem que $\mathrm{rg}(B) = 2$ i $\mathrm{rg}(B^2) = 0$. Per tant, tenim que els subespais $0 = F_0 \subset F_1 \subset F_2 = \mathbb{R}^4$ tenen dimensions $n_1 = 2$ i $n_2 = 4$.

Busquem una base d'F_1, per exemple $\{(1,3,0,0),(0,-5,2,1)\}$ i l'ampliem a una base d'F^2: $\{(1,3,0,0),(0,-5,2,1),(0,1,0,0),(0,0,1,0)\}$. Ens quedem els vectors que hem afegit: $u_1^1 := (0,1,0,0)$, $u_2^1 := (0,0,1,0)$ i en considerem les seves imatges per g: $g(u_1^1) = (-1,3,0,0)$ i $g(u_2^1) = (1,-7,4,2)$. Com que en total ja tenim 4 vectors i estem a $\mathbb{R}^4$, ja hem acabat. Si recordem de construir una taula segons les condicions de 4.1.9, ens queda una taula tal que:

$$\begin{array}{|c|c|} \hline u_1^1 & u_2^1 \\ \hline g(u_1^1) & g(u_2^1) \\ \hline \end{array} = \begin{array}{|c|c|} \hline (0,1,0,0) & (0,0,1,0) \\ \hline (-1,-3,0,0) & (1,-7,4,2) \\ \hline \end{array} \tag{4.1.8}$$

Podem formar la matriu U de canvi de base amb tots aquests vectors. En general, les columnes d'U són els vectors de la taula, començant per la primera columna i anar baixant i escombrant columnes. Aleshores, es té

$$\begin{gathered} U = \begin{pmatrix} 0 & -1 & 0 & 1 \\ 1 & -3 & 0 & 7 \\ 0 & 0 & 1 & 4 \\ 0 & 0 & 0 & 2 \end{pmatrix}, \\ \mathcal{J} = \mathcal{J}(0,2) \oplus \mathcal{J}(0,2) = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \end{gathered} \tag{4.1.9}$$

que si comprovem amb $\mathcal{J} = \mathcal{J}(0,2) \oplus \mathcal{J}(0,2)$, hauríem de trobar que $\mathcal{J} = U^{-1}AU$, també.

Existència de bases de Jordan

Si tenim un subespai $F \subseteq E$ tenim l'espai quocient E/F i, per a tot $u \in E$, podem considerar la classe d'u en l'espai vectorial quocient E/F: $[u] \equiv u + F := \{u + \omega \mid \omega \in E\}$.

Lema 4.1.12. *Sigui f un endomorfisme primari d'E amb polinomi mínim $\mu_f = (t-\lambda)^s$ i $F_m = \ker(f-\lambda)^m$. Per a $m > 2$, sigui $\{u_1, \ldots, u_r\}$ una família de vectors de F_m tals que la família de les seves classes $\{[u_1], \ldots, [u_r]\}$ en F_m/F_{m-1} és linealment independent. Aleshores la família de les classes $\{[g(u_1)], \ldots, [g(u_r)]\}$ és linealment independent en F_{m-1}/F_{m-2}, on $g = f - \lambda$.*

Demostració. Suposem que existeixen escalars α_i tals que $\alpha_1[g(u_1)] + \cdots + \alpha_r[g(u_r)] = 0$. Això equival a dir que $\alpha_1 g(u_1) + \cdots + \alpha_r g(u_r) \in F_{m-2} = \ker(g^{m-2})$. Per tant, tenim

$$0 = g^{m-2}(\alpha_1 g(u_1) + \cdots + \alpha_r g(u_r)) = g^{m-1}(\alpha_1 u_1 + \cdots + \alpha_r u_r) \iff \alpha_1 u_1 + \cdots + \alpha_r u_r \in \ker(g^{m-1}) \\ \implies \alpha_1 u_1 + \cdots + \alpha_r u_r = 0, \tag{4.1.10}$$

on $\sum_i^r \alpha_i u_i = 0$ en el quocient F_m/F_{m-1} i, per tant, per hipòtesi tenim $\alpha_i = 0, \forall i$. ■

El lema anterior ens dona condicions necessàries per tal que el subespai F_m admeti una base de Jordan que conté un conjunt de vectors donat, per a $m \geq 1$.

Teorema 4.1.13. *Siguin $m \geq 1$ i $u_1, \ldots, u_r$ vectors de F_m tals que les seves classes F_m/F_{m-1} són linealment independents. Aleshores, existeix una base de Jordan de F_m que conté els vectors $\{u_1, \ldots, u_r\}$.*

Demostració. La prova és per inducció sobre m. Si $m = 1$, notem que F^0 és l'espai vectorial trivial i, per tant, la condició que les classes de $\{u_1, \ldots, u_r\}$ siguin linealment independents en $F_1/F_0 = F_1$ és equivalent a demanar que els vectors $\{u_1, \ldots, u_r\}$ siguin linealment independents en F_1. Podem completar aquest conjunt de vectors a una base de F_1 que és, per tant, una base de F_1 formada per vectors propis d'f, per definició de $F_1 = \ker(f - \lambda)$. En particular, és una base de Jordan que conté $\{u_1, \ldots, u_r\}$.

Suposem que $m > 1$ i que el teorema és cert fins a $m - 1$ (inducció completa). Sigui $\{u_1, \ldots, u_r\}$ una família de vectors d'F_m tals que la família de les seves classes $\{[u_1], \ldots, [u_r]\}$ en F_m/F_{m-1} és linealment independent. Completem aquesta família a una base de F_m/F_{m-1} i ens queda

$$\{[u_1], \ldots, [u_r], [u_{r+1}], \ldots, [u_s]\}. \tag{4.1.11}$$

Ara, escollim representants $u_{r+1}, \ldots, u_s \in F_m$ de les classes que hem afegit. Aplicant el lema anterior a la família $\{u_1, \ldots, u_s\}$ podem afirmar que les classes $[g(u_1)], \ldots, [g(u_s)]$ són linealment independents en F_{m-1}/F_{m-2}. Per la hipòtesi d'inducció, existeix una base de Jordan $\{g(u_1), \ldots, g(u_s)\} \cup \mathfrak{B}$ de F_{m-1} que conté el conjunt de vectors $\{g(u_1), \ldots, g(u_s)\}$. Per construcció, el conjunt de vectors

$$\{u_1, \ldots, u_r\} \cup \{g(u_1), \ldots, g(u_s)\} \cup B \tag{4.1.12}$$

és una base de Jordan de F_m que conté $\{u_1, \ldots, u_r\}$. ■

El següent corol·lari és idèntic al teorema 4.1.7, però no deixa de ser una conseqüència del teorema anterior. Aleshores, mantindrem aquesta duplicació perquè val la pena.

Corol·lari 4.1.14 (Existència de bases de Jordan). *Sigui E un espai vectorial sobre $\mathbb{K}$ i $f \in \mathrm{End}(E)$. Si el polinomi mínim de f descompon completament en factors lineals, existeix una base de Jordan d'E respecte d'f.*

Demostració. Tenint en compte la descomposició primària, podem suposar que el polinomi mínim d'f és primari, de la forma $\mu_f(t) = (t - \lambda)^s$. Considerem els espais vectorials $F_m = \ker(f - \lambda\mathbb{I})^m$, per a $0 \leq m \leq s$. Sigui $\mathfrak{B}_{s-1}$ una base arbitrària de F_{s-1} i l'ampliem a una base $\mathfrak{B}_s = \mathfrak{B}_{s-1} \cup A_s$ de $E = F_s$. En particular, A_s és un conjunt de vectors tals que les classes en

F_s/F_{s-1} formen una base. Aplicant el teorema anterior, A_s es pot ampliar a una base de Jordan de $F_s = E$.

Anem a veure la unicitat. Tenim una successió estrictament creixent de subespais invariants $0 \subset F_1 \subset \cdots \subset F_s = E$. Les mides dels blocs elementals de Jordan queden determinats per les dimensions dels subespais F_i. En efecte, el nombre de blocs de Jordan d'alçada i és $\nu_i := 2n_i - n_{i-1} - n_{i+1}$, on $n_i = \dim F_i$. ■

4.2 Aplicacions

Formes de Jordan

Definició 4.2.1 (Forma de Jordan)**.** Si A és una matriu $n \times n$, es diu *forma de Jordan d'A* a l'única matriu de Jordan $\mathcal{J}$, llevat permutacions de blocs de Jordan, de la forma $\mathcal{J} = U^{-1}AU$, amb U invertible.

Observació 4.2.2. Per provar que dues matrius quadrades A i B no són equivalents (recordar 1.2.9) hauríem de demostrar que no existeix cap matriu invertible U tal que $A = UBU^{-1}$. En canvi, mirant la forma de Jordan obtenim una forma totalment directa de saber-ho.

Corol·lari 4.2.3. *Dues matrius A i B són equivalents si, i només si, tenen la mateixa forma de Jordan llevat de permutacions dels seus blocs. En altres paraules, tota matriu és equivalent a una en forma de Jordan.*

Demostració. Siguin A i B dues matrius tals que $A = UBU^{-1}$, on U és una matriu invertible. Siguin V_A i V_B les matrius de canvi a bases de Jordan, de manera que

$$\mathcal{J}_A = V_A^{-1}AV_A \text{ i } \mathcal{J}_B = V_B^{-1}BV_B, \tag{4.2.1}$$

on $\mathcal{J}_A$ i $\mathcal{J}_B$ són matrius de Jordan. Aleshores, tenim que

$$V_A\mathcal{J}_AV_A^{-1} = UV_B\mathcal{J}_BV_B^{-1}U^{-1} \iff J_A = V_A^{-1}UV_B\mathcal{J}_BV_B^{-1}U^{-1}V_A. \tag{4.2.2}$$

Si $W = V_A^{-1}UV_B$, podem escriure $\mathcal{J}_A = W\mathcal{J}_BW^{-1}$. Hem de tenir en compte, però, que les matrius de Jordan són úniques llevat de permutacions dels blocs i, per tant, $\mathcal{J}_A$ i $\mathcal{J}_B$ han de ser la mateixa matriu amb possibles permutacions de blocs. ■

Potències de matrius

Una altra aplicació de les formes de Jordan és en el càlcul de les potències de matrius. Ja hem vist que si tenim una matriu diagonal només cal fer la potència de cada element de la diagonal. Amb les formes de Jordan, aleshores, obtenim una fórmula per al càlcul.

Procés 4.2.4 (Càcul de potències de les matrius)**.**

- *Si tenim una matriu quadrada diagonal per blocs, de la forma $A = A_1 \oplus \cdots \oplus A_k$, aleshores la seva potència es calcula fent les potències de cada bloc: $A^n = A_1^n \oplus \cdots \oplus A_k^n$.*
- *D'altra banda, si $A = U\mathcal{J}U^{-1}$, aleshores tenim $A^n = U\mathcal{J}^nU^{-1}$. Per tant, per calcular una potència d'una matriu A quadrada en podem calcular la seva forma de Jordan $\mathcal{J}$ i buscar les potències de cadascun dels seus blocs elementals.*

- *El cas general satisfà:*

$$\mathcal{J}(\lambda; s)^n = \begin{pmatrix} \lambda^n & 0 & 0 & \cdots & 0 \\ n\lambda^{n-1} & \lambda^n & 0 & \cdots & 0 \\ \binom{n}{2}\lambda^{n-2} & n\lambda^{n-1} & \lambda^n & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ \binom{n}{s-1}\lambda^{n-s+1} & \binom{n}{s-2}\lambda^{n-s+2} & \cdots & n\lambda^{n-1} & \lambda^n \end{pmatrix}, \tag{4.2.3}$$

on hem usat que $\mathcal{J}(\lambda; s)^n = (\lambda\mathbb{I} + \mathcal{J}(0; s))^n$ i hem desenvolupat pel binomi de Newton.

Exemple 4.2.5. Per al cas 2×2 és fàcil comprovar que

$$\mathcal{J}(\lambda; 2)^n = \begin{pmatrix} \lambda & 0 \\ 1 & \lambda \end{pmatrix}^n = \begin{pmatrix} \lambda^n & 0 \\ n\lambda^{n-1} & \lambda^n \end{pmatrix} \tag{4.2.4}$$

Exemple 4.2.6 (Matrius exponencials). Considerem la matriu

$$A = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}, \tag{4.2.5}$$

que és prou senzilla. Calculant una mica, trobem:

$$A^2 = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}, A^3 = \begin{pmatrix} 1 & 2 \\ 2 & 3 \end{pmatrix}, \ldots \tag{4.2.6}$$

i aquí identifiquem els nombres de Fibonacci: $0, 1, 1, 2, 3, 5, 8, 13, 21, 34, \ldots$ Per tant, per trobar una fórmula explícita per e^A cal trobar una fórmula explícita del k-èsim nombre de Fibonacci, molt més fàcil a partir de la matriu de Jordan:

1. Per a una matriu $A = A_1 \oplus \cdots \oplus A_k$ diagonal per blocs, tenim $e^A = e^{A_1} \oplus \cdots \oplus e^{A_k}$.
2. Si $A = U\mathcal{J}U^{-1}$, aleshores tenim $e^A = Ue^J U^{-1}$. Per a $\mathcal{J} = \mathcal{J}(\lambda; s)$, amb $s \geq 1$, es té

$$e^J = \begin{pmatrix} e^\lambda & 0 & 0 & \cdots & 0 \\ \frac{e^\lambda}{1!} & e^\lambda & 0 & \cdots & 0 \\ \frac{e^\lambda}{2!} & \frac{e^\lambda}{1!} & e^\lambda & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ \frac{e^\lambda}{(s-1)!} & \cdots & \frac{e^\lambda}{2!} & \frac{e^\lambda}{1!} & e^\lambda \end{pmatrix}. \tag{4.2.7}$$

Cayley-Hamilton

Les formes de Jordan també ens permeten demostrar el teorema de Cayley-Hamilton com a corol·lari. L'expressarem d'una manera diferent, però equivalent, a 3.5.2.

Teorema 4.2.7. *Sigui A una matriu $n \times n$ sobre un cos $\mathbb{K}$, i sigui $p_A(t)$ el seu polinomi característic. Aleshores, $P_A(A) = 0$.*

Demostració. Suposem $A_1, \ldots, A_m$ els blocs de Jordan d'una matriu A. Aleshores, $p(A)$ és una matriu formada pels blocs $p(A_1), \ldots, p(A_m)$ a la diagonal, i p és el producte dels polinomis característics de cada bloc. És suficient, doncs, amb mostrar que $p_i(A_i) = 0$ per al polinomi característic p_i d'A_i. Si $A_i = \mathcal{J}(\lambda; k)$, aleshores $p_i(t) = (\lambda - t)^k$ i $p_i(A_i) = (\lambda\mathbb{I} - A_i)^k = 0$. ■

4.3

Formes de Jordan sobre $\mathbb{R}$

Sabem que per a matrius complexes tenim formes de Jordan, però sobre $\mathbb{R}$ no és sempre així (només quan tots els valors propis són reals). Si una matriu real té un o més valors propis complexos, la seva forma de Jordan pot ser substituïda.

Notació 4.3.1 ($\mathcal{J}(\alpha,\beta;s)$). Denotem per $\mathcal{J}(\alpha,\beta;s)$ a:

$$\mathcal{J}(\alpha,\beta;s) := \begin{pmatrix} C(\alpha,\beta) & 0 & \cdots & 0 & 0 \\ \mathbb{I}_2 & C(\alpha,\beta) & \cdots & 0 & 0 \\ 0 & \mathbb{I}_2 & C(\alpha,\beta) & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & \mathbb{I}_2 & C(\alpha,\beta) \end{pmatrix} \tag{4.3.1}$$

un bloc de Jordan real, donat per una matriu $2s \times 2s$ on hi ha una matriu $C(\alpha,\beta)$ a la diagonal i matrius $\mathbb{I}_2$ per sota de la diagonal.

Definició 4.3.2 (Matriu de Jordan real). És una matriu diagonal per blocs on els blocs són de la forma $\mathcal{J}(\lambda;s)$ (els blocs de Jordan *habituals*) o bé de la forma $\mathcal{J}'(\alpha,\beta;s)$ (els blocs de Jordan reals).

La prova de l'existència de bases de Jordan reals se segueix de la subsecció que hem fet abans amb força detall. En el cas de valors propis complexos conjugats, sempre es poden escollir parells de vectors que siguin complexos conjugats. Prenent com a base la part real i la part imaginària d'aquest parell, la matriu en aquesta vase és de la forma $C(\alpha,\beta)$.

Exemple 4.3.3. Considerem la matriu

$$A = \begin{pmatrix} 1 & 1 \\ -4 & 1 \end{pmatrix} \tag{4.3.2}$$

que té dos valors propis $\lambda_1 = 1 + 2i$ i $\lambda_2 = 1 - 2i$. Si la tractem com a matriu complexa (permetem vectors propis complexos) aleshores aquesta matriu diagonalitza: trobem vectors propis $v_1 = (1, -2i)$ i $v_2 = (1, 2i)$ i, per tant, si definim

$$C = (v_1|v_2) = \begin{pmatrix} 1 & 1 \\ -2i & 2i \end{pmatrix} \tag{4.3.3}$$

aleshores

$$C^{-1}AC = \begin{pmatrix} 1-2i & 0 \\ 0 & 1+2i \end{pmatrix}. \tag{4.3.4}$$

Si la tractem com a matriu real (i, per tant, la matriu de canvi de base ha de ser real) sabem que no diagonalitza, però té forma de Jordan real:

$$\mathcal{J} = \begin{pmatrix} 1 & 2 \\ -2 & 1 \end{pmatrix} \tag{4.3.5}$$

Per trobar el canvi de base, prenem $u_1 = \Re(v_1) = (1, 0)$ i $u_2 = \Im(v_2) = (0, 2)$. Aleshores,

$$U = (u_1|u_2) = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix} \text{ i } \mathcal{J} = U^{-1}AU. \tag{4.3.6}$$

Bibliografia

[CL09] Manuel Castellet i Irene Llerena. *Àlgebra lineal i geometria*. Vol. 1. Univ. Autònoma de Barcelona, 2009.

[Cre13] Teresa Crespo. *Àlgebra Lineal*. Notes de l'assignatura d'Àlgebra Lineal. Gen. de 2013.

Índex terminològic