TEMA 3. ESTRUTURA DAS APLICACIÓNS LINEAIS
Apuntamentos de clase da materia de Álxebra Linear e Multilinear.

Sexa K un corpo. Dada unha aplicación lineal $f : E \to F$ entre dous K-espazos vectoriais diferentes, pódense escoller bases dos dous espazos de maneira que a matriz correspondente sexa moi sinxela. En cambio, para endomorfismos dun mesmo espazo nos que se traballa coa mesma base, os problemas análogos son máis complicados.
En todo o capítulo, sexa K un corpo e E un espazo vectorial. Ao longo deste tema imos traballar tres cuestións fundamentais.

- **Diagonalización.** Caracterizar os endomorfismos que admiten unha base diagonal.
- **Álxebra de endomorfismos.** Introducir a álxebra de endomorfismos e a noción de polinomio anulador dun endomorfismo. Reformular a caracterización de diagonalización nestes termos.
- **Forma de Jordan.** En certos casos nos que o endomorfismo non diagonaliza, construír unha matriz *en forma reducida.* É o que se coñece como forma de Jordan.

1. Diagonalización

1.1. Vectores propios e valores propios

Definición. Sexa $f : E \to E$ un endomorfismo.

- Un vector $v \in E$ é un vector propio de f se $v \neq 0$ e existe un escalar $\lambda \in K$ tal que $f(v) = \lambda v$.
- Un escalar $\lambda \in K$ é un valor propio de f se existe un vector $v \neq 0$ para o cal $f(v) = \lambda v$.

A cada vector propio correspóndelle un único valor propio; pero para cada escalar hai moitos vectores propios que o teñen como valor propio.

Definición. Dado un endomorfismo $f \in \mathrm{End}(E)$ e un escalar $\lambda \in K$, pódese considerar o endomorfismo $f_\lambda := f - \lambda\mathbb{I} \in \mathrm{End}(E)$. O seu núcleo é

$$E_\lambda := \{v \in E \text{ con } f(v) = \lambda v\} = \ker(f - \lambda\mathbb{I}),$$

e é o subespazo vectorial de E formado polo vector cero e tódolos vectores propios de f con valor propio λ, en caso de haber algún. Chámase o subespazo propio de valor propio λ.

Exemplo. Consideremos a matriz

$$A = \begin{pmatrix} 3 & 4 \\ 6 & 1 \end{pmatrix}.$$

Entón, o vector $(1, 1)$ é un vector propio de valor propio 7; en xeral, calquera vector da forma (a, a), con $a \neq 0$, é un vector propio de valor propio 7. Por outro lado, -3 tamén é un vector propio, con $(-2, 3)$ un xerador do subespazo de vectores propios correspondente.

Proposición. Se $\lambda_1, \ldots, \lambda_r$ son valores propios dun endomorfismo, os subespazos propios correspondentes están en suma directa:

$$\sum_{i=1}^{r} E_{\lambda_i} = E_{\lambda_1} \oplus E_{\lambda_2} \oplus \ldots \oplus E_{\lambda_r}.$$

Demostración. Faremos a demostración por indución sobre r. Como para $r = 1$ o enunciado é trivialmente certo, suporémolo demostrado ata $r - 1$. Supoñamos entón que existen $v_i \in E_{\lambda_i}$ de maneira que $v_1 + \ldots + v_r = 0$. Aplicando $f_{\lambda_r} = f - \lambda_r \mathbb{I}$ a cada lado temos que

$$0 = f_{\lambda_r}(v_1 + \ldots + v_r) = (\lambda_1 - \lambda_r)v_1 + \ldots + (\lambda_{r-1} - \lambda_r)v_{r-1}.$$

Aplicando a hipótese de indución temos os vectores v_i, para $1 \leq i \leq r - 1$ son independentes, e como os coeficientes $\lambda_i - \lambda_r \neq 0$ xa que os valores propios son diferentes, chegamos a que $v_1 = \ldots = v_{r-1} = 0$. De aquí dedúcese que $v_r = 0$ e tódolos vectores teñen que ser nulos. □

É importante ter en conta que o feito de estar en suma directa tamén implica que a descomposición dun elemento como suma dos diferentes subespazos é única. Polo tanto, se $\lambda_1, \ldots, \lambda_r$ son valores propios dun endomorfismo $f \in \mathrm{End}(V)$ e para λ_i, $\mathcal{B}_i = \{v_{i,1}, \ldots, v_{i,t_i}\}$ é unha base do subespazo propio E_{λ_i}, entón a unión dos vectores $\cup_{i=1}^{r} \mathcal{B}_i$ é unha familia de vectores linealmente independentes.

1.2. Polinomio característico

Definición. O polinomio característico dunha matriz cadrada $A = (a_{ij}) \in \mathcal{M}_n(K)$ é o polinomio

$$\det(A - X\mathbb{I}) = \begin{vmatrix} a_{11} - X & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} - X & \ldots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \ldots & a_{nn} - X \end{vmatrix}.$$

O polinomio característico dun endomorfismo $f \in \mathrm{End}(E)$ dun espazo de dimensión finita defínese como o da matriz do endomorfismo dunha base $\mathcal{B}$ de E calquera.

En certas referencias defínese o polinomio característico como $\det(X\mathbb{I} - A)$, pero o único que cambia iso é o signo (facendo dese xeito que sempre sexa mónico). O seguinte resultado resume as propiedades principais do polinomio característico. Recordamos que para unha matriz $A \in \mathcal{M}_n(K)$, a traza de A, $\mathrm{Tr}(A)$, é a suma dos elementos da súa diagonal.

Proposición. Sexa f un endomorfismo e A a súa matriz asociada nunha base fixada. O polinomio característico cumpre as seguintes propiedades.

(a) $\mathrm{Char}(f; X)$ está ben definido (non depende da base do endomorfismo).

(b) Ten grao n e o coeficiente principal é $(-1)^n$.

(c) O coeficiente de X^{n-1} é $(-1)^{n-1}\mathrm{Tr}(A)$.

(d) O termo independente é $\det(A)$.

Demostración. (a) Imos comparar o polinomio característico de f en dúas bases $\mathcal{B}$ e $\mathcal{B}'$ con matrices asociadas A e B, respectivamente. Sexa $B = P^{-1}AP$, onde P é a matriz de cambio de base. Entón

$$\begin{aligned}\mathrm{Char}_{\mathcal{B}'}(f;X)\, &\det(P^{-1}AP - X\mathbb{I}_n) \\ &= \det(P^{-1}AP - P^{-1}(X\mathbb{I}_n)P) \\ &= \det(P^{-1}(A - X\mathbb{I}_n)P) \\ &= \det(P)^{-1}\det(A - X\mathbb{I}_n)\det(P) \\ &= \det(A - X\mathbb{I}_n).\end{aligned}$$

(b) O coeficiente principal é o produto dos coeficientes que acompañan a n en cada entrada da matriz, isto é, $(-1)^n$.

(c) O coeficiente de X^{n-1} é o seu coeficiente no polinomio $\prod_{i=1}^n (a_{ii} - X)$, que correspondente a coller a suma de tódolos elementos $(-1)^{n-1}a_{ii}$.

(d) O termo independente é o resultado de pór $X = 0$, co cal queda simplemente o determinante da matriz A.

□

Cando estea clara a identificación, falaremos indistintamente do polinomio característico da matriz ou do endomorfismo, isto é, trataremos $\mathrm{Char}(A;X)$ e $\mathrm{Char}(f;X)$ como dous obxectos análogos.

Proposición. Un escalar $\lambda \in K$ é un valor propio de f se e soamente se é unha raíz do seu polinomio característico.

Demostración. Un escalar $\lambda \in K$ é unha raíz do polinomio característico se e soamente se $\det(A - \lambda\mathbb{I}) = 0$, isto é, cando o núcleo do endomorfismo dado por $A - \lambda\mathbb{I}$ é non trivial. Isto é o mesmo que dicir que existe $v \neq 0$ con $(A - \lambda\mathbb{I})(v) = 0$, ou equivalentemente $Av = \lambda v$, que é precisamente a definición de valor propio. □

Definición. Sexa $\lambda \in K$ un valor propio do endomorfismo f. Defínense as seguintes nocións:

(a) A multiplicidade alxébrica de λ, $m_a(\lambda)$, é a súa multiplicidade como raíz do polinomio característico de f.

(b) A multiplicidade xeométrica de λ, $m_x(\lambda)$, é a dimensión do subespazo propio E_λ.

Exemplo. Consideremos o endomorfismo f que na base canónica ten por matriz

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & 4 \\ 0 & 0 & 1 \end{pmatrix}.$$

Entón, o seu polinomio característico é $\mathrm{Char}(A;X) = -(X-1)^3$, co que a multiplicidade alxébrica de 1 é 3. Para determinar a multiplicidade xeométrica, achamos o núcleo de $A - \mathbb{I}$, isto é

$$\begin{pmatrix} 0 & 2 & 3 \\ 0 & 0 & 4 \\ 0 & 0 & 0 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

Polo tanto, $\ker(A - \mathbb{I}) = \langle(1,0,0)\rangle$. Polo tanto, a multiplicidade xeométrica é 1.

En xeral, este fenómeno que observamos no exemplo anterior dáse máis en xeral: a multiplicidade xeométrica nunca é superior á alxébrica. Porén, a súa demostración require un lema sobre determinantes.

Lema. Sexa $A = (a_{ij}) \in \mathcal{M}_n(K)$ unha matriz da forma

$$A = \begin{pmatrix} X & Y \\ 0 & Z \end{pmatrix},$$

con $X \in \mathcal{M}_r(K)$ un bloque cadrado de orde r, $Y \in \mathcal{M}_{r\times(n-r)}$ e $Z \in \mathcal{M}_{(n-r)\times(n-r)}$. Entón, $\det(A) = \det(X) \cdot \det(Z)$.

Demostración. Recordemos que $\det(A) = \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) a_{1\sigma(1)} \cdots a_{n\sigma(n)}$. Se para algún $r+1 \leq i \leq n$ se ten que $\sigma(i) \in \{1, \dots, r\}$, entón $a_{i\sigma(i)} = 0$, co cal os únicos σ que temos que considerar son aqueles da forma $\sigma = \tau_1\tau_2$, con $\tau_1 \in S_r$ e $\tau_2 \in S_{n-r}$. Aquí, identificamos S_{n-r} coas permutacións de $\{r+1, \dots, n\}$. Entón,

$$\begin{aligned} \det(A) &= \sum_{(\tau_1,\tau_2)\in S_r\times S_{n-r}} \operatorname{sgn}(\sigma) a_{1\tau_1(1)} \cdots a_{r\tau_1(r)} a_{r+1\tau_2(r+1)} \cdots a_{n\tau(n)} \\ &= \Big(\sum_{\tau_1 \in S_r} \operatorname{sgn}(\tau_1) a_{1\tau_1(1)} \cdots a_{r\tau_1(r)} \Big) \Big(\sum_{\tau_2 \in S_{n-r}} \operatorname{sgn}(\tau_2) a_{r+1\tau_2(r+1)} \cdots a_{n\tau_2(n)} \Big) \\ &= \det(X)\det(Z). \end{aligned}$$

□

Proposición. As multiplicidades alxébrica e xeométrica dun valor propio λ cumpren as desigualdades

$$1 \leq m_x(\lambda) \leq m_a(\lambda).$$

Demostración. Por definición de valor propio, existe $v \neq 0$ en E_λ, co cal a desigualdade $m_x(\lambda) \geq 1$ está clara. Sexa agora $\{v_1, \dots, v_k\}$ unha base de E_λ, e sexa $v_{k+1}, \dots, v_n$ unha extensión a unha base de V. Entón, a matriz de f na base $\{v_1, \dots, v_k, v_{k+1}, v_n\}$ é a matriz por bloques dada por

$$A = \begin{pmatrix} \lambda \cdot \mathbb{I}_k & A_1 \\ 0_{(n-k)\times k} & A_2 \end{pmatrix},$$

onde $A_1 \in \mathcal{M}_{k\times(n-k)}$ e $A_2 \in \mathcal{M}_{(n-k)\times(n-k)}$. O seu polinomio característico é

$$\operatorname{Char}(A; X) = (\lambda - X)^k \cdot \det(A_2 - X\mathbb{I}_{n-k}),$$

co cal temos que a multipilicidade alxébrica de λ é polo menos k, e de feito a igualdade dáse cando λ non é unha raíz de $\det(A_2 - X\mathbb{I}_{n-k})$. □

1.3. Condicións de diagonalización

De cara á seguinte definición, recordamos que $\mathrm{GL}_n(K)$ é o conxunto das matrices $n \times n$ con coeficientes en K e con determinante diferente de 0.

Definición. Sexa f un endomorfismo de E. Dise que f diagonaliza se E ten unha base de vectores propios de f. Do mesmo xeito, unha matriz $A \in \mathcal{M}_n(K)$ diagonaliza se existe $P \in \mathrm{GL}_n(K)$ tal que $P^{-1}AP$ diagonaliza.

Proposición. O endomorfismo f diagonaliza se e soamente se cumpre simultaneamente as seguintes dúas condicións.

(a) O polinomio característico de f descompón en factores lineais sobre K:

$$\mathrm{Char}(f;X) = (-1)^n(X-\lambda_1)^{m_1}\cdots(X-\lambda_r)^{m_r},$$

sendo λ_i raíces diferentes e $m_1+\ldots+m_r = n$.

(b) Para todo valor propio λ_i, $m_a(\lambda_i) = m_x(\lambda_i)$.

Demostración. Supoñamos primeiro que diagonaliza. Sexa $\{v_{1,1},\ldots,v_{1,m_1}\}$ unha base ordenada correspondente ao primeiro vector propio, λ_1; $\{v_{2,1},\ldots,v_{2,m_2}\}$, unha base ordenada de λ_2; e así sucesivamente ata λ_r. Entón, a matriz nesa base é diagonal e ten na diagonal os valores propios λ_i repetidos m_i veces; polo tanto

$$\mathrm{Char}(f;X) = \prod_{i=1}^{r}(\lambda_i - X)^{m_i}.$$

Para cada índice k, o subespazo propio E_{λ_k} é o subespazo $\langle v_{k,1},\ldots,v_{k,m_k}\rangle$ e polo tanto a súa dimensión é m_k; isto é así porque os vectores $v_{k,j}$ pertencen a este espazo, son linealmente independentes por ser parte dunha base e polo resultado anterior, a multiplicidade xeométrica non pode ser superior á alxébrica. Alternativamente, se houbera outro vector $v = \sum_{i=1}^{r}\sum_{j=1}^{m_i} x_{ij}v_{ij} \in E_{\lambda_i}$, entón

$$f(v) = \sum_{i=1}^{r}\sum_{j=1}^{m_i} x_{ij}f(v_{ij}) = \sum_{i=1}^{r}\sum_{j=1}^{m_i} x_{ij}\lambda_i v_{ij} = \lambda_k v = \lambda_k \sum_{i=1}^{r}\sum_{j=1}^{m_i} x_{ij}v_{ij}.$$

Por unicidade da expresión dun vector nunha base, dedúcese que $\lambda_k x_{ij} = \lambda_i x_{ij}$ e polo tanto $x_{ij} = 0$ para todo $i \neq k$ e v só pode ter compoñentes non nulas nos vectores v_{kj}. Supoñamos agora que se cumpren as dúas condicións do enunciado. Entón, a unión das bases dos subespazos propios contén n vectores independentes (xa que vectores propios de valores propios diferentes son linealmente independentes) e polo tanto é unha base de vectores propios do espazo. $\square$

Deste resultado é evidente que se o polinomio característico ten n raíces diferentes, entón o endomorfismo diagonaliza. Nese caso, ademais,

$$E = E_{\lambda_1} \oplus \ldots \oplus E_{\lambda_n},$$

co cal temos unha descomposición do espazo nunha base de vectores propios.

1.4. Exemplos de diagonalización

O procedemento para diagonalizar unha matriz A ou un endomorfismo f é o seguinte.

(a) Encontrar o seu polinomio característico. Se non descompón en factores lineais, entón o endomorfismo non diagonaliza.

(b) Para cada unha das raíces do polinomio característico, achar a súa multiplicidade xeométrica e unha base de vectores propios. Se a multiplicidade xeométrica non coincide coa alxébrica para algún valor propio, entón o endomorfismo non diagonaliza.

(c) Se o polinomio característico descompón e tódalas multiplicidades alxébricas coinciden coas correspondentes xeométricas, entón a matriz $A = PDP^{-1}$, onde P é unha matriz que ten por columnas os vectores propios e D a matriz diagonal que ten por entradas os valores propios correspondentes aos vectores propios da mesma columna.

Imos facer tres exemplos.

- *O polinomio característico non descompón completamente.* Sexa $A \in \mathcal{M}_3(\mathbb{R})$ a matriz
$$A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{pmatrix}.$$
O polinomio característico é $\mathrm{Char}(A; X) = -(X-1)(X^2+1)$. Como o factor $X^2 + 1$ non descompón en $\mathbb{R}[X]$, a matriz non diagonaliza.

  En cambio, si diagonaliza en $\mathbb{C}$. Nese caso, os valores propios son 1, i e $-i$. Como cada un ten multiplicidade alxébrica 1, sabemos que as multiplicidades xeométricas tamén serán 1 e a matriz vai diagonalizar. En concreto, temos que $E_1 = \langle(1,0,0)\rangle$, $E_i = \langle(0,1,i)\rangle$ e $E_{-i} = \langle(0,1,-i)$. Polo tanto,
$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & i & -i \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & i & 0 \\ 0 & -i & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & i & -i \end{pmatrix}^{-1}.$$

- *Algunha multiplicidade xeométrica é menor que a correspondente alxébrica.* Sexa $B \in \mathcal{M}_3(\mathbb{R})$ a matriz
$$B = \begin{pmatrix} 4 & 3 & 1 \\ -3 & -2 & 1 \\ 0 & 0 & 1 \end{pmatrix}.$$
O polinomio característico é $\mathrm{Char}(B; X) = -(X-1)^3$, co cal 1 é o único valor propio e a súa multiplicidade alxébrica é 3. Para achar a multiplicidade xeométrica facemos $\ker(B - \mathbb{I})$, isto é, resolvemos
$$\begin{pmatrix} 3 & 3 & 1 \\ -3 & -3 & 1 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$
As condicións que obtemos son $x + y = 0$ e $z = 0$, e polo tanto o subespazo de vectores propios é o xerado por $\langle(1,-1,0)\rangle$. Como a multiplicidade xeométrica é 1, non coincide coa alxébrica e a matriz non é diagonalizable.

- *Cúmprense as dúas condicións de diagonalización.* Sexa $C \in \mathcal{M}_3(\mathbb{R})$ a matriz
$$C = \begin{pmatrix} 5 & -3 & -3 \\ 3 & -1 & -3 \\ 3 & -3 & -1 \end{pmatrix}.$$
O polinomio característico é $\mathrm{Char}(C; X) = -(X-2)^2(X+1)$, co cal os valores propios son 2, con multiplicidade 2; e -1, con multiplicidade 1. Imos calcular

agora as multiplicidades xeométricas e os subespazos de vectores propios. En concreto, temos que

$$\ker(C - 2\mathbb{I}) = \langle (1,1,0), (1,0,1) \rangle$$

e

$$\ker(C + \mathbb{I}) = \langle (1,1,1) \rangle,$$

polo que en ambos casos a multiplicidade alxébrica coincide coa xeométrica. Temos entón que

$$\begin{pmatrix} 5 & -3 & -3 \\ 3 & -1 & -3 \\ 3 & -3 & -1 \end{pmatrix} = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix} \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} 1 & 1 & 1 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix}^{-1}.$$

É importante observar que sempre hai liberdade á hora de escoller a base de $\ker(f_\lambda)$, o que se fai especialmente visible no caso no que a multiplicidade é maior ou igual que dous.

Antes de rematar a sección, imos salientar dous aspectos importantes. O primeiro é que nestes exemplos é moi claro como certas cantidades asociadas a un endomorfismo (ou a unha matriz) son invariantes por cambio de base. Unha delas é o determinante, que é o produto dos valores propios (termo independente do polinomio característico); outra delas é a traza, que é a suma dos valores propios (salvo signo, termo con X^{n-1} no polinomio característico).
O segundo aspecto a destacar é que a forma diagonal pode resultar útil para certos cálculos matriciais. Por exemplo, se $C = PDP^{-1}$, con D diagonal, temos que

$$C^n = (PDP^{-1}) \cdot (PDP^{-1}) \cdot (PDP^{-1}) \overset{n}{\cdots} (PDP^{-1}) \cdot (PDP^{-1}) = PD^nP^{-1}.$$

2. Polinomio mínimo e subespazos invariantes

2.1. Polinomio mínimo

Sexa f un endomorfismo. Podemos considerar as súas potencias $f^0 = \mathbb{I}$, $f^1 = f$, $f^2 = f \circ f$ e así sucesivamente, tendo $f^r = f \circ f^{r-1}$ para calquera n.

Definición. Sexa K un corpo. Unha álxebra F sobre K é un espazo vectorial cunha operación bilineal $F \times F \longrightarrow F$ (xeralmente chamada produto) que cumpre as seguintes propiedades, onde $x, y, z \in F$ e $a, b \in K$:

- Propiedade distributiva pola dereita: $x(y+z) = xy + xz$.
- Propiedade distributiva pola esquerda: $(y+z)x = yx + zx$.
- Compatibilidade coa multiplicación por escalares: $(ax)(by) = (ab)xy$.

Se a álxebra admite unha descomposición $F = \oplus_{n \geq 0} F_n$ con $F_r \cdot F_s \subset F_{r+s}$, a álxebra dise que é graduada.
Un morfismo de álxebras $\Phi : F \longrightarrow F$ é unha aplicación lineal que tamén respecta o produto, é dicir, $\Phi(xy) = \Phi(x)\Phi(y)$.

Exemplo. As matrices con entradas nun corpo K son unha álxebra. Do mesmo xeito, os endomorfismos dun espazo vectorial E tamén forman unha álxebra. O anel de polinomios $K[X]$ tamén constitúe unha álxebra, que como veremos a continuación, está moi relacionada coa álxebra de endomorfismos.
Outro exemplo que traballamos neste curso son os tensores covariantes, que ademais teñen unha gradación (considerando como produto o produto tensorial). No último tema introduciremos, tamén no contexto dos tensores, a álxebra exterior.

Definición. Sexa f un endomorfismo. A aplicación avaliación en f é o morfismo dado por

$$\Phi_f \,:\, K[X] \longrightarrow \mathrm{End}(E), \quad p(X) = \sum_{i=0}^{r} a_i X^i \mapsto p(f) = \sum_{i=0}^{r} a_i f^i.$$

É sinxelo comprobar que Φ_f é un morfismo de álxebras. O próximo obxectivo será estudar o seu núcleo. Para a imaxe de $p(X)$, pomos $\Phi_f(p(X))$ ou simplemente $p(f)$.

Definición. Un polinomio $p(X) \in \ker(\Phi_f)$ dise que é un polinomio anulador de f.

Proposición. O conxunto de polinomios anuladores diferentes de cero é non baleiro. Sexa $\mathrm{Min}(f;X)$ un polinomio non nulo, mónico e de grao mínimo no núcleo de Φ_f. Entón, $\mathrm{Min}(f;X)$ está ben definido (isto é, só hai un que cumpra esas condicións) e calquera outro elemento do núcleo de Φ_f é un múltiplo de $\mathrm{Min}(f;X)$.

Demostración. Se a dimensión de E é n, o espazo vectorial de endomorfismos ten dimensión n^2, co cal non tódalas potencias poden ser linealmente independentes e ten que existir unha combinación lineal non trivial dos f^i que dea 0. Isto demostra que o conxunto dos polinomios anuladores diferentes de cero non é baleiro.
Entre tódolos polinomios que anulan f, consideremos un de grao mínimo, $p(X)$. Se $q(X)$ é outro polinomio que anula f, podemos facer a división euclidiana e temos polinomios $a(X)$ e $b(X)$, con $\deg(b(X)) < \deg(p(X))$ e tal que

$$q(X) = p(X)a(X) + b(X).$$

Como $q(f) = 0$ e $p(f) = 0$, entón tamén pasa que $b(f) = 0$; pero temos que se $b \neq 0$, b é un polinomio non nulo de grao menor que $p(X)$ que anula f, o cal é unha contradición. Polo tanto, $b = 0$. Se impomos que o polinomio sexa mónico, está claro que só hai un de grao mínimo, xa que calquera outro sería múltiplo del e non é posible obter un polinomio mónico multiplicando outro mónico por unha constante. $\square$

No caso do polinomio mínimo, en vez de escribir $\Phi_f(\mathrm{Min}(f;X)) = 0$, indicaremos simplemente $\mathrm{Min}(f;f) = 0$, e o mesmo no caso do polinomio característico.

Exemplo. Se $E \neq \{0\}$ e $f = 0$, temos que $\mathrm{Min}(f;X) = X$. Se $f = \mathbb{I}$, entón $X - 1$. En cambio, se $E = \{0\}$, entón calquera polinomio anula o único endomorfismo g de E, polo que $\mathrm{Min}(g;X) = 1$. Este é ademais o único caso no que o polinomio mínimo é 1.
Consideremos os endomorfismos que teñen por representacións matriciais na base ordinaria

$$A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, \quad B = \begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix}, \quad C = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}.$$

Un polinomio anulador de A é $X^2 + 1$, que tamén é o polinomio mínimo. O polinomio mínimo de B é $(X - 2)(X - 3) = X^2 - 5X + 6$. O polinomio mínimo de C é $(X -$

$1)(X-2) = X^2 - 3X + 2$. Non é complicado ver que para unha matriz diagonal con entradas $\lambda_1, \ldots, \lambda_n$, o produto $\prod_{i=1}^n (X - \lambda_i)$ é un polinomio anulador, que é de feito o mínimo se eliminamos os factores repetidos.

Sexa agora $v \in E$ un vector calquera. Podemos considerar a aplicación

$$\Phi_{f,v} \,:\, K[X] \longrightarrow E, \quad p(X) \mapsto p(f)(v).$$

O núcleo de $\Phi_{f,v}$ está formado por aqueles polinomios tales que $p(f)(v) = 0$, e como sucedía no caso anterior, hai unha noción de polinomio mínimo para o par (f, v), $\mathrm{Min}(f, v; X)$, isto é, calquera outro que anule v será un múltiplo do mínimo.

Proposición. Sexa $\mathrm{Min}(f, v; X) = a_s X^s + \ldots + a_1 X + a_0$ o polinomio mínimo do par (f, v). Entón, os vectores $v, f(v), \ldots, f^{s-1}(v)$ son linealmente independentes, pero en cambio os vectores $v, f(v), \ldots, f^{s-1}(v), f^t(v)$, para calquera $t \geq s$, son linealmente dependentes.

Demostración. Se $v, f(v), \ldots, f^{s-1}(v)$ fosen linealmente dependentes habería un polinomio $p(X)$ de grao menor que s tal que $p(f)(v)$. Iso é contraditorio coa definición de $\mathrm{Min}(f, v; X)$.
Se $t = s$, $m_v(f)(v) = a_0 v + a_1 f(v) + \ldots + a_s f^s(v) = 0$, co cal os vectores son linealmente independentes. Para $t > s$ procedemos por indución, usando como hipótese que $f^{t-1}(v) \in \langle v, f(v), \ldots, f^{s-1}(u) \rangle$. De aquí temos entón que

$$f^t(v) \in \langle f(v), f^2(v), \ldots, f^s(v) \rangle = \langle v, f(v), \ldots, f^{v-1}(v) \rangle,$$

e a conclusión do enunciado é certa. $\square$

Das condicións da definición, temos que o polinomio mínimo $\mathrm{Min}(f; X)$ é un múltiplo de $\mathrm{Min}(f, v; X)$ para calquera $v \in E$.

Exemplo. Antes vimos que o polinomio mínimo de $B = \begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix}$ é $X^2 - 5X + 6$. En cambio, se consideramos o vector $e_1 = (1, 0)$, temos que $X - 2$ é anulador xa que

$$(B - 2\mathbb{I})e_1 = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Do mesmo xeito, $X - 3$ é un anulador para o vector $e_2 = (0, 1)$. Máis en xeral, coas ideas introducidas na seguinte sección, temos que unha condición necesaria para que un vector teña un polinomio mínimo de grao menor que o do endomorfismo é que pertenza a un subespazo invariante por f diferente do total (no caso dun endomorfismo que diagonaliza, que sexa combinación unicamente dalgúns vectores propios, pero non de todos eles).

2.2. Subespazos invariantes

Definición. Sexa $f \in \mathrm{End}(E)$. Un subespazo vectorial $F \subset E$ dise que é invariante por f se $f(F) \subset F$. Neste caso, f induce un endomorfismo de F, que se denota por $f|F$

$$f|F \,:\, F \longrightarrow F, \quad v \mapsto f(v)$$

e que se acostuma chamar restrición de f a F.

En particular, o subespazo xerado por un vector propio é un subespazo invariante de dimensión 1, ou máis en xeral r vectores propios xeran un subespazo invariante de dimensión r.

Proposición. Sexa $\mathrm{Min}(f|F;X)$ o polinomio mínimo da restrición de f a un subespazo invariante F.

(a) $\mathrm{Min}(f|F;X)$ divide $\mathrm{Min}(f;X)$.

(b) Se G é outro subespazo invariante e $(\mathrm{Min}(f|F;X),\mathrm{Min}(f|G;X))=1$, entón tense que $F\cap G=\{0\}$.

(c) Para todo $p(X)\in K[x]$, os subespazos $\ker(p(f))$ e $\mathrm{im}(p(f))$ son invariantes por f.

(d) Se $\mathrm{Min}(f;X)=p(X)q(X)$, con $(p(X),q(X))=1$, $E=\ker(p(f))\oplus\ker(q(f))$.

Demostración. (a) $\mathrm{Min}(f;X)$ cumpre que $\mathrm{Min}(f;f)(v)=0$ para todo $v\in E$. Polo tanto, $\mathrm{Min}(f;f)(v)=0$ para todo $v\in F$; o polinomio mónico de grao mínimo que cumpre esa propiedade é $\mathrm{Min}(f|F;X)$, que polo tanto ten que ser un divisor de $\mathrm{Min}(f;X)$.

(b) O subespazo $F\cap G$ tamén é invariante (comprobación rutineira), e polo apartado anterior, o seu polinomio mínimo divide tanto $\mathrm{Min}(f|F;X)$ como $\mathrm{Min}(f|G;X)$, e como son polinomios relativamente primos entre si, o polinomio mínimo de $F\cap G$ ten que ser 1. Iso quere dicir que $F\cap G=\{0\}$.

(c) Imos facer primeiro a comprobación para $\ker(p(f))$. Sexa $v\in\ker(p(f))$. Entón, temos que comprobar que $f(v)\in\ker(p(f))$, que se ve observando que

$$p(f)(f(v))=f(p(f)(v))=f(0)=0.$$

Se $v=p(f)(u)$ é un elemento na imaxe, entón

$$f(v)=f(p(f)(u))=p(f)(f(u))\in\mathrm{im}(p(f)),$$

co cal temos a conclusión desexada.

(d) Comezamos vendo que $\mathrm{im}(q(f))\subset\ker(p(f))$. Para iso, sexa $v=q(f)(u)$; entón $p(f)(v)=p(f)q(f)(u)=\mathrm{Min}(f;f)(u)=0$. Polo tanto, $v\in\ker(p(f))$. Polo primeiro teorema de isomorfismo aplicado ao endomorfismo $p(f)$, temos que

$$\begin{aligned}n&=\dim\ker(p(f))+\dim\mathrm{im}(p(f))\\&\le\dim\ker(p(f))+\dim\ker(q(f))\quad,\\&=\dim(\ker(p(f))\oplus\ker(q(f)))\le n\end{aligned}$$

onde a última igualdade é consecuencia do segundo apartado, xa que o polinomio mínimo de $\ker(p(f))$ é un divisor de $p(X)$ e o de $\ker(q(f))$ é un divisor de $q(X)$, co cal son coprimos. Polo tanto, tódalas inclusións teñen que ser igualdades e temos que $E=\ker(p(f))\oplus\ker(q(f))$.

□

Este resultado dinos que unha factorización en irredutibles do polinomio mínimo induce unha descomposición do espazo vectorial correspondente. Antes de poder demostrar o primeiro teorema de descomposición, necesitamos establecer que o polinomio mínimo da restrición de f a cada un deses espazos é, efectivamente, o factor correspondente de $\mathrm{Min}(f;X)$.

Lema. Se $\mathrm{Min}(f;X) = p(X)q(X)$, con $(p(X), q(X)) = 1$, entón a restrición de f a $\ker(p(f))$ e $\ker(q(f))$ cumpre que $\mathrm{Min}(f|\ker(p(f));X) = p(X)$ e $\mathrm{Min}(f|\ker(q(f));X) = q(X)$.

Demostración. Polo que vimos antes, $p(X)$ e $q(X)$ son polinomios anuladores das restricións de f a cada un dos subespazos; entón, o polinomio mínimo ten que ser un divisor deles. Sexan $r(x)$ e $s(X)$ os divisores correspondentes de $p(X)$ e $q(X)$ que anulan as restricións de f; en particular $r(X)s(X)$ é un divisor de $\mathrm{Min}(f;X)$. Ademais, $r(X)s(X)$ é un anulador de f: se $v \in E$ e $v = v_1 + v_2$, con $v_1 \in \ker(p(f))$ e $v_2 \in \ker(q(f))$, temos que

$$r(X)s(X)(v) = s(X)r(X)(v_1) + r(X)s(X)(v_2) = 0 + 0 = 0.$$

Polo tanto, $r(X)s(X)$ é un múltiplo de $\mathrm{Min}(f;X)$. Isto quere dicir que $r(X)s(X)$ e $\mathrm{Min}(f;X)$ son iguais, xa que non pode pasar que dous polinomios mónicos sexan cada un deles múltiplos do outro. Conclúese por tanto que $r(X) = p(X)$ e $s(X) = q(X)$. $\square$

Podemos agora enunciar o primeiro teorema de descomposición.

Proposición. Se o polinomio mínimo de $f \in \mathrm{End}(E)$ é

$$\mathrm{Min}(f;X) = m_1(X)^{n_1} \cdots m_r(X)^{n_r},$$

con $m_1(X), \ldots, m_r(X)$ factores irredutibles, o espazo E é suma directa de subespazos invariantes $E = E_1 \oplus \ldots \oplus E_r$, de forma que o polinomio mínimo da restrición de f a E_i é $m_i(X)^{n_i}$. Esta descomposición é a única que cumpre $E_i = \ker(m_i(f)^{n_i})$ para todo $i = 1, \ldots, r$.

Demostración. Dos resultados anteriores, a existencia da descomposición está clara. Supoñamos que temos unha descomposición arbitraria $E = E_1 \oplus \ldots \oplus E_r$ da que só sabemos que o polinomio mínimo da restrición de f a E_i é $m_i(X)^{n_i}$. Polo visto antes, isto quere dicir que $E^i \subset \ker(m_i(f)^{n_i})$. De aquí, temos que

$$\begin{aligned} n &= \dim E_1 + \ldots + \dim E_r \\ &\leq \dim\ker(m_1(f)^{n_1}) + \ldots + \dim\ker(m_r(f)^{n_r}) = n. \end{aligned}$$

$\square$

A importancia deste teorema radica en que para entender un endomorfismo f, é entón suficiente entender a súa restrición a cada un dos subespazos.

Proposición. Se λ é un valor propio de f, entón $x - \lambda$ divide $\mathrm{Min}(f;X)$.

Demostración. Se λ é un valor propio, entón o núcleo $\ker(f - \lambda)$ é un subespazo invariante non trivial. O seu polinomio mínimo é $X - \lambda$, co cal para calquera v dese espazo temos que $\mathrm{Min}(f, v; X) = X - \lambda$, e sabemos que o polinomio mínimo do endomorfismo $\mathrm{Min}(f;X)$ é un múltiplo de tódolos $\mathrm{Min}(f, v; X)$. $\square$

Proposición. Un endomorfismo diagonaliza se e soamente se o seu polinomio mínimo descompón en factores lineais non repetidos.

Demostración. Supoñamos primeiro que $\mathrm{Min}(f;X) = (X-\lambda_1)\cdots(X-\lambda_r)$, con tódolos λ_i diferentes entre si. Entón, polo primeiro teorema de descomposición, $E = E_1 \oplus \ldots \oplus E_r$, con $E_i = \ker(f - \lambda_i \mathbb{I})$ o subespazo de vectores propios de valor propio λ_i. Existe polo tanto unha base de vectores propios de E formada a partir de bases de cada un dos subespazos propios.

Vexamos agora o recíproco, supondo que E ten unha base de vectores propios, que podemos agrupar en función do valor propio: $\{v_{1,1}, \ldots, v_{1,m_1}\}$ para λ_1, e así ata chegar a $\{v_{r,1}, \ldots, v_{r,m_r}\}$ para λ_r. Nese caso, podemos pór $E_i = \langle v_{i,1}, \ldots, v_{i,m_i}\rangle$ e temos que $E = E_1 \oplus \ldots \oplus E_r$. Ademais, o polinomio mínimo de E_i é $X - \lambda_i$, co cal cada un dos $X - \lambda_i$ é un factor de $\mathrm{Min}(f;X)$. Porén, $(X-\lambda_1)\cdots(X-\lambda_r)$ é un polinomio anulador, xa que se $v = v_1 + \ldots + v_r$, con $v_i \in E_i$, temos que

$$(f-\lambda_1)\cdots(f-\lambda_r)(v) = 0,$$

xa que cada un dos factores $f - \lambda_i$ anula a compoñente v_i. □

Exemplo. Imos considerar a diferenza entre os endomorfismos dados polas matrices

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad C = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}.$$

Tense que $\mathrm{Char}(A;X) = \mathrm{Char}(B;X) = \mathrm{Char}(C;X) = -(X-1)^3$ e resulta doado ver que $(X-1)^3$ é un polinomio anulador para os tres endomorfismos. En particular, o polinomio mínimo de cada un deles é un divisor de $(X-1)^3$, polo que podemos ver que $\mathrm{Min}(A;X) = X - 1$, $\mathrm{Min}(B;X) = (X-1)^2$ e $\mathrm{Min}(C;X) = (X-1)^3$. En termos da descomposición en subespazos, isto quere dicir que se $E = \mathbb{R}^3$, temos que $E = \ker(A - \mathbb{I})$, $E = \ker((B-\mathbb{I})^2)$ e $E = \ker((C-\mathbb{I})^3)$; é dicir, mentres que no caso de A a propia matriz é suficiente para proporcionarnos unha descomposición do espazo, no caso da segunda e da terceira hai vectores que están no núcleo de $(B-\mathbb{I})^2$ ou $(C-\mathbb{I})^3$, pero que non pertencían aos núcleos de $B - \mathbb{I}$ ou $C - \mathbb{I}$, respectivamente. A matriz A é a única que diagonaliza das tres, pois o seu polinomio mínimo descompón en factores lineais.

2.3. Teorema de Cayley–Hamilton

Imos comezar cunha comprobación rutineira, considerando unha matriz xenérica e o seu polinomio característico,

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \quad \mathrm{Char}(A;X) = X^2 - (a+d)X + (ad - bc).$$

Entón, neste caso, podemos comprobar que $A^2 - (a+d)A + (ad-bc) = 0$, isto é,

$$\begin{pmatrix} a^2 + bc & b(a+d) \\ c(a+d) & d^2 + bc \end{pmatrix} - \begin{pmatrix} a^2 + ad & b(a+d) \\ c(a+d) & d^2 + bc \end{pmatrix} + \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}.$$

Este resultado particular das matrices 2×2 pode estenderse a matrices de orde calquera, aínda que a comprobación non é tan rutineira. É importante salientar que o polinomio

característico non ten por que ser o mínimo. Por exemplo, se consideramos $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}$, entón o polinomio característico é $-(X-1)^2(X-2)$, mentres que o polinomio mínimo é $(X-1)(X-2)$.
O primeiro resultado que amosa a relación entre os dous conceptos é o seguinte.

Proposición. Sexa λ un escalar. Entón λ é un cero de $\mathrm{Min}(f;X)$ se e soamente se é un cero de $\mathrm{Char}(f;X)$.

Demostración. Se λ é un cero de $\mathrm{Min}(f;X)$, entón $\mathrm{Min}(f;X) = (X-\lambda)m(X)$, e existe algún vector $v \in E$ tal que $m(f)(v) \neq 0$, xa que en caso contrario o polinomio mínimo sería $m(X)$. Entón o vector $w = m(f)(v)$ ten vector propio λ xa que

$$(f - \lambda\mathbb{I})(w) = (f - \lambda\mathbb{I})m(f)v = \mathrm{Min}(f;f)(v) = 0.$$

Polo tanto, λ é un valor propio e polo tanto un cero do polinomio característico.
Para o recíproco, notamos que se λ é un cero do polinomio característico entón é un valor propio e nese caso xa vimos que $X - \lambda$ é un factor de $\mathrm{Min}(f;X)$. $\square$

O seguinte obxectivo é ver que o polinomio característico sempre é un polinomio anulador. Imos dar primeiro unha versión simplificada deste feito no caso particular no que o polinomio característico descompón totalmente en factores lineais. Logo faremos a proba xeral.

Proposición. Sexa f un endomorfismo que cumpre que o seu polinomio característico descompón en factores lineais. Entón, o polinomio característico é un polinomio anulador de f.

Demostración. Sexa $\mathrm{Min}(f;X) = (X-\lambda_1)^{n_1}\cdots(X-\lambda_r)^{n_r}$ o polinomio mínimo e sexa $E = E_1 \oplus \ldots \oplus E_r$ a descomposición habitual en subespazos invariantes. Temos que a matriz correspondente a A é unha matriz con bloques diagonais, de maneira que se $\mathrm{Char}(f|E_1;X)$ é o polinomio característico da restrición de f a E_i, entón $\mathrm{Char}(f;X) = \mathrm{Char}(f|E_1;X)\cdots\mathrm{Char}(f|E_r;X)$. Cada $\mathrm{Char}(f|E_i;X)$ descompón en factores lineais e os seus ceros son ceros do polinomio mínimo de E_i, polo que $\mathrm{Char}(f|E_i;X) = (x-\lambda_i)^{m_i}$, sendo m_i a dimensión de E_i. Se demostramos que $n_i \leq m_i$ teremos que $\mathrm{Min}(f;X) \mid \mathrm{Char}(f;X)$ e podemos acabar. Para probar iso, se $m_i(X)^{n_i}$ é o polinomio mínimo da restrición de f a E_i, hai un $v_i \in E_i$ tal que $m_i(f)^{n_i}(v_i) = 0$ pero $m_i(f)^{n_i-1}(v_i) \neq 0$. Entón o polinomio mínimo de v_i é $m_i(X)^{n_i}$ e polo visto anteriormente $v_i, f(v_i), \ldots, f^{n_i-1}(v_i)$ son linealmente independentes. Polo tanto, n_i é menor que a dimensión de E_i, é dicir, que m_i. $\square$

Imos dar agora a demostración xeral do resultado, que é o que se coñece como teorema de Cayley–Hamilton.

Teorema. Toda matriz cadrada anula o seu polinomio característico.

Demostración. A proba baséase nos seguintes dous resultados, que demostraremos en primeiro lugar e que logo usaremos para concluír que o teorema de Cayley–Hamilton é certo.

(i) Sexa $A \in \mathcal{M}_n(K)$ de maneira que hai un vector tal que $v, Av, A^2v, \ldots, A^{n-1}v$ é unha base de K^n. Entón $\mathrm{Char}(A;A) = 0$.

(ii) Se A é unha matriz por bloques $\begin{pmatrix} X & Y \\ 0 & Z \end{pmatrix}$ de maneira que X e Z cumpren Cayley–Hamilton, entón A tamén o cumpre.

Para a primeira propiedade, observamos que se $v, Av, \ldots, A^{n-1}v$ é unha base, temos algunha relación da forma $A^n v + \sum_{j=0}^{n-1} \lambda_j A^j v = 0$; alternativamente, podemos pór $\sum_{j=0}^{n} \lambda_j A^j v$, con $\lambda_n = 1$. Polo tanto, na base $v, Av, \ldots, A^{n-1}v$, o endomorfismo representado pola matriz A escríbese como

$$\begin{pmatrix} 1 & 0 & \ldots & 0 & -\lambda_n \\ 0 & 1 & \ldots & 0 & -\lambda_{n-1} \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \ldots & 1 & -\lambda_1 \end{pmatrix}.$$

Nesta base temos que o polinomio característico é $\mathrm{Char}(A; X) = \sum_{i=0}^{n} \lambda_i X^i$. Multiplicando por A^k á esquerda, con $0 \leq k \leq n-1$, deducimos que

$$A^k \Big(\sum_{i=0}^{n} \lambda_i A^i \Big) v = \Big(\sum_{i=0}^{n} \lambda_i A^{k+i} \Big) v = \Big(\sum_{i=0}^{n} \lambda_i A^i \Big) A^k v = \mathrm{Char}(A; A)(A^k v) = 0,$$

xa que vimos ao principio da proba que $\Big(\sum_{i=0}^{n} \lambda_i A^i \Big) v = 0$ para todo v. Polo tanto, o polinomio característico anula tódolos elementos da base e é un polinomio anulador do endomorfismo.

A segunda propiedade é consecuencia directa do lema que demostramos sobre determinantes e que afirma que o determinante de A é $\det(X) \cdot \det(Z)$.

Para demostrar agora o teorema Cayley–Hamilton podemos proceder por indución sobre a dimensión de E. De feito, xa probamos o resultado para dimensión 0, 1 ou 2. Supoñamos logo que a dimensión de E é $n > 0$ e escollamos $v \in E$ diferente de 0. Entón, collemos o mínimo r tal que hai unha relación de dependencia lineal entre $v, Av, \ldots, A^r v$. Como $v \neq 0$, necesariamente $r \geq 1$. Se $r = n$, concluímos pola primeira parte. Senón, completamos $v, Av, \ldots, A^{r-1}v$ a unha base de K^n, na cal a matriz asociada é unha matriz por bloques $A = \begin{pmatrix} X & Y \\ 0 & Z \end{pmatrix}$, con X un bloque $r \times r$ e Z un bloque $(n-r) \times (n-r)$. Pola hipótese de indución, tanto X como Z cumpren o teorema de Cayley–Hamilton, polo que podemos concluír. $\square$

Un dos aspectos importantes do teorema de Cayley–Hamilton é que nos dá un polinomio mínimo que anula a matriz. Por exemplo, no caso da matriz

$$A = \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} \in \mathcal{M}_2(\mathbb{Q}),$$

que non ten valores propios en $\mathbb{Q}$, podemos ver que $\mathrm{Char}(A; X) = X^2 + 2X + 2$. Polo tanto $A^2 + 2A + 2\mathbb{I} = 0$. Iso pode ser útil, por exemplo, para achar a matriz inversa: $A(A + 2\mathbb{I}) = -2\mathbb{I}$, polo que $\mathbb{A}^{-1} = -\frac{1}{2}(A + 2\mathbb{I})$.

3. Forma de Jordan

Polo de agora, vimos que hai dúas obstrucións para que un endomorfismo diagonalice: que o polinomio característico non teña suficientes raíces, ou que para algún valor propio non haxa suficientes vectores propios (é dicir, que o polinomio mínimo teña algunha raíz múltiple). O obxectivo desta sección é describir unha maneira de proceder no segundo caso. Imos analizar un caso límite que ilustra este fenómeno. Sexa $\varepsilon \neq 0$ un número real. Entón, a matriz $\begin{pmatrix} \varepsilon & 1 \\ 0 & 0 \end{pmatrix}$ ten dous valores propios diferentes, ε e 0, con vectores propios asociados $(1, 0)$ e $(1, -\varepsilon)$, respectivamente. Entón,

$$\begin{pmatrix} \varepsilon & 1 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 0 & -\varepsilon \end{pmatrix} \begin{pmatrix} \varepsilon & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & \varepsilon^{-1} \\ 0 & -\varepsilon^{-1} \end{pmatrix}.$$

O problema cando ε tende a 0, o vector propio asociado a ε tende a $(1, 0)$, que é o mesmo vector propio que o asociado a 0. Polo tanto, a matriz de cambio de base deixa de ser invertible e temos unha singularidade. A forma de Jordan é unha alternativa a este problema.

A modo de notación, diremos que un polinomio descompón completamente cando ten tantas raíces como o seu grao, contando multiplicidades. Un corpo no que isto sempre sucede dise que é alxebricamente pechado. Por exemplo, $\mathbb{C}$ é alxebricamente pechado, pero non é o único exemplo.

3.1. Vectores propios xeneralizados

Antes de comezar, observamos que polo primeiro teorema de descomposición é suficiente tratar por separado cada un dos subespazos propios correspondentes a un endomorfismo.

Para introducir a forma de Jordan, imos comezar estendendo a definición de vector propio a través da seguinte xeralización.

Definición. Sexa λ un valor propio do endomorfismo f. Un vector propio xeneralizado de valor propio λ é un vector non cero $v \in V$ de maneira que $f_\lambda^k(v) = 0$ para algún $k \geq 1$. Un ciclo de vectores propios xeneralizados de valor propio λ (tamén chamado ciclo de Jordan) é unha sucesión de vectores non nulos $v_1, v_2, \ldots, v_r$ con $f_\lambda(v_r) = v_{r-1}, \ldots, f_r(v_2) = v_1$ e $f_\lambda(v_1) = 0$. O vector v_r chámase xerador do ciclo.

Exemplo. Sexa

$$A = \begin{pmatrix} 2 & 3 & 4 \\ 0 & 2 & 5 \\ 0 & 0 & 2 \end{pmatrix}.$$

Consideremos os vectores $v_3 = (0, 0, 1)$, $v_2 = (4, 5, 0)$ e $v_1 = (15, 0, 0)$. Entón, $(A - 2\mathbb{I})v_3 = v_2$, $(A - 2\mathbb{I})v_2 = v_1$ e $(A - 2\mathbb{I})v_1 = 0$. Polo tanto, (v_1, v_2, v_3) é un ciclo de Jordan para o endomorfismo representado por A.

Nun ciclo de Jordan de valor propio λ, cada vector v_k pertence a $\ker(f_\lambda^k) \setminus \ker(f_\lambda^{k-1})$, e o primeiro vector v_1 é un vector propio de valor propio λ. Dise que dous ciclos de Jordan son independentes se os subespazos que xeran están en suma directa.

Proposición. Sexa $v_1, \ldots, v_r$ un ciclos de Jordan de valor propio λ para un endomorfismo $f \in \mathrm{End}(E)$. Entón:

(a) Os vectores v_i son independentes.

(b) Para calquera escalar $\mu \in K$, o subespazo $\langle v_1, v_2, \ldots, v_r \rangle$ é invariante por f_μ.

(c) Se $\mu \neq \lambda$, entón f_μ é un automorfismo do espazo xerado polos r vectores v_i.

Demostración. (a) Para demostrar o primeiro apartado procedemos por indución sobre a lonxitude do ciclo. Se $\ell = 1$, é inmediato. Supoñámolo agora certo para ciclos e lonxitude $\ell-1$ e collamos unha combinación lineal da forma $x_1v_1+\ldots+x_\ell v_\ell = 0$. Aplicando f_λ, vemos que

$$f_\lambda(x_1v_1 + \ldots + x_\ell v_\ell) = x_2v_1 + \ldots + x_\ell v_{\ell-1}.$$

Como os vectores $v_1, \ldots, v_{\ell-1}$ son independentes pola hipótese de indución $x_2 = \ldots = x_\ell = 0$, e entón o coeficiente con x_1 tamén é 0.

(b) Temos que $f_\mu = f_\lambda + (\lambda - \mu)\mathbb{I}$, polo que podemos ver que $f_\mu = (\lambda - \mu)v_1$ e $f_\mu(v_i) = v_{i-1} + (\lambda - \mu)v_i$ para todo $2 \leq i \leq r$. Isto demostra a invariancia do subespazo.

(c) A matriz de f_μ na base formada por $v_1, \ldots, v_r$ ten os valores $\lambda - \mu$ na diagonal e uns na diagonal superior. Polo tanto, o seu determinante é $(\lambda-\mu)^r$, que é diferente de 0.

□

Chámaselle bloque de Jordan á matriz da restrición do endomorfismo f ao subespazo xerado polos v_i. É unha matriz da forma

$$J_{\lambda,\ell} = \begin{pmatrix} \lambda & 1 & 0 & \ldots & 0 & 0 \\ 0 & \lambda & 1 & \ldots & 0 & 0 \\ 0 & 0 & \lambda & \ldots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \ldots & \lambda & 1 \\ 0 & 0 & 0 & \ldots & 0 & \lambda \end{pmatrix} \in \mathcal{M}_r(K),$$

onde λ é o valor propio correspondentes aos vectores propios xeneralizados v_i.

Definición. Unha base de Jordan para un endomorfismo f é unha base formada pola unión de ciclos de Jordan para este endomorfismo.

3.2. A descomposición canónica de Jordan

De cara a introducir a descomposición canónica de Jordan, procedemos considerando unha táboa da seguinte maneira.

- Cada columna representa un ciclo de Jordan.
- Situamos as columnas en orde descendente segundo a súa lonxitude.
- Na columna i-ésima, con $1 \leq i \leq r$, situamos o vector propio $v_{i,1}$ na fila inferior, $v_{i,2}$ na seguinte fila, e así ata chegar a v_{i,ℓ_i} na fila ℓ_i-ésima.
- En particular, altura de cada columna é igual á lonxitude do ciclo.
- Se unha columna só ten un elemento quere dicir que o ciclo unicamente consiste dun vector propio.

Proposición. Sexa $g \in \mathrm{End}(E)$. Entón, para cada enteiro $k \geq 1$, temos as seguintes propiedades.

(a) $\ker(g^k) \subset \ker(g^{k+1})$.

(b) Se $\ker(g^k) = \ker(g^{k+1}$, entón $\ker(g^{k+1}) = \ker(g^{k+2})$.

(c) A aplicación $\varphi : \ker(g^{k+1})/\ker(g^k) \longrightarrow \ker(g^k)/\ker(g^{k-1})$ definida por $\varphi([v]) = [g(v)]$ está ben definida e é inxectiva.

Demostración. (a) Se $v \in \ker(g^k)$, entón $g^k(v) = 0$ e polo tanto $0 = g(g^k(v)) = g^{k+1}(v)$.

(b) Se $\ker(g^k) = \ker(g^{k+1})$ e $v \in \ker(g^{k+2})$, entón $g^{k+2}(v) = g^{k+1}(g(v)) = 0$. Polo tanto, $g(v) \in \ker(g^{k+1}) = \ker(g^k)$ e sucede que $g^k(g(v)) = 0 = g^{k+1}(v)$, de onde resulta que $v \in \ker(g^{k+1})$.

(c) En primeiro lugar observamos que se $v \in \ker(g^{k+1})$, entón $g(v) \in \ker(g^k)$ xa que $g^k(g(v)) = g^{k+1}(v) = 0$. Ademais, se $[u] = [v]$, entón $u - v \in \ker(g^k)$ e polo tanto $g(u-v) \in \ker(g^{k-1})$, co cal $[g(u)] = [g(v)]$. Isto asegura que φ está ben definida. A linealidade é inmediata a partir da definición. Para ver a inxectividade, se $\varphi([v]) = [g(v)] = [0]$, entón $g(v) \in \ker(g^{k-1})$. Iso quere dicir que $g^k(v) = g^{k-1}(g(v))$ e polo tanto $[v] = 0$.

□

O seguinte resultado di que os ciclos de valores propios pódense estruturar nun *edificio* no que os diferentes pisos son familias de vectores de $\ker(f_\lambda^j)$ de maneiras que as súas clases son independentes módulo o subespazo $\ker(f_\lambda^{j-1})$.

Lema. Sexa λ un valor propio de $f \in \mathrm{End}(E)$. Para $i = 1, \dots, r$, sexan $v_{i,1}, \dots, v_{i,\ell_i}$ ciclos de Jordan de valor propio λ de lonxitudes decrecentes: $\ell_1 \geq \ell_2 \geq \dots \geq \ell_r$. Se para cada $j = 1, \dots, \ell_i$ as clases $[v_{i,j}]$ son linealmente independentes no cociente $\ker(f_\lambda^j)/\ker(f_\lambda^{j-1})$, entón tódolos vectores $v_{i,j}$ son linealmente independentes en E.

Demostración. Procedemos por indución sobre ℓ_1. Se $\ell_1 = 1$, entón temos un único piso e a colección consta unicamente de vectores propios.
Consideramos agora unha familia de k' ciclos de lonxitudes ℓ'_i, con $\ell'_1 = \ell_1 + 1$. Collemos unha combinación lineal

$$\sum_{i=1}^{k'} \sum_{j=1}^{\ell'_i} x_{i,j} v_{i,j} = 0,$$

e veremos que tódolos coeficientes son cero. Aplicando o endomorfismo f_λ a cada lado, deducimos a igualdade

$$\sum_{i=1}^{k'} \sum_{j=1}^{\ell'_i} x_{i,j} f_\lambda(v_{i,j}) = \sum_{i=1}^{k} \sum_{j=1}^{\ell'_i - 1} x_{i,j+1} v_{i,j} = 0.$$

Isto é unha combinación lineal dos vectores da familia de $k \leq k'$ ciclos que se obteñen eliminando o último vector de tódolos ciclos da familia de partida. Dito doutra maneira, estamos eliminando o último piso dos bloques de Jordan. Esta nova familia ten o primeiro ciclos de lonxitude $\ell_1 = \ell'_1 - 1$. Ademais, segue cumprindo a condición sobre independencia lineal nos cocientes $\ker(f_\lambda^j)/\ker(f_\lambda^{j-1})$, xa que os vectores que hai en cada piso son algúns dos que xa había antes (e unha subfamilia dunha independente segue a ser independente).

Pola hipótese de indución, tódolos coeficientes desta combinación lineal son cero, e polo tanto tódolos coeficientes $x_{i,j}$ da combinación lineal inicial con $j > 1$ tamén son cero. $\square$

O seguinte resultado é o que adoita denominarse como segundo teorema de descomposición.

Teorema. Sexa $f \in \mathrm{End}(E)$ un endomorfismo tal que o seu polinomio característico descompón completamente. Para cada valor propio λ de multiplicidade m, o subespazo $\ker(f_\lambda^m)$ ten unha base de Jordan.

Demostración. Xa sabemos que hai unha cadea de subespazos $\ker(f_\lambda) \subset \ker(f_\lambda^2) \subset \dots$ que nalgún momento estabiliza (de feito, con expoñente $\leq m$). Sexa $\ell \leq m$ o menor enteiro positivo tal que $\ker(f_\lambda^\ell) = \ker(f_\lambda^{\ell+1})$. Consideramos como antes unha táboa, enchéndoa por pisos e comezando por arriba. No piso ℓ-ésimo pomos vectores de maneira que a súa clase sexa unha base de $\ker(f_\lambda^\ell)/\ker(f_\lambda^{\ell-1})$. Calculamos as súas imaxes por f_λ e colócanse no piso inferior. Grazas aos resultados anteriores, as clases destas imaxes tamén son independentes en $\ker(f_\lambda^{\ell-1})/\ker(f_\lambda^{\ell-2})$. Completamos estas clases ata que formen unha base de $\ker(f_\lambda^{\ell-1})/\ker(f_\lambda^{\ell-2})$ engadindo clases de $\ker(f_\lambda^{\ell-1})$, e desta maneira completamos o piso $\ell - 1$.
Repetimos o proceso ata chegar á base, procedendo dese mesmo xeito. Unha vez temos os vectores $v_{i,j}$ do piso j tales que as súas clases son unha base no cociente $\ker(f_\lambda^j)/\ker(f_\lambda^{j-1})$, calculamos as súas imaxes $f_\lambda(v_{i,j}) = v_{i,j-1}$ e póñense no piso inferior. As clases destes vectores son independentes no cociente $\ker(f_\lambda^{j-1})/\ker(f_\lambda^{j-2})$ e complétase o piso $(j-1)$-ésimo engadindo vectores de $\ker(f_\lambda^{j-1})$ de maneira que as clases de todos sexan unha base do cociente $\ker(f_\lambda^{j-1})/\ker(f_\lambda^{j-2})$.
O lema anterior asegura que, unha vez completada a táboa desta maneira, os vectores que a forman serán independentes e, por dimensión, serán unha base de $\ker(f_\lambda^m)$, que está formada por unha unión de ciclos de Jordan. $\square$

Nesta construción sucede o seguinte, que ilustra con máis precisión a forma da táboa que representa a forma de Jordan.

- A altura é a lonxitude dos ciclos máis longos.
- A anchura é a dimensión do subespazo $\ker(f_\lambda)$ dos vectores propios ordinarios (isto é, a multiplicidade xeométrica).
- A dimensión de $\ker(f_\lambda^k)$ é o número de vectores que hai ata o piso k incluído.
- O número de columnas de altura $\geq \ell$ é a dimensión de $\ker(f_\lambda^\ell/\ker(f_\lambda^{\ell-1})$.
- O número de columnas de altura exactamente igual a ℓ é

$$2 \dim \ker(f_\lambda^\ell) - \dim \ker(f_\lambda^{\ell-1}) - \dim \ker(f_\lambda^{\ell+1}).$$

Unha primeira consecuencia da existencia de bases de Jordan é o seguinte resultado.

Proposición. Un endomorfismo admite unha base de Jordan se e soamente se o seu polinomio característico descompón completamente. Se A e B son dúas matrices co mesmo polinomio característico, que descompón completamente, A e B representan o mesmo endomorfismo se e soamente se teñen a mesma forma de Jordan.

Demostración. A primeira parte da proposición é inmediata a partir do resultado anterior. Para a segunda, se dúas matrices teñen a mesma forma de Jordan, entón $P^{-1}AP = Q^{-1}BQ = J$, para algunhas $P, Q \in \mathrm{GL}_n(K)$. Por tanto $B = (PQ^{-1})^{-1}A(PQ^{-1})$ e ambas matrices son conxugadas. Para o recíproco, se ambas son conxugadas e se cumpre $B = P^{-1}AP$, entón teñen o mesmo polinomio característico, que por hipótese descompón completamente. Para cada valor propio λ o número e medida dos ciclos de Jordan de valor propio λ unicamente depende dos rangos das matrices $(A - \lambda\mathbb{I})^k$ e $(B - \lambda\mathbb{I})^k$ para todo $k \geq 0$. Como

$$(B - \lambda\mathbb{I})^k = (P^{-1}AP - \lambda\mathbb{I})^k = P^{-1}(A - \lambda\mathbb{I})^k P,$$

o rango destas matrices son os mesmos. □

3.3. Exemplos de cálculo de matrices de Jordan

Comezamos considerando a matriz

$$A = \begin{pmatrix} -2 & 2 & 1 \\ -7 & 4 & 2 \\ 5 & 0 & 0 \end{pmatrix}$$

Neste caso, $\mathrm{Char}(A; X) = -t(t-1)^2$, co cal os valores propios son 0 e 1. O valor propio 0 ten multiplicidade 0, co cal é suficiente con achar $\ker(A) = \langle(0, -1, 2)$. En cambio, $\ker(A + \mathbb{I}) = \langle(1, -1, 5)$, e polo tanto, o 1 ten multiplicidade xeométrica 1. Achamos entón $\ker(A + \mathbb{I})^2$, e é suficiente coller un vector que estea nese núcleo pero non no de $A + \mathbb{I}$. Sexa $v_3 = (2, 1, 0)$. Entón $v_2 = (A - \mathbb{I})v_2 = (1, -1, 5)$. Polo tanto, unha descomposición da matriz é

$$\begin{pmatrix} -2 & 2 & 1 \\ -7 & 4 & 2 \\ 5 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 1 & 2 \\ -1 & -1 & 1 \\ 2 & 5 & 0 \end{pmatrix} \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & 1 & 2 \\ -1 & -1 & 1 \\ 2 & 5 & 0 \end{pmatrix}^{-1}.$$

Sexa agora

$$B = \begin{pmatrix} 1 & 1 & 0 & 0 & 2 \\ 1 & 1 & 1 & -1 & -4 \\ 1 & 1 & 2 & -1 & -3 \\ 1 & 2 & 1 & 0 & -1 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

O polinomio característico é $\mathrm{Char}(B; X) = -(X - 1)^5$. Ademais,

$$B - \mathbb{I} = \begin{pmatrix} 0 & 1 & 0 & 0 & 2 \\ 1 & 0 & 1 & -1 & -4 \\ 1 & 1 & 1 & -1 & -3 \\ 1 & 2 & 1 & -1 & -1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}, \quad (B - \mathbb{I})^2 = \begin{pmatrix} 1 & 0 & 1 & -1 & -4 \\ 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & -1 & -4 \\ 2 & 0 & 2 & -2 & -8 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Podemos ver que $\dim\ker(B-\mathbb{I}) = 2$, $\dim\ker(B-\mathbb{I})^2 = 4$ e $\dim\ker(B-\mathbb{I})^3 = 5$. Sabemos entón que a base de Jordan estará formado por vectores $(v_1, v_2, v_3, v_4, v_5)$, de maneira que (v_1, v_2, v_3) e (v_4, v_5) son ciclos de Jordan. Comezamos achando v_3 encontrando un elemento de $\ker((B - I)^3)$, pero que non estea en $\ker((B - I)^2)$. Podemos tomar por exemplo $v_3 = (1, 0, 0, 0, 0)$. Entón, $v_2 = (B - \mathbb{I})v_3 = (0, 1, 1, 1, 0)$ e $v_1 = (B - \mathbb{I})v_2 =$

$(1,0,1,2,0)$. Para achar v_5, temos que completar v_2 a unha base de $\ker(B-\mathbb{I})^2/\ker(B-\mathbb{I})$. Por exemplo, podemos coller $v_5=(3,0,1,0,1)$ e entón $v_4=(B-\mathbb{I})v_5=(2,0,1,3,0)$. Polo tanto

$$\begin{pmatrix}1&1&0&0&2\\1&1&1&-1&-4\\1&1&2&-1&-3\\1&2&1&0&-1\\0&0&0&0&1\end{pmatrix}=\begin{pmatrix}1&0&1&2&3\\0&1&0&0&0\\1&1&0&1&1\\2&1&0&3&0\\0&0&0&0&1\end{pmatrix}\begin{pmatrix}1&1&0&0&0\\0&1&1&0&0\\0&0&1&0&0\\0&0&0&1&1\\0&0&0&0&1\end{pmatrix}\begin{pmatrix}1&0&1&2&3\\0&1&0&0&0\\1&1&0&1&1\\2&1&0&3&0\\0&0&0&0&1\end{pmatrix}^{-1}.$$

Unha das aplicacións habituais da forma de Jordan é achar os subespazos invariantes por un endomorfismo f. O primeiro teorema de descomposición dinos que podemos tratar cada valor propio por separado; por outro lado, os bloques de Jordan dan unha descomposición da matriz en suma directa de subespazos. A partir de aí, é inmediato ver que para cada un dos ciclos simplemente temos que decidir cantos elementos collemos: ningún; unicamente o vector propio; o vector propio e o seguinte elemento do ciclo; e así sucesivamente. Nunca sucederá que un elemento do ciclo diferente do vector propio forme por si só un subespazo invariante, senón que os hai que coller de forma ordenada: se $(v_1,\dots,v_\ell)$ é un ciclo de Jordan, entón podemos coller o subespazo xerado por $(v_1,\dots,v_k)$, con $0\le k\le\ell$.